

# A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters

By PAVEL SAKOV\* and PETER R. OKE, *CSIRO Marine and Atmospheric Research and Wealth from Oceans Flagship Program, GPO 1538, Hobart, Tasmania 7001, Australia*

(Manuscript received 5 May 2007; in final form 26 November 2007)

## ABSTRACT

The use of perturbed observations in the traditional ensemble Kalman filter (EnKF) results in a suboptimal filter behaviour, particularly for small ensembles. In this work, we propose a simple modification to the traditional EnKF that results in matching the analysed error covariance given by Kalman filter in cases when the correction is small; without perturbed observations. The proposed filter is based on the recognition that in the case of small corrections to the forecast the traditional EnKF without perturbed observations reduces the forecast error covariance by an amount that is nearly twice as large as that is needed to match Kalman filter. The analysis scheme works as follows: update the ensemble mean and the ensemble anomalies separately; update the mean using the standard analysis equation; update the anomalies with the same equation but half the Kalman gain. The proposed filter is shown to be a linear approximation to the ensemble square root filter (ESRF). Because of its deterministic character and its similarity to the traditional EnKF we call it the ‘deterministic EnKF’, or the DEnKF. A number of numerical experiments to compare the performance of the DEnKF with both the EnKF and an ESRF using three small models are conducted. We show that the DEnKF performs almost as well as the ESRF and is a significant improvement over the EnKF. Therefore, the DEnKF combines the numerical effectiveness, simplicity and versatility of the EnKF with the performance of the ESRFs. Importantly, the DEnKF readily permits the use of the traditional Schur product-based localization schemes.

## 1. Introduction

The Ensemble Kalman Filter (EnKF) was introduced by Evensen (1994) and has attracted a lot of attention in the literature (e.g. see references in Evensen, 2003). It makes it possible to apply Kalman filter to high-dimensional discrete systems, when the explicit storage and manipulation of the system state error covariance are impossible or impractical.

The EnKF may be characterized by the application of the analysis equation given by Kalman filter to an ensemble of forecasts. It has been recognized (Burgers et al., 1998) that a straightforward application of the analysis equation to each of the ensemble members results in an ensemble collapse, when the ensemble spread reduces too rapidly. The standard way to prevent the ensemble collapse in the EnKF is to update each ensemble member using independently perturbed observations (Burgers et al., 1998; Houtekamer and Mitchell, 1998); the resulting algorithm has become known as the perturbed observations EnKF and is currently commonly associated with the acronym EnKF. The use of perturbed observations in the EnKF results in the analysed error covariance matching the theoretical value given by

Kalman filter in a statistical sense only. It introduces sampling error, which makes the filter suboptimal, particularly for small ensembles (Whitaker and Hamill, 2002).

The traditional EnKF can be regarded as a Monte Carlo formulation of the Kalman Filter. By contrast, a number of deterministic ensemble-based filters have been proposed; and are referred to as the Ensemble Square Root Filters (ESRFs). A brief description of four different schemes may be found in Tippett et al. (2003); see also Pham et al. (1998), Ott et al. (2004), Evensen (2004) and Sakov and Oke (2008). Although an ESRF may show a significantly better performance over the EnKF in numerical experiments (e.g. Whitaker and Hamill, 2002), the EnKF still has a number of features that may be attractive in practice. The EnKF is simpler in some way as it requires only a repetitive application of the analysis equation to each ensemble member; and it readily permits the traditional Schur-product-based localization (Hamill and Whitaker, 2001; Houtekamer and Mitchell, 2001), which may be required for realistic applications when the system dimension exceeds the ensemble size (Oke et al., 2006). Note that applying the Schur-product-based localization to an ESRF currently requires using serial assimilation (Whitaker and Hamill, 2002).

In this work, we propose a simple modification to the EnKF, which results in an asymptotic matching of the analysed error covariance given by Kalman filter theory in cases when the

---

\*Corresponding author.  
e-mail: pavel.sakov@csiro.au  
DOI: 10.1111/j.1600-0870.2007.00299.x

correction is small; without perturbed observations. Because of the deterministic character of the proposed filter and its close relation to the EnKF we refer to it as the deterministic EnKF, or the DEnKF.

To test the performance of the DEnKF we apply it to three different systems based on small models with different characteristics. It is shown to perform as well as an ESRF and significantly better than the EnKF.

This paper is organized as follows: the theoretical background of the traditional EnKF is given in Section 2; the DEnKF is introduced in Section 3; followed by a series of numerical tests in Section 4. A discussion is presented in Section 5, followed by our conclusions in Section 6.

## 2. Background

The EnKF is based on the Kalman filter analysis equation:

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}(\mathbf{d} - \mathbf{H}\mathbf{x}^f), \quad (1)$$

where

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1} \quad (2)$$

is referred to as the Kalman gain;  $\mathbf{x}^a$  is the analysis;  $\mathbf{x}^f$  is the forecast;  $\mathbf{d}$  is the vector of observations;  $\mathbf{H}$  is the observation sensitivity matrix,  $\mathbf{H} = \nabla_{\mathbf{x}} \mathcal{H}(\mathbf{x}^f)$ , where  $\mathcal{H}(\mathbf{x})$  is the (non-linear) operator mapping the state vector space to the observation space;  $\mathbf{P}^f$  is the forecast error covariance matrix;  $\mathbf{R}$  is the observation error covariance matrix; the superscripts  $f$  and  $a$  denote forecast and analysis, respectively; and the superscript  $T$  denotes a matrix transpose. In the EnKF, the covariance matrix  $\mathbf{P}$  is stored and manipulated implicitly via an ensemble  $\mathbf{X}$  of model states,  $\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_m]$ , where  $m$  is the ensemble size, using the relation

$$\mathbf{P} = \frac{1}{m-1} \sum_{i=1}^m (\mathbf{X}_i - \bar{\mathbf{x}})(\mathbf{X}_i - \bar{\mathbf{x}})^T = \frac{1}{m-1} \mathbf{A} \mathbf{A}^T, \quad (3)$$

where  $\bar{\mathbf{x}}$  is the ensemble mean:

$$\bar{\mathbf{x}} = \frac{1}{m} \sum_{i=1}^m \mathbf{X}_i, \quad (4)$$

and  $\mathbf{A} = [\mathbf{A}_1, \dots, \mathbf{A}_m]$  is the ensemble of anomalies, or perturbations,

$$\mathbf{A}_i = \mathbf{X}_i - \bar{\mathbf{x}}. \quad (5)$$

The ensemble mean is considered the best estimate for the system state.

