TECHNICAL MEMORANDUM

664

On the use of EDA background error variances in the ECMWF 4D-Var

Massimo Bonavita, Lars Isaksen and Elias Holm

Research Department

January 2012

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

A hybrid assimilation system which uses sample statistics from an Ensemble of Data Assimilations (EDA) to estimate background error variances has been implemented at ECMWF. We show that the new system is beneficial in terms of deterministic forecast skill provided that random and systematic errors in the estimation of variances are properly accounted for. The mechanisms through which EDA sample variances influence the deterministic analysis are clarified. An interesting aspect is that the use of flow-dependent variances alone is able to introduce a significant degree of flow-dependency in the analysis increments.

Further possible improvements and extensions to the current methodology are also discussed.

1 Introduction

An accurate specification of the statistics of background errors is a fundamental prerequisite of any effective data assimilation scheme. Methods based on an ensemble of perturbed data assimilations (EDA), originally devised in the ensemble Kalman filter context (Houtekamer et al., 1996), have been employed in the specification of background error statistics for the variational analysis at ECMWF (Fisher, 2003) and Météo-France (Belo Pereira and Berre, 2006). However, these latter applications have relied on a static, climatological representation of the background-error matrix, thus implicitly neglecting the flow-dependency of background errors.

A characteristic of EDA systems 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 schemes it is common to represent the background error covariance matrix in operator form, decomposing it in standard deviations and correlations (Fisher, 2003):

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

Where $\mathbf{T}$ is a matrix representation of the balance operator, $\boldsymbol{\Sigma}_b^{1/2}$ is the diagonal matrix of background error standard deviations and $\mathbf{C}$ is the background error correlation operator (which is modelled in wavelet space in the ECMWF implementation: Fisher, 2003). In this paper we demonstrate how the EDA system is now used at ECMWF to extract flow-dependent information for the modelling of the background error standard deviations ($\boldsymbol{\Sigma}_b^{1/2}$).

In recent years, several attempts to use error standard deviation information from small (6 or 10 member) EDA systems have been reported in the literature. Some of the early studies (Kucukkaraca and Fisher 2006, Fisher 2007, Isaksen *et al.*, 2007) have tried to directly use the sample EDA spread as a proxy of the background error standard deviation fields which are input to a deterministic 4D-Var analysis. Despite the EDA system being designed to account for the main sources of background error, it was consistently found that the ensemble considerably underestimates the magnitude of background errors. This problem was simply countered by applying a global rescaling factor of order ~2 to the

sample EDA error estimates. However the impact of using these rescaled flow-dependent error estimates was found to be generally neutral in terms of average forecast skill scores.

More recently attention has shifted to evaluating the impact of sampling noise on the EDA estimated errors. As the EDA is designed to provide a purely Monte Carlo simulation of the errors of the reference assimilation system, sampling errors are expected to be a non negligible component of any EDA statistics for operationally affordable EDA sizes (10-100 members). Algorithms to filter the EDA sampling errors in a statistically optimal way have been demonstrated by Raynaud *et al*., 2008, 2009.

The background error estimates derived from the sample ensemble statistics are themselves affected by estimation errors which have to be dealt with and possibly minimized. It is useful in this context (Berre and Desroziers, 2010) to separate the estimation error into a random and a systematic component:

$$b - b^* = (b - E[b]) + E[b - b^*] \quad (2)$$

Where b is the ensemble estimate of any component of the background error and b^* is its true value. The 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 a large ensemble size. For affordable ensemble sizes this type of error limits the information on the forecast errors that we are able to extract from the EDA sample covariances. In particular, the dimension of the ensemble directly determines the range of spatial scales that can be robustly estimated from the sample statistics (Bonavita *et al.*, 2011).

The systematic error component is the part of the estimation error that gives an indication on how much our EDA error estimates deviate on average from the truth. As such it reflects the underlying outstanding deficiencies in the representation of the main sources of uncertainties in the EDA.

Both types of errors need to be addressed in order to be able to extract realistic estimates of background errors from an EDA system. In this paper we present the solutions adopted at ECMWF in order to successfully employ flow-dependent EDA error estimates in the deterministic 4D-Var analysis cycle. In section 2 we review the issue of sampling noise filtering of EDA variances. An algorithm based on Raynaud *et al.* 2009 and Bonavita *et al.*, 2011 is currently used, but an improved version, based on previous work by Berre *et al.*, 2007 is presented and demonstrated. An easily implementable extension that relaxes the homogeneity assumption implicit in the spectral filter is also proposed. In section 3 we deal with the systematic error component in the EDA error estimates. This is an aspect of the use of EDA error estimates that has received less attention in the literature, which has been mainly focused on innovation-based diagnostics (Berre *et al*., 2007, sec. 3.4; and the review paper of Berre and Desroziers, 2010). We present diagnostics of the EDA systematic errors in terms of their statistical, spatial and temporal distribution and an effective mechanism for taking these estimation errors into account is demonstrated in the context of the ECMWF assimilation cycle. Section 4 tries to shed some light on the mechanisms through which the use of flow-dependent error standard deviations affects the deterministic analysis. These involve changes in observation usage through the observation background error checks and, more importantly, a radical change in the way

observational information in the proximity of active weather systems is used in the analysis. Section 5 illustrates the impact that the use of filtered and bias-corrected EDA error standard deviations has on the skill of the ECMWF analysis and forecast system. In section 6 we offer our conclusions and we indicate some promising avenues for further development.

It is worth noting that the hybrid approach described in this paper aims at developing an on-line calibration of the **B** matrix model as is currently implemented at ECMWF. This is different from the "alpha control variable" hybrid method currently in use or under development in other major NWP Centres (Lorenc, 2003; Buehner 2005; Wang *et al*, 2008a, 2008b). In this latter method the effective **B** matrix used in the analysis update is a linear combination of the static **B** and a flow-dependent term that comes from sampling (and usually localizing) the covariances of a forecast ensemble.

2 Random errors in EDA Variances

The ECMWF EDA system tries to represent the evolution of the deterministic 4D-Var assimilation errors through an ensemble of lower resolution 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, it is reasonable to 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 dimensions which are viable in current NWP operational contexts require the use of ad-hoc filtering techniques (Raynaud *et al.*, 2008, 2009; Berre and Desroziers, 2010) on the raw ensemble estimates.

The filtering algorithm which is currently implemented in the operational ECMWF analysis closely follows Raynaud *et al*., 2009, with the modifications reported in Bonavita *et al*., 2011, and it is briefly recalled below for reference.

For sample variances, we can define $G^e(i)$ as the random component of the error in the ensemble estimated variance at grid point i,

$$G^e(i) \equiv \widetilde{B}_{ii} - E\left[\widetilde{B}_{ii}\right] \tag{3}$$

Where $\widetilde{B}_{ii}$ is the ensemble estimated background error variance at grid point i. It can be shown (Mallat *et al.*, 1998, Raynaud *et al.*, 2009) that the covariance of this sampling noise is a simple function of the expectation of the ensemble-based error covariance matrix:

$$E\left[G^e(i)G^e(j)\right] = \frac{2}{N-1}\left(E\left[\widetilde{B}_{ij}\right]\right)^2 \tag{4}$$

where N is the ensemble size. A consequence of Eq. (4) is that the sampling noise correlation length scale is smaller on average than the background error 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. 1, where the dashed line shows the power spectrum of the 10-member EDA short range (t+9h) forecast variance and the continuous line the power spectrum of the climatological sampling noise computed from Eq. (4) for temperature at model level 49 (~200 hPa). The separation in spectral space of the signal (the EDA variance) from the random error power spectrum is clear up to about wavenumber 50. This naturally leads to the definition of a truncation wavenumber (N_{trunc}), which is a function of model variable and vertical model level, and is derived by computing an effective correlation length scale of the sampling noise and retaining the spectral components of the signal which have larger length scales. The objective filtering technique then applies the following spectral low-pass filter:

$$\rho(n) = \cos^2\left(\frac{\pi}{2}\min(n, N_{trunc})/N_{trunc}\right) \tag{5}$$

Here n is the total spectral wavenumber and N_{trunc} is the truncation wavenumber of the filter. The application of this low-pass filter in spectral space is equivalent to a weighted spatial averaging of the error variances in grid-point space, which enables the large-scale signal of interest to be extracted while filtering out the small-scale sampling noise (Buehner and Charron, 2007, apply a similar technique to the filtering of background error correlations).

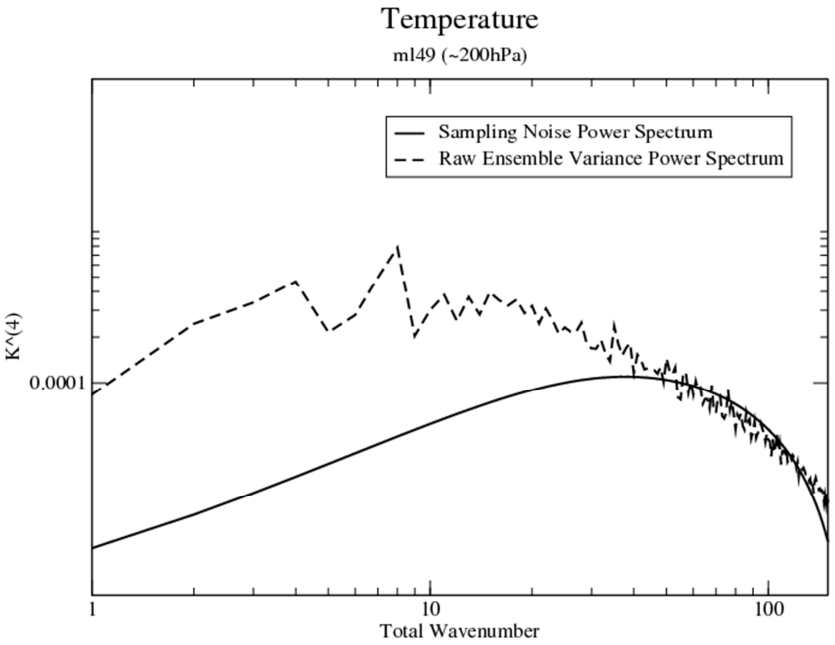

Figure 1: Power spectra of ensemble short range (t+9h) forecast variance (dashed line) and of the sampling noise (full line) associated to the estimation of temperature variances at model level 49 (~200 hPa) with a 10-member ensemble. Unit: K^4.

Fig. 2 shows an example of the raw EDA forecast standard deviation field for vorticity at model level 64 (~500 hPa) in the top panel; the objectively filtered version (middle panel); and, for comparison, the corresponding background error estimate from the "randomization" technique (Fisher and Courtier, 1995; Fisher, 2003), which was used in operations at ECMWF before the introduction of EDA based estimates. The "randomization" errors 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 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 clearer signature of the current synoptic situation than the randomization error map: EDA errors tend to be concentrated in the meteorologically active areas of sparse observational coverage.

The approach to sampling error filtering described above has proved effective and it is currently used in the operational ECMWF Integrated Forecast System (IFS). However it has some limitations: a) It is based on a sampled estimate of the error covariance matrix, which is not known accurately itself; b) it relies on a parametric representation of the filter (eq. 4) which may not be appropriate for all variables/model levels; c) it can be subject to inconsistent behaviour in response to small changes in the EDA variances' power spectra (see Bonavita *et al*., 2011). This has motivated us to develop an alternative approach, based on ideas first put forward in Berre *et al.*, 2007, which are recalled for convenience in the Appendix.

