TECHNICAL MEMORANDUM

726

Representing model error in ensemble data assimilation

Carla Cardinali, Nedjeljka Žagar[1], Gabor Radnoti, Roberto Buizza

Research Department

[2] University of Ljubljana and Center of Excellence SPACE-SI, Slovenia

Accepted for publication in *Nonlinear Processes in Geophysics*

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

Summary

The paper investigates a method to represent model error in the ensemble data assimilation (EDA) system. The ECMWF operational EDA simulates the effect of both observations and model uncertainties. Observation errors are represented by perturbations with statistics characterized by the observation error covariance matrix whilst the model uncertainties are represented by stochastic perturbations added to the physical tendencies to simulate the effect of random errors in the physical parameterizations (*ST*-method). In this work an alternative method (*XB*-method) is proposed to simulate model uncertainties by adding perturbations to the model background field. In this way the error represented is not just restricted to model error in the usual sense but potentially extends to any form of background error. The perturbations have the same correlation as the background error covariance matrix and their magnitude is computed from comparing the high-resolution operational innovation variances with the ensemble variances when the ensemble is obtained by perturbing only the observations (*OBS*-method). The *XB*-method has been designed to represent the short range model error relevant for the data assimilation window. Spread diagnostic shows that the *XB*-method generates a larger spread than the *ST*-method that is operationally used at ECMWF, in particular in the extra-tropics. Three-dimensional normal-mode diagnostics indicate that *XB-EDA* spread projects more than the spread from the other EDAs onto the easterly inertia-gravity modes associated with equatorial Kelvin waves, tropical dynamics and, in general, model error sources.

The background error statistics from the above described EDAs have been employed in the assimilation system. The assimilation system performance showed that the *XB*-method background error statistics increase the observation influence in the analysis process. The other EDA background error statistics, when inflated by a global factor, generate analyses with 30-50% smaller degree of freedom of signal. *XB-EDA* background error variances have not been inflated.

The presented EDAs have been used to generate the initial perturbations of the ECMWF Ensemble Prediction System (EPS) of which the *XB-EDA* induces the largest EPS spread, also in the medium-range, leading to a more reliable ensemble. Compared to *ST-EDA, XB-EDA* leads to a small improvement of the EPS ignorance skill score at day-3 and 7.

Keywords: Analysis uncertainty, Ensemble of initial conditions, Ensemble modelled background covariance matrix, Ensemble prediction

1 Introduction

Data assimilation systems combine observations and background state, usually a short-range forecast of 6 or 12 hours. Alternative approaches to a deterministic initial condition system based on ensemble methodologies such as the ensemble Kalman filter or the ensemble variational analysis are nowadays widely used in numerical weather prediction (NWP). They have the advantage of providing information on flow dependent background error covariances. Nevertheless, ensemble data assimilation systems (EDA) are usually implemented in the context of a perfect model and the failure of representing model error affects for example the computation of the background error variances, which tends to be underestimated. Different strategies have been tested to take model error into account by rescaling the ensemble of analyses with constant, isotropic factors (inflation), the multi-model and multi-physical parametrization approach (Houtekamer *et al.* 1996; Houtekamer *et al.* 2005), the stochastic physical tendency (Buizza *et al* 1999) and the backscatter stochastic kinetic energy schemes (Shutts 2005, Berner *et al.* 2008). Comparison among different model error representations have been performed by Houtekamer *et al. (*2009), which indicates that the inflation approach has the largest contribution in their model error simulation.

Several operational NWP centres such as Météo-France, UK Met-Office and Environment Canada have implemented an EDA system. Recently, Météo-France has implemented a 6-member EDA where not only the observation uncertainties are represented but also the model uncertainties are explicitly taken into account. This is done by inflating the background field with a latitude-level varying factor in the ensemble (see Raynaud et al 2009 and 2012). Their first operational EDA configuration (Berre *et al.*, 2007) has been used in operations since July 2008. Flow-dependent background error variances for the operational 4D-Var assimilation system (Raynaud *et al.*, 2011) were derived within a perfect model framework and the estimated variances were inflated 'off-line' (i.e. after the ensemble has been completed) by using a posteriori diagnostics (Desroziers and Ivanov, 2001). The inflation aims at representing model error contributions. Multiplicative inflation is in fact a simple and widely used technique to deal with unknown error sources. Because the off-line multiplicative inflation that was applied to the variances was not accounted for in the background perturbation update, the recently implemented operational EDA configuration includes the multiplicative inflation to enlarge the amplitude of forecast perturbations within the ensemble.

The Environment Canada ensemble system is based on a Kalman filter (Houtekamer and Mitchell 2005 and Houtekamer *et al* 2005) and has been used operationally since January 2005. It provides an ensemble of initial conditions for the medium-range EPS (Ensemble prediction System) and represents both observations and model sources of uncertainties. In particular the model error component has been extensively investigated. Initially a simplified and reduced amplitude form of the Canadian 3D-Var background error covariance was used to perturb the ensemble of background fields (Houtekamer *et al* 2005). Then, different ways to determine the covariance for the additive model error component have been investigated (Hamill and Whitaker, 2005) and model error perturbations are added to the ensemble analysis rather than to the background ensemble (Houtekamer and Mitchell 2005) to account for the data assimilation weakness. More recently, each member has a different model version (Meng and Zhang, 2007; Fujita et al. 2007) to represent uncertainties in model representation of physical processes. As already said above, Houtekamer and Mitchell (2005) concluded that the addition of isotropic model error perturbations to the ensemble of analyses is found to have the largest impact in terms of ensemble spread.

