TECHNICAL MEMORANDUM

636

Ensemble of Data Assimilations at ECMWF

L. Isaksen, M. Bonavita, R. Buizza, M. Fisher, J. Haseler, M. Leutbecher and Laure Raynaud[1]

Research Department

(1) Météo-France

December 2010

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications/

Contact: library@ecmwf.int

European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, Berkshire RG2 9AX, England

Abstract

The Ensemble of Data Assimilations (EDA) implemented in June 2010 is an important new component of ECMWF's operational suite. It provides valuable estimates of analysis and short-range forecast uncertainty. This paper describes the EDA method and its applications. The rationale for developing an EDA is discussed and the considerations that went into deciding the operational configuration, including resolution, ensemble size and model error formulation, are presented. Since its implementation, EDA based perturbations have been used to improve the initial perturbations for the Ensemble Prediction System (EPS): the methodology used to combine EDA-based perturbations and initial-time singular vectors and the impact of replacing the evolved singular vectors with EDA-based perturbations on the EPS performance is presented. Another important application of the EDA is to provide flow-dependent background error estimates for the deterministic 4D-Var system: the method under development and testing is described, and the significant progress that has been made in the past year is discussed.

1 Introduction

An Ensemble of Data Assimilations (EDA) system was introduced at ECMWF with cycle 36r2 of the Integrated Forecasting System (IFS) on 22 June 2010. The EDA consists of an ensemble of ten independent lower-resolution 4D-Var assimilations that differ by perturbing observations, sea surface temperature fields and model physics.

The main justification for implementing the EDA is that it quantifies analysis uncertainty.

- It is the first system implemented at ECMWF to provide estimates of analysis uncertainty that take model dynamics into account. A properly designed EDA will complement the data assimilation system with important information about the quality of the deterministic analysis.
- It can be used to estimate flow-dependent background errors in the deterministic 4D-Var assimilation system; this will potentially improve the medium-range forecast.
- It can improve the representation of initial uncertainties for the Ensemble Prediction System (EPS). With the introduction of EDA in IFS cycle 36r2, EDA-based perturbations replaced evolved singular vectors (SVs) to generate the EPS initial conditions. This change improved the EPS skill, especially in the tropics.

Other operational NWP centres are also exploring or have already implemented EDA-like methods. In appendix A these activities at Météo-France, UK MetOffice and Environment Canada are described.

If the perturbations of observations and model physics are correctly specified, the EDA will provide good estimates of analysis uncertainty. The theoretical explanation follows in Chapter 2.

The experimentation that led to the operational EDA design decisions is discussed in Chapter 3. Because four-dimensional data assimilation, similar to that used for the operational analysis, is an integral part of EDA, it has the potential to provide very valuable information about analysis uncertainty of the operational assimilation system. This is difficult to obtain by other means. The system will also provide short-range forecast error uncertainty. Many applications would benefit from accurate estimates of the uncertainty in analysis and short-range forecast errors. This could, e.g., provide guidance for the quality of ECMWF's short-range forecasts.

Although the EDA is based on perturbed lower-resolution versions of the operational analysis system (with outer loops at spectral triangular truncation T399 instead of T1279), its computing cost is significant, similar to running the deterministic analysis. Chapter 4 presents the operational implementation and address computational cost aspects.

The simulation of initial uncertainties is a crucial aspect of ensemble prediction. The replacement in the EPS of the evolved singular vectors with EDA-based perturbations has lead to improvements in the EPS spread/skill relationship (Buizza *et al.* 2008). This is described in detail in Chapter 5.

In data assimilation, one of the crucial aspects is the estimation of the background error variances. To a large degree these are static in the operational 4D-Var system. This is unrealistic, especially for severe weather events where the background error variances can be underestimated significantly. The EDA is able to produce flow-dependent estimates of analysis uncertainty and background error uncertainty based on the ensemble spread, measured as the standard deviation of the independent EDA members (Fisher 2003; Tan *et al.* 2007). This information gives an estimate of the error-of-the-day and can also be used for the estimation of seasonally-varying background errors. Recent research at ECMWF has shown a beneficial impact from using the EDA variances of the day in the operational high-resolution 4D-Var. This is described in more detail in Chapter 6. It is expected to be implemented operationally in the first half of 2011.

2 Theoretical Justification of the Perturbed-Observation EDA

The perturbed-observation ensemble Kalman filter is well understood (e.g. Burgers *et al.* 1998). In this context, it is clear that ensemble members must be presented with observations that have been perturbed by the addition of random noise drawn from the probability density function (p.d.f.) of observation error, and that further perturbations must be added during the forecasts to account for model error. The justification for using such perturbations in the EDA is perhaps less clear, since the EDA does not purport to use optimal covariance matrices. In this Chapter, we present an analysis of the perturbations used in the EDA, and show that if the perturbations are drawn from the true distributions of observation and model error, then the spread of the EDA about the control (unperturbed) analysis will be representative of analysis error. Note that no perturbations are applied to the background state directly. Doing so would result in double counting of model error perturbations. It is better to use stochastic model error representation during the time integration to represent model uncertainties.

A theoretical justification of the perturbed-observation EDA is presented in Appendix A of Žagar *et al*. (2005). The authors consider a very general analysis-forecast system:

$$\begin{aligned} x_a^k &= f_k(x_b^k, y^k) \\ x_b^{k+1} &= M_k(x_a^k) \end{aligned} \tag{2.1}$$

where k denotes analysis cycle, y^k is the vector of observations, x_a^k represents the analysed state and x_b^k is the background.

Subject to the assumption that the analysis f_k and the forecast model M_k are weakly nonlinear functions, the authors demonstrate that perturbed analyses, with perturbations drawn from the p.d.f. of analysis error, may be produced by the following perturbed system (denoted by ~):

$$\begin{aligned} \tilde{x}_a^k &= f_k(\tilde{x}_b^k, y^k + \eta^k) + \omega^k \\ \tilde{x}_b^{k+1} &= M_k(\tilde{x}_a^k) + \xi^k \end{aligned} \tag{2.2}$$

Here, η^k is a perturbation drawn from the p.d.f. of observation error, and ξ^k is a perturbation drawn from the p.d.f. of model error. (The perturbation ω^k accounts for the possibility that the analysis may itself be a source of error. That is, the analysis may be in error even if presented with the true background state and the true values of the observed quantities. This error does not appear in linear analysis theory, and will be ignored in the discussion below.)

Although, as shown by Žagar *et al*. (*op. cit*.), the EDA method is valid for weakly nonlinear systems, it is instructive to consider the linear case. Let us consider the following system:

$$\begin{aligned} x_a^k &= x_b^k + K_k(y^k - H_k x_b^k) \\ x_b^{k+1} &= M_k x_a^k \end{aligned} \tag{2.3}$$

where K_k and M_k are matrices. We will make no assumption about the optimality of the analysis, so that K_k here denotes a general gain matrix, and not specifically the Kalman gain.

It can be shown (subject to certain common assumptions about the independence of background, observation and model error) that the covariance matrices for this system evolve according to:

$$\begin{aligned} P_k^a &= (I - K_k H_k) P_k^b (I - K_k H_k)^T + K_k R_k K_k^T \\ P_{k+1}^b &= M_k P_k^a M_k^T + Q_k \end{aligned} \tag{2.4}$$

where R_k is the covariance matrix of observation error, and Q_k is the covariance matrix of model error.

Consider now the following perturbed system:

$$\begin{aligned} \tilde{x}_a^k &= \tilde{x}_b^k + K_k(y^k + \eta^k - H_k \tilde{x}_b^k) \\ \tilde{x}_b^{k+1} &= M_k \tilde{x}_a^k + \xi^k \end{aligned} \tag{2.5}$$

Let us denote the differences between the perturbed and unperturbed quantities as follows:

$$\begin{aligned}\varepsilon_a^k &= \tilde{x}_a^k - x_a^k \\ \varepsilon_b^k &= \tilde{x}_b^k - x_b^k\end{aligned} \tag{2.6}$$

Then, subtracting the perturbed and unperturbed equations (2.5 and 2.3) we find that the perturbations satisfy the following:

$$\begin{aligned}\varepsilon_a^k &= \varepsilon_b^k + K_k(\eta^k - H_k\varepsilon_b^k) \\ \varepsilon_b^{k+1} &= M_k\varepsilon_a^k + \zeta^k\end{aligned} \tag{2.7}$$

Suppose that the applied perturbations η^k and ξ^k have covariance matrices R_k and Q_k, respectively. Forming the covariance matrices (and with the same assumptions of statistical independence that were used to derive equation 2.4) we find that:

$$\begin{aligned}\overline{\varepsilon_a^k\left(\varepsilon_a^k\right)^T} &= \left(I - K_kH_k\right)\overline{\varepsilon_b^k\left(\varepsilon_b^k\right)^T}\left(I - K_kH_k\right)^T + K_kR_kK_k^T \\ \overline{\varepsilon_b^{(k+1)}\left(\varepsilon_b^{(k+1)}\right)^T} &= M_k\overline{\varepsilon_a^k\left(\varepsilon_a^k\right)^T}M_k^T + Q_k\end{aligned} \tag{2.8}$$

By comparing equations 2.4 and 2.8 we see at once that if $\overline{\varepsilon_b^k\left(\varepsilon_b^k\right)^T} = P_k^b$ for some k, then $\overline{\varepsilon_a^m\left(\varepsilon_a^m\right)^T} = P_m^a$ and $\overline{\varepsilon_b^m\left(\varepsilon_b^m\right)^T} = P_m^b$ for all $m \geq k$. That is, the perturbations of the analyses and backgrounds have covariances equal to the covariances of the corresponding analysis and background error for all subsequent analysis cycles.

Typically, the initial perturbation ε_b^0 provided to an EDA is not drawn from the p.d.f. of background error. In fact, it is common to "cold start" an EDA from identical backgrounds (i.e. from $\varepsilon_b^0 = 0$). We now show that for any initial perturbation, the covariance matrix of ε_b^k converges towards P_k^b (subject to an observability criterion). We consider two perturbed systems for which the observation and model error perturbations are identical, but whose initial perturbation ε_b^0 is different. That is, we consider one system to evolve according to equation 2.5 and the second to evolve according to:

$$\begin{aligned}\hat{x}_a^k &= \hat{x}_b^k + K_k\left(y^k + \eta^k - H_k\hat{x}_b^k\right) \\ \hat{x}_b^{k+1} &= M_k\hat{x}_a^k + \xi^k\end{aligned} \tag{2.9}$$

Subtracting the equations for the two systems gives:

$$\begin{aligned}\delta_a^k &= \left(I - K_kH_k\right)\delta_b^k \\ \delta_b^{k+1} &= M_k\delta_a^k\end{aligned} \tag{2.10}$$

Where $\delta_a^k = \hat{x}_a^k - \tilde{x}_a^k$ and $\delta_b^k = \hat{x}_b^k - \tilde{x}_b^k$.

Next, consider two unperturbed systems, governed by equation 2.3, but with different initial states x_b^k. We see that the equation for the difference between the unperturbed systems is identical to the equation 2.10 for the difference between the identically-perturbed systems. Hence, we have the following:

> *If* unperturbed analysis-forecast systems started from different initial states converge towards the same solution as $k \rightarrow \infty$, *then* the same is true of identically-perturbed systems.

The convergence of analyses started from different initial states towards the same solution has been demonstrated in practice by Fisher *et al*. (2005). It is a consequence of the effects of model error and unstable dynamics, which limit the information brought to the analysis by old observations.

Now suppose that one of the perturbed systems was started with an initial perturbation whose covariance matrix is P_k^b. We have already shown that, given the correct observation and model error perturbations, the background and analysis states for this system have correct perturbations for every subsequent step (that is $\overline{\varepsilon_a^m \left(\varepsilon_a^m\right)^T} = P_m^a$ and $\overline{\varepsilon_b^m \left(\varepsilon_b^m\right)^T} = P_m^b$ for all m). Since identically-perturbed analyses converge, it follows that for any other initial perturbation, and subject to the observability criterion above, the perturbed analysis must converge towards the system whose background and analysis perturbations are drawn from the distributions of background and analysis error.

Fisher *et al*. (*op. cit*.) show the asymptotic convergence of an EDA "cold-started" from identical backgrounds. At least for 500 hPa geopotential height, the global spread of the ensemble reaches an asymptotic value after roughly one week. The analysis above suggests that, provided the observations and model states (and other inputs, such as sea surface temperature) are perturbed according to their true error covariances, the spread of the ensemble after the first week will correctly represent the variance of analysis and background error.

3 Experimentation leading up to operational implementation

It is important to stress that the EDA is designed to serve both the needs of applications in the data assimilation system and the needs of the EPS. This has major implications for the design of the operational Ensemble of Data Assimilations.

To be able to use the system to calculate background error statistics and estimate flow-dependent background errors, the EDA must be designed to be sufficiently similar to the operational 4D-Var assimilation system. So it requires that the top level of the EDA model is equal to or above the top level of the operational 4D-Var system. Since physical parameterizations are particularly sensitive to the number and positioning of the vertical levels, it is beneficial that the same levels are used in the EDA and the high-resolution 4D-Var system (91 at present). The use in the EDA of the same vertical levels also avoids costly vertical interpolation steps that might have a negative impact on the estimation of the flow-dependent background error covariances. The horizontal resolution cannot be significantly lower than the operational 4D-Var resolution in order to get a flow-dependent impact on severe weather events, and analysis uncertainty estimates that describe the operational 4D-Var system.

The choice of horizontal resolution will be discussed further in Section 3.1. Also, for reasons of consistency, a 4D-Var configuration in the EDA is preferable to a 3D-Var configuration (Berre *et al*. 2009). It is beneficial to keep the assimilation as close as possible to the operational 4D-Var system, so the gain matrices of the two systems are comparable.

When EDA fields are used to calculate initial uncertainties for the EPS (Buizza *et al*. 2010), vertical interpolation from 91 to 62 levels and horizontal interpolation from T399 to T639 are required. It has been decided not to force the EPS and the EDA resolutions to be the same. This leaves more flexibility in the future system design of the EPS and the EDA. If the resolution of the EDA becomes an issue for the EPS, the situation will be revisited. However, the computer codes have been designed to cater for any resolution change, as this may be required for research experiments or future operational configurations.

Isaksen *et al*. (2007) describe the design of the EDA in detail, based on analyses run with a T255 outer loop and T95/T159 inner loops and 91 vertical levels. In the EDA, for each observation, perturbations are defined by randomly sampling a Gaussian distribution with zero mean and standard deviation equal to the estimate of the observation error standard deviation. For cloud-track wind observations, perturbations are horizontally correlated, based on Bormann *et al*. 2003. Sea surface temperature fields are also perturbed, with correlated patterns as currently used in the Seasonal Forecasting System (Vialard *et al*. 2005). At the very first assimilation cycle, the randomly-perturbed observations are the only source of difference between the perturbed analyses, while for the subsequent cycles differences will evolve in the first-guess fields and contribute to the spread of the analyses, as described in Chapter 2. Model error is simulated by stochastically perturbing the model tendencies using the same method applied in the EPS – the method used is described later in Section 3.2 and Chapter 5.

Isaksen *et al*. (op. cit.) demonstrated the EDA's ability to produce flow-dependent spread and deliver promising results for some severe meteorological events. The EDA is also able to produce a realistic horizontal distribution of average analysis error and background error, with small values over areas well constrained by observations, like the USA, Europe and Australia.

To test the robustness of the EDA method an ensemble has been run using only SYNOP data over land and radiosonde/pilot profile measurements, similar to an early pre-satellite scenario. This was compared against a full 2006 observing system scenario. Fig. 1 shows the snapshot of zonal wind ensemble spread for model level 49 (close to 200 hPa) for the two EDA configurations on a day in September 2006. It is evident that the EDA spread, that should represent background error uncertainty, is significantly higher for the Southern Hemisphere extratropics in the reduced observing system scenario. For the Southern Hemisphere the spread is reduced over Australia and South America and downstream of these well observed regions. For the Northern Hemisphere extratropics the spread is increased over the oceans, but only slightly increased over Japan, North America, Europe and downstream Russia. For the tropics the spread is also increased, especially over the oceans. All in all the response gives confidence in the EDA method. This confirms earlier results by Tan *et al*. (2007).

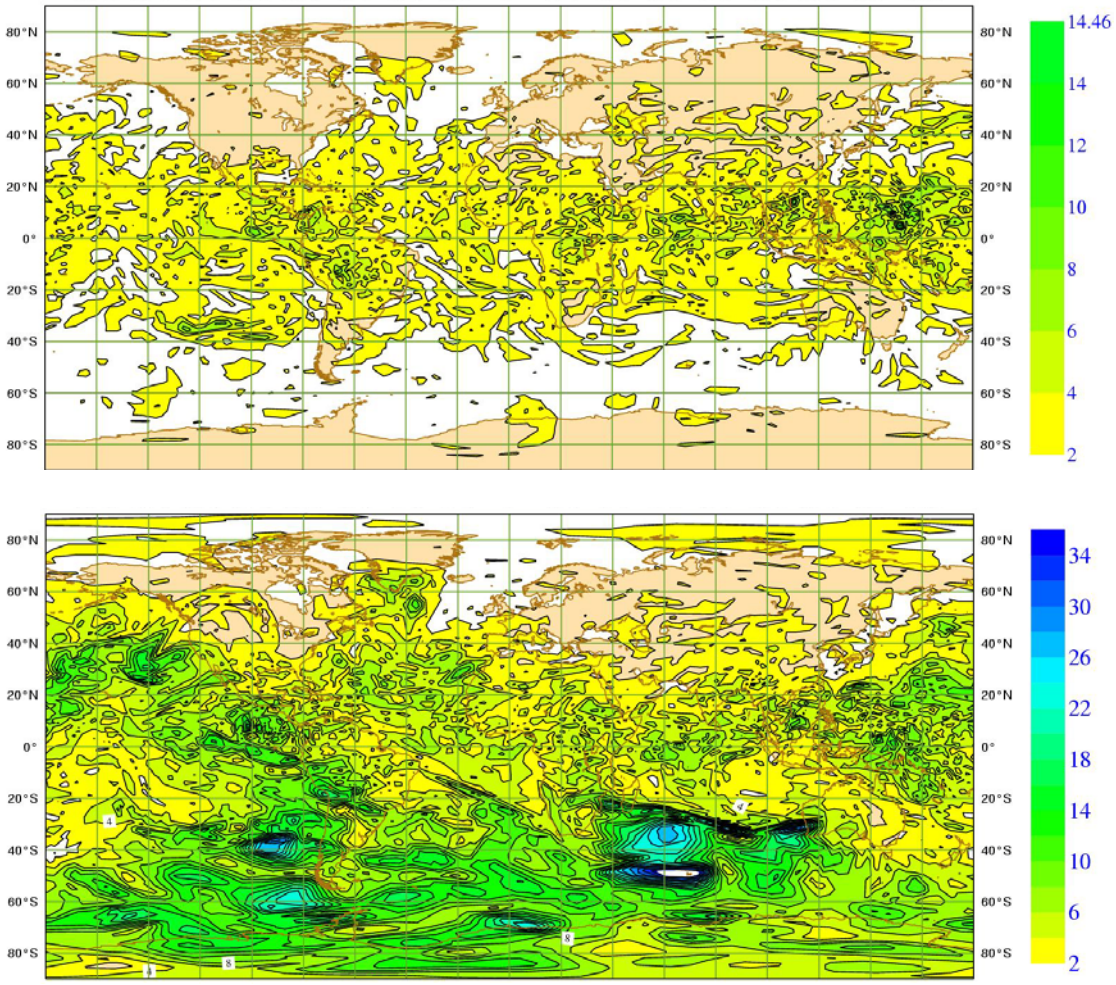

Figure 1: Standard deviation of 3h forecast ensemble spread for zonal wind (unit m/s) at model level 49 (near 200 hPa) valid on 2100 UTC 30 September 2006. Top panel: From an EDA that use all observations. Bottom panel: From an EDA that only use radiosondes, pilots and SYNOP observations over land.