It is seen from equation (A.5) in the Appendix that the noise power spectrum of the EDA background standard deviation field can be estimated directly from different realizations of the same EDA system:

$$P_n\left(S^e\right) = 0.5\left(P_n\left(S_i - S_j\right)\right) \qquad i \neq j = 1,2,\ldots,n \tag{6}$$

To this end, two 50 member EDA were run for 45 days with the same configuration as the target pre-operational 10-member EDA (see Section 5 for further details), over both a boreal winter (January-February 2009) and summer (August-September 2008) period. In the present case, we can select 5 independent 10-member EDA realizations (i.e., a set with members 1 to 10, another with members 11 to 20, etc.). From these 5 sets we can compute 10 independent sampling noise estimates (Eq. (5)) for a 10-member EDA over each period. These estimates are then averaged and the resulting noise power spectrum is used over the winter period of EDA runs to compute a sample (N_{sample}=45*2) of 'raw' time-dependent filters (from eq. A.4):

$$\rho_n\left(i\right) = 1 - \frac{P_n\left(S^e\right)}{P_n\left(S_i\right)} \qquad \text{i=1,….,Nsample} \tag{7}$$

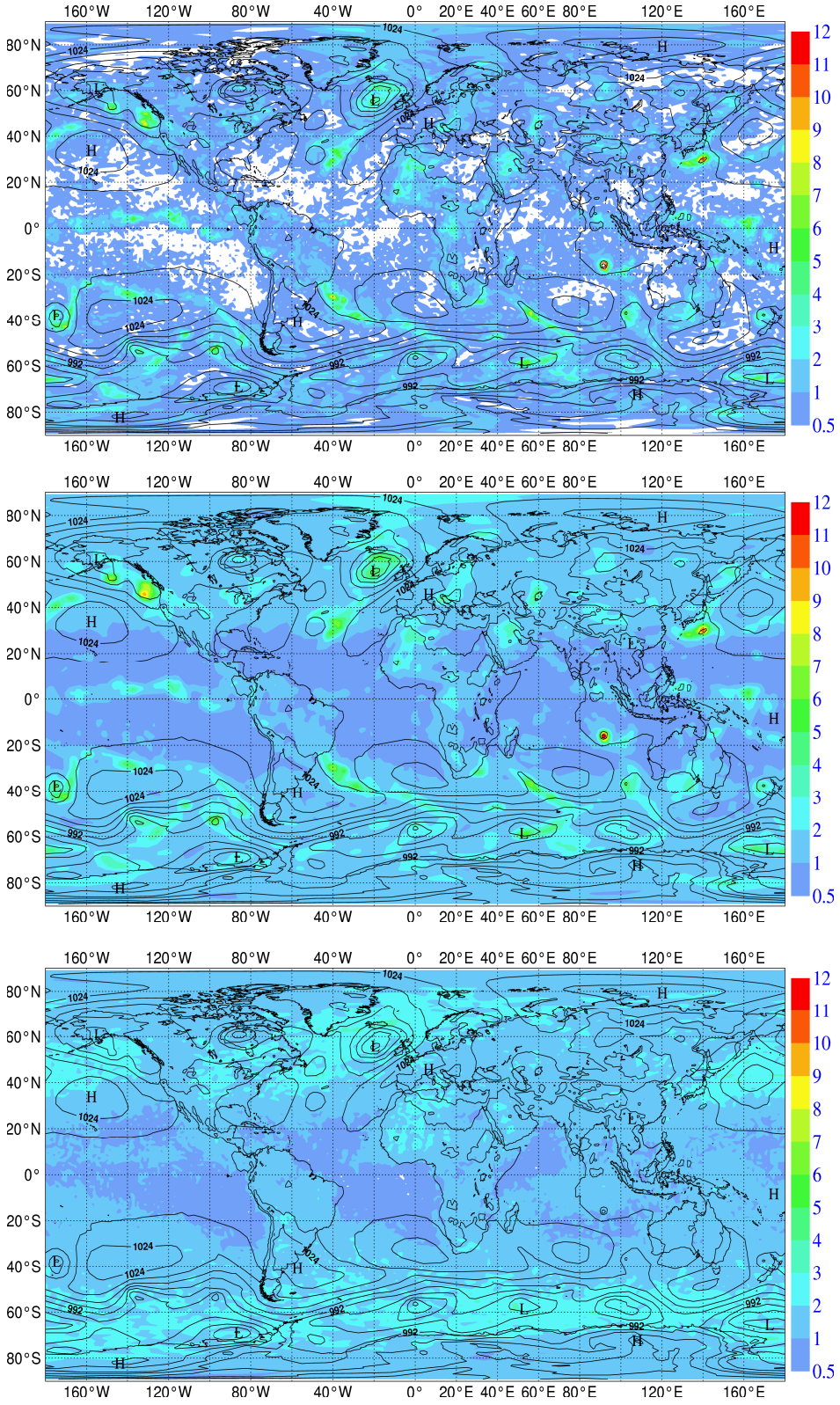

Figure 2: Standard deviations of vorticity at model level 81 (~900hPa), corresponding to 5 April 2010 at 09 UTC. Unit: 5.10^{-5} s^{-1}. Raw estimates from the 10 member ensemble (top); objectively filtered estimate (middle); background error estimate from the "randomization" technique valid at the same time. The mean sea level pressure analysis is overlaid, contour: 8 hPa.

The statistics of these time-dependent filters are shown in Fig.3, where their average values and standard deviations are shown as crosses with error bars (plots refer to vorticity at model level 39 (~100 hPa) and model level 64 (~500 hPa)). For comparison the filter based on Eq. (4), where modal values of the truncation wavenumber N_{trunc} have been used (see Bonavita *et al.*, 2011), is shown as a continuous line. It is apparent that a parametric approximation like the one expressed by Eq. (4) can not adequately represent the sample filter statistics for all model levels and variables. The proposed filter, based on the average values shown in Fig.3, relies instead entirely on the sample statistics of the raw time-dependent filters (ρ_n(i)). It should thus be able to more accurately retain the statistically significant signal of the EDA error estimates. Results of its application in assimilation experiments at full operational resolution are presented in Section 5.

Another appealing feature of the proposed spectral formulation of the noise filter is that it can easily be extended to a wavelet formulation. The wavelet transform defined in Fisher, 2003, is characterized by a convolution of the signal of interest with a set of radial basis functions $\{\psi_j(r), j=1,..K\}$ which are chosen to be spectrally band-limited, i.e. whose Legendre transforms satisfy:

$$\hat{\psi}_j(n) = 0 \qquad \text{if } n<N_{j-1} \text{ or } n>N_{j+1} \tag{8}$$

for some series of cut-off wavenumbers $N_1 ... N_K$. From our set of independent 10-member EDA realizations we can then compute a sample correlation estimate in wavelet space ($\rho_j(\lambda,\varphi)$, $j=1,..K$), i.e. a spatial correlation field for each waveband defined by the chosen wavelet decomposition. The wavelet filtered standard deviations would then result from the convolution of the raw standard deviation fields with the sample correlation estimates in wavelet space:

$$S^{filt} = \sum_{j=1}^{K} \rho_j \otimes S_j^{raw} \tag{9}$$

Where S_j^{raw} is the wavelet j component of the raw sampled EDA standard deviations.

An example of the structure of the wavelet filters based on the described methodology is presented in Fig. 4. This shows the sample filter for the wavelet band with cut-off numbers N_{j-1}=95, N_j=127. While the globally averaged filter for this wavenumber band has a low value ($\overline{\rho} = 0.143$), it is apparent from the plot that there are non-negligible geographical variations. Specifically, the sampled correlation coefficient is close to zero in the Tropics but is in the 0.3 range in large zones of the extra-Tropics. This implies that the current spectral filter over-estimates the statistical significance of the sample variances in the Tropics for these relatively short scales, and, conversely, under-estimates their significance in the extra-Tropics.

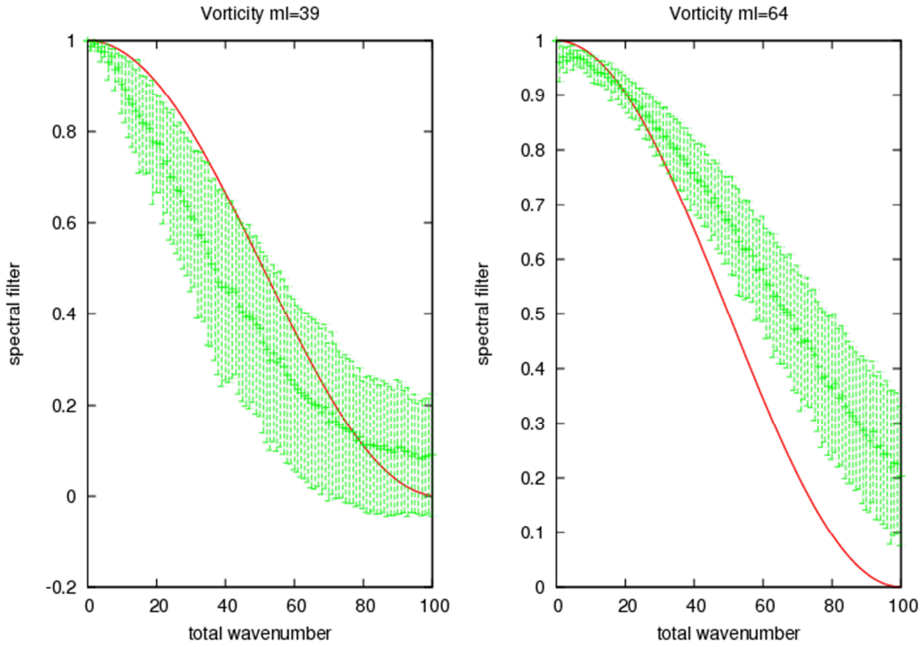

Figure 3: Average values of the statistical spectral filter with error bars (± 1 standard deviation) in green. Modal values of the Raynaud et al., 2009, filter (Eq. 4 in the text), in red. Plots are for vorticity at model level39 (~100 hPa, left) and model level 64(~500 hPa, right).

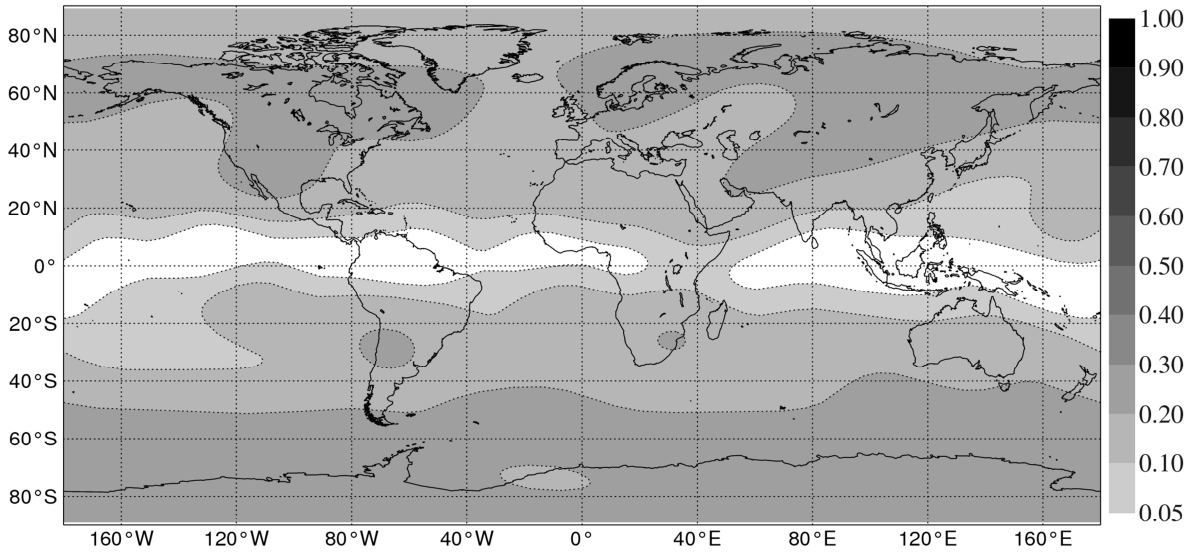

Figure 4: Average values of the wavelet filter for the wavelet band with cut-off numbers N_{j-1}=95, N_j=127 . Plot is for vorticity model level 64(~500 hPa, right).

3 Systematic errors in EDA variances

Errors in the deterministic background forecast can be considered to arise from three different sources: errors in the initial conditions, errors in the boundary conditions and errors in the model formulation. The ECMWF EDA tries to represent these sources of errors through the use of perturbed observations, perturbed SST fields and perturbed model physical tendencies, respectively. Consequently any deficiencies or approximations in the prescribed observation error statistics, SST perturbation fields and model error parameterizations, or any other missing sources of uncertainty, e.g. land surface, will cause the sampled EDA variances to be sub-optimal estimates of the analysis and background 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 true analysis/forecast errors. This state of affairs has been recognised 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 minimise the imbalance.

However the systematic errors of the EDA variances have a complex spatial and temporal structure which cannot be adequately represented by a global multiplicative inflation factor. Figure 5 shows the average over one month (August 2008) of the difference between the EDA background standard deviation and the diagnosed background mean errors (computed using the ECMWF operational analysis as a proxy for the truth) for vorticity at model level 19 (~10 hPa) and model level 64 (~500 hPa). It is clear that systematic errors present nonnegligible latitudinal and model level variations. In this example, at model level 19 the EDA appears to be statistically consistent in the Northern Hemisphere and the Tropics, while it is under-dispersive in the Southern Hemisphere, possibly because of polar vortex activity during the austral winter. On the other hand the EDA background forecast at model level 64 tends, on average, to be under-dispersive in the extra-Tropics while being closer to statistical consistency in the Tropical belt. This last feature can be interpreted as a consequence of the fact that the model error parameterization scheme employed in the EDA ("Stochastically Perturbed Parameterization Tendencies" method, SPPT, Palmer *et al.*, 2009) is more active in regions which have 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 in Fig. 6 for vorticity and temperature. Spread-error plots are constructed to represent the conditional distribution of the ensemble mean error given the ensemble standard deviation. They are obtained by dividing the sample EDA spread fields into quantiles according to their magnitude and plotting their mean value together with the mean values of the corresponding analysis/background errors. For a reliable EDA, the calibration curves (separately computed for the Northern Hemisphere, Southern Hemisphere and the Tropics) should lie on the diagonal (dashed black line) (Kolczynski *et al.*, 2009, 2011). Their distance to the diagonal shows the under/over-dispersiveness of the ensemble while their slope gives an indication of the conditional biases of the ensemble spread. 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. It is also apparent how tropical regions have, on average, different reliability characteristics than extra-tropical regions.

