Q. J. R. Meteorol. Soc. (2005), **131**, pp. 3269–3289 doi: 10.1256/qj.05.135

Ensemble Kalman filtering

By P. L. HOUTEKAMER* and HERSCHEL L. MITCHELL
Division de la Recherche en Météorologie, Environment Canada, Canada

(Received 24 June 2005; revised 28 December 2005)

SUMMARY

An ensemble Kalman filter (EnKF) has been implemented at the Canadian Meteorological Centre to provide an ensemble of initial conditions for the medium-range ensemble prediction system. This demonstrates that the EnKF can be used for operational atmospheric data assimilation.

We show how the EnKF relates to the Kalman filter. In particular, to make the ensemble approximation feasible, we have to use a fairly small ensemble with many less members than either the number of model coordinates, or the number of independent observations, or the (unknown) dimension of the dynamical system. To nevertheless obtain good results, we must (i) counter the tendency of the ensemble spread to underestimate the true error, and (ii) localize the ensemble covariances. The localization is severe and leads to imbalance in the initial conditions.

The operational EnKF is used to investigate to what extent our system respects the underlying hypotheses of both the Kalman filter and its ensemble approximation. In particular, we quantify the imbalance in the initial conditions and the magnitude of the model-error component. The occurrence of imbalance constrains the ways in which time interpolation can be performed and in which parametrized model error can be added. With this study we hope to obtain and provide guidance for further improvements to the EnKF.

KEYWORDS: Balance Data assimilation Model error

1. INTRODUCTION

The Kalman filter (KF) (Kalman 1960; Kalman and Bucy 1961) provides 'an elegant and comprehensive mathematical description of the data assimilation problem' (Daley 1991, p. 384). As suggested by Ghil *et al.* (1981) and Cohn (1982), the Kalman filter could in theory be used for the optimal estimation of atmospheric states. In practice, however, it suffers from two serious drawbacks (Ghil and Malanotte-Rizzoli 1991, p. 162; Daley 1991, p. 384). The first is the estimation of the error covariances of the forecast model, which is crucial but difficult (Dee 1995). The second is the computational expense. In spite of the continuing development of computational platforms, the application of the KF to atmospheric data assimilation is still not possible. It has become possible, however, to employ an ensemble approximation to the KF. The so-called ensemble Kalman filter (EnKF) (Evensen 1994) is now under development or in use for hydrologic data assimilation (Reichle *et al.* 2002), ocean data assimilation (Keppenne and Rienecker 2002), retrospective atmospheric analysis (Whitaker *et al.* 2004), convective-scale atmospheric data assimilation (Snyder and Zhang 2003; Dowell *et al.* 2004) and global atmospheric data assimilation (Houtekamer *et al.* 2005).

Our own experience with the EnKF is mostly limited to the field of large-scale atmospheric data assimilation. One would hope that a good theory could be applied equally to the different fields. In practice, however, the precise formulation of an algorithm is often the result of a consideration of different important factors. The relative importance of these factors may depend on the specific properties of the dynamical model and the observational network. For instance, for large-scale atmospheric assimilation the issue of balance is important, but reasonable results can still be obtained when no specific action is taken (Derber *et al.* 1991; Parrish and Derber 1992; Houtekamer *et al.* 2005). For the application of the EnKF in other fields, the issue of balance may either be not relevant or conversely extremely important. Similarly, in large-scale atmospheric

* Corresponding author: Division de la Recherche en Météorologie, 2121 Route Trans-Canadienne, Dorval, QC, H9P 1J3, Canada. e-mail: peter.houtekamer@ec.gc.ca

data assimilation, a certain quality of results can be obtained while assuming that the forecast model has no bias, however, we cannot assume that the second moment of the model error is zero (Dee 1995). In this paper, attention will be restricted to the application of the EnKF to large-scale atmospheric data assimilation. We are not in a position to make statements about the other areas of application.

At the Meteorological Service of Canada (MSC), the development of an EnKF has resulted in an operational implementation in the medium-range ensemble prediction system (EPS) as of 12 January 2005. This demonstrates that the EnKF has passed a certain baseline level of quality. In this study, the operational configuration is used to investigate the realism of the assumptions and principles that are behind the KF and its approximation with an ensemble. We hope that this study will help identify promising areas for future research. In fact, based on the results of this study, some revisions to the EnKF algorithm were implemented operationally on 13 December 2005 at the Canadian Meteorological Centre (CMC). The main objective of these revisions was to improve the balance of the background fields.

As we transformed our research code to be suitable in a complex operational environment, it became more difficult for us to test radically different algorithms. Since we continued to obtain reasonable results with our original EnKF formulation consisting of a pair of ensembles (Houtekamer and Mitchell 1998), we have not tested any of the more recently proposed alternative EnKF formulations, e.g. Tippett *et al.* (2003), Szunyogh *et al.* (2005). Such comparisons are perhaps best performed in a simpler experimental environment, as in the recent paper by Lawson and Hansen (2004). However, we do believe that the current study will also be of interest to people using other EnKF algorithms.

In the next section we describe the basic components of the EnKF. The EnKF that became operational at the Meteorological Service of Canada is presented in section 3. In section 4 we try to estimate the importance of different terms in that EnKF. As it appears that time interpolation of the model trajectory to the observations could be beneficial, we investigate in section 5 if the model integrations are sufficiently balanced. In section 6 we provide a discussion and new interpretation of the model-error term. We summarize our analysis of the EnKF algorithm in section 7.

2. Basic components of the Ensemble Kalman filter

In this section, we discuss the concepts related to the EnKF and the KF. The EnKF provides an approximation to the KF that improves as the ensemble size increases. For small ensemble sizes, however, the properties of the EnKF differ from those of the KF. In particular, in the EnKF, the equation for the gain is modified to permit localization of covariances.

(*a*) *Kalman filter equations*

In principle it is possible to estimate and correct for bias in the forecast model (Dee and da Silva 1998). For atmospheric data assimilation, it is nevertheless traditional to make the simplifying assumptions of having no bias in either the nonlinear forecast model $\mathcal{M}$, or the observations y^{o}. Letting q denote the model error and r denote the observational error, we thus assume:

$$Eq = 0, \tag{1}$$

$$Er = 0. \tag{2}$$

Here the symbol E denotes the expectation operator. In the case of a linear prediction model $\mathbf{M}$ and a linear interpolation operator $\mathbf{H}$, the KF provides the optimal solution to the analysis problem even for non-Gaussian model- and observation-error statistics (Gelb 1974; Ghil *et al.* 1981; Cohn 1982; Cohn and Parrish 1991; Ghil and Malanotte-Rizzoli 1991). The KF equations can be written as

$$x^{\mathrm{f}}(t) = \mathbf{M}x^{\mathrm{a}}(t-1), \tag{3}$$

$$\mathbf{P}^{\mathrm{f}}(t) = \mathbf{M}\mathbf{P}^{\mathrm{a}}(t-1)\mathbf{M}^{\mathrm{T}} + \mathbf{Q}, \tag{4}$$

$$\mathbf{K}(t) = \mathbf{P}^{\mathrm{f}}(t)\mathbf{H}(t)^{\mathrm{T}}(\mathbf{H}(t)\mathbf{P}^{\mathrm{f}}(t)\mathbf{H}(t)^{\mathrm{T}} + \mathbf{R}(t))^{-1}, \tag{5}$$

$$x^{\mathrm{a}}(t) = x^{\mathrm{f}}(t) + \mathbf{K}(t)(y^{\mathrm{o}}(t) - \mathbf{H}(t)x^{\mathrm{f}}(t)), \tag{6}$$

$$\mathbf{P}^{\mathrm{a}}(t) = (\mathbf{I} - \mathbf{K}(t)\mathbf{H}(t))\mathbf{P}^{\mathrm{f}}(t). \tag{7}$$

Here $\mathbf{M}$, or more precisely $\mathbf{M}(t, t-1)$, maps a model state at time $t-1$ to time t; $x^{\mathrm{a}}(t-1)$ is the best estimate of the true state at time $t-1$; $x^{\mathrm{f}}(t)$ is the best estimate of the true state at time t, given only the data available until time $t-1$; $\mathbf{Q}$, or more precisely $\mathbf{Q}(t, t-1)$, is the covariance matrix of the cumulative model error q between time $t-1$ and time t; $\mathbf{R}$ is the covariance matrix of the observational error r; $\mathbf{P}^{\mathrm{f}}$ is the covariance matrix of the forecast error; $\mathbf{P}^{\mathrm{a}}$ is the covariance matrix of the analysis error; and $\mathbf{K}$ is known as the Kalman gain.

In the extended Kalman filter (EKF), e.g. Gauthier *et al.* 1993, the reference trajectory is described using a nonlinear model $\mathcal{M}$, and a nonlinear forward operator $\mathcal{H}$, is used to interpolate $x^{\mathrm{f}}(t)$ to the observation space. Equations (3) and (6) are thus replaced by:

$$x^{\mathrm{f}}(t) = \mathcal{M}(x^{\mathrm{a}}(t-1)), \tag{8}$$

$$x^{\mathrm{a}}(t) = x^{\mathrm{f}}(t) + \mathbf{K}(t)\{y^{\mathrm{o}}(t) - \mathcal{H}(t)(x^{\mathrm{f}}(t))\}. \tag{9}$$

Equations (4), (5) and (7) are still evaluated with linear operators $\mathbf{M}$ and $\mathbf{H}$. These now, for the EKF, are tangent linear to the nonlinear $\mathcal{M}$ and $\mathcal{H}$ about the reference trajectory defined by $\mathcal{M}$.

