Use of analysis ensembles in estimating flow-dependent background error variance

Lars Isaksen, Mike Fisher and Judith Berner

European Centre for Medium-range Weather Forecasts
Shinfield Park, Reading, UK

Abstract

This paper will describe the present status of the ensemble data assimilation (EnDA) work at ECMWF. We will first give a brief description of the EnDA method, with special emphasis on the spectral backscatter (SPBS) scheme that recently has been included. The impact of SPBS on analysis error estimates and short range forecast error will be presented. Next we investigate the ensemble spread in detail. We compare the background error variances for the control assimilation that use the randomization method with the ensemble data assimilation based method. We discuss how the ensemble spread can be increased to more realistic values. Despite the large differences in background error estimates from the two methods, applying "Errors of The Day" modifications of variances have little impact on analyses and forecasts. We try to explain why the impact on the general forecast scores of the flow dependent background errors is so small.

Introduction

The streamlined ensemble data assimilation (EnDA) system developed at ECMWF has been used to run a range of EnDA experiments mainly to investigate how to achieve a more realistic spread. Most experiments have been done with a 10-member ensemble of 4D-Var (T255L91 outer loop, T95/T159 inner loop) with perturbed observations, SST and 2metre relative humidity input for the soil moisture analysis. In collaboration with the Predictability and Diagnostics section at ECMWF, the main investigation has been to explore and enhance the representation of model error in the EnDA system using the spectral backscatter scheme (SPBS). The SPBS parameterization scheme has been used in the trajectory job and the first-guess forecasts of the assimilation for each ensemble member. The scheme increases the spread in the tropics and to a lesser extent in the extra-tropics. The original SPBS version produced large vertical correlations which resulted in unrealistic J_b statistics; this has been successfully corrected, without loss of spread. Even more realistic spread has been achieved by assuring that the SPBS perturbations fulfil mass/wind balance constraint. Despite these improvements the EnDA estimation of analysis error and background error is still too low. A number of areas to alleviate the shortcomings of EnDA are currently being investigated: to account for correlated observation error for satellite radiances; to further development of the SPBS scheme; to explore the randomization elements of the new radiation scheme to treat each ensemble member differently. The increase the understanding of the EnDA method we are now performing experiments with 'conventional' data only.

To investigate the impact of ensemble-based flow dependent covariances in 4D-Var, we have used the 10 member ensemble of first-guess forecasts to construct a background error variance (sigma-b) estimate. This variance estimate is multiplied by a factor of $(1.5)^2$-$(2.5)^2$ and is then used in a separate 4D-Var assimilation without perturbed observations, SST or soil moisture. The sigma-b estimate is also used in the background

QC check for this 4D-Var assimilation system. The first preliminary results show that the impact of flow-dependent background error variances is neutral for hemispheric averaged 500 hPa forecast scores, averaged over a 20-30 cases. This may not be that surprising, we do expect that flow-dependent background errors will be most helpful when the error in the background forecast is extreme (very different than that implied by the operational background errors). If the ensemble estimate of B correctly captures these changes, the observations should be used more effectively in the DA system and the analyses (and subsequent forecasts) should be more accurate on average. This is being diagnosed at the moment.

Finally it should be noted that the use of ensemble perturbations to initialize the EPS system instead of the evolved singular vectors results in a small improvement in EPS performance in mid-latitudes, and a significant improvement in the tropics.

The ensemble data assimilation system used at ECMWF

The ensemble data assimilation system used at ECMWF run an ensemble of analyses with random observation and SST perturbations, and form differences between pairs of background fields. These differences will have the statistical characteristics of background error (but twice the variance). For most of the experimentation we used a 10 member ensemble of 4D-Var assimilations. One additional low-resolution 50 member ensemble was also run. Compared to previously run ensemble data assimilation experiments (Kucukkaraca and Fisher 2006: Fisher 2003) the present experiments have improved the representation of model error by using the new spectral backscatter (SPBS) method (developed by Judith Berner, ECMWF). We have also introduced the use of the variational bias correction method (Dee 2005). In the control assimilation the variational bias correction scheme is used in the standard way to cycle the biases. For the ensemble members the bias coefficients are fixed to the control assimilation values for each cycle. The variational bias correction is not part to the minimization problem for the ensemble members. When the ensemble data assimilation system was tested with active variational bias correction for the ensemble members it lead to artificially long correlations in the J_b statistics derived for the short range ensemble forecasts. Freezing the variational bias correction coefficients for each member improved this problem considerably. This set of ensemble experiments was the first that included GPS radio occultation data in addition to the full operational set of standard conventional and satellite observations. SST and observations were perturbed with observation error size random uncorrelated quantities. Only for feature tracking winds (SATOB) we have taken account of horizontal correlations following the work of Bormann *et al*. (2003). For most experiments we performed 45 days of T255/T159 ensemble data assimilations with ten members, a few experiments were performed at lower resolution. The main quantity of interest from the ensemble is the standard deviation of ensemble spread among the 10 members. Both the day-to-day variations and the long term average of the spread are of interest. This will be discussed further in the following sections.

Estimation of background errors using the operational randomization method and the ensemble data assimilation system

The ECMWF analysis system has for many years used a cycling algorithm (Fisher and Courtier, 1995) to determine the standard deviation of background errors that are used to define the background cost function, J_b. The algorithm consists of an initial estimate of analysis error standard deviation, based on the leading eigenvectors of the Hessian of the analysis cost function, followed by a simple inflation to account for the growth of error between one cycle and the next. The method is described in more detail by Fisher (2003) and Fisher in these proceedings. The main source of flow-dependence in background error standard deviation in this method is produced through the use of linearised versions of nonlinear balance equations in the analysis change of variable (Fisher, 2003). The standard deviations of the balanced components of divergence, temperature and surface pressure are not directly specified. They are determined implicitly through the action of the balance operators on the covariance matrix for vorticity. By using balance operators that are linearised about the background state, a significant degree of flow-dependence is achieved (see Fig. 6).

Despite these sources of flow-dependence in the current operational method for estimating background error standard deviation, it is believed that significantly improved estimates, that reflect the dynamical characteristics of the underlying flow, could be generated from the spread of an ensemble of analyses. An initial investigation of this possibility was conducted by Kucukkaraca and Fisher (2006), who investigated two case studies. This paper will to some degree investigate this issue using more recent ensemble data assimilation experiments.

In the experiments presented here, each member of the analysis ensemble consisted of an independent run of the analysis-forecast system for a period of 45 days during September-October 2006. For most experiments ten members and a control job were used, but we also performed a 50 member experiment at reduced analysis resolution. For each member, random perturbations with the statistical characteristics based on the prescribed observation error were added to the observations. Spatially-correlated perturbations were used in the case of atmospheric motion vectors, using Bormann *et al*.'s (2003) estimate of spatial correlation. Perturbations for all other observations were assumed to be spatially uncorrelated. In addition, sea-surface temperatures were perturbed to account for uncertainty in the prescribed field, and the soil moisture analysis via perturbed 2 metre relative humidity data input. As well as this, the short background forecasts that connects analysis cycles and the trajectory forecasts used for observation comparison were perturbed (to account for model error) using a spectral backscatter scheme designed by Berner (personal communication 2007), based on earlier backscatter schemes developed at ECMWF (Shutts 2005). Despite the attempt to add perturbations that account for the main sources of error, the analysis ensemble seriously underestimates the standard deviation of background error. For this reason, the spread of the ensemble must be inflated in order to produce a realistic estimate of the standard deviation of background error. A simple linear scaling by a factor of 2 or 2.5 is the only method that has been used for the experiments discussed here.

The Spectral Backscatter Scheme

We will briefly describe the characteristics of the Spectral Backscatter Scheme (SPBS, see Fig. 1) that accounts for some of the model errors that the perturbation of observations cannot represent. The main features are that it is stochastic and represents spatial and temporal correlations in a meteorologically sensible way. The rationale behind the scheme, is that a fraction of the dissipated energy is scattered upscale and acts as stream function forcing for the resolved-scale flow (based on earlier LES and CASBS backscatter schemes developed at ECMWF by Shutts 2005).

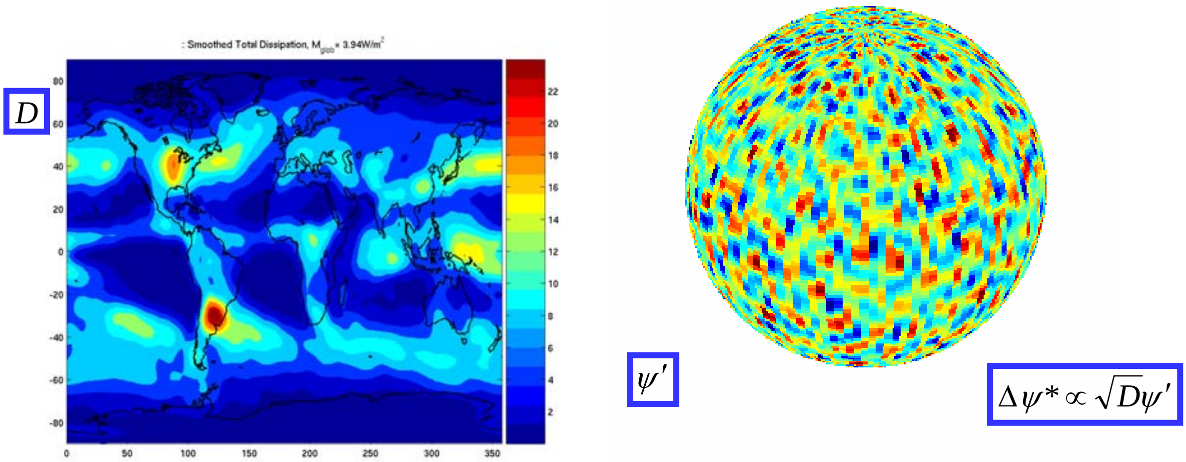

Figure 1: Schematic representation of the spectral backscatter scheme (SPBS) designed by Judith Berger. Total Dissipation rate (D) is calculated from numerical dissipation, convection, gravity/mountain wave drag. A spectral Markov chain(Ψ') is used to prescribe temporal and spatial correlations. The SPBS scheme is used to account for unresolved model errors for each ensemble member.

