

# Comparing hybrid data assimilation methods on the Lorenz 1963 model with increasing non-linearity

By MICHAEL GOODLIFF\*, JAVIER AMEZCUA and PETER JAN VAN LEEUWEN, *Department of Meteorology, University of Reading, Reading, UK*

(Manuscript received 9 December 2014; in final form 26 March 2015)

## ABSTRACT

We systematically compare the performance of ETKF-4DVAR, 4DVAR-BEN and 4DENVAR with respect to two traditional methods (4DVAR and ETKF) and an ensemble transform Kalman smoother (ETKS) on the Lorenz 1963 model. We specifically investigated this performance with increasing non-linearity and using a quasi-static variational assimilation algorithm as a comparison. Using the analysis root mean square error (RMSE) as a metric, these methods have been compared considering (1) assimilation window length and observation interval size and (2) ensemble size to investigate the influence of hybrid background error covariance matrices and non-linearity on the performance of the methods. For short assimilation windows with close to linear dynamics, it has been shown that all hybrid methods show an improvement in RMSE compared to the traditional methods. For long assimilation window lengths in which non-linear dynamics are substantial, the variational framework can have difficulties finding the global minimum of the cost function, so we explore a quasi-static variational assimilation (QSVA) framework. Of the hybrid methods, it is seen that under certain parameters, hybrid methods which do not use a climatological background error covariance do not need QSVA to perform accurately. Generally, results show that the ETKS and hybrid methods that do not use a climatological background error covariance matrix with QSVA outperform all other methods due to the full flow dependency of the background error covariance matrix which also allows for the most non-linearity.

*Keywords:* data assimilation, hybrid methods, flow dependence

## 1. Introduction

Hybrid data assimilation methods are becoming more widely used in Numerical Weather Prediction (NWP). These methods combine ideas from successful variational methods such as 4DVAR (Talagrand and Courtier, 1987) and sequential methods such as the ensemble transform Kalman filter (ETKF, Bishop et al., 2001) and the ensemble Kalman smoother (e.g. Evensen and Van Leeuwen, 2000; Yang et al., 2012). The motivation behind hybrid methods is to make use of a flow-dependent background error covariance matrix ( $\mathbf{P}^b$ ) in a variational setting. Although some of these hybrid methods are being used operationally now, several basic questions on their performance are still open.

Hybrid methods were introduced by Zupanski (2005) when they produced the maximum likelihood ensemble filter (MLEF) which obtains a maximum a posteriori estimation (MAP) in ensemble space. Gu and Oliver (2007) explore a method called randomised maximum likelihood which is a 3DVAR in ensemble space with each ensemble member

based on stochastic EnKF. Both of these papers are examples of using variational methods within an ensemble Kalman filter. Other methods which use ensemble methods to update variational methods have also been popular. Wang et al. (2008) looked into a hybrid ETKF-3DVAR data assimilation method for the Weather Research and Forecasting (WRF) model. The ETKF-3DVAR uses ensemble information within a variational framework. The  $\mathbf{P}^b$  from the ETKF is combined with the climatological background error covariance matrix in 3DVAR ( $\mathbf{B}_c$ ), at the start of each assimilation window. This is our motivation for the ETKF-4DVAR which is the same but uses 4DVAR framework instead of 3DVAR. Other examples such as Fairbarn et al. (2014) and Liu et al. (2008) looked at the four-dimensional ensemble variational method, 4DENVAR. 4DENVAR uses the four-dimensional covariance from an ensemble of model trajectories which alleviates the need for the tangent linear and adjoint model in the 4DVAR. In Fairbarn et al. (2014), both deterministic and stochastic versions ensemble of 4DENVARs (EDA-D and EDA-S) are tested against 4DVAR, deterministic-EnKF [DetEnKF, which is an approximation of the ETKF for small background error

\*Corresponding author.  
email: m.goodliff@pgr.reading.ac.uk

covariances, Sakov and Oke (2008)] and 4DVAR-BEN. 4DVAR-BEN is similar to ETKF-4DVAR but uses the ETKF generated background error covariance matrix. The root mean square error (RMSE) is then calculated for different observation densities and a different range of ensemble sizes on the Lorenz 2005 model (which uses 180-variables). Liu et al. (2008) compare 4DENVAR to 4DVAR on a one-dimensional shallow water model. They show that 4DENVAR needs more iterations over the 4DVAR to reach the same value of the cost function. This is because the sample background error covariance in 4DENVAR, calculated by the ensemble, has a higher condition number than the fixed background error covariance matrix used in 4DVAR. The ensemble  $\mathbf{P}^b$  is a sample estimator that may contain spurious long-distance correlations (a result of finite sample size); this can increase the condition number. However, as the adjoint model is not needed in the minimisation, the computing time for each iteration is smaller. To evaluate the performance of 4DENVAR, the absolute error was calculated (comparing with the truth) and compared with that of 3DVAR, 4DVAR and the EnKF. 4DENVAR outperforms all methods on this model with only a small absolute error. This paper shows some encouraging results for 4DENVAR. It suggests that this method may be a good choice for real atmosphere data assimilation.

On the operational side, Wang (2010) introduced a hybrid background error covariance formulation into the 3DVAR-GSI (gridpoint statistical interpolation) system. Since then, there have been many ideas to use hybrid methods to try and decrease forecast error. Kuhl et al. (2013) give an example of hybrid methods in operational models. Their implementation is very similar to the ETKF-4DVAR which we used in this paper. They use a weak constraint 4DVAR system (named NAVDAS-AR), but their implementation of the background error covariance matrix is a combination of a static error covariance matrix with a flow-dependent error covariance matrix based on an ensemble transform technique of 80 members. They found that a hybrid blend of static and flow-dependent error covariances significantly reduces forecast error in comparison to just using a static error covariance.

In this work, we systematically compare the ETKF and the proposed ETKF-4DVAR from Wang et al. (2008), the 4DENVAR and 4DVAR-BEN from Fairbairn et al. (2014) with 4DVAR and the smoother described in Yang et al. (2012), which we will refer to as the ensemble transform Kalman smoother (ETKS) on the Lorenz 63 model. We use different window lengths with different observation periods and different ensemble sizes for all methods. We focus on the influence of non-linearity. By using the Lorenz 1963 model, we avoid issues related to localisation in the methods that use an ensemble. While localisation is an important factor in the performance of data assimilation

methods in large-dimensional systems, and, in the context of hybrid methods, especially the time-dependence of localisation in an assimilation window, we isolate the influence of hybrid background covariances and the non-linearity of the data assimilation problem in this study. On our longest assimilation window experiments, we also use a quasi-static variational assimilation algorithm [QSVa (Swanson et al., 1998), for the variational-based methods]. QSVa attempts to find the minimum of a non-convex cost function coming from a long assimilation window by solving a series of minimisations in shorter assimilation sections of increasing length.

The layout of this paper is as follows. In Section 2, we give a description of each one of the methods used in this paper. In Section 3, using the analysis RMSE as a metric, these methods have been compared considering assimilation (1) window length and observation interval size, and (2) ensemble size and observation interval size. The final section concludes the paper and gives plans for future work.

