School of Mathematical
and Physical Sciences

University of Reading

Department of Mathematics and Statistics

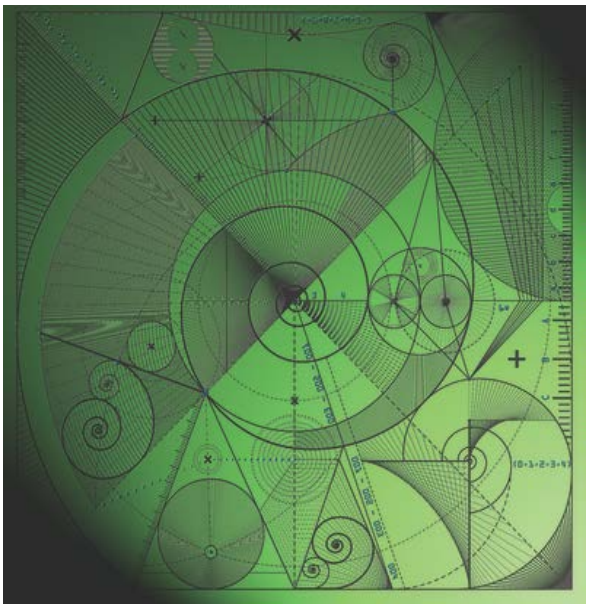
Preprint MPS-2013-14

7 November 2013

Estimating correlated observation errors with an ensemble transform Kalman filter

by

J. A. Waller, S. L. Dance, A. S. Lawless
and N. K. Nichols



Estimating correlated observation errors with an ensemble transform Kalman filter

J. A. Waller*, S. L. Dance, A. S. Lawless, and N. K. Nichols

School of Mathematical and Physical Sciences, University of Reading, Reading, Berkshire, United Kingdom

Abstract

For certain observing types, such as those that are remotely sensed, the observation errors are correlated and these correlations are state and time-dependent. In this work we develop a new method for diagnosing and incorporating spatially correlated and time-dependent observation error in an ensemble data assimilation system. The method combines an ensemble transform Kalman filter with a method that uses statistical averages of background and analysis innovations to provide an estimate of the observation error covariance matrix.

To evaluate the performance of the method we run identical twin experiments using the Lorenz '96 and Kuramoto-Sivashinsky models. Using our approach we are able to recover a good approximation of the true observation error covariance in cases where the initial estimate of the error covariance is incorrect. We are also able to capture spatial observation error covariances where the length-scale of the true covariance is changing slowly in time. We find that using the estimated correlated observation error in the assimilation improves the analysis.

1 Introduction

Data assimilation techniques combine observations with a model prediction of the state, known as the background, to provide a best estimate of the state, known as the analysis. The errors associated with the observations can be attributed to four main sources:

1. Instrument error.
2. Error introduced in the observation operator - these include modelling errors, such as the misrepresentation of gaseous constituents in radiative transfer models, and errors due to the approximation of a continuous function as a discrete function.
3. Errors of representativity - these are errors that arise where the observations can resolve spatial scales that the model cannot.
4. Pre-processing errors - errors introduced by preprocessing of the data such as cloud clearance for radiances.

For a data assimilation scheme to produce an optimal estimate of the state, the error covariances associated with the observations and background must be well understood and correctly specified [Houtekamer and Mitchell, 2005]. In practice many assumptions are violated and the analysis provided by the assimilation may be far from optimal. Therefore

*Corresponding Author.
email: j.a.waller@reading.ac.uk

to obtain an accurate analysis it is important to have good estimates of the observation and background error covariance matrices to be used in the assimilation.

In previous work much attention has been given to the estimation of the background error covariance matrix and as a result static background error covariance matrices are now often replaced with flow dependent matrices that reflect the 'errors of the day' [Bannister, 2008]. Until recently less emphasis has been given to understanding the nature of the observation error covariance matrix and the matrix is often assumed diagonal. The unknown errors, such as the representativity error, and any other possible unaccounted for correlations, are represented by inflating the error variance [Hilton et al., 2009, Whitaker et al., 2008], or by using techniques such as observation thinning [Buehner, 2010] or 'superobbing' [Daley, 1991].

Although correlated observation errors are not widely accounted for in operation assimilation schemes, methods do exist for calculating the observation error covariance matrix [Hollingsworth and Lönnberg, 1986, Desroziers et al., 2005]. Recent work [Stewart et al., 2009, 2013b, Bormann et al., 2002, Bormann and Bauer, 2010, Bormann et al., 2010] using these methods has shown that for certain observing instruments the observation error covariance matrix is correlated. When these correlated errors have been accounted for in the assimilation, it has been shown to lead to a more accurate analysis [Stewart et al., 2013a, Stewart, 2010, Healy and White, 2005], the inclusion of more observation information content [Stewart et al., 2008] and improvements in the UK Met Office skill score [Weston, 2011]. Indeed, Stewart et al. [2013a] and Healy and White [2005] show that even the use of a crude approximation to the observation error covariance matrix may provide significant benefit.

The importance of accounting for correlated errors in the assimilation has led to the development of new schemes that provide estimates of the observation error covariance matrix. Miyoshi et al. [2013] use the diagnostic of Desroziers et al. [2005] (hereafter denoted the DBCP diagnostic) embedded in a local ensemble transform Kalman filter to give an estimate of a static observation error covariance matrix. At each analysis step the DBCP diagnostic is applied to a subset of observations to give an estimate for the observation error covariance matrix. Their work focuses on convergence to a static observation error covariance matrix. However, in previous work it has been shown that representativity errors are time-dependent [Janjic and Cohn, 2006, Waller, 2013, Waller et al., 2013]. Therefore we consider if it is possible to estimate an observation error covariance matrix that varies in time.

In this paper we introduce a new method that combines an ensemble filter with the DBCP diagnostic. Our method uses statistics from observations over a short period of time to produce a slowly time varying estimate of the observation error covariance matrix that is then used in the assimilation.

In section 2 we describe the ensemble filter that can be used to provide a time varying estimate for correlated observation error. Our experimental design is given in section 3 and we present our numerical results in section 4. Finally in section 5 we conclude that, as demonstrated in experiments using simple models, it is possible to use the proposed method to provide an estimate of spatial observation error correlations that vary slowly in time.

2 Estimating the observation error covariance matrix with the ensemble transform Kalman filter

Data assimilation techniques combine observations $\mathbf{y}_n \in \mathbb{R}^{N^p}$ at time t_n with a model prediction of the state, the background $\mathbf{x}_n^f \in \mathbb{R}^{N^m}$, which is often determined by a previous forecast. The observations and background are weighted by their respective errors, to provide a best estimate of the state $\mathbf{x}_n^a \in \mathbb{R}^{N^m}$, known as the analysis. This analysis is then forecast using the possibly non-linear model $\mathcal{M}_n$ to provide a background at the next assimilation time,

$$\mathbf{x}_{n+1}^f = \mathcal{M}_n(\mathbf{x}_n^a). \tag{1}$$

We now give a brief overview of the ensemble transform Kalman filter (ETKF) [Bishop et al., 2001, Livings et al., 2008] that we will adapt and the notation that is used in this study. At time t_n we have an ensemble, a statistical sample of N state estimates $\left\{\mathbf{x}_n^i\right\}$ for $i = 1 \dots N$. These ensemble members are stored in a state ensemble matrix $\mathbf{X}_n \in \mathbb{R}^{N^m \times N}$ where each column of the matrix is a state estimate for an individual ensemble member,

$$\mathbf{X}_n = \begin{pmatrix} \mathbf{x}_n^1 & \mathbf{x}_n^2 & \dots & \mathbf{x}_n^N \end{pmatrix}. \tag{2}$$