An EDA has been operationally implemented at ECMWF in June 2010 (Isaksen et al 2010). The EDA ensemble consists of ten independent members of lower resolution (with respect to the high-resolution operational 4D-Var system) 4D-Var data assimilation systems with perturbed observations and perturbed model tendencies. In particular, the observation uncertainties are represented by perturbing the observations and the model uncertainties by adding stochastic perturbations to the model tendencies during the first 12-hour model evolution using the Stochastically Perturbed Parameterisation Tendency scheme (SPPT, see Palmer *et al* 2009 for a review).

The ECMWF EDA provides a flow-dependent or daily model background error covariance matrix that is supposed to improve the high resolution analysis system by better representing the daily dynamical synoptic features (Raynaud *et al*, 2008, 2009 and 2011; Buehner *et al*. 2010). Since its implementation, the EDA has been used together with the singular vectors to initialize the operational EPS (Molteni *et al* 1996; Buizza *et al*. 2007) and to improve the simulation of initial uncertainties (Buizza *et al* 2008), one of the fundamental aspects of the EPS design.

In this paper, a different way of representing model error in the operational ECMWF EDA is presented and compared to the standard SPPT method. The model uncertainties are represented by adding perturbations to the model background field. The magnitude of the perturbations varies with vertical level and with geographical latitude. They are estimated from a comparison between the innovation variance of the high resolution 4D-Var system, i.e. the difference between observation and background at the observation location, and the ensemble data assimilation variance (variance taken over an ensemble of assimilations over 3 weeks period) in which only observation uncertainties are represented. The model error representation is therefore similar to the one introduced by Raynaud et al (2012) at Météo-France that is referred to as the multiplicative perturbation method. The method described here is denoted as an additive perturbation method. The error represented is not restricted to the model error in the usual sense i.e. the error that would be present in the forecast even if the initial condition were exact, but is related to any form of error for example errors in the background covariance matrix coming from the operational ECMWF EDA. The present paper studies the EDA sensitivity to the different model error representations. The proposed method is compared to the operational one, which uses the stochastically perturbed parameterisation tendency scheme to simulate model uncertainties, and to the EDA obtained by representing only observation uncertainties. A fourth EDA has also been designed to just quantify the impact of background cycling in the EDA where only observations are perturbed. Observation uncertainties are always equally represented in all EDAs examined.

Section 2 describes the methodology used to simulate model uncertainties. Section 3 analyses the spread characteristics of the investigated EDAs by using a variety of diagnostics and with an evaluation of the background error covariance matrix provided by the EDAs. The EPS performance sensitivity to the EDAs is also discussed. Conclusions are drawn in Section 4.

2 Representation of uncertainties

2.1 *The XB-EDA*

An ensemble of analyses attempts to generate a representative sample of possible states of a dynamical system. The samples are generated by the same assimilation system. From the optimal solution of the analysis problem, $\mathbf{x}_a = f(\mathbf{x}_b, \mathbf{y})$, two input parameters can be identified: the observation vector y and the background vector x_b obtained from a short-range forecast, respectively. An ensemble of analyses can be generated by perturbing both input vectors. In particular the observations uncertainties can be represented by perturbing vector y, whilst the model uncertainties (at least the short-range model error) can be represented by perturbing the model state vector x_b. The perturbed analysis equation can be written as:

$$\tilde{\mathbf{x}}_a = f(\mathbf{x}_b + \boldsymbol{\zeta}, \mathbf{y} + \boldsymbol{\eta}) \tag{2.1}$$

where ζ and $\boldsymbol{\eta}$ are perturbations defined as

$$\begin{aligned} \boldsymbol{\zeta} &= f(\lambda, l, x)\mathbf{B}^{1/2}\tilde{\boldsymbol{\zeta}} \\ \boldsymbol{\eta} &= \mathbf{R}^{1/2}\tilde{\boldsymbol{\eta}} = \sigma_o \tilde{\boldsymbol{\eta}} \end{aligned} \tag{2.2}$$

with $f(\lambda, l, x)$ being a function of latitude (λ), model level (l) and model parameters (x). $\tilde{\boldsymbol{\zeta}}$ and $\tilde{\boldsymbol{\eta}}$ are samples of vectors drawn from a multi-dimensional Gaussian distribution with zero mean and identity covariance matrix. To achieve that the final perturbations $\boldsymbol{\zeta}$ and $\boldsymbol{\eta}$ have a covariance matrix specified by B and R, respectively, the square root of B and R, is applied to the sequence of normally distributed vectors $\tilde{\boldsymbol{\zeta}}$ and $\tilde{\boldsymbol{\eta}}$. B and R are the estimated background and observation and error covariance matrices; they are therefore only approximations of the true covariance matrices. When R is diagonal (i.e. uncorrelated observation errors) a simple multiplication by the observation error standard deviation σ_o is applied (Eq. 2.2). Only two sets of observations are perturbed with spatially correlated patterns. One is the Atmospheric Motion Vector (AMV) observation (Bormann *et al*. 2003) and the other set is the sea-surface temperature field (Vialard *et al*. 2005).