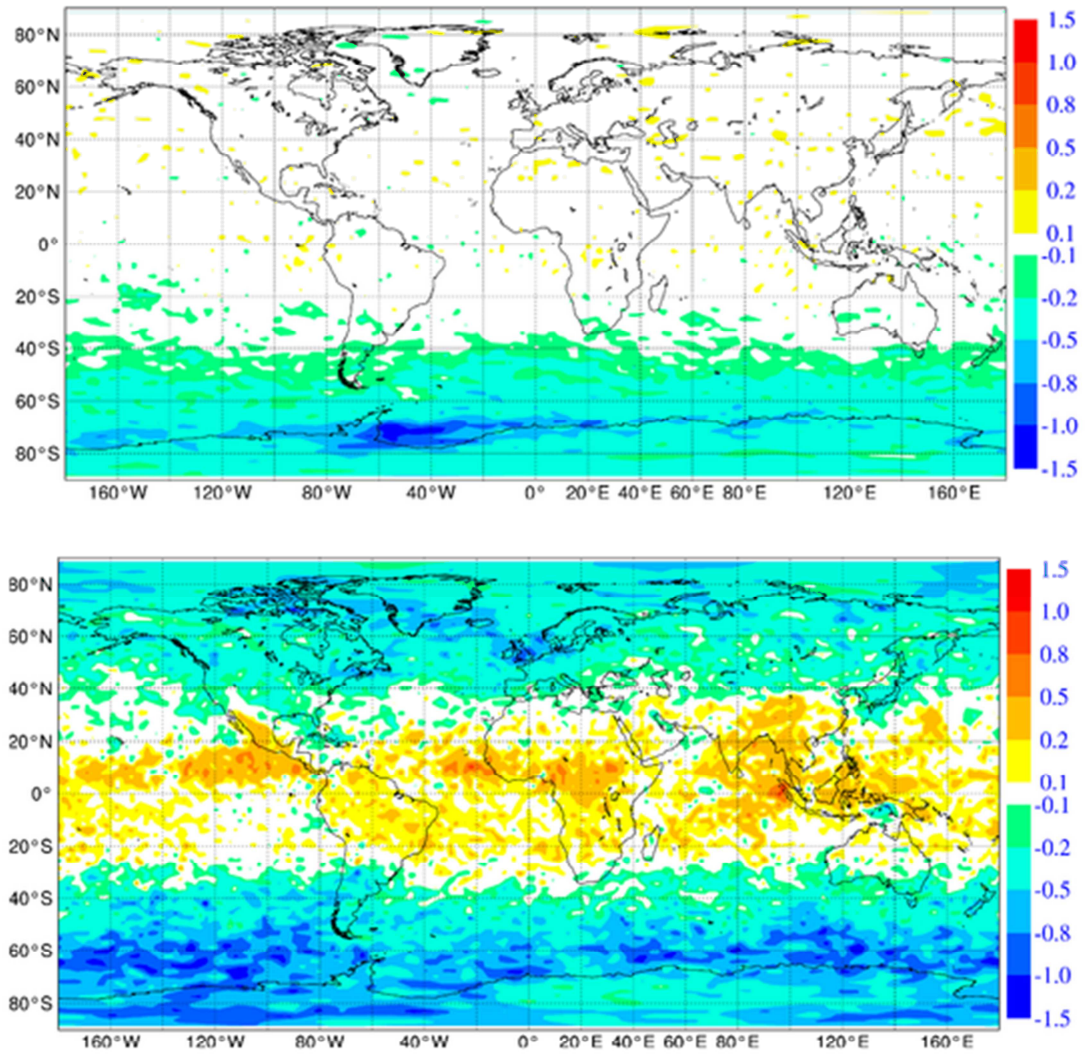

Figure 5: Difference between EDA background forecast spread and EDA mean forecast error for vorticity at model level19 (~10 hPa, top) and model level 64(~500 hPa, bottom), averaged over one month (August-September 2009). Unit: 10^{-5} s^{-1}.

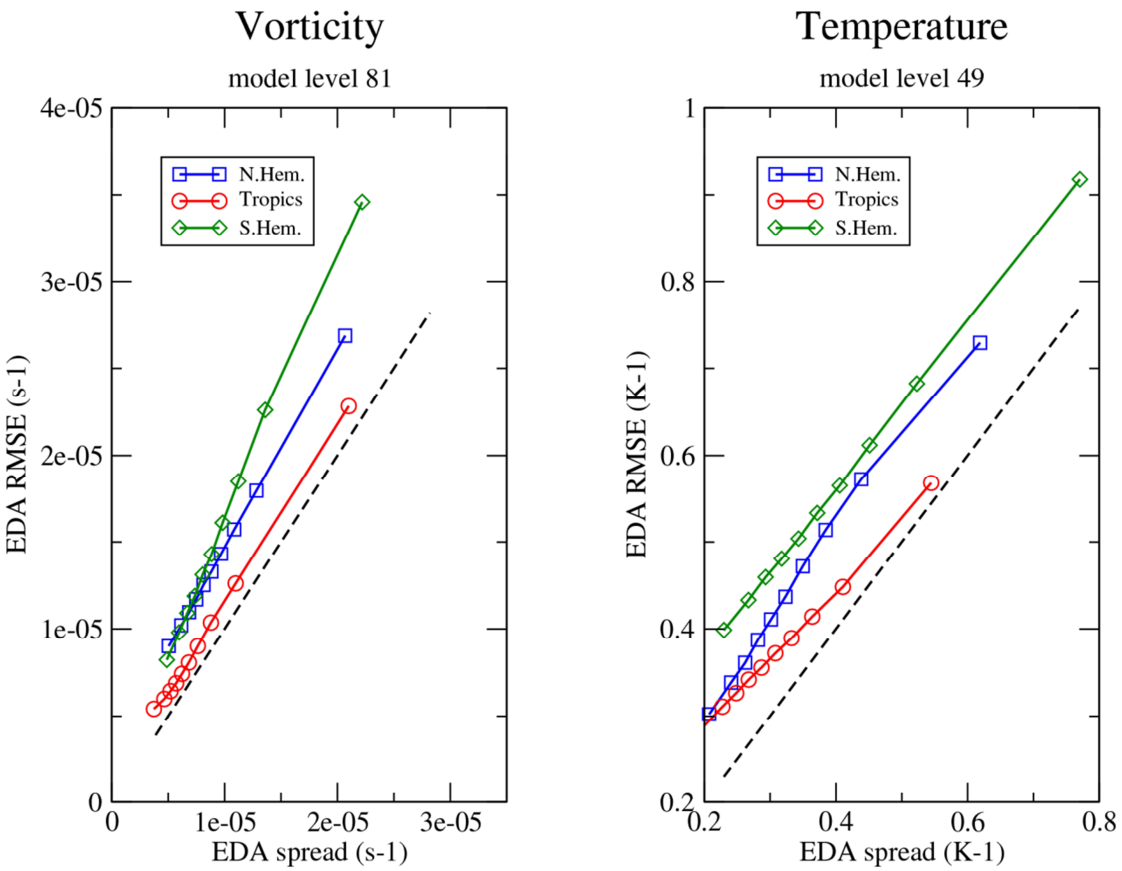

Figure 6: Spread-Error diagram for vorticity at model level 81 (~900hPa), left, and for temperature at model level 49 (~200hPa), right. Average over one month (August 2009).

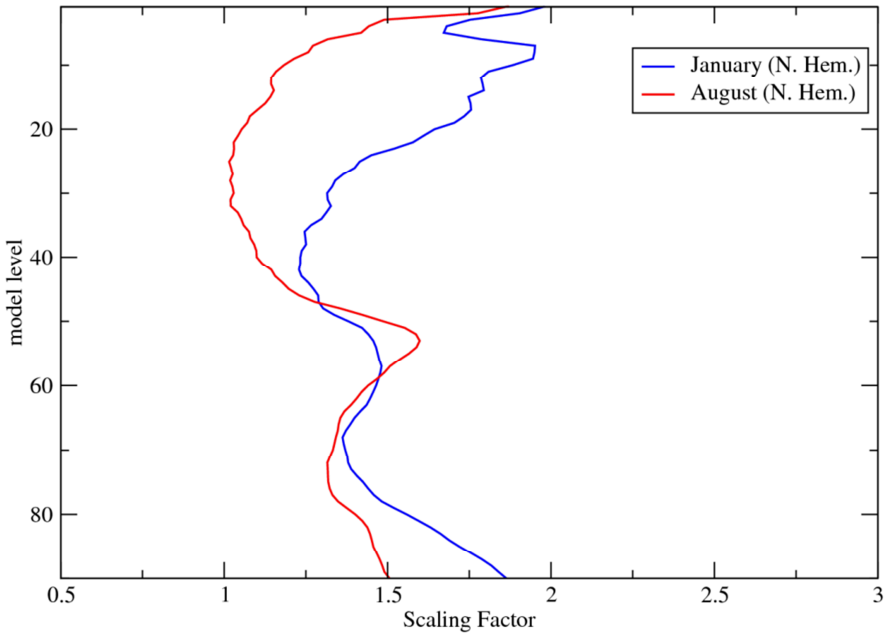

Figure 7: Scaling factors for vorticity in the Northern Hemisphere. Average over one month (January 2009, blue; August 2008, red).

The spread-error calibration coefficients do not change significantly on daily to weekly timescales. However they show an appreciable seasonal drift. In Fig. 7 the mean calibration factors over the whole EDA spread distribution, averaged over January and August, are shown for vorticity in the Northern Hemisphere. In the stratosphere the large differences shown are thought to be 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 more probably linked to model errors in the representation of a weakly coupled boundary layer and, also, to the lack of explicit perturbations of the soil parameters. Irrespective of the physical causes of the drift of the calibration factors, it is obvious that any statistical correction we might want to apply to the EDA spread must have a time-varying component.

In light of the above diagnostic findings, a rescaling step is currently applied to the raw (unfiltered) EDA estimates of background error variances. This is aimed at enforcing approximate statistical consistency between the EDA variance estimates and the diagnosed ensemble mean background errors according to (Leutbecher, 2009):

$$\frac{N}{N-1}\mathrm{E}\left[\frac{1}{N}\sum_{j=1}^{N}\left(x_j-\frac{1}{N}\sum_{j=1}^{N}x_i\right)^2\right]=\frac{N}{N+1}\mathrm{E}\left[\frac{1}{N}\sum_{j=1}^{N}x_j-y\right]^2 \qquad (10)$$

Here N is the ensemble size and y is the true state (in the present case, the operational ECMWF analysis is taken as 'truth'). Eq. (10) is enforced separately for each spread-error bin, variable and latitude band (North Hem., Tropics, South Hem.). The time-varying component is finally taken into account in the ECMWF implementation by computing a running mean of the scaling factors over the previous 5 days (i.e., the last available 10 analysis cycles) leading up to the nominal analysis date.

4 Effects of EDA variances on the 4D-Var analysis

EDA background error variances are used in two different phases of the ECMWF deterministic analysis: the background quality control of the observations and in the formulation of the initial B matrix of 4DVar (Eq. 1).

The observation quality control step is designed to discard marginal observations whose squared background departures exceed some predefined multiple of the sum of the background error and the observation error variances. Given the flow-dependency of the EDA background error variances, the new system is expected to be able to make more "intelligent" decisions in terms of observation rejection. While the overall observation usage is not significantly changed, there is some evidence of the positive impact of the different observation selection. In Fig. 8 we present forecast skill scores for geopotential at 1000 hPa in the southern extra-Tropics averaged over March-April 2010. The comparison is between an assimilation experiment which uses the EDA variances *both* in the observation first guess check and in the **B** matrix formulation (*fffn*); and an experiment which uses EDA variances for the **B** matrix *only (ffi4)*. Given the nature of the observation quality control procedure, it would certainly be possible to find specific cases where the use of EDA variances results in better observation selection decisions. However it is worth noting that a statistically significant improvement can be obtained on average, at least for some performance metrics.