Because the forcing is applied in the spectral wave domain it is possible to allow for scale selective forcing. The scheme include flow-dependent perturbation, because they are weighted with dissipation rates (see Fig.1 left panel for an example). Another advantage of the spectral representation is the consistency with the spectral dynamical core and the assurance of isotropic patterns on the sphere. The SPBS scheme injects energy in regions of large dissipation, which are the regions of large model errors. The method will not account for all types of model errors, but mainly for errors associated with convectively active regions and physically active regions. Therefore the SPBS will take better account of model errors in the tropical region than in the extra-tropical regions. This will be discussed further in the next section of the paper. It is clearly visible from Fig. 4. We have of course not got a full knowledge of the distribution and amplitude of model errors, but it is true that they are largest in the tropical region, partly because the initial conditions are less accurate here. So the SPBS is most likely doing a good job, even though it does not take account for all types of model errors.

Estimating analysis error from the average ensemble data assimilation spread

First we will look at analysis error estimates derived as the standard deviation values averaged over the last 40 days of the ensemble data assimilation experiments. This was done by Edit Hagel during her stay at ECMWF in the spring 2006. The results are presented in Fig. 2 upper panel. It is clearly producing a realistic horizontal distribution of analysis errors, with small values over USA, Europe and Australia. The largest analysis errors are found in the tropics and in outflow from South America in the region between Brazil and Argentina. As mentioned in the previous section, the ensemble spread is underestimated, so the ensemble spread has been multiplied by a factor of two in Fig. 2 upper panel.

The bottom panel in Fig. 2 shows the ratio of analysis error estimates between the ensemble that is using SPBS to improve representation of model error in the ensemble data assimilation and the ensemble data assimilation experiment without SPBS. The SPBS is increasing spread in most regions of the world, but mainly in the tropics and in the southern hemisphere extra-tropics. It is interesting that the SPBS seems to capture a well known model problem associated with outflow from the South American continent during September and October.

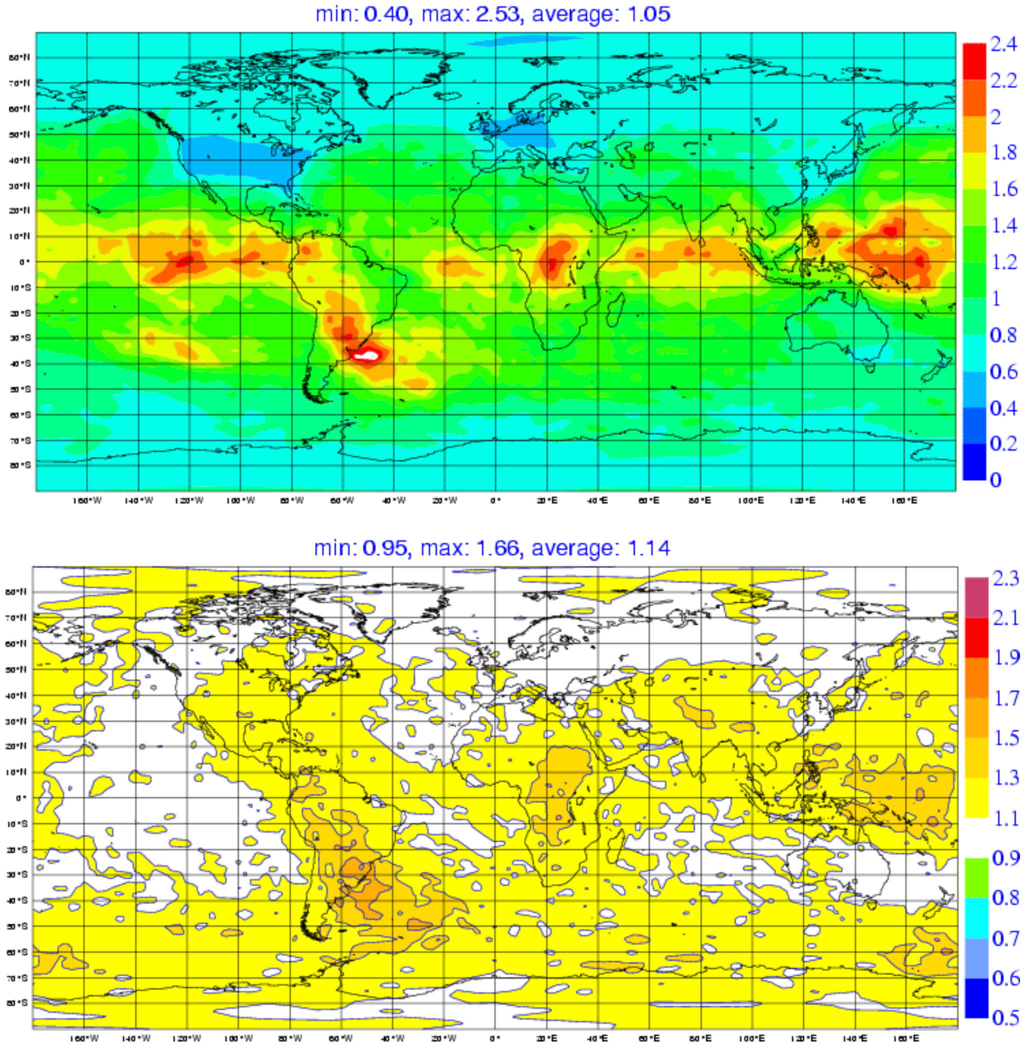

Figure 2: Upper panel: Static estimation of analysis error for 200hPa u-wind from the standard deviation of the last 40 days of 10 member ensemble data assimilation experiment with spectral backscatter scheme used. A scaling factor of two is required due to underestimation of the ensemble spread. Lower panel: The ratio of analysis error between the experiments with versus without spectral backscatter applied. It is a measure of the increased spread due to the Spectral Backscatter (SPBS) scheme. (Courtesy Edit Hagel).

The structure of background error variances generated from the operational randomization method and from the ensemble data assimilation spread

In this section we will investigate how the background error variances differ in the operational style control configuration and when they generated from a ten member ensemble data assimilation system.

Fig. 3 shows the zonal average temperature background error variances estimated by the randomization method and representative for the actual background errors used in 4D-Var J_b calculations on a specific day. The left panel shows the operational method and the right panel the ensemble based background error estimate. To make the colour scaling consistent between the left and the right panels, the values of the right panel have been rescaled by a factor 0.7. After this rescaling the structures and amplitudes are remarkably similar, at least when we look at day-to-day zonal mean background error estimates. There is slightly larger values in the tropics for the rescaled ensemble based estimates, resulting in a larger difference between tropics and extra-tropics for the ensemble based estimates. It should be noted that the right panel shows the background error variances as they are estimated from the randomization method, i.e. this is how they are used by the 4D-Var J_b calculations. That explains to some extend why they are fairly similar.

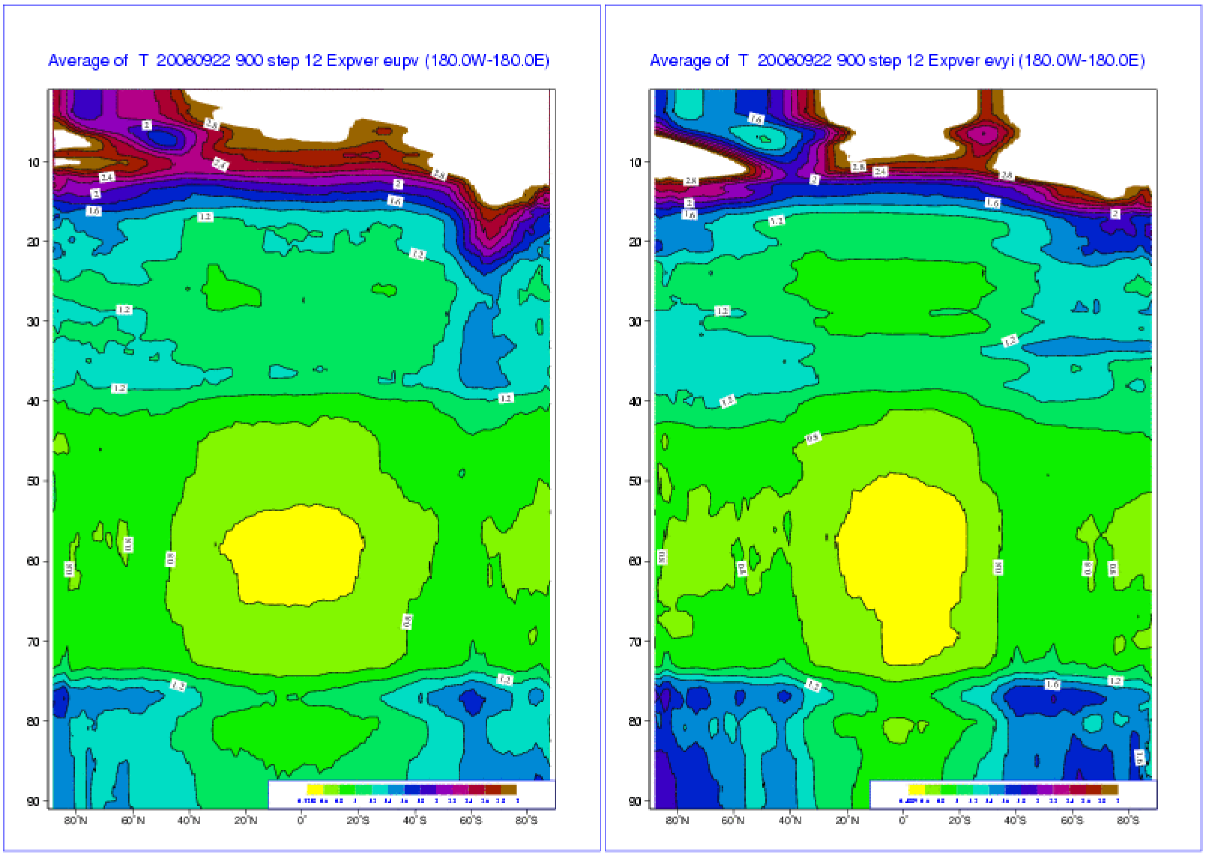

Figure 3: Zonal average temperature (K) background error variances estimated by the randomization method and representative for the actual background errors used in 4D-Var Jb calculations valid at 2100UTC on 22 Sep. 2006. Left panel: The operational method. Right panel: Flow dependent background errors based on 2.0 times the standard deviation of the 10 (T255 outer loop+T95/T159 inner loop) ensemble members. The values of the right panel have been rescaled by a factor 0.7.