The KF equations have a number of appealing properties. Two that are of particular interest are:

(i) The evolution and growth of the errors during the forecast is accounted for by Eq. (4) and the subsequent reduction of the error in the analysis due to the use of the Kalman gain is accounted for by Eq. (7). Thus, as can be seen from Eqs. (3)–(7), given x^{a} and $\mathbf{P}^{\mathrm{a}}$ at time $t-1$, the KF equations lead to these same quantities at time t, which allows the algorithm to be cycled in time. Pure three-dimensional methods such as Optimal Interpolation (Gandin 1965) or 3D variational methods (Parrish and Derber 1992) implement neither Eq. (4) nor Eq. (7). The four-dimensional variational method (Lewis and Derber 1985) does not implement Eq. (7).

(ii) When the model error projects only on to slow modes, as was argued by Phillips (1986), the KF produces balanced analyses automatically at each analysis time (Petersen 1973; Cohn and Parrish 1991).

To evaluate the covariance transport equation, Eq. (4), it is necessary to integrate the tangent linear forecast model once for each of the O(1 000 000) coordinates of the forecast model. This makes application of the KF equations for atmospheric data assimilation prohibitively expensive. The KF has consequently been used mainly as a pedagogical device (Cohn and Parrish 1991) or 'prototype algorithm' (Cohn 1997).

The success of a hypothetical implementation of the KF would depend on the realism of the hypothesis of no bias (Eqs. (1) and (2)) and on the quality of the covariance matrices **Q** and **R**. The use of the maximum-likelihood method for the on-line estimation of some parameters of a parametrized representation of the matrix **Q** was proposed and demonstrated by Dee (1995). This approach was applied in an EnKF by Mitchell and Houtekamer (2000). At the present time, however, it is not clear how to best describe model error.

(*b*) *Approximation with an ensemble*

In the EnKF (Evensen 1994), a small random ensemble with O(100) members is used to represent the best estimate of the state vector and information about its covariance. The ensemble mean states, $\overline{x_i^{\mathrm{f}}}$ and $\overline{x_i^{\mathrm{a}}}$, correspond to the KF estimates, x^{f} and x^{a}. The covariances $\mathbf{P}^{\mathrm{f}}$ and $\mathbf{P}^{\mathrm{a}}$ can be estimated from the spread of the ensembles x_i^{f} and x_i^{a}, respectively (Burgers *et al.* 1998). Performing the O(100) short-range integrations that are required to approximate Eq. (4) is within reach of modern parallel computers. As the ensemble that can feasibly be run becomes larger, the approximation to the KF becomes better.

Dropping the time index when it is t, we write the equations for the EnKF as:

$$x_i^{\mathrm{f}} = \mathcal{M}(x_i^{\mathrm{a}}(t-1)) + q_i, \quad i = 1, \ldots, N, \tag{10}$$

$$q_i \sim N(0, \mathbf{Q}), \tag{11}$$

$$\mathbf{P}^{\mathrm{f}}\mathcal{H}^{\mathrm{T}} \equiv \frac{1}{N-1}\sum_{i=1}^{N}(x_i^{\mathrm{f}} - \overline{x_i^{\mathrm{f}}})(\mathcal{H}x_i^{\mathrm{f}} - \overline{\mathcal{H}x_i^{\mathrm{f}}})^{\mathrm{T}}, \tag{12}$$

$$\mathcal{H}\mathbf{P}^{\mathrm{f}}\mathcal{H}^{\mathrm{T}} \equiv \frac{1}{N-1}\sum_{i=1}^{N}(\mathcal{H}x_i^{\mathrm{f}} - \overline{\mathcal{H}x_i^{\mathrm{f}}})(\mathcal{H}x_i^{\mathrm{f}} - \overline{\mathcal{H}x_i^{\mathrm{f}}})^{\mathrm{T}}, \tag{13}$$

$$\mathbf{K} = \mathbf{P}^{\mathrm{f}}\mathcal{H}^{\mathrm{T}}(\mathcal{H}\mathbf{P}^{\mathrm{f}}\mathcal{H}^{\mathrm{T}} + \mathbf{R})^{-1}, \tag{14}$$

$$y_i^{\mathrm{o}} = y^{\mathrm{o}} + r_i, \quad i = 1, \ldots, N, \tag{15}$$

$$r_i \sim N(0, \mathbf{R}), \tag{16}$$

$$x_i^{\mathrm{a}} = x_i^{\mathrm{f}} + \mathbf{K}(y_i^{\mathrm{o}} - \mathcal{H}x_i^{\mathrm{f}}), \quad i = 1, \ldots, N. \tag{17}$$

As can be seen from these equations, given an ensemble of analyses at time $t-1$, the EnKF algorithm yields an ensemble of analyses at time t, i.e. the EnKF, like the KF, can be cycled in time.

By means of Eq. (10), the EnKF uses the full nonlinear model $\mathcal{M}$, to transport the covariances. The replacement of a linear model **M** (in the extended Kalman filter the tangent linear model of $\mathcal{M}$) by the full nonlinear model would seem to be an improvement. The full model will properly deal with saturation of errors. The EnKF also offers more flexibility for dealing with model error. It can be sampled from a covariance matrix, as in Eq. (11), or also simulated using, for instance, different model versions. Unlike the KF, the EnKF (via Eqs. (12) and (13)) uses a random ensemble to estimate the error covariances. In fact, since only $\mathbf{P}^{\mathrm{f}}\mathcal{H}^{\mathrm{T}}$ and $\mathcal{H}\mathbf{P}^{\mathrm{f}}\mathcal{H}^{\mathrm{T}}$ are required explicitly, $\mathbf{P}^{\mathrm{f}}$ itself (a very large matrix) need never be calculated. With the observations perturbed in Eq. (15) in accordance with their uncertainty, the EnKF provides, by means of Eq. (17), a Monte Carlo-like estimate of the uncertainty in the analysis (Burgers *et al.* 1998). To avoid changing the best estimates, $\overline{x_i^{\mathrm{f}}}$ and $\overline{x_i^{\mathrm{a}}}$, we use a first-order-exact (i.e. with

imposed zero mean) sample of perturbations q_i and r_i. Second-order-exact sampling of random numbers has been discussed by Pham (2001).

There is also a family of non-stochastic filters (discussed in a unified framework by Tippett *et al.* (2003)) which are not based on Eqs. (10)–(17). These filters do not use perturbed observations, which, it is argued, can be a source of sampling error when ensembles are small. Instead, they deterministically transform the ensemble of background fields into an ensemble of analyses using Eq. (7). However, Eq. (7) is valid only when the gain $\mathbf{K}$ is optimal, which depends, in turn, on $\mathbf{Q}$ and $\mathbf{R}$ being accurately known. A more general, but more complicated, equation for $\mathbf{P}^{\mathrm{a}}$, that reduces to Eq. (7) when the gain is optimal, is given by, e.g. Cohn (1982, Eq. (2.10b)), Daley (1991, Eq. (13.3.19)) and Ghil and Malanotte-Rizzoli (1991, Eq. (4.13b)). The performance of stochastic and deterministic filters has been compared in a hierarchy of perfect-model scenarios by Lawson and Hansen (2004).

The use of a small ensemble, x_i^{f}, $i = 1, \ldots, N$, to estimate the covariances $\mathbf{P}^{\mathrm{f}}\mathcal{H}^{\mathrm{T}}$ and $\mathcal{H}\mathbf{P}^{\mathrm{f}}\mathcal{H}^{\mathrm{T}}$ results in a sampling error in the gain matrix $\mathbf{K}$. The use of the same ensemble in $\mathbf{P}^{\mathrm{f}}\mathcal{H}^{\mathrm{T}}$ and $\mathcal{H}\mathbf{P}^{\mathrm{f}}\mathcal{H}^{\mathrm{T}}$ for the estimation of $\mathbf{K}$ and in (17) for the estimation of the analysis error results in a dependent error estimate (i.e. use of the same sample to calculate the gain and estimate the error associated with using that gain) and, hence, a systematic underestimation of the uncertainty in the analysis, i.e. an underdispersive ensemble of analyses. In a data assimilation context, this would result in too much weight being given to the background with respect to the observations. This inbreeding problem can be countered by using a pair of ensembles (Houtekamer and Mitchell 1998; van Leeuwen 1999; Houtekamer and Mitchell 1999). Here the gain that is used for the assimilation of observations into one ensemble is obtained from the other ensemble.

Alternatively, with an N-member ensemble, one might assimilate the observations into each ensemble member using covariances calculated from all of the $N - 1$ remaining ensemble members (Hamill and Snyder 2000). This approach maximizes use of the available ensemble members but, if used with a direct algorithm, requires the computation of N different gain matrices.