3.1 Horizontal resolution in the EDA 4D-Var inner and outer loops

The impact of EDA resolution and ensemble size has been investigated. Fig. 2 shows examples of EDA based spread for 850 hPa zonal wind for a region near Japan that is dynamically very active. Only the inner loop resolution impact is considered in this case study. All examples use a T255 outer loop resolution. For the low-resolution assimilation system with one T95 inner loop (Fig. 2c), there is a clear flow dependence, but this is not the case when the operational "randomization" technique (Fisher and Courtier, 1995; Fisher, 2003) is used (Fig. 2a). The background error estimate from the "randomization" technique are computed from random samples of the static B matrix, used at the start of the 4D-Var analysis with a small flow-dependent component, which results from the application of the non-linear balance equation and the omega equation. Also, the low-resolution assimilation system (Fig. 2c) delivers less focussed, but still similar results compared to a higher-resolution system with two (T95 and T159) inner loops (Fig. 2b).

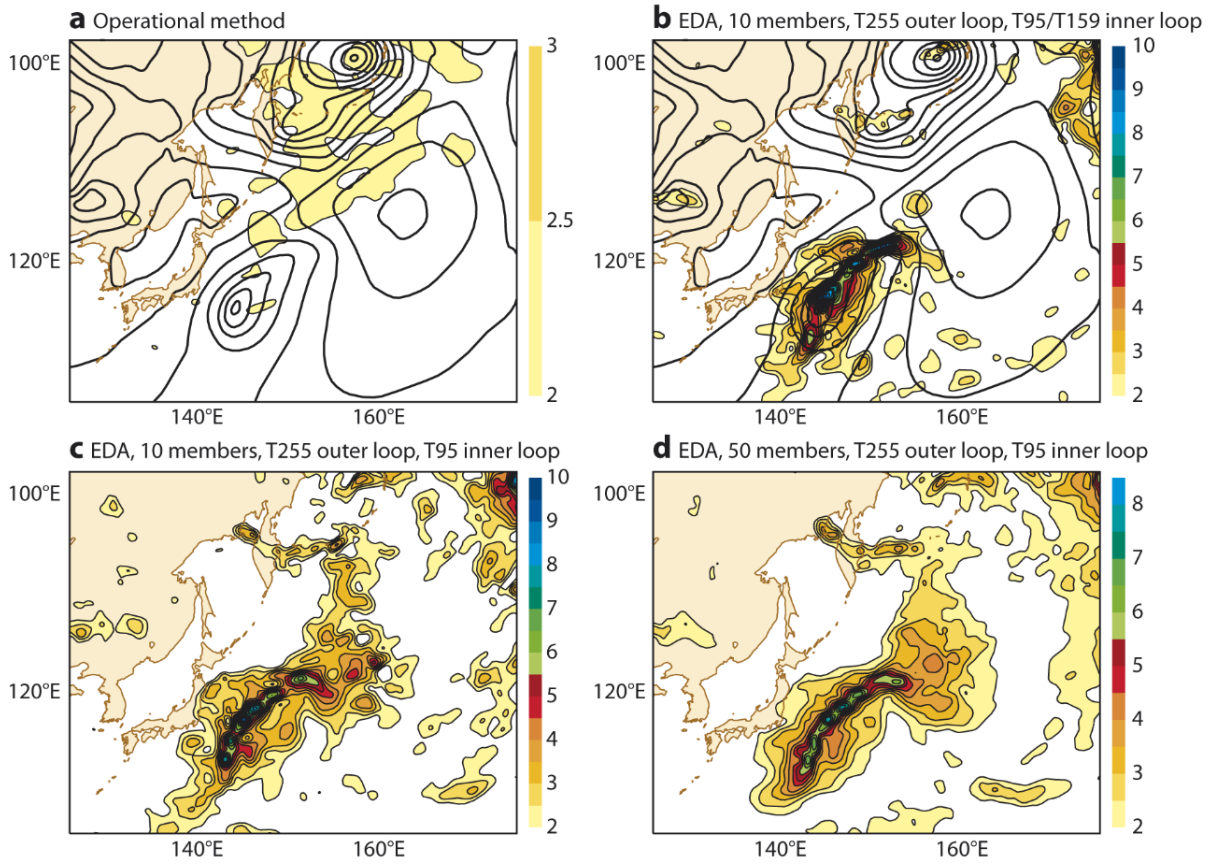

Figure 2: Estimates of 850 hPa zonal wind background error standard deviation valid at 0000 UTC on 16 October 2006 for a baroclinic area near Japan. (a) The operational cycling randomization method (max. 2.8 ms^{-1}). (b) The 10 member T255 outer loop with T95/T159 inner loop (global max. 13.8 ms^{-1}). (c) The 10 member T255 outer loop with T95 inner loop (global max. 13.9 ms^{-1}). (d) The 50 member T255 outer loop with T95 inner loop (global max. 12.0 ms^{-1}). The standard deviation in panels (b), (c) and (d) has been multiplied by a factor of 2 in order to make the average global EDA spread closer to innovation statistics.

Comparing Figures 2c and 2d, one can see the impact of increasing the number of ensemble members from 10 to 50. The patterns are very similar, but the standard deviation computed with the 50 member ensemble is smoother with fewer spikes and also avoids regions with very low variance (not shown). The 50 members are not only giving a statistically better sampling of the forecast errors, but they also describe flow-dependent features at a higher resolution. The results shown in Fig. 2 suggest that computer resources may be better spent on more members with a simpler inner-loop low-resolution version of the 4D-Var system (see also the discussion on ensemble size in Chapter 6). The benefit of a larger lower resolution ensemble also appears in global statistics, not only for severe weather events. But more research will have to be performed before the final conclusions can be drawn on this subject.

The capturing of flow dependence for all three ensemble-based versions shown in Fig. 2 is clearly visible. The smoothing property of using 50 members is also detectable, although the amplitudes and structures are surprisingly similar for the low- and higher-resolution analysis system. This may well be due to the fact that all systems for these experiments used a T255 outer loop resolution. It is at this

stage and resolution that the observations are perturbed. The uncertainty information is also propagated in time with the same T255 model resolution.

Variations of forecast uncertainty described by the EDA can be particularly pronounced for severe weather events. As an example, consider the Category-3 Hurricane Emily on 20 July 2005 just before it made landfall in Mexico. Fig. 3a shows the precipitation measured by the local weather radar. The EDA spread (i.e., the standard deviation of the ten T399 outer-loop members) for zonal wind at approximately 850 hPa is given in Fig. 3b. Typical standard deviations in the region would be 1-3 ms^{-1}, but the flow-dependent estimates from the EDA yield standard deviations up to 7 ms^{-1}. Note that the region of enhanced spread is concentrated in the vicinity of Hurricane Emily, identified by the mean sea level pressure contours. Fig. 3c is similar to Fig. 3b, but for the operational "Randomization method". It is clear this method lacks flow-dependence. Finally Fig. 3d shows the impact of using a very high resolution EDA (T95/T255 inner-loop and T799 outer-loop/forecasts). The higher resolution enables significantly higher (14 ms^{-1}) and more focussed background error estimates.

The ability of the EDA to identify regions of large background error associated with severe weather events has the potential to significantly improve quality control decisions and give higher weight to observations used in the analysis from such regions. This will be discussed in Chapter 6.

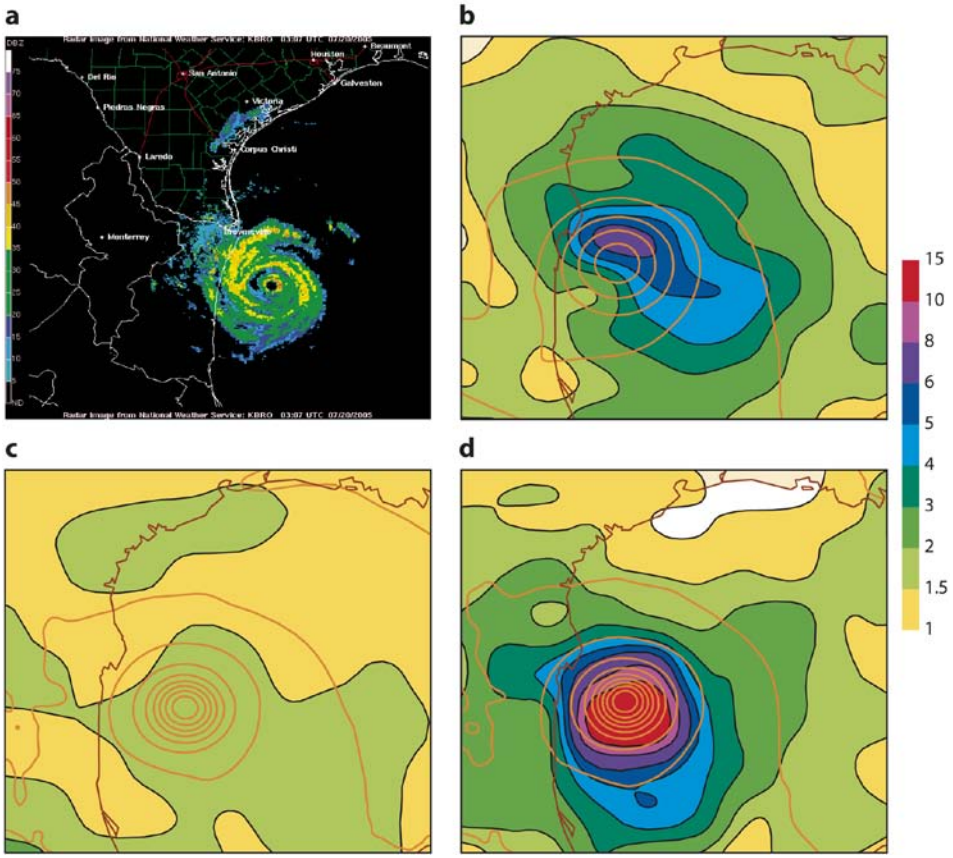

Figure 3: Atlantic hurricane Emily on 20 July 2005 near the coast of Mexico. (a) Radar image from National Weather Service. (b) Ensemble spread at t=6h, for zonal wind at 850 hPa near Hurricane Emily with T399 outer-loop and T95/T159 inner-loop; (c) Similar to (b), but using operational randomization method; (d) Similar to (b), but with T799 outer-loop and T95/T255 inner-loop. The mean sea level pressure contours for the respective analyses are overlaid (5 hPa intervals).

The performance and computing costs of a number of potentially suitable low-resolution configurations were investigated to choose an assimilation system that was close in skill to the operational system, but still significantly cheaper. The investigations described above, performed at T255, showed that a 10 member EDA gave similar ensemble spread patterns to a 50 member EDA, though with more noise. Because noise filtering methods are available it was decided to choose 10 members for the initial operational implementation.

The comparison of the performance of short-rage forecasts run from various assimilation systems that were candidates for the EDA configuration, with different combinations of inner/outer loop resolutions also provided objective evidence that the operational EDA choice of T399 outer-loop resolution and T95/T159 inner-loop resolution was a good compromise. Fig. 4 shows average root-mean-square-error of forecast scores during the first 48 hours calculated over 92 days. The left panel shows that it is very beneficial to increase the outer loop resolution from T255 to T399. In general it was found that an increase in outer-loop resolution from T255 to T399 improves the forecast scores significantly. This is especially true for Europe and the Northern Hemisphere (not shown). Increasing the inner loop resolution from T95 to T159 is also beneficial, but to a smaller extend. This gives the biggest impact for wind scores (see Fig. 4 right panel as an example). For severe weather events the benefit of increased inner loop resolution is even more evident. The investigations showed that a T95 final inner loop is not capable of representing tropical cyclones and other extreme events accurately. On the other hand, the use of a T255 inner loop added significant extra computing costs without a significant impact on EDA variance estimates. Reducing the number of iterations from 70 to 50 in the first 4D-Var minimization resulted in almost identical scores for all regions on the globe (not shown).

It was also seen that a higher-resolution outer loop was the main contributor to increased, more detailed ensemble spread and more accurate analysis uncertainty estimates. The EDA only runs 15-hour forecasts (not 10-day forecasts), so the cost of increasing the outer-loop/forecast resolution from T255 to T399 is relatively small (10-15%).

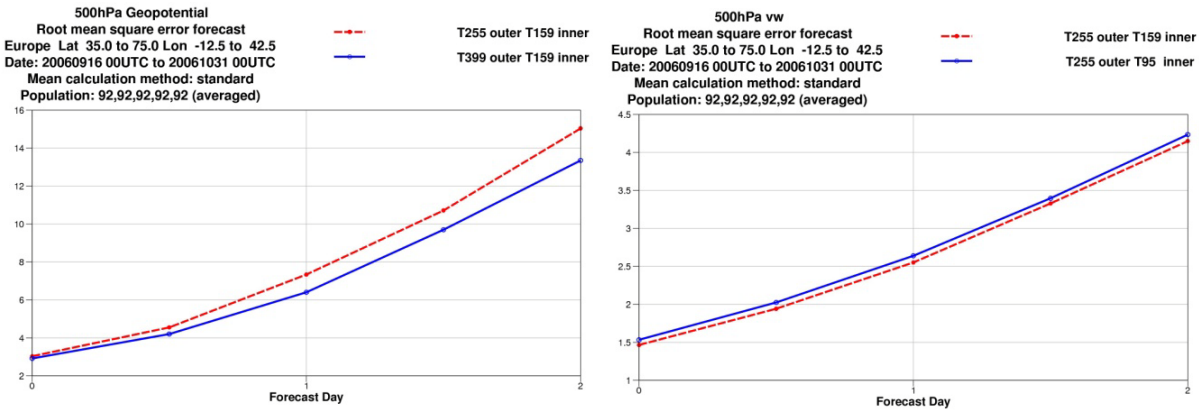

Figure 4: Root mean square error for 500 hPa 0-48h forecasts over Europe averaged over 92 days. Left panel: Geopotential height scores for T399 versus T255 outer loop resolution. Right panel: Vector wind scores for T95 versus T155 inner loop resolution, with the same T255 outer loop resolution.

3.2 Representation of model error in the EDA

In the EDA it is important to represent model error to account for the fact that the forecast model is not perfect. To simulate the impact of model uncertainty, the stochastically perturbed parametrization tendency (SPPT) scheme is used; this perturbs the total parametrized tendency of physical processes. Positive results have also been obtained with the stochastic backscatter (SPBS) scheme that is based on the idea of backscatter of kinetic energy from unresolved scales (see Palmer *et al.* 2009 for a recent review of ECMWF work on stochastic parametrization schemes). To date, however, the use of the SPPT scheme alone gave the best performance (work is in progress to assess the impact of also introducing a backscatter scheme, as discussed in Chapter 5).

On average, if the EDA spread is measured in terms of the 700 (850) hPa temperature standard deviation, the SPPT scheme increases the global average by 19% (23%), and if the EDA spread is measured in terms of the 700 (850) hPa vector wind, the standard deviation is increased by 33% (39%).

It is interesting to examine the geographical distribution of model error impact on short-range forecast uncertainty implied by the stochastic methods. Fig. 5 shows the zonally averaged EDA spread for temperature (Fig. 5a) and zonal wind (Fig. 5b) valid on 14 October 2008, when SPPT is used. The impact of SPPT on the spread can be assessed by calculating the ratio of EDA spread from an experiment with SPPT applied compared to EDA spread from an experiment without model error parametrization. It is clear that the increase in spread due to SPPT is significant for both temperature (Fig. 5c) and zonal wind (Fig. 5d) throughout the atmosphere. The largest SPPT impact for temperature is at the top of the planetary boundary layer, especially in the strato-cumulus regions.

For wind the largest impact is near 700 hPa in the tropics. The SPBS directly perturbs only the wind field, and the increased spread in temperature (Fig. 5e) is small.

The increased wind spread for the SPBS scheme (Fig. 5f) is mainly located in the planetary boundary layer. It is clear that the SPPT scheme provides more widespread perturbations than the SPBS. The larger level of spread looks reasonable, but this needs to be evaluated more systematically. It has been shown that in the EPS the SPPT scheme has a large positive impact on probabilistic skill, while the additional impact of SPBS is smaller (Palmer *et al.*, o*p. cit.)*

Fig. 5 confirms the SPBS and SPPT methods complement each other in the EDA, but the latter is the more effective of the two methods. Further testing using both methods together is planned when revised version of the two schemes have been implemented in the EPS. This is linked to the medium term goal of unifying the model error representation in the EPS and EDA.