Fig. 4 is similar to Fig.3, but for U-wind component. For this case there are large differences between the operational randomization based background error variances and the ensemble based ones. Especially in the tropics we see much larger values for the latter method. The plots, like Fig. 3, show the diagnosed variances estimated by the randomization method. Again we have rescaled the ensemble based values (right panel) by 0.7. The reason for the larger differences between the two methods for wind is because the SPBS scheme mainly adds perturbations for vorticity, so the impact is largest on the wind field. Due to the balance constraint in J_b it is also the case that vorticity perturbations in the background error variances read in to the analysis feeds directly through to the J_b calculations. We believe that this large amplitude difference between tropics and extra-tropics is one of the reasons for the problems causing lack of forecast improvement for assimilations/forecast done from analyses using flow dependent ensemble based background error estimates. The amplitude difference between tropics and extra-tropics is now too large. We are investigating how to compensate for that, e.g., by taking account of correlated radiance errors.

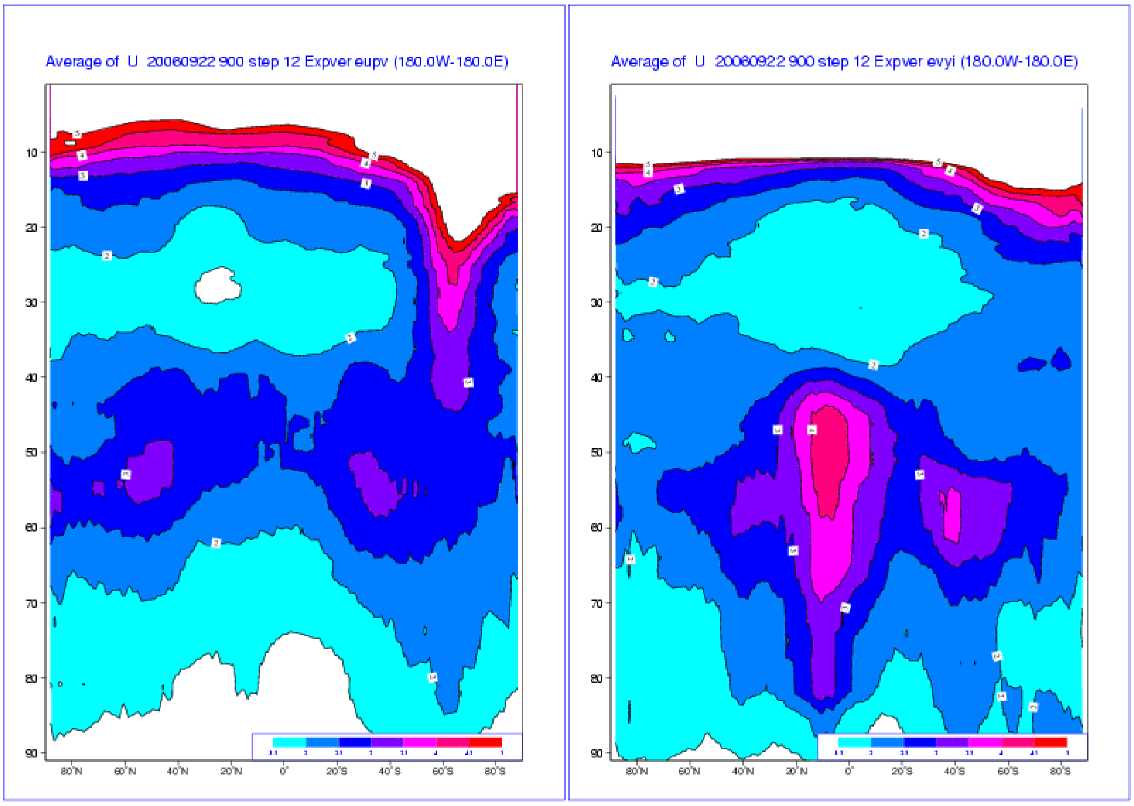

Figure 4: Similar to Fig. 3, but zonal averages for u-wind (m/s).

Fig. 5 shows a snap shot of background error estimates for 850 hPa U-wind on a day in October 2006 from respectively the operational randomization method and 2.0 times the standard deviation of the 10 ensemble data assimilation members, similar to the two methods shows in Figs. 3 and 4. It is clear that the day-to-day background error estimates for wind have very different structure and amplitude, calculated from operational randomization method and from ensemble spread. Note that for Fig. 5 we have not rescaled the values. It is seen that the ensemble based method results in much more flow dependent background error variance estimate. The ensemble based method really captures the dynamically active regions like extra-tropical lows and troughs - an ability that to a large extend lacks for the operational randomization based method, despite the additional flow dependent mass/wind balance constraints (Fisher 2003).

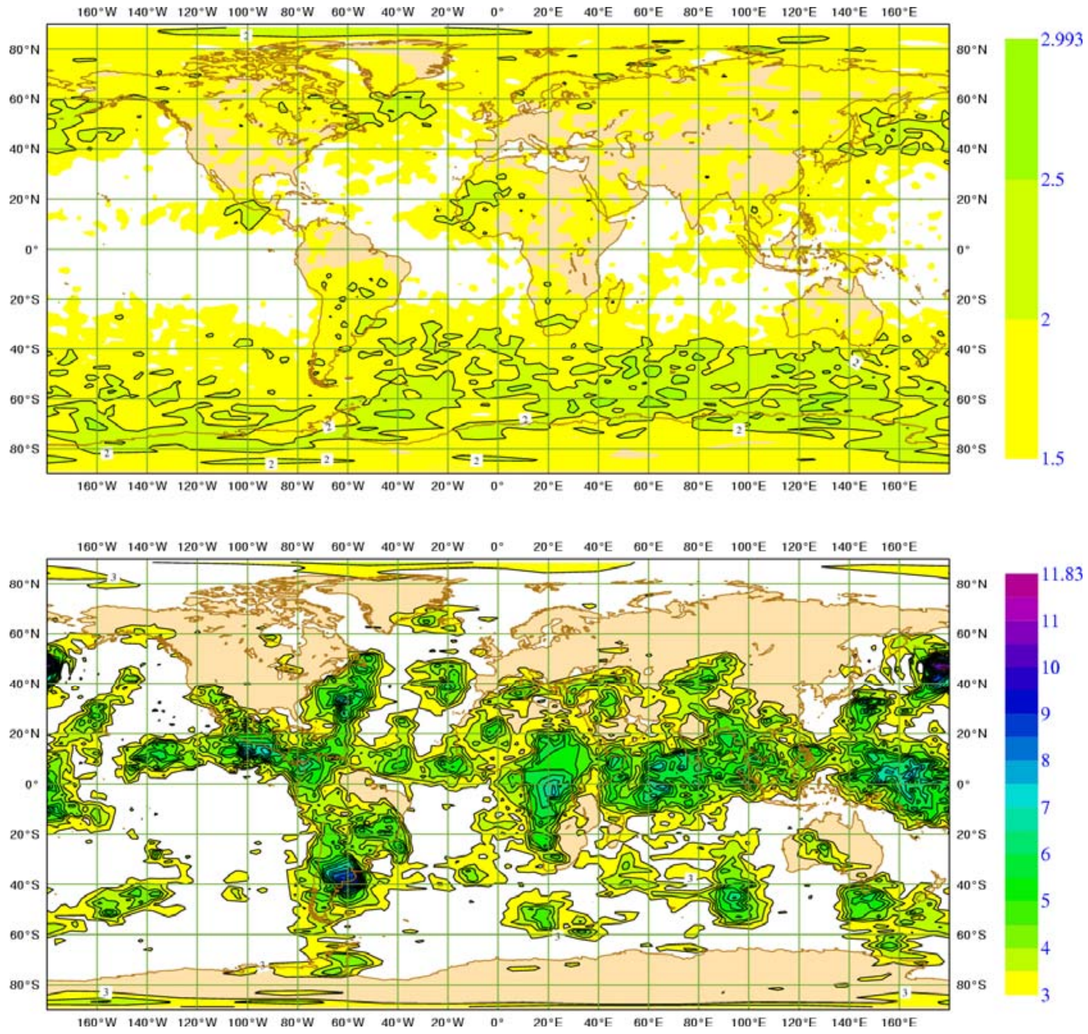

Figure 5: Background error estimates for 850 hPa U-wind at 0000UTC 16 October 2006 as estimated by the randomization method. Upper panel: The operational randomization method (maximum value 2.99 m/s). Lower panel: Input vorticity background error fields are defined as 2.0 times the standard deviation of the10 ensemble members, using T255 outer loop and T95/T159 inner loop (maximum value 11.83 m/s).

Fig. 6 shows the Northern Hemisphere extra-tropical region for the same field as presented in Fig. 5 with the 500 hPa geopotential height field overlaid. In the operational system (left panel) the randomization method is applied, but the main contributor to flow dependence is a non-linear balance and Omega equation mass/wind constraint. This is most effective at jet stream level, but in the present case we do see some impact in the North-East Pacific and near South Greenland. The ensemble based spread method (right panel) produce significantly larger flow dependence and the large amplitudes are almost completely associated with low pressure systems and troughs around the Globe. The largest values are seen east of Japan where an extra-tropical low is developing.