## 2. The hybrid methods

### 2.1. 4DVAR

The four-dimensional variational assimilation (4DVAR) is a method of estimating a set of state variables (or parameters). This is done by adjusting the model variables until the analysis trajectory balances the observations with the first guess. As the name would suggest, the 4DVAR method not only uses three-dimensional space, but also includes the time domain; i.e. the solution is a trajectory. If we consider the model to be perfect (strong constraint), this problem reduces to finding an initial condition  $\mathbf{x}$  from the minimisation of the following cost function:

$$J(\mathbf{x}) = \frac{1}{2} (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}_c^{-1} (\mathbf{x} - \mathbf{x}_b) + \frac{1}{2} \sum_{i=0}^p (\mathbf{y}_i - \mathcal{H}_i(\mathcal{M}_i(\mathbf{x})))^T \mathbf{R}_i^{-1} (\mathbf{y}_i - \mathcal{H}_i(\mathcal{M}_i(\mathbf{x})))^T \quad (1)$$

which measures the difference in the model trajectory and the observations. Here, for  $p$  observations,  $\mathcal{M}_i = \mathcal{M}_{0 \rightarrow i}$  is the non-linear model operator for time 0 to time  $i$ ,  $\mathbf{R}_i$  represents the observation error covariance matrix at time  $i$ ,  $\mathbf{x} \in \mathbb{R}^N$  is the initial state,  $\mathbf{y}_i \in \mathbb{R}^L$  is an observation set at time  $i$ ,  $\mathbf{x}_b$  is the background state, and  $\mathcal{H}_i$  represents the non-linear observation operator at time  $i$ . The background error covariance matrix is given as  $\mathbf{B}_c$ . To find a minimum

value, one can set the gradient of  $J(\mathbf{x})$  to zero and find the root of this equation iteratively

$$\nabla J(\mathbf{x}) = \mathbf{B}_c^{-1}(\mathbf{x} - \mathbf{x}_b) - \sum_{i=0}^p \mathbf{M}_i^\top \mathbf{H}_i^\top \mathbf{R}_i^{-1} (\mathbf{y}_i - \mathcal{H}_i(\mathcal{M}_i(\mathbf{x})))^\top = 0, \quad (2)$$

with  $\mathbf{M}$  being the linearised model around the current iteration and  $\mathbf{M}^\top$  its adjoint. If  $J(\mathbf{x})$  is not convex, this can be a local minimum and not global. Once the optimal initial conditions are reached, the model is run throughout the assimilation window to produce a forecast into the next window. This future forecast will provide the next assimilation windows background state. The version of 4DVAR which we use in our experiments is non-incremental 4DVAR.

*2.1.1. A quasi-static variational approach.* As the assimilation windows lengths increase on the Lorenz 1963 model, so does the non-linear error growth. This makes calculating the global minimum of the cost function even more difficult in the traditional variational framework. Pires et al. (1996) suggest using quasi-static variational assimilation, and this is then built upon in Swanson et al. (1998) and Swanson et al. (2000). This involves taking a smaller section of the assimilation window, using all observations in that section to find the minimum, and then gradually increasing the length of the section using the previous section minima as a first guess until we have covered the whole window.

## 2.2. Ensemble Transform Kalman Filter

The ETKF was introduced in Bishop et al. (2001) and modified in Wang et al. (2004). The latter is the version we use in our experiments. The ETKF is a type of ensemble square root filter (EnSRF, Tippett et al., 2003), in which the background ensemble of perturbations is transformed into the analysis via post-multiplication by a matrix of weights as:

$$\mathbf{X}^a = \mathbf{X}^b \mathbf{W}^a, \quad (3)$$

where  $\mathbf{X}^b$  is given by

$$\mathbf{X}^b = [\mathbf{x}^1 - \bar{\mathbf{x}}, \dots, \mathbf{x}^M - \bar{\mathbf{x}}]. \quad (4)$$

Here,  $\mathbf{X}^a, \mathbf{X}^b \in \mathbb{R}^{N \times M}$  are the analysis and background ensemble of perturbations respectively,  $M$  is the ensemble size and  $\mathbf{W}^a$  is a matrix of weights. This weight matrix is calculated as follows. The ensemble analysis error covariance matrix equation is

$$\mathbf{P}_e^a = (\mathbf{I} - \mathbf{K}_e \mathbf{H}) \mathbf{P}_e^b, \quad (5)$$

where  $\mathbf{K}_e = \mathbf{P}_e^b \mathbf{H}^\top (\mathbf{H} \mathbf{P}_e^b \mathbf{H}^\top + \mathbf{R})^{-1}$  is the Kalman gain matrix, and  $\mathbf{P}_e^a$  and  $\mathbf{P}_e^b$  are ensemble covariances, i.e.

$$\mathbf{P}_e^b = \frac{1}{M-1} \mathbf{X}^b \mathbf{X}^{b\top}, \quad (6)$$

Making a substitution of  $\mathbf{D} = \mathbf{Y}^b (\mathbf{Y}^b)^\top + \mathbf{R}$ , where  $\mathbf{Y}^b = \mathbf{H} \mathbf{X}_e^b$ , we get the equation

$$\mathbf{X}_e^a (\mathbf{X}_e^a)^\top = (\mathbf{I} - \mathbf{X}_e^b (\mathbf{X}_e^b)^\top \mathbf{H}^\top (\mathbf{D})^{-1} \mathbf{H}) \mathbf{X}_e^b (\mathbf{X}_e^b)^\top, \quad (7)$$

$$= \mathbf{X}_e^b (\mathbf{I} - (\mathbf{Y}^b)^\top (\mathbf{D})^{-1} \mathbf{Y}^b) (\mathbf{X}_e^b)^\top, \quad (8)$$

$$\mathbf{X}_e^a = \mathbf{X}_e^b (\mathbf{I} - (\mathbf{Y}^b)^\top \mathbf{D}^{-1} \mathbf{Y}^b)^{\frac{1}{2}}. \quad (9)$$

Use the Sherman-Morrison-Woodbury identity to find

$$\mathbf{I} - (\mathbf{Y}^b)^\top \mathbf{D}^{-1} \mathbf{Y}^b = (\mathbf{I} + (\mathbf{Y}^b)^\top \mathbf{R}^{-1} \mathbf{Y}^b)^{-1}. \quad (10)$$

The ETKF takes the eigenvalue decomposition

$$(\mathbf{Y}^b)^\top \mathbf{R}^{-1} \mathbf{Y}^b = \mathbf{C} \mathbf{\Gamma} \mathbf{C}^\top, \quad (11)$$

where  $\mathbf{\Gamma}$  is an  $M \times M$  real, non-negative, diagonal matrix and  $\mathbf{C}$  is an  $M \times M$  orthonormal matrix. Using eq. (11), eq. (10) can now be written as

$$\mathbf{I} - (\mathbf{Y}^b)^\top \mathbf{D}^{-1} \mathbf{Y}^b = \mathbf{C} (\mathbf{I} + \mathbf{\Gamma})^{-1} \mathbf{C}^\top. \quad (12)$$

$\mathbf{C} (\mathbf{I} + \mathbf{\Gamma})^{-\frac{1}{2}}$  satisfies eq. (5), but to avoid bias  $\mathbf{C}^\top$  must be included [see Wang et al. (2004) or any other orthonormal matrix Livings et al. (2008)]. This gives the weight matrix,  $\mathbf{W}^a$  for eq. (3),

$$\mathbf{W}^a = \mathbf{C} (\mathbf{I} + \mathbf{\Gamma})^{-\frac{1}{2}} \mathbf{C}^\top. \quad (13)$$

To optimise the performance of ETKF we use inflation such that  $\mathbf{X}^b = \mathbf{X}^b (1 + \rho)$ , where  $\rho \in \{0, 0.025, \dots, 0.375, 0.4\}$ . For each combination of ensemble size and observation period, an optimal was found. More information about the implementation of the ETKF can be found in the appendix of Amezcua et al. (2012).

## 2.3. Ensemble Transform Kalman Smoother

While filters modify (update) state variables at observation times, smoothers modify the whole trajectories of state variables with the information obtained from observations. The weights in the ETKF are used only at observation time. As each ensemble member arises from a forward integration from the previous observation time, these weight matrices should describe a closer estimation of the truth throughout the whole trajectory. So, this smoother applies the weight at each observation time (from the ETKF) and applies it to the ensemble trajectories between the last observation and the current one. This is similar to the 4D-EnKF by Hunt et al. (2004). This method does not require any extra model runs as this smoother simply multiplies the entire ETKF ensemble with an extra

weight matrix. The difference between the original ensemble Kalman smoother (Evensen and Van Leeuwen, 2000) and the smoother used in this paper is that we do not perturb the observations.

#### 2.4. ETKF-4DVAR

Wang et al. (2008) was the pilot study of the ETKF-3DVAR hybrid data assimilation method. In that paper, they conducted an identical-twin OSSE (Observing System Simulation Experiment) with a coarse grid spacing (200 km) of the WRF model, comparing the ETKF-3DVAR to traditional 3DVAR. They show that the hybrid provides an approximately 15–20% more accurate analysis than the 3DVAR. It was shown that the improvement in RMS analysis error of the hybrid over the 3DVAR was larger over data-sparse regions. They suggest that future work should involve looking into ETKF-4DVAR.