The EnKF processes each ensemble member using the analysis eq. (1):

$$\mathbf{X}_i^a = \mathbf{X}_i^f + \mathbf{K}(\mathbf{d} + \mathbf{D}_i - \mathbf{H}\mathbf{X}_i^f), \quad i = 1, \dots, m, \quad (6)$$

where  $\mathbf{D}_i$  is a synthetic vector of perturbations of observations  $\mathbf{d}$ . The ensemble average of (6) yields the analysis eq. (1), provided that the ensemble average of  $\mathbf{D}_i$  is zero:

$$\mathbf{D}\mathbf{1} = \mathbf{0}, \quad (7)$$

where  $\mathbf{D} = [\mathbf{D}_1, \dots, \mathbf{D}_m]$ , and  $\mathbf{1} = [1, \dots, 1]^T$  is the vector with all components equal to 1. To calculate the analysed error covariance, we first subtract the analysis eq. (1) from (6); this yields the equation for the update of an ensemble anomaly

$$\mathbf{A}_i^a = \mathbf{A}_i^f + \mathbf{K}(\mathbf{D}_i - \mathbf{H}\mathbf{A}_i^f), \quad i = 1, \dots, m, \quad (8)$$

which may be written in a matrix form for the full ensemble as

$$\mathbf{A}^a = \mathbf{A}^f + \mathbf{K}(\mathbf{D} - \mathbf{H}\mathbf{A}^f). \quad (9)$$

The analysed error covariance carried by the ensemble can then be calculated using (3):

$$\begin{aligned} \mathbf{P}^a &= \frac{1}{m-1} \mathbf{A}^a \mathbf{A}^{aT} \\ &= \frac{1}{m-1} [\mathbf{A}^f + \mathbf{K}(\mathbf{D} - \mathbf{H}\mathbf{A}^f)] [\mathbf{A}^f + \mathbf{K}(\mathbf{D} - \mathbf{H}\mathbf{A}^f)]^T \\ &= \mathbf{P}^f - \mathbf{P}^f \mathbf{H}^T \mathbf{K}^T - \mathbf{K} \mathbf{H} \mathbf{P}^f + \mathbf{K} \mathbf{H} \mathbf{P}^f \mathbf{H}^T \mathbf{K}^T \\ &\quad + \frac{1}{m-1} \mathbf{K} \mathbf{D} \mathbf{D}^T \mathbf{K}^T + \frac{1}{m-1} (\mathbf{I} - \mathbf{K} \mathbf{H}) \mathbf{A}^f \mathbf{D}^T \mathbf{K}^T \\ &\quad + \frac{1}{m-1} \mathbf{K} \mathbf{D} \mathbf{A}^{fT} (\mathbf{I} - \mathbf{H}^T \mathbf{K}^T). \end{aligned} \quad (10)$$

Without perturbing the observations,  $\mathbf{D} = \mathbf{0}$ , the analysed error covariance carried by the ensemble becomes

$$\begin{aligned} \mathbf{P}^a &= \mathbf{P}^f - \mathbf{P}^f \mathbf{H}^T \mathbf{K}^T - \mathbf{K} \mathbf{H} \mathbf{P}^f + \mathbf{K} \mathbf{H} \mathbf{P}^f \mathbf{H}^T \mathbf{K}^T \\ &= (\mathbf{I} - \mathbf{K} \mathbf{H}) \mathbf{P}^f (\mathbf{I} - \mathbf{H}^T \mathbf{K}^T), \end{aligned} \quad (11)$$

which is less than the value given by Kalman filter

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K} \mathbf{H}) \mathbf{P}^f, \quad (12)$$

and therefore results in a premature reduction in the ensemble spread (Burgers et al., 1998; Houtekamer and Mitchell, 1998). The traditional way of handling this problem in the EnKF is to add synthetic perturbations  $\mathbf{D}$  (Burgers et al., 1998; Houtekamer and Mitchell, 1998). If  $\mathbf{D}$  is defined so that

$$\frac{1}{m-1} \mathbf{D} \mathbf{D}^T = \mathbf{R}, \quad (13)$$

then in (10)

$$-\mathbf{K} \mathbf{H} \mathbf{P}^f + \mathbf{K} \mathbf{H} \mathbf{P}^f \mathbf{H}^T \mathbf{K}^T + \frac{1}{m-1} \mathbf{K} \mathbf{D} \mathbf{D}^T \mathbf{K}^T = \mathbf{0}.$$

If there was a solution for  $\mathbf{D}$  that satisfies (7), (13) and

$$\mathbf{A}^f \mathbf{D}^T \mathbf{K}^T = \mathbf{0}, \quad (14)$$

then the ensemble transformation (9) would result in the analysed error covariance (10) exactly matching the theoretical value (12). However, such a solution does not, generally, exist. Consequently, the traditional EnKF only satisfies these conditions approximately, in a statistical sense. If  $\mathbf{D}$  contains random normally distributed elements such that eqs (7) and (13) are satisfied in a statistical sense, then, if rows of  $\mathbf{D}$  are not correlated with rows of  $\mathbf{A}$ , eq. (14) is also satisfied in a statistical sense, so that arguably  $\mathbf{P}^a = (\mathbf{I} - \mathbf{K} \mathbf{H}) \mathbf{P}^f + O(m^{-1/2})$  (Burgers et al., 1998).

The above considerations form a theoretical basis for the traditional EnKF, in which the analysis eq. (6) for the  $i$ th ensemble member is contaminated with random Gaussian noise  $\mathbf{D}_i$  with zero mean and variance equal to that of the observation error. (Note that in the case of the correlated observation errors, when  $\mathbf{R}$  is non-diagonal, one needs to generate correlated Gaussian noise with the covariance  $\mathbf{R}$ .) The downside of this approach is that it introduces a sampling error and makes the traditional EnKF suboptimal, particularly for small ensembles.

In this work, we propose a simple modification to the traditional EnKF that makes the analysed error covariance asymptotically match the theoretical value (12).

### 3. The deterministic EnKF

Because

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

the expression for the analysed error covariance without perturbed observations (11) can be written as

$$\mathbf{P}^a = \mathbf{P}^f - 2\mathbf{K} \mathbf{H} \mathbf{P}^f + \mathbf{K} \mathbf{H} \mathbf{P}^f \mathbf{H}^T \mathbf{K}^T.$$

If  $\mathbf{K} \mathbf{H}$  is small in some sense, then the quadratic term  $\mathbf{K} \mathbf{H} \mathbf{P}^f \mathbf{H}^T \mathbf{K}^T$  is much smaller than the linear term  $2\mathbf{K} \mathbf{H} \mathbf{P}^f$ ; and one can asymptotically match the theoretical covariance up to quadratic terms by halving the Kalman gain  $\mathbf{K}$ . This argument forms the theoretical basis for the DEnKF. The proposed analysis scheme is as follows.

- (i) Given the forecast ensemble  $\mathbf{X}^f$ , calculate the ensemble mean, or forecast  $\mathbf{x}^f$  by (4), and the ensemble anomalies  $\mathbf{A}^f$  by (5).
- (ii) Calculate the analysis  $\mathbf{x}^a$  by using the Kalman analysis eq. (1).
- (iii) Calculate the analysed anomalies by

$$\boxed{\mathbf{A}^a = \mathbf{A}^f - \frac{1}{2} \mathbf{K} \mathbf{H} \mathbf{A}^f.} \quad (15)$$

- (iv) Calculate the analysed ensemble by offsetting the analysed anomalies by the analysis:

$$\mathbf{X}^a = \mathbf{A}^a + [\mathbf{x}^a, \dots, \mathbf{x}^a]. \quad (16)$$

This scheme is equivalent to applying the Kalman analysis equation to each ensemble member using half the Kalman gain and no perturbed observations, with a subsequent readjustment of the ensemble mean to the explicitly calculated analysis (1). Similarly to the ESRF, the analysed error covariance in the DEnKF does not depend on the particular realization of the observation or assimilation procedures. Therefore, it also represents a ‘deterministic’ filter.