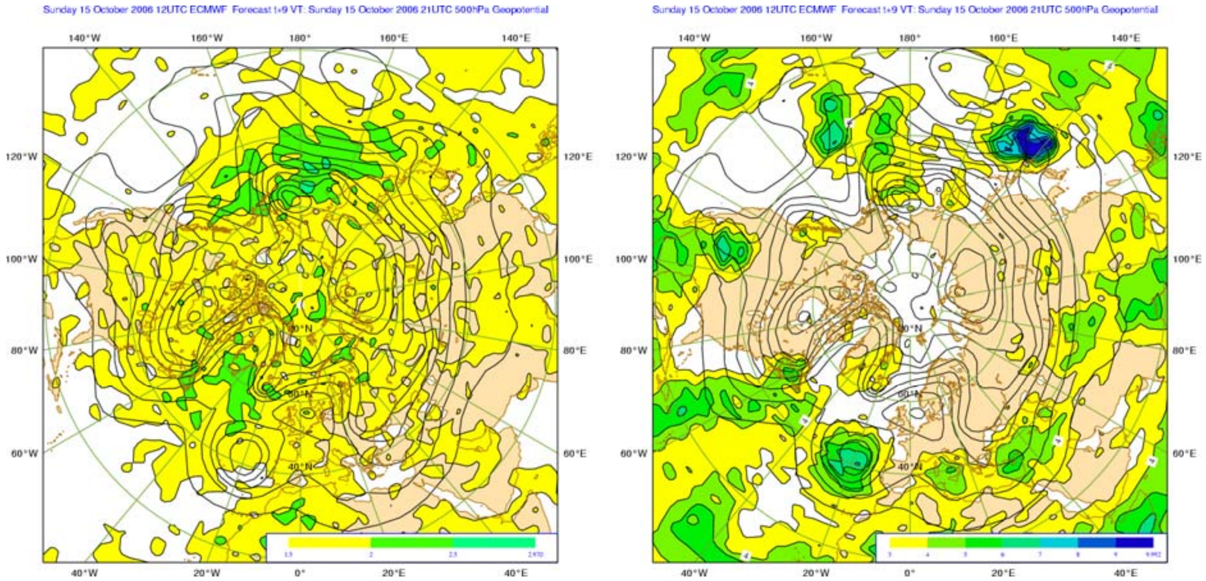

Figure 6: Like Fig. 5, but for Northern Hemisphere extra-tropical region. Both panels have 500 hPa geopotential height field overlaid (8dm contour interval used). Upper panel: Operational randomization method (maximum value 2.97 m/s). Lower panel: 2.0 times the standard deviation of 10 ensemble members using T255 outer loop and T95/T159 inner loop (maximum value 9.99 m/s).

Fig. 7 shows plots of 2.0 times the standard deviation of the ensemble spread for the three systems: 10 member T255 outer loop with T95/T159 inner loop; 10 member T255 outer loop with T95 inner loop; 50 member T255 outer loop with T95 inner loop. The operational method is shows in the top right panel. The raw ensemble based spread fields for vorticity are used for quality control calculations and is the basis for calculation of all the background error variances used in J_b. It is clear from Fig. 7 that these raw fields differ even more from the operational randomization methods smooth fields. This will result in dramatically different quality control decisions, an issue we will have to investigate further. It may well be an advantage to use smoother fields for the quality control decisions, because flow dependent errors should affect all the observations, say in the vicinity of a low pressure system, not just at the “bulls eye like” patterns produced by the the raw spread fields.

Figs. 7 and 8 highlights another interesting thing. It is seen that a low resolution assimilation system (lower left panel) delivers almost the same result as a higher resolution system with two outer loops. This suggests that the computer time may be better spent on more member with a simple low resolution version of the 4D-Var system. Comparing the lower left and lower right panels one can see the impact of increasing the number of ensemble members from 10 to 50. The patterns are very similar, but the result of the 50 member ensemble is smoother with less spikes and also avoiding regions with very low variance. Fifty members are basically giving a statistically better sampling of the forecast errors. This result also favours more members over higher resolution. But case studies of extreme events will have to be performed before the final conclusions can be drawn on this subject.

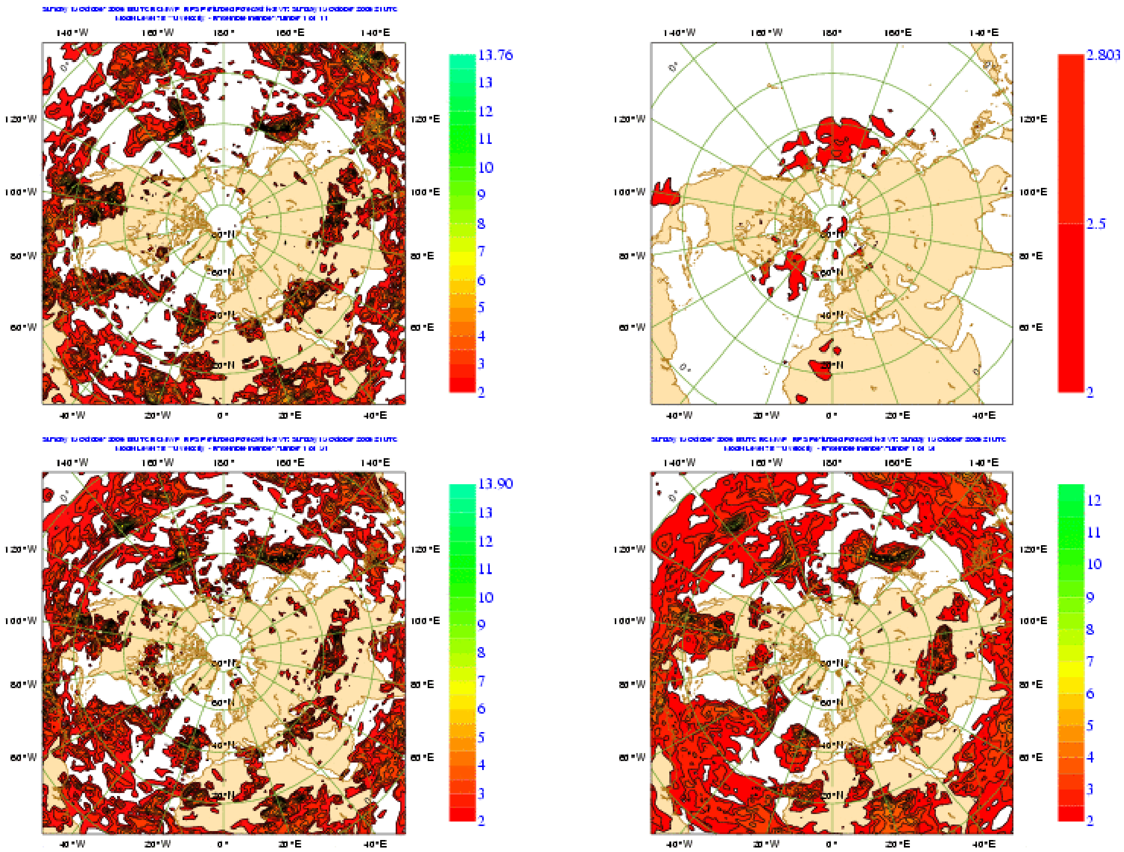

Figure 7: All panels represent 850hPa zonal wind background error estimates valid at 0000UTC 16 October 2006. The ***top right panel*** *shows the operational cycling randomization method (max. 2.8m/s). The* ***other three panels*** *all show 2.0 times standard deviation of the ensemble members.* ***Top left panel*** *represents the 10 member T255 outer loop with T95/T159 inner loop (max. 13.8m/s).* ***Bottom left panel*** *represents the 10 member T255 outer loop with T95 inner loop (max. 13.9m/s).* ***Bottom right panel*** *represents the 50 member T255 outer loop with T95 inner loop (max. 12.0m/s).*

Fig. 8 is a zoom of Fig. 7. Fig. 8 just focuses on the very dynamically active region near Japan. The nice capturing of flow dependence for the ensemble based versions is even more clearly visible here. The smoothing property of using 50 members is also marked. 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 used T255 outer loop resolution. It is at this stage and resolution the observations are perturbed.

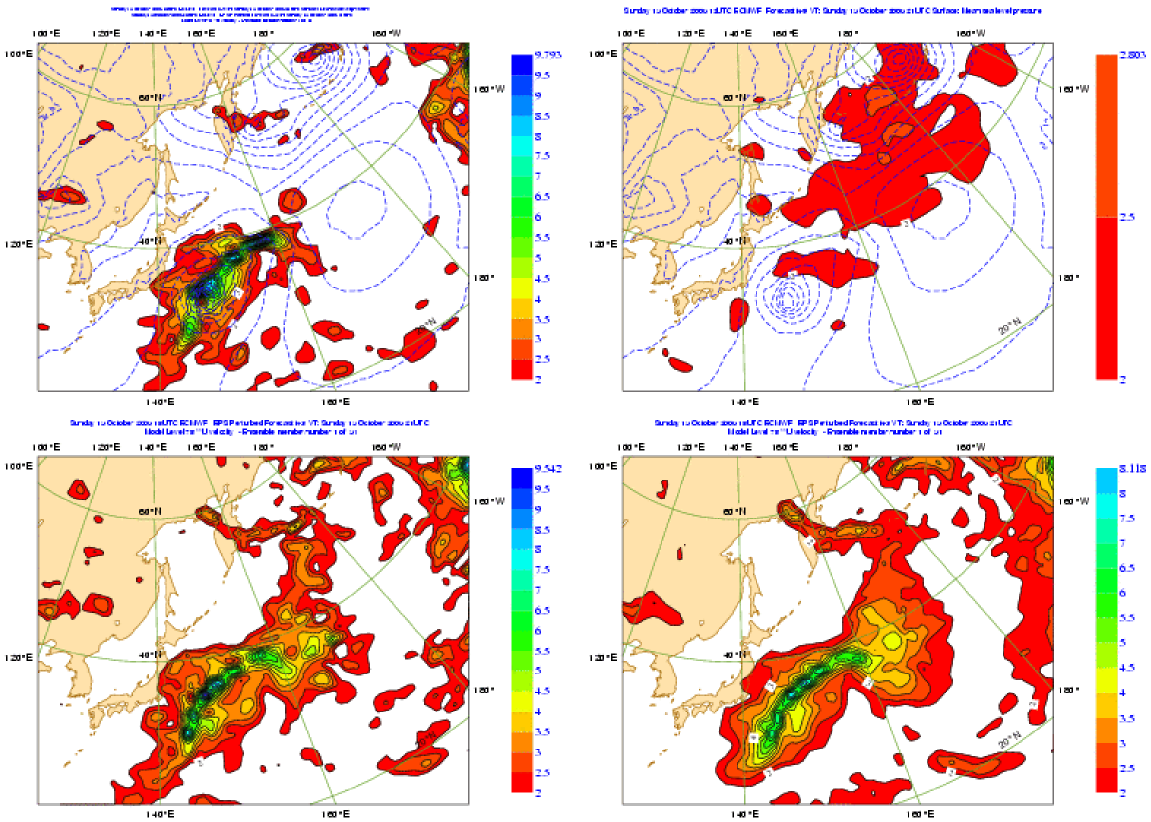

Figure 8: Similar to Fig. 7, but for a region near Japan.