The ETKF-4DVAR uses a similar cost function to eq. (1) written as

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^\top \tilde{\mathbf{B}}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2} \sum_{i=0}^p (\mathbf{y}_i - \mathcal{H}_i(\mathcal{M}_i(\mathbf{x})))^\top \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathcal{H}_i(\mathcal{M}_i(\mathbf{x}))), \quad (14)$$

where all variables are the same as in the 4DVAR version except the background error covariance matrix,  $\mathbf{B}_c$  is replaced by  $\tilde{\mathbf{B}}$  which is given by the equation,

$$\tilde{\mathbf{B}} = \beta \mathbf{B}_c + (1 - \beta) \mathbf{P}^b. \quad (15)$$

The climatological background error covariance matrix  $\mathbf{B}_c$  is the same as the one used in 4DVAR, but  $\mathbf{P}^b$  represents the background error covariance matrix generated by the ETKF as in eqs. (6) and (4) and  $\beta \in [0, 1]$  represents a weighting scalar. When  $\beta$  is set to 1 in eq. (15), we find the traditional 4DVAR. When  $\beta = 0$ , only the ensemble error covariance matrix is used, and this method is known as 4DVAR-BEN, and when  $\beta = 0.5$ , the method will be referred to as the ETKF-4DVAR.

#### 2.5. 4DENVAR

4DENVAR is a variational method in which an ensemble of model trajectories is used to estimate space and time correlations, avoiding the need for an adjoint model following original ideas of Van Leeuwen and Evensen (1996). Further details on this method can be found in Liu et al. (2008), Fairbairn et al. (2014), Lorenc (2003, 2011),

Buehner et al. (2010a, 2010b), Wang and Lei (2014) and Tian et al. (2008). This method seeks to minimise the cost function

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^\top (\mathbf{P}^b)^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2} \sum_{i=0}^p (\mathbf{y}_i - \mathcal{H}_i \mathcal{M}_i(\mathbf{x}))^\top \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathcal{H}_i \mathcal{M}_i(\mathbf{x})), \quad (16)$$

which is similar to the four-dimensional variational data assimilation method (4DVAR) except  $\mathbf{B}_c$  has been replaced by  $\mathbf{P}^b$ , which is identical to the 4DVAR-BEN cost function. 4DENVAR differs in the minimisation, and the gradient of eq. (16) is almost identical to 4DVAR's without the need for a model adjoint. Our background covariance matrix is calculated by

$$\mathbf{P}_i^b = \frac{1}{M-1} \mathbf{X}_i \mathbf{X}_i^\top. \quad (17)$$

where

$$\mathbf{X}_i = [\mathbf{x}_i^1 - \tilde{\mathbf{x}}_i, \dots, \mathbf{x}_i^M - \tilde{\mathbf{x}}_i], \quad (18)$$

with  $\mathbf{x}_i^j$  represents the value of ensemble member  $j$  at time  $t$  and  $\tilde{\mathbf{x}}$  represents the recentred 4DVAR analysis. Thus, the cost function gradient can be written as

$$\nabla J(\mathbf{x}) = (\mathbf{P}_0^b)^{-1}(\mathbf{x} - \mathbf{x}_b) - \sum_{i=0}^p (\mathbf{P}_0^b)^{-1} \mathbf{P}_0^b \mathbf{M}_i^\top \mathbf{H}_i^\top \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathcal{H}_i \mathcal{M}_i(\mathbf{x})). \quad (19)$$

Using the transformation,

$$\mathbf{P}_0^b = \mathbf{X}_0 \mathbf{X}_0^\top, \quad (20)$$

$$\mathbf{P}_0^b \mathbf{M}_i^\top \mathbf{H}_i^\top = \mathbf{X}_0 \mathbf{X}_0^\top \mathbf{M}_i^\top \mathbf{H}_i^\top \quad (21)$$

$$= \mathbf{X}_0 (\mathcal{H}_i \mathcal{M}_i \mathbf{X}_0)^\top \quad (22)$$

$$= \mathbf{X}_0 (\mathcal{H}_i \mathbf{X}_i)^\top, \quad (23)$$

we can replace  $\mathbf{P}_0^b \mathbf{M}_i^\top \mathbf{H}_i^\top$  with  $\mathbf{X}_0 (\mathcal{H}_i \mathbf{X}_i)^\top$  to remove the need for the model and observation operator adjoints. This ensemble of model trajectories at the initial time step,  $t_0$ , is set to the same distribution as the analysis background error covariance matrix of our ETKF at the end of the previous assimilation window, this is then centred around the end point of the 4DENVAR analysis at the end of the previous assimilation window. From this description, it becomes clear that the method closely resembles an iterative variant of the ensemble smoother as presented in Van Leeuwen and Evensen (1996).

A brief summary of each method can be found in the Appendix 1.

### 3. Experiments

#### 3.1. Lorenz 1963

For our experiments, we use the Lorenz (1963) model. This is a simple dynamical model which exhibits chaotic behaviour for certain choices of parameters. The Lorenz equations are given by the non-linear coupled ODE system

$$\frac{dx}{dt} = -\sigma(x - y), \quad (24)$$

$$\frac{dy}{dt} = \rho x - y - xz, \quad (25)$$

$$\frac{dz}{dt} = xy - \beta z, \quad (26)$$

where  $x = x(t)$ ,  $y = y(t)$ ,  $z = z(t)$  are the state variables and  $\sigma$ ,  $\rho$  and  $\beta$  are parameters. In these experiments, they are chosen to have the values 10, 28 and  $8/3$ , respectively. To ensure we start from a point in the attractor, 20 consecutive 4DVARs were run over 50 time steps ( $\Delta t = 0.01$ ) and the final analysis point,  $\mathbf{x}_a$ , was taken to be the initial state,  $\mathbf{x}_0$ , of our experiments. Thus, in these experiments the initial state is

$$\mathbf{x}_0 = (-3.12346395, -3.12529803, 20.69823159)^\top.$$

#### 3.2. Generation of the background error covariance matrix

The climatological background error covariance matrix was generated by an iterative method using 4DVAR with one observation at the end of the assimilation window. Each assimilation window is the same length as the observation period used in each experiment, and these assimilation windows are run over a 5000 time step cycle starting from an arbitrary  $\mathbf{B}_c$ . This was thought to be long enough to let the system evolve and provide the appropriate correlations among variables. At the end of the cycle, we calculate a new  $\mathbf{B}_c$  matrix using,

$$\mathbf{B}_c = \overline{(\mathbf{x}^f - \mathbf{x}^t)(\mathbf{x}^f - \mathbf{x}^t)^\top}, \quad (27)$$

from  $t = 500$  to  $t = 5000$  to ignore the transient and where  $\mathbf{x}^f$  is the forecast state and  $\mathbf{x}^t$  is the truth state. With this new  $\mathbf{B}_c$ , we restart the cycle for another iteration. This is done for 10 iterations which is long enough to show convergence. This method is described in Yang et al. (2006). Figure 1 shows results from experiments with a fixed versus non-fixed  $\mathbf{B}_c$  for both traditional 4DVAR and 4DVAR with QSVAs. A fixed  $\mathbf{B}_c$  means that, for all experiments, the  $\mathbf{B}_c$  was generated using the method described above with observation period of 12 for each experiment. A non-fixed  $\mathbf{B}_c$  is generated using the method described above with the corresponding observation period taken as in each experiment. It can be seen that 4DVAR

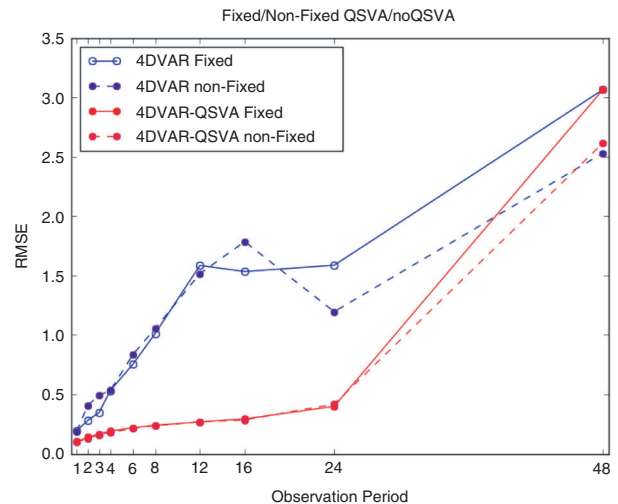


Fig. 1. Comparison of standard 4DVAR with 4DVAR-QSVA both fixed and non-fixed background error covariance.