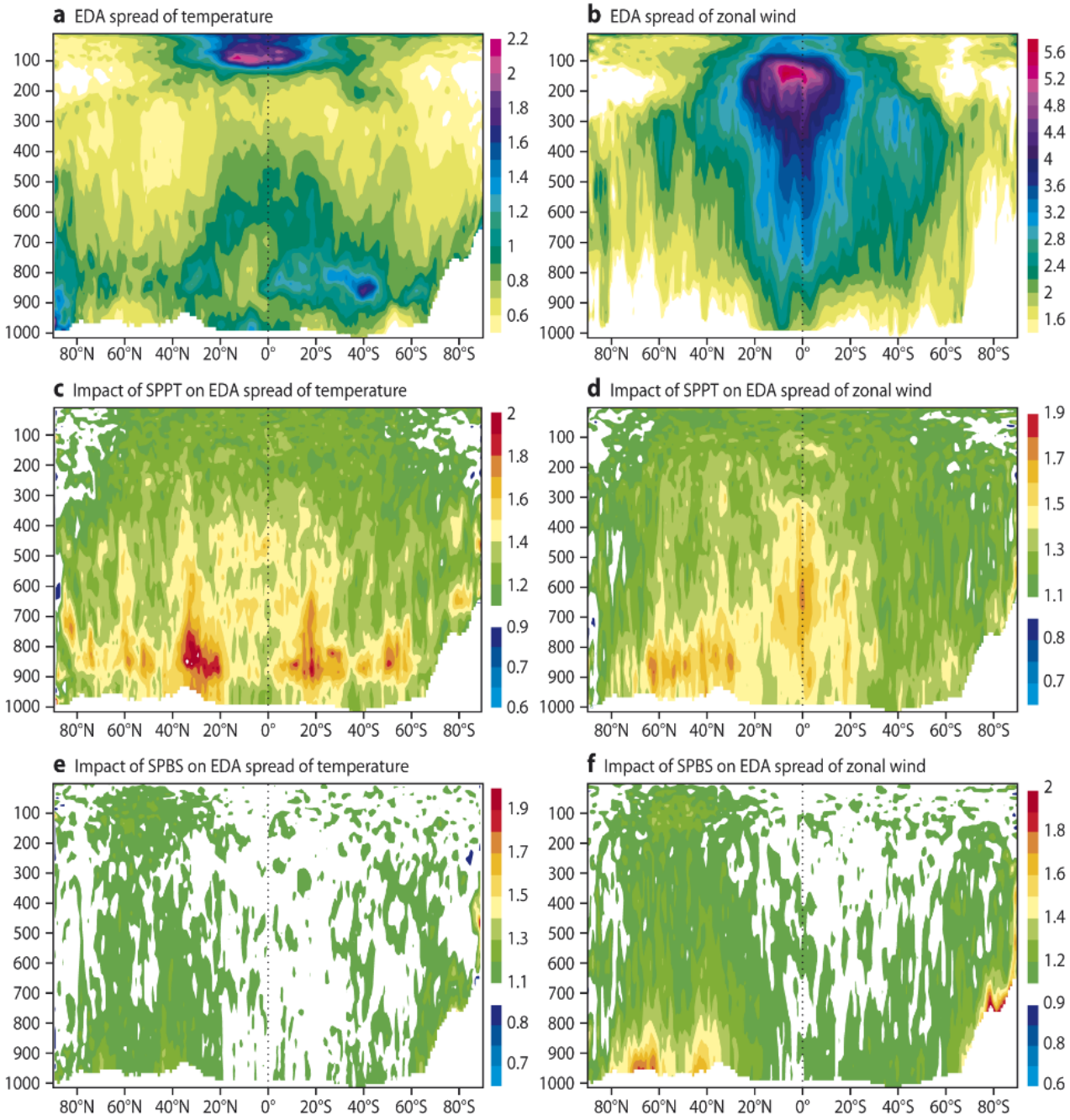

Figure 5: Impact of model error representation on EDA spread. Zonal mean values for temperature (a) and zonal wind (b) on 14 October 2008 using the operational SPPT configuration. The ratio of EDA spread for SPPT applied compared to EDA spread without model error parametrization for temperature (c) and zonal wind (d). The same ratios for temperature (e) and zonal wind (f) when SPBS is used. As expected, the ratios are almost always greater than one.

4 The Operational implementation of the Ensemble of Data Assimilations

Based on the investigations described in chapter 3, it was decided to use the following configuration for the operational EDA introduced 22 June 2010, with cycle 36r2:

- The EDA is run at T399L91 resolution (i.e. spectral triangular truncation T399 with linear grid and 91 hybrid vertical levels of the ECMWF model) with a control (unperturbed) analysis and 10 perturbed analyses and models.
- The analyses are 12-hour 4D-Var, with two minimizations, first at T95, followed by T159 with advanced linearized physics. The first minimization is run with 50 iterations, which is less than the 70 used for the operational deterministic 4D-Var suite.
- The EDA is run twice daily, with the midnight analyses using observations from 2100 UTC to 0900 UTC (incl.) and the midday analyses using observations from 0900 UTC to 2100 UTC (incl.). Only short 15-hour forecasts, required for analysis cycling, are run.
- The observations used are those which have already been extracted for the operational high-resolution delayed cut-off 12-hour 4D-Var analysis, so the EDA analyses can run as soon as these observations become available.
- Unperturbed observations are used for the control analysis, while the other members use observations which have been modified by a random perturbation drawn from a Gaussian distribution with zero mean and standard deviation equal to the observation error estimates used in 4D-Var.
- For Atmospheric Motion Vector (AMV) observations, the perturbations are horizontally correlated, based on error correlations from a study by Bormann *et al.* (2003).
- Sea surface temperature fields are also perturbed with correlated errors using the methodology currently applied in the ECMWF seasonal ensemble forecasting system (Vialard *et al.* 2005).

The data assimilation tasks are run with 48 MPI processes and 8 OpenMP threads on 6 64 processor nodes, while the forecasts are run with 16 MPI processes and 4 threads on 1 node. Fig. 6 shows the computational costs for one ensemble member, and for the operational configuration with one control and 10 ensemble members. The main computational cost of EDA is for the minimization steps (**min 1** and **min 2** in Fig. 6), followed by the observation handling (the major part of **trajectory 1** and **trajectory 3** in Fig. 6). The high resolution related costs of EDA (forecast and approximately three times the cost of **trajectory 2** in Fig. 6) is rather small in the operational configuration. The last two bars in Fig. 6 are included for comparison. The third bar shows the cost of the operational delayed cut-off 12-hour 4D-Var data assimilation. It is run at T1279 resolution, with 3 inner loops of T159, T255 and T255, followed by a 60-hour forecast. Its first minimization is run with 70 iterations. The data assimilation tasks and the forecast are run using 384 MPI processes and 8 OpenMP threads on 48 nodes. The fourth bar shows the cost of running a 10-day forecast at T1279 resolution. It should be

kept in mind that the EDA, due to its inherent parallelism, is much more scalable than the deterministic 4D-Var. This makes the method very suitable for modern and next generation super computers. For more details, see Isaksen *et al.*, 2010.

At the start of each cycle, all the members wait until the observations have been pre-processed and a single Observation Data Base (ODB) has been constructed in the control family. Only the control analysis updates the variational bias correction and the emissivity files. Each member copies the control ODB and takes its first guess from its own forecast from the previous cycle. The members then run an analysis and a short forecast independently, before synchronising with the control again at the start of the next cycle.

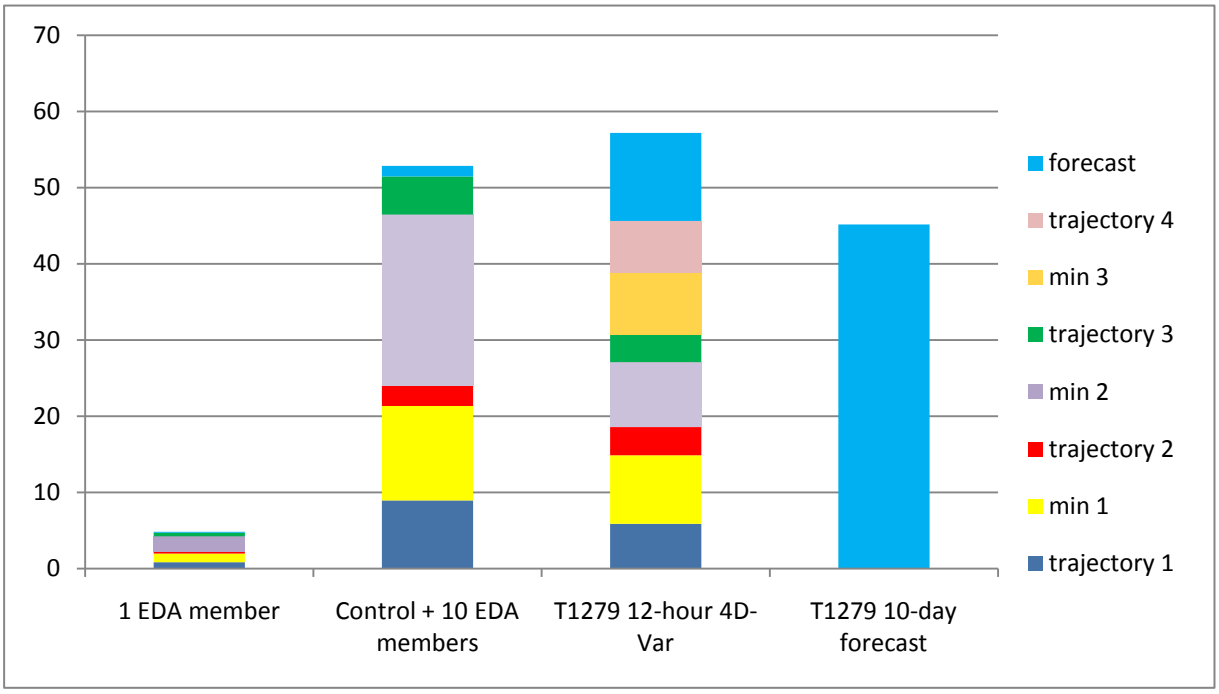

Figure 6: Real computational cost, measured as (Elapsed time in hours) x (nodes used) for:
1st bar: a single EDA member
2nd bar: the control analysis and 10 EDA members, with two outer loops
3rd bar: the T1279 12-hour 4D-Var delayed cut-off analysis followed by a 60-hour forecast
4th bar: the T1279 10-day forecast

Fig. 7 shows the tasks in an ensemble of data assimilations with a control and 3 members. At each analysis time, both the analysis and forecast families have a control family and an 'endann' family for each ensemble member. Note that the control and each member perform their own surface analysis (surf_anal). It is simple to change the resolution or number of members of a data assimilation ensemble.

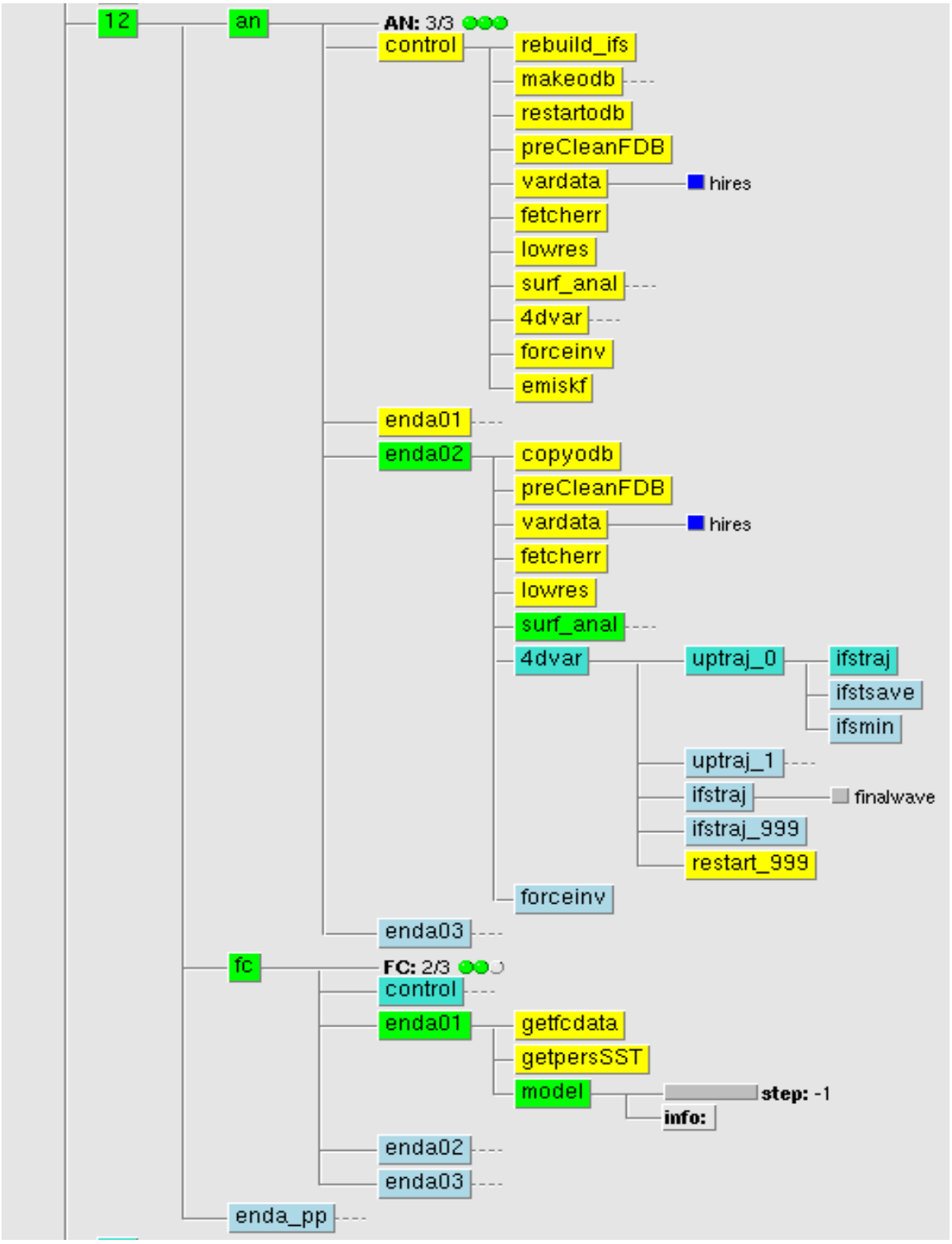

Figure 7: Computational tasks and how they are connected in a 3-member ensemble of data assimilations

5 Use of EDA-based perturbations in the Ensemble Prediction System

The simulation of initial uncertainties is one of the key aspects in ensemble prediction. At ECMWF, since the implementation of the first version of the EPS (Palmer *et al.* 2007), these uncertainties have been simulated with singular vectors (SVs, Buizza and Palmer 1995), perturbations characterized by the fastest growth, measured using a total energy norm, over a finite time interval. With the implementation of the EDA in model cycle 36r2 (22 June 2010), the methodology used to generate the EPS initial perturbations was changed, with EDA-based perturbations replacing evolved singular vectors to generate the EPS initial conditions (Buizza *et al.* 2008, Buizza *et al.* 2010).

5.1 Operational implementation of EDA perturbations in the EPS

Before 22 June 2010, the operational EPS used SVs growing over two different time periods (EVO-SVINI configuration). Initial-time SVs optimally growing during the first 48 hours of the forecast (SVINI) were used to sample directions likely to contribute most to forecast uncertainty, and evolved SVs (EVO) growing optimally during the 48 hours leading to the analysis time, were used to represent uncertainties that have been growing during the past data assimilation cycles. To improve the geographical sampling of the initial uncertainties, the initial-time and evolved SVs were computed separately over the Northern and the Southern Hemisphere extratropics, and for up to six local regions in the tropics. The initial-time and evolved SVs for the different areas were re-scaled to have initial amplitude comparable to the analysis error estimate given by the high-resolution data assimilation system.

In the new configuration (EDA-SVINI), EDA-based perturbations are used instead of the evolved SVs to represent uncertainties that have been growing during the data assimilation cycles. The initial conditions are defined by adding to the unperturbed analysis an EDA-based perturbation and a linear combination of initial-time SVs (see Appendix B for more details). The initial perturbations are symmetric, with the EPS even members having the opposite perturbation of the odd members.

The initial-time SV component is identical to the one used in the old configuration except for a reduction of the amplitude by 10%. The coefficients that determine the linear combination and the amplitude of the SVs are computed with a Gaussian sampling method (Leutbecher and Palmer, 2008). Perturbations over the extratropics (NH and SH) combine the 50 leading SVs, and perturbations in each of the targeted tropical regions combine the 5 leading SVs. Only the dry part of the model state is perturbed by the initial SVs, i.e. temperature, vorticity, divergence and surface pressure.

Each EDA-based initial perturbation is defined by the difference between one perturbed and the unperturbed 6-hour forecast (the first-guess) started from the previous EDA cycle (i.e. from EDA 4D-Var analyses run during the 12-hour period preceding the most recent analysis used to generate the unperturbed analysis). Differences between +6h first-guesses from the preceding EDA assimilation cycle are used instead of differences between analyses, since the EDA suite runs with a 12-hour delayed mode. To obtain EDA differences between analyses, it would be required to run an EDA for the ‘early delivery’ 6-hour 4D-Var suite, which would be costly and more time critical. Results indicated that the use of 6-hour forecast perturbations instead of analyses perturbations did not degrade

the probabilistic skill of the system. EDA-based perturbations are added to all the upper-level fields, including the specific humidity component, and to the surface pressure.

Since March 2008, when the 15-day variable resolution ensemble was merged with the monthly prediction system (Vitart *et al.* 2008), a key component of the seamless EPS/monthly system has been the re-forecast suite (Hagedorn *et al.* 2010). This re-forecast suite includes 5-member EPS forecasts starting from ERA-Interim analyses for the same calendar day in the past 18 years: these 90 forecasts are used to generate monthly anomaly products and to compute EPS products such as the Extreme Forecast Index (EFI). The new EDA-SVINI configuration is used also for the re-forecast suite, but since the EDA is not available for the past years, the re-forecast suite has to use the EDA-based perturbations computed for the current year. More precisely, to allow for an earlier generation of the re-forecasts, the EDA-based perturbations of the current year but for 14 days earlier are used. Experimentation has indicated that the EDA-SVINI re-forecast EPS, despite using EDA perturbations from another year, has spread and skill characteristics closer to the operational EPS than the old EVO-SVINI re-forecast EPS.

5.2 Impact of replacing evolved singular vectors by EDA perturbations

The impact of the replacement of the evolved SVs with EDA-based perturbations in the EPS is briefly summarized hereafter. Ensembles run with the current operational resolution T639vT319 (i.e. T639L62 up to day 10 and T319L62 afterwards) in the EVO-SVINI and EDA-SVINI configurations are compared for a whole season (88 cases, with initial conditions from the 5th of October to 31st of December 2009 at 1200 UTC). These ensembles have been run with the same model cycle (36r2) and using the same unperturbed initial conditions.

In terms of initial structures, the EDA-SVINI perturbations are less localized both horizontally and in the vertical, and provide a more homogeneous coverage of the globe. This is true especially over the tropics which were sampled, by design, only in a limited fashion by the initial perturbations of the old EVO-SVINI system. The EDA-based perturbations, computed at T399 resolution, have smaller scales, a less evident vertical tilt and slower growth than T42 SV-based (Buizza *et al.* 2008). This explains why an ensemble with only EDA-based initial perturbations has too little spread and a poorer performance than an EDA-SVINI ensemble.