The implementation of Eqs. (10)–(17) permits analysis increments only in the space spanned by the N-member ensemble x_i^{f} of background fields. If the distribution of the ensemble is Gaussian, with negligible curvature, as is common when errors are small and error dynamics are linear, then the EnKF (like the KF) will provide a balanced ensemble of analyses. In the case of large errors, non-Gaussianity of the distribution and significant curvature of the local balanced space can lead to analyses that are off the attractor (Lawson and Hansen 2004). Nonlinear filtering methods have been applied in the context of simple models (Anderson and Anderson 1999; Miller *et al.* 1999), but are too computationally demanding for application in global atmospheric data assimilation.

The property of optimality, however, is lost in the approximation with an ensemble even if errors are small. This is illustrated by Lorenc (2003), who provides an example in which the ensemble of analyses has a bigger uncertainty than the ensemble of guess fields. In that case, a more accurate ensemble of analyses would have been obtained by giving no weight to the observations or by localizing their impact.

(*c*) *Need for localization and the impact on balance*

Having a small ensemble and a high-dimensional system, we cannot expect to closely fit a large number of observations. Therefore, unless the ensemble is truly large, with a size O(10 000), it is necessary to localize the impact of the observations (Houtekamer and Mitchell 1998; Hamill *et al.* 2001; Lorenc 2003).

The impact of observations can be smoothly reduced for increasingly distant grid points using a Schur or Hadamard product (Gaspari and Cohn 1999; Houtekamer and Mitchell 2001). In our algorithm, we modify Eq. (14) to permit both a horizontal and a vertical localization as follows:

$$\mathbf{K} = \{\rho_{\mathrm{V}} \circ \rho_{\mathrm{H}} \circ (\mathbf{P}^{\mathrm{f}} \mathcal{H}^{\mathrm{T}})\}\{\rho_{\mathrm{V}} \circ \rho_{\mathrm{H}} \circ (\mathcal{H}\mathbf{P}^{\mathrm{f}} \mathcal{H}^{\mathrm{T}}) + \mathbf{R}\}^{-1}. \tag{18}$$

Here ρ_{H} and ρ_{V} are the correlation functions used for horizontal and vertical localization, respectively, and $\circ$ denotes the Schur product.

Alternatively, one could divide the global problem into a number of local problems that can be solved independently. This strategy, proposed by Ott *et al.* (2004), leads to a more efficient parallelization.

Localization of any type can potentially cause imbalance in the initial conditions (Cohn *et al.* 1998; Mitchell *et al.* 2002; Lorenc 2003), because the analyses will not be in the balanced space defined by the guess fields.

3. Canadian implementation

On 12 January 2005, the CMC implemented the EnKF to provide the initial conditions for the medium-range EPS. The first operational medium-range forecasts were generated from the EnKF analyses valid at 0000 UTC 13 January. Since January 1996, the EPS had been using parallel data assimilation cycles based on the Optimal Interpolation method and a spectral forecast model to obtain the initial conditions for an ensemble of medium-range forecasts. That EPS, the medium-range ensemble-forecast component of which was not changed in January 2005, had been described by Houtekamer *et al.* (1996) and more recently by Pellerin *et al.* (2003).

In the current paper, we investigate properties of the operational EnKF. In this section, we briefly describe its configuration. For a more detailed discussion and evaluation of a pre-operational version, the reader is referred to Houtekamer *et al.* (2005).

(*a*) *The model*

One of the two forecast models used in the Canadian EPS (Pellerin *et al.* 2003) is a low-resolution version of the Global Environmental Multiscale grid-point model (Côté *et al.* 1998a,b). For the EnKF, we use this model with the same 300×150 horizontal grid and with the same 28 η-levels and model top at 10 hPa. We found that the problem of narrow vertical structures near the top of the model (Houtekamer *et al.* 2005, Fig. 2) was alleviated by using cubic-Lagrangian, instead of cubic spline, vertical interpolations to find the field values at the upstream locations in the semi-Lagrangian time-stepping scheme used in the forecast model. For this, we follow the variational analysis group at our Centre which uses this interpolation procedure in the tangent linear model. For now, the regional and global deterministic forecasts at our Centre continue to be performed using cubic spline vertical interpolations.

(*b*) *The observations*

The EnKF has been implemented with a 6 h cycle, with analyses valid at 0000, 0600, 1200 and 1800 UTC each day. In the EnKF data assimilation procedure, it is assumed that all observations are valid exactly at one of these central analysis times. We only assimilate observations that have passed the background check and variational quality control procedures of our Centre's operational variational deterministic analysis.

In a pre-operational assimilation cycle with the EnKF, we had a problem near 135°W, 63°S in the South Pacific with the assimilation of radiance data from Advanced Microwave Sounding Unit-A (AMSU-A) channel 3. This channel, which is used only over the open ocean, is sensitive to surface conditions in addition to the atmospheric temperature profile. The RTTOV-7 forward operator (Saunders *et al.* 2002), that we use in the assimilation of radiance data, models the effect of ocean foam on AMSU-A channel 3 observations assuming that ocean foam has an emissivity of 1 and that the fractional foam coverage F_{foam} is given by (S. English 2004, personal communication)

$$F_{\text{foam}} = 0.195 \times 10^{-4}\ W^{2.55}, \tag{19}$$

where W is the intensity of the surface winds. Due to the use of this formula, RTTOV-7 is very sensitive to intense surface winds for AMSU-A channel 3. In the presence of such winds in a region of significant uncertainty, and lacking direct measurements of surface winds, the EnKF could act to amplify the wind speed, irrespective of the wind direction. That is, because the analysis is formulated in terms of the two horizontal wind components, the modelled sensitivity of this channel to surface wind speed may actually lead to an increased ensemble spread for the wind components. Subsequent channel 3 observations will consequently receive more weight. This unstable process could lead to surface winds stronger than had been considered in the development of Eq. (19). To control the problem, we subsequently limited the surface wind speed that is provided to the RTTOV-7 operator to at most 10 m s^{-1} and, following a suggestion by J.-N. Thépaut (2004, personal communication), we started assimilating the available surface wind observations.

The AMSU-A radiances are primarily sensitive to the atmospheric temperature profile. The operational EnKF also assimilates four AMSU-B radiance channels. These are primarily sensitive to atmospheric humidity. Quality-controlled AMSU-B radiances became available for assimilation by the EnKF in 2004 when they began being assimilated by the CMC operational variational deterministic analysis.

As discussed in Houtekamer *et al.* (2005, section 2(d)), each AMSU channel is assigned the (approximate) pressure where that channel peaks for the purposes of the vertical localization. Due to the highly variable nature of humidity profiles, we had anticipated difficulty assigning these pressures to the four AMSU-B channels. In practice, since humidity is largely confined to the lower atmosphere, our use of the natural logarithm of pressure as the vertical coordinate for the localization implies that our vertical localizing function is broad in the lower atmosphere. Consequently, results were less sensitive to the precise value assigned to the peak pressure for each channel than had been expected. For the purposes of the vertical localization, the four AMSU-B channels (channels 2–5) were assigned peak pressures of 800, 475, 525 and 600 hPa, respectively. As in the operational variational deterministic analysis, channels 2 and 5 are not used over land and sea-ice. Further information about the pre-processing and quality control of AMSU radiances can be found in Anselmo and Deblonde (2005).

(*c*) *Model error*

The ensemble of model-error realizations q_i is added to the ensemble of predictions, as indicated by Eqs. (10) and (11). In adding the model-error component to the prediction error immediately after the predictions are generated, we follow the standard Kalman filter formulation, Eq. (4).

Our current description of model error is not flow dependent. In fact, our model-error parametrization, as discussed in Mitchell and Houtekamer (2000) and in

Houtekamer *et al.* (2005), is based on the forecast-error description used by the 3D-Var at our Centre (Gauthier *et al.* 1999). Due to this choice, our EnKF implementation is similar in some ways to a hybrid scheme (Hamill and Snyder 2000) with the importance of the homogeneous isotropic statistics depending on the size of our parametrized model error.

To add model error, we start by generating a random 3D perturbation field for stream function. Addition of this field should, by design, not degrade the balance of the meteorological field. In a subsequent step, however, we add a random 3D field of unbalanced temperature perturbations (Gauthier *et al.* 1999; Houtekamer *et al.* 2005). Addition of this field does, as expected, have a negative impact on the balance of a state.

We introduced the unbalanced temperature component at an early stage of our project because our forecast model did not perfectly respect the no-bias condition of Eq. (1). Having an unbalanced temperature component in the formulation of the background error enabled us to control a temperature bias near the top of the model. Recent experiments have shown that the unbalanced temperature component continues to have a beneficial impact.

Our current model-error description does not include a component for humidity. Consequently, the ensemble spread for humidity in our EnKF is too small (Houtekamer *et al.* 2005).

(*d*) *The method*

We use a double EnKF configuration consisting of a pair of 48-member ensembles. To avoid having to store and invert very large matrices, batches of up to 300 observations are assimilated sequentially, as described in Houtekamer and Mitchell (2001). After a given batch of observations has been assimilated, the resulting pair of analysis ensembles is used as the pair of background ensembles for the assimilation of the next batch of observations, until all observations have been assimilated.

For the covariance localization functions, ρ_{H} and ρ_{V}, we use a piecewise rational function (Gaspari and Cohn 1999, Eq. (4.10)). In the formulation of ρ_{V}, the natural logarithm of pressure is used as the vertical coordinate. The localization is such that covariances are forced to zero in two units of ln p. Similarly the horizontal impact of an observation drops to zero at 2800 km.