(with and without QSVAs) with a fixed  $\mathbf{B}_c$  does just as well as with a non-fixed  $\mathbf{B}_c$  matrix in nearly all observation period sizes. For all our experiments, all variables ( $x$ ,  $y$  and  $z$ ) will be directly observed ( $\mathcal{H} = \mathbb{I}$ ), with an observation error covariance matrix of  $\mathbf{R} = \sigma^2 \mathbb{I} = \mathbb{I}$  (uncorrelated observations), with  $\sigma^2 = 1$ .

#### 3.3. Does the initial point change the result?

Before comparison of the hybrid methods are made, we test whether initial conditions are important for both the ETKF and 4DVAR. Four different initial states were selected from the Lorenz 1963 attractor, and then run over 10 000 assimilation windows, each window is 24 time steps using 4DVAR and the ETKF (with inflation).

Doing short data assimilation experiments may be misleading and give information of the performance of the method only in a specific neighbourhood of the attractor. Thus, we use 10 000 assimilation windows to fully model the state of the system. The initial points are given as:

$$\mathbf{x}_0^1 = (-3.12346395, -3.12529803, 20.69823159)^\top$$

$$\mathbf{x}_0^2 = (-12.2275757, -13.28328434, 28.50731193)^\top$$

$$\mathbf{x}_0^3 = (4.87127426, 8.78267131, 11.57329377)^\top$$

$$\mathbf{x}_0^4 = (14.7894135, 10.17205532, 39.60479722)^\top$$

Both methods are run with eight different observation periods [at every 1,2,3,4,6,8,12,24 time step(s)]. The ETKF has 20 ensemble members with the initial ensemble being generated from a Gaussian distribution perturbed from the initial state [with  $\mathbf{x}^t + \mathbf{B}_c^{1/2} \boldsymbol{\eta}$ , where  $\boldsymbol{\eta} \sim (\mathbf{0}, \mathbb{I})$ ]. After running these methods for this time scale, we notice that the RMSE are all equivalent for each method. This concludes that,

after a long enough time frame, the initial state of the experiment does not change the result.

### 3.4. Improvements from QSVa

In Fig. 1, the quasi-static variational approach (QSVa) experiments show the influence of QSVa for a 48 time step window for different observation periods. We first minimise over the first 12 time steps, then the first 24 time steps, then the first 36 time steps and finally the whole window. QSVa perform much better than 4DVAR, suggesting that local minima are present. Note that an observation period of 48 time steps, which has only one observation at the end of the window, 4DVAR and 4DVAR-QSVa are the same. This is because with one observation per window, the smaller sections which the QSVa method uses have no observations, thus not improving the first guess.

### 3.5. Ensemble size versus observation interval

For complex models, ensembles can be expensive to run. Hence, we first compare each method over 10 000 consecutive assimilation windows (each one being 24 time steps in length), observation periods of  $\Delta t = \{1, 2, 3, 4, 6, 8, 12, 24\}$  using an ensemble size of 5, 10, 20, 50 or 100 members for the ETKF. This will show if a large ensemble size is needed to give the most accurate estimation of the true state of the system. As 4DVAR is unaffected by changing the ensemble size of the ETKF, it is not necessary for comparison.

Fairbarn et al. (2014) show that an ensemble of at least five members is needed for EDA-D to have a lower analysis RMSE over 4DVAR in the perfect and imperfect model. Here, in Fig. 2, we show how accurate each hybrid method (along with the ETKF) is when the number of ensemble members is varied.

The true nature of ensemble methods is that more ensemble members will more accurately represent the systems behaviour, giving a smaller RMSE. Without inflation, we do see a decrease in RMSE with more ensemble members, but as inflation is used in this experimentation, this is not always the case. Looking at all panels in Fig. 2, it can be seen that for small observation intervals, ensemble size above 50 members produces the lowest RMSE for all methods except for the ETKF-4DVAR (which is the only method which uses the climatological background error covariance matrix). With a lot of ensemble members and a lot of observations, an ensemble size at 50 or above seems to provide the lowest RMSE until our observation period gets to six time steps. This is because the trajectory of the system is very closely observed. As the frequency of observations decreases, more ensemble members are expected to give a more accurate estimate of the state of the system. Nonetheless, in Fig. 2a we notice that for long observation periods

(e.g. 12, 24 in all panels), the largest RMSE actually corresponds to the largest ensemble ( $M = 100$ ). This counter-intuitive result has been observed (and explained) before in small models (e.g. Lorenz, 1963) under largely non-linear error growth (as it is the cast with infrequent observations). Lawson and Hansen (2004) and Anderson (2010) show that with large ensemble sizes using a deterministic ensemble Kalman filter, the system can suffer from ‘ensemble clustering’. This can be alleviated with random rotations (e.g. Pham et al., 1998; Amezcuca et al., 2012) but at the expense of losing information of the trajectory. After a certain amount of members, increasing the size will not increase accuracy. After an interval size of 12, ETKF suffers ensemble clustering causing a higher RMSE for 50 ensemble members. Figure 3 shows the degree of clustering given different observation periods. This is calculated by

$$CD = \frac{Tr(\mathbf{P}_{M-1})}{Tr(\mathbf{P}_M)}, \quad (28)$$

where  $\mathbf{P}_{M-1}$  is the ensemble spread ignoring the furthest member from the truth. As  $CD \rightarrow 0$ , we have a higher clustering of ensemble members. This is because as  $CD \rightarrow 0$ , this implies  $Tr(\mathbf{P}_{M-1}) \rightarrow 0$ ; hence, all the variance is coming from a single member. The figure shows that, as observation period increases, the probability of the ensemble clustering increases. With a small observation period, we have random ensemble clusters this turns into systematic clustering at higher observation periods. This is also the case with the ETKS (see Fig. 2b), but as the whole trajectory between observations is weighted, we see lower RMSE in comparison to Fig. 2a as the space between observations increases. Using this increase in accuracy, the method which seems to be affected the most is 4DENVAR in Fig. 2e. It can be seen that with a small ensemble size with the added iterational aspect of the variational framework, we get a decrease in RMSE in comparison to the ETKF. Lastly, Fig. 2e shows no significant change in RMSE with increased ensemble size except with single observation windows. This could be due to the variational aspect of the hybrid methods and how, in this system, the ETKF is quite accurate even with small ensemble sizes. Thus, any ensemble size over 5 would be enough for this method. In the following experiments, we will use 20 ensemble members as this seems to be the maximum amount before we suffer from clustering. In Lorenz 1963, we do not need to use localisation.

### 3.6. Assimilation window length and observation interval

In this section, a detailed comparison of ETKF-4DVAR (with  $\beta = 0.5$ ), 4DVAR-BEN, 4DENVAR, ETKF, ETKS