It is possible to calculate the ensemble mean,

$$\bar{\mathbf{x}}_n = \frac{1}{N}\sum_{i=1}^{N} \mathbf{x}_n^i, \tag{3}$$

and subtracting the ensemble mean from the state ensembles gives the ensemble perturbation matrix

$$\mathbf{X}'_n = \begin{pmatrix} \mathbf{x}_n^1 - \bar{\mathbf{x}}_n & \mathbf{x}_n^2 - \bar{\mathbf{x}}_n & \dots & \mathbf{x}_n^N - \bar{\mathbf{x}}_n \end{pmatrix}. \tag{4}$$

This allows us to write the ensemble covariance matrix as

$$\mathbf{P}_n = \frac{1}{N-1}\mathbf{X}'_n {\mathbf{X}'_n}^T. \tag{5}$$

For the ensemble transform Kalman filter (ETKF) the analysis at time t_n is given by,

$$\bar{\mathbf{x}}_n^a = \bar{\mathbf{x}}_n^f + \mathbf{K}_n(\mathbf{y}_n - \mathcal{H}_n(\bar{\mathbf{x}}_n^f)), \tag{6}$$

where $\bar{\mathbf{x}}_n^a$ is the analysis ensemble mean and $\bar{\mathbf{x}}_n^f$ is the forecast ensemble mean. The possibly non-linear observation operator $\mathcal{H} : \mathbb{R}^{N^p} \to R^{N^m}$ maps the state space to the observation space . The Kalman gain matrix,

$$\mathbf{K}_n = \mathbf{P}_n^f \mathbf{H}_n^T (\mathbf{H}_n \mathbf{P}_n^f \mathbf{H}_n^T + \mathbf{R}_n)^{-1}, \tag{7}$$

is a matrix of size $N^m \times N^p$ where $\mathbf{H}_n$ is the observation operator linearised about the background state. The observation error covariance matrix is denoted by $\mathbf{R}_n \in \mathbb{R}^{N^p \times N^p}$ and $\mathbf{P}_n^f \in \mathbb{R}^{N^m \times N^m}$ is the forecast error covariance matrix. When the forecast error covariance is derived from climatological data and assumed static, it is often denoted as $\mathbf{B}_n$ and known as the background error covariance matrix.

Previously it has been assumed that the observation error covariance matrix $\mathbf{R}$ is diagonal. However, with recent work showing that $\mathbf{R}$ is correlated and state dependent, it is important to be able to gain accurate estimates of the observation error covariance matrix. Here we propose a new method that combines the DBCP diagnostic with an ensemble transform Kalman filter (ETKF) [Bishop et al., 2001, Livings et al., 2008], to provide an estimate of time varying correlation observation error matrices that can be used within the assimilation scheme. We begin by describing the diagnostic proposed in Desroziers et al. [2005].

2.1 The DBCP diagnostic

The DBCP diagnostic described in Desroziers et al. [2005] makes use of the background (forecast) and analysis innovations to provide an estimate of the observation error covariance matrix. The background innovation, $\mathbf{d}^b = \mathbf{y} - \mathcal{H}(\mathbf{x}^f)$, is the difference between the observation $\mathbf{y}$ and the mapping of the forecast vector, $\mathbf{x}^f$, into observation space by the observation operator $\mathcal{H}$. The analysis innovations, $\mathbf{d}^a = \mathbf{y} - \mathcal{H}(\mathbf{x}^a)$, are similar to the the background innovations, but with the forecast vector replaced by the analysis vector $\mathbf{x}^a$. Making the tangent linear hypothesis on the observation operator, taking the outer product of the analysis and background innovations and assuming that the forecast and observation errors are uncorrelated results in

$$E[\mathbf{d}^a \mathbf{d}^{bT}] \approx \mathbf{R}. \tag{8}$$

This is valid if the observation and forecast errors used in the gain matrix,

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^T (\mathbf{H}\mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1}, \tag{9}$$

to calculate the analysis, are the exact observation and forecast errors [Desroziers et al., 2005]. However, provided that the correlation length-scales in $\mathbf{P}^f$ and $\mathbf{R}$ are sufficiently different, it shown that a reasonable estimate of $\mathbf{R}$ can be obtained even if the $\mathbf{R}$ and $\mathbf{P}^f$ used in $\mathbf{K}$ are not correctly specified. It has also been shown that the method can be iterated to estimate $\mathbf{R}$ [Mènard et al., 2009, Desroziers et al., 2009].

Much of the previous work using the DBCP diagnostic to estimate observation error covariance matrices has considered variational data assimilation methods [Bormann and Bauer, 2010, Bormann et al., 2010, Stewart et al., 2009, Stewart, 2010, Stewart et al., 2013b, Weston, 2011]. As ensemble filters and hybrid methods are becoming more important in operational data assimilation [Buehner et al., 2010, Miyoshi et al., 2010, Clayton et al., 2012] we consider the use of the DBCP diagnostic when using an ensemble data assimilation method with flow dependent forecast error statistics.

2.2 The ensemble transform Kalman filter with R estimation

We first give a brief overview of the proposed method, the ensemble transform Kalman filter with $\mathbf{R}$ estimation (ETKFR), before discussing it in further detail. The idea is to estimate the observation error covariance matrix within the ETKF. We use the ETKF to provide the samples of the background and analysis innovations to be used in the DBCP diagnostic. After the initial ensemble members, forecast error covariance matrix and observation error covariance matrix are specified, the filter is split into two stages, a spin-up phase where the matrix $\mathbf{R}$ remains static and a 'varying estimate' stage. The spin-up stage is run for a predetermined number of steps, N^s, and is an application of the standard ensemble transform Kalman filter. In the second stage the DBCP diagnostic is used to provide a new estimate of $\mathbf{R}$ that is then used within the assimilation. We note that for any assimilation that is running continuously the spin-up stage need only be run once to determine the initial samples required to estimate $\mathbf{R}$. We now present in detail the method that we have developed. Here the observation operator, $\mathbf{H}$, is chosen to be linear, but the method could be extended to account for a non-linear observation operator $\mathcal{H}$ (e.g. Evensen [2003]).

Initialisation - Begin with an initial ensemble $\left\{\mathbf{x}_0^i\right\}$ for $i = 1 \dots N$ at time $t = 0$ that has an associated initial covariance matrix $\mathbf{P}_0^f$. Also assume an initial estimate of the observation error covariance matrix $\mathbf{R}_0$; it is possible that this could just consist of the instrument error.

Step 1 - The first step is to use the full non-linear model, $\mathcal{M}_n$, to forecast each ensemble member, $\mathbf{x}_{n+1}^{f,i} = \mathcal{M}_n(\mathbf{x}_n^{a,i})$.

Step 2 - The ensemble mean and covariance are calculated using (3) and (5).

Step 3 - Using the ensemble mean the background innovations at time t_n, calculate and store $\mathbf{d}_n^b = \mathbf{y}_n - \mathbf{H}\bar{\mathbf{x}}_n^f$.

Step 4 - The ensemble mean is updated using,

$$\bar{\mathbf{x}}_n^a = \bar{\mathbf{x}}_n^f + \mathbf{K}_n(\mathbf{y}_n - \mathbf{H}_n\bar{\mathbf{x}}_n^f), \tag{10}$$

where $\mathbf{K}_n$ is the Kalman gain $\mathbf{K}_n = \mathbf{X'}_n^f \mathbf{Y'}_n^{fT} \mathbf{S}_n^{-1}$ of size $N^m \times N^p$. Here $\mathbf{S}_n$ is the invertible matrix $\mathbf{S}_n = \mathbf{Y'}_n^f \mathbf{Y'}_n^{fT} + \mathbf{R}_n$, where $\mathbf{Y'}_n^f = \mathbf{H}_n \mathbf{X'}_n^f$ is defined as the matrix containing the mapping of the ensemble perturbations into observation space.

