Ensemble of Data Assimilations and Uncertainty Estimation

Massimo Bonavita

ECMWF, Reading, UK
Massimo.Bonavita@ecmwf.int

ABSTRACT

The background error covariance matrix (**B**) plays a fundamental role in modern data assimilation systems. Advanced data assimilation methods use an evolving, flow-dependent model for **B**, which can account for the non-stationary and non-homogeneous character of the background errors. An evolving **B** is used in the Kalman Filter which, for linear systems, is known to be the statistically optimal solution. The Kalman Filter is, however, computationally intractable for realistic Numerical Weather Prediction applications. This practical difficulty has led to two different classes of solutions: the variational approach; and low-rank, Monte-Carlo approximations to the Kalman Filter (Ensemble Kalman Filter). Both methods have advantages and disadvantages. This has motivated the research effort to blend the two approaches towards some 'hybrid' configuration which tries to avoid the drawbacks of either technique. This paper presents a review of hybrid data assimilation in contemporary global operational NWP with an emphasis on the solutions being pursued at ECMWF.

1. Introduction

Atmospheric data assimilation (DA) works by blending a short range (t+6-12h) forecast (the background) with more recent observations collected over a time window. To a large extent, the skill of an assimilation system is determined by how accurate the characterization of the errors in the background is. In fact, the weight given in the analysis to new observations depends on the ratio of the computed/assumed variances of the background and the observation errors. From this perspective, the background error variances control the influence which past observations have on the current analysis update. Further to this, and just as importantly, the spatial and temporal extrapolation of the observational information onto the analysis grid relies on the spatial and temporal structures of the background error covariances.

It is commonly observed that forecast errors have large spatial and temporal variability. This is also true for background errors, whose spatial distribution and magnitude is closely connected with the spatial distribution of available observations and the instabilities of the atmospheric flow: Larger background errors will be prevalent in sparsely observed areas and/or where active whether systems are prevalent. Statistically optimal background error estimates need then to reproduce these variabilities and to be flow-dependent and evolving. For linear systems with Gaussian error distributions, the Kalman Filter (KF) provides the optimal solution to the assimilation problem by computing and evolving estimates of both the state and its error covariances. However, the standard KF is only a computationally viable solution for very low-dimensional systems, and certainly not for realistic Numerical Weather Prediction (NWP) applications. Two approaches have been followed in

atmospheric data assimilation to overcome the 'curse of dimensionality' of the KF: Variational techniques and, more recently, the Ensemble Kalman Filter. Although they make quite radical and different approximations to the standard KF, both techniques have been found to perform adequately and, as recent results suggest, with comparable skill in real NWP data assimilation systems (Miyoshi *et al.*, 2010; Buehner *et al.*, 2010a, b). It is then natural to consider whether a hybrid system that combines uncertainty information from an ensemble assimilation system with the high resolution analysis of a variational scheme would provide better results than either method in isolation. Early results at a few operational NWP Centers seem to confirm that this is indeed the case. Hybrid data assimilation systems can be designed in a number of different ways: the most common approaches in current operational NWP will be described below, together with some indication of recent directions of development in this fast evolving field.

2. The Kalman Filter framework

The data assimilation problem can be coherently described from a Bayesian perspective (Lorenc, 1986; Wikle and Berliner, 2007, for further details). Denoting X as the state and Y as our observational data, the full probability model that describes our knowledge of the system can always be factored using the definition of conditional probability: $p(x,y) = p(x \mid y)p(y) = p(y \mid x)p(x)$ or (*Bayes' Rule*):

$$p(x \mid y) = \frac{p(y \mid x)p(x)}{p(y)} \tag{1}$$

Eq. (1) is interpreted in this way: In order to infer the distribution of the state given the data (*posterior distribution*, *p(x|y)*), we need only form the product of the distributions of measurement errors conditioned on the state (*data distribution, p(y|x)*) and our prior knowledge about the state (*prior distribution, p(x)*). The marginal distribution $p(y) = \int p(y \mid x)p(x)dx$ can be thought of as a normalising constant.

In the multivariate case, assuming the prior to be a normally distributed background forecast, $p(\mathbf{x}) \sim \mathcal{N}(\mathbf{x}_b, \mathbf{P}^b)$ and the data distribution to be normal with known covariance $\mathbf{R}$, $p(\mathbf{y}|\mathbf{x}) \sim \mathcal{N}(\mathbf{H}\mathbf{x}, \mathbf{R})$, the posterior distribution will also be Gaussian:

$$p\left(x \mid y\right) \sim \mathcal{N}\left(x_b + \mathbf{K}\left(y - \mathbf{H}x_b\right), \left(\mathbf{I} - \mathbf{KH}\right)\mathbf{P}^{\mathbf{b}}\right) \tag{2}$$

Here $\mathbf{K} = \mathbf{P^b H^T}(\mathbf{R}+\mathbf{H}\mathbf{P^b}\mathbf{H^T})^{-1}$ is the *Kalman Gain* matrix. This is the basis of Optimal Interpolation algorithms (Gandin, 1963). The problem of finding the mean (and mode) of the distribution in Eq. (2) can be recast in an equivalent variational form, (*3D-Var*), by finding the minimum of the following cost function:

$$J(\mathbf{x}) = \left(\mathbf{x} - \mathbf{x_b}\right)^{\mathbf{T}} \left(\mathbf{P^b}\right)^{-1} \left(\mathbf{x} - \mathbf{x_b}\right) + \left(\mathbf{y} - \mathbf{H}\mathbf{x_b}\right)^{\mathbf{T}} \mathbf{R}^{-1} \left(\mathbf{y} - \mathbf{H}\mathbf{x_b}\right) \tag{3}$$

The variational form of Eq. (3) offers many practical advantages for large scale NWP analysis problems: it avoids observation selection; allows more realistic models of $\mathbf{P}^b$; the use of additional constraints in $J(\mathbf{x})$; the direct use of radiances; more sophisticated algorithms for the quality control of observations.

Let us now introduce the time dimension in our problem: We wish to estimate a set of states $X_{0:t}=[X_0,X_1,..,X_t]$ over a time interval given observations over the same time interval $Y_{0:t}=[Y_1,Y_2,..,Y_t]$:

$$p(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t}) \propto p(\mathbf{y}_{1:t} \mid \mathbf{x}_{0:t})\, p(\mathbf{x}_{0:t}) \tag{4}$$

Two assumptions are commonly made to simplify the structure of the conditional probability in Eq. (4):

a) The Markov property is assumed to hold for the prior distribution, i.e.:

$$p(\mathbf{x}_{0:t}) = p(\mathbf{x}_0)\, p(\mathbf{x}_1 \mid \mathbf{x}_0) \ldots p(\mathbf{x}_t \mid \mathbf{x}_{t-1})$$

b) Observations are statistically independent:

$$p(\mathbf{y}_{1:t} \mid \mathbf{x}_{0:t}) = p(\mathbf{y}_1 \mid \mathbf{x}_1) \ldots \mathrm{p}(\mathbf{y}_t \mid \mathbf{x}_t)$$

These two assumptions allow one to write (4) as:

$$p(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t}) \propto p(\mathbf{x}_0)\, p(\mathbf{x}_1 \mid \mathbf{x}_0)\, p(\mathbf{y}_1 \mid \mathbf{x}_1) \ldots p(\mathbf{x}_t \mid \mathbf{x}_{t-1})\, p(\mathbf{y}_t \mid \mathbf{x}_{t-1}) \tag{5}$$

This form of the posterior distribution is clearly amenable to a *sequential* solution: starting from same prior knowledge $p(\mathbf{x}_0)$, each new batch of observations successively modifies the posterior distribution.

Let us restrict our attention to the *Filtering problem*, i.e., we assume $p(x_{t-1}|y_{1:t-1})$ to be known and we want to find out how a new batch of observations at time t modifies our estimate of the state of the system $p(x_t|y_{1:t})$. Eq. (5) suggests a two-step procedure:

1. Compute the background distribution at time t, (*forecast step*):

$$p(\mathbf{x}_t \mid \mathbf{y}_{1:t-1}) = \int p(\mathbf{x}_t \mid \mathbf{x}_{t-1})\, p(\mathbf{x}_{t-1} \mid \mathbf{y}_{1:t-1})\, \mathrm{d}\mathbf{x}_{t-1} \tag{6}$$

2. Compute the analysis distribution (*analysis step*):

$$\begin{aligned} p(\mathbf{x}_t \mid \mathbf{y}_{1:t}) &= p(\mathbf{x}_t \mid \mathbf{y}_t, \mathbf{y}_{1:t-1}) \propto p(\mathbf{y}_t \mid \mathbf{x}_t, \mathbf{y}_{1:t-1})\, p(\mathbf{x}_t \mid \mathbf{y}_{1:t-1}) \\ &= p(\mathbf{y}_t \mid \mathbf{x}_t)\, p(\mathbf{x}_t \mid \mathbf{y}_{1:t-1}) \end{aligned} \tag{7}$$

Here the Bayes' rule and the Markov assumption have been used.

If the forecast model and observation operators are *linear* (*affine* model operator $\mathcal{M}$ and observation operator $\mathcal{H}$ would be a sufficient requirement) and have *Gaussian* error distributions:

$$\mathbf{x}_t = \mathbf{M}_{t-1,t}\mathbf{x}_{t-1} + \boldsymbol{\eta}_{t-1,t} \quad \boldsymbol{\eta} \sim N\left(\mathbf{0}, \mathbf{Q}_{t-1,t}\right) \tag{8}$$

$$\mathbf{y}_t = \mathbf{H}_t\mathbf{x}_t + \boldsymbol{\varepsilon}_t \qquad \boldsymbol{\varepsilon}_t \sim N\left(\mathbf{0}, \mathbf{R}_t\right) \tag{9}$$

Then the background distribution will also be Gaussian with mean and covariance given by:

$$\mathbf{x}_{t|t-1} = \mathbf{M}_{t-1,t}\mathbf{x}_{t-1} \tag{10}$$

$$\mathbf{P}_{t|t-1} = \mathbf{M}_{t-1|t}\mathbf{P}_{t|t-1}\mathbf{M}^{\mathrm{T}}{}_{t-1|t} + \mathbf{Q}_{t-1|t} \tag{11}$$

The analysis distribution will be Gaussian with mean and covariance given by:

$$\mathbf{x}_{\mathrm{t|t}} = \mathbf{x}_{\mathrm{t|t-1}} + \mathbf{K}_{\mathrm{t}}\left(\mathbf{y}_{\mathrm{t}} - \mathbf{H}_{\mathrm{t}}\mathbf{x}_{\mathrm{t|t-1}}\right) \tag{12}$$

$$\mathbf{P}_{\mathrm{t|t}} = \left(\mathbf{I} - \mathbf{K}_{\mathrm{t}}\mathbf{H}_{\mathrm{t}}\right)\mathbf{P}_{\mathrm{t|t-1}} \tag{13}$$

$$\mathbf{K}_{\mathrm{t}} = \left(\mathbf{P}_{\mathrm{t|t-1}}\right)^{-1}\mathbf{H}_t^T\left(\mathbf{R}_{\mathrm{t}} + \mathbf{H}_{\mathrm{t}}\mathbf{P}_{\mathrm{t|t-1}}\mathbf{H}_t^T\right)^{-1} \tag{14}$$

We have thus recovered the classical Kalman Filter equations (Kalman 1960). This set of equations provide the optimal (minimum variance *and* maximum likelihood) solution to the filtering problem for linear systems with Gaussian error distributions.

We may also be interested in the distribution $p(\mathbf{x}_t|\mathbf{y}_{1:T})$ for t=1,…,T, i.e. we want to estimate the state using observations both before and after time t (*smoothing distribution*). Under the same hypothesis used for the KF, a Kalman smoother (KS) can be derived (Cosme et al., 2012). Two aspects need only be emphasised here: 1) The Kalman smoother differs from the Kalman filter only by the use of cross-covariances in time to correct the state at time t using observations at future times; 2) At the end of the assimilation window (t=T) the KS and KF solutions coincide.