Updating the ensemble anomalies by eq. (15) results in the analysed error covariance that is equal to

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K} \mathbf{H}) \mathbf{P}^f + \frac{1}{4} \mathbf{K} \mathbf{H} \mathbf{P}^f \mathbf{H}^T \mathbf{K}^T. \quad (17)$$

This expression has an extra term compared with (12). The term is quadratic over  $\mathbf{K} \mathbf{H}$  and positive semi-definite. The main question is how significant is the discrepancy it introduces for the performance of the DEnKF. This question will be partially answered in the next section in experiments with three small models.

Because the DEnKF linearly approximates the theoretical error covariance, the ensemble transform (15) must represent a linear approximation to the ensemble transform of an ESRF. This is indeed the case because (15) can be rewritten as

$$\mathbf{A}^a = \mathbf{A}^f \left( \mathbf{I} - \frac{1}{2} \frac{\mathbf{A}^{fT} \mathbf{H}^T \mathbf{M}^{-1} \mathbf{H} \mathbf{A}^f}{m-1} \right), \quad \mathbf{M} \equiv \mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R},$$

in which the expression in brackets represents the first two terms of the expansion of the square root in the ESRF solution

$$\mathbf{A}^a = \mathbf{A}^f \left( \mathbf{I} - \frac{\mathbf{A}^{fT} \mathbf{H}^T \mathbf{M}^{-1} \mathbf{H} \mathbf{A}^f}{m-1} \right)^{1/2} \quad (18)$$

(Evensen, 2004) into the Taylor series:

$$\mathbf{A}^a = \mathbf{A}^f \left[ \mathbf{I} - \frac{1}{2} \frac{\mathbf{A}^{fT} \mathbf{H}^T \mathbf{M}^{-1} \mathbf{H} \mathbf{A}^f}{m-1} - \frac{1}{8} \left( \frac{\mathbf{A}^{fT} \mathbf{H}^T \mathbf{M}^{-1} \mathbf{H} \mathbf{A}^f}{m-1} \right)^2 - \dots \right]. \quad (19)$$

### 4. Numerical tests

In this section we compare the performance of the DEnKF with that of the traditional EnKF and an ESRF for three small models, namely the Linear Advection (LA) model of Evensen (2004); the non-linear Lorenz-40 (L40) model of Lorenz and Emanuel (1998); and a 1.5-layer reduced-gravity quasi-geostrophic (QG) model with double-gyre wind forcing and bi-harmonic friction. We first give a detailed description of the models and the configurations used in the tests, and then describe the results. In all experiments with the LA and L40 models we use a symmetric (‘spherical simplex’) ensemble transform Kalman filter (ETKF) (Bishop et al., 2001; Wang et al., 2004) to represent the ESRF, while in the experiments with the QG model we use the serial scheme by Whitaker and Hamill (2002) that permits localization.

#### 4.1. The LA model

The LA model, described below, is based on that of Evensen (2004). The dimension of the state vector  $\mathbf{x}$  is 1000; the signal propagates (advects) in the positive direction by one element at each time step without changing its shape; and the model domain is periodic:

$$\begin{aligned} \mathbf{x}(t) &= [\mathbf{x}_1(t), \dots, \mathbf{x}_{1000}(t)], \quad t = 1, 2, \dots; \\ \mathbf{x}_i(t+1) &= \begin{cases} \mathbf{x}_{i-1}(t), & i = 2, \dots, 1000; \\ \mathbf{x}_{1000}(t), & i = 1, \end{cases} \end{aligned}$$

where  $\mathbf{x}_i(t)$  is the  $i$ th component of the state vector at the  $t$ th time step.

To generate a state vector sample, a sum of 25 sine curves with random amplitude and phase and a random offset is calculated:

$$\mathbf{s}_i = \sum_{k=0}^{25} a_k \sin\left(\frac{2\pi k}{1000}i + \varphi_k\right), \quad i = 1, \dots, 1000,$$

where  $a_k$  and  $\varphi_k$  are random numbers uniformly distributed in the intervals  $(0, 1)$  and  $(0, 2\pi)$ , correspondingly. This sample is then normalized to have a variance of 1:

$$\mathbf{x}(1) = \frac{\mathbf{s}}{[(\mathbf{s} - \bar{\mathbf{s}})^T(\mathbf{s} - \bar{\mathbf{s}})]^{1/2}},$$

where  $\bar{\mathbf{s}}$  denotes the state average. To generate the initial ensemble, a specified number of samples are generated using this procedure. The ensemble mean field is subsequently subtracted from each member, and another random sample ('climatology') is added to each member. The true field is defined as a sum of yet another random sample and the climatology.

Therefore, by construction all possible model state vectors belong to a subspace with dimension of 51 (referred to hereafter as the model dimension) of the full model state vector space with dimension of 1000. Also by construction, the initial root mean square error (RMSE) of the ensemble mean is equal to 1. Note that because the ensemble mean has been subtracted from each member *after* the normalization, each ensemble anomaly has an initial variance that is slightly different from 1.

Four observations of the true field are conducted and assimilated into the model at every fifth time step,  $t = 1, 6, 11, \dots$ , at equidistant locations  $i = \{125, 375, 625, 875\}$ . Each observation is contaminated with random normally distributed uncorrelated noise with variance of 0.01.

#### 4.2. The L40 model

The L40 model (Lorenz and Emanuel, 1998) is a strongly non-linear model with a state vector dimension of 40. Lorenz and Emanuel (1998) argue that it roughly imitates the evolution of an unspecified scalar meteorological quantity (such as temperature or vorticity) along a latitude circle. This model has been used for testing ensemble-based assimilation methods in a number of earlier studies (Anderson, 2001; Whitaker and Hamill, 2002; Ott et al., 2004; Lawson and Hansen, 2004). It contains 40 coupled ordinary differential equations in a domain with cyclic boundary conditions:

$$\dot{\mathbf{y}}_i = (\mathbf{y}_{i+1} - \mathbf{y}_{i-2})\mathbf{y}_{i-1} - \mathbf{y}_i + 8, \quad i = 1, \dots, 40;$$

$$\mathbf{y}_0 = \mathbf{y}_{40}, \quad \mathbf{y}_{-1} = \mathbf{y}_{39}, \quad \mathbf{y}_{41} = \mathbf{y}_1.$$

The consecutive model states are obtained by integrating these equations forward by intervals of  $\Delta t = 0.05$ , so that the model states  $\mathbf{x}(t)$  relate to the solution of the above system as

$$\mathbf{x}(t) = \mathbf{y}(0.05t), \quad t = 1, 2, \dots$$

The model has an estimated fractal dimension of 27.1, the doubling time of the leading Lyapunov exponent of 0.42 (which corresponds to approximately 8 time steps), a mean of 2.34 and standard deviation of 3.66 (Lorenz and Emanuel, 1998).

Following Lorenz and Emanuel (1998), each model time step in our tests is conducted by a single step of the standard fourth-order Runge–Kutta integrator. The ensemble members are initialized by random sampling from a set of 10 000 model states obtained during one continuous integration at  $t = 1000, 1001, \dots, 11\,000$ . The true field is initialized by one more randomly chosen state from this set. Following Whitaker and Hamill (2002) and Ott et al. (2004), at every time step we conduct 40 observations of the true field at the node locations,  $i = 1, \dots, 40$ ; each observation is contaminated with random normally distributed uncorrelated noise with a variance of 1. Common for assimilation with non-linear models, an ensemble inflation is applied to the ensemble at the end of each assimilation step by multiplying the ensemble anomalies by the inflation factor  $\delta$  with typical values between 1.00 and 1.10:  $\mathbf{A}^a \leftarrow \delta \cdot \mathbf{A}^a$ .