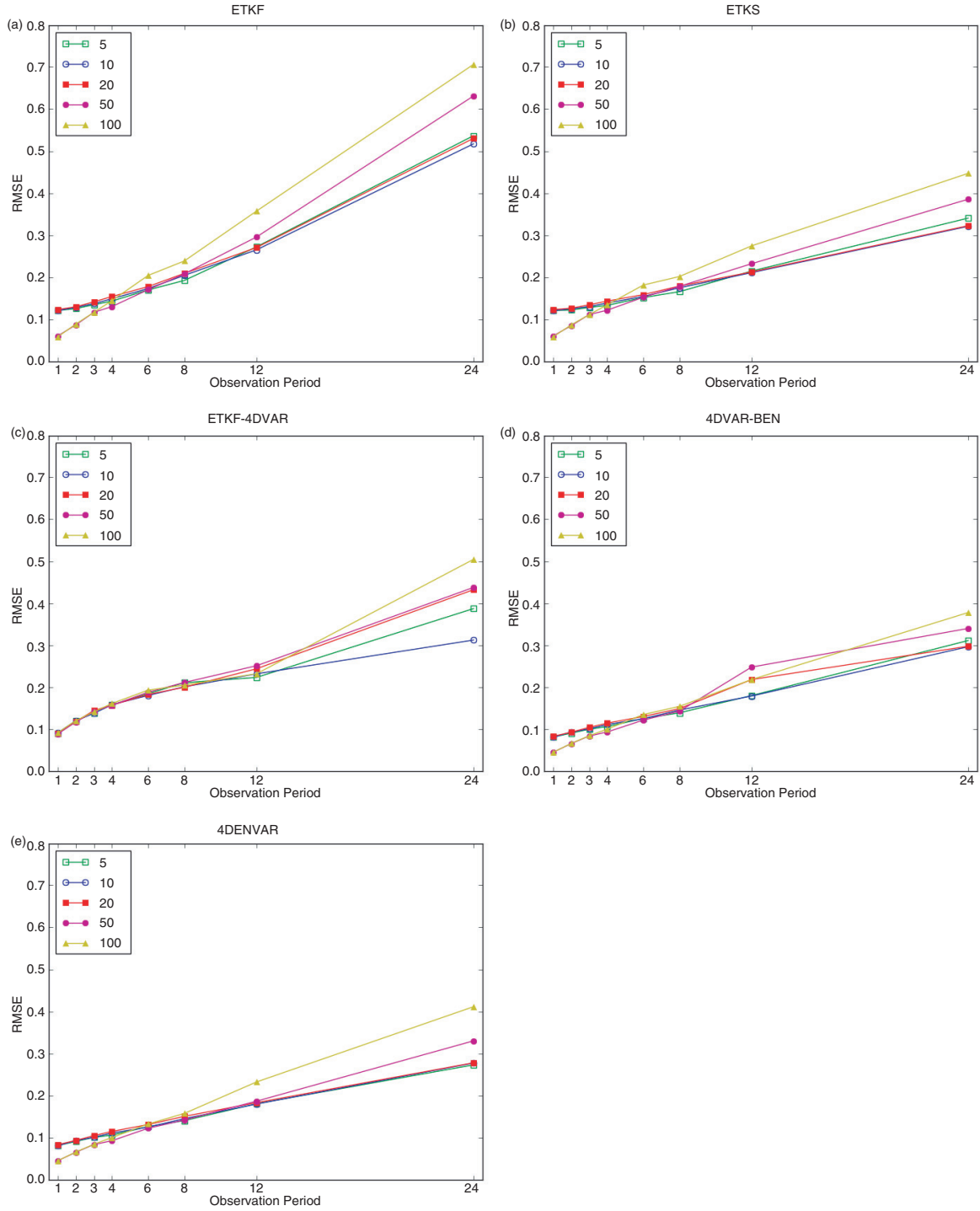


Fig. 2. RMSE (y-axis) for different hybrid methods (and the ETKF) as a function of ensemble size (different lines) and observation periods (x-axis). Each panel shows a different data assimilation method (Each y-axis has a different scale).

and 4DVAR over different assimilation window lengths, with different observation periods is made. In this experiment, windows of lengths 12, 24, 36 and 48 time steps are used over 10000 assimilation windows, we also used the quasi-static variational assimilation method for window

lengths of 36 and 48 time steps (QSVA only needs 1000 assimilation windows, since the variation in RMSE among cycles is lower than without QSVA). The observation interval sizes used are generated using the multiplication factors of the window length so as to give us an observation

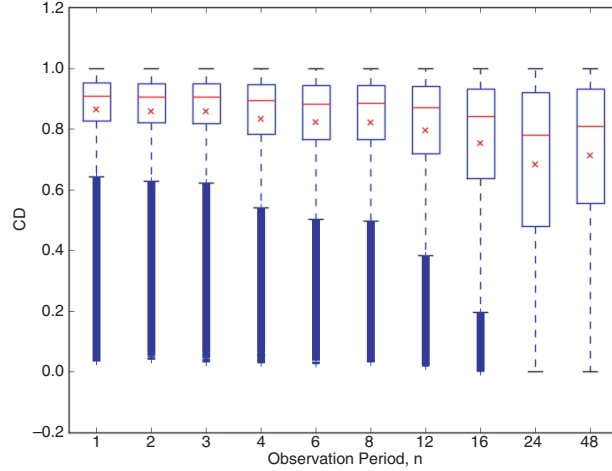


Fig. 3. Cluster Degree (CD, y-axis) for different observation periods for a 50 member ETKF with no inflation.

at the end of the assimilation window. For example, in a 48 time step assimilation window, we have 10 different observation periods, those are with an observation at every 1,2,3,4,6,8,12,16,24,48 time step(s). The ETKF will have 20 ensemble members with the initial ensemble being generated from a Gaussian distribution perturbed from the initial state [with  $\mathbf{x}^t + \mathbf{B}^t \eta$  where  $\eta \sim (\mathbf{0}, \mathbf{I})$ ].

Figure 4 shows the performance of the different methods for two assimilation window length (12 and 24 steps) and different observation periods. First, we notice that in a smaller assimilation window (Fig. 4a), the 4DVAR-BEN and 4DENVAR have the smallest RMSE. The mean RMSE is calculated from the equation,

$$\text{RMSE} = \frac{1}{T} \sum_{t=1}^T \sqrt{\frac{(\bar{\mathbf{x}}^a - \mathbf{x}^t)^\top (\bar{\mathbf{x}}^a - \mathbf{x}^t)}{N}}, \quad (29)$$

where  $\mathbf{x}^t$  is the true state of the system and  $\bar{\mathbf{x}}^a$  is the mean analysis state. Mean in ensemble methods refers to sample mean, and in variational methods, it refers to the trajectory which minimises the cost function and  $T$  is the window length. It can be seen that the two methods which use a climatological background error covariance matrix perform less accurately than the other methods for this small window length, except in the case where we have an observation at every time step. As the ETKF provides a flow-dependent background error covariance matrix (Fig. 5 shows ensemble background error covariance matrices before and after observation given different observation periods), 4DENVAR and 4DVAR-BEN give the lowest RMSE over small observation interval sizes. This is consistent with linear theory, the variational method gives the best linear unbiased estimator (BLUE) on the time window, and the  $\mathbf{P}^b$  part ensures an optimal  $\mathbf{B}$  matrix.

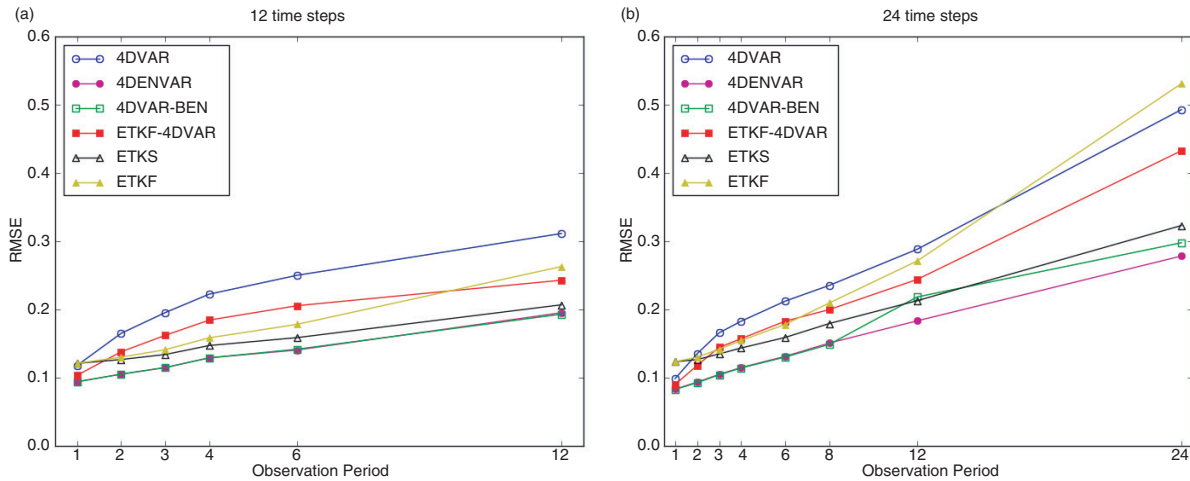


Fig. 4. Performance of each data assimilation method analysis as a function of window lengths and observation periods (x-axis); y-axis is the RMSE (scale in both panels are different). In Fig. 4a, 4DENVAR is under 4DVAR-BEN.