The KF and its extension to weakly non-linear systems (Extended Kalman Filter) remain to this day the "gold standard" of data assimilation for global NWP. This is because the current operational model resolutions (~10-30 Km), the global and dense observation distribution, the high quality of current forecast models and the nature of the variables that are commonly part of the analysis control vector make deviations from linearity sufficiently small to be handled in an incremental fashion. On the other hand it is well known that the KF is computationally intractable except for very low dimensional systems. A contemporary NWP model state has N~(10^8-10^9) variables so that the dimensionality of the error covariance matrices in Eq.(8-14) is N*N. This makes both the storage and the computation of these matrices unfeasible.

It is then clear that any practical approximation to the KF to be used in global NWP needs to tackle the fundamental issue of the dimensionality of the state error covariance matrices. Two different approaches to this problem have found widespread use in atmospheric data assimilation: 1) 4D-Var assimilation techniques; 2) Reduced-rank Kalman Filter techniques.

3. 4D-Var and the Kalman Filter

Assuming the forecast model used in the background computation to be *perfect* (*strong constraint* 4D-Var), the *smoothing* problem of finding the model trajectory that best fits the observations over an assimilation interval (t=0,1,…,T) for a given a background state $\mathbf{x_b}$ and its error covariance **B**, is solved finding the minimum of the following cost function:

$$J(\mathbf{x}_0) = (\mathbf{x}_b - \mathbf{x}_o)^T (\mathbf{B})^{-1} (\mathbf{x}_b - \mathbf{x}_o) + \sum_{t=0}^{T} (\mathbf{y}_t - H_t M_{0\to t}(\mathbf{x}_0))^T \mathbf{R}_t^{-1} (\mathbf{y}_t - H_t M_{0\to t}(\mathbf{x}_0)) \quad (15)$$

The model trajectory that minimizes this cost function can be shown (Fisher et al., 2005) to coincide, for linear Gaussian systems, to the Kalman Smoother solution over the assimilation interval for the same $\mathbf{x_b}$, **B** and to the Kalman filter solution at the end of the interval (t=T). However what the 4D-Var technique cannot easily provide is an estimate of the state error covariances. While it can be shown (Thepaut et al., 1996) that the 4D-Var solution implicitly evolves background error covariances *over the assimilation window* (currently 12 hours at ECMWF), its inability to re-use them in successive analysis cycles effectively means that information from past observations is only carried forward by the state estimates. This implies that the background error covariances used at the beginning of the assimilation window are static, approximately isotropic and the background error standard deviations are largely homogeneous. This inability to evolve state error covariances in time is a serious limitation of strong constraint 4D-Var.

A generalization of strong constraint 4D-Var which relaxes the perfect model assumption through an explicit representation of the model error covariances is possible (*weak constraint* 4D-Var, Trémolet, 2006). This approach would allow the use of a longer assimilation window which would then make the 4D-Var analysis solution less sensitive to the estimates of the background state and its errors. For very long windows ($\Delta T \geq 24h$) the problem in this method is effectively shifted from that of estimating **B** to one of having a realistic representation of **Q** (see Trémolet, 2007, and Fisher in these proceedings, for further details).

4. Reduced-rank Kalman Filter

The representation of $\mathbf{P^b}$ in 4D-Var is full rank. This is achieved through the use of a *parameterized* model of $\mathbf{P^b}$ at the start of the assimilation window. This model is calibrated off-line using a statistical sample of perturbation fields which are meant to represent the relevant background errors (Parrish and Derber, 1992; Fisher, 2003). This initial estimate of $\mathbf{P^b}$ is then *implicitly* evolved during the assimilation window by the use of the forecast model tangent linear and adjoint operators ($\mathbf{MP^bM^T}$; Thepaut *et al.*, 1996).

Another possible approach is to represent and evolve the background error matrix *explicitly*. Given its huge dimensionality for realistic NWP simulations, the only feasible option is to use a reduced-rank representation of $\mathbf{P^b}$ ($\mathbf{B}$) and $\mathbf{P^a}$ ($\mathbf{A}$). A low rank approximation of $\mathbf{P^b}$ can be written as:

$$\mathbf{B} = \mathbf{X^b}(\mathbf{X^b})^{\mathbf{T}} \tag{16}$$

Where $\mathbf{X^b}$ has dimensions Nxm and m<<N. Substituting (16) in (11), (13) and (14), one gets:

$$\mathbf{B} = \mathbf{M}_{t-1}\mathbf{P^a}\mathbf{M}^{\mathbf{T}}{}_{t-1} + \mathbf{Q}_{t-1,t} = \mathbf{M}_{t-1}\mathbf{X}^b\left[\mathbf{A}_{\text{mxm}}\right](\mathbf{M}_{t-1}\mathbf{X^b})^{\mathbf{T}} + \mathbf{Q}_{t-1,t} \tag{17}$$

$$\mathbf{A} = \mathbf{X^b}\left[\mathbf{I}_{\text{mxm}} + \left(\mathbf{HX^b}\right)^{\mathbf{T}}\mathbf{R}\left(\mathbf{HX^b}\right)\right](\mathbf{X^b})^{\mathbf{T}} \tag{18}$$

$$\mathbf{K} = \mathbf{P^b}\mathbf{H}^{\mathbf{T}}\left[\mathbf{HP^b}\mathbf{HT} + \mathbf{R}\right]^{-1} = \mathbf{X^b}\left(\mathbf{HX^b}\right)^{T}\left[\left(\mathbf{HX^b}\right)\left(\mathbf{HX^b}\right)^{\mathbf{T}} + \mathbf{R}\right]^{-1} \tag{19}$$

From (17-19) it is apparent that N (the dimension of the state) has disappeared from the Kalman Filter equations and it has been substituted by the (considerably smaller) m. This is the basis for reducing the computational cost of the KF. However the KF update for the state (Eq. 12) in this framework can be written as:

$$\begin{aligned}\mathbf{x_a} - \mathbf{x_b} &= \mathbf{K}\left(\mathbf{y} - \mathbf{Hx_b}\right) = \mathbf{P^b}\mathbf{H}^{\mathbf{T}}\left[\mathbf{HP^b}\mathbf{HT} + \mathbf{R}\right]^{-1}\left(\mathbf{y} - \mathbf{Hx_b}\right) = \\ &= \mathbf{X^b}\left(\mathbf{HX^b}\right)^{T}\left[\left(\mathbf{HX^b}\right)\left(\mathbf{HX^b}\right)^{\mathbf{T}} + \mathbf{R}\right]^{-1}\left(\mathbf{y} - \mathbf{Hx_b}\right)\end{aligned} \tag{20}$$

From (20) it is clear that analysis increments are confined to the subspace spanned by $\mathbf{X^b}$, which has rank no larger than m<<N. This can cause the reduced-rank KF analysis to be severely *rank-deficient*.

Reduced-rank KF has become popular in atmospheric NWP following the introduction of the Ensemble Kalman Filter (EnKF; Evensen, 1994; Burgers *et al.*, 1998). EnKF is a Monte Carlo approximation of the KF which crucially does not require the Tangent Linear and Adjoint operators of the model $\mathcal{M}$ and the observation operator $\mathcal{H}$.

In the EnKF the columns of $\mathbf{X^b}$ are the ensemble perturbations with respect to the ensemble mean (scaled by a $\sqrt{(N_{ens}-1)}$ factor, N_{ens} being the ensemble size), so that the subspace spanned by $\mathbf{X^b}$ still has dimension $N_{ens}-1 << N$.

The common solution of spatially localizing the $\mathbf{B}$ sampled from the ensemble (Schur product covariance localization, Hamill *et al.*, 2001; Local analysis, Evensen, 2003; Ott et al., 2004; Adaptive localization, Anderson 2006, Bishop and Hodyss, 2009) has proved effective in controlling sampling noise and implicitly increasing the ensemble size (because different linear combinations of ensemble perturbations are used at different grid points), but it comes at a cost. With $\mathbf{B}$ localization the analysis dynamical balance may be degraded (Mitchell *et al.*, 2002), the asymptotic convergence of the EnKF to the KF is not guaranteed and, more importantly, theoretical and practical problems arise for the assimilation of non-local observations (such as satellite radiances; Campbell *et al.*, 2010). This last

issue is highlighted in Figs. 1,2, which show the relative performance of the ECMWF implementation of the EnKF and ECMWF 4D-Var for an assimilation cycle making use of conventional observations only (Fig. 1) and the full observing system (Fig. 2). It is apparent that, at the time of writing, the ECMWF EnKF is competitive with 4D-Var only when conventional observations are assimilated, but not when the full observing system with satellite radiance observations is used.

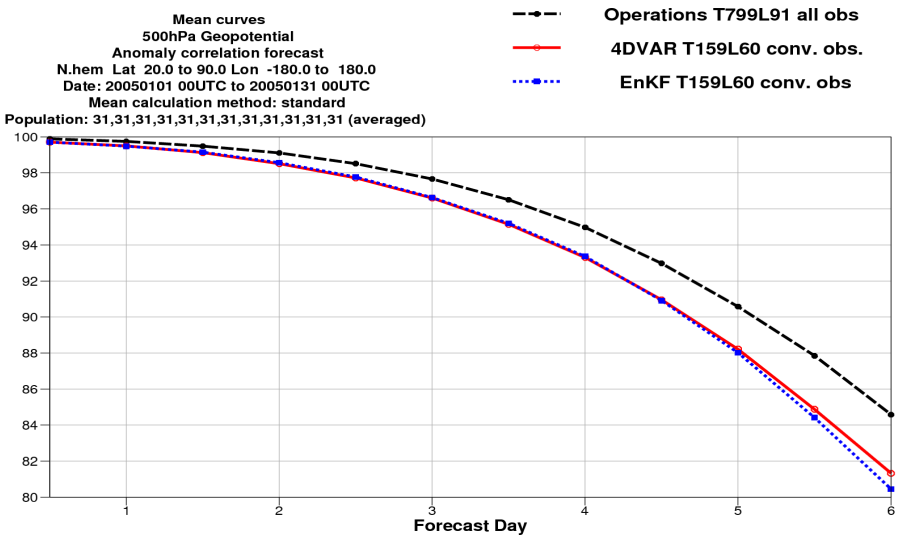

Figure 1: Forecast skill scores (Anomaly Correlation of geopotential height at 500 hPa) of: Forecasts from 4D-Var assimilation cycle using conventional observations only,at reduced resolution (T159L60, red line);Forecasts from EnKF assimilation cycle using conventional observations only,at reduced resolution (T159L60, dotted blue line); Forecasts from operational 4D-Var with full observing system (T799L91 forecasts and T255L91 analysis outer loop). Scores are averaged over the period 20050101-20050131 for the Northern Hemisphere.