Step 5 - Rather than calculate the ensemble perturbation update explicitly, the analysis perturbations are calculated as

$$\mathbf{X'}_n^a = \mathbf{X'}_n^f \boldsymbol{\Upsilon}_n, \tag{11}$$

where $\boldsymbol{\Upsilon}_n$ is the symmetric square root of $(\mathbf{I} - \mathbf{Y'}_n^{fT} \mathbf{S}_n^{-1} \mathbf{Y'}_n^f)$ [Livings et al., 2008].

Step 6 - The analysis mean is then used to calculate the analysis innovations, $\mathbf{d}_n^a = \mathbf{y}_n - \mathbf{H}(\bar{\mathbf{x}}_n^a)$.

Step 7 - If $n > N^s$, where N^s is the specified sample size, update $\mathbf{R}$ using

$$\mathbf{R}_{n+1} = \frac{1}{N^s - 1} \sum_{c=n-N^s+1}^{c=n} \mathbf{d}_c^a \mathbf{d}_c^{bT}. \tag{12}$$

Then symmetrise the matrix, $\mathbf{R}_{n+1} = \frac{1}{2}(\mathbf{R}_{n+1} + \mathbf{R}_{n+1}^T)$. Otherwise keep $\mathbf{R}_{n+1} = \mathbf{R}_0$.

Many of the steps in the proposed method are identical to the ETKF. Step 7, along with the storage of the background and analysis innovations in steps 3 and 6, are the additions to the ETKF that provide the estimate of the observation error covariance matrix.

After initialisation, the method consists of two stages, a static $\mathbf{R}$ stage and a 'varying estimate' stage. In the spin-up phase the algorithm is an application of the standard ETKF. In this stage the method is iterated for a number of assimilation steps N^s. Once these assimilation steps are completed, a new time-averaged estimate of $\mathbf{R}$ is calculated using the DBCP diagnostic from equation (8),

$$\mathbf{R}_{n+1} = \frac{1}{N^s - 1} \sum_{c=n-N^s+1}^{c=n} \mathbf{d}_c^a \mathbf{d}_c^{bT}. \tag{13}$$

Once this initial estimate of $\mathbf{R}$ has been calculated, the method moves into the 'varying estimate' stage. In this stage it is possible to include and update the estimate of $\mathbf{R}$ used within the assimilation. We continue running the ETKF using the updated $\mathbf{R}_n$ in place of our initial guess for $\mathbf{R}$. After the forecast and analysis stages we calculate a new estimate for the observation error covariance matrix $\mathbf{R}$ by removing the oldest samples for $\mathbf{d}^b$ and $\mathbf{d}^a$ and replacing them with those calculated in the current assimilation step.

At every assimilation step $\mathbf{R}$ is updated using the latest information, with the oldest information being discarded. Although this does not give a completely time-dependent estimate of $\mathbf{R}$ it should give a slowly time varying estimate that should take into account the most recent information relating to the observations. We now explain how the method is tested.

3 Experimental set up

3.1 The models

In order to demonstrate the ETKFR method we use two different models, the Lorenz '96 model [Lorenz, 1996, Lorenz and Emanuel, 1998] and the Kuramoto-Sivashinsky equation Kuramoto [1978], Sivashinsky [1977].

3.1.1 The Lorenz '96 model

The Lorenz '96 model is a method that has been widely used to test state estimation problems [Anderson, 2001, Ott et al., 2004, Fertig et al., 2007]. The model emulates the behaviour of a meteorological variable around a circle of latitude. The model consists of N^m variables $X_1, \ldots, X_{N^m}$ on a cyclic boundary, that is $X_{-1} = X_{N^m-1}$, $X_0 = X_{N^m}$ and $X_{N^m+1} = X_1$, which are governed by,

$$\frac{dX_j}{dt} = X_{j-1}(X_{j+1} - X_{j-2}) - X_j + F. \tag{14}$$

The first term on the right hand side simulates advection, whereas the second simulates diffusion, and the third is a constant forcing term. The solution exhibits chaotic behaviour for $N^m \geq 12$ and $F > 5$ and to ensure that we see chaotic behaviour in our solution we choose $N^m = 40$ and $F = 8$.

3.1.2 The Kuramoto-Sivashinsky equation

The Kuramoto-Sivashinsky (KS) equation,

$$\frac{\partial u}{\partial t} = -u\frac{\partial u}{\partial x} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^4 u}{\partial x^4}, \tag{15}$$

is a non-linear, non-dimensional partial differential equation where u is a function of time, t, and space, x. The equation produces complex behaviour due to the presence of the second and fourth order terms. The equation can be solved on both bounded and periodic domains and, when the domain is sufficiently large, the solutions exhibit multi-scale and chaotic behaviour [Gustafsson and Protas, 2010, Eguíluz et al., 1999]. This chaotic and multi-scale behaviour makes the KS equation a suitable low dimensional model that represents a complex fluid dynamic system. The KS equation has been used previously for the study of state estimation problems using both ensemble and variational methods [Protas, 2008, Jardak et al., 2000].

3.2 Twin experiments

To analyse the ensemble transform Kalman filter with **R** estimation (ETKFR) we run a series of twin experiments. We first describe how we calculate the observations, then we provide details for the experiments using the Lorenz '96 and Kuramoto-Sivashinsky models.

3.2.1 The observations

To create observations we must add errors from a specified distribution to the truth runs of the models. We choose $\mathbf{R}^t = \mathbf{R}^D + \mathbf{R}^C$ to consist of uncorrelated errors $\mathbf{R}^D = \sigma_D^2 I$, where σ_D^2 is the diagonal error variance, and correlated errors $\mathbf{R}^C = \sigma_C^2 C$, where σ_C^2 is the correlated error variance and C is a correlation matrix. We use direct observations with added uncorrelated observation error, which are calculated by adding pseudo-random samples from $\mathcal{N}(0, \sigma_D^2 \mathbf{I})$ to the values of the truth. We then add correlated error to our observations. As the correlation function we use the SOAR function,

$$\rho(i,j) = \{\cos(2ba\sin(\frac{\theta_{i,j}}{2}) + \frac{\sin(2ba\sin(\frac{\theta_{i,j}}{2}))}{Lb}\} e^{\frac{-2a\sin(\frac{\theta_{i,j}}{2})}{L}}, \tag{16}$$

where ρ is the correlation between two points i and j on a circle and $\theta_{i,j}$ is the angle between them [Thiebaux, 1976]. The constants L, b and a determine the lengthscale of the correlation function. We choose the SOAR function to approximate our correlated error because at large correlation length-scales the SOAR resembles the observation error covariance structure found in Bormann et al. [2002]. The SOAR function is used to determine a circulant covariance matrix C. To specify a true correlated observation error covariance matrix we multiply the circulant matrix by the correlated error variance which is chosen to be σ_C^2. The true observation error covariance matrix, $\mathbf{R}^t$, is obtained by adding the diagonal and correlated error covariance matrices $\mathbf{R}^D$ and $\mathbf{R}^C$. Having a specified observation error covariance matrix allows us to determine how well the method is working as the estimated matrix can be compared to the truth.