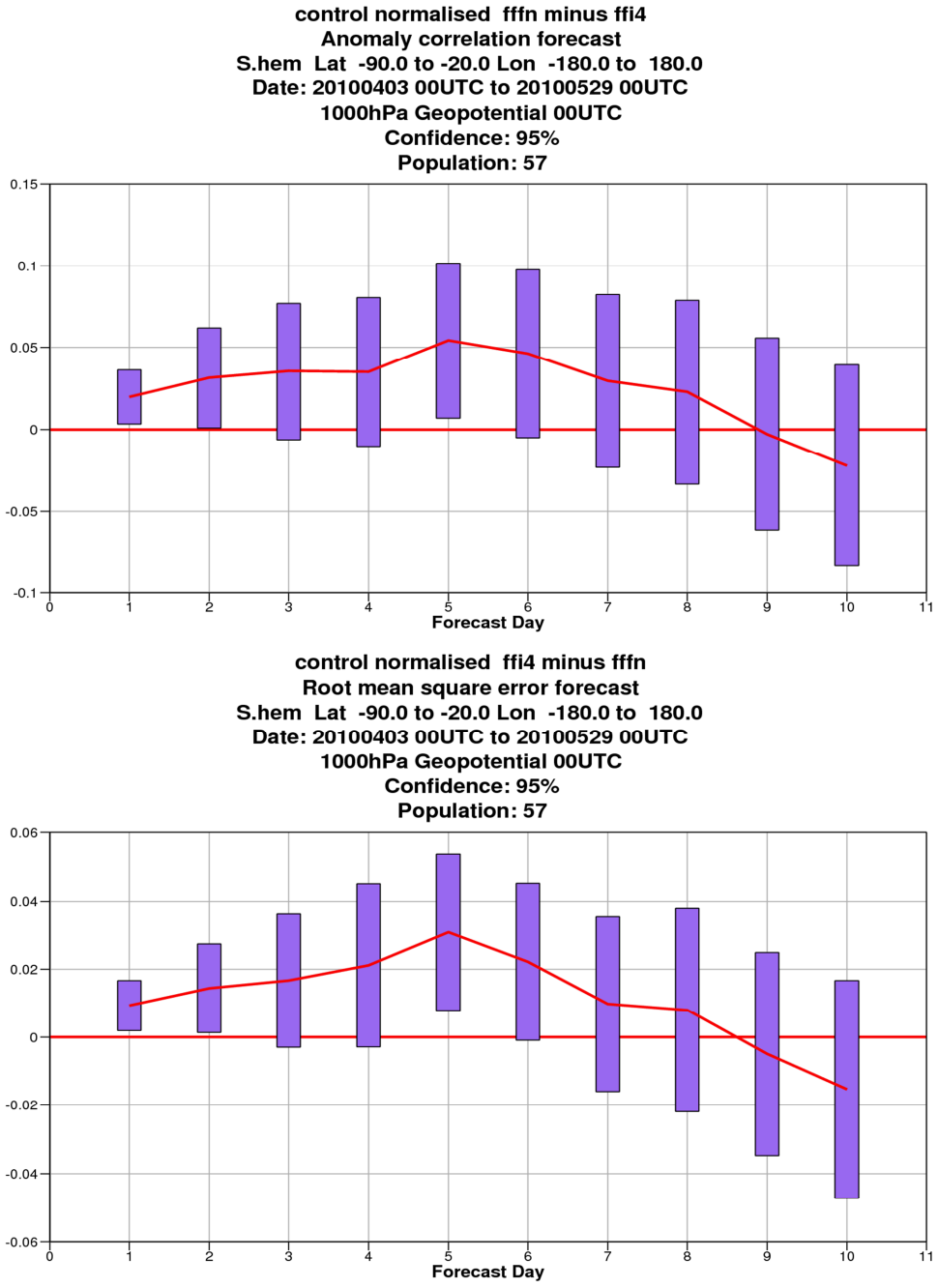

Figure 8: Forecast skill score (Anomaly Correlation, first line; Root mean square error, second line, of geopotential at 1000 hPa) of forecasts from analysis cycle using EDA variances for both observation QC and 4DVar B matrix estimation (fffn); and analysis cycle using EDA variances for 4DVar B matrix estimation only (ffi4). Scores are averaged over the period 20100402-20100529 for the Sothern Hemisphere.

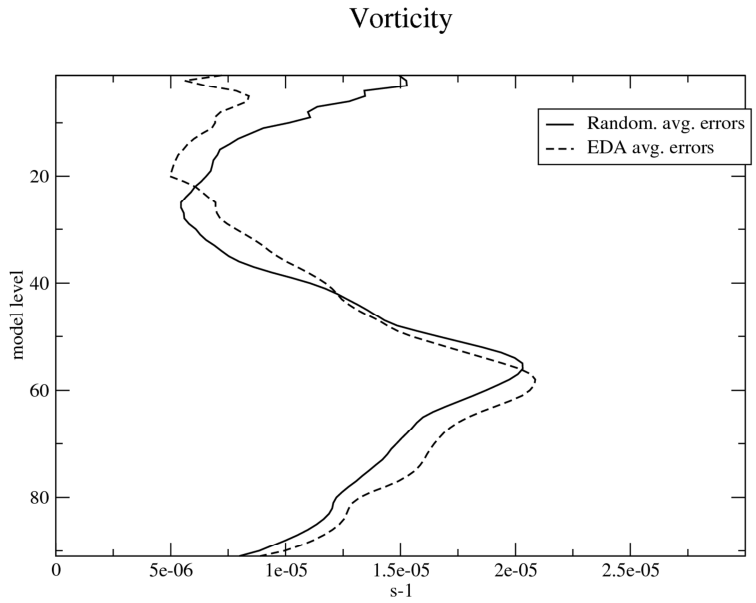

Figure 9: Vertical distribution of background error standard deviations computed with the randomization method (continuous line) and estimated from the ECMWF EDA (dashed line).

EDA variances also affect the 4D-Var analysis through changes to its **B** matrix. Their impact can be investigated through single observation experiments. The key analysis variable to consider here is vorticity, which determines the analysis of the balanced part of the flow in the ECMWF formulation of the mass-wind balance (Bouttier *et al.*, 1997). The average vertical profiles of vorticity background error standard deviations from the randomization technique and the EDA scaled variances are quite similar (Fig. 9). Their 3-dimensional spatial distributions at the beginning of the analysis time-window are however markedly different. An example is given in Fig. 10 where the first row depicts the synoptic situation over North America in terms of 500 hPa geopotential height and mean sea level pressure (MSLP) fields valid on 30 September 2009, 21 UTC; in the second row we present the EDA standard deviation of vorticity at ml 78 (~850 hPa) overlaid on the 500 hPa geopotential field; in the third row the randomization estimate of vorticity standard deviation overlaid on the 500 hPa geopotential field. The values of the diagnosed vorticity errors are broadly similar (~1-2*10^{-5} s^{-1}) but a clear difference is evident along the south-eastern US seaboard. In this region, the EDA forecasts an area of large vorticity spread corresponding to a secondary development on the eastern flank of the 500 hPa trough, while the randomization errors are approximately homogeneous. We can expect that an observation located in proximity to this region would have a different impact on the resulting analysis depending on which background error estimate we use. Fig. 11 shows the analysis increments at the beginning of the assimilation window which are generated by a single temperature observation with a departure ΔT=+1K located close to the vorticity spread maximum (34N,74W at 900 hPa). The first row presents the increments for temperature at model level 81 (~900 hPa), the second row the increments for vorticity at model level 78 (~850 hPa). The first column refers to an assimilation experiment which makes use of the randomization errors, the second column to an experiment which uses the EDA estimates of background error standard deviations.

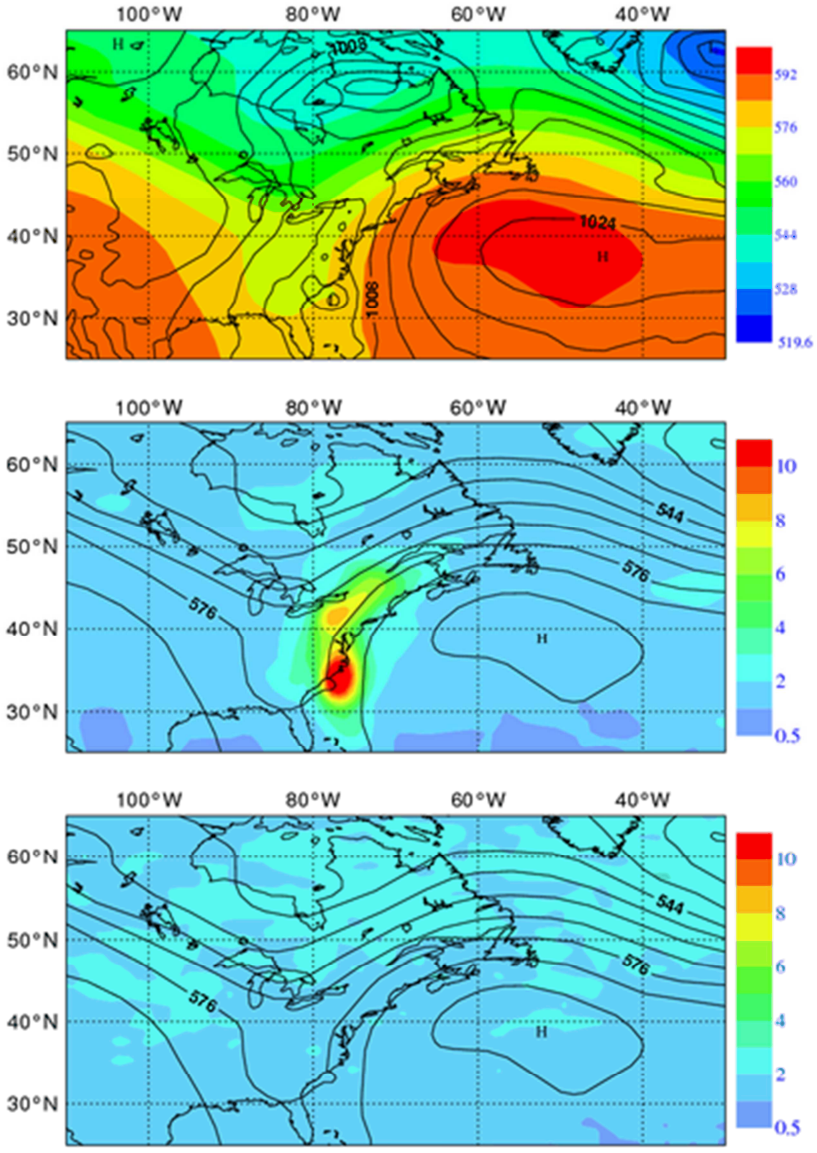

Figure 10: Geopotential at 500 hPa and MSLP background fields valid on 30 September 2009, 21 UTC (first row). Geopotential at 500 hPa and vorticity background error standard deviations estimated from the ECMWF EDA (second row). Geopotential at 500 hPa and vorticity background error standard deviations computed with the randomization method (third row).

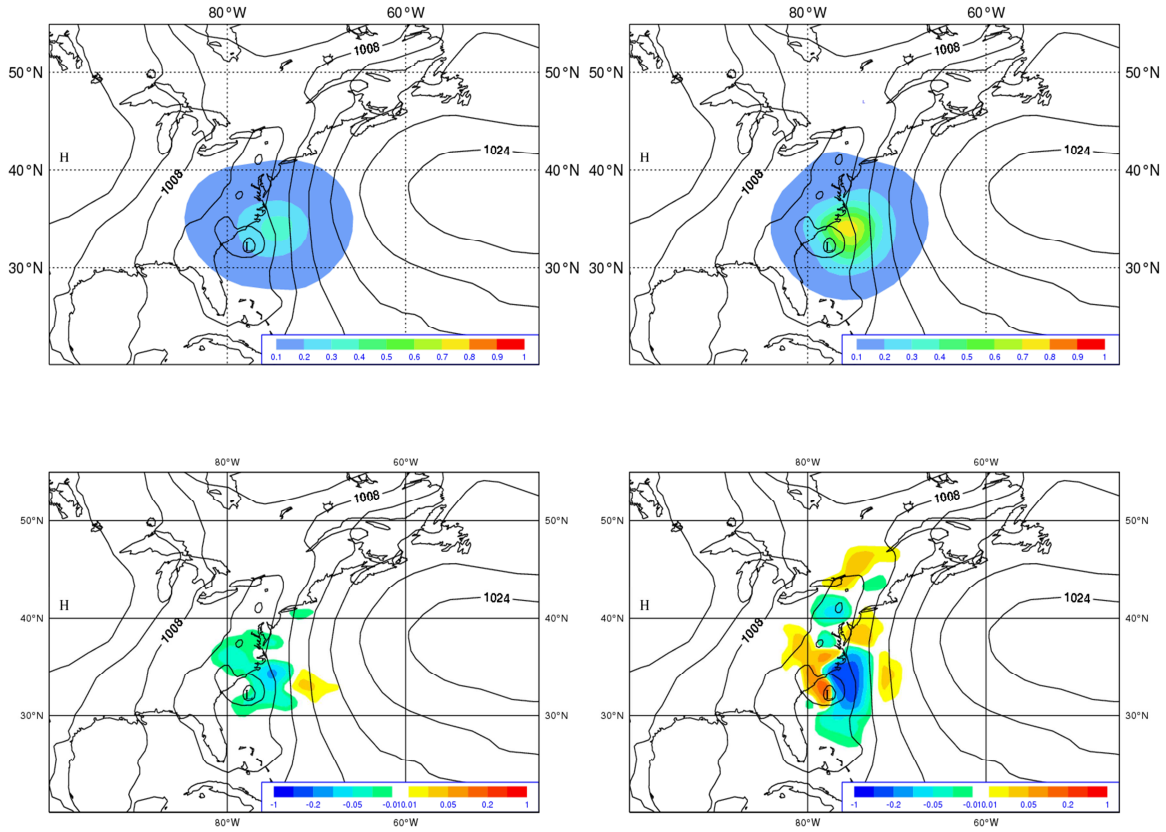

Figure 11: Single observation (observation departure ΔT=+1K at 900 hPa) analysis increments for temperature at model level 81 (first row) and vorticity at model level 78 (second row) valid on 30 September 2009, 21 UTC. The first column refers to an assimilation experiment making use of the "randomization" error estimates, the second column to an assimilation experiment making use of estimates from the ECMWF EDA. Scale intervals are in degrees for the temperature increments, in 10^{-5} s^{-1} for the vorticity increments.