The L40 model is a chaotic non-linear model, and as such represents an excellent framework for testing the filter performance with a dynamic model; however, it does not represent 'typical' conditions of a GCM due to its low dimension (40) and relatively high model subspace dimension (about 27). We therefore introduce a third model that is described below.

#### 4.3. The QG model

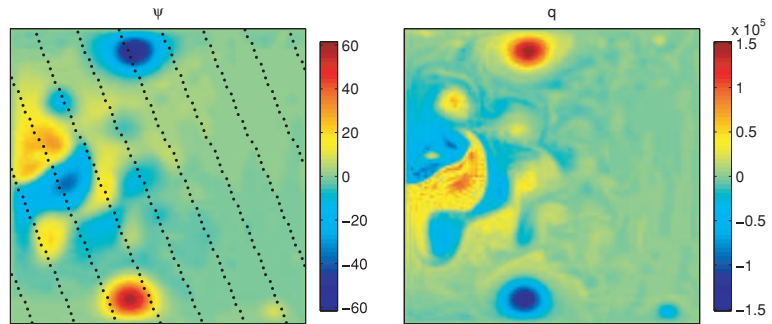
The QG model is a derivative of the 1.5-layer reduced-gravity quasi-geostrophic model with double-gyre wind forcing and bi-harmonic friction. It is a non-linear model with dimension of about  $1.6 \times 10^4$  and model subspace dimension of the order of  $10^2 - 10^3$ , which, we believe, is more representative of a realistic atmospheric or oceanic data assimilating system.

The QG model is a numerical approximation of the following equation:

$$q_t = -\psi_x - \varepsilon J(\psi, q) - A\Delta^3\psi + 2\pi \sin(2\pi y), \quad (20)$$

where  $q = \Delta\psi - F\psi$ ,  $J(\psi, q) \equiv \psi_x q_y - \psi_y q_x$ ,  $\Delta \equiv \partial^2/\partial x^2 + \partial^2/\partial y^2$ . This equation is similar to eq. (3) from Jelloul and Thierry (2003), which was obtained by non-dimensionalizing the initial potential vorticity equation.  $\psi$  may be interpreted as either a stream function or surface elevation. For the coefficients in (20) we use values of  $F = 1600$ ,  $\varepsilon = 10^{-5}$ , and  $A = 2 \times 10^{-12}$ . The model domain represents a  $1 \times 1$  square,  $0 \leq x \leq 1, 0 \leq y \leq 1$ , discretized by a  $129 \times 129$  grid. We use boundary conditions  $\psi = \Delta\psi = \Delta^2\psi = 0$ ; second-order centred differencing approximations for the derivatives and Laplacian; and Arakawa approximation for the Jacobian. In many papers involving quasi-geostrophic models different or more complicated boundary conditions are used; however, the simplistic boundary conditions above are sufficient for the purpose of this paper. Terms with  $\Delta^2$  and  $\Delta^3$  are calculated by consecutive application

Fig. 1. A typical state of the QG model; the dots show an example of the observation locations.



of the employed approximation for the Laplacian. We integrate in time by the fourth-order Runge–Kutta scheme with the time step of 1.25. The integration is conducted for the potential vorticity  $q$ . To calculate  $\psi$  from a given  $q$  we solve the Helmholtz equation  $\Delta\psi - F\psi = q$  using an iterative multigrid solver with the latest estimate for  $\psi$  as the initial guess.

The model has dimension of  $127 \times 127 = 16\,129$  (excluding the boundary points) and a typical doubling time of the leading eigenmode between 25 and 50 (which corresponds to between 5 and 10 assimilation cycles). A typical example of the  $\psi$  and  $q$  fields from the QG model are shown in Fig. 1.

The data assimilation is conducted at every fourth time step. At each assimilation cycle, 300 observations of  $\psi$  with observation variance of 4.0 are used. The observations are distributed uniformly over the state vector length, with a random offset, different for each assimilation cycle, resulting in 10 densely populated equidistant parallel tracks (Fig. 1). This observation network is motivated by the typical distribution of satellite altimetry for oceanic applications. All experiments are conducted with an ensemble of 25 members. This ensemble size is intentionally chosen to be much smaller than the model subspace dimension since this is typical of many realistic oceanic or atmospheric applications (Oke et al., 2006). During data assimilation, ensemble inflation and localization are applied. The ensemble inflation is applied in the same way as for the L40 model, by multiplying the analysed ensemble anomalies by a given factor; the factors from 1.0 to 1.18 with the increment of 0.02 are used. For localization we use the Gaussian localization function. It is applied to the state error covariance matrix by means of a Schur product by multiplying each element  $\mathbf{P}_{ij}$  of the covariance matrix by the factor  $\rho_{ij} = \exp(-0.5 r_{ij}^2 / r_0^2)$ , where  $r_{ij}$  is the horizontal distance between elements  $i$  and  $j$  in grid space, and  $r_0$  is the localization radius. The localization radii of 5–45 with the increment of 5 are used. The initial ensemble is formed by 25 random samples of 2000 fields collected from a long model run to the time of  $t = 5 \times 10^5$ , with the true field being initialized from another random field from this set.

For each combination of the ensemble inflation and localization radius we conduct 10 runs of 1200 steps each, during which 301 assimilation cycles are conducted. We compare results obtained in these runs for three different schemes, the DEnKF, the

serial ESRF (Whitaker and Hamill, 2002), and the traditional EnKF.

In choosing the parameters of the data assimilating system based on the eq. (20), we aim to achieve stable runs (without divergence) for some range of the inflation factor and localization radius. We find that our assimilating system cannot achieve a stable performance with model parameters equivalent to those that allowed long stable runs of the stand-alone model. We attribute this less stable behaviour of the assimilating system compared to the stand-alone model to the dynamic inconsistency of the corrections introduced, firstly, by localization (e.g. Oke et al., 2006) and, secondly, by the model's non-linearity. We find that it is possible to achieve a stable performance of the assimilating system by increasing the dissipation for the ensemble by a factor of 10, to  $A = 2 \times 10^{-11}$  and, importantly, by reducing the time step from 1.5 (used in free model runs) to 1.25.