3.2.2 The Lorenz '96 model

In this study we solve the system of equations using MATLAB's (version R2008b) `ode45` solver, which is based on an explicit Runge-Kutta formula [Dormand and Prince, 1980]; this uses a relative error tolerance of 1e-3 and an absolute error tolerance of 1e-6. To generate the true solution the Lorenz equations are started from initial conditions where $X_j = 8, j = 1 \ldots 40$ with a small perturbation of 0.001 added to variable X_{20}. The numerical model provides output at intervals of $\Delta t = 0.01$ until a final time of $T = 50$. A slightly perturbed initial condition, created by adding pseudo-random samples from the distribution $\mathcal{N}(0, \sigma_b^2 \mathbf{I})$, where σ_b^2 is the forecast error variance, to the true initial condition, is used. From this initial condition the $N = 500$ ensemble members are created by adding pseudo-random samples from the initial forecast error distribution, which is chosen to also be $\mathcal{N}(0, \sigma_b^2 \mathbf{I})$. A large number of ensemble members is used to minimise the risk of ensemble collapse and to help obtain an accurate forecast error covariance matrix. For the purposes of this initial study we wish to avoid using the techniques of covariance inflation and localisation so as not to contaminate the estimate of $\mathbf{R}$. We take 20 equally spaced direct observations, calculated as described in Section 3.2.1, at each assimilation step. Constants for equation (16) are chosen to be $L = 6$ and $b = 3.6$. We then consider time-dependent $\mathbf{R}$ where b varies linearly with time according to $b(t) = \alpha t + \beta$. The frequency varies between experiments, with the chosen frequencies being observations available every 5 and 30 time steps, that is every 0.05 and 0.3 time units respectively.

3.2.3 The Kuramoto-Sivashinsky equation

In this study the KS equation is solved using an exponential time differentiating Runge-Kutta 4 (ETDRK4) numerical scheme. Details of this scheme, along with code to solve the

KS equation are given in Cox and Matthews [2000] and Kassam and Trefethen [2005]. The truth is defined by the solution to the KS equation on the periodic domain $0 \leq x \leq 32\pi$ from initial conditions $u = \cos(\frac{x}{16})(1 + \sin(\frac{x}{16}))$ until time $T = 10000$, using $N = 256$ spatial points and a time step of $\Delta t = 0.25$. The assimilation model is run at the same spatial and temporal resolution as the truth with $\Delta t = 0.25$ and $N = 256$. We use a slightly perturbed initial condition, created by adding pseudo-random samples from the distribution $\mathcal{N}(0, \sigma_b^2 \mathbf{I})$ to the true initial condition. For the KS equation we choose to use $N = 1000$ ensemble members as we are estimating a large number of state variables. The $N = 1000$ ensemble members are created by adding pseudo-random samples from the initial forecast error distribution, which is chosen to be $\mathcal{N}(0, \sigma_b^2 \mathbf{I})$, to the initial condition. We take 64 equally direct observations, calculated as described in Section 3.2.1, at each assimilation step. Constants for equation (16) are chosen to be $L = 15$ and $b = 3.8$; for some experiments b is chosen to vary linearly in time according to $b(t) = \alpha t + \beta$. The frequency varies between experiments, with the chosen frequencies being observations available every 40 and 100 time steps.

To show the effect of the method on the analysis we also run the ETKF without the $\mathbf{R}$ estimation. In this case the $\mathbf{R}$ used in the assimilation consists of the diagonal of the true observation error covariance matrix, $\mathbf{R}_0 = diag(\mathbf{R}^t)$.

We note that the ETKFR has been also been tested with different frequencies of observations in both time and space. The results presented here have been selected to demonstrate the different behaviours of the method under certain conditions. The method was also run with fewer ensemble members and appears to work well so long as the filter does not diverge.

4 Results

We present the results from all our experiments in Tables 1 and 2. We give details of the matrix used as the true observation error covariance matrix $\mathbf{R}^t$. Tables 1 and 2 show whether the ensemble transform Kalman filter with $\mathbf{R}$ estimation (ETKFR) is used, or whether we run a standard ETKF using a diagonal observation error covariance. We also provide the frequency of the observations and the variances for the initial forecast, diagonal and correlated error variances. We give the time-averaged analysis root mean square error (RMSE). The analysis RMSE at each time is calculated using,

$$RMSE = \sqrt{\frac{\sum_{j=1}^{N^m} ([\mathbf{x}_n^a]_j - [\mathbf{x}_n^t]_j)^2}{N^m}}, \tag{17}$$

where $[\mathbf{x}_n^t]_j$ is the j^{th} element of the true state vector at time t_n and $[\mathbf{x}_n^a]_j$ is the j^{th} element of the analysis at time t_n. The RMSEs at every analysis step are averaged to give the time-averaged analysis RMSE. The time-averaged analysis RMSE allows us to compare the performance of the filter for each experiment. To provide an idea of how well the ETKFR is performing we also give the RMSE of the covariance estimate at the final time. The RMSE of the row of the estimated observation error covariance matrix is calculated by comparing one row of the true observation error covariance matrix to the average calculated covariance structure. As the true observation error covariance matrix is isotropic and homogeneous the estimated observation correlation structure is calculated by averaging the permuted rows of the covariance matrix. The permuted rows are rows that have been shifted so that the variance lies in the same column. Rank histograms [Hamill, 2000] (shown in Waller [2013]) were considered to give information about the ensemble

spread. If the ensemble spread is not maintained the analysis and the estimation of the observation error covariance matrix may be affected.

4.1 Results with a static R and frequent observations

We begin by considering the case when the true observation error covariance matrix is static.

In Experiments 1L and 1K we use the standard EKTF for the assimilation. We begin by setting the true matrix $\mathbf{R}^t$ to $\mathbf{R}^t = \mathbf{R}^D + \mathbf{R}^C$, where $\mathbf{R}^D = 0.1\mathbf{I}$ and $\mathbf{R}^C = 0.1C$. The Lorenz '96 and KS models are each run for 1000 assimilation steps. We assume that $\mathbf{R}$ is diagonal, with $\mathbf{R}_0 = diag(\mathbf{R}^t)$. The standard ETKF is used to gain an estimate of $\mathbf{R}$ after the assimilation. The background and analysis innovations are calculated throughout the assimilation window. After the final assimilation these can be used to give an estimate of the true observation error covariance matrix. This allows us to check that the filter is working correctly and that the observation error covariance matrix can be calculated using the DBCP diagnostic.

The time-averaged analysis RMSEs for these experiments using the Lorenz and KS models are given in Tables 1 and 2, Experiments 1L and 1K. To show the performance of the DBCP diagnostic we plot in Figure 1 the estimated observation correlation structure. We see from the time-averaged analysis RMSE that the assimilation performs well and the rank histogram (not shown) suggests that the ensemble is well spread. From Figure 1 we see that even where it is assumed that $\mathbf{R} = diag(\mathbf{R}^t)$ the DBCP diagnostic still gives a good estimate of the true covariance with approximately correct length-scales. This suggests that even if a correlated observation error covariance is initially unknown it should be possible to use the DBCP diagnostic to estimate it. This is consistent with the results using 4D-Var [Stewart et al., 2009, 2013b, Bormann and Bauer, 2010, Bormann et al., 2010].

We now consider what happens where we estimate $\mathbf{R}$ within the ETKFR assimilation scheme described in Section 2 with $N^s = 100$ for the Lorenz model and $N^s = 250$ for the KS model. We find that with the chosen number of samples, for the estimated covariance matrix to be full rank, and hence used in the assimilation, it is necessary to compensate for the sampling error. To compensate for the sampling error we make the matrix isotropic and homogeneous by taking the mean of the shifted rows of the estimated matrix. The shifted rows of the matrix have been permuted so that the variance lies in the same column. This averaged row is then used to reconstruct a circulant matrix. This makes the assumption that all the observations have the same correlation structure. It may also be possible to overcome sampling error by increasing the number of samples; however, this reduces the time-dependence of the estimation and hence we choose to compensate for the sampling error by making the matrix isotropic and homogeneous. We verify that the method proposed is able to improve the analysis by including improved estimates of $\mathbf{R}$ in the assimilation scheme.