Two aspects of the analysis increments derived from the two background error models are of interest. In the first place, there is a large discrepancy in the magnitude of the analysis increments: in the randomization experiment the maximum (minimum) temperature (vorticity) increment is 0.37 K (-1.30*10^{-5} s^{-1}) , respectively; in the 'EDA errors' experiment the corresponding value is 0.749 K (-5.0*10^{-5} s^{-1}). Further to this, the shape of the analysis increments is clearly affected by the structure of the vorticity error standard deviations: while the increments of the randomization experiment have a predominantly isotropic structure, the increments of the 'EDA errors' experiment have, especially for vorticity, a non-negligible state-dependent component (note their meridional elongation).

In Fig.12 we show the cross-section of the EDA standard deviation of vorticity (left panel) and the operational randomization estimate (right panel). As in the map views, the EDA points to a well defined area of errors linked to the synoptic development described above, while the randomization estimate is flat. The resulting cross sections of vorticity analysis increments (Fig. 13) also reflect the much stronger (note the different scales used in the two plots) and more localized impact of the temperature observation in the EDA experiment.

These effects can also be demonstrated when the full observing system is used.

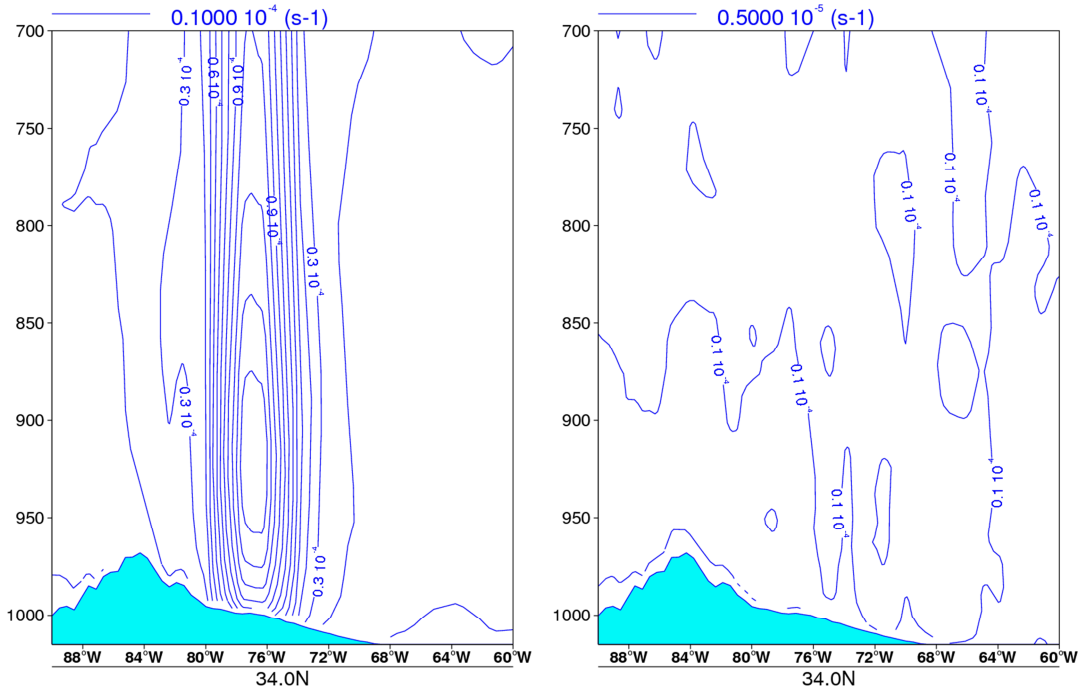

Figure 12:Zonal cross section of the vorticity background error standard deviations estimated from the ECMWF EDA (left, interval=1.$^{-5}$ s^{-1}) and from the randomization method (right, interval=0.5^{-5} s^{-1}). Vertical scale in hPa.

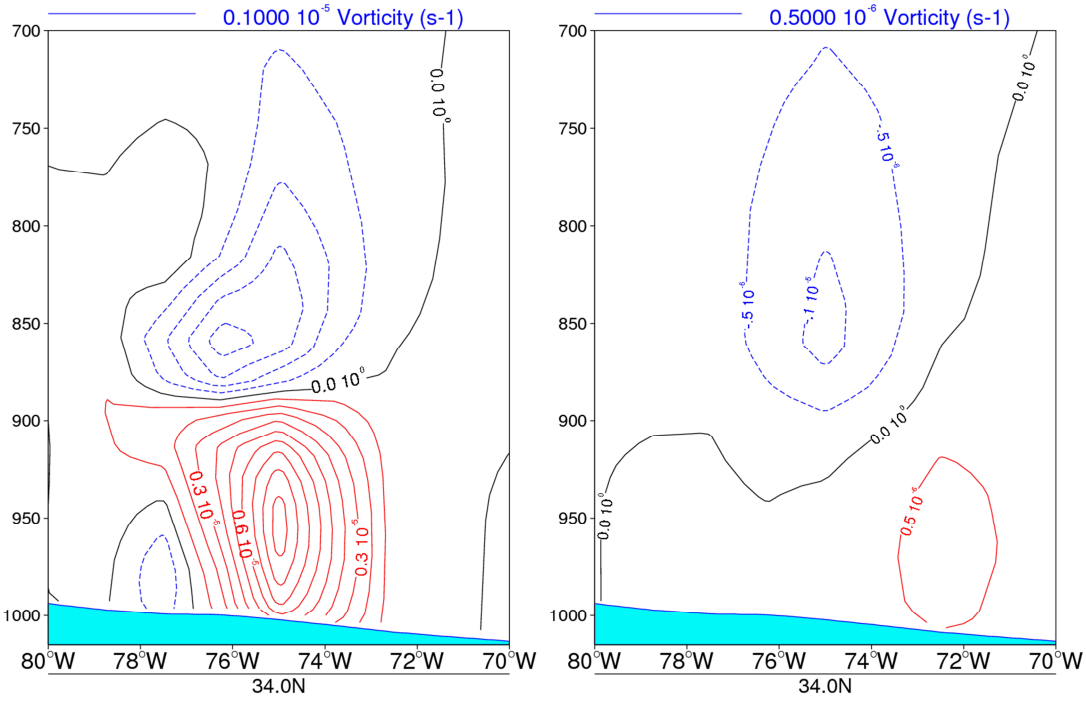

Figure 13: Zonal cross sections of analysis increments of vorticity for the single observation experiment described in the text. The left column refers to an assimilation experiment making use of error estimates from the ECMWF EDA, the right column to an experiment making use of the "randomization" error estimates. Scale intervals are in 0.1 10^{5} s^{-1} for the EDA experiment, 0.05*10^{-5} s^{-1} for the randomization experiment.*

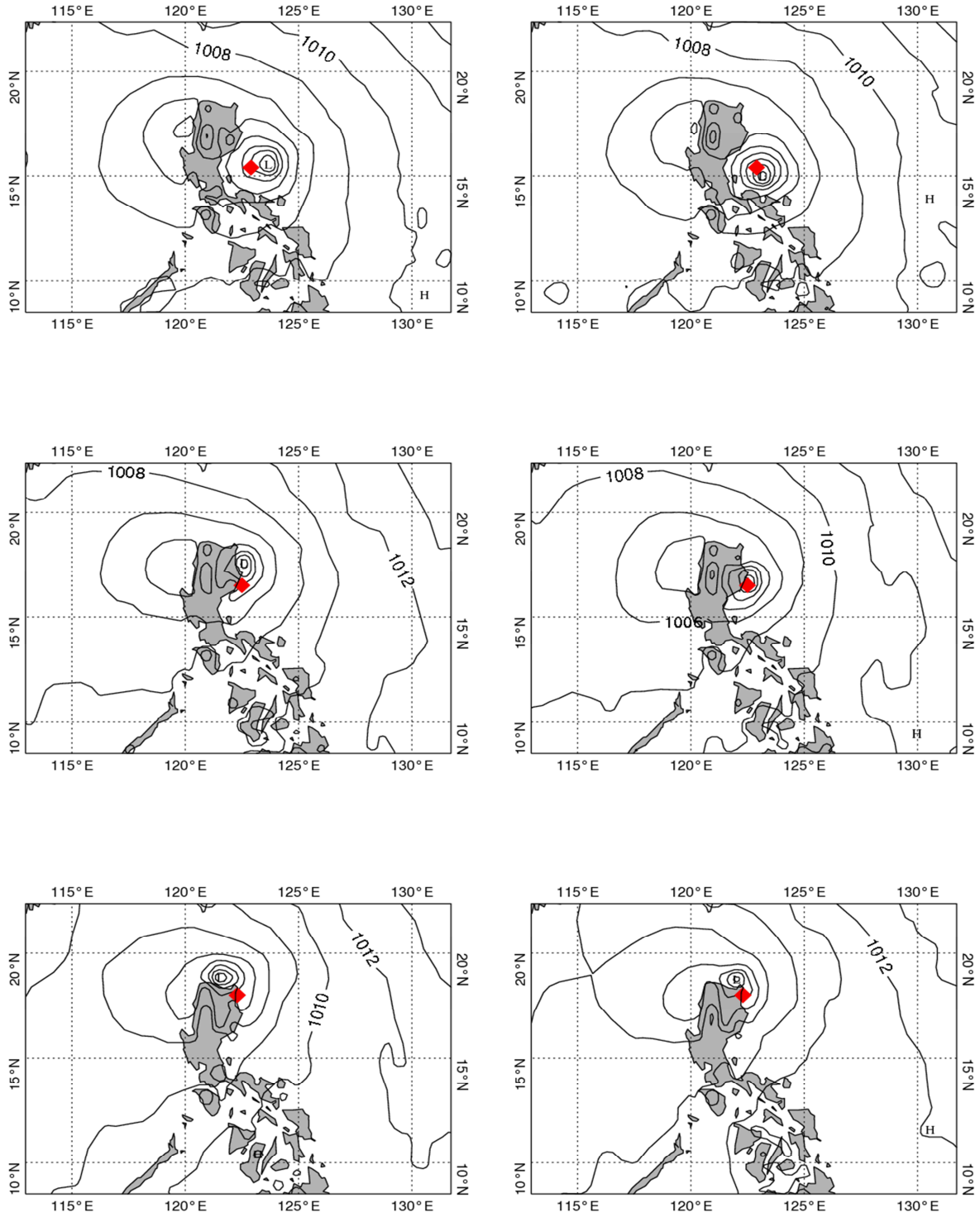

Figure 14: Analysed mean sea level pressure valid on 8 May 2011 at 12 UTC (first line); 9 May 2011 at 00 UTC (second line); 9 May 2011 at 12 UTC (third line). First column shows fields from the operational ECMWF analysis, second column from an experiment using EDA error estimates. Diamond symbol denotes independent best estimate of Storm position (see text for further details).

An interesting example is given by the tropical cyclone Aere which affected the north-eastern part of the Philippines on the 8-9 May 2011. Fig. 14 shows the evolution of the storm through the mean sea level pressure fields from the operational ECMWF analysis (which at the time was still using background error estimates computed with the randomization technique) on the left column, and from an experiment making use of the EDA error estimates (right column). The diamond symbol locates the estimated position of the storm according to the relevant operational advisories issued by the Joint Typhoon Warning Center (JTWC) of the U.S. Air Force and Navy, located at Pearl Harbour, Hawaii. The experiment which makes use of the EDA variances is clearly more accurate in analysing the cyclone position, but what is more interesting is the different use of the available observations (mainly surface pressure observations from land stations, Fig. 15). The first line of Fig. 16 shows (shaded contours) the estimated uncertainty of the background forecast for the logarithm of surface pressure and the background mean sea level pressure field (continuous isolines). The second line shows the analysis increments for the mean sea level pressure field. For both lines the first column shows fields from the operational ECMWF analysis, the second column from an experiment using EDA error estimates. One can notice that the estimated uncertainty in the operational background is much smaller than in the EDA experiment and tends to be axis-symmetric with the predicted position of the cyclone. This reflects the flow-dependent effect of the application in the current **B** matrix formulation of a linearized version of the non-linear balance equation and the omega equation dynamical balance constraint (Fisher, 2003). The effect of this is to increase the diagnosed background uncertainty in dynamically active regions, but since the linearization is computed around the background state, the error estimate will also reflect inaccuracies in the background state estimate. This is a fundamental limitation of the previous formulation of the background error variances: its flow-dependent error structures are simply due to the flow-dependent nature of the linearized balance operators, but they still originate from random samples of the climatological B matrix.