The blending of EDA-based perturbations and the initial-time SVs combines the benefits of both sets of diverse perturbations, and provides a superior performance to the operational EVO-SVINI EPS. The effect of the modified initial perturbations is detectable over the extratropics during the first 48 hours and over the tropics for the first week. The EDA-SVINI perturbations start with a larger initial amplitude than the EVO-SVINI perturbations, but after 24 hours they have amplitude close to that of the EVO-SVINI perturbations, especially over the extra-tropical regions where the SVINI component starts dominating the perturbation growth. After 48 hours, over the extratropics the difference becomes, on average, smaller and gradually disappears in the medium-range (say around forecast day 7). The difference in the characteristics (amplitude, scale, growth rate, coverage) of the initial-time perturbations affects the average ensemble spread. This can be seen, e.g., in Fig. 8 that shows the

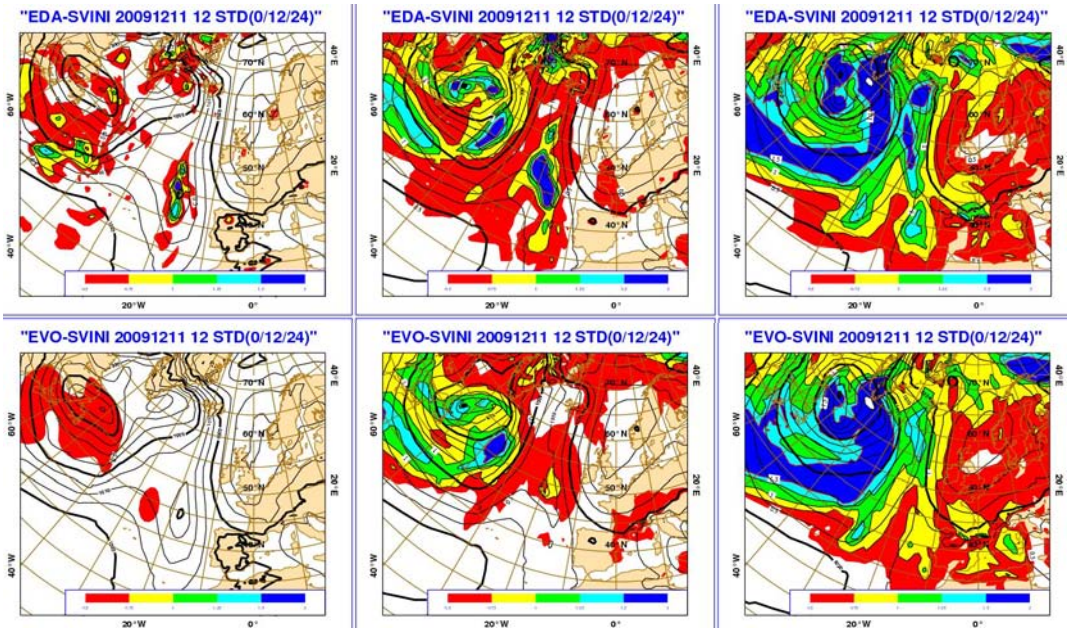

Figure 8: Ensemble-mean (black contours) and standard deviation (coloured shading) in terms of mean-sea-level-pressure (MSLP) of the EDA-SVINI (top panels) and the EVO-SVINI (bottom panels) ensembles at initial time (left panels), at t+12h (middle panels) and at t+24h (right panels). The contour interval for the ensemble-mean fields is 5 hPa; the shading for the standard deviation is for 0.5, 0.75, 1.0, 1.25, 1.50 and 3 hPa.

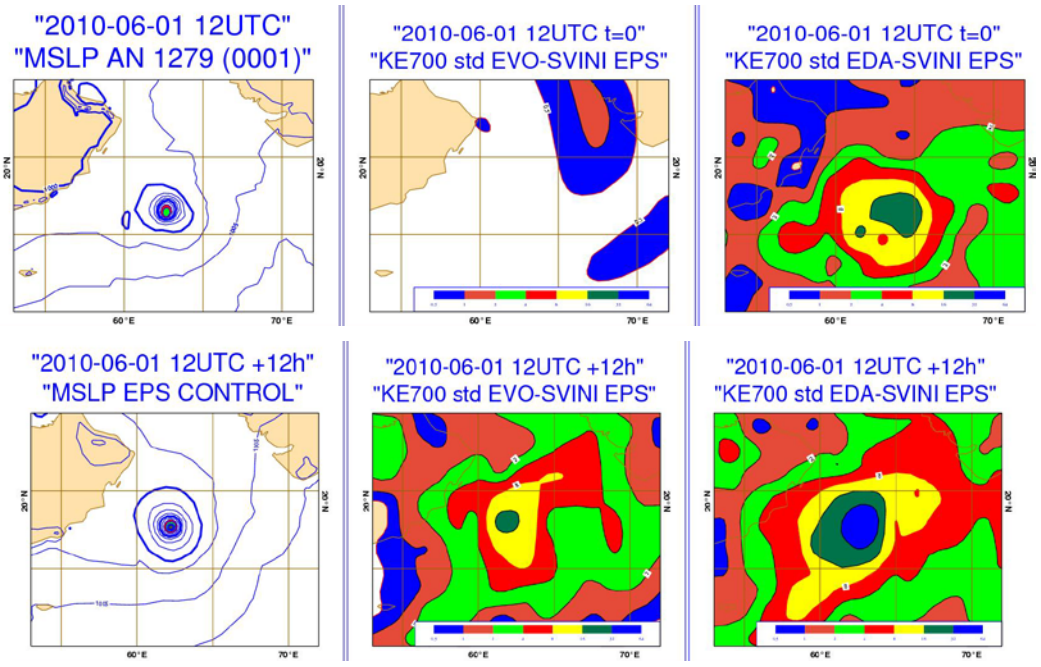

Figure 9: Mean sea level pressure for tropical cyclone Phet, south of Oman, valid at 1200 UTC 1 June 2010. Analysis (top-left panel) and 12h EPS control forecast (bottom-left panel), and ensemble standard deviation expressed in terms of the kinetic-energy per unit mass at 700 hPa at initial time (top) and at 12h (bottom) of the EVO-SVINI (middle panels) and the EDA-SVINI (right panels). Contour isolines every 2.5 hPa for the analysis and 0.5/1/2/4/8/16/32/64 m^2s^{-2} for kinetic energy.

ensemble spread at initial time, and at t+12h and t+24h for the EPS started at 1200 UTC 11 December 2009. The EDA-SVINI EPS (Fig. 8, top row) has a larger initial spread at 20ºW where the EDA perturbed analyses differ slightly in the positioning and intensification of a low-pressure system, which propagates in time and leads to larger spread at t+24h west of Spain.

The tropics is the region where the old and the new EPS differ the most, due to the fact that by design the old EVO-SVINI ensemble included only initial-time SVs targeted for few regions of the tropics, where tropical depressions were observed. The perturbed initial conditions of the EVO-SVINI and of the EDA-SVINI and their standard deviations have been compared for a few selected cases. Fig. 9 shows the EPS standard deviation in terms of mean-sea-level-pressure of the EVO-SVINI and the EDA-SVINI initial conditions in one of these cases, when the tropical cyclone Phet is present in the analysis south of Oman at (130ºE, 15ºN).

The EDA-SVINI EPS has a much larger spread, due to the fact that the depression in each EDA perturbed member have a slightly different location and/or intensity. By contrast, the EVO-SVINI standard deviation is very small, different from zero only in a region where initial-time (T42) SVs were used to compute the initial perturbations.

Overall, the EDA-SVINI ensemble has a better tuned ensemble spread and a higher skill, especially over the tropics. As an example, some statistics averaged over 88 cases for the 850 hPa temperature (T850) are shown in Fig. 10 for the Northern and Southern hemispheres extratropics, and the tropics. Over the extratropics, there is a clear increase in ensemble spread during the first 48 hours: this reduces the spread under-estimation of the old ensemble system by about 50%. Fig. 10 shows that there is a small positive reduction of the error of the ensemble-mean and a positive increase of the skill of probabilistic scores measured by the continuous rank probability skill score (CRPSS). Differences in the root-mean-square error of the ensemble-mean forecasts are statistically significant at the 5% level up to forecast day 6 over Northern Hemisphere extratropics (day 10 over Southern Hemisphere extratropics).

As expected, the impact is more evident over the tropics. In terms of probabilistic forecasts, the EDA-SVINI ensemble has statistically significant (at the 5% level) higher rank probability skill score (RPSS) over Northern Hemisphere extratropics up to forecast day 6, and over the tropics up to forecast day 8. Similar conclusions can be drawn from other probabilistic accuracy measures, such as the area under the relative operating characteristic curve or the Brier skill score (not shown).

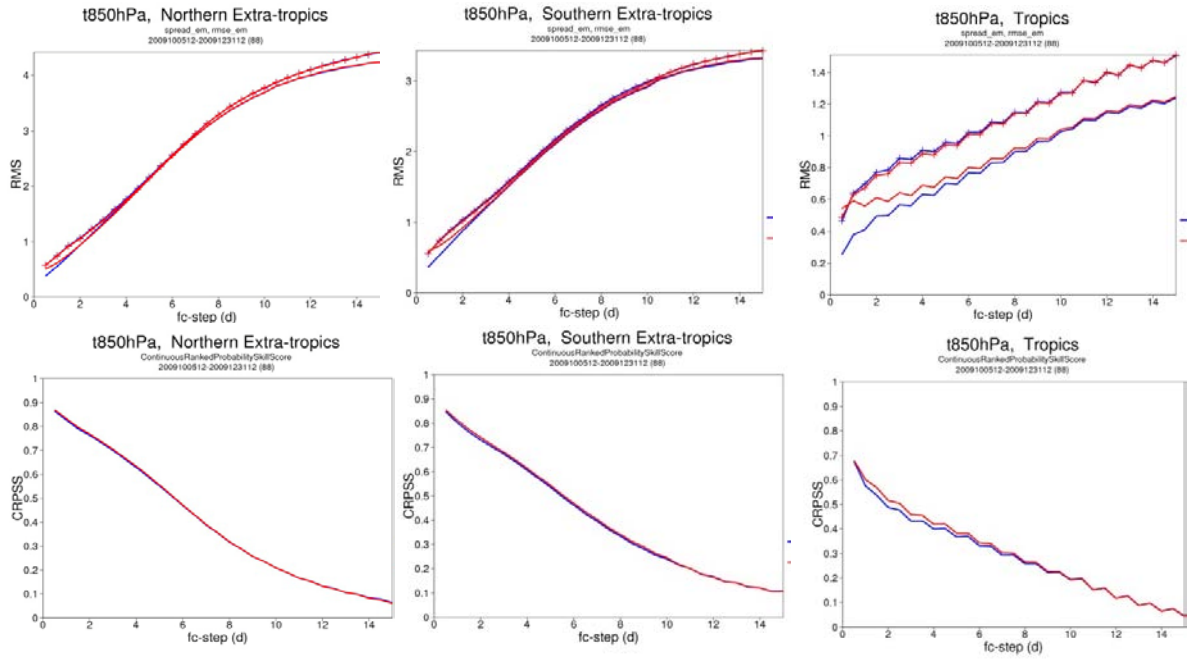

Figure 10: Average (88 cases, from 5/10 to 31/12/2009) statistics for the 850 hPa temperature over the Northern Hemisphere (left column) and the Southern Hemisphere (middle column) extratropics, and the tropics (right column). Top panels: root-mean-square error of the ensemble-mean forecast of the EDA-SVINI (red line with symbols) and the EVO-SVINI (blue line with symbols) ensembles, and standard deviation of the EDA-SVINI (red line) and the EVO-SVINI (blue line) ensembles. Bottom panels: continuous rank probability skill score (CRPSS) of the EDA-SVINI (red line) and the EVO-SVINI (blue line) ensembles.

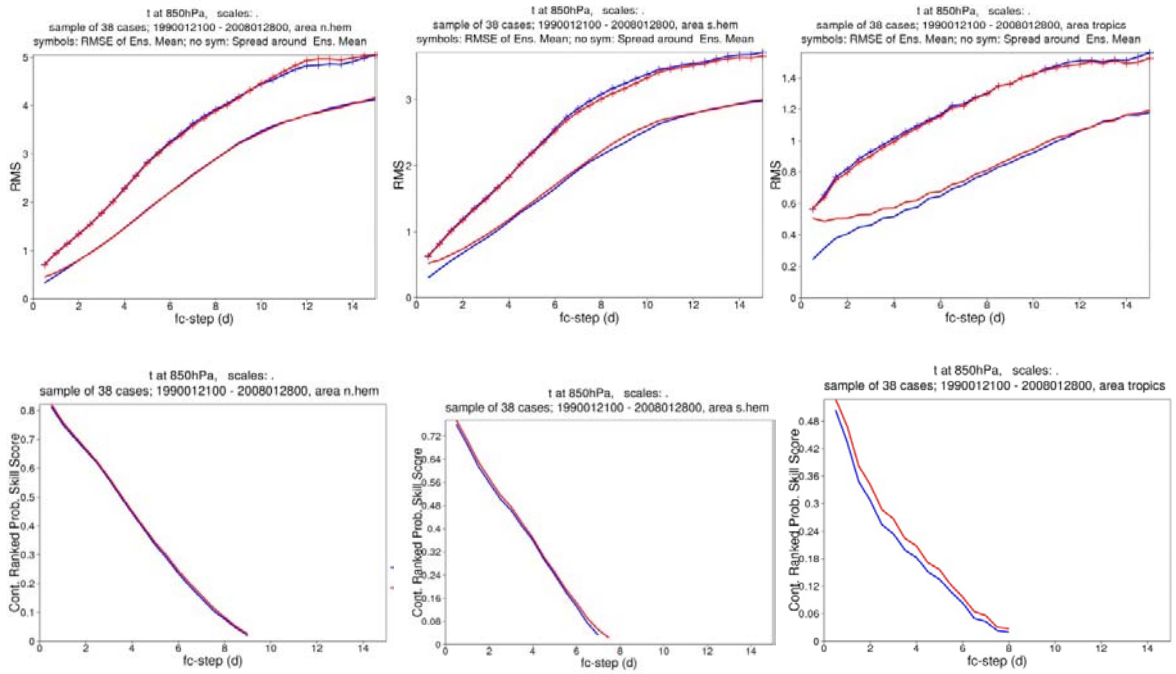

Figure 11: Average statistics of the re-forecast suite (38 ensembles with 5-member, with initial conditions 21st and 28th of January 1990-2008) for the 850 hPa temperature over the Northern Hemisphere (left column) and the Southern Hemisphere (middle column) extratropics, and the tropics (right column). Top panels: root-mean-square error of the ensemble-mean forecast of the EDA-SVINI (red line with symbols) and the EVO-SVINI (blue line with symbols) ensembles, and standard deviation of the EDA-SVINI (red line) and the EVO-SVINI (blue line) ensembles. Bottom panels: continuous rank probability skill score (CRPSS) of the EDA-SVINI (red line) and the EVO-SVINI (blue line) ensembles.

5.3 Impact of a reduced singular vector perturbation amplitude on the relationship between spread and error in the EPS

The implementation of initial perturbations from the EDA in the EPS had a positive impact on the skill of the EPS overall, but it has also resulted in an EPS which is overdispersive in terms of the large-scale flow in the extratropics. The overdispersion can be addressed by re-tuning the relative weight given to the EDA- and the SV-based components of the EPS initial perturbations.

Experiments were run for four different singular vector perturbation amplitudes for 21 cases in the period Oct–Dec 2009. The amplitude of the EPS initial perturbations is defined by the value of the γ parameter (Palmer *et al.* 2007): experiments have been performed with γ=0.0126 as operational EPS in cycle 36r2, γ=0.0100 reduced by 21%, γ=0.0080 reduced by 37% and γ=0.0060 reduced by 52%. These experiments use cycle 36r2, horizontal resolution T639, the operational Stochastically Perturbed Parameterization Tendency scheme (SPPT) scheme for representing model uncertainty and initial perturbations based on EDA and initial singular vectors. When 500 hPa geopotential height variances averaged over the extratropics are considered, the reduction of the singular vector perturbations by 37% results in the best agreement between ensemble spread and ensemble mean root-mean-square error. However, in terms of the spatial and temporal distribution of the ensemble variance, the largest reduction (by 52%) of the singular vector perturbation amplitude yields the most reliable distribution of spread (Fig. 12). The impact of reducing the singular vector perturbation amplitude on probabilistic skill for 500 hPa geopotential height is very close to neutral for reductions up to 37%. However, for the lowest perturbation amplitude, the skill in terms of the Continuous Ignorance Score is lower than for the higher perturbation amplitude. For temperature at 850 hPa, the skill measured in terms of the Continuous Ignorance Score drops already for the moderate reduction in amplitude by 21% and further decreases in perturbation amplitude lead to larger reductions in skill. For both variables, the CRPSS is insensitive to the singular vector perturbation amplitude.

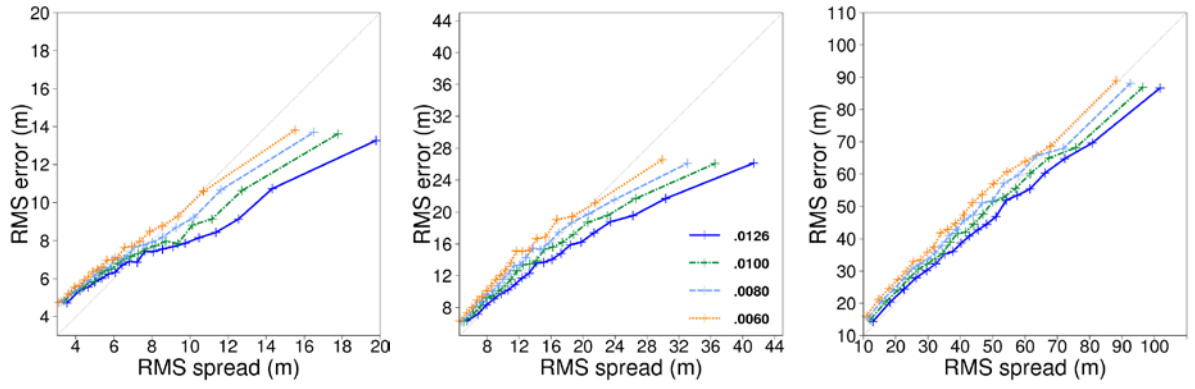

Figure 12: Ensemble mean RMS error stratified by predicted ensemble standard deviation (spread-reliability diagram) for 500 hPa geopotential height in the Northern Mid-latitudes (35°N–65°N) for four different amplitudes of the singular vector initial perturbations γ=0.0126, 0.0100, 0.0080, 0.0060 at 24h (left panel), 48h (middle panel) and 120h (right panel). Analysis uncertainty has been accounted for.