The impact of flow dependent background error variance on assimilation and forecast performance

In the previous sections we have investigated how the background error variances differ in the operational style control configuration and when they generated from an ensemble data assimilation system. We saw that the latter system generated considerably more flow dependent variances. We will now investigate how analyses and forecasts are influenced by the choice of background error variances. We ran a control assimilation plus two assimilations that used ensemble based background errors. The assimilation period was from 20 Sep. to 31 Oct. 2006, because the five first days of the ensemble runs were used for warm-up and excluded from statistics. The assimilation system was using 12h 4D-Var with T319 outer loop and three inner loops (T95/T159/T255). We performed three different assimilations:

- 4D-Var control (static J_b + randomization method applied for background error variance)
- 4D-Var with flow dependent background error variances. Defined as 2.0*StDev(10 member Ensemble 3h forecasts)
- 4D-Var with flow dependent background error variances. Defined as 2.5*StDev(10 member Ensemble 3h forecasts)

The general forecast impact measured by 500 hPa scores for the three systems was very similar (see Fisher, these proceedings Fig. 3). We will not discuss this further here, but refer to the discussion in the next section and in Fisher (these proceedings).

Fig. 9 shows the mean sea level pressure analysis difference between assimilations performed with flow dependent background error and the operational method, respectively. It is clearly seen that the analysis in most parts of the region are very similar, actually a healthy sign, but they differ dramatically near a tropical cyclone south of Japan.

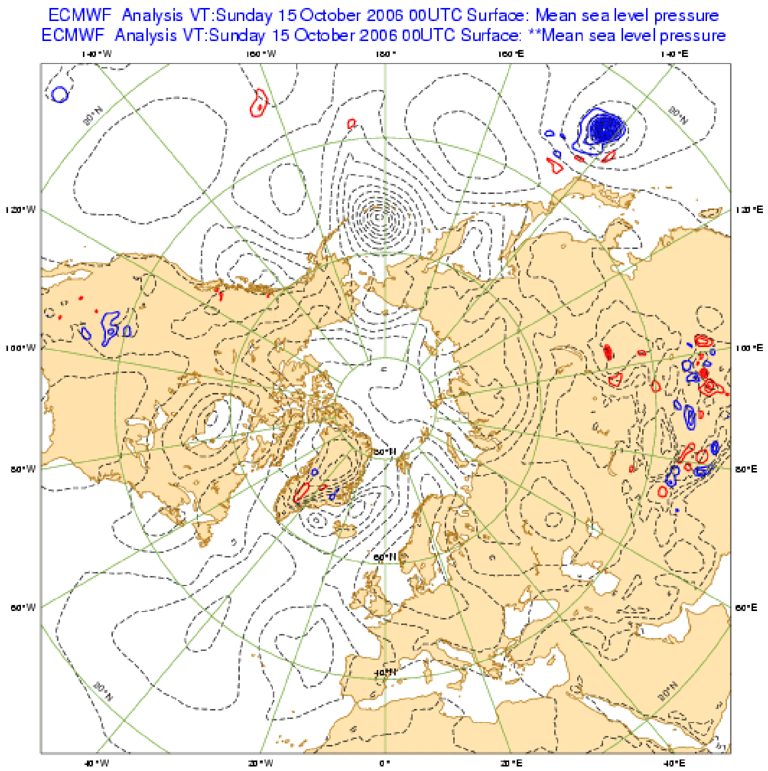

Figure 9: Mean Sea Level pressure field (5hPa contiurs) valid on 15 October 2006. Red (positive) and blue (negative) curves show analysis differences (1hPa contours) for analyses where the back ground error variances are based on ensemble spread and the operational randomization method, respectively.

Fig. 10 shows the 24h and 48h mean sea level pressure forecast difference between assimilations performed with flow dependent background error and the operational method, respectively. It is clear that the two methods differ most in meteorologically dynamical regions (troughs, near lows and fronts) and in the southern hemisphere extra-tropics where errors are more variable. In general area averaged sense the two methods produce very similar analyses and short range forecasts, as one would expect. But it is promising that the two methods mainly differ in regions where flow dependence and analysis uncertainty is largest. It is also good to see that the analysis differences are retained and typically grow in the forecast. The largest differences are typically seen near tropical cyclones, like in this case where we can have a hurricane in the Atlantic ocean.

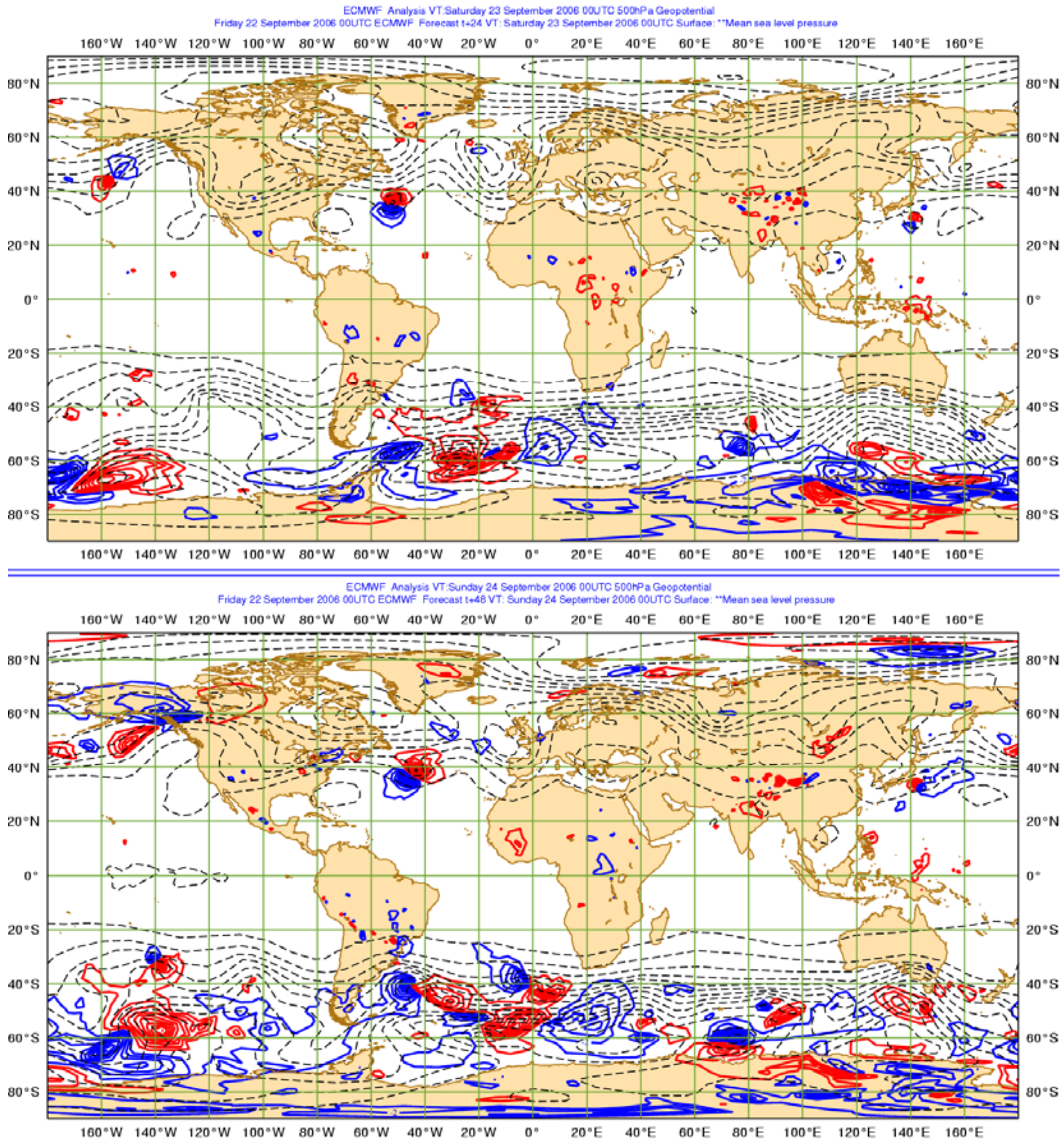

Figure 10: Similar to Fig. 9, but for 24h and 48h mean sea level pressure forecast differences from analyses valid at 0000UTC 22 September 2006 is shown with solid curves (1hPa contours, red=positive, blue=negative). The Dashed curves show analysis fields for the 500hPa height field (8dm contours)..

The 24h and 48h MSLP forecast differences for assimilations with background error estimates based on ensemble spread *2.0 and operational randomization method, respectively, show the same pattern.

We also investigated the impact on mean sea level pressure 24h and 48h forecasts of increasing the EnDA based background error variances from a factor 2.0 to a factor 2.5. The impact of this is small (not shown). Almost identical difference pattern is seen, whereas the amplitudes differ to some degree.

Why does flow dependent background errors and REDNMC (ratio of background error variances versus observation error variances) not influence the general circulation and scores very much?

In the variational assimilation system at ECMWF the variable REDNMC is used to perform a blunt retuning of all background error variances. The background error variances are multiplied by the REDNMC factor, so a small REDNMC results is smaller back ground error variances and therefore relatively speaking, larger observation error variances. REDNMC therefore has a major impact on how tight the analysis fits the observations. When we perform major assimilation system upgrades or recalculate J_b statistics it is necessary, on more less trail and error basis, to determine the REDNMC that delivers a reasonable fit to observations, the best innovation statistics and forecasts. In the context of this paper we have explored the impact on performance by investigating a large range of REDNMC values.

It has already been shown by Fisher (these proceedings) that the area averaged impact of changing REDNMC (in the range 0.6-1.4) is very small on forecast scores. Fisher (these proceedings, Fig. 4) shows the scores for 500hPa geopotential height scores over the Northern Hemisphere. Fisher (these proceedings, Fig. 3) also showed the very small difference in general forecast scores when the background error standard deviations are estimated from the operational cycling method or when they are estimated from the spread of a the ten member ensemble data assimilation described above. Fisher investigated the reasons for this using a simple idealized model. We will investigate the issue further here.