On the other hand the errors diagnosed by the EDA are, by design, constructed to estimate the real analysis errors, thus implicitly taking into account the observation network distribution and the model instabilities. In the present case they act to extrapolate the observational information from the land based stations into the more uncertain areas to the north-east of the cyclone, thus helping achieve a better positioning of the analysed storm.

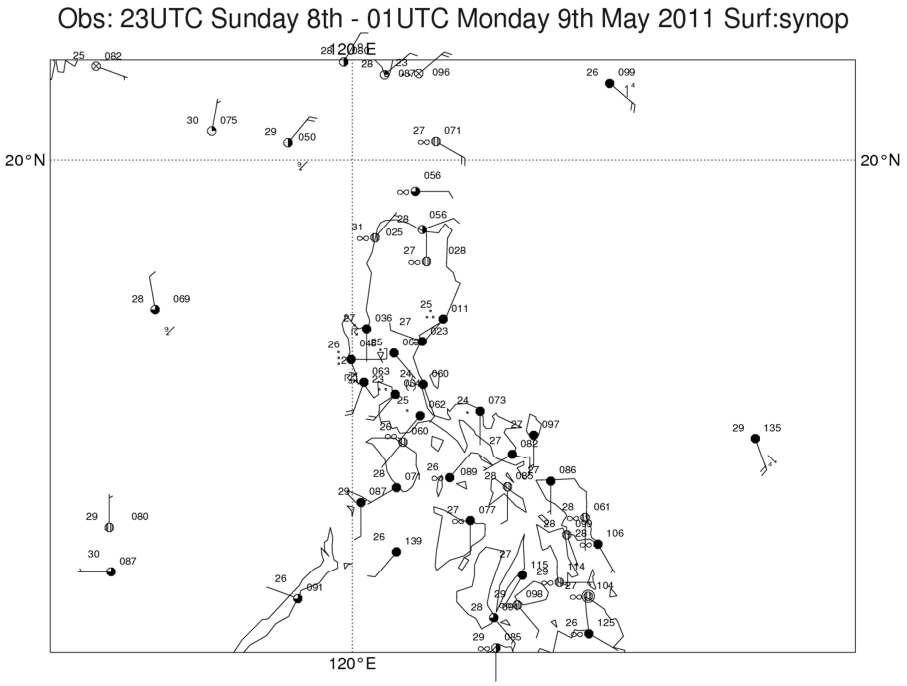

Figure 15: Available surface observations from land and ship stations on 9 May 2011 at 00 UTC;

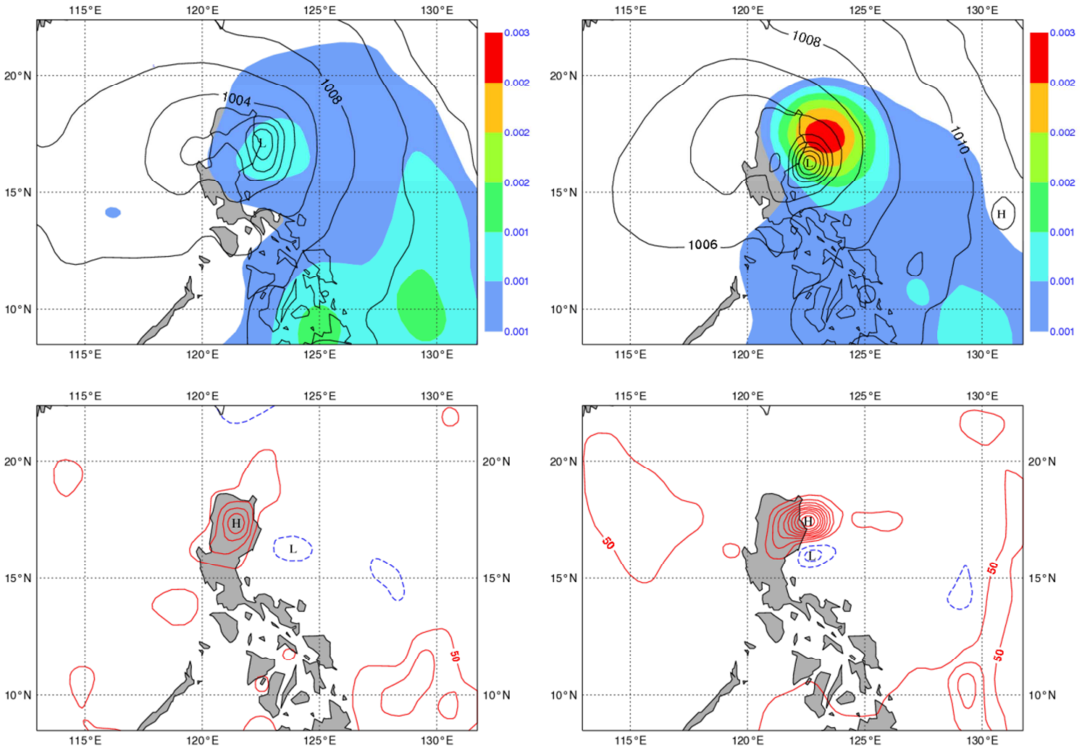

Figure 16: First line: Background mean sea level pressure forecast valid on 9 May 2011 at 00 UTC (solid line, units: hPa) superimposed on background error estimates for the logarithm of surface pressure (shaded contours). Second line: Surface pressure analysis increments valid on 9 May 2011 at 00 UTC (solid line indicate positive increments, dashed line negative increments; isolines of 50 Pa). First column shows fields from the operational ECMWF analysis cycle, second column from an experiment using EDA error estimates.

5 Assimilation experiments using EDA variances

The impact of the use of EDA error variances in the deterministic analysis has been tested in two long (~3 months) assimilation experiments over the winter (11/01/2010 to 30/03/2010) and summer (2/08/2010 to 30/10/2010) seasons. The experiments make use of the variances computed from the 10 member operational ECMWF EDA system, as described in detail by Isaksen *et al*., 2010. The sample variances are filtered and rescaled as discussed in sections 2 and 3. The filtering process applied in these experiments is the one originally proposed by Raynaud *et al.,* 2009. The EDA error variance estimates are used as proxies of the background errors both in the 4D-Var minimization step (for vorticity and the balanced part of the other control variables) and in the observation background quality control check (for all variables). The assimilation experiments are run with the ECMWF Integrated Forecasting System (IFS) cycle 36R4, run at the operational resolution (T1279L91, with 3 incremental 4D-Var loops run at T159-T255-T255 resolution). The controls differ from the experimental runs only in that they use the current quasi-static background errors estimates derived from the "randomization" technique.

Fig. 17 and 18 present vertical cross-sections of the RMS reduction of geopotential height errors at forecast lengths ranging from t+12h to t+192h for the winter EDA variance experiment (*ffg8*) against the control (*fezj*) and the boreal summer experiment (*ffge*) against its control (*0051*), respectively. Fig. 19 and 20 present the corresponding RMS error reductions for wind vector over the winter and summer period respectively. All forecasts are verified against own analysis. The improvement in forecast skill is apparent for both parameters at most forecast ranges, latitude bands and pressure levels. It is interesting to note that improvements tend to be larger in the winter hemisphere. This is consistent with the effect of EDA variances on the deterministic analysis as illustrated in section 4: In regions and times of the year where synoptic activity is stronger the impact of the use of EDA variances will be more significant.

The only statistically significant deterioration in scores affects the tropical winds in the first 24-48 hours of the forecast range. This may be caused by the spatially homogeneous nature of the spectral filter that we use. As was shown in fig. 4, the error variance field typically has inhomogeneous correlation length scales and this requires the use of a filter with both spectral and spatial resolution. A filter of this type, based on the wavelet technique (Fisher, 2003), has been described in Section 2 and it is currently being evaluated. Another possible cause of this degradation lies in the fact that the current use of EDA error estimates in the 4D-Var B matrix is limited to vorticity, while the unbalanced components of the analysis control vector are still estimated from the static, climatological B. This can introduce local imbalances in the analysed wind field, which are more evident in the Tropics where the fraction of error variance of the total wind field explained by the balanced components is smaller.

Further experimentation has also been conducted to evaluate the alternative spectral filtering methodology described in section 3. Results of two assimilation experiments run over the period 10 April 2010 – 22 June 2010, one using the Raynaud et al., 2009, methodology (experiment fffn), the other with the proposed non-parametric algorithm (experiment fha8) are shown in Fig. 21 for geopotential height forecast RMS errors. Statistically significant improvements for the experiment

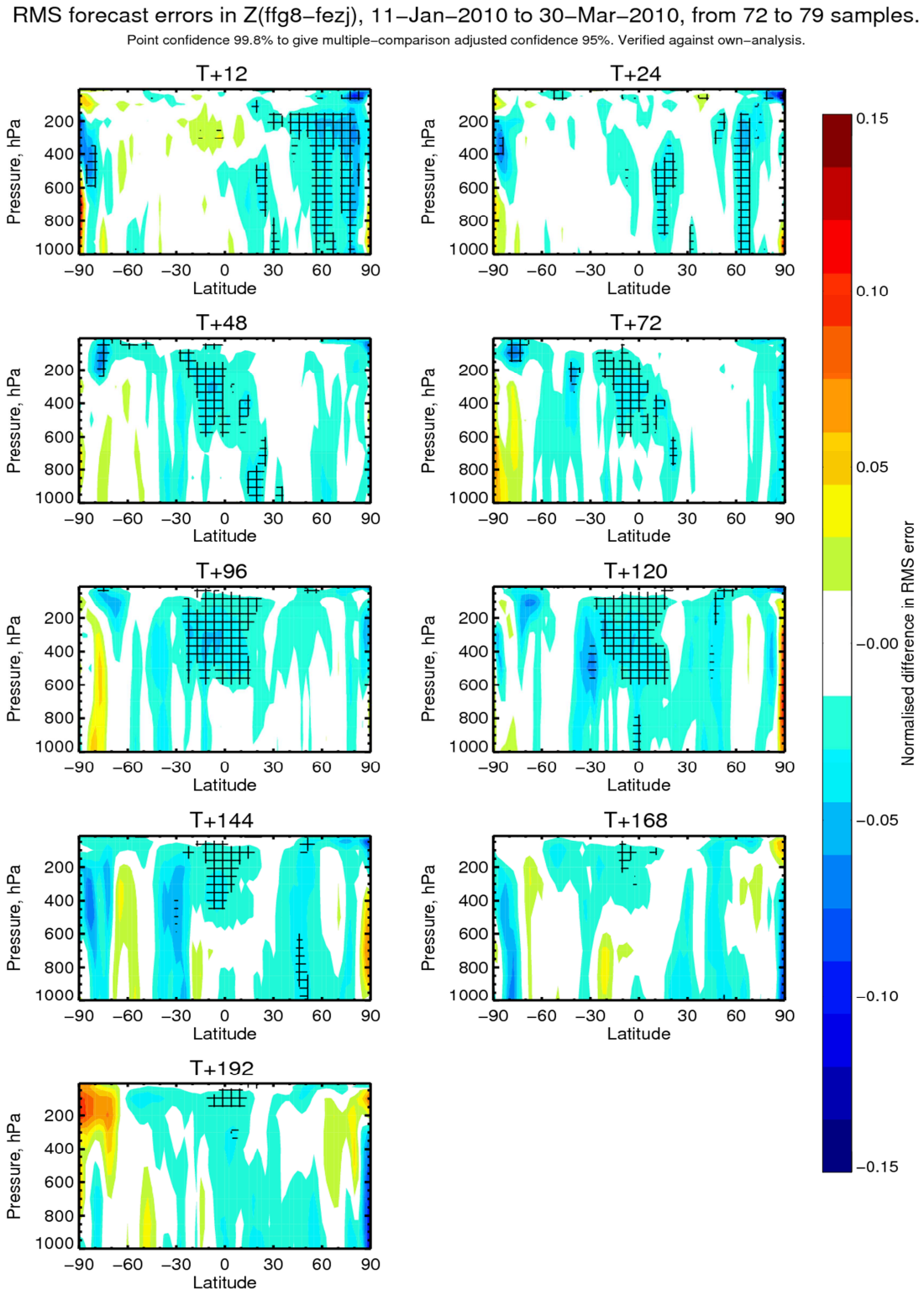

Figure 17: Meridional averages of Root Mean Squared Error reduction of Geopotential forecasts from deterministic 4D-Var analyses with background error variances estimated from a 10 member EDA (experiment ffg8) versus using randomization errors (experiment fezj). Forecasts are verified against own analysis. Scores are averaged over the period 2010111-20100330. Crosses indicate statistical reliability of the results at the 95% confidence level.