As implemented in January 2005, the operational EnKF did not include any explicit balancing operator but, as will be seen later, we have recently reconsidered this.

4. Length of increments

In this section, we first measure the length of the forecast and analysis increments using a dry total-energy norm as in Mitchell *et al.* (2002, Eq. (9)). Subsequently, we quantify and order the different approximations that were made in our implementation of the EnKF. These include (i) the impact of not initializing, (ii) the assumption of no bias in the model or the observations, (iii) the use of a small ensemble and consequent sampling error, and (iv) the approximation that all observations are valid at the central time of the 6 h assimilation window.

(*a*) *Basic components*

Using the default operational parameters of the EnKF and fields valid at 0000 UTC 1 August 2004, we find for the analysis increment of the first ensemble member a length of 2.65 m s^{-1} in terms of the total energy norm. The subsequent 6 h model integration has a length of 4.62 m s^{-1}. To the resulting predicted field we add an

TABLE 1. IMPACT OF VARIOUS COMPONENTS IN AN ASSIMILATION CYCLE

Analysis increment (A)	2.65
6 h prediction	4.62
Model-error perturbation	1.39
Total impact of 6 h prediction and model-error perturbation (F)	4.81
Impact of digital filter on a 6 h forecast (I)	0.35
Impact of physical parametrizations on a 6 h forecast	1.52

All values are computed using an energy norm (m s^{-1}) and reflect changes to the first ensemble member over the course of a single 6 h cycle.

independent perturbation field of length 1.39 m s^{-1} which reflects the model error. In total, the distance between the analysis field valid at 0000 UTC 1 August 2004 and the subsequent trial field valid 6 h later is 4.81 m s^{-1}. For ease of reference, the lengths of the various increments are listed in Table 1. We thus verify the first part of the criterion of Hollingsworth *et al.* (1986) for a successful data assimilation cycle which states that the forecast increment, here of length 4.81 m s^{-1}, should be larger than the analysis increment, here of length 2.65 m s^{-1}.

It would appear that the length of the model error, here 1.39 m s^{-1}, is fairly significant compared to the length of the total forecast, here 4.81 m s^{-1}. It has, however, been argued by Houtekamer *et al.* (2005) that the 'model' error in the EnKF collectively accounts for a multitude of error terms in the entire EnKF data assimilation system. It is not clear at this point which component of the model error should be addressed with the highest priority.

(*b*) *Initialization*

Since 4D data assimilation methods are intrinsically more consistent with the atmospheric flow than 3D methods, they improve the possibility of producing balanced analysis increments (Mitchell *et al.* 2002). Therefore, one may hope that there is no need for an explicit balancing operator, which may itself not be entirely consistent with the model dynamics. Consequently, in the January 2005 operational implementation of the EnKF, we did not include such an operator.

To measure the impact of initialization, if it were applied, an additional integration has been performed from the 0000 UTC 1 August 2004 initial condition with a digital filter finalization technique (Fillion *et al.* 1995). Here, the filter is applied to a 6 h integration consisting of eight unfiltered time steps (i.e. a nine-point time series implementation), to obtain a filtered state valid at $t = 3$ h. The filter uses a 6 h cut-off period. From the filtered state at $t = 3$ h, we subsequently run the model out to 6 h.

The initialization increment is measured as the difference, evaluated at 0600 UTC 1 August 2004, between the 6 h forecast with and without the digital filter. Its value of 0.35 m s^{-1} is well below the 2.65 m s^{-1} of the analysis increment. Thus our EnKF respects the second part of the criterion of Hollingsworth *et al.* (1986) which states that the analysis increment should be bigger than the initialization increment.

As noted above, a balancing operator may not be entirely consistent with the model dynamics and consequently could have an impact even if it were applied to a state obtained from a long and balanced integration of a forecast model. In the case of digital filter finalization, which does not require any adiabatic assumption at any stage, the impact on a dynamically balanced model state is small (Fillion *et al.* 1995, section 3.1). The application of a digital filter will be further discussed in section 5 in the context of a hypothetical EnKF which includes time interpolation of the model trajectory to the observations.

(*c*) *Bias*

From the outset, with Eqs. (1) and (2), we have assumed that the forecast model and the observations have no bias. As discussed in Kailath (1968) and Daley (1992a), in an optimal linear system, without bias, the innovation vectors are not serially correlated. Here, because the observational network is non-stationary, we will not use innovation vectors but analysis increments.

To detect bias, we produce a sequence of ensemble mean analysis increments, $\overline{x^{\mathrm{a}}}(t) - \overline{x^{\mathrm{f}}}(t)$, with successive increments calculated from analyses one day apart. The sequence goes from 0000 UTC 1 August 2004 until 0000 UTC 30 August 2004 and thus contains 30 values.

We assume that each increment consists of (i) a random component s, with fixed zero-mean distribution $N(0, \sigma^2)$ and no serial correlation, and (ii) a bias-induced constant component C, which is identical for each increment:

$$\overline{x^{\mathrm{a}}}(t) - \overline{x^{\mathrm{f}}}(t) = C + s(t). \tag{20}$$

As noted by Dee and da Silva (1998, Eq. (12)), the mean analysis increment C is a linear function of the bias in the observations and in the model. Generally it will underestimate the forecast bias, if the observations are unbiased.

With these assumptions, we expect for the first analysis increment, valid at 0000 UTC 1 August 2004, a squared length L_1^2 of $\|C\|^2 + \sigma^2$. For the average over n analysis increments, we obtain an expected squared length L_n^2 of $\|C\|^2 + (1/n)\sigma^2$. The experimentally obtained lengths of the mean analysis increments are displayed in Fig. 1. Equating the observed first length of 1.65 m s^{-1} to L_1 and equating the observed length of 0.54 for the mean over 30 increments to L_{30}, we obtain $\|C\| = 0.46$ m s^{-1} and $\sigma = 1.58$ m s^{-1}. With these values, we can subsequently compute the intermediate values of L_n. From Fig. 1, we note that the modelled values agree rather well with the observed values.

From this simple model and analysis, we find that the ensemble-mean analysis increment, which has a total length of 1.65 m s^{-1}, has a fixed bias component of 0.46 m s^{-1} (almost 30%). This is larger than the initialization increment of 0.35 m s^{-1} that was obtained for a single member. It is troubling to see that the assumption of no bias is not better respected in our EnKF implementation.

As mentioned in Dee and da Silva (1998, p. 276) and in Thiébaux and Morone (1990), systematic forecast errors may be transient in nature. (This transience may be due to regime dependence (Tibaldi and Molteni 1990).) Consequently, the total systematic-error component will be bigger than the fixed bias-component estimated using Eq. (20). In fact, being aware of an underprediction of the diurnal cycle in our model, we used successive increments one day apart so that the time averaging would not filter this diurnal bias component.

As mentioned in Houtekamer *et al.* (2005), we performed some experiments with a bias-correction algorithm. We found, however, that it was more difficult to obtain and interpret experimental results in this more complex environment and decided to rely on ongoing research on the forecast model for the future reduction of bias.

Lacking any specific procedure to estimate and correct the bias, we need to use a somewhat larger model-error term to maintain realistic ensemble statistics. It is, therefore, not surprising that this term, with an amplitude of 1.39 m s^{-1} for a single member, is larger than the bias term.

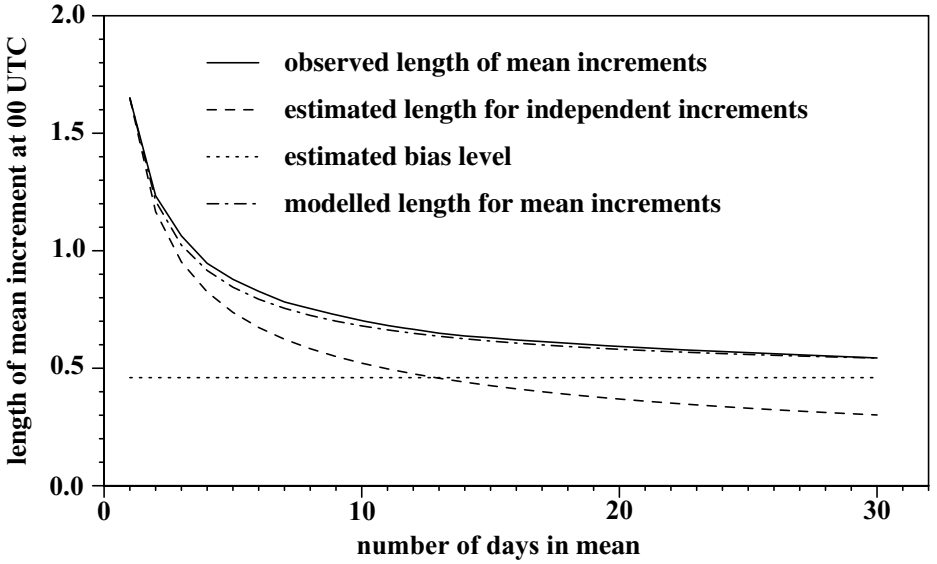

Figure 1. The solid line is the length of the mean analysis increments computed from a sequence of analyses (all valid at 0000 UTC) over a 30-day period. The dashed line is the expected length for independent analysis increments. The dotted line shows the estimated length of the bias-induced constant component. The dashed-dotted line is the modelled length calculated from Eq. (20) with the parameters $\|C\|$ and σ estimated from the solid line (as described in the text). All lengths are in m s^{-1}.