5.4 Future revisions of the representation of model uncertainties

Further upgrades of the representation of model uncertainties are planned for the EPS and EDA. It is planned to maintain a consistent methodology for representing model uncertainty in EDA and EPS in general, because it is the same deficiencies that need to be represented. The first step will be to introduce a multi-scale pattern in the SPPT and to activate the Stochastic Kinetic Energy Backscatter scheme (SPBS). However, the implementation of the upgrade is planned first for the EPS and subsequently for the EDA for a number of technical reasons. The ability to propagate the multi-scale random pattern from one assimilation cycle to the next needs to be developed. The use of EDA for providing flow-dependent background error variances has been tested with the operational single-scale version of SPPT, because no EDA runs with the multi-scale version were available at the time when the testing in the EPS started.

In September 2009 (IFS cycle 35r3) the SPPT scheme, operational since October 1998, has been revised to use random patterns that are smooth in space and time. These patterns are based on auto-recursive AR(1) processes in spectral space with correlation scales of 500 km and 6h and a standard deviation of 0.5 (Palmer *et al.* 2009). Following on from the promising initial results with a two-scale pattern for SPPT, a generic multi-scale pattern has been included in the SPPT code in order to be able to represent model errors on multiple space and time scales. Recent experiments have tested a 3-scale version of SPPT with a slow, medium and fast component contributing to the random pattern. The three components have correlation scales of 6 h and 500km for the fast scale, 3 day and 1000 km for the medium scale and 30 day and 2000 km for the slow scale. The fast scale has the same correlation scales as the single-scale version of SPPT used operationally since cycle 35r3. Preliminary testing with the seasonal prediction system helped to constrain the standard deviation for the slow scale. The variance of the medium scale is determined by assuming a power-law relation between correlation time and variance. The standard deviations of the fast, medium and slow scale are 0.52, 0.18 and 0.06, respectively. Furthermore, the tests with the 3-scale SPPT version also use the vertically-consistent treatment of supersaturation and a tighter clipping of the distribution at 2 standard deviations to avoid sign reversal of tendencies.

The SPBS scheme is described in Palmer *et al.* (2009). For SPBS, the spectral power distribution of the forcing pattern is set to be of the form of a power law, with its slope determined from coarse-graining. The decorrelation time used in the evolving AR(1) process is set to approximately 7 hours for all wavenumbers. The SPBS forcing is restricted to wavenumbers 1 to 159 for pragmatic reasons primarily. Coarse graining results suggest that for wavenumbers larger than 159 the decorrelation time is much shorter than 7 hours, implying only a small energy input and hence small impact at those scales. The practical advantages of restricting the forcing to a limited wavenumber band are a reduction in computational cost of SPBS by approximately 25% and avoidance of spin-up of the diagnosed dissipation rate used in SPBS due to feedback with the forcing. The backscatter ratios for the operational EPS and EDA resolutions are 0.095, 0.064 and 0.043 for T639, T399 and T319, respectively. These backscatter ratios imply a similar energy input at all three resolutions and the resulting stream function forcing is in line with coarse graining results.

The impact of the joint use of SPBS and the multi-scale SPPT is being evaluated in the EPS first, before introducing it in the EDA. The experiments use EDA initial perturbations and initial singular vector perturbations with reduced amplitude (γ=0.0060). It was decided to use the halved amplitude for the singular vector perturbations because the reduction has a significant positive impact on the spread-reliability (Fig. 12). Furthermore, preliminary experiments using γ=0.0126 showed that the joint use of SPBS and multi-scale SPPT instead of the single-scale SPPT alone would exacerbate the overdispersion of the large-scale flow in the extratropics. Model uncertainty in EDA has been represented by the single-scale version of SPPT only for reasons explained above. Results of this experimentation indicates that the upgrade of the model uncertainty representation together with the halving of the amplitude of the singular vector perturbations has a positive impact on probabilistic skill and the relationship of spread and ensemble mean error.

5.5 Other EPS future developments

The availability of EDA-based perturbations has opened the possibility to improve other areas of the EPS.

A first area of investigation is the possibility to include surface perturbations. Following an assessment of the spread of the EPS in the boundary layer and for surface variables, the potential use of EDA-based perturbations to perturb surface variables (e.g. for soil moisture and soil temperature) is being investigated. Preliminary results indicate, e.g., that the addition of EDA-based soil moisture and temperature perturbations can improve the spread of the EPS over land in the early forecast range. The plan is to investigate the impact of adding perturbations to other relevant land surface variables.

A second area that could lead to improvements is the use of a larger ensemble of perturbed analyses (25 or 50 instead of 10). Experiments to investigate the potential impact of this upgrade will start as soon as it is clearer which EDA configuration proves to give the most cost-effective improvements in the estimation of the background error statistics in the ECMWF high-resolution data assimilation system.

6 Use of the EDA in 4D-Var for flow dependent QC and background error estimation

It is widely recognised that an accurate specification of the statistics of background errors is a fundamental prerequisite of a successful data assimilation system. Methods based on an ensemble of perturbed observations assimilations, like the EDA described in Chapters 2 to 4, have been successfully employed in the specification of background error statistics at ECMWF (Fisher 2003) and Météo-France (Belo Pereira and Berre 2006). However, these applications have relied on a static, climatological representation of the background error matrix, thus implicitly discarding to account for the flow-dependency of covariance errors.

An attractive characteristic of EDA methods is their ability to provide estimates of day-to-day background error statistics that represent the current meteorological situation, thus possibly overcoming one of the main limitations of covariance models currently in use in variational data

assimilation systems. Early attempts to use flow-dependent error information from a 6 or 10 member EDA have been reported in Kucukkaraca and Fisher (2006), Fisher (2007), Isaksen *et al.* (2007) and, more recently, in Raynaud *et al.* (2009, 2010).

It must be kept in mind, however, that the day-to-day background errors derived from the sample ensemble statistics are themselves affected by estimation errors which have to be dealt with and possibly minimized. It has proved useful in this context (Berre and Desroziers, 2010) to separate the estimation error into a random and a systematic component:

$$b - b^* = \left(b - E\left[b\right]\right) + E\left[b - b^*\right] \tag{6.1}$$

where b is the ensemble estimate of any component of the background error and b^* is its true value. The first (random) component is the part of the estimation error which averages out to zero under the application of the statistical expectation operator. We would expect it to be negligible for an infinite ensemble size. For affordable ensemble sizes this type of error effectively limits the information on the forecast errors that we are able to extract from the EDA sample statistics.

The second (systematic) error component is the part of the estimation error that tells us how much our EDA error estimates deviate on average from the truth. As such it gives an indication of underlying problems in the representation of the main sources of uncertainties in the EDA.

6.1 Random errors in EDA based variances

The ECMWF EDA tries to represent the evolution of the deterministic 4D-Var assimilation cycle errors through an ensemble of lower resolution deterministic 4D-Var analysis cycles which make use of perturbed observations, perturbed sea surface temperature (SST) fields and perturbed model physical tendencies. Being a purely stochastic method, one can expect the accuracy of error estimates sampled from the ensemble to increase proportionally to the square root of the ensemble size. This slow rate of convergence and the limited (O(10-100)) ensemble sizes which are viable in current NWP operational contexts require the application of ad-hoc filtering techniques (Raynaud *et al.*, 2008, 2009; Berre and Desroziers, 2010) to the raw ensemble estimates.

Limiting our attention to sample variances, if we define $G^e(i)$ as the random component of the estimation error in the estimated ensemble variance at grid point i,

$$G^e\left(i\right) \stackrel{\text{def}}{=} \tilde{B}_{ii} - E\left[\tilde{B}_{ii}\right] \tag{6.2}$$

Then it can be shown (Mallat *et al.* 1998, Raynaud *et al.* 2008) that the covariance of this sampling noise is a simple function of the expectation of the ensemble-based error covariance matrix:

$$E\left[G^e\left(i\right)G^e\left(j\right)\right] = \frac{2}{N-1}\left(E\left[\tilde{B}_{ij}\right]\right)^2 \tag{6.3}$$

where N is the ensemble size.

An immediate consequence of this equation is that the sampling noise correlation length scale is smaller on average than the background errors length scale. The error variance field, which is the actual signal of interest here, tends to show variations on even larger length scales, reflecting observation density variations and coherent synoptic structures. These arguments provide the basis for using a spectral filtering technique to disentangle the sampling error from the sampled EDA variances. An example is given in Fig. 13, where the green line shows the power spectrum of the 10-member EDA short range (t+9h) forecast variance and the red line the power spectrum of the climatological sampling noise computed from Eq. 6.3 for vorticity at model level 64 (~500 hPa). The separation in spectral space of the signal (the EDA variance) from the random error power spectrum is clear up to about wavenumber 80. It is apparent from Fig. 13 that the objective filtering procedure allows a truncation wavenumber to be defined beyond which the sampled error estimates are not deemed to be statistically significant. This implies that, at least in terms of forecast error estimation, the effective spatial resolution of the ensemble is typically lower than the nominal resolution at which the ensemble forecasts are actually run (Spectral triangular truncation T399 in the present case, similar to the operational implementation).

The objective filtering technique proposed by Raynaud *et al.* (2009), which has been implemented at Météo- France and ECMWF, applies the following spectral low-pass filter:

$$\rho(n) = \left[\cos\left(0.5 * \pi * \min\left(n, N_{trunc}\right) / N_{trunc}\right)\right]^2 \tag{6.4}$$

where n is the total spectral wavenumber and N_{trunc} is the truncation wavenumber of the filter.

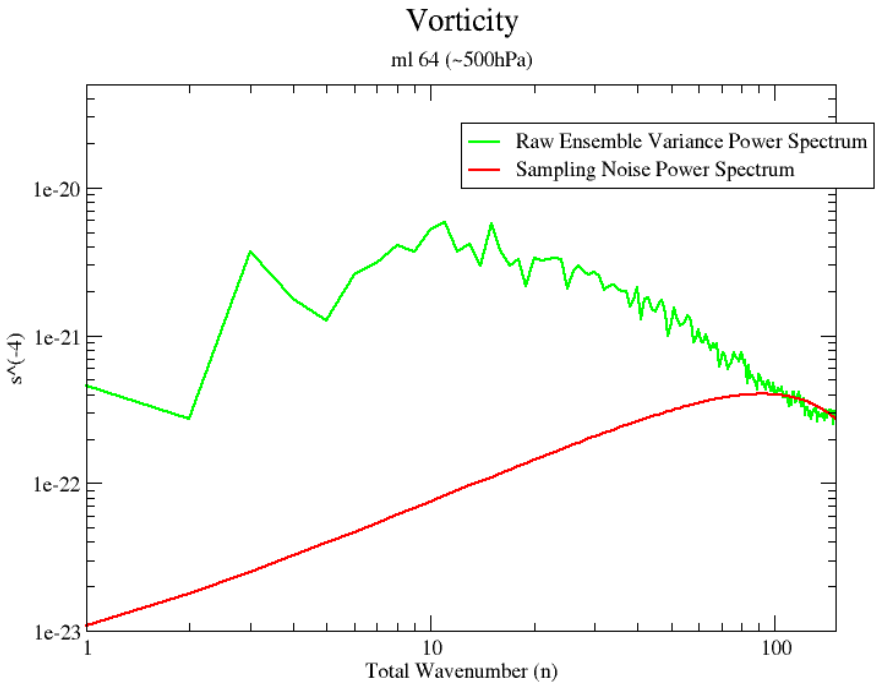

Figure 13: Power spectra of ensemble short range (t+9h) forecast variance (green) and of the climatological sampling noise (red) associated to the estimation of vorticity variances at model level 64 (~500 hPa) with a 10-member ensemble. Unit: s^{-4}.

The application of this low-pass filter in spectral space is equivalent to a spatially weighted averaging technique in grid-point space, which enables the large-scale signal of interest to be extracted while filtering out the small-scale sampling noise. The truncation wavenumber (N_{trunc}), which is a function of model variable and vertical model level, is derived by computing the effective length scale of the sampling noise and retaining the signal spectral components which have larger length scales.

Fig. 14 (top panel) shows an example of the raw EDA background standard deviation field for vorticity at model level 64 (~500 hPa); the objectively filtered version (middle panel); and, for comparison in the bottom panel, the corresponding background error estimate from the "randomization" technique (Fisher and Courtier, 1995; Fisher, 2003), which is currently used in operations at ECMWF. This technique, as mentioned earlier, is based on random sampling of the static B matrix, with a weak flow-dependent component due to the use of the non-linear balance equation and the omega equation, linearized around the background state. The filtered variance map shows, as expected, the same large scale features present in the raw EDA spread map, 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 have a much clearer signature of the current synoptic situation than the randomization error map: EDA errors are clearly concentrated in the meteorologically active areas of sparse observational coverage, as we would expect them to be.

6.2 Sensitivity of sample standard deviations and filtering to ensemble size

A diagnostic study has been performed to investigate the impact of ensemble size on sample standard deviation and filtering properties (see Bonavita *et al.* 2010 for further details). The results presented are based on the analysis of a 50-member EDA experiment run for 45 days with the same configuration as the operational 10-member EDA described in Chapter 4, over both the boreal winter (January-February 2009) and summer (August-September 2008) seasons.

The total standard deviation on each model level (i.e. the global spatial average of the ensemble background standard deviations) is the first global diagnostic to check to evaluate the impact of larger ensembles. Fig. 15 shows the sample ensemble variance power spectra for temperature at model level 49 (~200 hPa, continuous lines) for a particular date and the corresponding climatological sampling noise spectra (dashed lines) from a 10-member ensemble (red), a 20-member ensemble (blue), and a 50 member ensemble (black), respectively. As expected, the portion of the spectrum corresponding to the large scales features (up to wavenumber 40 approximately) is virtually identical for all ensemble sizes. For smaller spatial scales, however, the sampling noise contribution to the total variance spectra is much smaller for the larger ensembles. The climatological noise spectra estimate is also correspondingly smaller, 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 is around T70 for the 10-member ensemble, T90 for the 20-member ensemble, and more than T159 for the 50-member ensemble.

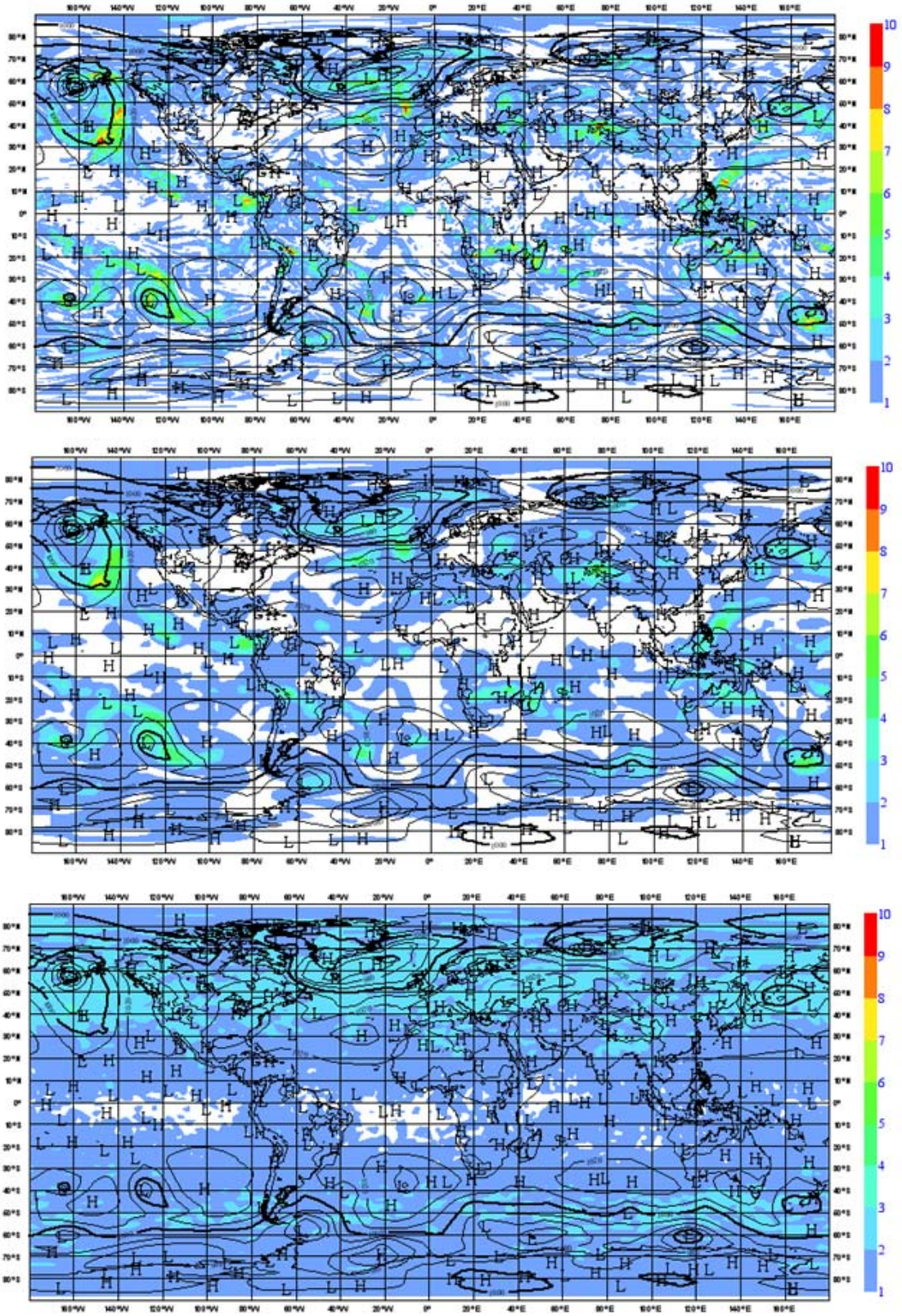

Figure 14: Standard deviations of the EDA background vorticity at model level 64 (~500 hPa), valid at 0900 UTC 10 January 2009. Unit: 5.10^{-5} s^{-1}. Top panel: raw estimates from the 10 member ensemble. Middle panel: objectively filtered estimate. Bottom panel: background error estimate from the operational "randomization" technique. The mean sea level pressure analysis is overlaid, contour interval 10 hPa.