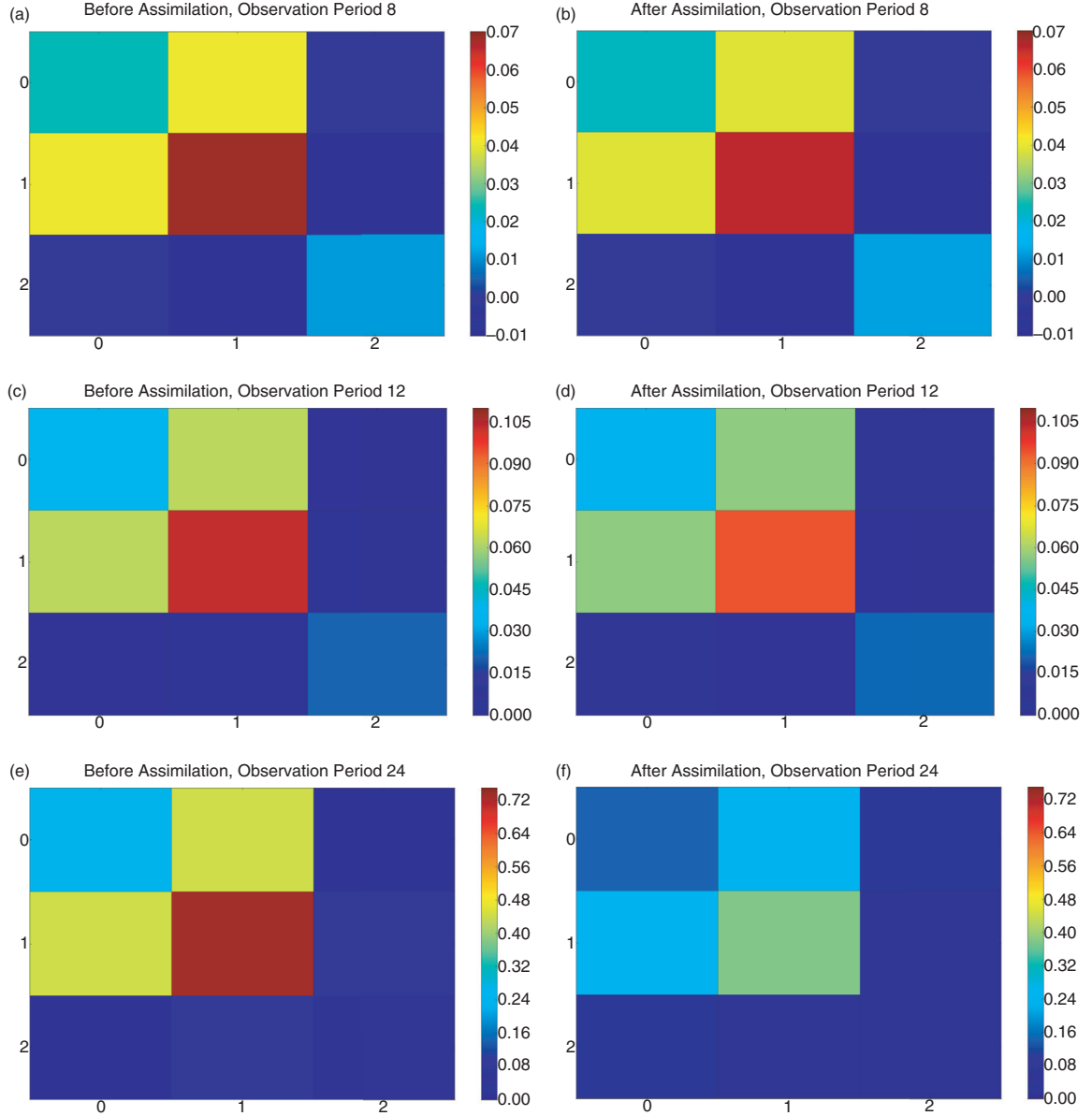


Fig. 5. (a), (c) and (e) are the ensemble-based background error covariances calculated before assimilation, and (b), (d) and (f) are the ensemble-based background error covariance matrices after assimilation. The ensemble-based background error covariance matrices are taken at observation time for different observation periods.

Figure 6 shows a comparison of all our methods over 36 and 48 time step windows without QSVA (Fig. 6a and b) and with QSVA (Fig. 6c and d). In the figures, we show how both the QSVA and non-QSVA versions of the variational methods perform. Figure 6a shows the ETKF, ETKS and 4DENVAR having a much lower RMSE than the methods which use an adjoint because the non-linearity of the system render the linearisation around the first guess via tangent linear and adjoint models no longer suitable. This is investigated further in Fig. 7, which shows that variational methods have an optimal window length, given

fixed observation periods, which gives the lowest RMSE of that method. Looking at each plot, we see that the optimal window length (in our experiments), with an observation period of 2, 3, 6 or 12 time steps, is 24 time steps. This seems to be the point where the variational methods have enough data from observations (as opposed to 12 time steps which has less observations) to improve performance without having the highly non-linear error growth (which occurs in long windows such as 48 time step windows).

To account for the non-linear error growth, we also ran a comparison with QSVA on the 36 and 48 time step windows

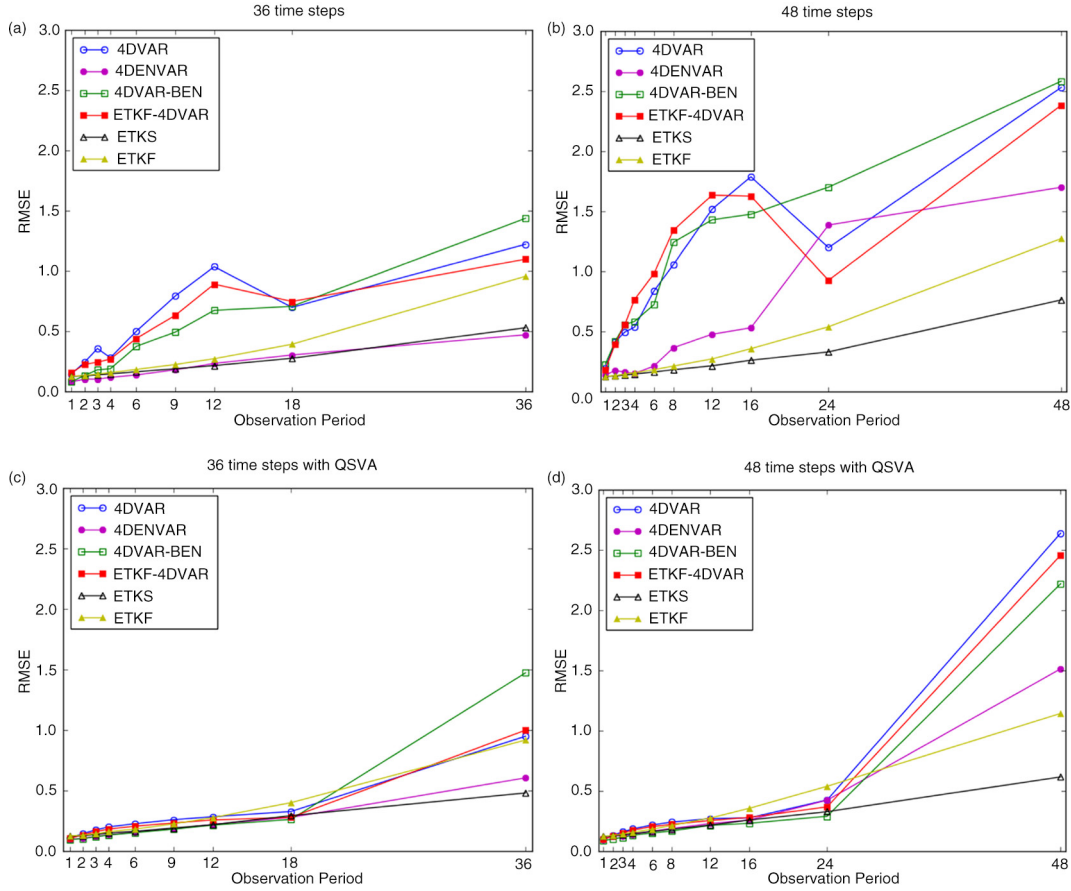


Fig. 6. Performance of each data assimilation method analysis as a function of window lengths and observation periods ( $x$ -axis);  $y$ -axis is the RMSE. These plots show longer windows without (top row) and with (bottom row) QSVA.