(*d*) *Sampling error due to the finite ensemble size*

The main approximation in the EnKF, with respect to the KF, is the use of an ensemble of finite size to estimate covariance information. To measure the severity of this approximation, one should ideally, for a range of ensemble sizes, perform sufficiently long data assimilation cycles with well-adjusted parameters for the horizontal and vertical localization. Such a project, however, is beyond what we can currently afford to run on our computer.

Instead, we have used a much reduced experimental design. As in Houtekamer *et al.* (2005), we generated sets of perturbed background fields, in this instance valid at 0000 UTC 27 July 2004, using an isotropic covariance matrix $\mathbf{P}^{\mathrm{f}}_{\mathrm{3D}}$. Here we used sets having 2×12, 2×24, 2×48, 2×96 and 2×192 members. For each set of N members, we performed a single analysis leading to an ensemble mean analysis $\overline{x^{\mathrm{a}}_N}$. For each analysis, we used the same, operational, values for the horizontal and vertical localization. Thus, the focus in these experiments is on the sampling error due to the finite ensemble size. The impact of the localization, made necessary by the finite ensemble size, is not considered here.

For each $\overline{x^{\mathrm{a}}_N}$, we would like to know the error, with respect to an unfortunately unobtainable analysis $\overline{x^{\mathrm{a}}_\infty}$. We speculate, however, that the error will decrease to zero with the root of the ensemble size as is commonly the case with Monte Carlo methods. Further assuming that the errors in the estimates $\overline{x^{\mathrm{a}}_N}$ and $\overline{x^{\mathrm{a}}_M}$ are independent, we obtain for the distance between two different ensemble-mean analyses:

$$\|\overline{x^{\mathrm{a}}_N} - \overline{x^{\mathrm{a}}_M}\| \propto \left(\frac{1}{N} + \frac{1}{M}\right)^{0.5}. \tag{21}$$

TABLE 2. ERROR DUE TO LIMITED ENSEMBLE SIZE

	$\|\overline{x_N^a} - \overline{x_M^a}\|$ (m s^{-1})				
$M \setminus N$	2×12	2×24	2×48	2×96	2×192
2×12	–	1.077	1.000	0.959	0.943
2×24	(1.096)	–	0.776	0.717	0.691
2×48	(1.001)	(0.775)	–	0.532	0.500
2×96	(0.949)	(0.707)	(0.548)	–	0.362
2×192	(0.922)	(0.671)	(0.500)	(0.387)	–
$2 \times \infty$	(0.895)	(0.633)	(0.448)	(0.317)	(0.225)

The distance, computed using an energy norm, between ensemble-mean analyses. Here the two ensemble mean analyses have been obtained using N and M members. The values below the dashed diagonal, enclosed in brackets, have been obtained using a fit to the values above the diagonal.

We computed the distance for all available pairs of M and N. The results are presented in the upper-right triangle of Table 2. A least-squares procedure was used to obtain the proportionality constant in Eq. (21). The fitted values are shown in parentheses in the lower-left triangle of Table 2. The desired error estimate for each N is obtained by setting $M = 2 \times \infty$ in Eq. (21) and given in the bottom row of Table 2. As can be seen in Table 2, the estimated error for an ensemble of the operational size of $N = 2 \times 48$ is 0.448 m s^{-1}.

Our experience, with our experimental environment, is that going to an ensemble size of $N = 2 \times 48$ leads to a clear improvement in the quality of the EnKF. Beyond this size, however, it becomes more difficult to observe improvement from using bigger ensembles. Apparently, as the sampling error becomes ever smaller, other sources of error, such as discussed in the other subsections of this section, start to dominate over the sampling error.

(*e*) *Time interpolation*

In our current implementation, we assume that all observations that are taken during a 6 h time window are valid at the central analysis time. To estimate the impact of this, we measure the difference between the 3 h and the 6 h state of the filtered forecast. The value of 2.48 m s^{-1} is comparable to the magnitude of the analysis increment itself. Similarly we obtain a value of 2.40 m s^{-1} for the difference between the state at 6 and 9 h. The approximation that the observations are valid at the central analysis time may have been appropriate when radiosonde and other synoptic observations were predominant, but given the ever-increasing amount of satellite and other asynoptic data, is equivalent to assuming negligible model evolution in the 6 h window. Such an approximation is clearly severe and is, in fact, inconsistent with the earlier condition, from Hollingsworth *et al.* (1986), that the forecast increment must be bigger than the analysis increment.

The recent positive experience with four-dimensional variational assimilation (4D-Var) underlines the importance of performing time interpolation (Rabier *et al.* 1998, 2000). Hunt *et al.* (2004) demonstrate how an expansion of the state vector permits time interpolation in an EnKF.

In the context of a 6 h data assimilation window, we can extend the state vector by including the model states at 3 and 9 h. Using a linear interpolation in time, the most severe errors will now occur at 4.5 and 7.5 h. Using the states of the filtered run, we find an error of 0.44 m s^{-1} at both 4.5 and 7.5 h. While this error is little more than a

sixth of the error when there is no time interpolation, it is still of the same order as the initialization increment.

Further extending the state vector to include the model states at 4.5 and 7.5 h, we now obtain maximum time interpolation errors at 3.75, 5.25, 6.75 and 8.25 h. At these times, we find errors of 0.16, 0.17, 0.17 and 0.17 m s^{-1}, respectively. This puts them well below the initialization increment.

The cost of the data assimilation step in our EnKF implementation is dominated by the computation of $\mathbf{P}\mathcal{H}^{\mathrm{T}}$, i.e. Eq. (12), and by the multiplication by this term (Houtekamer and Mitchell 2001, Eq. (8)). The cost of these operations is $\mathrm{O}(N_{\mathrm{model}}N_{\mathrm{obs}}N)$, where N_{model} is the number of model coordinates and N_{obs} is the number of observations (Houtekamer and Mitchell 2001). The above results suggest the use of five time levels (at 3.0, 4.5, 6.0, 7.5 and 9 h) which would increase N_{model}, and thus also the computational cost of the algorithm, by a factor of 5. This algorithm yields an ensemble of analysed five-time-level trajectories. The ensemble of central, i.e. 6 h, analyses serves as initial conditions for the subsequent ensemble of short-range forecasts required to continue the data assimilation cycle.

We wish to avoid the cost associated with the computation of analysed five-time-level trajectories of which only one time level is used. This can be done by performing the analysis in a joint state-observation space, following Tarantola (1987) and Anderson (2001). Here, the forward interpolation $\mathcal{H}(x^{\mathrm{f}}(t_{\mathrm{obs}}))$ would be precomputed using all available time levels (P. Gauthier 2005, personal communication) and appended to the smallest state vector required to continue the data assimilation cycle, i.e. $x(t = 6\ \mathrm{h})$. This gives a joint state-observation vector, $(x(t = 6\ \mathrm{h}),\ \mathcal{H}(x^{\mathrm{f}}(t_{\mathrm{obs}})))$. This algorithm would have a cost $\mathrm{O}\{(N_{\mathrm{model}} + N_{\mathrm{obs}})N_{\mathrm{obs}}N\}$ which, given that we have $N_{\mathrm{obs}} \ll N_{\mathrm{model}}$, would be almost identical to the cost of our current algorithm that does not include time interpolation. It should be noted that, as batches of observations are assimilated sequentially by the EnKF, each $\mathcal{H}(x^{\mathrm{f}}(t_{\mathrm{obs}}))$ will have to be updated with the rest of the state vector until the batch containing the corresponding observation has been assimilated.

Pre-computing $\mathcal{H}(x^{\mathrm{f}}(t_{\mathrm{obs}}))$ leads to a more modular data assimilation algorithm. Given that the computation of $\mathcal{H}(x^{\mathrm{f}}(t_{\mathrm{obs}}))$ for the different members of the ensemble is independent, and given that N is O(100), this part of the algorithm can be parallelized by assigning the complete state vector for one or more ensemble members to each processor. This allows $\mathcal{H}$ to be applied to complete global fields, unlike our current algorithm where $\mathcal{H}$ is applied to guess-fields that have been partitioned into fairly small tiles (Houtekamer and Mitchell 2001, section 4). Moreover, in the analysis algorithm itself, the tile structure, introduced to permit multi-tasking, is also simplified: successive points of the horizontal grid can now be assigned to successive processors as though each of these points defined a separate 'tile'.

5. Relationship between imbalance and time interpolation

As suggested in the previous section, time interpolation error is a major source of error in our system. An attractive and simple approach would be to have a data assimilation cycle with a length of only one short time step. However, this would require well-balanced initial conditions and, as pointed out by Hunt *et al.* (2004), 'in an operational setting, frequent switching between assimilation and model evolution may be costly and detrimental to the accuracy of the numerical time integration'. Balanced initial conditions and a smooth model trajectory are also required if time interpolation of that trajectory is to be beneficial. The balance requirement in this context is substantially

more severe than in our previous study (Mitchell *et al.* 2002) where we only considered observations at the central analysis time and consequently were only concerned about the degree of imbalance in the 6 h forecast.