We begin by assuming that the initial observation error covariance matrix consists of only the diagonal error $\mathbf{R}^D$. The time-averaged analysis RMSEs for these experiments using the Lorenz and KS models are given in Tables 1 and 2, Experiments 2L and 2K. We see that the time-averaged analysis RMSEs are lower than Experiments 1L and 1K and this, together with an improved ensemble spread, shows that overall the assimilation scheme performs better than the case where the observation error covariance matrix is assumed diagonal. In Figure 2 we plot the true covariance (solid) as well as the first estimate of the covariance calculated using the first N^s background and analysis innovations (dashed) and the last

estimate of the covariance calculated using the last N^s background and analysis innovations (dot-dash). We note that as $\mathbf{R}$ is estimated using fewer samples than Experiments 1K and 1L we expect the results to be more noisy due to the increased sampling error. By comparison with Figure 1 we see that the first estimates of the covariance structure are similar to that calculated in Experiment 1L and 1K. This is because all the innovations used as samples were calculated assuming that $\mathbf{R} = \mathbf{R}^D$. We see that the last estimates of the covariance structure are closer to the true covariance structure. Overall the method performs well and this suggests that updating the estimate of $\mathbf{R}$ at each assimilation step in the ETKF improves the estimation of a static $\mathbf{R}$. It also suggests that it should be possible to gain a time-dependent estimate of correlated observation error.

4.2 Static R, infrequent observations

We now consider the case where observations are less frequently available. We begin by running the standard ETKF with $\mathbf{R} = diag(\mathbf{R}^t)$. Observations are available only every 30 timesteps (6 times less frequent) for the Lorenz '96 model and 100 time steps (2.5 times less frequent) for the KS model; therefore we have statistics from 166 and 400 assimilation steps. Again we calculate the background and analysis innovations throughout the assimilation and use these after the assimilation has ended to provide an estimate of the observation error covariance structure. As there are fewer assimilation steps there are also fewer background and analysis innovations that can be used to estimate $\mathbf{R}$.

The time-averaged analysis RMSE and covariance RMSE for these experiment are given in Tables 1 and 2, Experiments 3L and 3K. The time-averaged analysis RMSE suggests that the analyses are not as accurate as Experiments 1L and 2L and Experiments 1K and 2K. This is because with less frequent observations the model solutions have more time to diverge from the true solution before being constrained by the observations. However, the assimilation still provides a good analysis and the rank histograms suggest that the ensemble spread is maintained. For Experiment 3L the covariance function is still estimated well. However, the estimate of the covariance function for Experiment 3K is not as good as Experiment 1K. This suggests that the larger temporal spacing between observations may affect how well the DBCP diagnostic estimates $\mathbf{R}$.

We now estimate $\mathbf{R}$ within the ETKFR scheme and then reuse the estimated $\mathbf{R}$ at the next assimilation step. As our initial error covariance we choose the diagonal error covariance, $\mathbf{R} = \mathbf{R}^D$. We see from Experiment 4L Table 1 and Experiment 4K Table 2 that the time-averaged analysis RMSE is lower than the case when $\mathbf{R}$ was assumed diagonal and fixed throughout the assimilation. We now consider if the ETKFR can give an improved estimate of the covariance structure. Considering the covariance estimate RMSE we see that for Experiment 4L the method works well and a good estimate of the the covariance matrix is provided (not shown). We find that for Experiment 4K the structure of the covariance function is improved from the initial diagonal, but does not match the truth as closely as in experiment 2K. This is partly because some of the background and analysis innovations were calculated using the diagonal $\mathbf{R}$. As we are considering a static observation error matrix we expect the estimated $\mathbf{R}$ to improve with every assimilation step. If the assimilation is run for longer period of time, and the estimate of $\mathbf{P}^f$ is accurate, we would expect the estimated $\mathbf{R}$ to converge to the truth [Mènard et al., 2009]. We now return to the the case of more frequent observations, but consider a time-dependent true $\mathbf{R}$.

4.3 Time dependent R

We now consider the case when the true $\mathbf{R}$ is time-dependent and only consider the scheme where $\mathbf{R}$ is estimated and used within the assimilation. We choose the correlation to be the SOAR function as described by equation (16). To create time-dependence we vary the length-scale with time according to $b(t) = \alpha t + \beta$, the chosen α and β are detailed in Tables 1 and 2. We set the variance of the correlated error matrix to $\sigma_C^2 = 0.1$. We show that it is possible to use the ETKF and DBCP diagnostic to estimate a time varying observation error covariance matrix. To show how well the filter is performing we give the time-averaged analysis RMSE in Tables 1 and 2 Experiments 5L and 5K.

We see that the time-averaged analysis RMSE is low suggesting that the assimilation is working well and the rank histogram shown in Waller [2013] suggests that the ensemble spread is maintained. We now show how well the DBCP diagnostic estimates the true observation error covariance matrix. For Experiment 5K we plot the estimates at every 100 assimilation steps in Figure 3. We see that the first estimate of $\mathbf{R}$ captures the true correlation structure well. Considering the estimates at each of the times plotted we see that the true correlation structure is well approximated. The ETKF with $\mathbf{R}$ estimation gives a good estimate of a slowly time-varying observation error covariance matrix. Results for the Lorenz '96 model support these conclusions. Experiments 6L and 6K show that the method also performs well when the covariance length-scale decreases with time.

We now consider how well the method performs when the magnitudes and ratios of the forecast, uncorrelated and correlated error variances are varied. Experiments 7L,7K, 8L, 8K, 9L, 9K, 10L and 10K use the time-varying observation error covariance matrix used in Experiments 5L and 5K with σ_b^2, σ_D^2, σ_C^2 all chosen to be one of 0.01, 0.1 or 1. In each of these experiments we see that the assimilation is providing good estimates of the true state and the correlated observation error matrix. This suggests that the initial magnitudes and ratio of the forecast and observation errors does not affect the performance of the method.

Finally we consider the case where the covariance structure in $\mathbf{R}$ varies more quickly with time. To show how well the filter is performing we give the time-averaged analysis RMSE and covariance RMSE in Tables 1 and 2 Experiments 11L and 11K. Again we see the filter is performing well. We find that for Experiment 11L the covariance structure is well estimated and closely resembles the truth. For Experiment 11K the estimate of the correlation structure is not as accurate; however, the variance is well estimated and the correlation length-scale is approximately correct (not shown).

5 Conclusions

For a data assimilation scheme to produce an optimal estimate of the state the error covariances associated with the observations and background must be well understood and correctly specified. As the observation errors have been found to be correlated and time-dependent it is necessary to determine if the observation error covariance matrix can be estimated within an assimilation scheme. In this work we have introduced an ensemble transform Kalman filter with observation error covariance matrix estimation. This is an ETKF where analysis and background innovations are calculated at each analysis time step and the most recent set of these innovations is used to estimate the matrix $\mathbf{R}$ using the DBCP diagnostic. This estimate of $\mathbf{R}$ is then used in the next assimilation step. The method has been developed to allow a slowly time-varying estimate of the observation error

covariance matrix to be calculated.

We showed it is possible to obtain a good estimate of $\mathbf{R}$ using the DBCP diagnostic. We then showed that estimating $\mathbf{R}$ within the ETKF worked well, with good estimates obtained, the ensemble spread maintained and the analysis RMSE reduced compared to the case where the matrix $\mathbf{R}$ is always assumed diagonal. We also showed that the method does not work as well where the observations are less frequent, although this may be dependent on the model. However the method still produces a reasonable estimate of $\mathbf{R}$, maintains the ensemble variance and the time-averaged analysis RMSE is lower than where a diagonal $\mathbf{R}$ is used.