(see Fig. 6). Figure 6a and b is compared with Fig. 6c and d to investigate how well a QSVA version of our methods compare with our original non-QSVA methods. It can be seen that QSVA greatly decreases the RMSE, which implies that QSVA has helped considerably in finding the global minimum over a local one. It can be seen on the 36 time step plots (Fig. 6a and c) that 4DENVAR performs very well with and without QSVA. 4DENVAR has a slightly lower RMSE without QSVA than 4DVAR-QSVA and ETKF-4DVAR-QSVA and also has a very similar RMSE to 4DVAR-BEN-QSVA. This tells us that 4DENVAR at this level of non-linear error growth does not need to use these advanced techniques in order to achieve lower RMSE over the traditional methods. At 48 time step windows, the non-linearity of the error growth is too large for 4DENVAR and QSVA improves all of our variational methods. It is noteworthy that the ETKS has the most robust low-RMSE performance of all methods.

Comparisons of future forecasts trajectories with the truth show similar results. For variational methods, the forecast trajectory over the next window was generated by running the model forward in time from the end of the

analysis trajectory in the current window. For sequential methods, we used the same approach but ran the model forward from the ensemble mean. The ETKF and ETKS both have the same values at observation time, so they produce the same forecast trajectories. We find that on shorter forecast windows, the non-climatological hybrid methods outperform the other methods, while, in longer windows, the ETKF and ETKS generally outperform the other methods.

#### 4. Conclusion

In this paper, the traditional data assimilation methods 4DVAR, ETKF and the ETKS were compared against three hybrid methods, 4DENVAR, ETKF-4DVAR and 4DVAR-BEN. It has been shown that over smaller data assimilation windows, when the error growth is close to linear (we used 12 time steps), variational methods which use the ensemble-generated background error covariance matrix (4DENVAR and 4DVAR-BEN) outperform all other methods for all observation periods. We notice that sequential methods also do well in comparison to the variational methods. As the window length increases (to 24 time

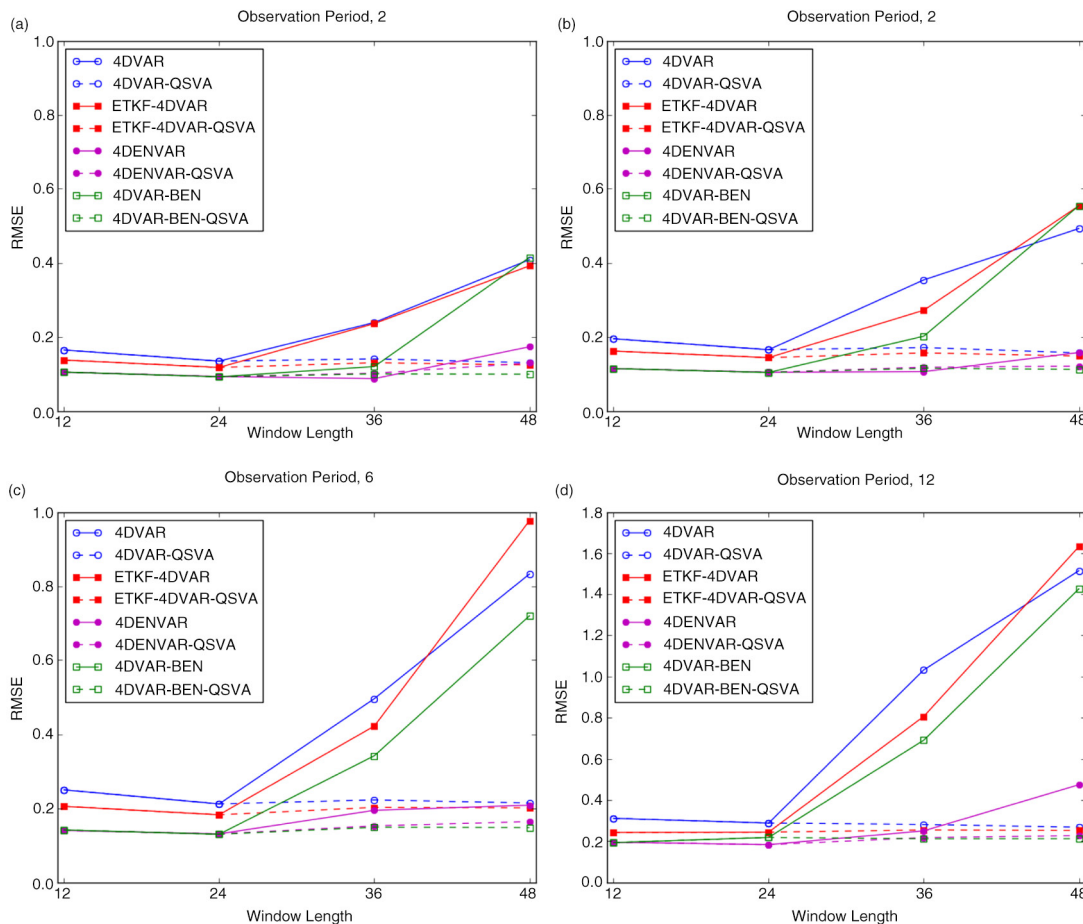


Fig. 7. RMSE (y-axis) for 4DVAR and different variational hybrid methods as a function of window length (x-axis). Each panel shows a different observation period. Note: (d) has a longer y-axis than the others.

steps, which has weakly non-linear error growth), we see an improvement in our variational methods which use a climatological background error covariance matrix, and this is because there is an optimal window length for climatological 4DVAR (Kalnay et al., 2007). Sequential methods produce the same RMSE as they assimilate every time an observation occurs, but for high observation periods, we see that the ETKF is less accurate than all other methods. It is almost certain that the optimal inflation in this case is outside the values we investigated.

Still, it can be seen that variational methods which use a flow-dependent background error covariance matrix from the ETKF are more accurate than any other method, and we also notice that the methods which do not use an adjoint are the most accurate over all observation interval sizes. At 36 time steps windows, it was shown that 4DENVAR provides a very accurate analysis in comparison to the other variational methods, managing to achieve a lower RMSE compared to the other methods even after using quasi-static variational assimilation framework on the other methods. The longest window size (48 time steps) shows sequential

methods outperforming every variational method (when using QSVA, 4DENVAR-QSVA and 4DVAR-BEN-QSVA also comparably well), with the smoother being more accurate than the filter. The method which does not use an adjoint (4DENVAR) is the most accurate variational method over all observation period lengths, although the ETKS has the best overall performance. It is striking how much the performance of variational methods improves when QSVA is used, for longer windows this is related to the existence of local minima.

NWP models are considerably bigger than the Lorenz 1963 model. There are, however, similarities and conclusions which we can extract and will apply to both. All models can be tested in terms of window length and for observation period sizes. In NWP models, these parameters are hard to test due to the complexity of the model (such as number of variables, fast and slow variables etc.), which is where testing these parameters on a smaller model comes in handy. Advantages of 4DENVAR in NWP comes from not having to use an adjoint model and not having to tune a climatological background error covariance matrix since an ensemble method can generate this flow

dependently. This becomes particularly important for long assimilation windows. As most forecasting centres use variational methods for NWP, Fig. 7 shows that increasing window length may not necessarily increase the accuracy of the assimilation. This can be helpful for plans to optimise assimilation parameters. NWP agencies push to increase assimilation window length, which increases the non-linearity for the system. In this case, using an ETKS or 4DENVAR will be a huge improvement over other methods. It should be noted, however, that these methods will need localisation. Another issue to take into consideration in NWP is model error. While the ETKF works rather well, even under the effect of model error, variational methods with model error have not been explored extensively. The next logical step forward would be to research the influence of non-linearity in hybrid data assimilation methods in a system with model error, and where localisation is necessary, this will be the focus of a future paper.

## 5. Acknowledgements

MG and PJvL acknowledge the support of the Natural Environment Research Council (NERC) via the National Centre for Earth Observations (NCEO) and also the DIAMET project (NE/I0051G6/1). JA and PJvL acknowledge the support of NERC (NE/I0125484/1).

## 6. Appendix 1.

---

ETKF	Sequential method which updates trajectories at observation time. The background error covariance matrix is generated from ensemble members, $\mathbf{B}_e$ .
ETKS	Similar to the ETKF but the smoother version which applies the weight to the whole ensemble trajectory. The background error covariance matrix is calculated from ensemble members, $\mathbf{B}_e$ .
4DVAR	Traditional non-incremental strong constraint 4DVAR. Uses an adjoint model and a full climatologic background error covariance matrix, $\mathbf{B}_e$ .
4DVAR-ETKF	Uses ETKF to generate a flow-dependent background error covariance matrix in the 4DVAR framework. Combines flow-dependent $\mathbf{B}_e$ matrix with the climatological $\mathbf{B}_e$ matrix.
4DVAR-BEN	The same method as 4DVAR-ETKF but uses a full flow-dependent $\mathbf{B}_e$ matrix without the climatology.
4DENVAR	4DVAR framework but replaces the adjoint model in the cost function gradient with cross state covariances. Background error covariance matrix is generated from an ensemble of model trajectories.