We begin this section by looking at the evolution of the imbalance in an unfiltered model integration. Next, we consider two ways of obtaining a balanced integration: (i) applying a balancing operator, and (ii) relaxing the localization operator that is the cause of the imbalance.

(*a*) *Evolution of the imbalance*

Rapid oscillations in the surface pressure are often taken to be indicative of imbalance (e.g. Williamson and Temperton 1981). To quantify this type of fast gravity-wave imbalance in a given analysis, we estimate the second difference $D^2(p_s)$ of the surface pressure p_s, from the time series obtained from a 24 h integration starting from that analysis:

$$D^2(p_s(t)) = p_s(t + \Delta t) - 2p_s(t) + p_s(t - \Delta t). \tag{22}$$

Here we take $\Delta t = 45$ min, i.e. one model time step. The global r.m.s. value $\|D^2(p_s)\|$ is subsequently obtained using

$$\|D^2(p_s)\|^2 = (D^2(p_s),\, D^2(p_s)) = \frac{1}{S}\int_S D^2(p_s) D^2(p_s)\, \mathrm{d}S, \tag{23}$$

where S is the surface area of the globe.

The above quantity is computed for the 24 h forecast of member 1 of the ensemble of analyses valid at 0000 UTC 1 August 2004. The analyses were, again, obtained with the EnKF configuration that was implemented in January 2005. In particular, the impact of observations drops to zero at a horizontal distance of 2800 km and at two units of $\ln p$.

It can be seen from Fig. 2 that the oscillations in surface pressure are strongly damped by the model. The initial estimate of $\|D^2(p_s(45\text{ min}))\| = 2.61$ hPa is more than an order of magnitude bigger than the final estimate of $\|D^2(p_s(23\text{ h }15\text{ min}))\| = 0.16$ hPa. This significant initial imbalance would probably cause problems if the EnKF were to run with a short data assimilation cycle.

In our current configuration, we have a 6 h window for the observations with the window extending from 3 to 9 h after the preceding analysis. We note that the estimated 1.05 hPa for the second difference at 3 h is twice as large as the value of 0.50 hPa obtained at 6 h. Consequently, considering that the first observations are valid at about 3 h, an explicit initialization could be beneficial for an EnKF (Lorenc 2003) that employs time interpolation.

To investigate this further, an additional filtered integration with a length of 24 h has been performed from the same initial condition. That is, from the filtered state at $t = 3$ h, we subsequently run the model out to 24 h. Evaluating Eqs. (22) and (23), we can obtain a time series consisting of four unfiltered points (valid at $t = 0$, 0.75, 1.5, and 2.25 h), and 29 subsequent points that are expected to be on a balanced trajectory. Due to the use of different time levels in Eq. (22), the first estimate of $D^2(p_s)$ that might be expected to indicate that balance has been achieved should occur at step 5 ($t = 3.75$ h). It can be seen from Fig. 2 that this is indeed the case, and in fact, that the digital filter is highly effective in removing the oscillations in the surface pressure field.

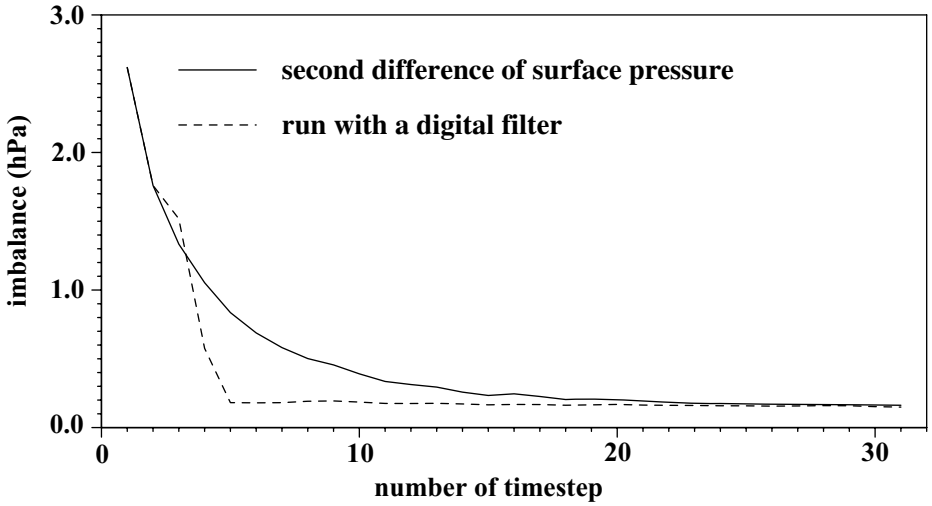

Figure 2. The evolution of the norm of the second difference of surface pressure, Eqs. (22) and (23), over the course of a one-day integration with a 45 min time step. The solid line is for an initial condition that has been integrated in time without the imposition of any balance procedure. The dashed line corresponds to an integration from the same initial condition, but with the application of a digital filter finalization technique, as described in the text.

TABLE 3. INITIAL IMBALANCE

	$\|D^2(p_s(45\text{ min}))\|$ (hPa)				
r_h (km)	$r_z = 2$	$r_z = 4$	$r_z = 6$	$r_z = 8$	$r_z = 100$
2 800	2.618	2.913	3.100	3.236	3.622
5 600	2.600	2.725	2.783	2.815	2.925
8 400	2.744	2.520	2.381	2.345	2.423
11 200	2.883	2.420	2.316	2.133	2.145
280 000	1.996	1.558	1.223	1.153	0.454

The second difference of surface pressure estimated at $t = 45$ min for different values of the localization parameters r_h and r_z. Horizontal correlations are forced to zero at a distance of r_h km and vertical correlations at r_z units of ln p.

(*b*) *Obtaining balance by relaxing the localization*

As mentioned in section 2(c), the localization operator in the EnKF will cause imbalance. One would expect that using a bigger ensemble and a less severe localization would lead to improved balance and a more accurate analysis. Here, to investigate only the impact of a less severe localization on the amount of imbalance, we generated the analyses valid at 0000 UTC 1 August 2004, using less severe horizontal and vertical localizations as specified in Table 3. The reference value of $\|D^2(p_s(45\text{ min}))\| = 2.618$ hPa, corresponds to the first value for the unfiltered run in Fig. 2. It can be seen that to slightly improve on this value, we need both to more than double the value of the horizontal localization parameter r_h and to simultaneously relax the vertical localization. Even if we remove the vertical localization and relax r_h to 11 200 km, the imbalance in the surface pressure only decreases by about 20%. It should be noted here that a significantly larger ensemble would be necessary to obtain a high-quality analysis with these values of r_h and r_z. Eventually, with values of r_h and r_z that correspond to virtually no localization, we obtain analyses that are almost as balanced as the guess fields, the

value of 0.454 being not significantly different from the value of 0.502 for time step 8 of the solid curve in Fig. 2.

To get a rough estimate of how many members we would need in the absence of covariance localization, we here provide a simple estimate based on the hypothesis that the current ensemble size of 96 members provides just enough degrees of freedom to fit the data in the area covered by the localizing function. With a zero impact of observations at 2800 km in the horizontal and 2 units of ln p in the vertical, we assume that an observation is effectively seen in a disc with horizontal diameter 2800 km and a vertical extent of 2 units in ln p. With a model top at 10 hPa, we must cover a total distance of 4.6 units in ln p. In the horizontal, we obtain 83 semi-independent areas (by dividing the total surface area of the sphere by the area covered by a single disc). From these three numbers, we conclude that the localizations lead to an effective dimensionality of over 10 000. We therefore estimate that we would need over 10 000 ensemble members to obtain an accurate and balanced analysis with an EnKF that does not include covariance localization. Such a configuration is beyond what we can currently handle on the computer.

With our current implementation of covariance localization and the parameters we can afford to use with it, we are consequently unable to obtain reasonably balanced analyses. To alleviate this problem, and to permit time interpolation in a subsequent implementation, the digital filter finalization was implemented in the operational EnKF configuration in December 2005.

It can be concluded from the above results that, as more and more batches of observations are assimilated, the sequential EnKF will generate increasingly unbalanced background fields. This might lead to an increasingly suboptimal assimilation of the observations. New inherently more balanced ways of localizing covariances (J. Kepert 2005, poster presentation at the Prague Symposium) are currently being developed and tested. These may have a beneficial impact on EnKF results.

6. Re-interpretation of model error

In Houtekamer *et al.* (2005), it was noted that our model-error term represents a multitude of terms including (i) errors in the forward interpolation operator, (ii) errors in the specification of the statistics of the observations, and (iii) errors due to the parametrization of unresolved dynamical and physical processes. The first two of these are not related to the forecast model at all; in fact, it is the parametrization errors that are generally considered to be dominant. It is for that reason that, in our EnKF, we have followed standard Kalman filter theory, adding the model error to the prediction error to obtain the forecast error, where our terminology for these various errors is as defined by Daley (1992b, section 2). It could be argued, however, that it would be more realistic to add model error at every time step as it occurs. This would allow the model error to subsequently evolve with the model dynamics. It would also be more consistent with the system simulation approach, using multiple model versions, employed in our medium-range EPS. In practice, neither the covariance that could account for the accumulation of model error over the course of a short-range forecast nor the covariance associated with the model error incurred over a single time step is well known.