The magnitude of the final perturbation $\boldsymbol{\zeta}$ is determined by $f(\lambda, l, x)$ which is estimated by comparing the variance of the innovation vector d (over 3 weeks) with the ensemble data assimilation variance, VAR(EDA) (estimated over an ensemble of assimilations over 3 weeks period), in the case the ensemble data assimilation is obtained by only perturbing the observations The coefficient $f(\lambda, l, x)$ is meant therefore to compensate for the discrepancy between the background error as obtained from the innovation d on the one hand, and the a priori background error covariance matrix B on the other. The innovation vector is the difference between the observation vector y and the background counterpart of the observation computed by using the nonlinear observation operator ($H(x_b)$). Under the assumption of un-biasedness of the errors and de-correlation between the background and observation errors, the background error variance, as obtained from the innovation, is $Var(\mathbf{d}) - \sigma_o^{\,2}$. The scalar function $f(\lambda, l, x)$ is hence defined as:

$$\tilde{f} = f(\lambda, l, x) = \sqrt{\frac{Var(\mathbf{d}) - \sigma_o^{\,2} - Var(\mathbf{EDA})}{Var(\mathbf{d}) - \sigma_o^{\,2}}} \tag{2.3}$$

where σ_0^2 is the prescribed observation error variance. If $Var(\mathbf{d}) - \sigma_o^{\,2}$ is less or equal to the EDA variance it is imposed that $f(\lambda, l, x) = 0$. The perturbation amplitude modulation hence varies in the interval [0,1]. The innovation variances have been computed for 10 hPa pressure layers for atmospheric measurements located between the surface and 50 hPa (wind observations), between surface and 5 hPa (temperature observations), and surface and 300hPa when humidity observations are considered. Three latitude bands, namely Northern Hemisphere (20°N, 90°N), Southern Hemisphere (20°S,90°S) and Tropics (20°S,20°N) and a three-week data set have been considered. For the u and v component of the wind all conventional observations (radiosondes, pilots, synops, aircrafts and

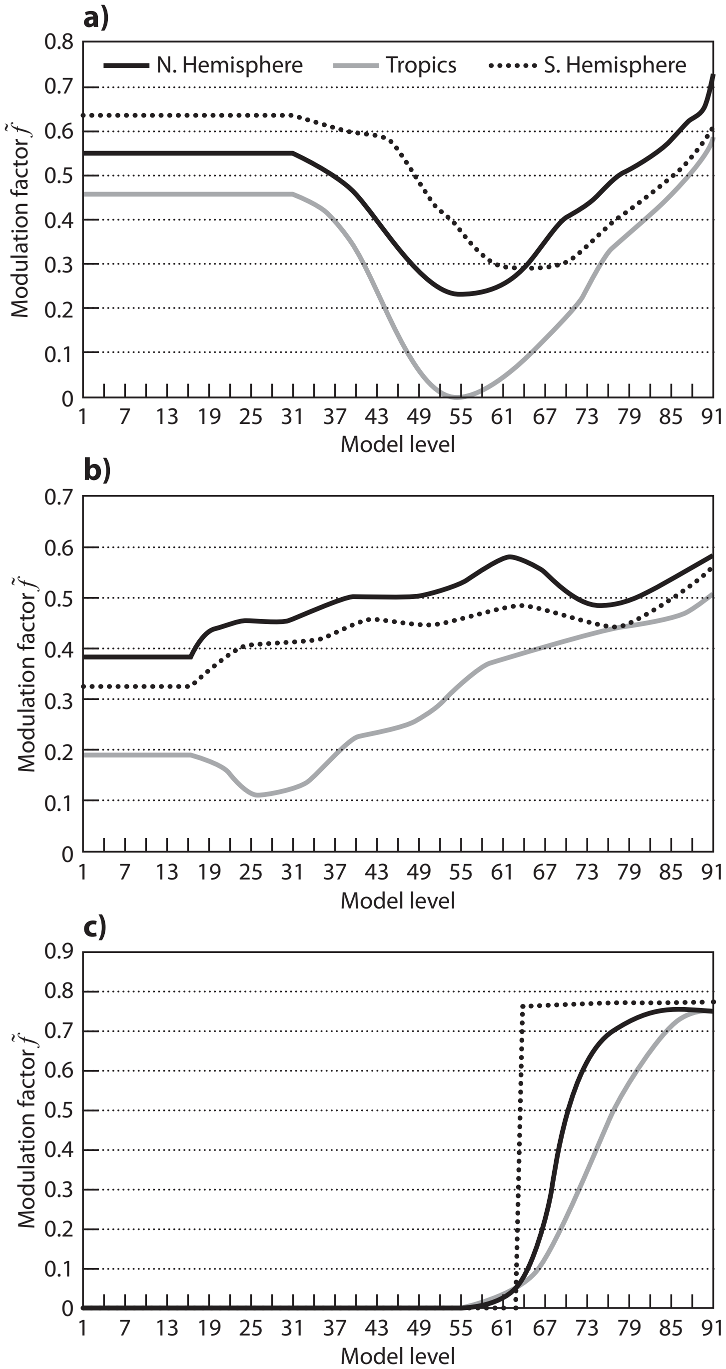

Figure 1 Perturbation modulation factor $\tilde{f}$ as a function of latitude band (North Hemisphere solid black line, South Hemisphere dotted line and Tropics solid grey line) and vertical model levels (model level 1 is 0.01 hPa and 91 is on average ~1000hPa) for a) zonal wind, b) temperature and c) humidity. The modulation factor is estimated over a 3 weeks period.

profilers), Atmospheric Motion Vectors (AMV) and scatterometer observations have been used to compute the innovation variance. For temperature conventional (radiosondes and aircrafts) and AMSU-A observations, for humidity radiosondes and All-Sky (SSM/I and TMI radiances) observations and, finally, for surface pressure all land and ocean stations have been used. The variation of $\tilde{f}$ with latitude band, model level and for model parameters u, T and q, is shown in Fig.1. Figure 1a shows that for the u-component $\tilde{f}$ decreases in the troposphere, when pressure increases, down to zero in the tropics and down to 0.3 in the extra-tropics (level 1 at 0.01 hPa, identifies the top of the atmosphere). If observations are unavailable to estimate the innovation variance, the modulation factor $\tilde{f}$ is kept constant i.e. from model level 1 to 30 for wind, from model level 1 to 18 for temperature and model level 1 to 55 for humidity. Similar results for the modulation factor are obtained for the v component of the wind (not shown). In the lower troposphere close to the surface the modulation factor globally increases on average up to 0.6. For temperature (Fig.1b) its magnitude increases with the increase of pressure on average for the three latitude bands from 0.3 to 0.5, the tropical modulation factor always being the smallest.