---

## References

- Amezcuca, J., Ide, K., Bishop, C. H. and Kalnay, E. 2012. Ensemble clustering in deterministic ensemble Kalman filters. *Tellus A*. **64**, 18039. DOI: 10.3402/tellusa.v64i0.18039.
- Anderson, J. L. 2010. A non-Gaussian ensemble filter update for data assimilation. *Mon. Weather Rev.* **138**(11), 4186–4198. DOI: 10.1175/2010MWR3253.1.
- Bishop, C., Etherton, B. and Majumdar, S. 2001. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Weather Rev.* **129**(3), 420–436. DOI: 10.1175/1520-0493(2001)129 <0420:ASWTET > 2.0.CO;2.
- Buehner, M., Houtekamer, P. L., Charette, C., Mitchell, H. L. and He, B. 2010a. Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experiments. *Mon. Weather Rev.* **138**(5), 1550–1566. DOI: 10.1175/2009MWR3157.1.
- Buehner, M., Houtekamer, P. L., Charette, C., Mitchell, H. L., and He, B. 2010b. Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations. *Mon. Weather Rev.* **138**(5), 1567–1586. DOI: 10.1175/2009MWR3158.1.
- Evensen, G. and Van Leeuwen, P. 2000. An ensemble Kalman smoother for nonlinear dynamics. *Mon. Weather Rev.* **128**(6), 1852–1867. DOI: 10.1175/1520-0493(2000)128 <1852:AEKSFN > 2.0.CO;2.
- Fairbarn, D., Pring, S., Lorenc, A. and Roulson, I. 2014. A comparison of 4DVAR with ensemble data assimilation methods. *Q. J. Roy. Meteorol. Soc.* **140**, 281–294.
- Gu, Y. and Oliver, D. S. 2007. An iterative ensemble Kalman filter for multiphase fluid flow data assimilation. *SPE J.* **12**(4), 438–446.
- Hunt, B., Kalnay, E., Kostelich, E. J., OTT, E., Patil, D. J. and co-authors. 2004. Four-dimensional ensemble Kalman filtering. *Tellus A*. **56**(4), 273–277. DOI: 10.1111/j.1600-0870.2004.00066.x.
- Kalnay, E., Li, H., Miyoshi, T., Yang, S.-C. and Ballabrera-Poy, J. 2007. 4-D-Var or ensemble Kalman filter? *Tellus A*. **59**(5), 758–773. DOI: 10.1111/j.1600-0870.2007.00261.x.
- Kuhl, D. D., Rosmond, T. E., Bishop, C. H., McLay, J. and Baker, N. L. 2013. Comparison of hybrid ensemble/4DVar and 4DVar within the NAVDAS-AR data assimilation framework. *Mon. Weather Rev.* **141**(8), 2740–2758. DOI: 10.1175/MWR-D-12-00182.1.
- Lawson, W. and Hansen, J. 2004. Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. *Mon. Weather Rev.* **132**(8), 1966–1981. DOI: 10.1175/1520-0493(2004)132 <1966:IOSADF > 2.0.CO;2.
- Liu, C., Xiao, Q. and Wang, B. 2008. An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technical formulation and preliminary test. *Mon. Weather Rev.* **136**(9), 3363–3373. DOI: 10.1175/2008MWR2312.1.
- Livings, D. M., Dance, S. L. and Nichols, N. K. 2008. Unbiased ensemble square root filters. *Physica D*. **237**(8), 1021–1028. DOI: 10.1016/j.physd.2008.01.005.
- Lorenc, A. 2003. The potential of the ensemble Kalman filter for NWP – a comparison with 4D-Var. *Q. J. Roy. Meteorol. Soc.* **129**(595B), 3183–3203. DOI: 10.1256/qj.02.132.

- Lorenc, A. C. 2011. Developments of variational data assimilation. In: *ECMWF Seminars*, ECMWF. 6–9 September 2011.
- Lorenz, E. N. 1963. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–148. DOI: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.
- Pham, D., Verron, J. and Roubaud, M. 1998. A singular evolutive extended Kalman filter for data assimilation in oceanography. *J. Mar. Syst.* **16**(3–4), 323–340.
- Pires, C., Vautard, R. and Talagrand, O. 1996. On extending the limits of variational assimilation in nonlinear chaotic systems. *Tellus A.* **48**(1), 96–121. DOI: 10.1034/j.1600-0870.1996.00006.x.
- Sakov, P. and Oke, P. R. 2008. A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. *Tellus A.* **60**(2), 361–371. DOI: 10.1111/j.1600-0870.2007.00299.x.
- Swanson, K., Palmer, T. and Vautard, R. 2000. Observational error structures and the value of advanced assimilation techniques. *J. Atmos. Sci.* **57**(9), 1327–1340. DOI: 10.1175/1520-0469(2000)057<1327:OESATV>2.0.CO;2.
- Swanson, K., Vautard, R. and Pires, C. 1998. Four-dimensional variational assimilation and predictability in a quasi-geostrophic model. *Tellus A.* **50**, 369–389.
- Talagrand, O. and Courtier, P. 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. 1. Theory. *Q. J. Roy. Meteorol. Soc.* **113**(478), 1311–1328. DOI: 10.1256/smsqj.47811.
- Tian, X., Xie, Z. and Dai, A. 2008. An ensemble-based explicit four-dimensional variational assimilation method. *J. Geophys. Res. Atmos.* **113**(D21). DOI: 10.1029/2008JD010358.
- Tippett, M., Anderson, J., Bishop, C., Hamill, T. and Whitaker, J. 2003. Ensemble square root filters. *Mon. Weather Rev.* **131**(7), 1485–1490. DOI: 10.1175/1520-0493(2003)131<1485:ESRF>2.0.CO;2.
- Van Leeuwen, P. and Evensen, G. 1996. Data assimilation and inverse methods in terms of a probabilistic formulation. *Mon. Weather Rev.* **124**(12), 2898–2913. DOI: 10.1175/1520-0493(1996)124<2898:DAAIMI>2.0.CO;2.
- Wang, X. 2010. Incorporating ensemble covariance in the grid-point statistical interpolation variational minimization: a mathematical framework. *Mon. Weather Rev.* **138**(7), 2990–2995. DOI: 10.1175/2010MWR3245.1.
- Wang, X., Barker, D. M., Snyder, C. and Hamill, T. M. 2008. A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part I: Observing system simulation experiment. *Mon. Weather Rev.* **136**(12), 5116–5131. DOI: 10.1175/2008MWR2444.1.
- Wang, X., Bishop, C. and Julier, S. 2004. Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble? *Mon. Weather Rev.* **132**(7), 1590–1605. DOI: 10.1175/1520-0493(2004)132<1590:WIBAEO>2.0.CO;2.
- Wang, X. and Lei, T. 2014. GSI-based four-dimensional ensemble variational (4DEnsVar) data assimilation: formulation and single-resolution experiments with real data for NCEP global forecast system. *Mon. Weather Rev.* **142**(9), 3303–3325. DOI: 10.1175/MWR-D-13-00303.1.
- Yang, S.-C., Baker, D., Li, H., Cordes, K., Huff, M. and co-authors. 2006. Data assimilation as synchronization of truth and model: experiments with the three-variable Lorenz system. *J. Atmos. Sci.* **63**(9), 2340–2354. DOI: 10.1175/JAS3739.1.
- Yang, S.-C., Kalnay, E. and Hunt, B. 2012. Handling non linearity in an ensemble Kalman filter: experiments with the three-variable Lorenz model. *Mon. Weather Rev.* **140**(8), 2628–2646. DOI: 10.1175/MWR-D-11-00313.1.
- Zupanski, M. 2005. Maximum likelihood ensemble filter: theoretical aspects. *Mon. Weather Rev.* **133**(6), 1710–1726. DOI: 10.1175/MWR2946.1.