We next considered a case where $\mathbf{R}$ varied slowly with time. We showed that the method worked well where the true $\mathbf{R}$ was defined to slowly vary with time. The time-averaged analysis RMSE was low and the ensemble spread was maintained. The estimates of the correlation structure were good, suggesting that the method is capable of estimating a slowly time-varying observation error covariance matrix. A case where the length-scale of the observation error covariance varied more quickly was also considered, and the ETKFR produced reasonable estimates of the observation error covariance matrix. We also showed that the ability of the method to approximate the correlation structure was not sensitive to the forecast error variances or the true magnitude of the observation error variance. This suggests that the method would be suitable to give a time-dependent estimate of correlated observation error. We note that the effectiveness of the method will depend on how rapidly the synoptic situation and hence correlated error is changing and how often observations are available. The correlated error will also be dependent on the dynamical system. For models designed to capture rapidly developing situations, where representativity error and hence correlated error is likely to change rapidly, assimilation cycling and observation frequency within the assimilation is expected to be more frequent and hence more data is available for estimating the observation error.

Acknowledgements

This work was funded by ESA, NERC as part of the National Centre for Earth Observation and the Met Office through a CASE studentship.

References

J. L. Anderson. An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, 129:2884–2903, 2001.

R. N. Bannister. A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances. *Quarterly Journal of the Royal Meteorological Society*, 134:1951–1970, 2008.

C. Bishop, B. Etherton, and S. Majumdar. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Monthly Weather Review*, 129:420–436, 2001.

N. Bormann and P. Bauer. Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. I: Methods and application to ATOVS data. *Quarterly Journal of the Royal Meteorological Society*, 136: 1036–1050, 2010.

N. Bormann, S. Saariene, G. Kelly, and J. Thepaut. The spatial structure of observation errors in atmospheric motion vectors from geostationary satellite data. *Monthly Weather Review*, 131:706–718, 2002.

N. Bormann, A. Collard, and P. Bauer. Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. II: Application to AIRS and IASI data. *Quarterly Journal of the Royal Meteorological Society*, 136:1051–1063, 2010.

M. Buehner. Error statistics in data assimilation: Estimation and modelling. In W. Lahoz, B. Khattatov, and R. Mènard, editors, *Data Assimilation Making Sense of Observations*, chapter 4, page 99. Springer, 2010.

M. Buehner, P. Houtekamer, C. Charette, H. Mitchell, and B. He. Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations. *Monthly Weather Review*, 138:1567–1586, 2010.

A. M. Clayton, A. C. Lorenc, and D. M. Barker. Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quarterly Journal of the Royal Meteorological Society*, 2012. Early View. doi: 10.1002/qj.2054.

S. M. Cox and P. C. Matthews. Exponential time differencing for stiff systems. *Journal of Computational Physics*, 176:430–455, 2000.

R. Daley. *Atmospheric Data Analysis*. Cambridge University Press, 1991.

G. Desroziers, L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, 131:3385–3396, 2005.

G. Desroziers, L. Berre, and B. Chapnik. Objective validation of data assimilation systems: diagnosing sub-optimality. *In Proceedings of ECMWF Workshop on diagnostics of data assimilation system performance, 15-17 June 2009*, 2009.

J. Dormand and P. Prince. A family of embedded Runge-Kutta formulae. *Journal of Computational and Applied Mathematics*, 6:19 – 26, 1980.

V. M. Eguíluz, P. Alstrøm, E. Hernández-García, and O. Piro. Average patterns of spatiotemporal chaos: A boundary effect. *Phys. Rev. E*, 59:2822–2825, 1999.

G. Evensen. The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, 53:343–367, 2003.

E. J. Fertig, J. Harlim, and B. R. Hunt. A comparative study of 4D-VAR and a 4D ensemble Kalman filter: perfect model simulations with Lorenz-96. *Tellus A*, 59:96–100, 2007.

J. Gustafsson and B. Protas. Regularization of the backward-in-time Kuramoto-Sivashinsky equation. *Journal of Computational and Applied Mathematics*, 234:398–406, 2010.

T. M. Hamill. Interpretation of rank histograms for varifying ensemble forecasts. *Monthly Weather Review*, 129:550–560, 2000.

S. B. Healy and A. A. White. Use of discrete Fourier transforms in the 1D-Var retrieval problem. *Quarterly Journal of the Royal Meteorological Society*, 131:63–72, 2005.

F. Hilton, A. Collard, V. Guidard, R. Randriamampianina, and M. Schwaerz. Assimilation of IASI radiances at european NWP centres. *In Proceedings of Workshop on the assimilation of IASI data in NWP, ECMWF, Reading, UK, 6-8 May 2009*, pages 39–48, 2009.

A. Hollingsworth and P. Lönnberg. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, 38A:111–136, 1986.

P. L. Houtekamer and H. L. Mitchell. Ensemble Kalman filtering. *Quarterly Journal of the Royal Meteorological Society*, 133:3260–3289, 2005.

T. Janjic and S. E. Cohn. Treatment of observation error due to unresolved scales in atmospheric data assimilation. *Monthly Weather Review*, 134:2900–2915, 2006.

M. Jardak, I. M. Navon, and M. Zupanski. Comparison of sequential data assimilation methods for the Kuramoto-Sivashinsky equation. *International Journal for Numerical Methods in Fluids*, 62:374–402, 2000.

A. Kassam and L. Trefethen. Fourth-order time-stepping for stiff pdes. *SIAM J. Sci. Computing*, 25:1214–1233, 2005.

Y. Kuramoto. Diffusion-induced chaos in reaction systems. *Progress of Theoretical Physics*, 64:346–367, 1978.

D. M. Livings, S. L. Dance, and N. K. Nichols. Unbiased ensemble square root filters. *Physica D*, 237:1021–1028, 2008.

E. Lorenz. Predictability: a problem partly solved. *In Predictability.Proc 1995. ECMWF Seminar.*, pages 1–18, 1996.

E. Lorenz and K. Emanuel. Optimal sites for supplementary weather observations: Simulations with a small model. *Journal of Atmospheric Sciences*, 55:399–414, 1998.

R. Mènard, Y. Yang, and Y. Rochon. Convergence and stability of estimated error variances derived from assimilation residuals in observation space. *In Proceedings of ECMWF Workshop on diagnostics of data assimilation system performance, 15-17 June 2009*, 2009.

T. Miyoshi, Y. Sato, and T. Kadowaki. Ensemble Kalman filter and 4D-Var intercomparison with the Japanese operational global analysis and prediction system. *Monthly Weather Review*, 138:2846–2866, 2010.

T. Miyoshi, E. Kalnay, and H. Li. Estimating and including observation-error correlations in data assimilation. *Inverse Problems in Science and Engineering*, 21:387–398, 2013.

E. Ott, B. Hunt, I. Szunyogh, A. Zimin, E. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, and J. Yorke. A local ensemble kalman filter for atmospheric data assimilation. *Tellus A*, 56:415–428, 2004.

B. Protas. Adjoint-based optimization of PDE systems with alternative gradients. *Journal of Computational Physics*, 227:6490–6510, 2008.

G. Sivashinsky. Non-linear analysis of hydrodynamic instability in laminar flames. *Acta Astronautica*, 4:1177–1206, 1977.

L. M. Stewart. *Correlated observation errors in data assimilation*. PhD thesis, University of Reading, 2010. http://www.reading.ac.uk/maths-and-stats/research/theses/maths-phdtheses.aspx.

L. M. Stewart, S. L. Dance, and N. K. Nichols. Correlated observation errors in data assimilation. *International Journal for Numerical Methods in Fluids*, 56:1521–1527, 2008.