For humidity (Fig.1c) $\tilde{f}$ rapidly grows with the atmospheric pressure level up to 0.8 (South Hemisphere) towards the surface. Concerning surface pressure, the correction (2.3) (not shown) is globally constant and around 0.4. During the cycling, $\tilde{f}$ has been re-computed for retuning purposes using Eq 2.3 every three days and by using the past 3-days' variance sample. However, the modulation factor has stabilized rather fast after two days of cycling.

The background is perturbed at the start of each assimilation window and the innovation vector d= (y-Hx^b) is computed along the trajectory starting from the perturbed background to correctly take into account the background changes (H is the non-linear model and observation operator) and to produce a balanced perturbed field. Figure 2 schematically represents the realization of the described *XB-EDA*

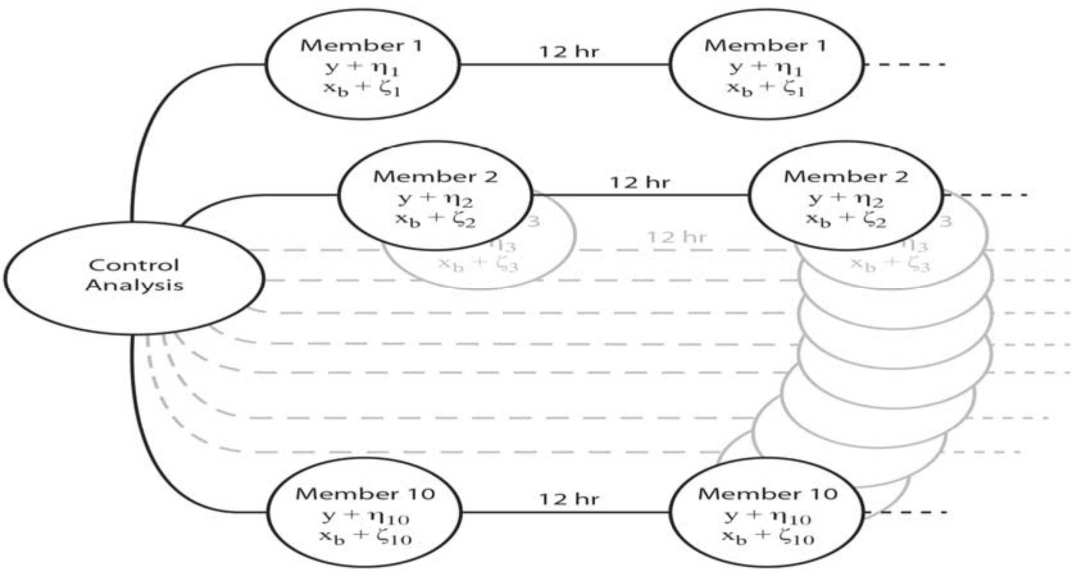

Figure 2. Schematic diagram of the XB-EDA realization

ensemble. From a control (unperturbed) analysis the two set of perturbations, η for the observations and ζ for the background, are respectively added at the beginning of the 4D-Var assimilation window to create ten different initial conditions (members). The two sets of perturbations are recomputed and added to the observations and the background field at every analysis cycle.

The ζ perturbation accounts for the short-range model error sources including the fraction of the analysis error that is due to the model error. Sources of error are therefore not only related to physical parameterisations but also to the dynamics, the spatial and temporal discretization, the linearized physical process and the misspecification of the probability distribution of errors in the observations and the background model.

2.2 *The other-EDA*

In all EDAs presented here the observations uncertainties are represented by perturbing the observations whilst the model uncertainties are produced differently. Table 1 shows all the EDA configurations investigated. In particular *OBS-EDA* is the ensemble analysis where only observation uncertainties are represented (different observation values generate different background fields during the cycling process). The *OBS-OBS* ensemble is similar to the *OBS-EDA* one but the background fields are not cycled. At every cycle the 10-member background fields are replaced with the background field of the control i.e. unperturbed analysis (Fig. 2). *OBS-OBS* is performed to evaluate the impact of the cycling by the background. The model error representation in *ST-EDA* is based on the SPPT scheme (Buizza *et al*. 1999, Palmer et al. 2009). Since November 2009, the operational EPS uses SPPT and the stochastic back scatter schemes (SPBS, Shutts, 2005; Berner *et al*. 2008) to simulate model uncertainties. The SPBS scheme simulates the inverse energy cascade due to the interaction between the unresolved and the resolved scales, and aims to compensate for the over-dissipation occurring in numerical models. The SPPT scheme, designed to simulate random model errors due to physical parameterisations, is still assumed to explain the largest source of model error in the EPS. The operational EDA configuration does not use the SPBS scheme but only the SPPT. In particular, SPPT model uncertainty is simulated using the 1-scale version of the stochastically perturbed parametrization tendency scheme, which perturbs the total parametrized tendency of physical processes. Since the 1-scale version use a time-scale of 6-hours, there is no need to cycle the model error perturbation across different data-assimilation cycles. In the 1-scale version of the SPPT, the perturbations to the physical tendencies are defined to have a spatial correlation length of 500 km and a time-correlation of 6 hours, as in the original SPPT scheme (Buizza et al 1999).