Fig. 11 (upper panel) shows the average analysis increment differences for 500hPa geopotential height for assimilations using REDNMC=0.4 and REDNMC=1.4, respectively. This plot is green or blue (negative values) almost every on the globe because a larger REDNMC value results in relatively smaller observation errors, which typically causes larger analysis increments. It is worth noting that we see red and yellow (positive values) in the American profiler region, Germany and in parts of China and Japan. This is surprising we consider that this is average statistics for a month of assimilations. One conclusion that could be drawn from is that it not possible in the present assimilation system to achieve optimal use of data with just one REDNMC value, it needs to be geographically varying. Alternatively, some of the specified observation errors may be unrealistic, causing the positive regions in Fig. 11. This will have to be investigated, but no matter what the cause is, it can partly explain why it is difficult to see improvement in the general forecast scores when REDNMC is modified. Fig. 11 (lower panel) shows the average 24h forecast error differences for the assimilations presented in the upper panel. It is seen that already after 24h significant areas are positive, especially in the dynamic flow dependent jet stream regions. The signal is clearest in the Northern Hemisphere extra-tropics, where the uncertainties and noise level is lower. It looks like we can conclude that the neutral general area averages forecast impact is not a uniform phenomenon, but a sum of negative and positive contributions. These results make one doubt that it is possible to find a REDNMC value that fits all purposes and that flow dependent specification of background really have a big impact on assimilation performance, but the overall results looks neutral due to averaging of positive and negative impact on different regions. Basically the problem is more complex to solve and require addition diagnostics and experimentation to get a deeper understand of what is going on.

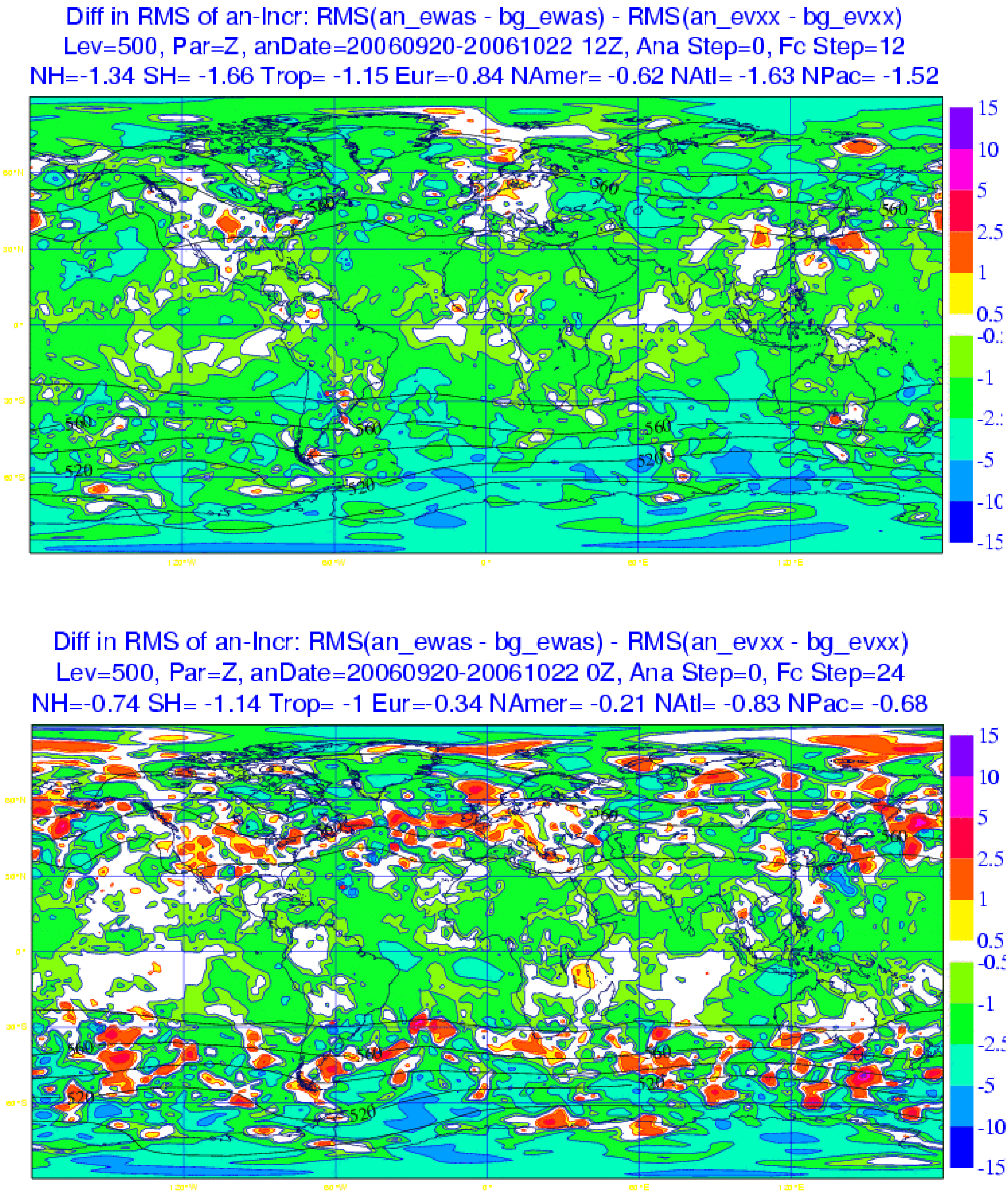

Figure 11: Upper panel: Average analysis increment differences for 500hPa geopotential height between experiments ewas (REDNMC=0.4) and evxx (REDNMC=1.4). Lower panel is similar to uppert panel, but for 24h forecast error differences.

We also investigated the zonal mean 24h forecast errors for temperature and wind for assimilations with background error estimates based on 2.0 times ensemble spread and the operational randomization method, respectively. We do not show the figure here, because they have almost completely the same pattern and amplitudes. This confirms the small forecast impact of changing the background error variances.

Next we investigated innovation statistics for the various REDNMC assimilations in more detail.

It is clear from Fig. 12 that REDNMC impacts analysis fit for conventional data, but has very little impact on the background fit. When REDNMC is small (0.4, left panel) background error variances are smaller relative to the observation errors. This results, as expected, in a worse fit to observations, whereas the right panel figures (REDNMC=1.4) show a better fit to observations. The important measure of the performance on an assimilation system is the innovations (observation minus background), so already from Fig. 12 one get the hint that varying REDNMC has little impact on forecast performance.

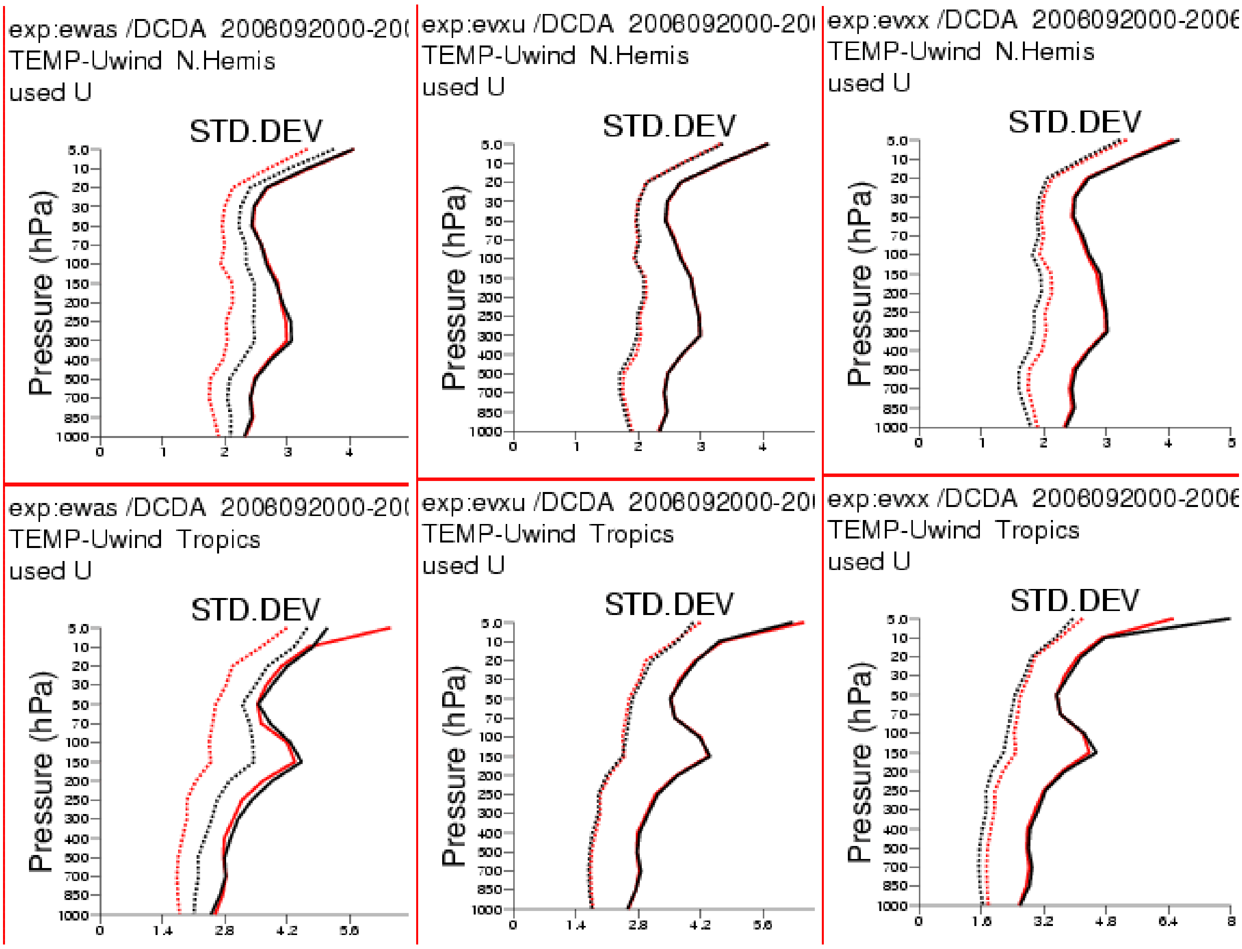

Figure 12: Standard deviation of Innovation statistics for radiosonde zonal wind component (m/s) for the period 20 Sep. to 22 Oct 2006. Solid curves represent obs-background statistics and dotted lines represent obs-analysis statistics. Top panels show Northern Hemisphere extra-tropics (20N-90N) statistics. Bottom panels show Tropics (20S-20N). Left column plots are for REDNMC=0.4, middle column plots are for REDNMC=1.0 (default value) and right column plots are for REDNMC=1.4. The red curves represent the operational reference.