L. M. Stewart, J. Cameron, S. L. Dance, S. English, J. R. Eyre, and N. K. Nichols. Observation error correlations in IASI radiance data. Technical report, University of Reading, 2009. Mathematics reports series, www.reading.ac.uk/web/FILES/maths/obs_error_IASI_radiance.pdf.

L. M. Stewart, S. L. Dance, and N. K. Nichols. Data assimilation with correlated observation errors: experiments with a 1-D shallow water model. *Tellus A*, 65, 2013a.

L. M. Stewart, S. L. Dance, N. K. Nichols, J. R. Eyre, and J. Cameron. Estimating interchannel observation-error correlations for IASI radiance data in the Met Office system. *Quarterly Journal of the Royal Meteorological Society*, 2013b. doi: 10.1002/qj.2211.

H. Thiebaux. Anisotropic correlation functions for objective analysis. *Monthly Weather Review*, 104:994–1002, 1976.

J. A. Waller. *Using observations at different spatial scales in data assimilation for environmental prediction*. PhD thesis, University of Reading, Department of Mathematics and Statistics, 2013. http://www.reading.ac.uk/maths-and-stats/research/theses/maths-phdtheses.aspx.

J. A. Waller, S. L. Dance, A. S. Lawless, N. K. Nichols, and J. R. Eyre. Representativity error for temperature and humidity using the Met Office UKV model. *Quarterly Journal of the Royal Meteorological Society*, 2013. Early View. DOI: 10.1002/qj.2207.

P. Weston. Progress towards the implementation of correlated observation errors in 4d-var. Technical report, Met Office, UK, 2011. Forecasting Research Technical Report 560.

J. S. Whitaker, T. M. Hamill, X. Wei, Y. Song, and Z. Toth. Ensemble data assimilation with the NCEP global forecast system. *Monthly Weather Review*, 136:463–482, 2008.

List of Figures

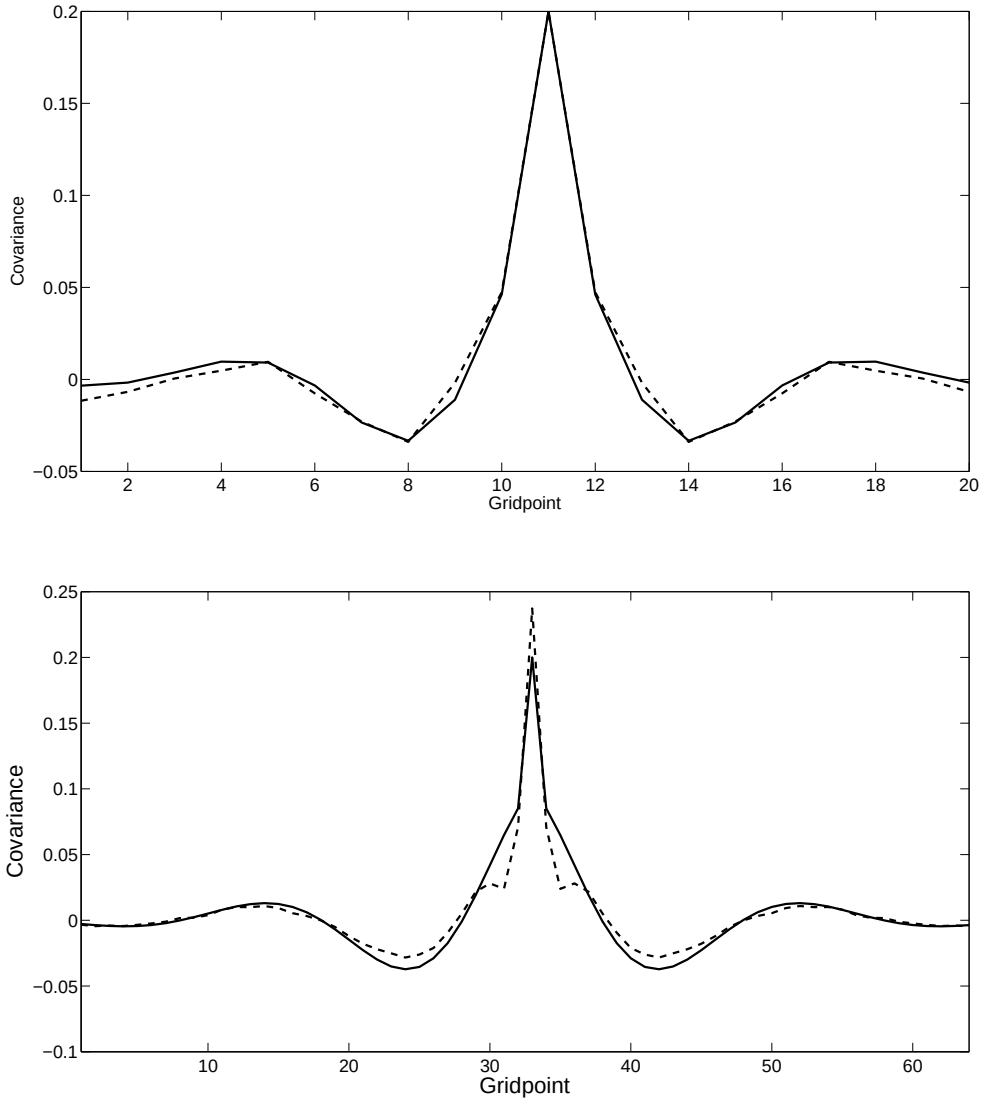


Figure 1 – Rows of the true (solid) and estimated (dashed) covariance matrices a) Experiment 1L. Observation error covariance RMSE: 0.002. b) Experiment 1K. Observation error covariance RMSE: 0.010.