EDA	Methodology	Inflation
OBS	Perturbation added to observations	Y
OBS-OBS	Perturbation added to observations; Members background fields are from Control An.	Y
ST	Perturbation added to observations; Perturbation added to physical parameter tendency	Y
XB	Perturbation added to observations; Perturbation added to background	N

Table 1: Ensemble data assimilation configurations.

Because of the presumed ensemble analysis underdispersivity (to be confirmed later) a global inflation factor (=1.4) has been applied to the background error standard deviation in the *OBS, OBS-OBS* and *ST-EDAs* methods to increase the ensemble spread and to penalize the model background further with respect to the observations in the assimilation process. The static background covariance matrix has, in fact, always been inflated in the ECMWF 4DVar assimilation system to avoid the excessive weight given to the background with respect to the observations. Indeed, studies on the observational influence in the analysis system have shown that globally and for a given assimilation cycle only 15% of the information was provided by the observations while the remaining 85% were due to the background (Cardinali *et al* 2004, Cardinali 2013). Unfortunately, the inflation is a constant that does not vary with respect to the parameters, with respect to the geographical location or weather situation and the resulting ensemble spread is simply globally amplified.

3 Results

In this section EDAs with different model error representations are compared and diagnosed. Each EDA includes 10 perturbed and 1 unperturbed 12-hour 4D-Var assimilations (Rabier *et al*. 2000, Janiskova *et al*. 2002; Tompkins and Janiskova, 2004; Lopez and Moreau, 2005) at the resolution of T_L399L91 (spectral triangular truncation with 399 wave numbers and linear grid, and on 91 vertical levels) for the model forecast and T_L159L91 for the minimization calculation, respectively.

The 4 EDA schemes (Tab. 1) have been run for the period 20081001-20081115, with twice daily 12-hour assimilation cycles using observations from (2100-0900] UTC and (0900-2100] UTC.

3.1 *EDA spread*

In Fig. 3 the averaged spread of the four data assimilation ensembles is compared for the zonal wind component. The average spread has been computed over the period 20081005-20081115 (the first 5 days of the EDA computation have not been included in the evaluation to take into account 'spin-up') from 6-hour forecasts according to the expression:

$$Spread(EDA) = E\left(\sqrt{\frac{\sum_{i=1}^{N}(m_i - \bar{m})^2}{N-1}}\right) \tag{3.1}$$

where m_i is the i^{th}-ensemble member, N=10 and $\bar{m}$ is the ensemble mean. Expectation stands for averaging over longitude and over the selected period.

When only observation uncertainties are represented (u wind component *OBS-EDA,* Fig. 3a), the spread is mainly confined to the upper stratosphere (above model level 20 i.e. ~10 hPa) and to the troposphere (below model level 40 i.e. ~110 hPa) in the tropics. Very little spread is accomplished poleward of 40°N or 40°S. To understand how much of the spread is due to the cycling of *OBS-EDA* over successive assimilation windows, Fig. 3b shows the *OBS-OBS* spread. The *OBS-OBS* spread is mainly confined to the stratosphere and with a smaller amount to the tropical troposphere. When