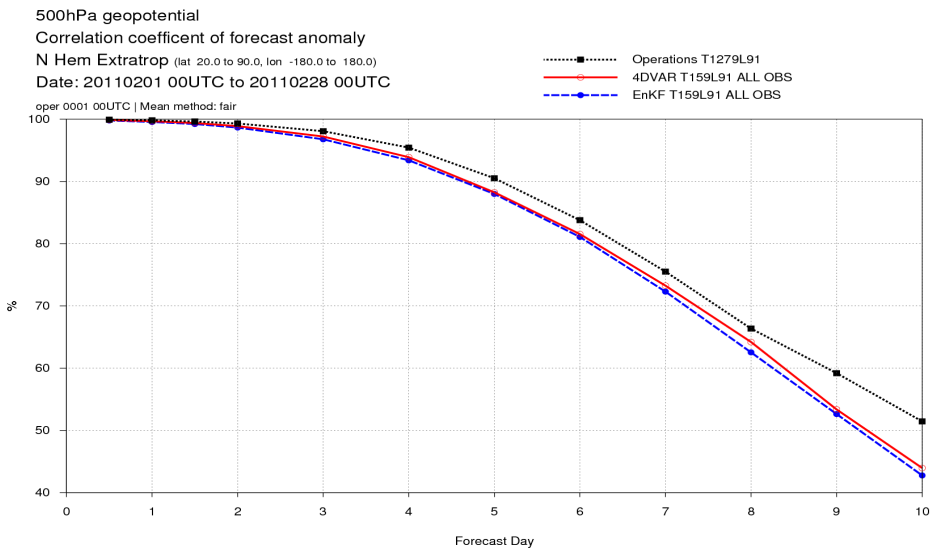

Figure 2: Forecast skill scores (Anomaly Correlation of geopotential height at 500 hPa) from: *Forecasts from 4D-Var assimilation cycle with full observing system at reduced resolution (T159L60, red line); Forecasts from EnKF assimilation cycle with full observing system at reduced resolution (T159L60, dashed blue line); Forecasts from operational 4D-Var with full observing system (T799L91forecats and T255L91 outer loop; dotted black line). Scores are averaged over the period 20110201-20110228 for the Northern Hemisphere.*

5. Hybrid approaches

The basic idea of the various hybrid data assimilation techniques which have been proposed in the atmospheric data assimilation literature (starting with the seminal paper of Hamill and Snyder, 2000) is to use flow-dependent error estimates from an ensemble data assimilation system (such as an EnKF) in a deterministic 3D-Var or 4D-Var analysis. In this way flow-dependent error covariance information can be integrated into a variational algorithm keeping at the same time the full rank representation of **B**. This type of solution would be more robust than a pure EnKF for the limited ensemble sizes which are currently affordable and would also allow a more consistent use of non-local observations (Campbell *et al.*, 2010).

A problem that hybrid DA shares with the EnKF is the noisiness and rank-deficiency of the **B** sampled from the ensemble background. Two different approaches are commonly used to tackle it in operational environments.

The "Alpha control variable" method (Barker 1998, Lorenc 2003; see also Barker in this Proceedings for further details) is being used or is under development at the UK Met Office, NCEP (USA) and CMC (Canada). In this method the background error covariance is conceptually modelled as a linear combination of a climatological estimate ($\mathbf{B}_c$) and a spatially localised covariance estimate sampled from the ensemble:

$$\mathbf{B} = \beta_c^2 \mathbf{B}_c + \beta_e^2 \mathbf{P}_e \circ \mathbf{C}_{loc} \tag{21}$$

In (21) $\mathbf{P}_e \circ \mathbf{C}_{loc}$ represents the Schur (element-wise) product of the sampled ensemble covariance $\mathbf{P}_e$ with a predefined spatial localization matrix $\mathbf{C}_{loc}$. In a variational framework the background error covariance is not represented explicitly as in (21) but is modelled in terms of a change of variable from the state variables to the analysis control variables. In this framework the error covariance in (21) is implicitly represented by:

$$\delta \mathbf{x} = \beta_c \mathbf{B}_c^{1/2} \mathbf{v} + \beta_e \mathbf{X}' \circ \boldsymbol{\alpha} \tag{22}$$

Here the standard analysis control variable **v** has been *augmented* by the alpha control variable (**X**' are the ensemble perturbations). The alpha control variable is itself constrained by an additional term in the Jb term of the 4D-Var cost function:

$$J = J_b + J_o + J_c = \frac{1}{2}\boldsymbol{v}^T\boldsymbol{v} + \frac{1}{2}\boldsymbol{\alpha}^T \boldsymbol{C}_{loc}^{-1}\boldsymbol{\alpha} + J_o + J_c \tag{23}$$

$\boldsymbol{C}_{loc}$ is an empirical spatial localization operator which serves a similar purpose as the **B** localization in the EnKF. The two constants $\boldsymbol{\beta}_c$ and $\boldsymbol{\beta}_e$ determine the relative weight given to the climatological and the ensemble **B** respectively.

The other common approach to hybrid DA, developed at Météo-France and ECMWF, is essentially a flow-dependent extension of the way the climatological **B** matrix is computed (Fisher, 2003). It is described in detail in the following paragraph.

6. Ensemble of Data Assimilations

The ECMWF Ensemble of Data Assimilations (EDA) is a system of N (N=10 in September 2011) independent, reduced resolution, 4D-Var assimilation cycles which differ by perturbing observations, sea-surface temperature fields and model physics. A theoretical analysis shows that if the perturbations are drawn from the true distributions of observation and model error, then the spread of the EDA analysis/background about the control (unperturbed) analysis will be representative of the analysis/background error (Isaksen *et al.*, 2010).

The EDA is already being used for the specification of climatological, static background error statistics in 4D-Var at ECMWF (Fisher, 2003) and Météo-France (Belo Pereira and Berre, 2006). The climatological **B** matrix used in the ECMWF 4D-Var can be written in this form (Derber and Bouttier, 1999:

$$\mathbf{B} = \mathbf{T}^{-1}\boldsymbol{\Sigma}_b^{1/2}\mathbf{C}\boldsymbol{\Sigma}_b^{1/2}\mathbf{T}^{-T} \tag{24}$$

Here **T** is a matrix representation of the balance operator, $\boldsymbol{\Sigma}_b^{1/2}$ is the diagonal matrix of background error standard deviations and **C** is the background error correlation operator (which is modeled in wavelet space in the ECMWF implementation: Fisher, 2003). The basic idea of a hybrid assimilation system based on the EDA is that of extending its use to make flow-dependent estimates of the background error standard deviations ($\boldsymbol{\Sigma}_b^{1/2}$) and correlation operator (**C**). We review the use of EDA derived background error standard deviations first, trying to clarify the mechanisms through which they affect the 4D-Var analysis and their impact on the ECMWF deterministic forecast skill scores. The use of EDA estimates for the online estimation of the error correlations is currently an active area of research and will be discussed next.

7. Use of EDA Variances

The background error standard deviations are modeled in grid point space in Eq. (24) and can thus be directly sampled from the EDA background fields:

$$\boldsymbol{\Sigma}_b(i,j,k) = \frac{1}{N_{EDA}-1}\sum_{l=1}^{N_{EDA}}\left(\mathbf{x}_b^l(i,j,k) - \overline{\mathbf{x}}_b(i,j,k)\right)^2 \tag{25}$$

Here the triplet (i,j,k) identifies the grid point.

The sampled variance estimates in Eq. (25) are affected by two problems: a) *sampling noise* due the small ensemble size (N_{EDA}=10), and b) *systematic errors* due to incorrect specification of error sources in the EDA (i.e., mis-specification of **R**, **B**, **Q**, uncertainties in the boundary conditions).

The sampling error can be ameliorated by increasing the ensemble size, but it must be recalled that the standard error for the sampled variance only decreases with the inverse square root of the ensemble size:

$$\hat{\sigma}_{\Sigma_b} = \sqrt{\frac{2}{N_{EDA}-1}}\,\Sigma_b \tag{26}$$

This means that some kind of spatial filtering needs to be applied to the raw sampled $\mathbf{\Sigma}_b$ to extract the statistically significant signal. The key insight in the filtering of EDA variances has been the recognition that sampling noise is small scale with respect to the background error variance field (Raynaud *et al.*, 2008). This allows the application of an objective filtering procedure in spectral space, as illustrated in Fig. 3. An example of the effect of the objective filter on the sampled EDA variances is given in Fig.4. Fig. 4 also displays (bottom row) the estimated background errors obtained with the randomization method (Fisher and Courtier, 1995) which was previously used at ECMWF. The "randomization" errors are computed from random samples of the static **B** matrix with a small flow-dependent component deriving from the application of the non-linear balance equation and the omega equation linearized around the background state (Fisher, 2003). The filtered variance map (second row of Fig.4) shows the same large scale features present in the raw EDA spread map (first row of Fig.4), but most of the smaller scale features which are indistinguishable from sampling noise have been removed. It is also apparent how the EDA error estimates present a clearer signature of the current synoptic situation than the randomization error map: EDA errors tend to be concentrated in meteorologically active areas of sparse observational coverage.

The discussion above highlights the fact that the presence of sampling noise effectively limits the spatial scales of background errors that can be robustly estimated from the EDA sample variance. The effective spatial resolution of the EDA error estimates is thus predominantly constrained by the ensemble size and it is usually much coarser than the nominal resolution of the EDA background fields (Bonavita *et al.*, 2010). An example is given in Fig. 5, where the power spectra of the EDA temperature background variance (continuous lines) and of the corresponding sampling noise (dashed

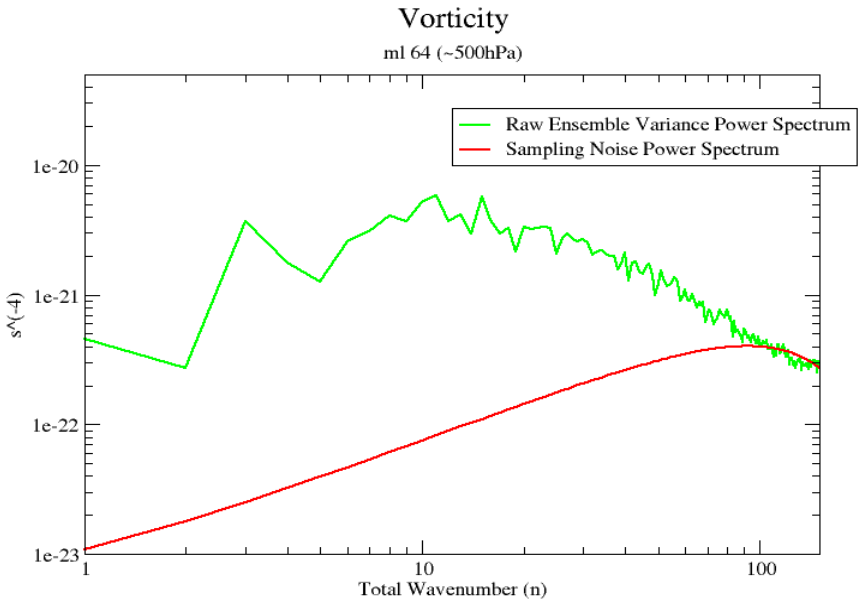

Figure 3: Power spectra of EDA short range (t+9h) forecast variance (green line) and of the sampling noise (red line) associated to the estimation of vorticity variances at model level 64 (~500 hPa) with a 10-member ensemble. Unit: s^{-4}.