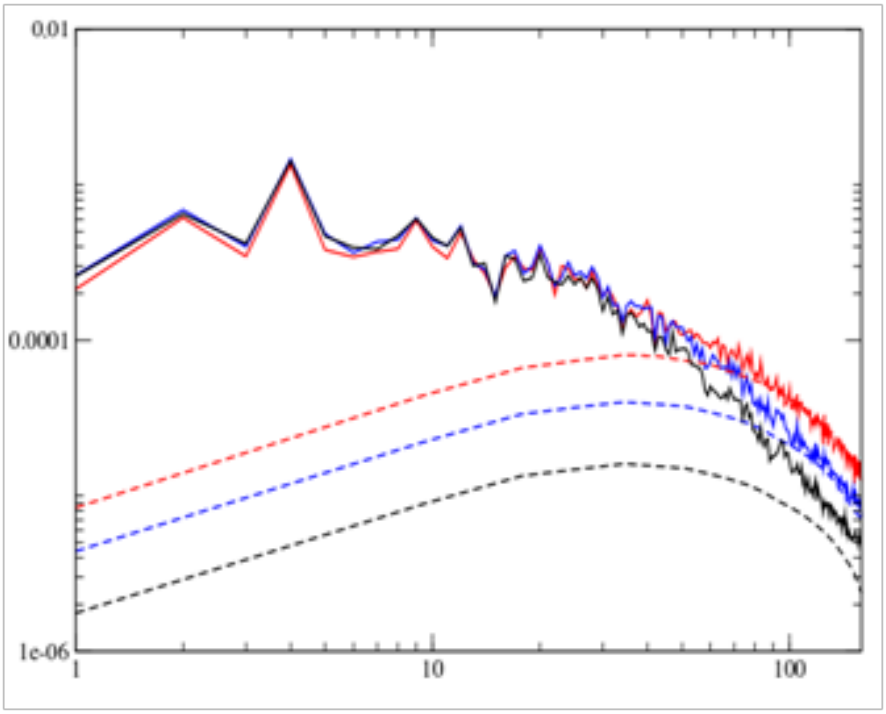

Figure 15 Power spectra of ensemble first guess variance valid on 21 Jan 2009 0900 (continuous lines) and of the climatological sampling noise (dashed lines) associated to the estimation of temperature variances at model level 49 (~200 hPa) for a 10-member ensemble (red), a 20-member ensemble (blue) and a 50 member ensemble (black). Unit: °C^4.

It is noteworthy that even for the 20 member ensemble the effective spatial resolution of the ensemble sample error estimates is far coarser than the nominal ensemble resolution (i.e., ~T90 vs. T399). One needs to use a 50 member ensemble to be able to extract a relatively noise free signal over the whole spectral range, up to T159, as shown by the black curves in Fig. 15. This implies that, in the EDA framework, the most effective way to increase the ensemble sampling error resolution is by increasing its size. This conclusion is based on the assumption that sampling error is the dominant form of error in the ensemble estimates of first guess errors.

A fairly accurate estimate of background error variances can be obtained with a limited number of ensemble members (10 in the present case). In particular, the uncertainty in the background state associated with extra-tropical synoptic scale features can be well captured by an ensemble of only 10 members. However, our investigation has shown that there are some caveats attached to the use of small ensembles. The filtering technique employed to extract the statistically significant signal from the ensemble variances drastically reduces the effective spatial resolution of the ensemble with respect to the nominal one (for vorticity at model level 64, effective truncation wavenumber is around T80, versus the T399 nominal resolution). Furthermore, this aggressive filtering inevitably affects the spectrum of the signal we would like to retain. These much coarser filtered ensemble variances may still have enough detail to be able to resolve the large scale uncertainties in the extra-tropical flow. Conversely, when we are faced with more localised anomalies such as those associated with a tropical disturbance, it can be argued that a filtered 10-member ensemble is only partially able to resolve the error features of interest and the advantage of a larger ensemble is more noticeable. These

investigations on the effective spatial resolution of the ensemble point to further possible gains with larger ensemble sizes, in terms of local error characterization.

6.3 Systematic errors in EDA variances

The use of perturbed observations, perturbed SST and perturbed model physical tendencies in the ECMWF EDA has the objective of characterizing the main sources of errors in the high resolution analysis cycle. Consequently any deficiencies or approximations in the prescribed observation error statistics, SST perturbation fields and model error parameterizations will cause the sampled EDA variances to be sub-optimal estimates of the analysis and forecast errors. This type of estimation error will not be alleviated by an increase in ensemble size and will translate into systematic differences between the EDA sampled variances and the analysis/forecast errors that can be diagnosed a-posteriori. This has been identified at an early stage during the development of the ECMWF EDA (Kucukkaraca and Fisher 2006; Fisher 2007; Isaksen *et al.* 2007): globally averaged EDA spread values were diagnosed to be underestimated by approximately a factor of two, so that a global inflation factor of the same magnitude was applied to the sample EDA standard deviations.

However the systematic errors of the EDA variances have a complex spatial and temporal structure which can hardly be accommodated with a global multiplicative inflation factor. Fig. 16 shows the average over one month (January 2009) of the difference between the EDA background standard deviation and the diagnosed EDA background mean errors (computed using the operational analysis as a proxy for the truth) for vorticity at model level 30 (~50 hPa) and 78 (~850 hPa). It is clear that systematic errors present non-negligible latitudinal and model level variations. In this example, at model level 30 the EDA appears to be nearly calibrated in the Southern Hemisphere and the tropics, while it is under-dispersive in the Northern Hemisphere. On the other hand the EDA short range forecast at model level 78 tends, on average, to be underdispersive in the extratropics while being closer to being reliable in the tropical belt. This last feature can be interpreted as a consequence of the fact that the model error parameterization scheme tends to be more active in regions which are likely to produce larger physical tendencies, of which the tropical troposphere is an example.

The relationship between EDA spread and analysis/background errors can be investigated more quantitatively by the use of Spread-Error plots, an example of which is given, for vorticity near 500 hPa, in Fig. 17. These are obtained dividing the sample EDA spread fields into deciles according to their magnitude and plotting their mean value together with the RMS values of the corresponding background mean errors. For a reliable EDA the calibration curves (separately computed for the Northern Hemisphere extratropics, Southern Hemisphere extratropics and the tropics) should lie on the diagonal (dashed black line). Their distance to the diagonal gives us an immediate information about the under/overdispersiveness of the ensemble, while their slope gives an indication of the conditional biases of the ensemble spread. The slope of the calibration curves with respect to the diagonal suggests that different calibration factors should be applied to the sampled EDA spread distribution.

The Spread-Error calibration coefficients do not change significantly on daily and weekly timescales. However they show a significant seasonal drift. Fig. 18 shows the mean calibration factors for vorticity in the Northern Hemisphere extratropics, averaged over one winter month and one summer month, respectively. In the stratosphere the large differences shown are most likely related to the polar

vortex activity in the winter months, while in the lower tropospheric levels the larger under-dispersiveness of the winter EDA is linked to model error issues in a weakly coupled boundary layer and the lack of explicit perturbations of the soil parameters and atmospheric parameters near the surface. Irrespective of the physical causes of the drift of the calibration factors, it is clear that any statistical correction we wish to apply to the EDA spread will benefit from a slowly time-varying component. In the ECMWF implementation this is achieved by computing a running mean of the calibration factors over the past five days (e.g., the most recent, available ten analysis cycles).

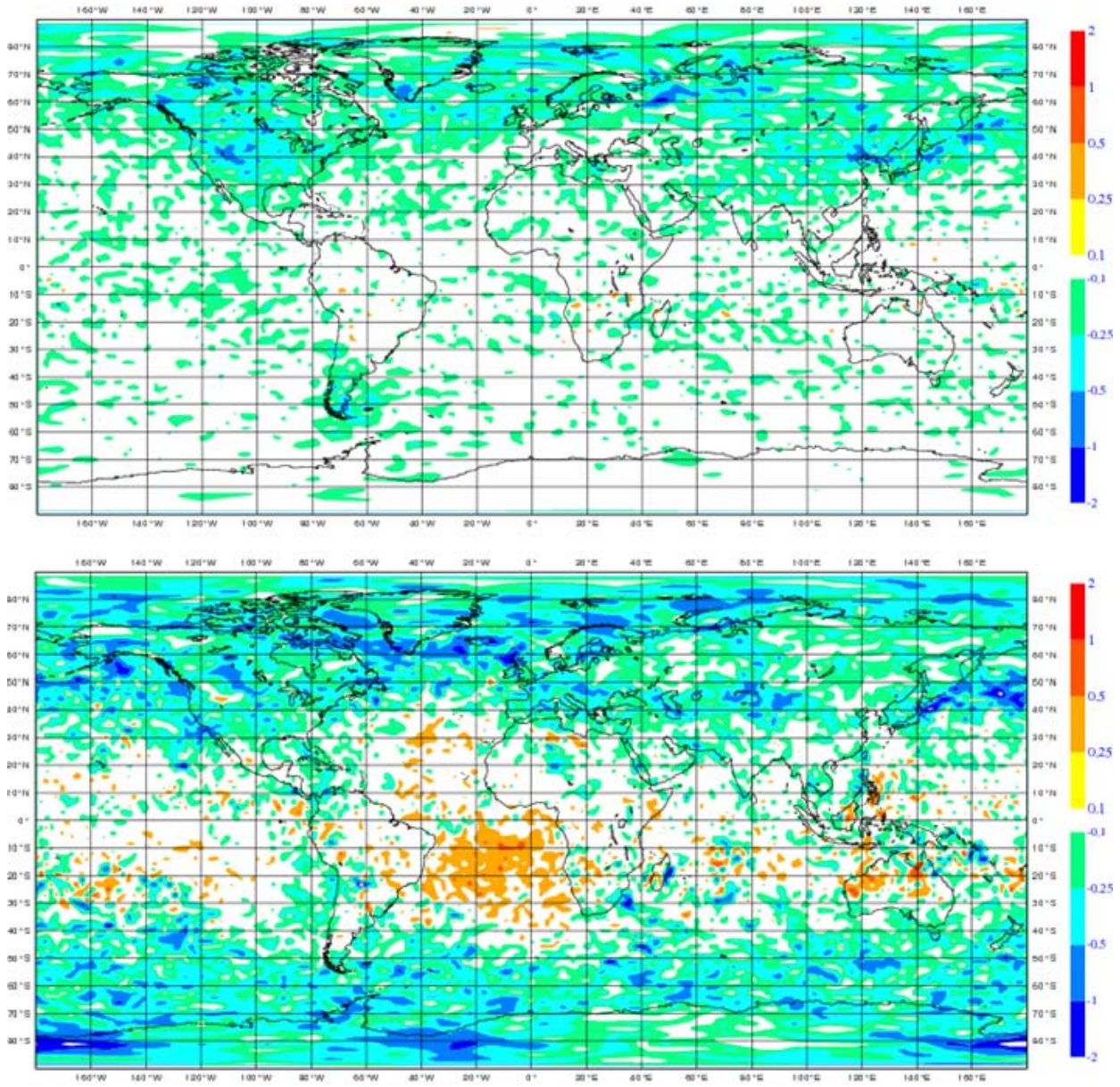

Figure 16: Difference between EDA background forecast spread and EDA background mean error for Vorticity at model level 30 (~50 hPa, top) and model level 78(~850 hPa, bottom), averaged over one month (January 2009). Unit: 10^{-5} s^{-1}.

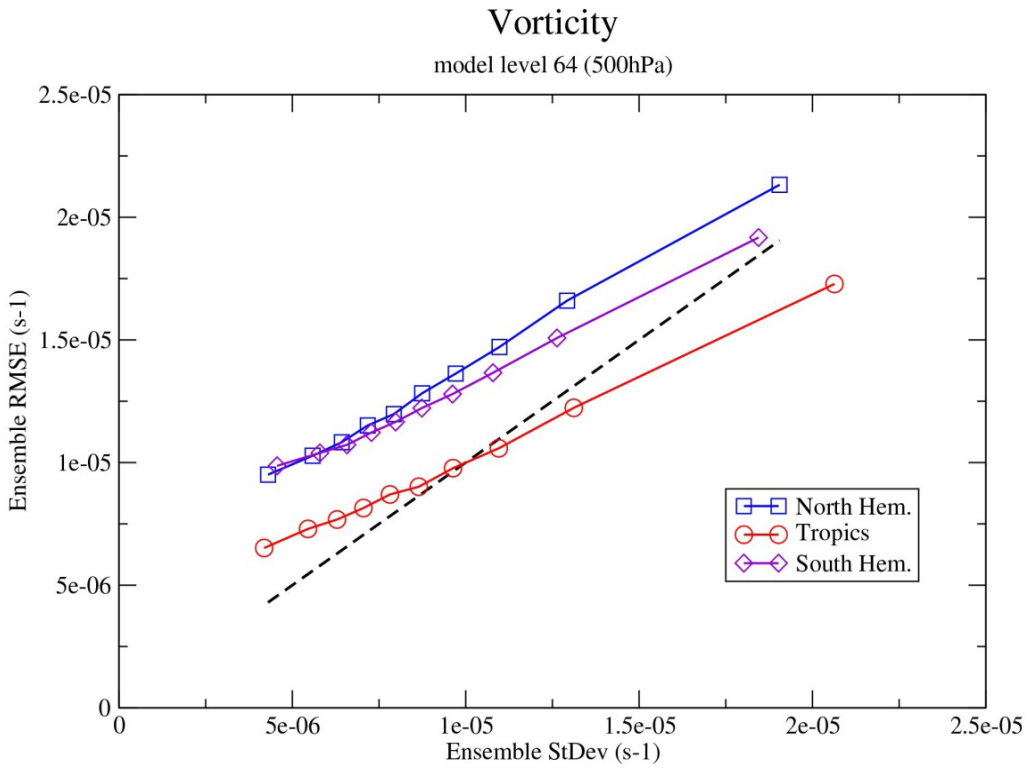

Figure 17: Spread-Error diagram for vorticity at model level 64 (~500 hPa), averaged over one month (January 2009). Northern Hemisphere extratropics (blue), Southern Hemisphere extratropics (purple), tropics (red). Unit: s^{-1}

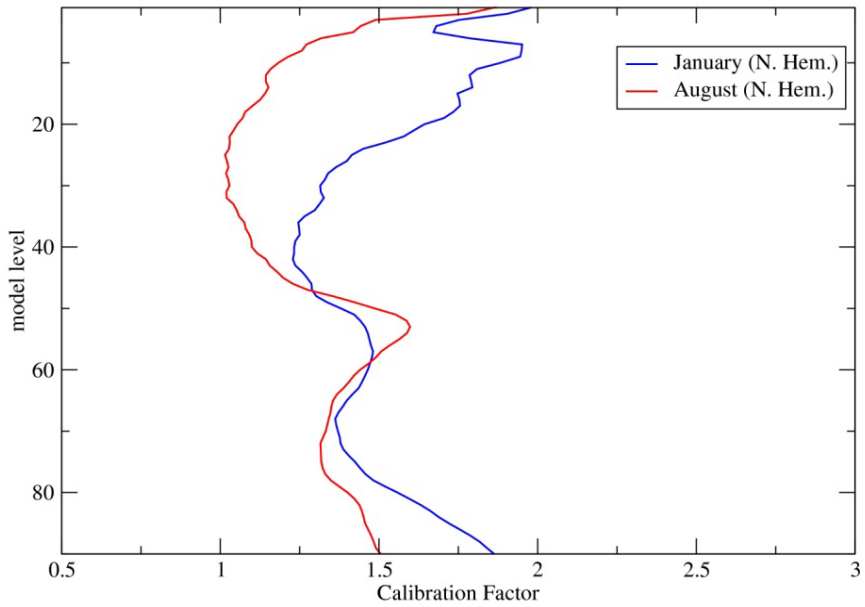

Figure 18: Calibration factors for standard deviation of vorticity for the Northern Hemisphere extratropics. Averaged over one month (January 2009, blue; August 2008, red).

6.4 4D-Var assimilation experiments using flow-dependent EDA variances

A 10 member EDA cycle was run alongside the deterministic 4D-Var analysis cycle for the assimilation experiments performed. This EDA provided estimates of flow-dependent background errors. The above discussion has highlighted the need of applying a spectral filter and a statistical recalibration step to the raw ensemble variances sampled from the EDA background forecast. This makes sure that the background error estimates have realistic values and are not excessively contaminated by sampling noise. The background errors derived from the recalibrated EDA variances are used in the 4D-Var analysis, both in the minimization step and in the first guess quality control check of the observations, replacing the "randomization" technique currently used in the operational configuration. A schematic representation of the process is shown in Fig. 19.

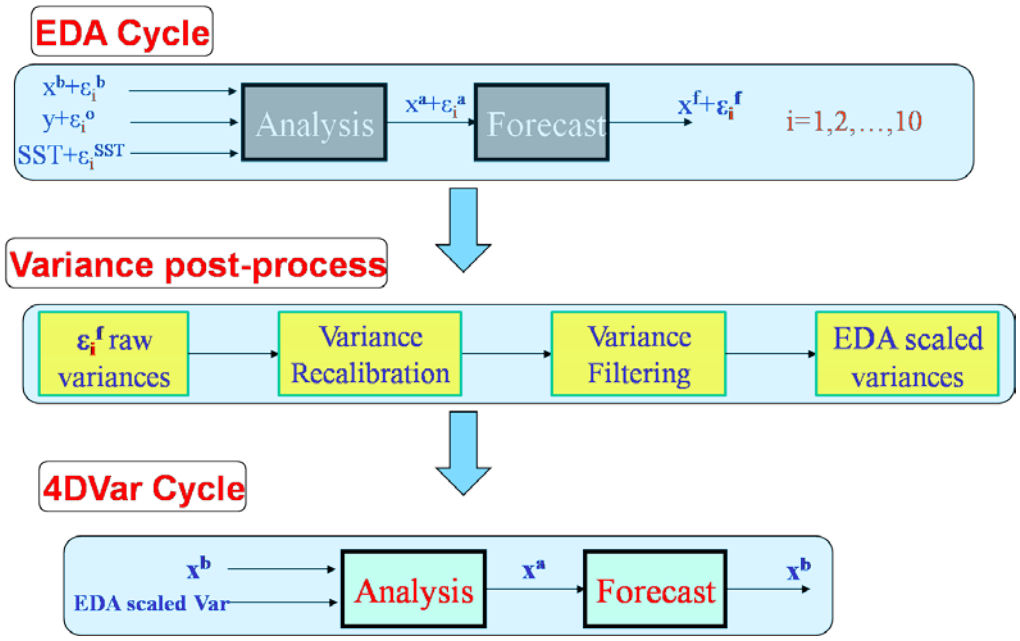

Figure 19: Flow diagram of the use of EDA variances as background errors in the deterministic 4D-Var assimilation cycle.