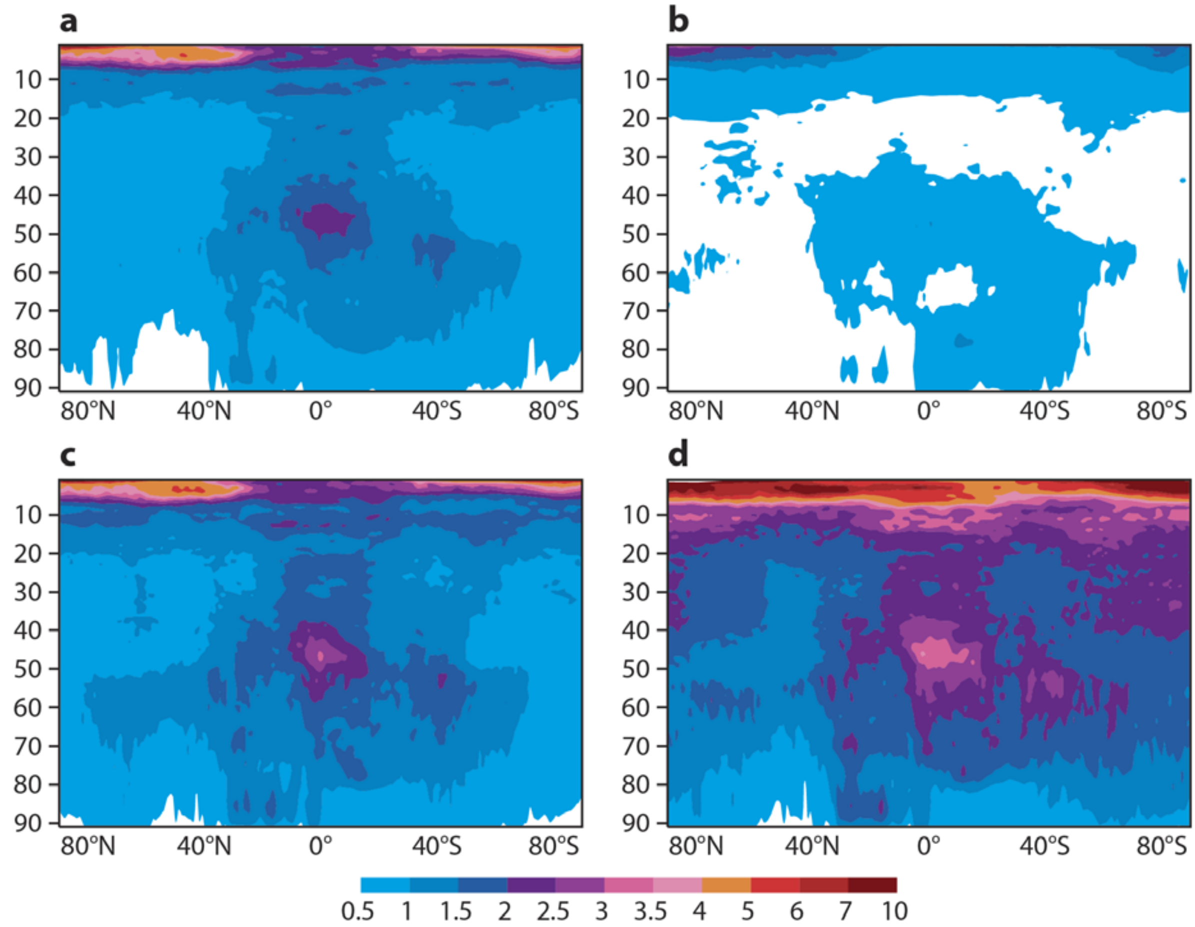

Figure 3 Zonally averaged cross section for period 20081005 to 20081115 for the u-component of the EDA spread. a) OBS-EDA, b) OBS-OBS-EDA, c) ST-EDA and d) XB-EDA. Vertical coordinate is model level. Contours are in ms^{-1}

model errors are explicitly represented in the ensemble the spread increases according to the method applied. Compared to *OBS-EDA*, *ST-EDA* presents 10% larger spread in the tropics and the mid-latitudes (Fig.3c). The *XB-EDA* ensemble (Fig. 3d) shows the largest and more globally distributed spread. It amounts to ~1 ms^{-1} and 4 ms^{-1} larger than *ST-EDA* in the tropics and the stratosphere, respectively. It is also remarkably different in the high and medium extra-tropical troposphere and differences exceed 4-5 ms^{-1} in the mid-latitude. Blank areas are values between 0 and 0.5.

Figure 4a shows the difference of spreads in the *OBS* and *OBS-OBS EDAs* and visualizes the impact of cycling over successive assimilation windows. Figure 4b shows difference of spreads in *XB* and *OBS EDAs* and visualizes the impact of the perturbation of the background (first eq. of 2.2). Most of the former difference is located in the stratosphere and in the tropics while the second extends to a large part of the troposphere, especially in the Southern hemisphere. The figures also suggest that the differences are everywhere positive (blank areas are values between 0 and 0.5).

Figure 5 shows the reduction of globally averaged spread of the *OBS-*, *OBS-OBS* and *ST-EDAs* relative to the *XB-EDA* for each model level and the u component of the wind. The spread in *XB-EDA* is a function of the amplitude of the perturbation applied to the background field. The spread loss with respect to the *XB* ensemble is decreasing with the increase of model level for *ST* and *OBS EDA* whereas it is constant for the *OBS-OBS EDAs*. Close to the surface (i.e. model level 91) the first two EDAs lose 20% and 40% spread and close to the top of the atmosphere 40-50% on average at al latitudes, respectively.

In all EDAs the largest spread is located in the stratosphere where also the largest loss with respect to *XB* is observed. Similar results are obtained for the temperature field but the magnitude of the spread loss is 25% smaller (not shown).

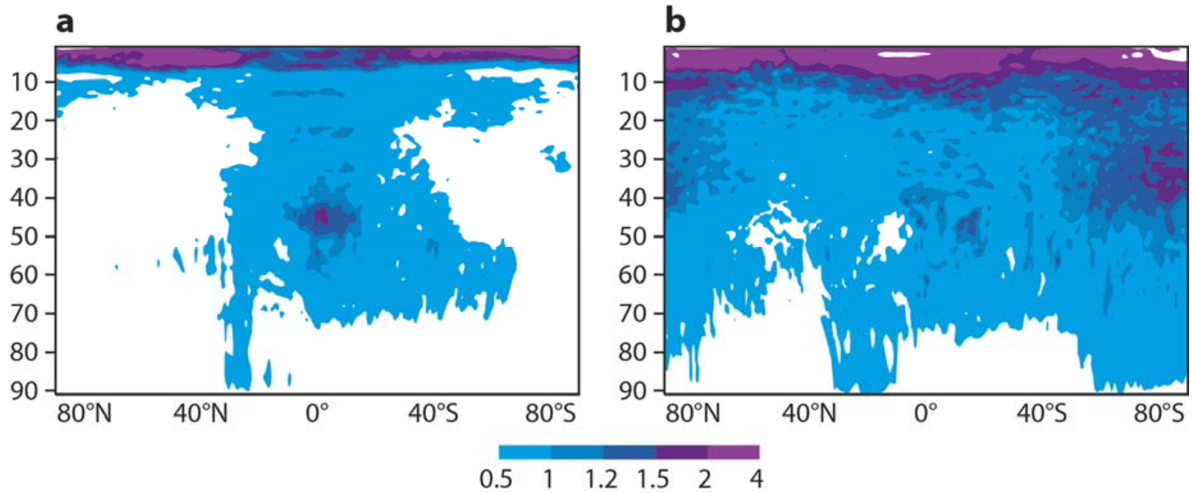

Figure 4 Zonally-averaged cross section from 20081005 to 20081115 for the u-component of the EDA spread. a difference in spreads between the OBS-EDA and the OBS-OBS-EDA spreads, b) difference between the XB-EDA and the OBS-EDA spreads. Vertical coordinate is model level. Contours in ms^{-1}