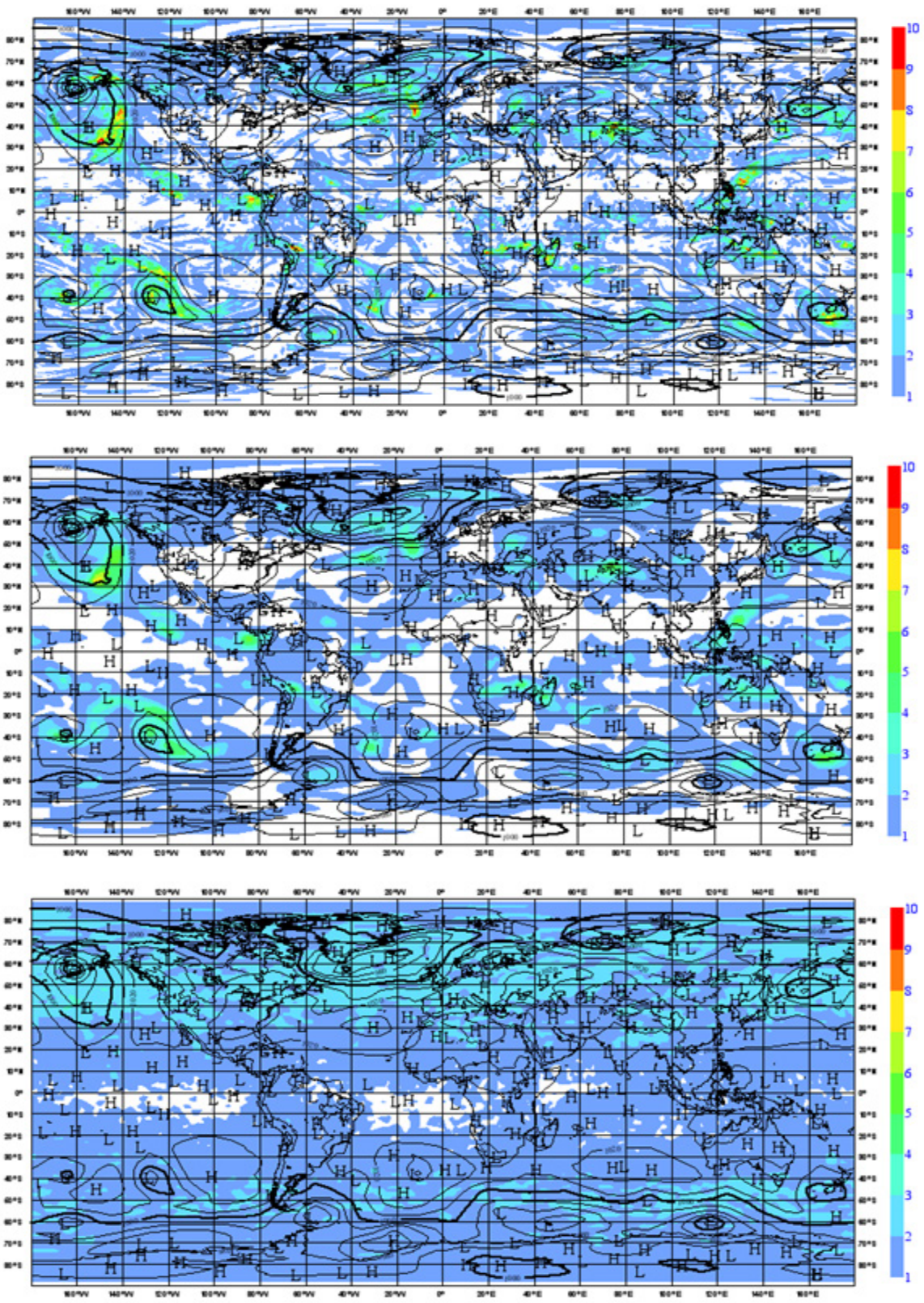

Figure 4: Standard deviations of vorticity at model level 64 (~500hPa), valid on 16 January 2009 at 21 UTC. Unit: 5.10^{-5} s^{-1}. Raw estimates from the 10 member ensemble (top); objectively filtered estimate (middle); background error estimate from the "randomization" technique (bottom row). The mean sea level pressure analysis is overlaid, contours: 8 hPa.

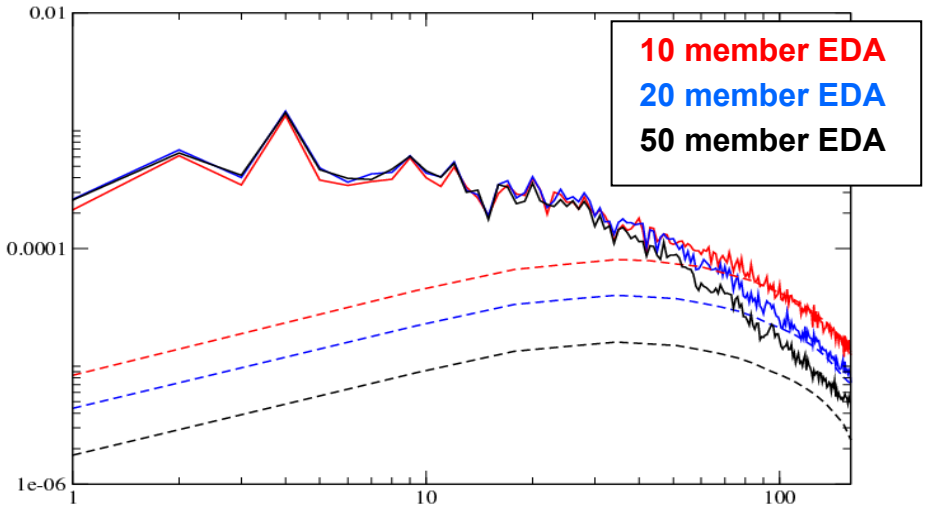

Figure 5: Power spectra of ensemble first guess variance valid on 21 Jan 2009, 09UTC (continuous lines) and of the sampling noise (dashed lines) associated with the estimation of temperature variances at model level 49 (~200 hPa) for a 10-member ensemble (red), a 20-member ensemble (blue) and a 50 member ensemble (black). Unit: $°C^4$.

lines) at model level 49 (~200 hPa) is shown for a 10-member ensemble (red), a 20-member ensemble (blue) and a 50 member ensemble (black). It is clear that for smaller spatial scales (high wavenumbers) the sampling noise contribution to the total variance spectra is smaller for the larger ensembles. The climatological noise spectra estimate is reduced even further, thus allowing the filter to retain finer scale structures from the original signal. In this context, the effective spatial resolution of the statistically significant error estimates that we can compute from the ensemble goes from aro total wavenumber 70 for the 10-member ensemble to ~90 for the 20-member ensemble to >159 for the 50-member ensemble.

The filtering procedures discussed so far operate in spectral space, which implies that the resulting low-pass filter is spatially homogeneous. On the other hand, both the sampling noise and the background error variance can be expected to have different dominant spatial scales in different regions of the Globe (e.g., tropics and extra-tropics) and in different synoptic situations. From these considerations a type of filter based on a wavelet expansion (Fisher, 2003) can provide a degree of spatial heterogeneity together with a degree of spectral resolution. Further details of this research can be found in Bonavita *et al.*, 2012.

The other crucial aspect to consider in the use of EDA estimates of error variances is the possible presence of non negligible systematic differences with the true background errors. This type of errors reflects the underlying outstanding deficiencies in the representation of the main sources of uncertainties in the EDA (i.e., observation error covariances, model error parameterizations, boundary condition uncertainties) and cannot be corrected by an increase in ensemble size. The systematic errors of the EDA variances usually have an inhomogeneous spatial structure (Fig. 6). The relationship between EDA spread and analysis/background errors can be investigated by the use of spread-error plots, an example of which is given in Fig. 7. Spread-error plots are constructed to represent the conditional distribution of the ensemble background mean Root Mean Squared error given the ensemble standard deviation. They are obtained by dividing the sample EDA spread into

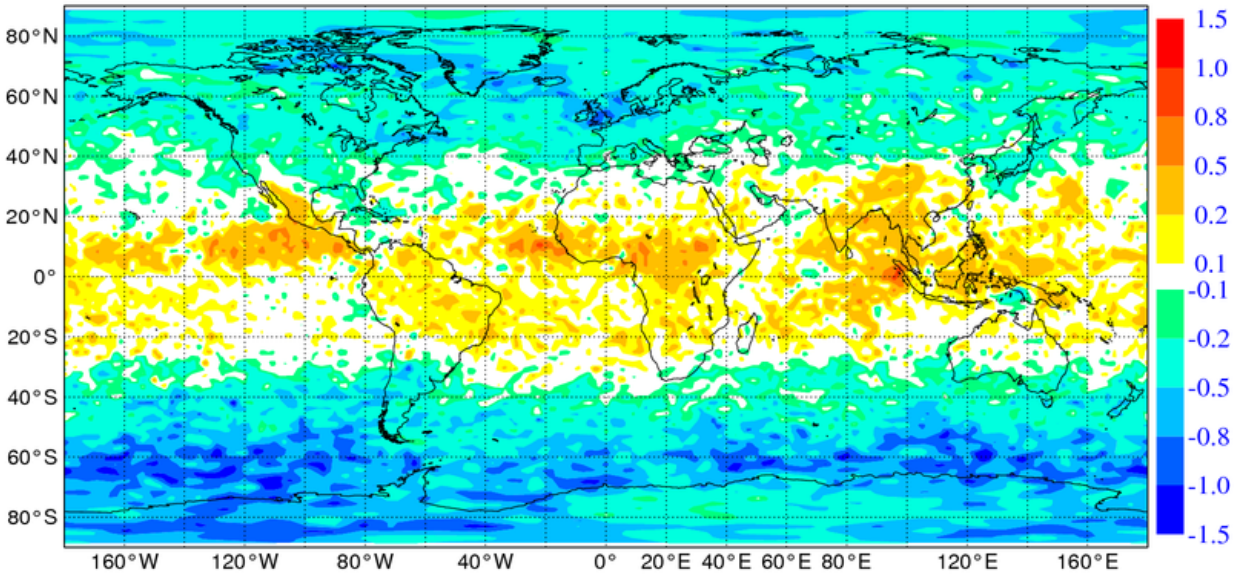

Figure 6: Difference between EDA background forecast spread and EDA mean forecast error for vorticity at model level 64 (~500 hPa, top) averaged over one month (August-September 2009). Unit: 10^{-5} s^{-1}

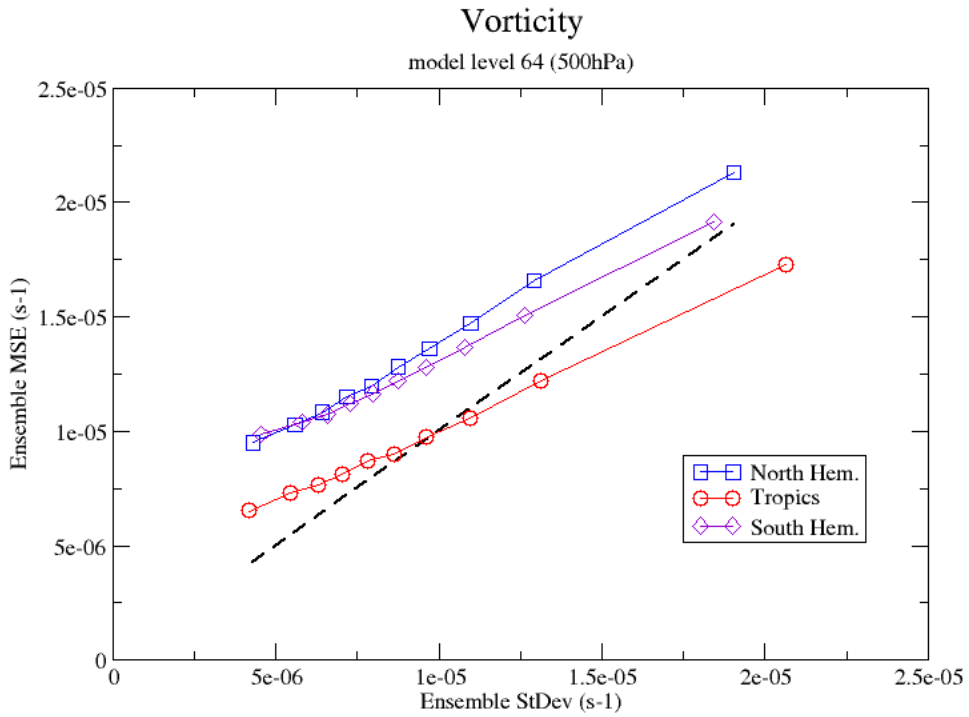

Figure 7: Spread-Error diagram for vorticity at model level 64 (~500hPa). Average over one month (August 2009).

quantiles according to their magnitude and plotting their mean value together with the mean values of the corresponding analysis/background errors. For a reliable EDA, the calibration curves (separately computed for the northern extra-tropics, southern extra-tropics and the tropics) should lie on the diagonal (dashed black line) (Kolczynski *et al.*, 2009, 2011). Their distance to the diagonal shows the under/over-dispersiveness of the ensemble while their slope gives an indication of the conditional biases of the ensemble spread (Bonavita *et al*., 2012).