#### 4.4. Results

Results using the LA model and the L40 model are shown in Fig. 2. It shows the behaviour of the RMSE of the analyses and the ensemble spread during the initial stage of one particular realization of the system. Here we define the ensemble spread as the ensemble mean of the root mean squared deviation of the anomalies. The top three panels in Fig. 2 refer to the LA model, while the bottom three panels refer to the L40 model. The LA model runs are conducted with a 55-member ensemble and no inflation; while the L40 runs are conducted with a 35-member ensemble and inflation factor of 1.02. These ensemble sizes are only slightly larger than the model dimensions (equal to 51 for the LA model and approximately 27 for the L40 model), and should result in a relatively small residual. After the initial transient period, for both the ESRF and DEnKF the RMSE and the ensemble spread become approximately equal; while for the EnKF one can observe the filter collapse. This demonstrates a consistency in this particular case between the actual analysis error covariance (represented by RMSE) and its representation by the ensemble for both ESRF and DEnKF, but not for EnKF. The performance of the ESRF and DEnKF is very comparable, with a slightly smaller residual achieved by the ESRF for the LA

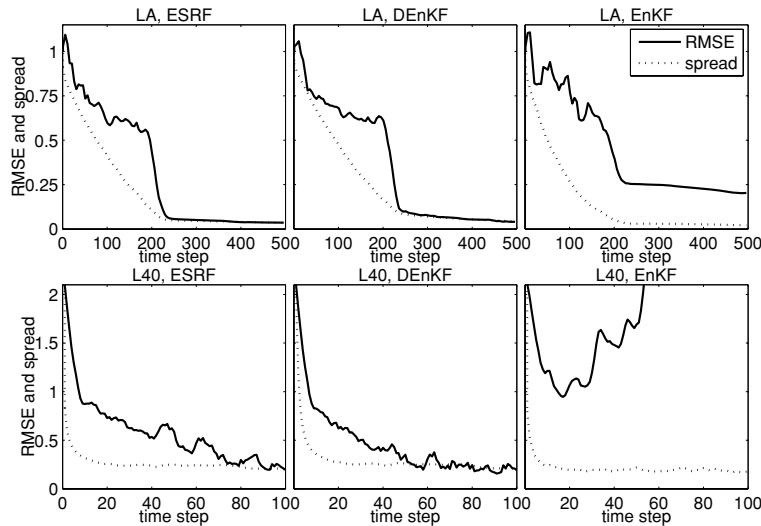


Fig. 2. An example of the RMSE and ensemble spread of the ESRF, DEnKF and EnKF for LA and L40 models.

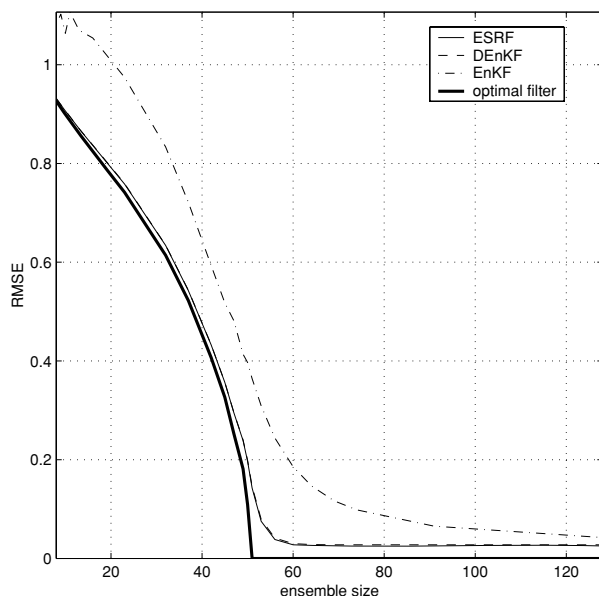


Fig. 3. RMSE for the LA model using the ESRF, DEnKF and EnKF for the LA model, averaged for the time interval  $t = [900, 1000]$ , and over 50 realizations.

model, and a slightly faster convergence achieved by the DEnKF for the L40 model

Figure 3 shows the mean value of the RMSE over the time interval  $t = [900, 1000]$ , averaged over 50 random realizations of the LA model, versus the ensemble size. Each realization refers here to a simulation with a different truth and initial ensemble, ran for from  $t = 0$  to 1000, with data assimilation performed every four steps, to the total of 251 assimilation cycles. Figure 3 also shows the best achievable value of RMSE error for a given ensemble, calculated as

$$\sigma_{\min} = \|\mathbf{X}\mathbf{s} - \mathbf{x}^t\|, \quad \mathbf{s} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{x}^t,$$

where  $\mathbf{X}$  is the ensemble matrix, and  $\mathbf{x}^t$  is the true solution; and averaged over 50 realizations. The performance of the DEnKF is almost identical to that of the ESRF, and is much better than the performance of the EnKF. Both the DEnKF and the ESRF perform nearly optimally, closely following the best possible RMSE.

The DEnKF shows a similarly good performance with the non-linear L40 model. Figure 4 depicts the average RMSE over a long run of the L40 model versus ensemble size and inflation factor for the ESRF, DEnKF and EnKF. Each run uses the same true field, and for a given ensemble size each filter starts from the same ensemble. The RMSE values are averaged over time steps  $t = 1000, \dots, 300\,000$ . Each panel in Fig. 4 shows the RMSE computed as a function of the ensemble size and the inflation factor. The white cells correspond to experiments in which the filter did not converge, here defined as the runs with RMSE greater than 1.

Figure 5 shows the best RMSE achieved in the runs from Fig. 4 for a given filter and ensemble size, for all inflation factors involved. Once again, the DEnKF and ESRF show almost identical performance, with the ESRF marginally better than the DEnKF, and both being substantially better than the EnKF.

An important characteristic of a data assimilating scheme is its stability towards the filter divergence. This stability is not necessarily directly related to the scheme's performance in terms of RMSE under normal conditions, but rather characterizes its ability to continue operation when the system turns out to be in an inconsistent state in the sense of Julier and Uhlmann (1997).

Because the previous experiments with the L40 model were not particularly challenging in regard to the filter divergence, we change the configuration of the system to increase the divergence rate by assimilating fewer observations, with smaller errors, and less often. Specifically, we use 10 equally spaced observations with error variance of 0.3, and assimilate every two time steps. This configuration substantially increases both the non-linearity

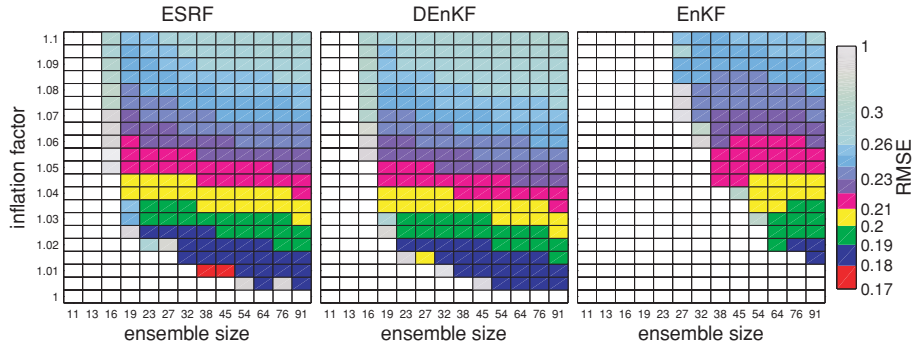


Fig. 4. RMSE for the L40 model using the ESRF, DEnKF and EnKF averaged over a long model run.