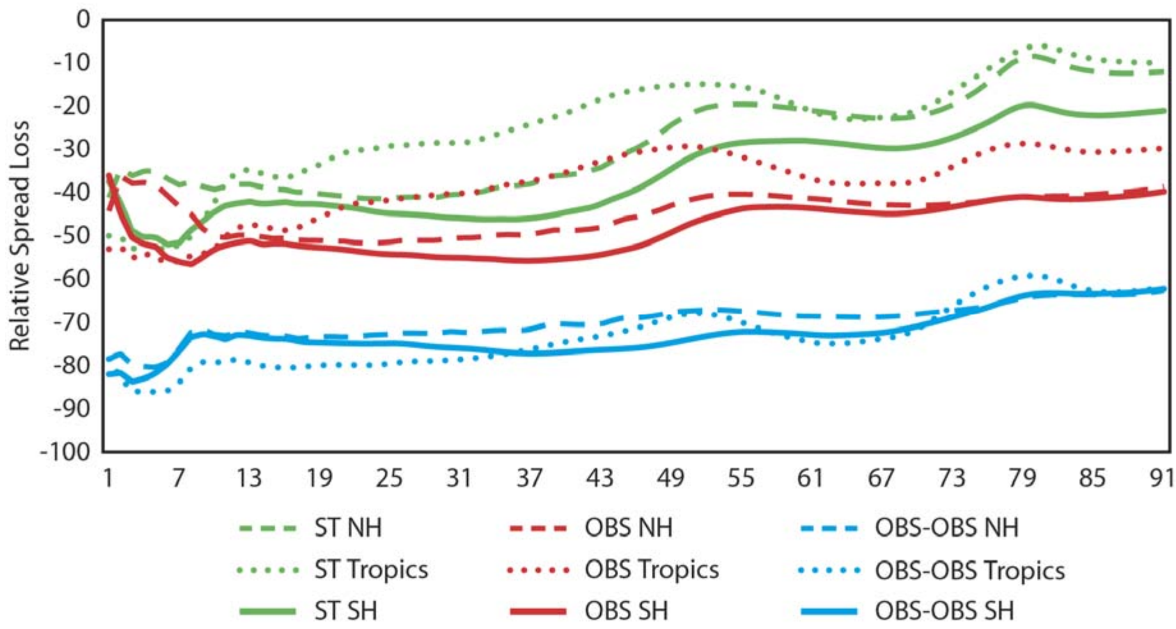

Figure 5 Zonal wind field relative average spread loss of OBS, ST, OBS-OBS EDA with respect to XB EDA at every model level and for Northern (dashed line) and Southern (solid line) Hemisphere and Tropics (dotted line).

3.2 *Spread case study*

An example of spread differences among the *OBS-, ST- and XB-EDAs* in physical space is presented for 20-21 October 2008 in Fig.6. In this period an intense baroclinic development took place over the northern west Pacific. In addition to a mature-stage cyclone moving toward the Gulf of Alaska, a deep cyclogenesis took place about 4000 km westward at about 50^{o}N, 180^{o}W (not shown). All methods produce spread associated with the two cyclones but differences exist in the structure and magnitude of the spread. The comparison of 6-hour forecast vorticity spread at 850 hPa valid on 21 October at 12 UTC (Fig. 6) shows that the *OBS-EDA* spread associated with the western cyclone is not only smaller than in the other two EDAs but also located on the north and north-eastern side of the system (Fig. 6a). The *ST-* and *XB-EDA*s both contain spread over a larger area. For the mature-stage cyclone in the eastern Northern Pacific, the spread in all three experiments has the typical comma shape (East Pacific, around 45°N-150°W) of frontal systems associated with baroclinic development. The *ST-* and *XB-EDAs* are the most similar although *XB-EDA* (Fig 6c) has a larger maximal amplitude and occupies a larger area. It is worth noticing that *XB-EDA* is the only ensemble showing some degree of uncertainty in the polar region where only very few observations are available and consequently the analysis uncertainties should be larger.