An example of the background error estimates derived from the recalibrated EDA variances is shown in Fig. 20. Recalibrated EDA standard deviations of the zonal component of wind are shown over the European area for two dates (23, 24 January 2009, both 2100 UTC). The pressure minimum visible in the Gulf of Biscay in the top panel is associated with the Klaus storm, which was responsible for widespread damage during its passage over France. It is apparent that the EDA identifies an area of increased uncertainty connected with the storm, associated with the larger values of standard deviation diagnosed for the u-wind component. After 24 hours (bottom panel) the weakened storm has reached the central Mediterranean Sea. An area of larger standard deviation is still visible, though values are reduced, as we would expect from the fact that the storm has crossed a densely observed region. The next extra-tropical synoptic system visible in the near Atlantic (bottom panel) is also interesting: the regions of increased EDA spread indicate that there is considerable uncertainty over the exact position of the new low and the associated fronts.

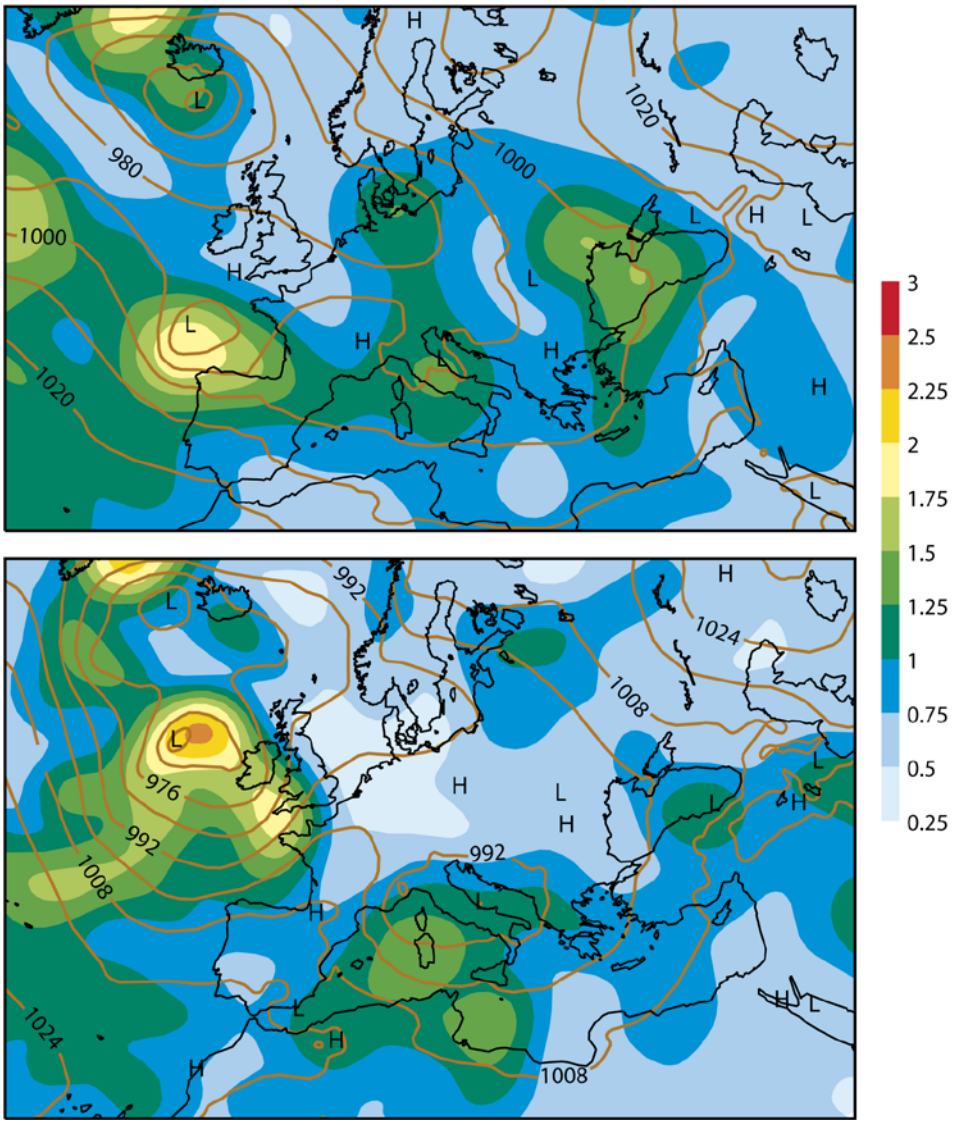

Figure 20: Standard deviations of zonal wind component at model level 78 (~850 hPa), for 9h forecast valid at 2100 UTC 23 January 2009 (top panel) and 2100 UTC 24 January 2009. Unit:ms^{-1}. Analysed mean sea level pressure field (8 hPa contours), valid at the same time, has been superimposed.

Two extended assimilation experiments using the recalibrated EDA variances in the deterministic analysis have been performed over both a boreal winter (January-February 2009) and a summer (August-September 2008) period. Forecast scores are shown in Fig. 21 for the 500 hPa geopotential height in the extratropics, and in Fig. 22 for the 200 hPa and 100 hPa vector wind in the tropics. The experiments have been run at the current operational resolution (T1279L91) using the latest model version available at the time (Cycle 36R2) and verified against the operational analysis Statistically significant improvements are seen in the first 1-2 day forecasts in the extratropics, more noticeably in the Southern Hemisphere. Improvements in the Tropics seem to be more evenly distributed over the entire forecast range and are statistically significant for the boreal summer experiment.

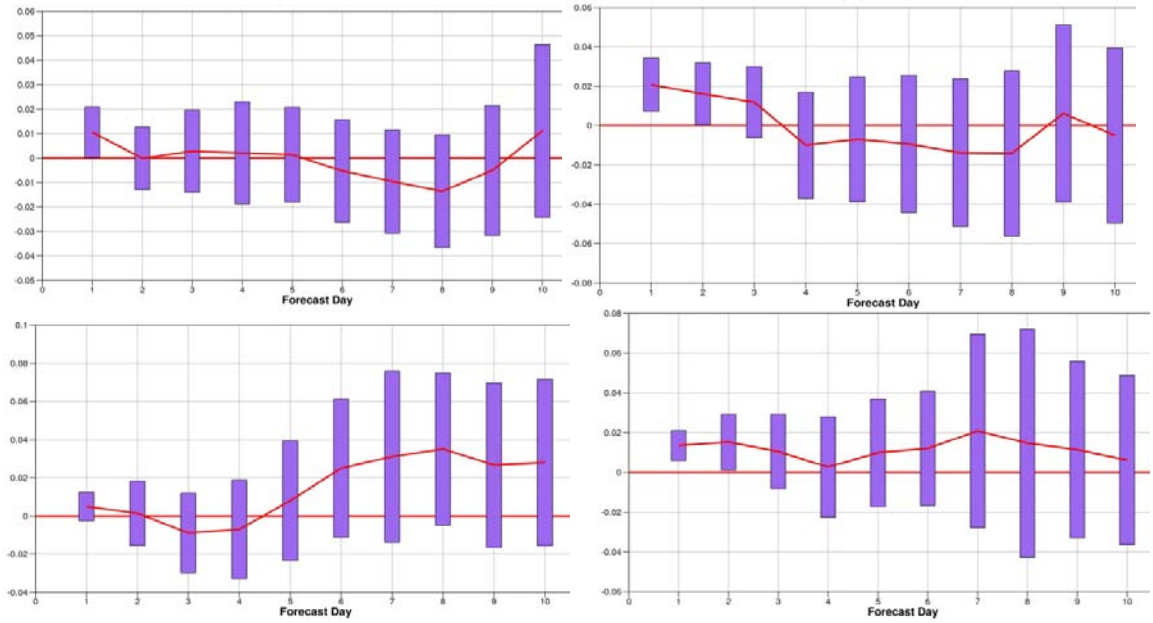

Figure 21: Forecast skill score differences for 500 hPa geopotential height RMSE of forecasts from deterministic 4D-Var analyses with background error variances estimated from a 10 member EDA versus the operational configuration that use the randomization method. Scores are averaged over the period 20090107-20090216 for the top two rows and for the period 20080807–20080916 for the bottom two. First column refers to Northern Hemisphere extratropics, second column to Southern Hemisphere extratropics.

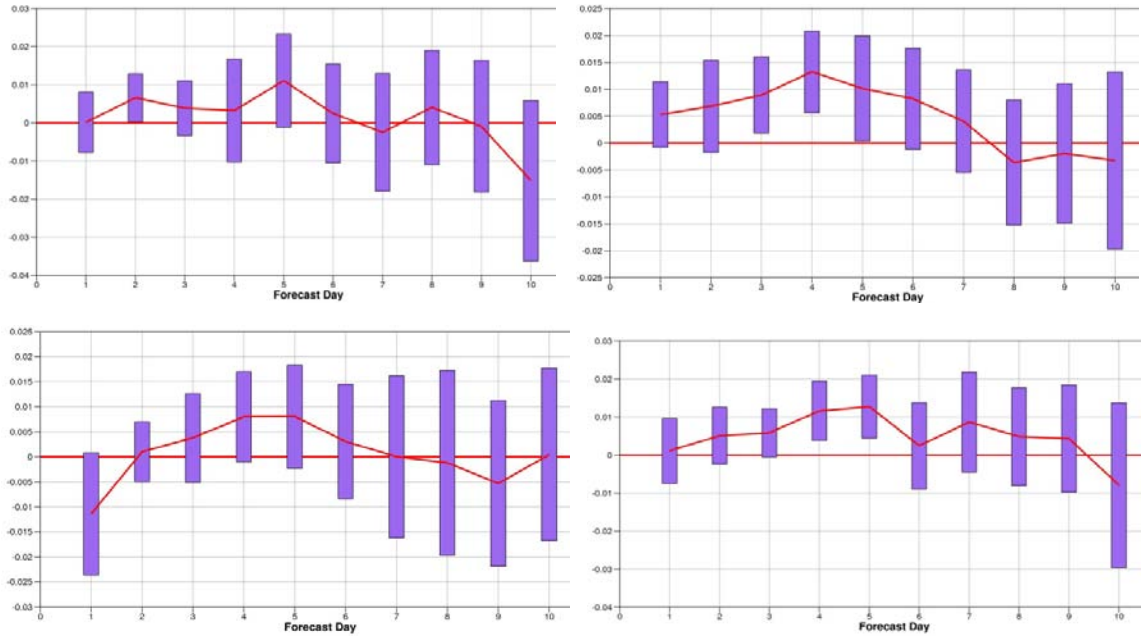

Figure 22: Similar to Fig. 21, but for Tropical Wind Vector scores at 200 hPa, top panels, and 100 hPa, bottom panels. Scores are averaged over the period 20090107-20090216 for the first column and for the period 20080807–20080916 for the second column.

6.5 Further developments in use of EDA in 4D-Var

The algorithm described provides a viable and promising framework for the utilization in the ECMWF deterministic analysis of cycled background error estimates sampled from the EDA. It will very likely be implemented in operations in its present form. It can however be further extended and improved upon in various aspects.

In terms of reducing the random error component of the EDA estimates it is clear that a larger EDA ensemble would be beneficial. This does not mean that we would be able to eliminate the spectral filtering of the EDA variances, since the sampling error for the estimated standard deviations would only reduce proportionally to the square-root of the ensemble size. However, as discussed in Section 6.2, a larger ensemble would allow the retention of higher wavenumber information.

The ability of a larger ensemble to represent finer spatial details of background error fields should translate into more accurate analysis and forecasts. Preliminary results with a 20-member assimilation experiment show that increasing the sample size from 10 to 20 members has a generally positive impact on forecast scores in the Northern Hemisphere and, more markedly in the Southern Hemisphere (Fig. 23), while the impact is neutral in the tropics (not shown).

The calibration step has been implemented to tackle the problem of insufficient variance in the EDA for specific parameters and latitude bands. The under-dispersiveness of the EDA is thought to arise from deficiencies in the current formulation of the observation error matrix and the representation of model error. The improvement of these aspects of the EDA is a long term research goal. A short-term objective is the development of the current formulation of the calibration step. The present implementation uses the operational analysis as the verifying truth for the calibration. This is a rather strong assumption which may be justified in densely observed regions, but not elsewhere. Current research effort is devoted to mitigate this assumption by looking both at the EDA a-priori diagnostics and at the statistics of observation increments.

Further extensions to the current methodology will also consider the use of EDA information in the estimation of background error correlations. Being a much larger-dimensional problem, sampling issues are expected to play an even bigger role than in the estimation of background error variances. This would probably imply the need to make use of a larger ensemble and to apply local spatial averaging techniques similar to those that have been implemented for the sampled EDA variances. In this context, an attractive possibility would be to extend the wavelet diagonal approach (Fisher, 2003) already implemented in the IFS, to make use of EDA flow-dependent correlation structures.

Quality control is an area that requires further study. The current EDA uses the same quality control parameters for the ensemble members as are used for the control analysis, despite the additional perturbations that are added to the observations used by the members, and the fact that the background for each member is perturbed with respect to the control. This results in more observations being rejected by the members than by the control. The difference in quality control rejections is small (roughly 1%) in the case of radiosonde observations, but significant (approximately 10%) in the case of atmospheric motion vectors (AMVs). It is clear that, as a result of the additional perturbations, the variance of differences between observations and the background is twice as large in the members as

in the control analysis, so that first-guess quality-control limits should according to theory be doubled. However, the background quality control accounts for very few rejections of conventional wind, temperature and pressure observations in the current system, which relies on a Huber-norm method (Tavolato and Isaksen, 2009). In this method, the probability density function (p.d.f.) of observation error is represented as a combination of Gaussian and exponential functions. Since, as discussed in Chapter 2, the members should use the same observation error variances as are used in the control, it is not clear how (or if) the parameters of the Huber-norm quality control should be adjusted to avoid excessive observation rejections.

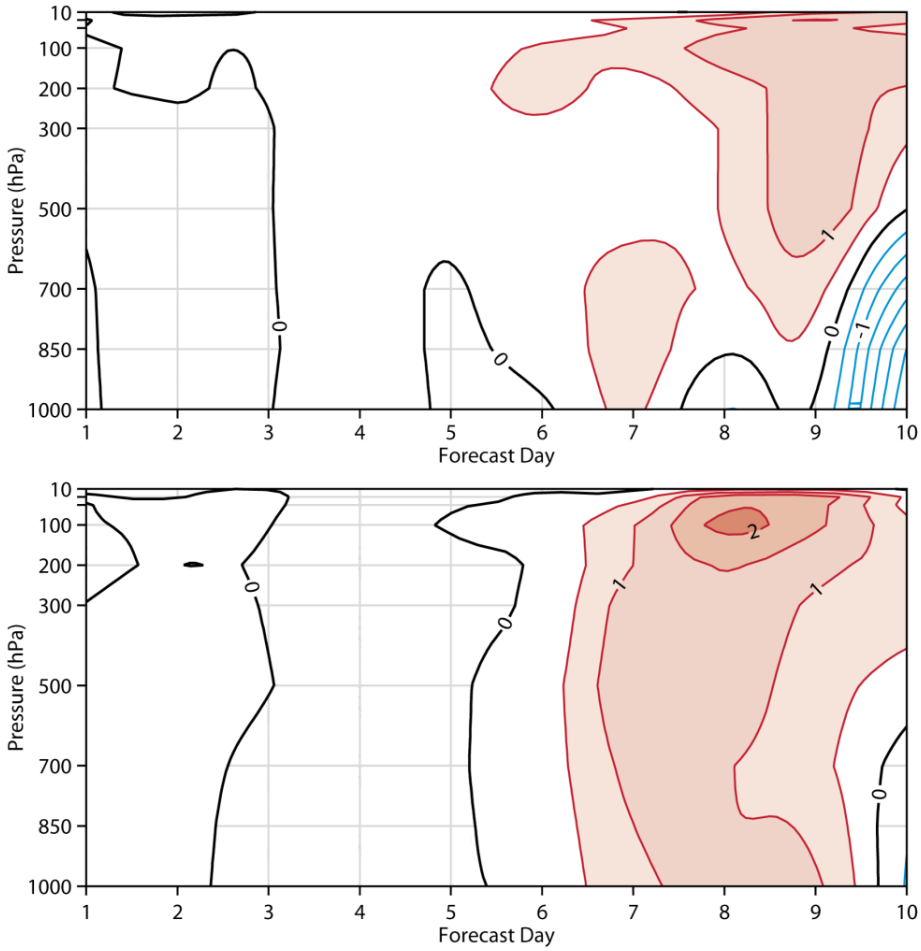

Figure 23: Forecast skill scores difference (Anomaly correlation, Geopotential height) of forecasts from deterministic 4D-Var analyses with background error variances estimated from a 20 member EDA versus using variance estimates from a 10 member EDA. Scores are averaged over the period 20080807–20080916. Positive differences contoured in red, negative in blue (contour interval=0.5). Top panel: Northern Hemisphere extratropics. Bottom panel: Southern Hemisphere extratropics.

7 Summary and plans

This paper has shown that the EDA is a soundly based approach that provides analysis and short-range forecast error uncertainty estimates of the underlying assimilation system. EDA is computationally expensive, so compromises have to be taken with respect to ensemble size and resolution. The paper has discussed these aspects extensively. The conclusions are that a simpler lower resolution inner-loop analysis combined with an outer-loop resolution slightly higher than the inner-loop resolution of the high-resolution 4D-Var system is sufficient to provide useful flow-dependent background error variance estimates. Additional computational resources are most likely better spent on increasing the ensemble size. The operational version implemented in June 2010 use T95/T159 inner-loop, T399 outer-loop/forecast resolution and 10 ensemble members. The same arguments seem to be valid for the use of EDA in the EPS. In the present implementation only the balanced part of the background error core variables are flow-dependent, via the vorticity field. Further investigations will be made, along the lines pursued at Météo-France, to introduce EDA based flow-dependence for humidity, ozone and unbalanced analysis variables.

In this paper we have documented the ability of the EDA to improve the EPS by incorporating knowledge of analysis uncertainty into the initial perturbations. Experimental results also indicate that improvements in the skill of the operational deterministic 4D-Var are to be realized through the use of EDA-derived, flow-dependent variances of background error. In addition to these improvements in forecast skill, we expect the estimates of analysis uncertainty provided by the EDA to become a valuable diagnostic output of the Centre's data assimilation system. They will also allow the development of new diagnostics (e.g. of EPS skill) that take analysis uncertainty into account.

The EDA, although already useful, is undergoing rapid development. In the short term, we expect to take account of horizontal correlations for radiance observation errors, to use the combined SPPT/SPBS model error formulation, and to improve the perturbations applied to the surface observations and parameters. Further investigations will be performed to assess the benefit of more ensemble members, various model error representations and different assimilation configurations. These investigations will have to account for likely future computer architectures, which may favour increased ensemble size over higher resolution. In general it is a clear advantage for EDA, compared to deterministic 4D-Var, that EDA is very scalable on today's and future computer architectures. This is one of the reasons why EDA is central in ECMWF's data assimilation strategy.