As discussed in Houtekamer *et al.* (2005, section 4(a)), our parametrized model error has been adjusted so that the innovation amplitude predicted by the EnKF agrees fairly well with the observed innovation amplitude. We have already noted that our parametrized model error has a magnitude of 1.39 m s^{-1}. To put this number in perspective, we ran our dynamical model with the physical parametrizations switched off.

This led to a difference of 1.52 m s^{-1} in the state vector at 6 h. Given that this is the total contribution of the physical parametrizations, these parametrizations would have to be severely in error to produce an error of 1.39 out of a total of 1.52. Having confidence in the model's physical parametrizations, we take this as support for the view that a sizable portion of the model error originates in the analysis step.

Only under rather restrictive conditions (Cohn and Parrish 1991) will the total model-error covariance, $\mathbf{Q}(t, t-1)$, be proportional to the model-error covariance that occurs at each time step. Even in this case, however, it does not follow that individual model-error fields would be non-propagating; that can only happen if the forecast model is simply the identity operation. In that special case, it would be possible to obtain the same result either by adding model-error fields at different time levels or by adding a single model-error field at the end of the forecast. In a realistic situation, however, a model-error field, once added to the state vector, will evolve in some non-trivial way with the dynamics and physics of the numerical model. This implies that, in an EnKF that uses time interpolation, model error cannot be added only to the final time-level as in Eq. (10).

An alternative which has the practical advantage that it permits time interpolation is to displace the regular addition of random fields, i.e. replace Eqs. (10) and (11) with

$$x_i^{\mathrm{f}} = \mathcal{M}\{x_i^{\mathrm{a}}(t-1) + d_i\}, \quad i = 1, \ldots, N, \tag{24}$$

$$d_i \sim N(0, \mathbf{D}), \tag{25}$$

where d_i represents the data assimilation error. The corresponding covariance matrix $\mathbf{D}$ plays a role similar to that of the original matrix $\mathbf{Q}$. The displaced regular addition of random fields d_i will regularly inflate the ensemble covariances so that they will remain sufficiently large. In their EnKF, Hunt *et al.* (2004) modify the analysis-error covariance, rather than the forecast-error covariance, to account for model-error (see their section 4). In the context of our EnKF, this could correspond to the belief that the lack of realism of, say, the observation-error statistics for which we assumed a $N(0, \mathbf{R})$ distribution, leads to a degradation of the quality of the analysis that needs somehow to be corrected for in the Kalman filter equations. Alternatively, one could take the viewpoint that since the model error represents a multitude of terms, not all of which relate to the forecast model, it should more appropriately be referred to as system error and the exact stage of the EnKF procedure where it should be added is therefore somewhat arbitrary. Moreover, adding it after the analysis, instead of after the forecast, allows the 'system error' to evolve, and even amplify, due to the model dynamics. Since we obtained better results with Eqs. (24) and (25), these equations replaced Eqs. (10) and (11) in the operational EnKF configuration in December 2005.

7. Conclusions

In this paper, we have reviewed concepts related to the Kalman filter and its approximation with an ensemble. The KF, which is derived in the context of linear dynamics, provides the optimal minimum variance solution to the data assimilation problem. If the model error projects only on to slow modes then, since imbalance is damped by the forecast model (see, e.g. Fig. 2), the KF solution will be increasingly balanced as the data assimilation cycle continues. In the case of small errors, the basic EnKF, given by Eqs. (10)–(17), converges to the KF solution as the ensemble size increases. For large errors, the local curvature of the balanced space, which is not considered by the EnKF, may lead to unbalanced errors. For a finite ensemble size,

the EnKF solution is not optimal. The use of a small ensemble (O(100) members) necessitates the introduction of measures, not required in the KF itself, to increase the dimensionality of the ensemble. An EnKF in which the effect of observations is localized is unbalanced even in the limit of small errors.

The focus in this paper is on large-scale global atmospheric data assimilation. The EnKF, which was implemented operationally at the Meteorological Service of Canada in January 2005, is briefly described and used to examine some aspects of EnKF performance. We measured the magnitude of the forecast component F, the analysis component A, and, if initialization is performed, of the initialization component I. The relative sizes of these components were found to be ordered in accordance with the Hollingsworth *et al.* (1986) criteria for a properly functioning data assimilation system, i.e. $F > A > I$.

A 30-day sequence of ensemble-mean analysis increments has been examined and the results (Fig. 1) indicate the presence of a fixed bias-induced component slightly larger than the initialization component I. This is troubling as the EnKF has been designed assuming that the forecast model and the observations have no bias (Eqs. (1) and (2)).

A set of experiments was performed to evaluate the sampling error due to the finite ensemble size. The results (Table 2) are consistent with the expected $N^{-0.5}$ decrease in error, where N is the ensemble size. With our current ensemble size (2×48 members), the error is apparently of the same order as that of other error terms.

As seen in Fig. 2, the forecast model strongly damps the oscillations associated with imbalance. Consequently, if it is assumed that the observations are valid at 6 h (the central time of the analysis window), there is perhaps no need for an explicit balance operation. On the other hand, given the dominant size of the forecast component F, and the ever growing quantity of asynoptic observations, we found that time interpolation of the model trajectory to the observations should not be neglected.

Figure 2 indicates that time interpolation of the model trajectory to the observations can only be implemented with a more balanced short-range forecast. The results of Table 3 indicate that moderate relaxation of the localization unexpectedly yields little or no reduction in imbalance. Implementation of a digital filter finalization would provide sufficient balance. In addition, to permit time interpolation, we displaced the addition of model error so that, instead of adding the model error to the ensemble of predictions, we now add it to the ensemble of analyses, as discussed in section 6.

In view of its large amplitude, the parametrized model-error term could be split into a parametrized data assimilation error and a simulated model-physics error. One could, for instance, use stochastic physics (Buizza *et al.* 1999; Shutts 2005) or an ensemble of differently configured forecast models (Houtekamer *et al.* 1996) to simulate the occurrence of model error with an ensemble of integrations. To the extent that these approaches account for some of the flow-dependent intermittent sources of model error, the homogeneous isotropic term, which accounts for unexplained error, could be reduced, as suggested in Mitchell and Houtekamer (2000).

The proposed EnKF algorithm, with time interpolation of the model trajectories to the observations, with the introduction of parametrized errors at the initial time and with a digital filter finalization, resembles current 4D-Var implementations. For instance, the CMC implemented 4D-Var operationally on 15 March 2005 for the high-resolution deterministic forecast system (Laroche *et al.* 2005). In that algorithm, the adjoint technique is used to accomplish time interpolation, background-error covariances are specified at the initial time, and a digital filter finalization is applied to the reference trajectory. Neither algorithm can be expected to perform well in the presence of

significant bias in the model or the observations, of large amplitude model error of unknown origin or of poorly known observational error statistics. We believe that a reduction, or better handling, of these problems will be highly beneficial to the performance of both 4D data assimilation algorithms.

ACKNOWLEDGEMENTS

The first author would like to thank the participants of the Fourth WMO International Symposium on Assimilation of Observations in Meteorology and Oceanography held in Prague for many fruitful discussions. The authors thank Luc Fillion for his internal review of the manuscript and Jeffrey Whitaker and an anonymous reviewer for their official reviews.

REFERENCES

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

Anderson, J. L. and Anderson, S. L. 1999 A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Weather Rev.*, **127,** 2741–2758

Anselmo, D. and Deblonde, G. 2005 'The assimilation of AMSU and SSM/I brightness temperatures in clear skies at the Meteorological Service of Canada'. Pp. 281–290 in Proceedings of 14th international ATOVS study conference, 25–31 May 2005, Beijing, China. http://cimss.ssec.wisc.edu/itwg/itsc/itsc14/proceedings/8_7_Anselmo.pdf

Buizza, R., Miller, M. and Palmer, T. N. 1999 Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Q. J. R. Meteorol. Soc.*, **125,** 2887–2908

Burgers, G., van Leeuwen, P. J. and Evensen, G. 1998 Analysis scheme in the ensemble Kalman filter. *Mon. Weather Rev.*, **126,** 1719–1724

Cohn, S. E. 1982 'Methods of sequential estimation for determining initial data in numerical weather prediction'. PhD thesis, Courant Institute of Mathematical Sciences, New York University

1997 An introduction to estimation theory. *J. Meteorol. Soc. Jpn.*, **75,** 257–288

Cohn, S. E. and Parrish, D. F. 1991 The behavior of forecast error covariances for a Kalman filter in two dimensions. *Mon. Weather Rev.*, **119,** 1757–1785

Cohn, S. E., da Silva, A., Guo, J., Sienkiewicz, M. and Lamich, D. 1998 Assessing the effects of data selection with the DAO physical-space statistical analysis system. *Mon. Weather Rev.*, **126,** 2913–2926

Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M. and Staniforth, A. 1998a The operational CMC-MRB Global Environmental Multiscale (GEM) model. Part I: Design considerations and formulation. *Mon. Weather Rev.*, **126,** 1373–1395

Côté, J., Desmarais, J.-G., Gravel, S., Méthot, A., Patoine, A., Roch, M. and Staniforth, A. 1998b The operational CMC-MRB Global Environmental Multiscale (GEM) model. Part II: Results. *Mon. Weather Rev.*, **126,** 1397–1418