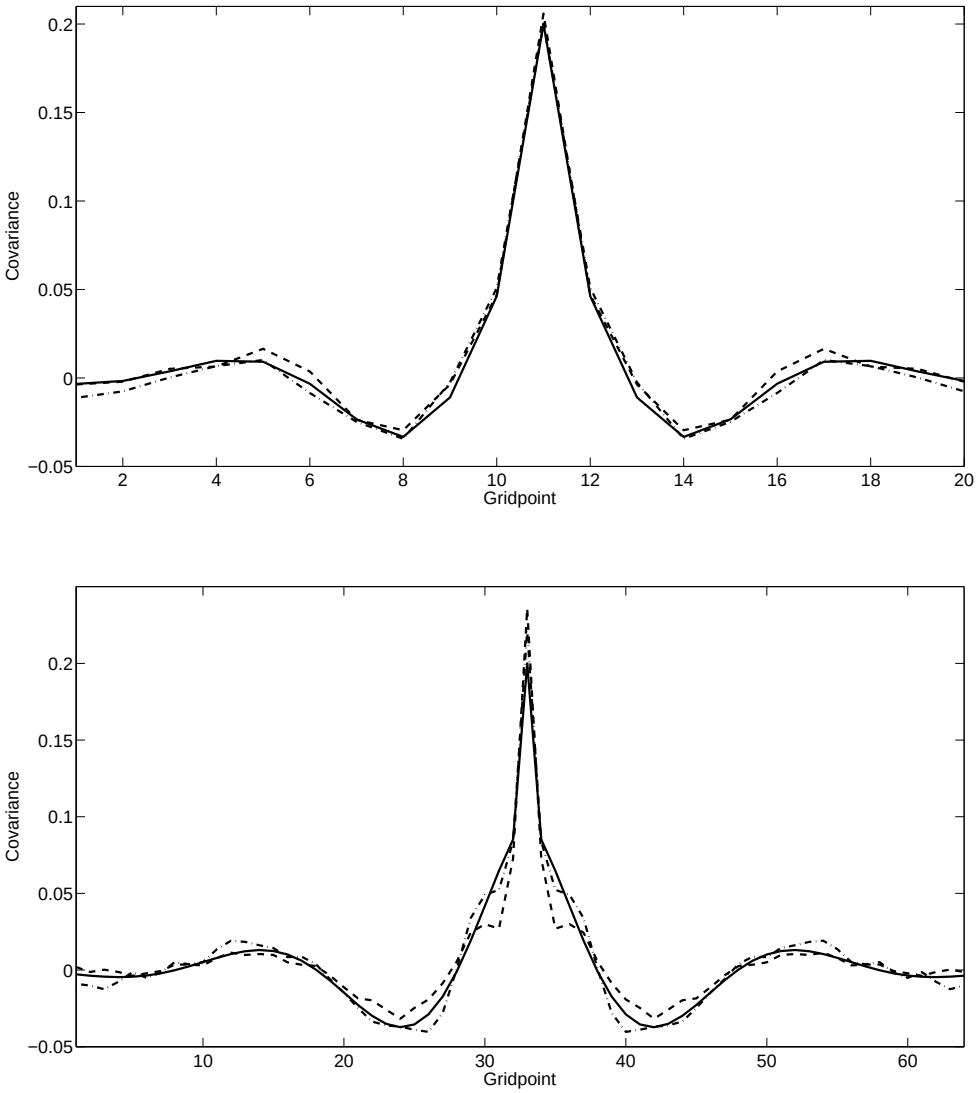


Figure 2 – Rows of the true (solid) and estimated covariance matrices a) Experiment 2L. Covariance calculated using the first 100 background and analysis innovations (dashed), observation error covariance RMSE: 0.007. Covariance calculated using the last 100 background and analysis innovations (dot-dashed), observation error covariance RMSE 0.004. b) Experiment 2K. Covariance calculated using the first 250 background and analysis innovations (dashed), observation error covariance RMSE 0.010. Covariance calculated using the last 250 background and analysis innovations (dot-dashed), observation error covariance RMSE: 0.006.

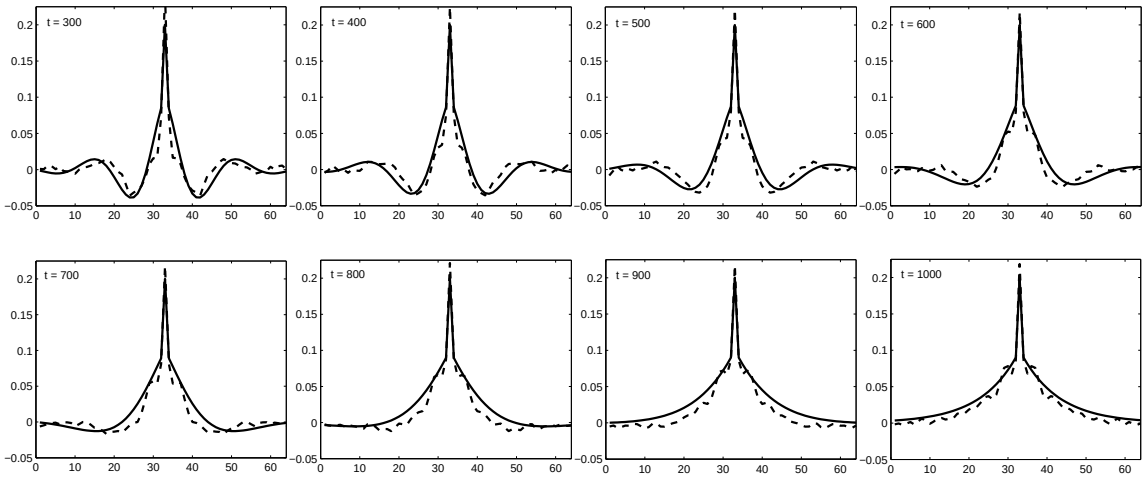


Figure 3 – Rows of the true (solid) and estimated (dashed) covariance matrices (covariance function plotted against observation point) every 100 assimilation steps from 300 to 1000 for Experiment 5K with a time-dependent $\mathbf{R}$, where b varies from 3.7 to 4.0, frequent observations and initial forecast, instrument and representativity error variances set to 0.1. Observation error covariance RMSE for final covariance estimate is 0.008.

List of Tables

Table 1 – Details of experiments executed using the Lorenz '96 model to investigate the performance of the ETKF with observation error covariance estimation

Exp. No.	True R	Assimilation Method	Obs Freq (time steps)	σ_b^2, σ_D^2, σ_C^2	Time Av analysis RMSE	Covariance RMSE
1L	SOAR + RI	ETKF ($\mathbf{R} = diag\mathbf{R}^t$)	5	0.1	0.115	0.005
2L	SOAR + RI	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	5	0.1	0.110	0.004
3L	SOAR + RI	ETKF ($\mathbf{R} = diag\mathbf{R}^t$)	30	0.1	0.065	0.008
4L	SOAR + RI	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	30	0.1	0.063	0.008
5L	Time dep. ($\alpha = -3 \times 10^{-4}, \beta = 3.6$)	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	5	0.1	0.106	0.010
6L	Time dep. ($\alpha = 3 \times 10^{-4}, \beta = 3.3$)	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	5	0.1	0.108	0.006
7L	Time dep. ($\alpha = -3 \times 10^{-4}, \beta = 3.6$)	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	5	0.01	0.033	0.001
8L	Time dep. ($\alpha = -3 \times 10^{-4}, \beta = 3.6$)	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	5	1.0	0.384	0.094
9L	Time dep. ($\alpha = -3 \times 10^{-4}, \beta = 3.6$)	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	5	0.1,1.0,1.0	0.414	0.095
10L	Time dep. ($\alpha = -3 \times 10^{-4}, \beta = 3.6$)	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	5	1.0,0.1,0.1	0.107	0.009
11L	Time dep. ($\alpha = 1 \times 10^{-3}, \beta = 3.6$)	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	5	0.1	0.108	0.009

Table 2 – Details of experiments executed using the Kuramoto-Sivashinsky equation to investigate the performance of the ETKF with observation error covariance estimation

Exp. No.	True R	Assimilation Method	Obs Freq (time steps)	σ_b^2, σ_D^2, σ_C^2	Time Av analysis RMSE	Covariance RMSE
1K	SOAR + RI	ETKF ($\mathbf{R} = diag\mathbf{R}^t$)	40	0.1	0.273	0.010
2K	SOAR + RI	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	40	0.1	0.251	0.010
3K	SOAR + RI	ETKF ($\mathbf{R} = diag\mathbf{R}^t$)	100	0.1	0.375	0.020
4K	SOAR + RI	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	100	0.1	0.357	0.021
5K	Time dep. ($\alpha = 3 \times 10^{-4}, \beta = 3.7$)	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	40	0.1	0.255	0.008
6K	Time dep. ($\alpha = -3 \times 10^{-4}, \beta = 4.0$)	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	40	0.1	0.252	0.014
7K	Time dep. ($\alpha = 3 \times 10^{-4}, \beta = 3.7$)	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	40	0.01	0.060	0.001
8K	Time dep. ($\alpha = 3 \times 10^{-4}, \beta = 3.7$)	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	40	1.0	0.704	0.044
9K	Time dep. ($\alpha = 3 \times 10^{-4}, \beta = 3.7$)	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	40	0.1,1.0,1.0	0.703	0.080
10K	Time dep. ($\alpha = 3 \times 10^{-4}, \beta = 3.7$)	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	40	1.0,0.1,0.1	0.261	0.010
11K	Time dep. ($\alpha = 4 \times 10^{-4}, \beta = 3.7$)	ETKFR ($\mathbf{R}_0 = \mathbf{R}^D$)	40	0.1	0.252	0.036