The data assimilation plans at ECMWF consist of a long-window weak-constraint 4D-Var supplemented by a high-resolution, short cut-off, most likely 6h 4D-Var assimilation system. EDA will clearly be important for the latter, as described in Chapter 6. It is also envisaged that the deterministic long-window weak-constraint could benefit from an accompanying long-window EDA to provide flow-dependent model error information and to improve the long-window trajectory.

A significant effort is ongoing to develop the use of flow-dependent background error in the deterministic 4D-Var system. We expect to replace the current system for estimating background errors with a system based on the EDA early in 2011. Correlation information derived from the EDA is expected to be incorporated into the deterministic 4D-Var in 2011-12. The best way to provide error estimates for the next re-analysis will also be investigated.

In the EPS, the possibility of using EDA analysis fields instead of 6h forecasts to provide EPS perturbations will be examined (although this places further real-time scheduling constraints on the EDA) and the use of the EDA to generate surface initial perturbations will be investigated. The implementation of the EDA and its use in the EPS has linked more closely the ECMWF approaches followed to estimate analysis and forecast uncertainties. The use in the EPS of EDA-based perturbations has introduced knowledge of analysis uncertainty into the EPS that was missing in the previous version of the EPS. The use in the EDA of the stochastic model error schemes that have been used for the past decade in the EPS provides a unified, consistent approach to the simulation of model uncertainties in the ECMWF analysis and forecast systems.

Further, closer links between the EDA and the EPS should lead to further improvements of both systems in the future. Integration of the methodologies used in the EPS and EDA is a long-term goal. Consistency in the representation of model error by the two systems is an important step towards this goal. In the future, further convergence of the EDA and EPS may be possible, for example by adopting similar resolutions and/or ensemble sizes. The current need to add singular vector perturbations to those provided by the EDA to define the EPS initial perturbations can be seen as a sign that the current initial and model error perturbations are not fully realistic. It is hoped that improvements in the representation of model error, and the use in 4D-Var of EDA-derived background error covariances, may eventually eliminate the need for these additional perturbations. It is also expected that the EDA will benefit from synergy with the recently started Ensemble Kalman Filter (EnKF) activities at ECMWF. This will primarily be in the areas of model error design; to compare the EDA perturbation calibrations with EnkF inflation factors; and to compare the estimated benefit of flow-dependence for reduced and full observing system impact studies in 4D-Var with EDA and EnKF.

References

Barker, D. M., 1999: Var scientific development paper 25: the use of synoptic-dependent error structure in 3DVAR. UK MET Met. Office Technical Reports.

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.

Berre, L., G. Desroziers, L. Raynaud, R. Montroty and F. Gibier, 2009: Use of consistent operational ensemble variaitonal assimilaiton to estimate and diagnose analysis and background error covariances. *ECMWF Workshop on Diagnostics of data assimilation system performance*, 15 - 17 June 2009, 47-56.

Berre, L. and G, Desroziers, 2010: Filtering of background error variances and correlations by local spatial averaging: a review. *Mon. Wea. Rev.*, doi:10.1175/2010MWR31111.1.

Bishop, C.H., B. J. Etherton and S. J. Majumdar, 2001: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon. Wea. Rev.*, **129**, 420-436.

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. ECMWF Technical Memorandum, **632** (available online at: http://www.ecmwf.int/publications/). To appear in *Quart. J. Roy. Meteor. Soc.*

Bormann, N., S. Saarinen, G. Kelly and J.-N. Thépaut, 2003: The spatial structure of Observation Errors in Atmospheric Motion Vectors from Geostationary Satellite Data, *Mon. Wea. Rev.*, **131**, 706–718.

Buizza, R., and T.N. Palmer, 1995: The singular-vector structure of the atmospheric general circulation. *J. Atmos. Sci.*, **52**, 9, 1434-1456.

Buizza, R., M. Leutbecher and L. Isaksen, 2008: Potential use of an ensemble of analyses in the ECMWF Ensemble Prediction System. *Quart. J. Roy. Meteor. Soc.*, **134**, 2051-2066.

Buizza, R., M. Leutbecher, L. Isaksen and J. Haseler, 2010: The use of EDA perturbations in the EPS. *ECMWF Newsletter*, **123**, 22-27 (available online at: http://www.ecmwf.int/publications/)

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

Fisher, M. and Courtier, P., 1995: Estimating the covariance matrices of analysis and forecast error in variational data assimilation. *ECMWF RD Tech Memorandum*, **220**, pp 26 (available online at: http://www.ecmwf.int/publications/).

Fisher, M. and Andersson, E., 2001: Developments in 4D-Var and Kalman Filtering. ECMWF Tech Memorandum, **347**, pp 36 (available online at: http://www.ecmwf.int/publications/).

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

Fisher, M., Leutbecher, M. and Kelly, G. A., 2005: On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc*., **131**, 3235-3246. (Also available as ECMWF Tech. Memo 447.)

Fisher, M., 2007: The sensitivity of analysis errors to the specification of background error covariances. . *Proceedings of the ECMWF Workshop on flow-dependent aspects of data assimilation* ECMWF, pages 27–36 (available online at: http://www.ecmwf.int/publications/).

Gagnon, N., Pellerin, G., Charron, M., Houtekamer, P., Mitchell, H., Candille, G., Spacek, L., Deng, X.-X., He, B., and Verret, R., 2007: An update on the CMC ensemble medium-range forecast system. *Proceedings of the ECMWF 11th Workshop on Meteorological Operational Systems*, 12-16 November 2007, ECMWF, Shinfield Park, Reading, RG2-9AX, UK, 57-68 (available also from http://www.ecmwf.int/publications/library) .

Hagedorn, R., Buizza, R., Hamill, M. T., Leutbecher, M., and Palmer, T. N., 2010: Comparing TIGGE multi-model forecasts with re-forecast calibrated ECMWF ensemble forecasts. Mon. Wea. Rev., under revision.

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

Houtekamer, P. L., and Mitchell, H. L., 2005: Ensemble Kalma filtering. *Q. J. R. Meteorol. Soc*., **131**, 3269-3289.

Houtekamer, P. .L, H. L. Mitchell, and X. Deng, 2009: Model error representation in an operational ensemble Kalman filter. *Mon. Wea. Rev*., **137**, 2126-2143.

Isaksen, L., M. Fisher and J. Berner, 2007: Use of analysis ensembles in estimating flow-dependent background error variance. *Proceedings of the ECMWF Workshop on flow-dependent aspects of data assimilation*. ECMWF, pages 65–86 (available online at: http://www.ecmwf.int/publications/).

Isaksen, L., J. Haseler, R. Buizza and M. Leutbecher, 2010: The new Ensemble of Data Assimilations. *ECMWF Newsletter*, **123**, pages 17-21.

Kucukkaraca, E., and M. Fisher, 2006: Use of analysis ensembles in estimating flow-dependent background error variances. *ECMWF Technical Memorandum*, 492 (available online at: http://www.ecmwf.int/publications/).

Leutbecher, M. and T. N. Palmer, 2008: Ensemble Forecasting. *J. Comp. Phys*., **227**, 3515–3539.

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP – a comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc*., **129**, 3183-3204.

Palmer, T N, Buizza, R., Leutbecher, M., Hagedorn, R., Jung, T., Rodwell, M, Vitart, F., Berner, J., Hagel, E., Lawrence, A., Pappenberger, F., Park, Y.-Y., van Bremen, L., Gilmour, I., and Smith, L., 2007: The ECMWF Ensemble Prediction System: recent and on-going developments. A paper presented at the 36th Session of the ECMWF Scientific Advisory Committee. *ECMWF Research Department Technical Memorandum*, **540**, ECMWF, Shinfield Park, Reading RG2-9AX, UK, pp. 53 (available online at: http://www.ecmwf.int/publications/).

Palmer, T.N., R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G.J. Shutts, M. Steinheimer and A Weisheimer, 2009: Stochastic parametrization and model uncertainty. *ECMWF Research Department Technical Memorandum*, **598**, ECMWF, Shinfield Park, Reading RG2-9AX, UK, pp. 42 (available online at: http://www.ecmwf.int/publications/).

Pannekoucke, O., L, Berre, and G. Desroziers, 2007 : Filtering properties of wavelets for the local background error correlations. *Quart. J. Roy. Meteor. Soc.*, **133**, 363-379.

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

Raynaud, L., L. Berre, and G. Desroziers, 2009: Objective filtering of ensemble-based background-error variances. *Quart. J. Roy. Meteor. Soc*., **135**, 1177–1199.

Raynaud, L., L. Berre, and G. Desroziers, 2010: An extended specification of flow-dependent background-error variances in the Météo-France global 4D-Var system. *Quart. J. Roy. Meteor. Soc*., submitted.

Tan, D.G.H., E. Andersson, M. Fisher and L. Isaksen, 2007: Observing system impact assessment using a data assimilation ensemble technique: application to the ADM-Aeolus wind profiling mission. *Q. J. R. Meteorol. Soc.*, **133**, 381-390.

Tavolato, C. and L. Isaksen, 2009: Huber norm quality control in the IFS. *ECMWF Newsletter*, **122**, 27-31. (available online at: http://www.ecmwf.int/publications/)

Vialard J., F. Vitart, M. A. Balmaseda, T. N. Stockdale and D. L. T. Anderson, 2005: An ensemble generation method for seasonal forecasting with an ocean-atmosphere coupled model. *Mon. Wea. Rev.*, **133**, 441–453.

Vitart, F., Buizza, R., Alonso Balmaseda, M., Balsamo, G., Bidlot, J. R., Bonet, A., Fuentes, M., Hofstadler, A., Molteni, F., and Palmer, T. N., 2008: The new VAREPS-monthly forecasting system: a first step towards seamless prediction. *Q. J. Roy. Meteorol. Soc*., **134**, 1789-1799.

Žagar, N., Andersson, E. and Fisher, M., 2005: Balanced tropical data assimilation based on a study of equatorial waves in ECMWF short-range forecast errors. *Q. J. R. Meteorol. Soc*., **131**, 987–1011. (Also available as ECMWF Tech. Memo **437**, available at: http://www.ecmwf.int/publications/)

Appendix A

Brief review of other operational centre's activities and plans in this area

A.1 Météo-France

An ensemble variational data assimilation system has been developed for the French global operational model Arpège (Berre *et al*. 2007). It is based on the EDA method developed at ECMWF with explicit perturbations of observations (consistent with observation errors) while the background states are implicitly perturbed (due to the assimilation cycling of perturbed observations). The current ensemble is formed of six independent 4D-Var assimilations with one outer loop at truncation T107, and is run in real time at a lower and unstretched horizontal resolution compared to the deterministic system (T399L70 vs T798L70).

This ensemble has been used routinely since July 2008 to provide space- and time-varying estimates of background error variances to the deterministic operational 4D-Var assimilation, in replacement of quasi-static variances. Ensemble-based variances are currently specified in the minimization of the cost function for all analysis variables (vorticity and unbalanced parts of divergence, temperature and surface pressure) except specific humidity. An objective spatial filtering of the variances has been implemented in the EDA to remove the sampling noise (Raynaud *et al*. 2009).

Moreover, the ensemble of data assimilations has been coupled to the Arpège ensemble prediction system since December 2009 to provide initial state perturbations in combination with a set of dry singular vector perturbations.

Near future developments include the use of ensemble-based variances for specific humidity and the use of ensemble-based variances in the observation quality control check. Moreover, an online adaptive inflation of background perturbations is currently under development to account for model error within the ensemble.

Finally, the extraction of useful flow-dependent information about the correlations is under investigation following works by Fisher and Anderson (2001) and Pannekoucke *et al*. (2007) on a wavelet diagonal approach.

A.2 UK MetOffice

The hybrid Var-EnKF data assimilation system at the UK MetOffice is based on the concept of the alpha control variable (Barker 1999; Lorenc 2003). In this approach the control variables are augmented by a linear combination of the ensemble short range forecast perturbations. The coefficients of these linear combinations, which are generally space-dependent, are then constrained by the inclusion of an additional term in the 4D-Var cost function.

In the UK MetOffice implementation the ensemble perturbations are derived from the Met Office Global and Regional Ensemble Prediction System (MOGREPS) which runs 24 members at 60 Km resolution. In this EPS the analysis perturbations are computed through a linear combination of the

ensemble forecast perturbations based on the Ensemble Transform Kalman Filter (ETKF) method (Bishop *et al.* 2001).

A.3 Meteorological Service of Canada (MSC)

An Ensemble Kalman filter has been running operationally at the Meteorological Service of Canada (MSC) since January 1992 (Houtekamer *et al*. 2005), and used to generate the initial conditions of the MSC Ensemble Prediction System. The latest version of the MSC EnKF (operational since 2007) has 96 members (Houtekamer *et al.* 2009) and includes different schemes designed to simulate the effect of model error. The latest modification of the model error strategy was implemented in 2007, when a perturbed tendency perturbation scheme and a stochastic back-scatter scheme were introduced (Gagnon et al. 2007, Houtekamer *et al.* 2009). Since then, the MSC-EPS has been using four different methods to simulate model error: the use of different parameterization schemes in each ensemble member, a stochastic tendency perturbation scheme, a stochastic backscatter scheme and a parameterized system error scheme. This latter scheme (Houtekamer and Hitchell 2005) simulates the effect of model error sources that degrade the quality of the initial conditions and that are not taken into account by the MSC procedure used to generate the ensemble initial conditions, by adding a random perturbation field is added to the initial conditions. The 96 members of the MSc EnKF are configured in 4 batches of 24 members: the assimilation of observations in each member of one batch is based on error statistics computed using the 72 members of the other three.

Appendix B

Construction of the EDA-SVINI EPS initial perturbations using a combination of EDA-based and SV-based perturbations

The initial conditions are defined by adding to the unperturbed analysis a linear combination of initial-time and evolved SVs. Consider an EPS starting at day/time DDHH. In order to optimize the use of computational resources, the SVs are computed along a forecast trajectory, defined by the 6-to-54 hour forecast started at (DDHH-6), i.e. from a 6-hour earlier analysis. The evolved SVs are the t+54h SVs started at 2-days earlier. The initial perturbations are symmetric, with the EPS even members having the opposite perturbation of the odd members.

In the old EVO-SVINI configuration, the j-th initial conditions were defined by adding to the unperturbed analysis $a_0(d)$, defined by interpolating the ECMWF operational analysis to the EPS resolution, a linear combination of initial-time and evolved SVs:

(B.1.a) $$a_j(d) = a_0(d) + \delta SV_j(d)$$

(B.1.b) $$\delta SV_j(d) = \sum_{s=1}^{S} \sum_{k=1}^{N_{SV,s}} [\beta_{j,k,s} SV_{k,s}(d-48h, 48h) + \alpha_{j,k,s} SV_{k,s}(d,0)]$$

The δSV_j perturbations were symmetric: the even ones had the opposite sign of the odd ones (e.g. $\delta SV_2 \equiv -\delta SV_1$). The $(\alpha_{j,k,s}, \beta_{j,k,s})$ coefficients determined the linear combination (and the amplitude) of the SVs, where s=1,S denotes the different areas for which SVs are computed (s=1 for NH, s=2 for SH, and s≥3 for the tropical areas, for which no evolved SVs are used). The $(\alpha_{j,k,s}, \beta_{j,k,s})$ coefficients were computed following a Gaussian sampling method: over the extratropics (NH and SH) using the 50 leading initial-time and the leading 50 final-time SVs, and over the tropics using the 5 leading initial-time SVs for each tropical target area. The initial-time SVs $SV_{k,s}(d,0)$ are the SVs valid for day d, at step 0, computed along a forecast trajectory started from a 6-hour earlier analysis. The evolved SVs were the +48h SVs started at 2-days earlier on day (d-48h).

In the new EDA-SVINI EPS, the *j*-th initial conditions are defined by adding to the unperturbed analysis an EDA-based perturbation and a linear combination of initial-time SVs:

(B.2.a) $$a_j(d) = a_0(d) + \delta PA_j(d) + \delta SV_j(d)$$

(B.2.b) $$\delta PA_{2k-1}(d) = [\, fg_{MOD(k,N_{EDA})}(d-6h, 6h) - fg_0(d-6h, 6h)]$$
$$\delta PA_{2k}(d) = -\delta PA_{2k-1}, k = 1,..., N_{pf}$$

(B.2.c) $$\delta SV_j(d) = \sum_{s=1}^{S} \sum_{k=1}^{N_{SV,S}} [\alpha_{j,k,s} SV_{k,s}(d, 0, -6h)]$$

Symmetric δSV_j perturbations are still used, but in the new EPS they are generated using the initial-time SVs only, with the $(\alpha_{j,k,s})$ coefficients determined as before. Symmetric EDA-based perturbations are also used, with the odd ones having the opposite sign of the even ones (e.g. $\delta PA_2 \equiv -\delta PA_1$). EDA-based initial perturbations are defined as the difference between the a perturbed first-guess $fg_j(d-6h, 6h)$ (i.e. the +6h forecasts started 6 hours earlier from the *j*-th perturbed analysis) and the un-perturbed first-guess $fg_0(d-6h, 6h)$ (started from the un-perturbed EDA member). In Eq. (B.2.b) the modulo function appears as the number of EPS members could be larger than 2*NEDA, which is the number of symmetric EDA-based perturbations that can be generated with NEDA perturbed analyses. Since there are only 10 EDA perturbed members in the new EDA-SVINI EPS and 50 EPS perturbed members, some EPS members share the same EDA-based perturbations since, e.g., $\delta PA_{21} \equiv \delta PA_1$, $\delta PA_{22} \equiv \delta PA_2 = -\delta PA_1$, and so forth.

Thus by definition, EPS members 1 and 2 have opposite perturbations:

(B.3)
$$a_1(d) = a_0(d) + \delta PA_1(d) + \delta SV_1(d)$$
$$a_2(d) = a_0(d) + \delta PA_2(d) + \delta SV_2(d) = a_0(d) - \delta PA_1(d) - \delta SV_1(d)$$

and so forth till member 20. Then EPS members 21 and 22 use

(B.4)
$$a_{21}(d) = a_0(d) + \delta PA_{21}(d) + \delta SV_{21}(d) = a_0(d) + \delta PA_1(d) + \delta SV_{21}(d)$$
$$a_{22}(d) = a_0(d) + \delta PA_{22}(d) + \delta SV_{22}(d) = a_0(d) - \delta PA_1(d) - \delta SV_{21}(d)$$

and so forth till member 40, and then similarly to member 50.