To reduce the impact of the systematic errors in the estimated variances, a rescaling step is currently applied to the raw (unfiltered) EDA estimates. This is aimed at enforcing approximate statistical consistency between the EDA variance estimates and the diagnosed ensemble mean background errors following Leutbecher, 2009:

$$\frac{N}{N-1}\mathbf{E}\left[\frac{1}{N}\sum_{j=1}^{N}\left(x_j-\frac{1}{N}\sum_{j=1}^{N}x_i\right)^2\right]=\frac{N}{N+1}\mathbf{E}\left[\frac{1}{N}\sum_{j=1}^{N}x_j-y\right]^2 \qquad (27)$$

Here N is the ensemble size, E[ ] represents the statistical expectation operator and y is the true state (in the present case, the operational ECMWF analysis is taken as 'truth'). Eq. (27) is enforced separately for each spread-error bin, variable and latitude band (northern extra-tropics, tropics, southern extra-tropics). Time-averaging is also applied by computing a running mean of the scaling factors over the previous 5 days (i.e., the last available 10 analysis cycles) leading up to the nominal analysis date.

8. Impact of EDA Variances on 4D-Var

The use of EDA variances as background errors estimates impacts the ECMWF 4D-Var analysis in two distinct ways: The first guess quality control of the observations and the formulation of the **B** matrix.

The observation quality control step is designed to flag marginal observations whose squared background departures exceed some predefined multiple of the sum of the background error and the observation error variances. Given the flow-dependency of the EDA background error variances, the new system is expected to be able to make more accurate decisions in terms of observation acceptance/rejection. While global observation usage is practically unchanged, there are indications (see Bonavita et al., 2012, for further details) that the observation selection based on EDA variances has a statistically significant positive impact on analysis quality, because they differ most near active weather systems.

EDA variances also affect 4D-Var by providing a flow-dependent estimate of the diagonal matrix of background error standard deviations ($\mathbf{\Sigma}_b^{1/2}$) in the **B** model of Eq. 24. Since the $\mathbf{\Sigma}_b^{1/2}$ matrix is defined in grid point space, its EDA derived estimate can introduce a significant degree of heterogeneity and anisotropy in the analysis increments. This point is illustrated in Fig. 8-9. They represent the result of an assimilation experiment in which a single simulated temperature observation with a background departure $\Delta T=+1K$ is assimilated at the start of the 4D-Var window. The

observation is located at (34N,74W), 900 hPa, in proximity of the vorticity error maximum diagnosed by the EDA spread (middle panel in Fig. 8). Note how, in contrast, the vorticity error estimates from the randomization method (bottom panel in Fig. 8) are predominantly homogeneous and isotropic. The analysis increments derived from the two background error models (Fig. 9) reflect to some degree their different spatial structures and intensities. A large difference in the magnitude of the analysis increments is visible: in the randomization experiment the maximum (minimum) temperature (vorticity) increment is 0.37 K ($-1.30*10^{-5}$ s^{-1}), respectively; in the 'EDA errors' experiment the corresponding value is 0.749 K ($-5.0*10^{-5}$ s^{-1}). Further to this, the shape of the analysis increments is clearly affected by the structure of the vorticity error standard deviations: while the increments of the randomization experiment have a predominantly isotropic structure, the increments of the 'EDA errors' experiment have, especially for vorticity, a non-negligible state-dependent component.

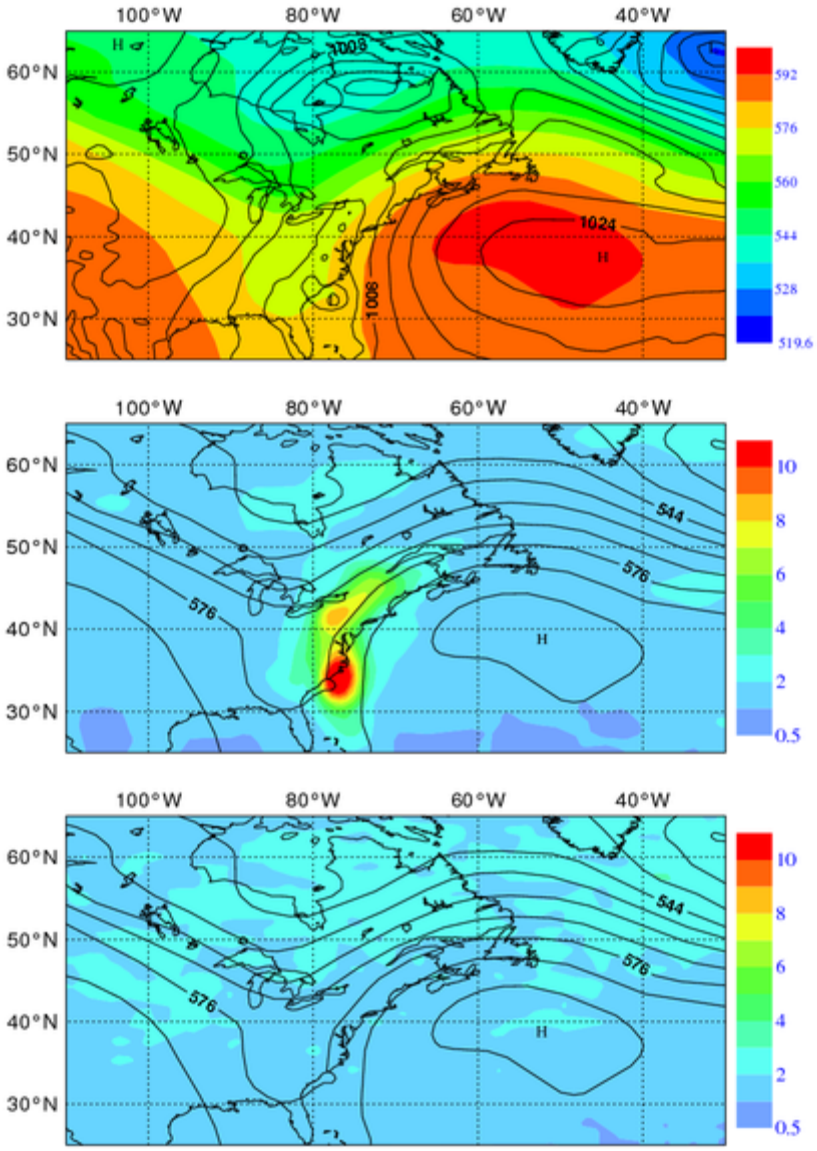

Figure 8: Geopotential at 500 hPa (shaded contours) and MSLP background fields valid on 30 September 2009, 21 UTC (first row). Geopotential at 500 hPa and vorticity background error standard deviations estimated from the ECMWF EDA (second row). Geopotential at 500 hPa and vorticity background error standard deviations computed with the randomization method (third row).

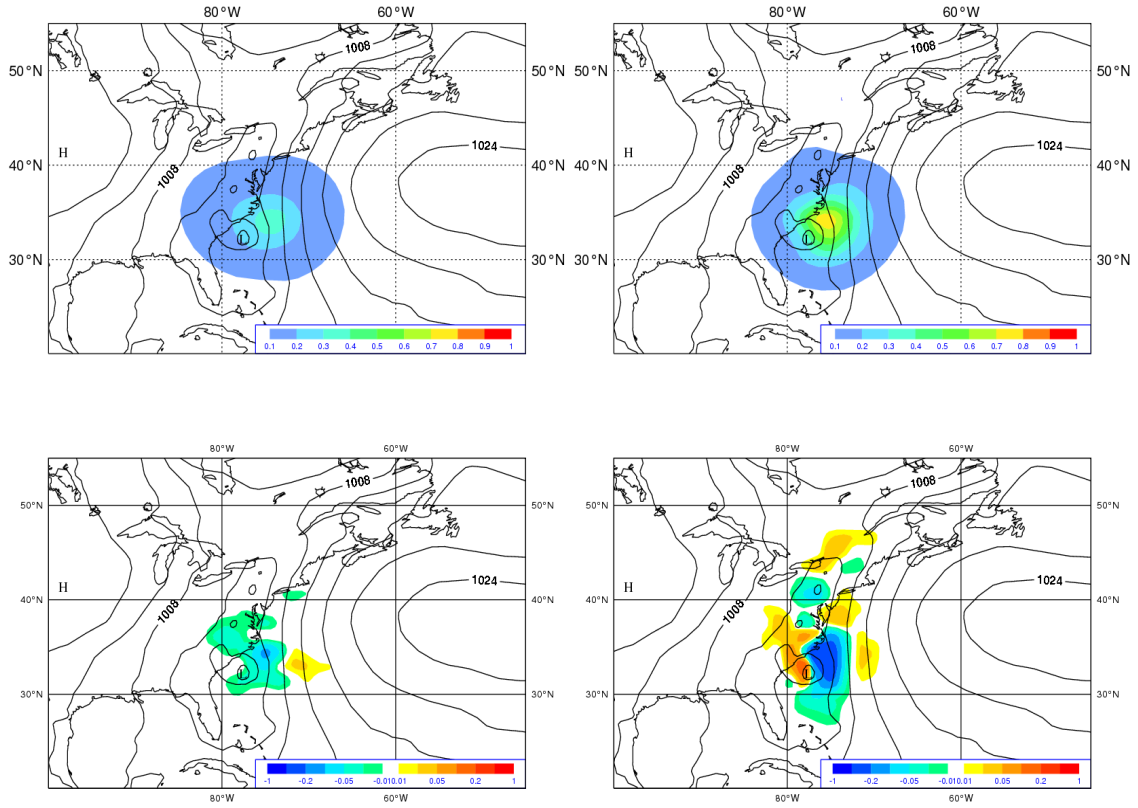

Figure 9: Single observation (observation departure ΔT=+1K at 900 hPa) analysis increments for temperature at model level 81 (first row) and vorticity at model level 78 (second row) valid on 30 September 2009, 21 UTC. The first column refers to an assimilation experiment making use of the "randomization" error estimates, the second column to an assimilation experiment making use of estimates from the ECMWF EDA. Scale intervals are in degrees for the temperature increments, in 10^{-5} s^{-1} for the vorticity increments.

The impact of the use of EDA variances in the 4D-Var analysis has been tested in two long (~3 months) assimilation experiments over the winter (11/01/2010 to 30/03/2010) and summer/autumn (2/08/2010 to 30/10/2010) seasons. The assimilation experiments are run with the ECMWF Integrated Forecasting System (IFS) cycle 36R4, at the operational resolution (T1279L91, with 3 incremental 4D-Var loops run at T159-T255-T255 resolution). The controls differ from the experimental runs only in that they use the previous quasi-static background errors estimates. Fig. 10 and 11 present vertical cross-sections of the RMS reduction of geopotential height errors at forecast lengths ranging from t+12h to t+192h for the winter EDA variance experiment (*ffg8*) against the control (*fezj*) and the boreal summer experiment (*ffge*) against its control (*0051*), respectively. All forecasts are verified against own analysis. The improvement in forecast skill is apparent at most forecast ranges, latitude bands and pressure levels. It is interesting to note that improvements tend to be larger in the winter hemisphere. This is consistent with the notion that in regions and times of the year where synoptic activity is stronger the impact of the use of EDA variances will be more significant.