Daley, R. 1991 *Atmospheric data analysis.* Cambridge University Press

1992a The lagged innovation covariance: A performance diagnostic for atmospheric data assimilation. *Mon. Weather Rev.*, **120,** 178–196

1992b Estimating model-error covariances for application to atmospheric data assimilation. *Mon. Weather Rev.*, **120,** 1735–1746

Dee, D. P. 1995 On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Weather Rev.*, **123,** 1128–1145

Dee, D. P. and da Silva, A. M. 1998 Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.*, **124,** 269–295

Derber, J. C., Parrish, D. F. and Lord, S. J. 1991 The new global operational analysis system at the National Meteorological Center. *Weather and Forecasting*, **6,** 538–547

Dowell, D. C., Zhang, F., Wicker, L. J., Snyder, C. and Crook, N. A. 2004 Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. *Mon. Weather Rev.*, **132,** 1982–2005

Evensen, G. 1994 Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**(C5)**,** 10143–10162

Fillion, L., Mitchell, H. L., Ritchie, H. and Staniforth, A. 1995 The impact of a digital filter finalization technique in a global data assimilation system. *Tellus*, **47A,** 304–323

Gandin, L. S. 1965 Objective analysis of meteorological fields (English translation). Israel program for scientific translations, Gidromet

Gaspari, G. and Cohn, S. E. 1999 Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.*, **125,** 723–758

Gauthier, P., Courtier, P. and Moll, P. 1993 Assimilation of simulated wind lidar data with a Kalman filter. *Mon. Weather Rev.*, **121,** 1803–1820

Gauthier, P., Buehner, M. and Fillion, L. 1999 'Background-error statistics modelling in a 3D variational data assimilation scheme: Estimation and impact on the analyses'. Pp. 131–145 in Proceedings of ECMWF workshop on diagnosis of data assimilation systems, 6–10 September 1999, Reading, UK. ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, UK

Gelb, A. 1974 *Applied optimal estimation.* M.I.T. Press, Cambridge, Ma, USA

Ghil, M. and Malanotte-Rizzoli, P. 1991 Data assimilation in meteorology and oceanography. *Adv. Geophys.*, **33,** 141–266

Ghil, M., Cohn, S., Tavantzis, J., Bube, K. and Isaacson, E. 1981 Applications of estimation theory to numerical weather prediction. Pp. 139–224 in *Dynamic meteorology—data assimilation methods.* Eds. L. Bengtsson, M. Ghil, and E. Källén. Springer-Verlag, Heidelberg

Hamill, T. M. and Snyder, C. 2000 A hybrid ensemble Kalman filter-3D variational analysis scheme. *Mon. Weather Rev.*, **128,** 2905–2919

Hamill, T. M., Whitaker, J. S. and Snyder, C. 2001 Distance dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Weather Rev.*, **129,** 2776–2790

Hollingsworth, A., Shaw, D. B., Lönnberg, P., Illari, L., Arpe, K. and Simmons, A. J. 1986 Monitoring of observation and analysis quality by a data assimilation system. *Mon. Weather Rev.*, **114,** 861–879

Houtekamer, P. L. and Mitchell, H. L. 1998 Data assimilation using an ensemble Kalman filter technique. *Mon. Weather Rev.*, **126,** 796–811

1999 Reply to 'Comment on "Data assimilation using an ensemble Kalman filter technique"'. *Mon. Weather Rev.*, **127,** 1378–1379

2001 A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Weather Rev.*, **129,** 123–137

Houtekamer, P. L., Lefaivre, L., Derome, J., Ritchie, H. and Mitchell, H. L. 1996 A system simulation approach to ensemble prediction. *Mon. Weather Rev.*, **124,** 1225–1242

Houtekamer, P. L., Mitchell, H. L., Pellerin, G., Buehner, M., Charron, M., Spacek, L. and Hansen, B. 2005 Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Weather Rev.*, **133,** 604–620

Hunt, B. R., Kalnay, E., Kostelich, E. J., Ott, E., Patil, D. J., Sauer, T., Szunyogh, I., Yorke, J. A. and Zimin, A. V. 2004 Four-dimensional ensemble Kalman filtering. *Tellus*, **56A,** 273–277

Kailath, T. 1968 An innovations approach to least squares estimation—Part I: linear filtering in additive white noise. *IEEE Trans. Automat. Control*, **13,** 646–655

Kalman, R. E. 1960 A new approach to linear filtering and prediction problems. *Trans. ASME, Ser. D, J. Basic Eng.*, **82,** 35–45

Kalman, R. E. and Bucy, R. S. 1961 New results in linear filtering and prediction theory. *Trans. ASME, Ser. D, J. Basic Eng.*, **83,** 95–108

Keppenne, C. L. and Rienecker, M. M. 2002 Initial testing of a massively parallel ensemble Kalman filter with the Poseidon isopycnal ocean general circulation model. *Mon. Weather Rev.*, **130,** 2951–2965

Laroche, S., Gauthier, P., Tanguay, M., Pellerin, S. and Morneau, J. 2005 'Evaluation of the operational 4D-Var at the Meteorological Service of Canada'. Preprint, 17th Am. Meteorol. Soc. conference on numerical weather prediction, Washington, DC, USA. http://ams.confex.com/ams/WAFNWP34BC/techprogram/paper_93941.htm

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. Weather Rev.*, **132,** 1966–1981

Lewis, J. M. and Derber, J. C. 1985 The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, **37A,** 309–322

Lorenc, A. C. 2003 The potential of the ensemble Kalman filter for NWP—A comparison with 4D-Var. *Q. J. R. Meteorol. Soc.*, **129,** 3183–3203

Miller, R. N., Carter, Jr., E. F. and Blue, S. T. 1999 Data assimilation into nonlinear stochastic models. *Tellus*, **51A,** 167–194

Mitchell, H. L. and Houtekamer, P. L. 2000 An adaptive ensemble Kalman filter. *Mon. Weather Rev.*, **128,** 416–433

Mitchell, H. L., Houtekamer, P. L. and Pellerin, G. 2002 Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Mon. Weather Rev.*, **130,** 2791–2808

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

Parrish, D. F. and Derber, J. C. 1992 The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Weather Rev.*, **120,** 1747–1763

Pellerin, G., Lefaivre, L., Houtekamer, P. and Girard, C. 2003 Increasing the horizontal resolution of ensemble forecasts at CMC. *Nonlinear Processes Geophys.*, **10,** 463–468

Petersen, D. P. 1973 Transient suppression in optimal sequential analysis. *J. Appl. Meteorol.*, **12,** 437–440

Pham, D. T. 2001 Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Mon. Weather Rev.*, **129,** 1194–1207

Phillips, N. A. 1986 The spatial statistics of random geostrophic modes and first-guess errors. *Tellus*, **38A,** 314–332

Rabier, F., Thépaut, J.- N. and Courtier, P. 1998 Extended assimilation and forecast experiments with a four-dimensional variational assimilation system. *Q. J. R. Meteorol. Soc.*, **124,** 1861–1887

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.- F. and Simmons, A. 2000 The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.*, **126,** 1143–1170

Reichle, R. H., McLaughlin, D. B. and Entekhabi, D. 2002 Hydrologic data assimilation with the ensemble Kalman filter. *Mon. Weather Rev.*, **130,** 103–114

Saunders, R. W., Brunel, P., Chevallier, F., Deblonde, G., English, S. J., Matricardi, M. and Rayer, P. J. 2002 RTTOV-7—Science and validation report. Met Office Forecasting Research Technical Report No. 387, Exeter, UK

Shutts, G. 2005 A kinetic energy backscatter algorithm for use in ensemble prediction systems *Q. J. R. Meteorol. Soc.*, **131,** 3079–3102

Snyder, C. and Zhang, F. 2003 Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Weather Rev.*, **131,** 1663–1677

Szunyogh, I., Kostelich, E. J., Gyarmati, G., Patil, D. J., Hunt, B. R., Kalnay, E., Ott, E. and Yorke, J. A. 2005 Assessing a local ensemble Kalman filter: perfect model experiments with the National Centers for Environmental Prediction global model. *Tellus*, **57A,** 528–545

Tarantola, A. 1987 *Inverse problem theory.* Elsevier, Amsterdam

Thiébaux, H. J. and Morone, L. L. 1990 Short-term systematic errors in global forecasts: their estimation and removal. *Tellus*, **42A,** 209–229

Tibaldi, S. and Molteni, F. 1990 On the operational predictability of blocking. *Tellus*, **42A,** 343–365

Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M. and Whitaker, J. S. 2003 Ensemble square root filters. *Mon. Weather Rev.*, **131,** 1485–1490

van Leeuwen, P. J. 1999 Comment on 'Data assimilation using an ensemble Kalman filter technique'. *Mon. Weather Rev.*, **127,** 1374–1377

Whitaker, J. S., Compo, G. P., Wei, X. and Hamill, T. M. 2004 Reanalysis without radiosondes using ensemble data assimilation. *Mon. Weather Rev.*, **132,** 1190–1200

Williamson, D. L. and Temperton, C. 1981 Normal mode initialization for a multilevel grid-point model. Part II: Nonlinear aspects. *Mon. Weather Rev.*, **109,** 744–757