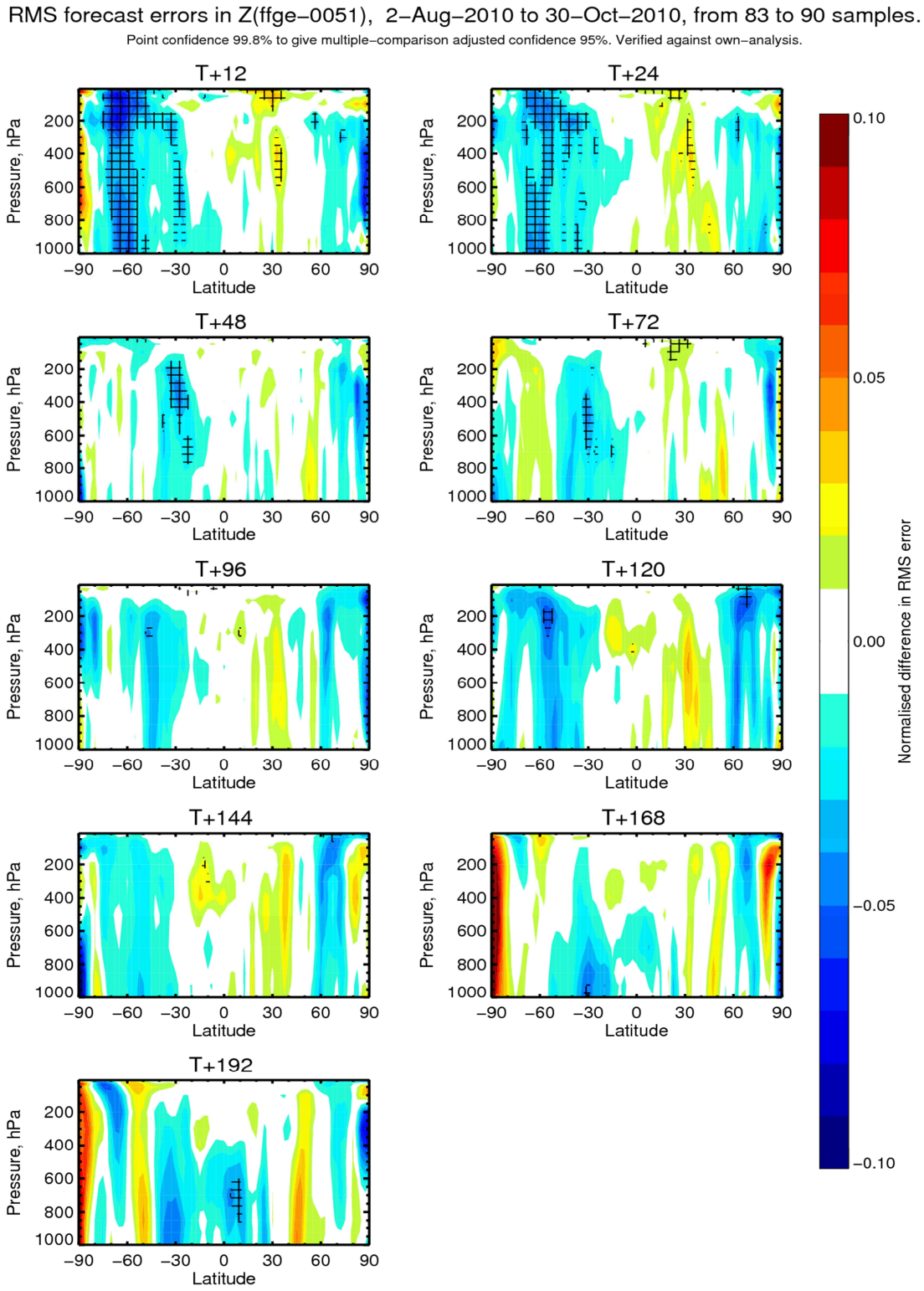

Figure 18: Meridional averages of Root Mean Squared Error reduction of Geopotential forecasts from deterministic 4D-Var analyses with background error variances estimated from a 10 member EDA (experiment ffge) versus using randomization errors (experiment 0051). Forecasts are verified against own analysis. Scores are averaged over the period 2010802-20101030. Crosses indicate statistical reliability of the results at the 95% confidence level.

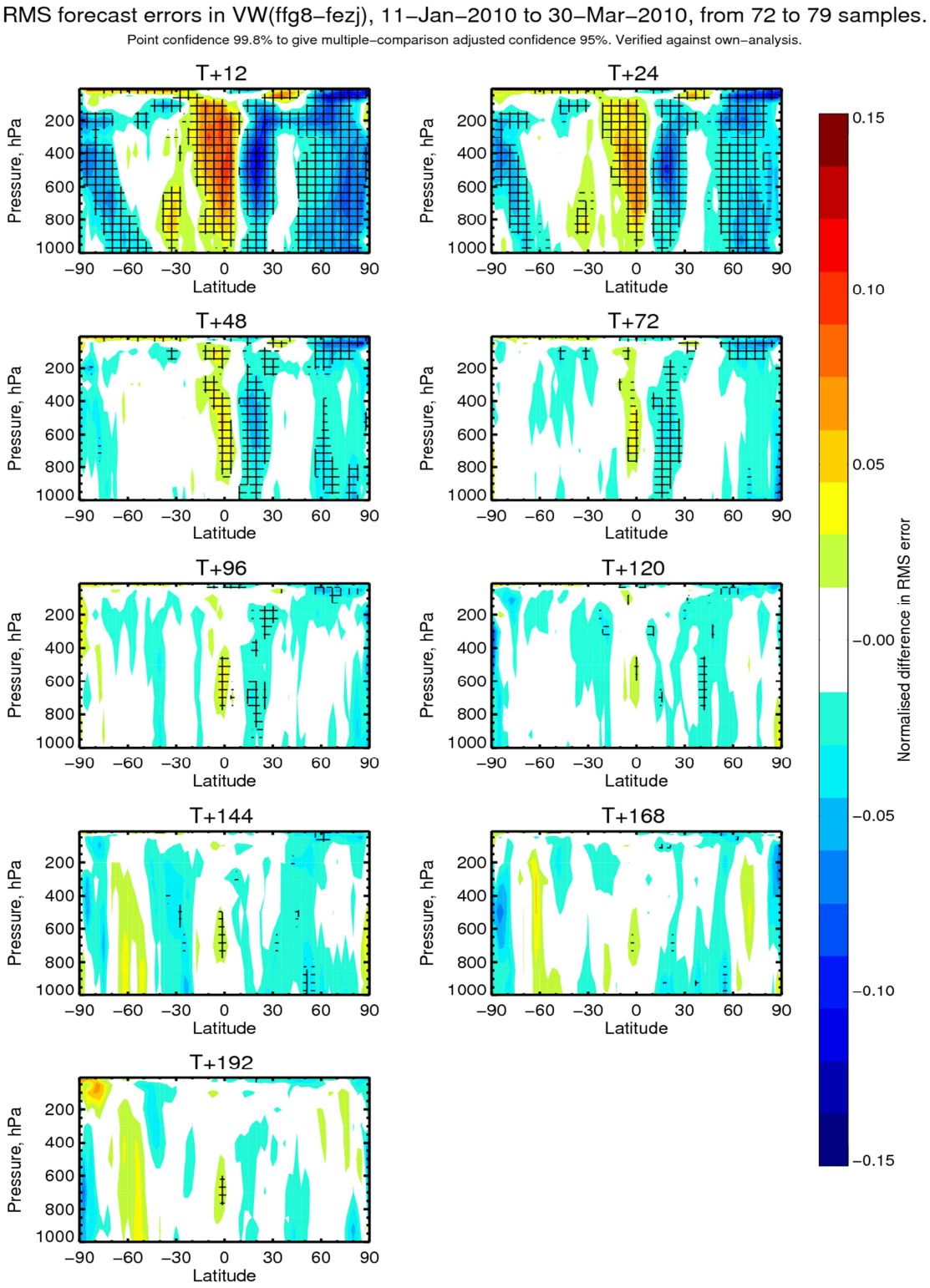

Figure 19: Meridional averages of Root Mean Squared Error reduction of Wind Vector forecasts from deterministic 4D-Var analyses with background error variances estimated from a 10 member EDA (experiment ffg8) versus using randomization errors (experiment fezj). Foracasts are verified against own analysis. Scores are averaged over the period 2010111-20100330. Crosses indicate statistical reliability of the results at the 95% confidence level.

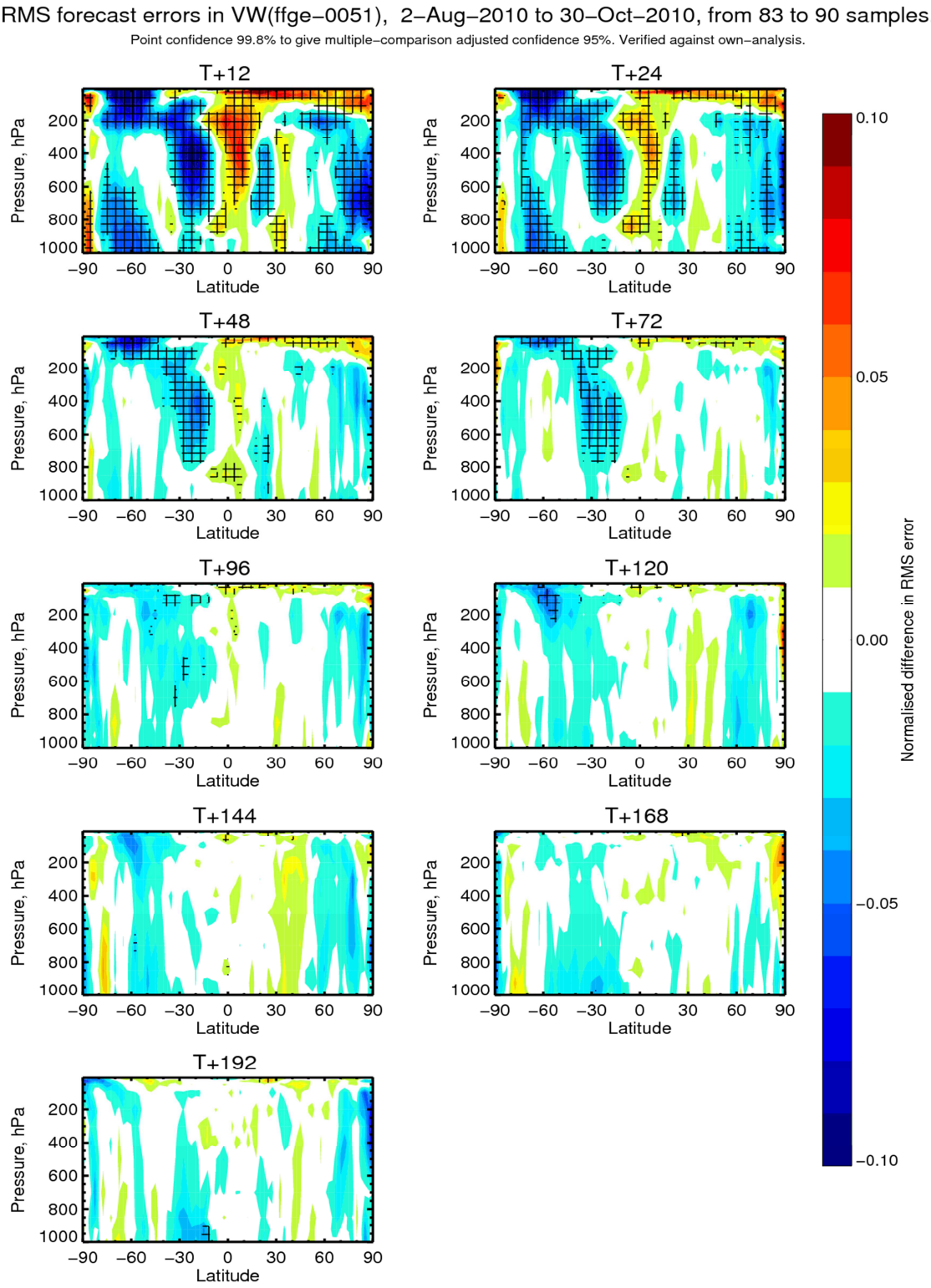

Figure 20: Meridional averages of Root Mean Squared Error reduction of Wind Vector forecasts from deterministic 4D-Var analyses with background error variances estimated from a 10 member EDA (experiment ffge) versus using randomization errors (experiment 0051). Forecasts are verified against own analysis. Scores are averaged over the period 2010802-20101030. Crosses indicate statistical reliability of the results at the 95% confidence level.