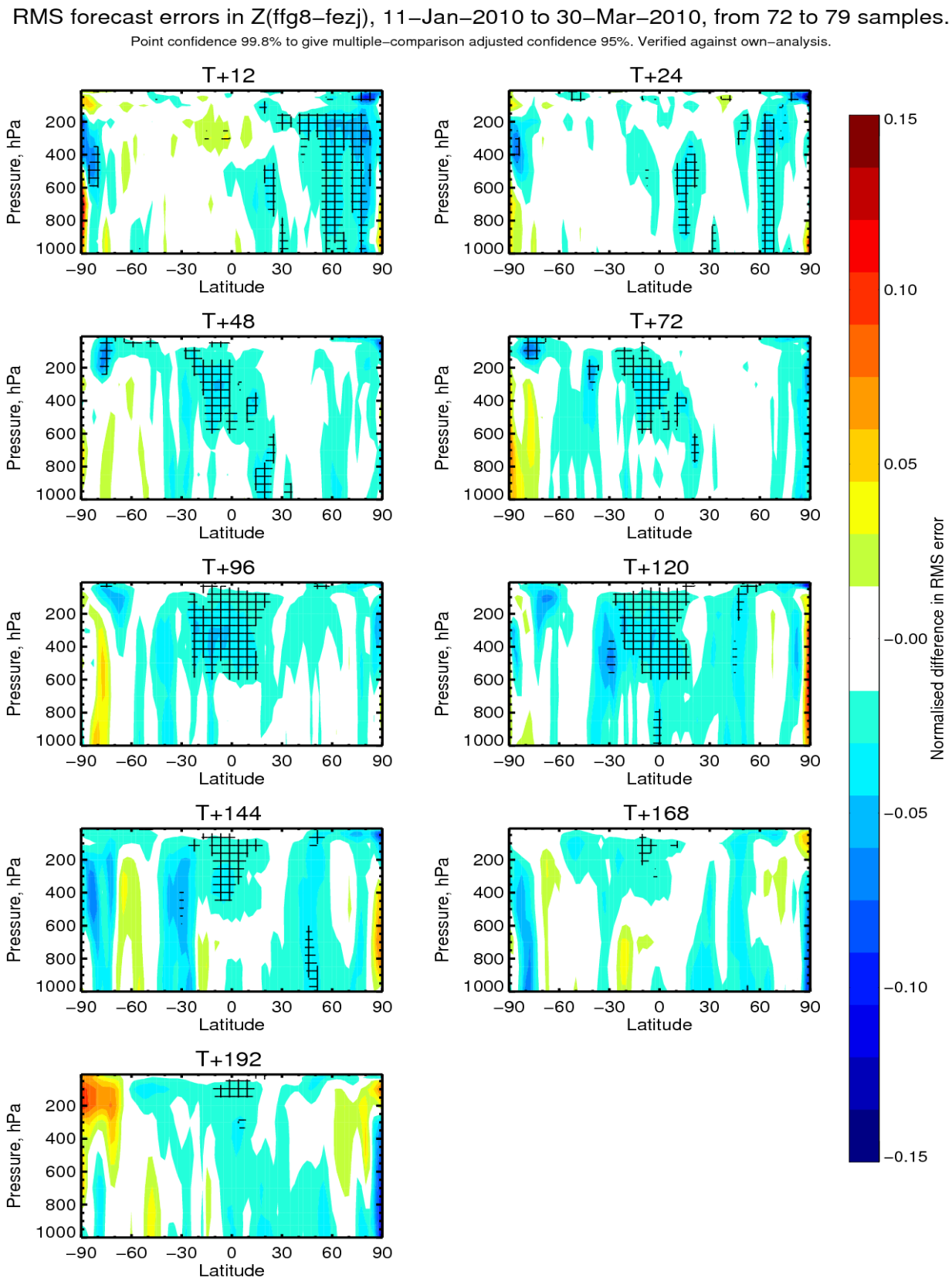

Figure 10: Meridional averages of Root Mean Squared Error reduction of Geopotential forecasts from deterministic 4D-Var analyses with background error variances estimated from a 10 member EDA (experiment ffg8) versus using randomization errors (experiment fezj). Forecasts are verified against own analysis. Scores are averaged over the period 2010111-20100330. Crosses indicate statistical reliability of the results at the 95% confidence level.

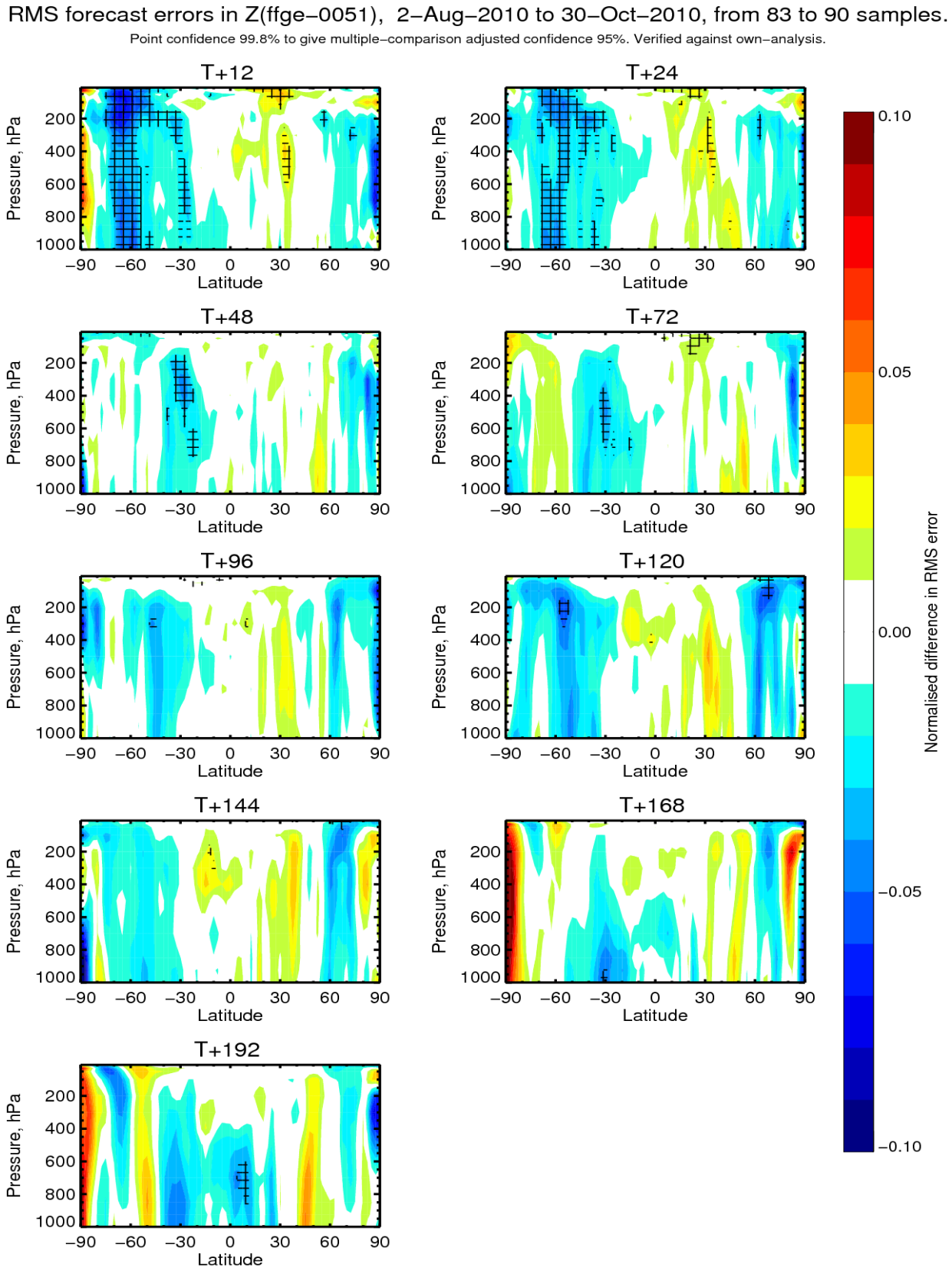

Figure 11: Meridional averages of Root Mean Squared Error reduction of Geopotential forecasts from deterministic 4D-Var analyses with background error variances estimated from a 10 member EDA (experiment ffge) versus using randomization errors (experiment 0051). Forecasts are verified against own analysis. Scores are averaged over the period 2010802-20101030. Crosses indicate statistical reliability of the results at the 95% confidence level.

9. Use of EDA Covariances

In the Gaussian framework used in linear estimation theory (and incremental 4D-Var) errors are completely characterized in terms of their variance-covariance matrix. We have seen in the previous paragraph how EDA variances are used to provide flow-dependent estimates of error variances. In a similar way, EDA perturbations can be used to construct online estimates of background error covariances.

EDA perturbations can provide realistic online estimates of background error covariances. An example is given in Figs. 12-14 where the background error spread and correlation length scale are diagnosed from EDA perturbations for the case of hurricane Fanele, near the coast of Madagascar (19-23 January 2009). It is apparent from these plots how the EDA can provide useful information not only on the size of background errors but also on their covariance structures. It is also clear, however, that the much larger dimensionality of the covariance estimation problem makes the EDA correlation length scale estimates more sensitive to sampling errors: Note how the EDA estimated error correlation length scale is noisier than the corresponding EDA estimated error standard deviation and more sensitive to the ensemble size (20 member EDA estimates in Fig. 13, 50 member EDA estimates in Fig. 14).

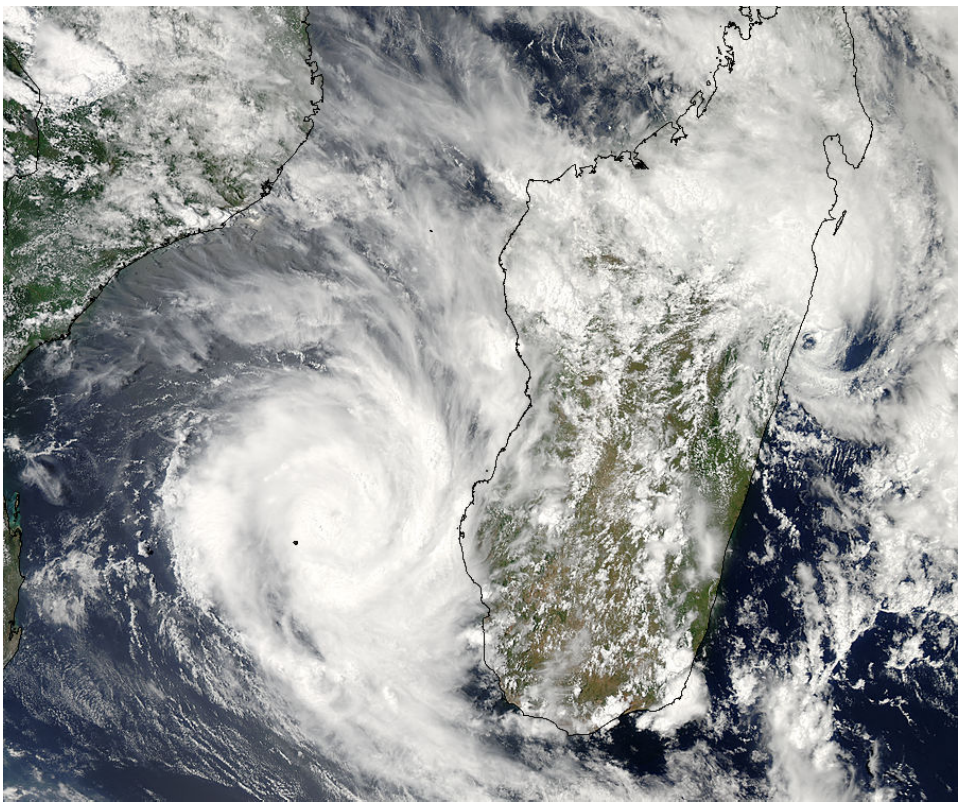

Figure 12: MODIS image from the Aqua Satellite of Hurricane Fanele (19 January 2009, 11UTC).