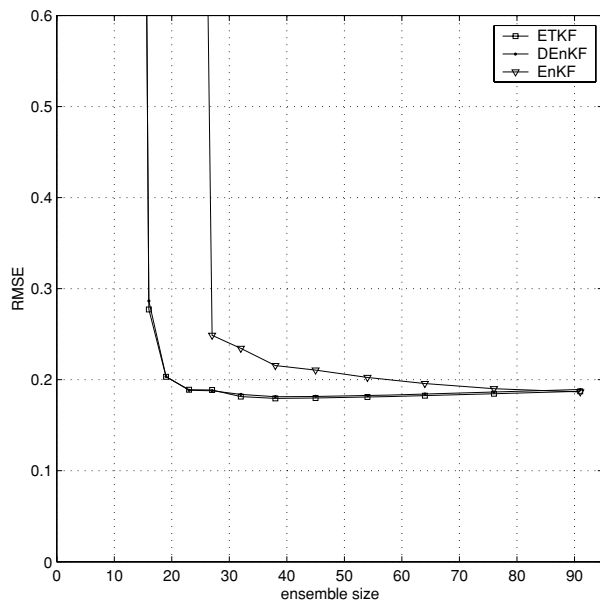


Fig. 5. The best RMSE from Fig. 4 for the L40 model for a given ensemble size.

of the system and the magnitude of the analysis correction compared to the configuration described above (40 observations with error variance of 1, assimilating every time step). An average reduction of RMSE after assimilation for the ETKF-based system

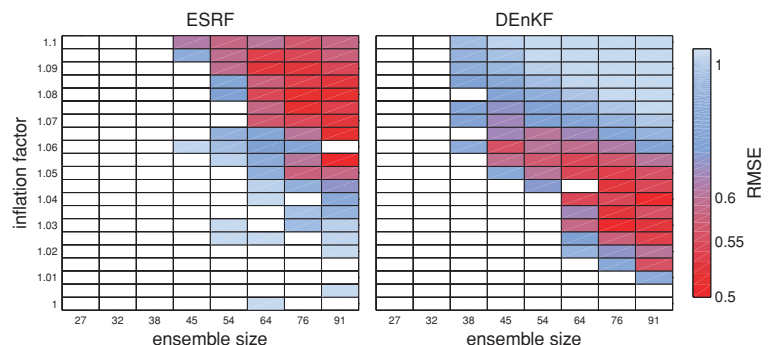
with the above configuration at steps 200–500 is about 0.10 compared to about 0.02 in the former configuration.

To investigate the convergence of the filter from the initial ensemble, we compare the average RMSE of the ESRF and the DEnKF over time interval  $t = [200, 500]$ , averaged over the same set of 50 realizations. The results of this experiment are shown in Fig. 6. The white cells indicate that at least in one realization out of 50 the model RMSE exceeds 10 at some point, which we qualify as the filter divergence. The figure shows that for convergence in the described configuration the ESRF requires much more inflation and a larger ensemble than the DEnKF. We conclude that although the DEnKF only approximates the ESRF in the case of small corrections and performs as well as ESRF in these situations, it may also be more stable in regard to filter divergence than the ESRF when the corrections are large.

Finally, we compare the performance of the DEnKF, ESRF and EnKF in a more ‘realistic’ environment in a series of experiments with the QG model. Figure 7 shows the analysis RMSE for the three systems as a function of the localization radius and inflation factor. Each value of RMSE represents an average over 10 realizations, for assimilation cycles 52–301 of each realization. The white cells correspond to the cases when the system became unstable or diverged in at least one realization out of 10.

The results presented in Fig. 7 show a very similar performance of both the ESRF- and DEnKF-based systems for a dense observation network. The performance of the EnKF-based system is slightly worse, with fewer configurations converging and

Fig. 6. Comparison of convergence from the initial ensemble for the ESRF and DEnKF with L40 model in difficult conditions. Shows the mean RMSE for time interval  $t = [200, 500]$ , averaged over 50 realizations. In these experiments 10 observations with observation error variance of 0.3 are assimilated every two time steps.





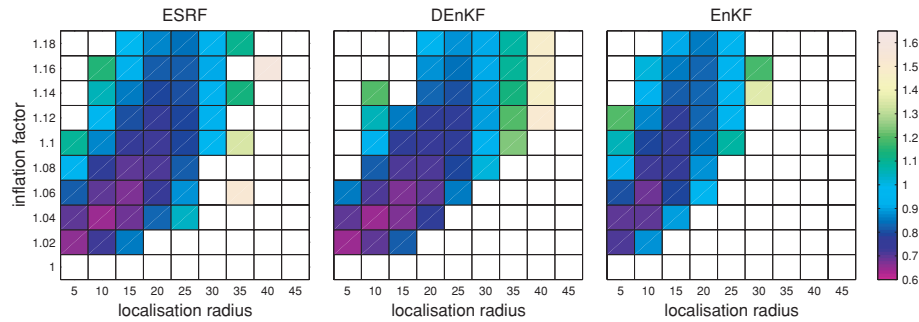


Fig. 7. Comparison of the RMSE of the analysis of  $\psi$  for the DEnKF, ESRF and EnKF for the QG model as a function of the localization radius and inflation factor; for the ensemble size of 25, averaged over 10 realizations and assimilation cycles 52–301 within each realization. White cells correspond to experiments where at least 1 realization out of 10 diverged.

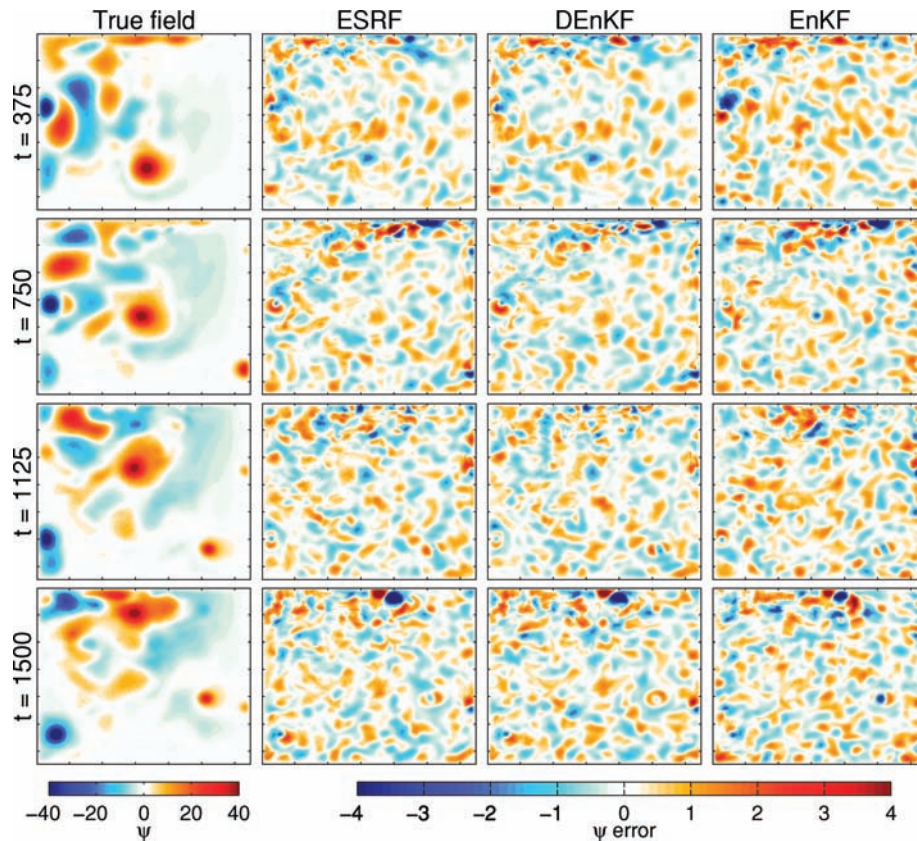


Fig. 8. Snapshots of the true field and the difference (error) fields between the analyses of ESRF, DEnKF and EnKF-based systems and the true field obtained from a particular run with the QG model; for time equal to 375, 750, 1125 and 1500; with the localization radius of 15 and inflation factor of 1.06 (ESRF, DEnKF) or 1.10 (EnKF).

greater RMSEs for those configurations that do converge. From a practical point of view, the superior performance of the DEnKF and ESRF here may not be very significant.