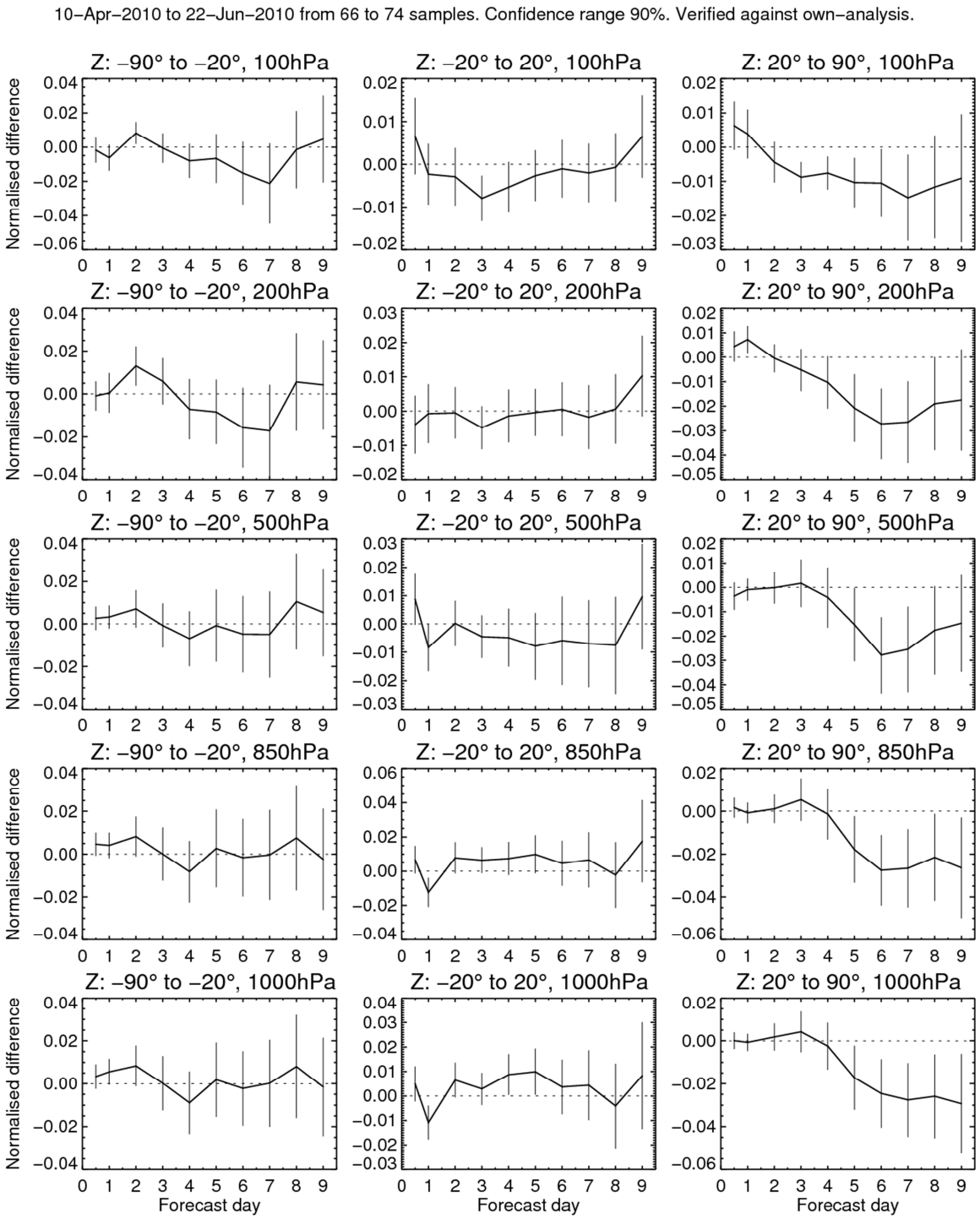

Figure 21: Forecast skill score differences of RMSE of Geopotential height at 100, 200, 500, 850 and 1000 hPa. Forecasts from deterministic 4D-Var analyses with background error variances estimated from a 10 member EDA using the non-parametric filtering algorithm (fha8) versus using the current Raynaud et al., 2009, filtering algorithm (ffffn). Scores are computed against own analysis and averaged over the period 20100410-20100630. First column refers to the Southern extra-Tropics, second column to the Tropics, third column to the Northern extra-Tropics.

using the new filtering scheme are apparent in the Northern extra-Tropics at the 95% confidence level, while the impact is overall neutral in other regions. Verification results for other parameters (e.g., wind, relative humidity) give similar results.

6 Summary and outlook

The use of cycled background error variance estimates sampled from an EDA system is a first step towards the ultimate goal of a hybrid assimilation system with a fully flow-dependent representation of background error covariances. It has been successfully implemented in the ECMWF 4D-Var resulting in substantial improvements in the skill of the ensuing deterministic forecast.

A key aspect is that the sample variances computed from the EDA are themselves affected by both random and systematic errors which need to be accounted for. Random errors are an unavoidable consequence of sampling due to the small dimension of the available ensemble (10 members). Techniques already available in the literature (Raynaud et al., 2009) or of partially new conception (Sec. 2) have been implemented and compared in terms of their effectiveness in extracting the statistically significant part of the signal. Systematic errors reflect, on the other hand, basic inadequacies of our modelling of the error sources in the system. While the long term research effort at ECMWF is aimed at understanding and addressing the basic causes of these deficiencies, it has been found useful to mitigate their impact through an adaptive statistical rescaling technique. This is somewhat analogous to current practice in Ensemble Kalman Filter applications in atmospheric numerical weather prediction, but with the advantage that the atmospheric balances imposed by the variational analysis are not compromised, since the rescaling procedure is performed on the EDA background variances, not on the background fields themselves.

The main mechanisms through which flow-dependent EDA variances influence the deterministic analysis have been elucidated in section 4 through single-observation experiments and the analysis of a case study focused on the tropical storm Aere. They essentially involve a better use of the observations, both at the quality control stage and by giving larger weight in the analysis to observations close to areas where the largest uncertainties are diagnosed. The use of EDA variances also introduces a non negligible degree of flow-dependency in the analysis increments. We expect this aspect of the analysis to be further improved once correlation information from the EDA is exploited.

The results of extensive numerical experimentation indicate that the use of EDA variances significantly improves the skill of ECMWF forecasts, particularly in the winter hemisphere. The only notable exception is for tropical winds in the early forecast range. It is hypothesised that this deterioration is partly attributable to the global, homogeneous nature of the filtering technique which does not take into account the spatial variations of the background errors' statistics. A solution to this problem is currently being pursued in the framework of the wavelet representation of background errors (Fisher, 2003). A wavelet based filtering technique would allow the noise model to incorporate some spatial detail at the cost of relinquishing the full spectral resolution of the current filter.

A number of further improvements to the current use of EDA variances in 4D-Var can also be envisaged. The increase of the EDA size is an obvious one and it has been discussed in Bonavita *et al.*, 2011. Other developments include the use of EDA variances in each of the EDA members' analysis step; the computation of EDA variances in radiance space in order to extend the first guess check from conventional observations to radiances from satellite-borne sensors; the use of EDA variances for the diagnosis of error variances of the unbalanced components of the ECMWF 4D-Var (along the lines discussed in Raynaud *et al.*, 2011).

On a more basic level, further extensions to the current methodology will be directed at 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 similarly to what has been implemented for the sampled EDA variances. In this context too, the basic idea would be to implement a flow-dependent extension of the wavelet diagonal correlation model (Fisher, 2003) already in use at ECMWF. In combination with the EDA-derived background errors this would lead to a fully flow-dependent error covariance model.

REFERENCES

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.

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

Berre, L., O. Pannekoucke, G. Desroziers, S. E. Stefanescu, B. Chapnik and L. Raynaud, 2007: A variational assimilation ensemble and the spatial filtering of its error covariances: increase of sample size by local spatial averaging. *Proceedings of the ECMWF Workshop on flow-dependent aspects of data assimilation* ECMWF, pages 151-168.
(Available from: http://www.ecmwf.int/publications/)

M. Bonavita, L. Raynaud and L. Isaksen, 2011: Estimating background-error variances with the ECMWF Ensemble of Data Assimilations system: the effect of ensemble size and day-to-day variability. *Q. J. R. Meteorol. Soc.*, **137**: 423–434 .

F. Bouttier, J. Derber and M. Fisher, 1997: The 1997 revision of the Jb term in 3D/4D Var. *ECMWF Technical Memorandum,* 238. (Available from: http://www.ecmwf.int/publications/)

Buehner, M., 2005: Ensemble-derived stationary and flowdependent background error covariances: Evaluation in aquasi-operational NWP setting. *Q. J. R. Meteorol. Soc.*, **131**: 1013–1043.

Buehner, M. and M. Charron, 2007:. Spectral and spatial localization of background-error correlations for data assimilation. *Q. J. R. Meteorol. Soc.*, **133**: 615-630.

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

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 from: http://www.ecmwf.int/publications/)

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 from: http://www.ecmwf.int/publications/)

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

Houtekamer, P. L., Louis Lefaivre, Jacques Derome, Harold Ritchie, Herschel L. Mitchell, 1996: A System Simulation Approach to Ensemble Prediction. *Mon. Wea. Rev*., **124**, 1225–1242.

Hsu, H. P., 1996: *Probability, random variables and random processes*. McGraw-Hill, New York,

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 from: http://www.ecmwf.int/publications/)

Isaksen, L., M. Bonavita, R. Buizza, M. Fisher, J. Haseler, M. Leutbecher and Laure Raynaud, 2010: Ensemble of data assimilations at ECMWF. *ECMWF Technical Memorandum* 636. (Available from: http://www.ecmwf.int/publications/)

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

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

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

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

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

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 Technical Memorandum*, 598. (Available from: http://www.ecmwf.int/publications/)

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

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

Raynaud, L., L. Berre, and G. Desroziers, 2011: An extended specification of flow-dependent background-error variances in the Météo-France global 4D-Var system. *Q. J. R. Meteorol. Soc.* **137**, 607–619.

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. Weather Rev*., **133**, 441–453.

Wang, X., D. Barker, C. Snyder, T. M. Hamill, 2008a: A hybrid ETKF-3DVar data assimilation scheme forthe WRF model. Part I: observing system simulation experiment. *Mon. Wea. Rev*., **136**, 5116-5131.

Wang, X., D. Barker, C. Snyder, T. M. Hamill, 2008b: A hybrid ETKF-3DVar data assimilation scheme for the WRF model. Part II: real observation experiments. *Mon. Wea. Rev*., **136**, 5132-5147.

APPENDIX

Assume we have n independent and identically distributed (i.i.d.) realizations of the EDA system, each of N_{ens} size. The signal of interest here is the standard deviation of the EDA analysis or background forecasts, which we can express for the EDA set j as the sum of a noise-free component plus an uncorrelated sampling error:

$$S_j = S_* + S_j^e \qquad j=1,2,\ldots,n \qquad \text{(A.1)}$$

Then the correlation between the standard deviations of any couple of EDA realizations will be:

$$\rho\left(S_i, S_j\right) = \frac{\mathrm{cov}\left(S_i, S_j\right)}{\sqrt{Var\left(S_i\right)Var\left(Sj\right)}} = \frac{Var\left(S_*\right)}{Var\left(S_*\right)+Var\left(S^e\right)} \qquad \text{i,j=1,2,…,n} \qquad \text{(A.2)}$$

where we exploited the i.i.d. assumption and the assumed uncorrelatedness of the sampling noise and the signal (i.e., $cov(S_*,S^e_i)= cov(S_*,S^e_j)= cov(S^e_i,S^e_j)=0$ for $i\neq j$).

A standard result of estimation theory (Hsu, 1996) implies that for any given EDA realization i the best linear mean square estimator $\hat{S}_*$ of the signal S_* will be given by the product of the sample estimate S_i with the correlation field in (A.2):

$$\hat{S}_* = \rho\left(S_*, S_i\right)S_i = \frac{\mathrm{cov}\left(S_*, S_i\right)}{\sqrt{Var\left(S_*\right)Var\left(S_i\right)}} S_i = \frac{Var\left(S_*\right)}{Var\left(S_*\right)+Var\left(S^e\right)} S_i = \rho\left(S_i, S_j\right)S_i \qquad \text{(A.3)}$$

In order to reduce the sampling errors in the estimation of the correlation field, we follow Berre *et al.*, 2007 and consider the signal and noise variances as functions of total wavenumber only (i.e., the signal and noise 'Power Spectra', denoted by P_n). This will allow the averaging to run over all spectral zonal numbers allowed for a given total wavenumber. The correlation spectrum will then be written as:

$$\rho_n\left(S_*, S_j\right) = \frac{P_n\left(S_*\right)}{P_n\left(S_*\right)+P_n\left(S^e\right)} = \frac{P_n\left(S_i\right)-P_n\left(S^e\right)}{P_n\left(S_i\right)} = 1-\frac{P_n\left(S^e\right)}{P_n\left(S_i\right)} \qquad \text{(A.4)}$$

Under the same assumptions of the sampling noise being a random, uncorrelated process, its variance (and power spectrum) can be estimated from (A.1) as:

$$Var\left(S^e\right) = 0.5Var\left(S_i - S_j\right) \qquad i\neq j=1,2,\ldots,n \qquad \text{(A.5)}$$

Since $Var\left(S_i^e - S_j^e\right) = Var\left(S_i^e\right)+Var\left(S_j^e\right) = 2Var\left(S^e\right)$; an equivalent expression holds for the corresponding power spectra.