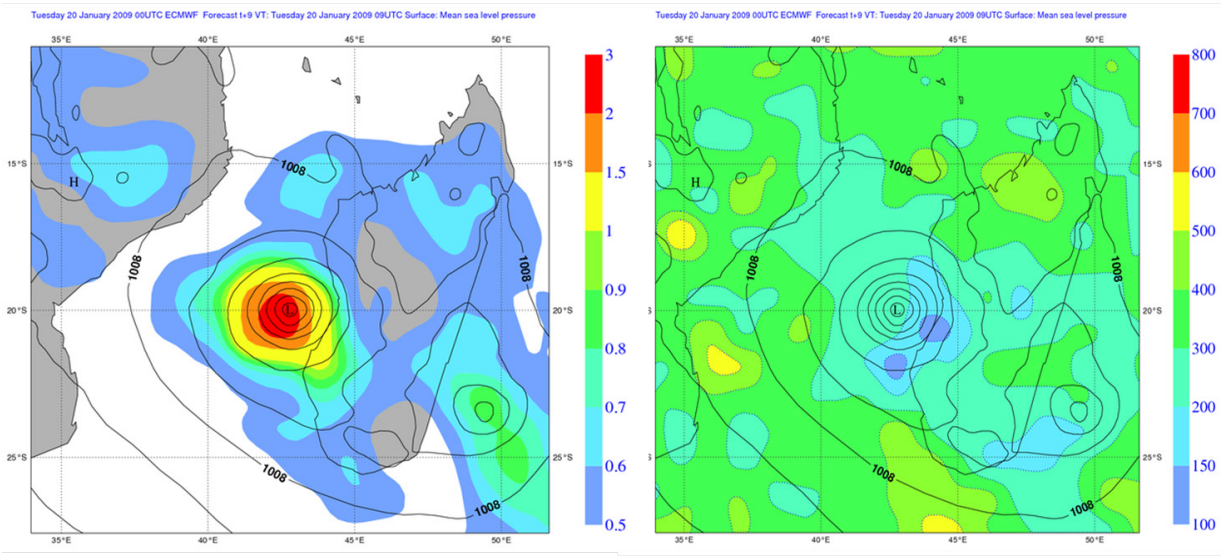

Figure 13: EDA estimate of surface pressure background error standard deviation (shaded contours, left panel) and EDA estimate of background error correlation length scale (shaded contours, right panel) in the area affected by hurricane Fanele. Plots are derived from a 20 member EDA, valid on 20 January 2009 at 09 UTC. Mean Sea Level Pressure contours (4 hPa interval) are superimposed.

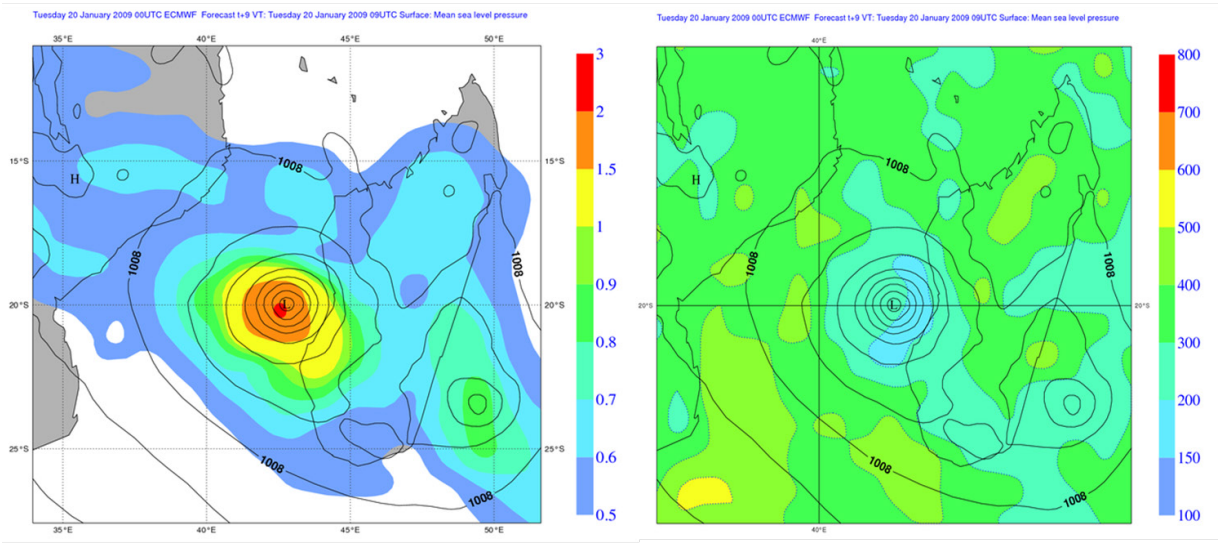

Figure 14: EDA estimate of surface pressure background error standard deviation (shaded contours, left panel) and EDA estimate of background error correlation length scale (shaded contours, right panel) in the area affected by hurricane Fanele. Plots are derived from a 50 member EDA, valid on 20 January 2009 at 09 UTC. Mean Sea Level Pressure contours (4 hPa interval) are superimposed.

The online estimation of the full covariance structures clearly requires some form of regularization when using a small dimensional ensemble (10-100 members). In the case of the ECMWF 4D-Var, a form of regularization is implicit in the use of the background error covariance model of Eq. (24). This is because: a) the balance operator **T** used in the **B** model represents linear, first order balances only; b) the use of the wavelet formulation for the correlation structures (**C**) implies a local spatial averaging of the correlation functions (Pannekoucke *et al.*, 2007). The calibration of this simplified **B** matrix is currently done offline aggregating one to two months of EDA short range forecast perturbations. A possible strategy to introduce a flow-dependent component in the estimated covariances is to reduce the training period on which the sample statistics are computed. This idea has been explored at Météo France with encouraging preliminary results (Varella *et al.*, 2011). An example is given in Figs. 15-16. Fig. 15 shows the diagnosed background error correlation length-scales for wind at 500 hPa using a three week calibration period, while Fig. 16 shows the same quantity estimated from a four day calibration period. While the three-week estimate mainly shows the expected "climatological" features (i.e., shorter correlations in the extra-Tropical storm tracks, longer correlations in the Tropics), the four-day estimate reflects more transient, synoptic features. This is apparent, for example, in the reduced correlation length-scales over the eastern seaboard of the US, connected with an area of cyclonic development (Fig. 17).

A related approach to introduce a degree of flow-dependency in the covariance structures in **B** is to represent it as a convex combination of the climatologically calibrated **B** and a $\mathbf{B}_{ens}$ matrix computed from the most recent EDA background perturbations. This is conceptually similar to what is done in the alpha control variable method (see Eq. 21) , with the fundamental difference that the flow-dependent component of **B** is still a parameterized **B** instead of a **B** directly sampled from the ensemble and localized (in space and, possibly, in scale). The use of a linear combination of a sampled $\mathbf{B}_{ens}$ and predefined climatological target **B** is a standard regularization tool widely used in statistics and it is the solution currently being pursued at ECMWF.

10. Summary and outlook

The atmospheric data assimilation problem essentially consists of blending a-priori information on the state and its errors with newly available observational data, according to some statistical optimality principle. It has been shown that, for systems (forecast model and observations) characterized by linear error evolution and Gaussian error distributions, the Bayesian solution to the DA problem leads to the Kalman Filter equations. In current global NWP data assimilation the hypotheses of Gaussian error distribution and linear error evolution between successive analysis update are reasonable. This is due to the current models' accuracy and typical resolution (10-20Km), the choice of analysis control variables, the density and homogeneity of the global observing system. Unfortunately, the dimensionality of the problem in realistic applications makes the direct use of the KF equations unfeasible. The core of the problem lies in the size of the background error covariance matrix **B** which makes it impossible to store it and evolve it explicitly.

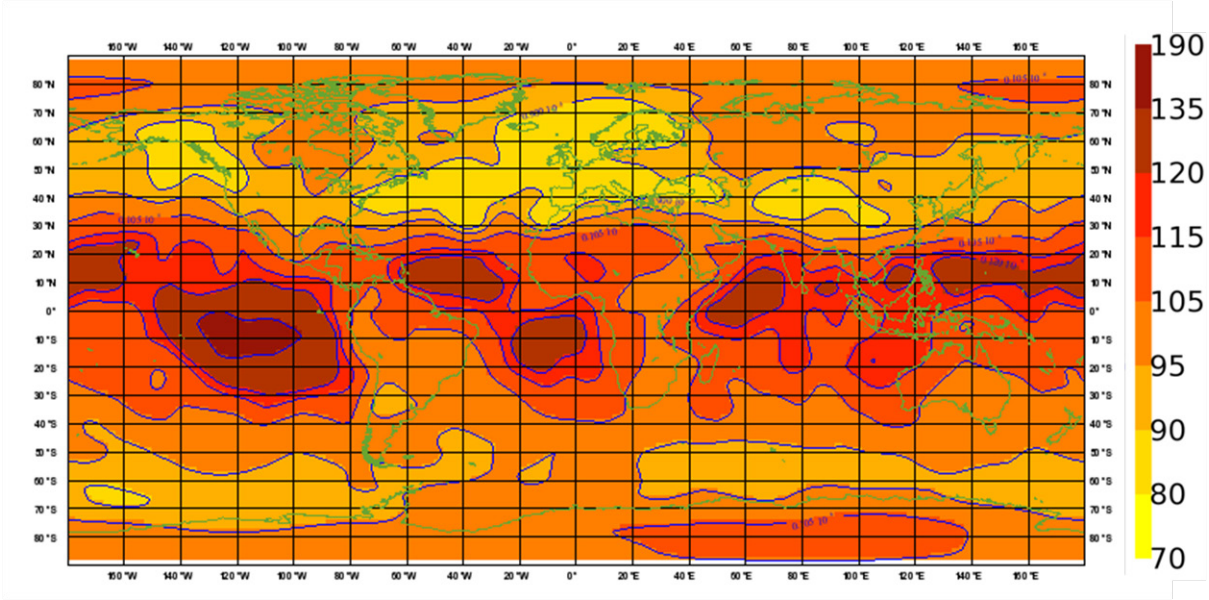

Figure 15: EDA "climatological" estimate of background error correlation length scales for wind. Results are computed from 6 member EDA perturbations averaged over three weeks (15/2/2010 to 7/3/2010). (Courtesy of L. Berre, Meteo-France).

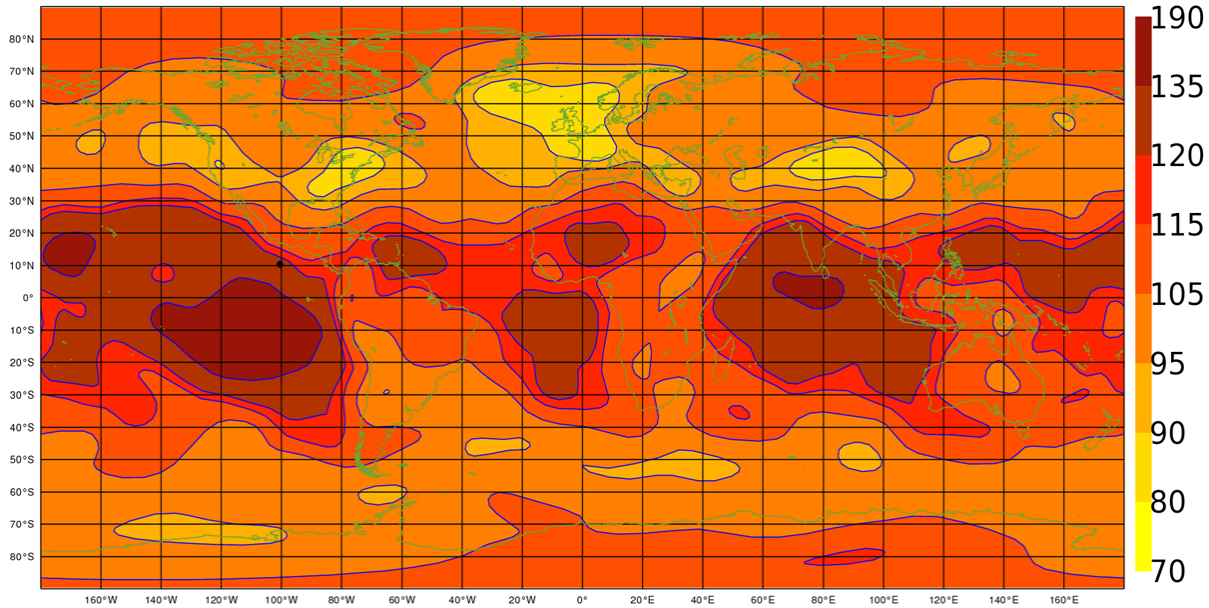

Figure 16: EDA "flow-dependent" estimate of background error correlation length scales for wind. Results are computed from 6 member EDA perturbations averaged over 4 days (24/2/2010 to 27/2/2010). (Courtesy of L. Berre, Meteo-France).