For almost all radiance data REDNMC in the range 0.4 to 1.4 does not influence the analysis or the background fit to observations. One typical example, for NOAA-16 AMSU-A, is shown in Fig. 13. The only noticeable difference is the bias correction statistics for the left column, small REDNMC, for channels 10-14 that peaks in the stratosphere. This will discussed further below.

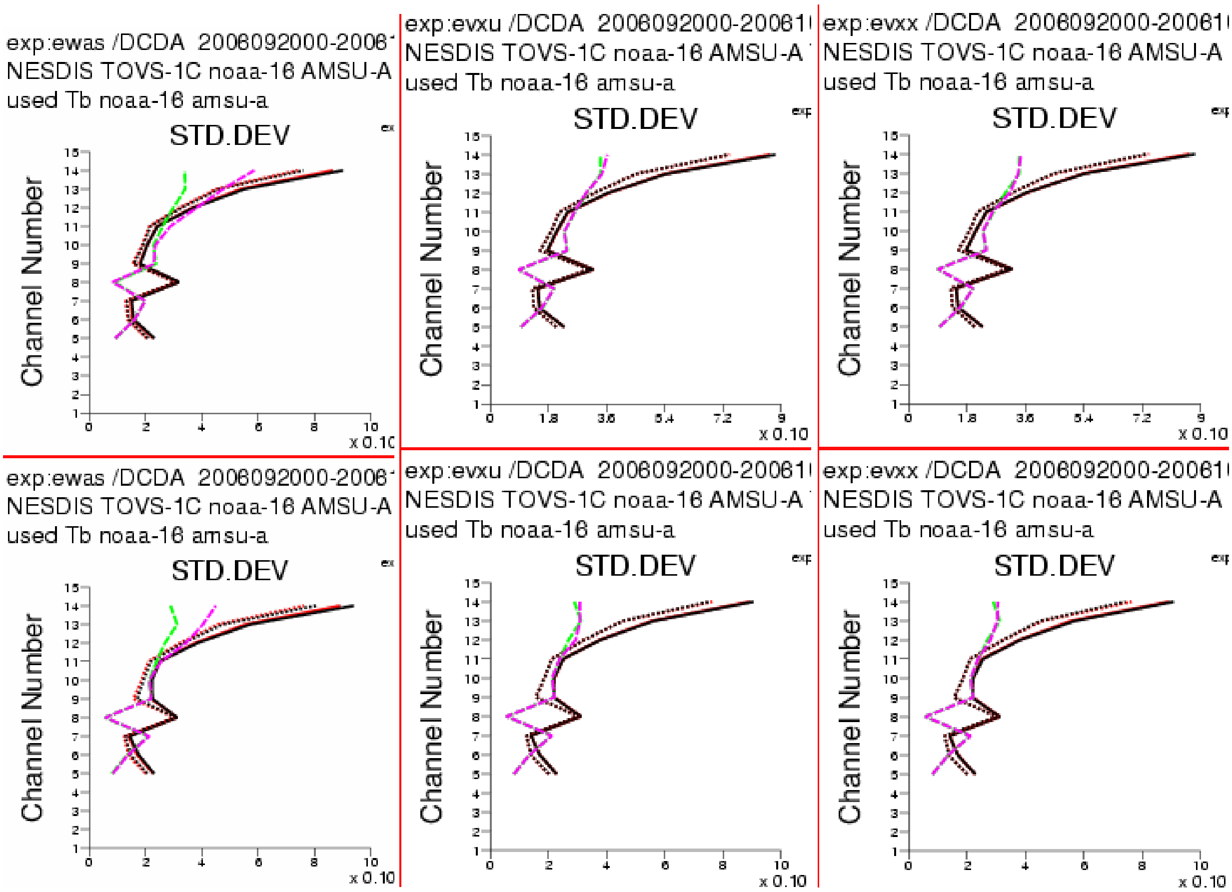

Figure 13: Similar to Fig. 12, but for NOAA-16 AMSU-A channels 5-14. The scale is 0.1K. The dashed curves represent statistics for bias corrections.

The results shown in Figs. 12 and 13 are from assimilation experiment that all used operational style randomized background error variances. The difference between red and black curves in the middle panels is because the ensemble described in the start of this paper was used to generate J_b statistics, whereas the operational system use an older ensemble for its J_b statistics.

For the assimilation where the background error variances were determined by 2.0 times the ensemble spread, as described in second section of the paper, the departure statistics for radiosonde zonal wind is presented in Fig. 14. For the Northern Hemisphere extra-tropics the analysis and background fit is very similar to the control assimilation (red curve), but for the tropical region the larger ensemble data assimilation spread in this region (see Figs. 3 and 4) results in a bigger effective REDNMC. This results in a better analysis fit to the observations, like we saw in Fig. 13. But the background fit in this case is now worse than what we saw in Fig. 13. This is not a very desirable result.

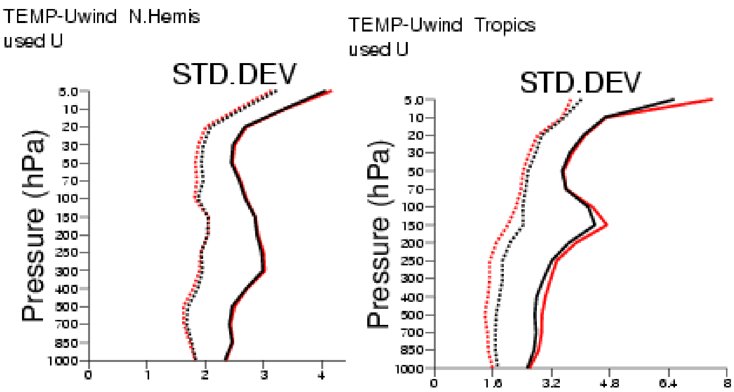

Figure 14: Similar to Fig. 12. Black curves represent operational static Jb and red curves flow dependent Jb. Left panel is for the Northern Hemisphere extra-tropical region. Right panel is for the tropical region.

These experiments show that varying REDNMC by up to a factor 3.5 (1.4/0.4) has very little impact on general forecast performance. It looks like the main reason for this is related with the assimilation of satellite data.

Fig. 15 shows the observation departures and bias statistics (black curves) for the whole 45 day assimilation period for NOAA-15 AMSU-A channel 13. The top panel is for REDNMC=1.4 and the bottom panel is for REDNMC=0.4.

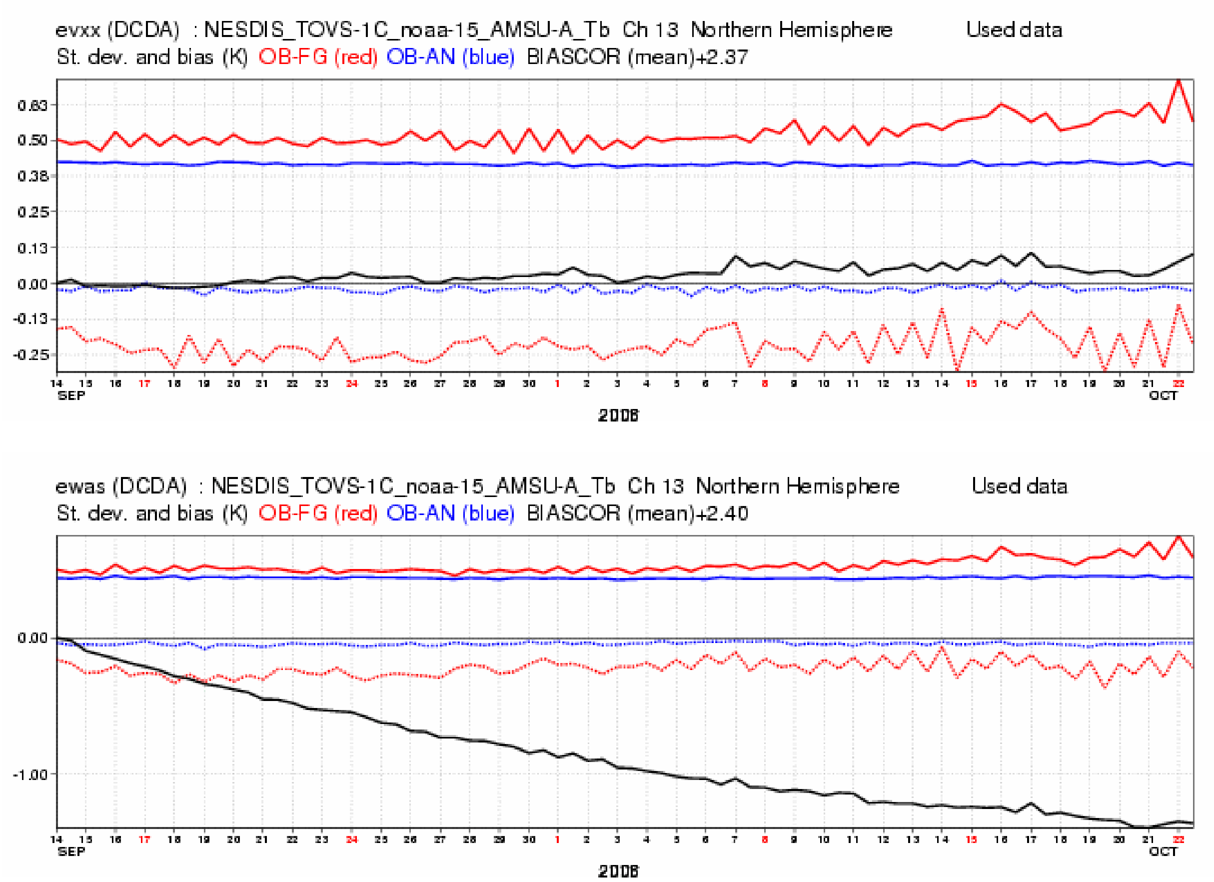

Figure 15: Observation departures and bias statistics (black curves) for period 15 Sep.-31 October 2006 for NOAA-15 AMSU-A channel 13. Top panel: from assimilation experiment using REDNMC=1.4. Bottom panel: from assimilation experiment with REDNMC=0.4.