Examples of  $\psi$  from the ‘true’ model run at  $t = 375, 750, 1125$  and 1500 are shown in Fig. 8, along with the corresponding error fields for analyses using the ESRF, DEnKF and EnKF-based systems. Results for each assimilation run are drawn from the configuration with the smallest RMSE using a localization radius

of 15. This corresponds to an inflation factor of 1.06 for ESRF, 1.06 for DEnKF, and 1.10 for EnKF-based system.

The true and analysed fields in Fig. 8 are quantitatively similar for all systems. The magnitude of the errors are an order of magnitude smaller than the magnitude of the signal in the true field. Recall that the magnitude of the observation error variance is 4. The characteristic spatial length-scales of the error is smaller than that of the signal. The most significant errors occur along



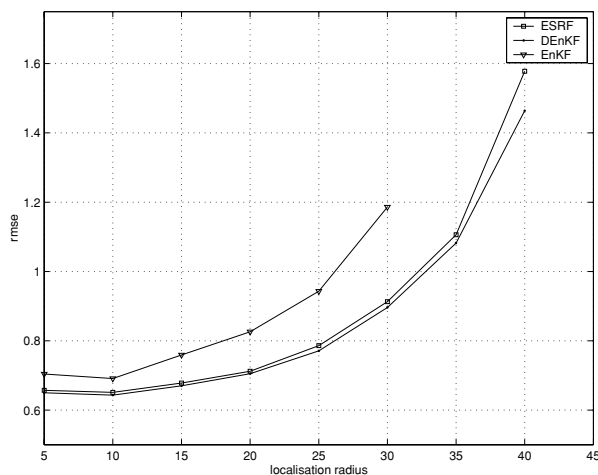


Fig. 9. The best RMSE for a given localization radius; for the same set of experiments as in Fig. 7.

the northern boundary; and are likely to be due to errors in the propagation of fast-moving waves. These results demonstrate that all three assimilation schemes tested here *can* result in a well-constrained analyses.

Probably the most important conclusion that we draw from this series of experiments relates to the range of different configurations (localization length and inflation factor) that reliably yields a satisfactory result for each scheme. Suppose we consider experiments satisfactory if the RMSE is less than 0.8. Figure 7 indicates that for the ESRF- and DEnKF-based systems, there is quite a broad range of configurations that satisfy this criterion. By contrast, the range of satisfactory configurations for the EnKF-based system is much narrower. Similar conclusions can be drawn from the experiments with the L40 model presented in Fig. 4. In practice, this may mean that more tuning experiments may be required for the EnKF; and that the performance of the EnKF-based system may be more sensitive to changes in the configuration.

The results in Fig. 7 demonstrate that the EnKF-based system produces slightly inferior results, compared to the ESRF- and DEnKF-based systems and that this difference in performance increases when the configuration of the assimilation (localization

and inflation) is less optimal. This can be seen more clearly from Fig. 9 that shows the best RMSE for each localization radius in the experiments shown in Fig. 7. The EnKF-based system also requires a greater inflation, which is consistent with the analysis of Whitaker and Hamill (2002). They found that an EnKF system with a small ensemble is more likely to underestimate the background error variance than to overestimate it.

The need for greater inflation for the EnKF is clearly demonstrated in Fig. 10, showing the difference between the average RMSE shown in Fig. 7 and the corresponding average ensemble spread (RMSE minus ensemble spread). Experiments with near zero difference correspond to black cells; experiments when the ensemble spread exceeds RMSE correspond to blue cells; and experiments when the ensemble spread is smaller than RMSE correspond to green cells. We note that for a given localization radius the smallest RMSE corresponds to experiments when the difference between the ensemble spread and the RMSE is near zero. We believe that this demonstrates the consistency of the data assimilation system; that is that the system performs best when the ensemble-based estimate of the background error variance matches the actual background error variance.

While the performance of ESRF and DEnKF-based systems in this experiment are quite similar, there are some subtle differences. The ESRF performs more reliably for bigger inflation factors at small localization radii, while DEnKF is able to perform reliably for bigger localization radii (40 versus 30 for ESRF).

Considering the difference in performance of the DEnKF- and ESRF-based systems, there are indications of a more robust performance of the DEnKF-based system at larger localization radii, and of the ESRF-based system at smaller localization radii. We believe that these differences are marginal in practice; and that at optimal or nearly optimal parameters in the case of small increments the performance the two systems is nearly identical.

## 5. Discussion

The numerical experiments described above with three quite different models show good performance and robustness of the

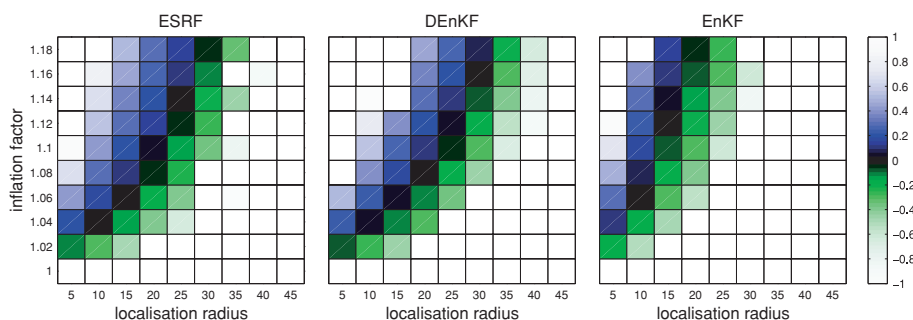


Fig. 10. Difference between the ensemble spread and the RMSE for the QG model for the same set of experiments as in Fig. 7.

DEnKF. In all tests the DEnKF demonstrates performance equal or very close to that of the ESRF, and a substantially better performance than the traditional EnKF. Importantly, the DEnKF also appears to be quite robust, since it converges as well as or even better than the ESRF, particularly in situations when the analysis correction is not small.

A possible reason for the robustness of the DEnKF can be seen from eq. (17), which shows that the DEnKF always *overestimates* the analysed error covariance, and the difference between the analysis error covariance of the DEnKF and the Kalman filter becomes larger when the magnitude of the analysis correction is larger. Such behaviour helps to avoid or reduce a possible underestimation of the estimated analysis error covariance compared to the actual analysis error covariance, which often leads to a divergence of the filter (Julier and Uhlmann, 1997). The variability in the magnitude of the overestimation of the analysis error covariance, which increases for larger corrections, can be seen as effectively an adaptive adjustment of the inflation, which decreases the risk of filter divergence. Interestingly, a linear approximation of the ETKF solution, which is formally equivalent to (18),

$$\begin{aligned} \mathbf{A}^a &= \mathbf{A}^f \left( \mathbf{I} + \frac{\mathbf{A}^{fT} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{A}^f}{m-1} \right)^{-1/2} \\ &= \mathbf{A}^f \left[ \mathbf{I} - \frac{1}{2} \frac{\mathbf{A}^{fT} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{A}^f}{m-1} \right. \\ &\quad \left. + \frac{3}{8} \left( \frac{\mathbf{A}^{fT} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{A}^f}{m-1} \right)^2 - \dots \right] \end{aligned} \quad (21)$$

(Bishop et al., 2001; Wang et al., 2004), *underestimates* the analysed error covariance because of the alternating signs in the expansion and therefore is not as robust as the DEnKF in regard to filter convergence.