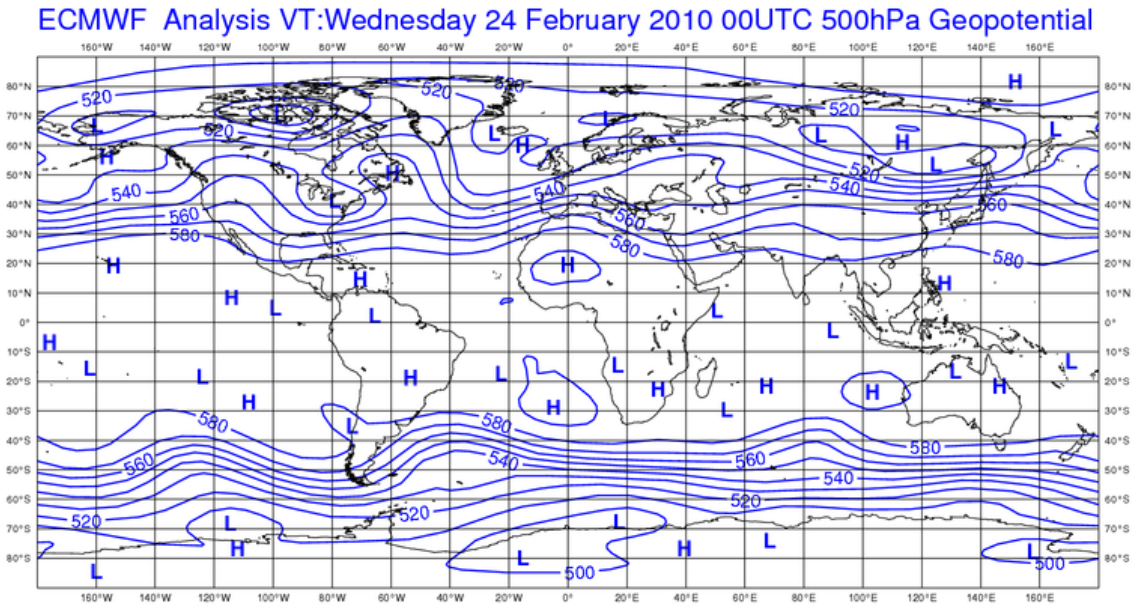

Figure 17: ECMWF analyzed geopotential height at 500 hPa. Time average from 24/2/2010, 00UTC, to 27/2/2010, 00UTC.

4D-Var and the EnKF are two computationally viable approximations to the KF which have received most attention in the atmospheric DA community.

4D-Var solves the KF equations inside the assimilation window, but it is not able to cycle the error estimates from one assimilation update to the next. For this reason it needs to start from a static, climatological estimate of **B** in a simplified, parameterized form.

The EnKF, on the other hand, converges to the KF in the limit of infinitely large ensemble size. However, for practical ensemble sizes (100), the flow-dependent estimates of **B** that can be sampled from the ensemble are severely rank-deficient.

Hybrid data assimilation techniques try to bridge the gap between 4D-Var and the EnKF. From a 4D-Var perspective, hybrid DA tries to incorporate the flow-dependent background error information from an EnKF (or an ensemble of 4D-Vars, such as the EDA) into the full rank **B** estimate used at the start of the assimilation window. Two common approaches to hybrid DA have been described in this paper: the alpha control variable method and the EDA method.

A recurring and crucially important theme in hybrid DA is how to deal with sampling errors in the ensemble **B** due to the relatively small size of computationally affordable ensembles. Common approaches are a) to regularize the ensemble estimate of **B** by combining it with a climatological estimate and b) to apply filtering to the ensemble **B** in grid point and/or in spectral space.

From a broader perspective it seems clear that the use of hybrid DA techniques is becoming more widespread. Hybrid DA techniques are attractive for the theoretical reasons discussed above, but also for more practical considerations. In the first place, it is usually conceptually straightforward to adapt a pre-existing variational system to make use of online **B** estimates from an ensemble; secondly, and possibly more importantly, the distributed architecture of current and near future high performance computers favors running multiple, independent instances of analysis and forecast models. This makes ensemble-based analysis and forecast systems the tool of choice to adequately exploit the future increased computational resources (see also Isaksen, 2012, in these proceedings).

REFERENCES

Anderson, J. L., 2006: Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D*, **230**, 99-111.

Barker, D. M., 1998: Var scientific development paper 25: The use of synoptically-dependent error structures in 3DVAR. UK MET Office Tech. Rep., 2 pp.

Belo Pereira, M. and L. Berre, 2006: The use of an ensemble approach to study the background-error covariances in a global NWP model. *Mon. Wea. Rev.*, **134**, 2466–2489.

Bishop CH., Hodyss D. 2009. Ensemble covariances adaptively localized with ECORAP. Part 2: a strategy for the atmosphere. Tellus, **61A**: 97-111.

Bonavita M., L. Raynaud and L. Isaksen, 2010: Estimating background-error variances with the ECMWF Ensemble of Data Assimilations system: the effect of ensemble size and day-to-day variability. *Q. J. R. Meteorol. Soc., 137: 423–434*.

Bonavita M., L. Isaksen and E. Holm, 2012: On the use of EDA background error variances in the ECMWF 4D-Var. *Q. J. R. Meteorol. Soc.*. Early Online Release, doi: 10.1002/qj.1899

Buehner, Mark, P.L. Houtekamer, C. Charette, H. L. Mitchell and Bin He, 2010a: Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part I: Description and Single-Observation Experiments, *Mon. Wea. Rev.,* **138**, 1550-1566.

Buehner, Mark, P.L. Houtekamer, C. Charette, H. L. Mitchell and Bin He, 2010b: Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part II: One-Month Experiments with Real Observations, *Mon. Wea. Rev.,* **138**, 1567-1586.

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

Campbell, W.F., Bishop, C.H. and D. Hodyss, 2010: Vertical Covariance Localization for Satellite Radiances in Ensemble Kalman Filters, *Mon Wea. Rev.* **138**: 282–290.

Cosme, E., J. Verron, P. Brasseur, J. Blum and D. Auroux, 2012: Smoothing Problems in a Bayesian Framework and Their Linear Gaussian Solutions. *Mon. Wea. Rev.*, **140**, 683–695.

Derber, J.C. and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. Tellus, **51A**, 195-221

Evensen G. 1994. Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys Res* 99(C5): 10 143–10 162

Evensen, G., 2003: The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, **53**, 343-367.

Fisher, M., 2003: Background error covariance modeling. *Proceedings of the ECMWF Seminar on recent developments in data assimilation for atmosphere and ocean*, ECMWF, pages 45–63. (Available from: http://www.ecmwf.int/publications/)

Fisher, M. and Courtier, P., 1995: Estimating the covariance matrices of analysis and forecast error in variational data assimilation. *ECMWF Technical Memorandum* 220.
(Available from: http://www.ecmwf.int/publications/)

Fisher, M., M. Leutbecher and G. A. Kelly, 2005: On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation, Q. J. R. Meteorol. Soc., 131, 3235-3246.

Gandin, L.S., 1963: Objective analysis of meteorological fields, in: Gidrometeorologicheskoe Izdatel'stvo (GIMIZ) (Israel Program for Scientific Translations, Jerusalem, Trans.) Leningrad.

Hamill, T. M., C. Snyder, 2000: A Hybrid Ensemble Kalman Filter–3D Variational Analysis Scheme. *Mon. Wea. Rev.*, **128**, 2905–2919.

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

Isaksen, L., M. Bonavita, R. Buizza, M. Fisher, J. Haseler, M. Leutbecher and Laure Raynaud, 2010: Ensemble of data assimilations at ECMWF. *ECMWF Technical Memorandum* No. 636. (Available from: http://www.ecmwf.int/publications/)

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

Kolczynski W. C., Jr., D. R. Stauffer, S. E. Haupt, A. Deng, 2009: Ensemble Variance Calibration for Representing Meteorological Uncertainty for Atmospheric Transport and Dispersion Modeling. *J. Appl. Meteor. Climatol*., **48**, 2001-2021.

Kolczynski W. C., Jr., D. R. Stauffer, S. E. Haupt, N. S. Altman, A. Deng, 2011: Investigation of Ensemble Variance as a Measure of True Forecast Variance. *Mon. Weather Rev*., Early Online Release, doi: 10.1175/MWR-D-10-05081.1

Leutbecher, M., 2010: Diagnosis of ensemble forecasting systems. *Proceedings of the ECMWF Seminar on Diagnosis of Forecasting and Data Assimilation Systems,* 7-10 September 2009, ECMWF, pages 235-266. (Available from: http://www.ecmwf.int/publications/)

Lorenc, A.C.,1986: Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **112**, 1177–1194.

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.

Mitchell, H. L., P.L. Houtekamer and Gérard Pellerin, 2002: Ensemble Size, Balance, and Model-Error Representation in an Ensemble Kalman Filter. *Mon. Wea. Rev*., **130**, 2791-2808.

Miyoshi, T., Y. Sato, T. Kadowaki, 2010: Ensemble Kalman Filter and 4D-Var Intercomparison with the Japanese Operational Global Analysis and Prediction System. *Mon. Wea. Rev*., **138**, 2846–2866.

Ott E, Hunt BR, Szunyogh I, Zimin AV, Kostelich EJ, Corazza M., Kalnay E, Patil DJ, Yorke JA. 2004. A local ensemble Kalman Filter for atmospheric data assimilation. *Tellus*, **56A**: 415-428.

Pannekoucke, O. L. Berre and G. Desroziers, 2007: Filtering properties of wavelets for local background-error correlations. *Q. J. R. Meteorol.Soc*., **133**, 363–379.

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

Raynaud, L., L. Berre, and G. Desroziers, 2008: Spatial averaging of ensemble-based background-error variances. *Q. J. R. Meteorol. Soc*., **134**, 1003–1014.

Thépaut, J.-N., Courtier, P., Belaud, G. and Lemaıtre, G., 1996: Dynamical structure functions in a four-dimensional variational assimilation: A case study. *Q. J. R. Meteorol. Soc.*, **122**,535–561

Trémolet Y. 2006. Accounting for an imperfect model in 4D-Var. *Q.J. R. Meteorol. Soc*. **132**, 2483–2504.

Trémolet Y. 2007. Model-error estimation in 4D-Var. *Q.J. R. Meteorol. Soc.* **133**, 1267–1280.

Varella H, Berre L, Desroziers G. 2011. Diagnostic and impact studies of a wavelet formulation of background-error correlations in a global model. *Q. J. R. Meteorol. Soc*., **137**, 1369–1379.

Wikle, C.K., and L. M. Berliner, 2007: A Bayesian tutorial for data assimilation, *Physica D*, **230**, 1-16.