It is clear from Fig. 15 that the value of REDNMC influences the satellite biases. It is especially the case for stratospheric channels. This is linked to the interaction between the variational bias correction and the relative weight of background and observation controlled by REDNMC. In the case where REDNMC is smaller, the background error is smaller and the observation error relatively larger, therefore the variational bias correction has more freedom to modify the bias. This is clear from the bottom panel of Fig. 15 where the bias gradually is adjusted from cycle to cycle until it reaches a plateau. On the other hand in can be seen (not shown) that static versus flow dependant background errors based on the ensemble data assimilation method does not influence the evolution of the satellite biases. This is not surprising because both configurations have similar REDNMC values.

Our main conclusions are that general circulation is in ECMWF's assimilation system determined by the broad temperature/wind structures that are well described by the high volume satellite data. Flow dependence is not very important for these structures that have a slow error growth. The forecast model can generate and evolve cyclones on its own from these accurate but broad initial conditions. We should only expect significant impact of flow dependent background errors for extreme events that due to their rarity has a small impact on measures like general 500hPa height field scores. Despite this the work is still justified because it is important to predict the extreme events well.

Assimilation experiments with surface pressure data only (Thépaut (2006))

We will now show an example where REDNMC does affect the general scores dramatically. Jean-Noël Thépaut (2006) investigated the impact of assimilating only surface pressure observations in 3D-Var and 4D-Var. He performed a set of assimilation experiments for the period 2004120400-2005022512, where the 12 first days used for warm-up and excluded from the statistics. The four different assimilations were 4D-Var control; 3D-Var "surface pressure only". With retuned REDNMC=2.7; 4D-Var "surface pressure only" with default REDNMC=1.; and 4D-Var "surface pressure only" with retuned REDNMC=2. The experiments with default REDNMC resulted in poorly performing assimilations because the background errors were much larger in the system that only used surface pressure data. Increasing REDNMC compensated for this by increasing the background error variances relative to the observation error variances. Fig. 16 shows the dramatic performance improvement by doubling REDNMC (red curve versus green curve).

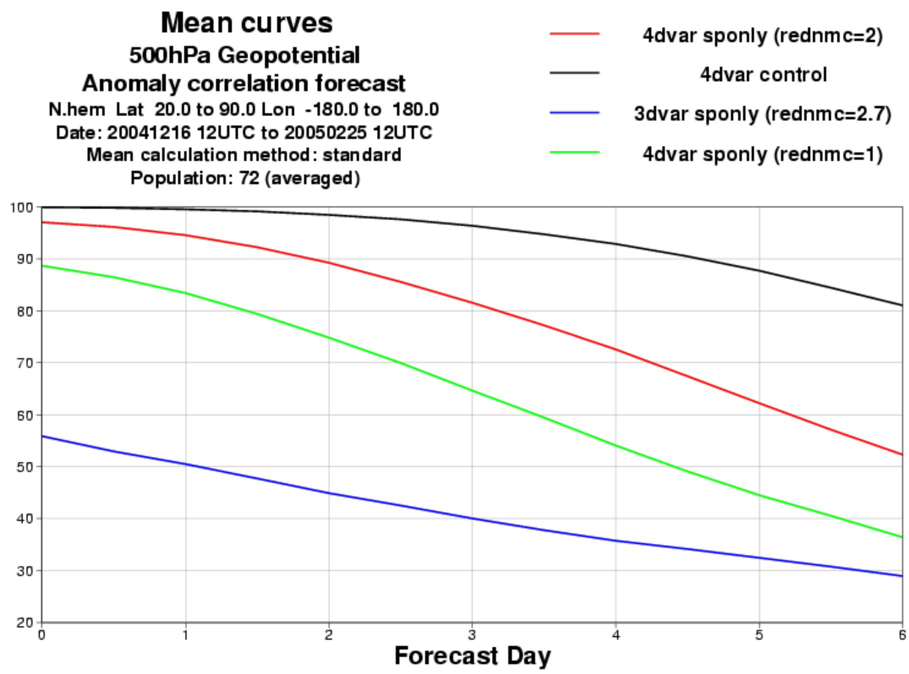

Figure 16: 500hPa geopotential Northern Hemisphere height field anomaly correlation forecast scores for 4D-Var control assimilation with all data types (black curve), 3D-Var surface pressure data only assimilation with retuned REDNMC (blue), 4D-Var surface pressure data only assimilation with default REDNMC=1 (green) and 4D-Var surface pressure data only assimilation with retuned REDNMC=2 (red). (This figure is from Thépaut (2006) - courtesy J.-N. Thépaut).

The main conclusion from this work is that tuning the statistics of the assimilation system when the Observing System is substantially degraded is essential. It highlights the need for an adaptive covariance model in these circumstances.

Conclusions

We have developed a streamlined ensemble data assimilation system at ECMWF. This is due to the considerable range of applications that now benefit from such a system. The lack of ensemble spread is the main problem with our ensemble data assimilation system and it a well known problem for most ensemble based systems. We have tried to remedy some of this by including an improved representation of model error to account better for the errors that cannot be represented by just perturbing observations. A recently developed spectral backscatter scheme has in this been investigated thoroughly in the context of ensemble data assimilation. The results have been very promising. First it can be concluded that the perturbations introduced by the SPBS are more realistic and balanced than the previously used stochastic physics from the operational ensemble prediction system at ECMWF. The SPBS perturbations are retained to a large extent by the analysis, whereas the previously used stochastic physics perturbations caused extensive amount of gravity waves to be generated, especially in the tropics. This made it impossible to use in the ensemble data assimilation system, because the gravity waves resulted in unrealistic J_b statistics with very large unrealistic variances. The SPBS scheme has not got these problems and has been used in our ensemble data assimilation experiments with success. The issues of concern with respect to the use of SPBS in the ensemble data assimilation has been that the method mainly corrects for model errors in the tropics, so the spread has been increased more in the tropics than in the extra-tropics. It may well be that the model errors are larger in the tropics than elsewhere, but this leads to a change in the ratio between tropical and extra-tropical variances for the calculated J_b statistics.

The use of flow dependent background error variances based on the ensemble DA spread does not improve the general scores, like 500hPa anomaly correlations for large domains. But we do see impact near tropical cyclones, troughs and extra-tropical cyclones. This is promising, because these are cases where you would hope and expect flow dependence to be important. Fisher (these proceedings) concluded that "Investigations with very simple zero- and one-dimensional analogues confirm this result, and suggest that for a factor-of-two range of background error standard deviation around the true value, the expected analysis error remains within a few percent of its optimal value, resulting in a reduction of forecast skill of only a few hours." In this second part of this paper we have further investigated the causes for the lack of general impact on the set of full system experiments performed with varying REDNMC (controls ratio of observation error and background errors). Our main conclusions are that general circulation is in ECMWF's assimilation system determined by the broad temperature/wind structures that are well described by the high volume satellite data. Flow dependence is not very important for these structures that have a slow error growth. The forecast model can generate and evolve cyclones on its own from these accurate but broad initial conditions. We should only expect significant impact of flow dependent background errors for extreme events that due to their rarity has a small impact on measures like general 500hPa height field scores. Despite this the work is still justified because it is important to predict the extreme events well. Finally we can conclude that despite the ensemble data assimilation being under-dispersive, it is beneficial to use ensemble data assimilation based estimates of short range forecast errors in the EPS system (see Leutbecher *et al*. these proceedings).

The next step, which is already in progress, will be to take account of horizontally correlated radiance errors along the lines already applied in the ensemble data assimilation for AMV (geostationary satellite atmospheric feature tracking winds). Another line of work is to improve the SPBS representation of model error in the ensemble DA. It would be beneficial to run a research mode ensemble DA system in real time to learn from daily monitoring and this would also allow calculation of seasonal variance estimates. One question is the resolution for such a system; we expect something like T399 outer loop/T159 inner loop, 91 levels and 10 members.

So far we have only investigated the impact of flow dependent variances, and even that in a very crude way. When the above mentioned improvements have been implemented, the next step will be to explore how we can benefit from the flexible wavelet J_b formulation used at ECMWF to ease introduction of scale selective flow dependent variances, correlation lengths and structures.

Acknowledgements

Jean-Noël Thépaut is acknowledged for allowing us to include some of his results on surface pressure data only assimilations. Edit Hagel is acknowledged for producing the analysis error estimate Fig. 2.

References

Bormann, N., Saarinen, S., Kelly, G. and Thépaut, J.-N. 2003: The Spatial Structure of Observation Errors in Atmospheric Motion Vectors from Geostationary Satellite Data. *Monthly Weather Review* **131**, 706-718

Dee, D., 2005: Bias and data assimilation. *Q.J.Roy.Met.Soc*, **131,** 3323-3343.

Kucukkaraca E. and M. Fisher, 2006: Use of analysis ensembles in estimating flow-dependent background error variances. *ECMWF Research Department Technical Memorandum* 492. 18 pages.

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

Schutts G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Q.J.Roy.Met.Soc*, **131,** 3079-3102.

Fisher, M. and Courtier, P., 1995: Estimating the covariance matrices of analysis and forecast error in variational data assimilation. *ECMWF Research Department Technical Memorandum* 220. 29 pages.

Jean-Noël Thépaut (2006): Assimilating only surface-pressure observations in 3D and 4D-Var. *ECMWF/GEO Workshop on Atmospheric Reanalysis*, 19-22 June 2006, 107-112.

Fisher, M. (2003): Background Error Covariance Modelling. In proceedings from “Recent developments in data assimilation for atmosphere and ocean”, *ECMWF Seminar*, 8-12 September 2003. 45-63.