The multiple of 1/2 in the ensemble transformation (15) can also be obtained from the serial ESRF proposed by Whitaker and Hamill (2002). When assimilating a single observation, the ensemble transformation in ESRF becomes equivalent to using a modified gain  $\tilde{\mathbf{K}} = \alpha \mathbf{K}$ , where

$$\alpha = \left[ 1 + \left( \frac{\mathbf{R}}{\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R}} \right)^{1/2} \right]^{-1},$$

with all matrices in this equation becoming scalars. In the case when the analysis correction is small,  $\mathbf{H} \mathbf{P}^f \mathbf{H}^T \ll \mathbf{R}$ , it yields  $\alpha \approx 1/2$ .

This serial analogue of the DEnKF makes it possible to quantify the assumption that the analysis corrections are ‘small’ for the DEnKF to be valid. Specifically, for each observation, the observation error variance must be much greater than the forecast error variance of the corresponding state vector element. Due to the coefficient of 1/8 at the second term of the expansion

$$\alpha = \frac{1}{2} + \frac{1}{8}x - \dots, \quad x \equiv \frac{\mathbf{R}}{\mathbf{H} \mathbf{P}^f \mathbf{H}^T},$$

in practice a weaker criteria

$$\mathbf{R} > \mathbf{H} \mathbf{P}^f \mathbf{H}^T \quad (22)$$

might be sufficient for the DEnKF to be valid. Apart from the initial period at the start of each assimilation run, this assumption is typically valid in the experiments presented above. The only experiment when the assumption (22) is not satisfied is the experiment on the filter divergence, presented in Fig. 6. In this experiment  $R = 0.3 \approx \mathbf{H} \mathbf{P}^f \mathbf{H}^T$ , but even in these conditions the performance of the DEnKF in terms of RMSE is very close to that of the ESRF.

One of the advantages of the traditional EnKF is that it readily permits the use of localization schemes based on the Schur product (an element-by-element multiplication of matrices) (Houtekamer and Mitchell, 2001; Hamill and Whitaker, 2001). In this case, the covariance matrix  $\mathbf{P}^f$  in the analysis eq. (1) is replaced by  $\rho \circ \mathbf{P}^f$ , where  $\rho$  is a correlation matrix, with elements  $\rho_{ij}$  calculated based on the distance in physical space between elements of  $i$  and  $j$  of the state vector. Because ESRFs update the ensemble anomalies using explicitly calculated ensemble transformations with no direct reference to the forecast error covariance  $\mathbf{P}^f$ , this technique is directly applicable to the ESRF only in the case of serial processing of observations, when the ensemble transformation can be written in terms of scaled down Kalman gain (Whitaker and Hamill, 2002). Another possible approach to localization in the ESRF is to use a window in physical space around a specified point and conduct data assimilation using the local representation of the state vector (Ott et al., 2004). In the DEnKF, the ensemble update written in the form (15) readily permits the use of the Schur product-based localization schemes, similarly to the EnKF.

## 6. Conclusions

We propose a new, deterministic modification of the traditional EnKF, which we refer to as the DEnKF. The DEnKF can be viewed as a linear approximation to the ESRF in the case when the analysis corrections are small. In terms of the analysis scheme, the DEnKF is closer to the traditional EnKF than to the ESRF. As a consequence, it allows the use of the traditional Schur product-based localization schemes in the same way as the EnKF. In regard to the performance of the DEnKF, a number of experiments with three small models show that the DEnKF performs equally as well as the ESRF, and substantially better than the EnKF. Furthermore, our experiments indicate that the DEnKF is quite robust; it may be less susceptible to the filter divergence than ESRFs and may perform well with a wider range of system configurations than the EnKF. Overall, we conclude that the DEnKF combines the performance of the ESRF with the simplicity and versatility of the EnKF and therefore represents an interesting alternative to these two established ensemble-based Kalman filters.

## 7. Acknowledgments

This research is funded by Australia's CSIRO through appropriation funding and by the US Office of Naval Research Ocean Modelling Program through Grant N000140410345.

## References

- Anderson, J. L. 2001. An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.* **129**, 2884–2903.
- Bierman, G. J. 1977. *Factorization Methods for Discrete Sequential Estimation*. Academic Press, New York, 68–112.
- Bishop, C., Etherton, B. and Majumdar, S. J. 2001. Adaptive sampling with the ensemble transform Kalman filter. Part I: theoretical aspects. *Mon. Wea. Rev.* **129**, 420–436.
- Burgers, G., van Leeuwen, P. J. and Evensen, G. 1998. Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.* **126**, 1719–1724.
- Evensen, G. 1994. Sequential data assimilation with a nonlinear quasigeostrophic model using Monte-Carlo methods to forecast error statistics. *J. Geophys. Res.* **99**, 10 143–10 162.
- Evensen, G. 2003. The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dyn.* **53**, 343–367.
- Evensen, G. 2004. Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dyn.* **54**, 539–560.
- Hamill, T. and Whitaker, J. 2001. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.* **129**, 2776–2790.
- Houtekamer, P. L. and Mitchell, H. L. 1998. Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.* **126**, 796–811.
- Houtekamer, P. L. and Mitchell, H. L. 2001. A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.* **129**, 123–137.
- Jelloul, M. B. and Thierry, H. 2003. Basin-mode interactions and selection by the mean flow in a reduced-gravity quasigeostrophic model. *J. Phys. Oceanogr.* **33**, 2320–2332.
- Julier, S. J. and Uhlmann, J. K. 1997. A new extension of the Kalman filter to nonlinear systems. In: *The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Controls*, Vol. **3068**, 182–193.
- Lawson, W. G. and Hansen, J. A. 2004. Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. *Mon. Wea. Rev.* **132**, 1966–1981.
- Lorenz, E. N. and Emanuel, K. A. 1998. Optimal sites for supplementary weather observations: simulation with a small model. *J. Atmos. Sci.* **55**, 399–414.
- Oke, P. R., Schiller, A., Griffin, D. A. and Brassington, G. B. 2005. Ensemble data assimilation for an eddy-resolving ocean model of the Australian region. *Q. J. R. Meteorol. Soc.* **131**, 3301–3311.
- Oke, P. R., Sakov, P. and Corney, S. P. 2006. Impacts of localization in the EnKF and EnOI: experiments with a small model. *Ocean Dyn.* **57**, 32–45.
- Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., and co-authors 2004. A local ensemble Kalman filter for atmospheric data assimilation. *Tellus* **56A**, 415–428.
- Pham, D. T., Verron, J. and Roubaud, M. C. 1998. A singular evolutive extended Kalman filter for data assimilation in oceanography. *J. Mar. Syst.* **16**, 323–340.
- Sakov, P. and Oke, P. R. 2008. Implications of the form of the ensemble transformation in the ensemble square root filters. *Mon. Wea. Rev.*, in press.
- Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M. and Whitaker, J. S. 2003. Ensemble square root filters. *Mon. Wea. Rev.* **131**, 1485–1490.
- Wang, X., Bishop, C. H. and Julier, S. J. 2004. Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble? *Mon. Wea. Rev.* **132**, 1590–1605.
- Whitaker, J. S. and Hamill, T. M. 2002. Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.* **130**, 1913–1924.