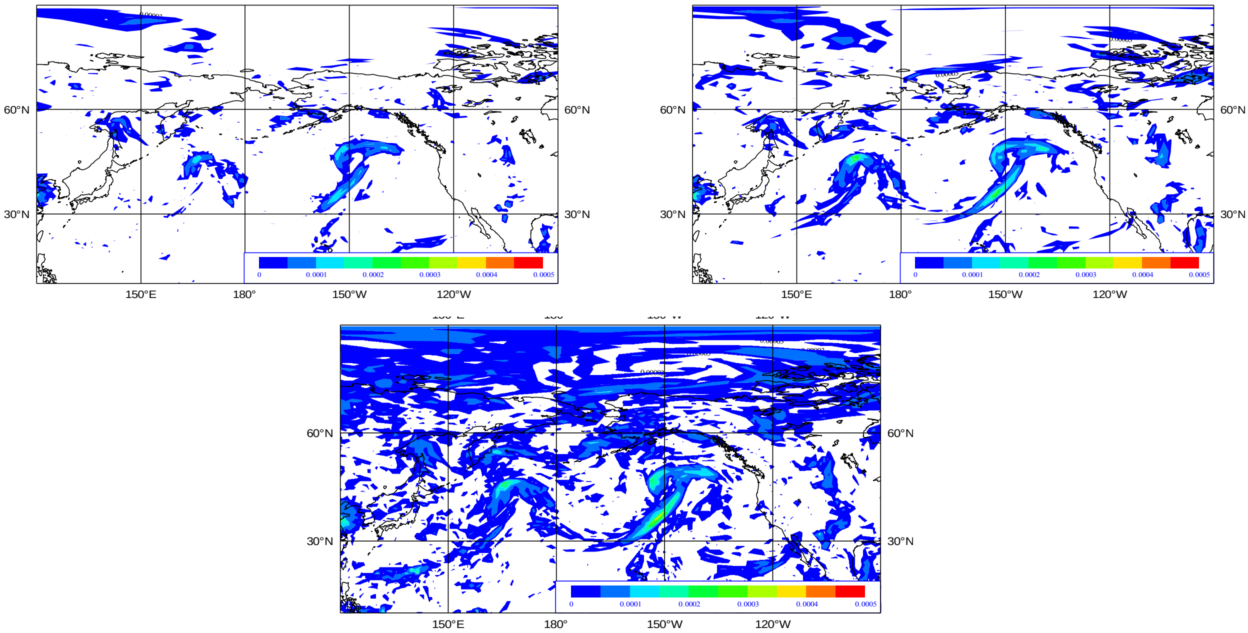

Figure 6: 6-hour forecast of vorticity field spread at 850 hPa valid on 21 October 2008 at 12 UTC. a) OBS-EDA (top-left), b) ST-EDA(top-right) and c) XB-EDA(bottom-left).

3.3 *Modal diagnosis of the ensemble spread*

The discussion of the results in the previous sections is complemented hereafter with the diagnosis of the ensemble spread in terms of normal modes as described in Žagar et al. (2011). The ten members of the four ensemble experiments are projected onto a set of three-dimensionally orthogonal vectors which are eigensolutions of the Navier-Stokes equations linearized about an horizontally homogeneous stable state of rest. This projection allows the attribution of the ensemble spread according to various horizontal and vertical scales as well as linearly balanced (quasi-geostrophic) and unbalanced (inertio-gravity, IG) parts of the flow. In Žagar et al. (2011), the method was applied to the analysis of the ensemble spread of the DART-CAM system whereas Žagar et al. (2013) used the method to study the balance properties of the ECMWF EDA system. In the ECMWF EDA for July 2007, it was found that about 50% of the short-range forecast-error variance was associated with the IG modes and that the eastward-propagating IG component was dominant on all scales. Both results were associated with the majority of EDA variance being present in the tropics. On the other hand, the ensemble spread of the DART-CAM ensemble was characterized by a prevalence of the westward-moving IG modes which was found to be related to the covariance inflation. These studies suggest that the normal mode function (NMF) expansion is a useful diagnostic of EDA systems.

In the present study, we followed Žagar et al. (2013) to analyze model levels under 10 hPa (model levels 19 to 91, totaling 73 levels) in order to avoid very large spread in the mesosphere (see Fig. 3) that projects strongly on the leading vertical modes and can obscure the interpretation of the results. For the presented diagnostics, 6-hours forecast starting at 18 UTC in the period 20081018 to 20081116 are used (30 samples) for 10 ensemble members. The analyzed data on the N64 Gaussian grid are projected onto 85 zonal wavenumbers, 50 vertical modes and 40 meridional modes for each motion type, namely balanced, eastward inertio-gravity (EIG mode) and westward inertio-gravity (WIG mode).

In modal space, the ensemble spread based on N ensemble members is defined as

$$S_\nu = \left[\frac{1}{N-1} \sum_{i=1}^{N} gH_\nu \left[\chi_{\upsilon,i} - \overline{\chi_\upsilon} \right] \cdot \left[\chi_{\upsilon,i} - \overline{\chi_\upsilon} \right]^* \right]^{1/2} \tag{3.2}$$

where $\chi_{\nu,i}$ is a non-dimensional complex projection coefficient for an ensemble member i while ν is a four-indices modal index which contains information about the zonal wavenumber, the meridional mode, the vertical mode and the wave type. The overbar stands for averaging over N ensemble members, i.e. $\overline{\chi_\upsilon}$ is the ensemble mean, $\overline{\chi_\upsilon} = \frac{1}{N}\sum_{i=1}^{N}\chi_{\upsilon,i}$. Each vertical mode is characterized by a value of 'the equivalent depth' H which couples horizontal and vertical motions. The spread computed by Eq. (3.2) applies at a single 6-hr forecast range and the results are presented as time averages over 30 samples. For details of the projection procedure see Žagar et al. (2011) and references therein. One difference between the normal-mode diagnostics and other spread evaluation methods consists in analyzing simultaneously the mass and wind fields allowing the physical interpretation of balance relations.

Spectra of the balanced, EIG and WIG energy spread as a function of the zonal wavenumber are shown in Fig. 7. In agreement with what has been presented so far, the ensemble spread in the *XB-EDA* dominates over the *ST-EDA* and *OBS-EDA* at all scales and for all three motion types. The *ST-EDA* spread is closer to the *OBS-EDA* spread than to the *XB-EDA* spread. In all experiments, the EIG spread dominates over WIG at largest horizontal scales and in *XB-EDA* it is greater than the WIG spread on all scales due to the equatorial Kelvin modes (not shown). The *XB-*, *ST-* and *OBS-EDAs* have a smaller percentage of their spread in the largest scale with respect to *OBS-OBS-EDA*. Below zonal wavenumber 10, there is around 30% of the total spread for all experiments, varying from 28% for *ST-EDA* to 34% for *OBS-OBS-EDA* (not shown).

On average, the total spread in *XB-EDA* is between 1.7 and 1.6 times greater than the *OBS-EDA* spread. The *ST-EDA* spread is around 1.25 times of the OBS-EDA spread. Both *XB*- and *ST-EDA* add relatively more spread in the IG part than in the balanced part with respect to *OBS-EDA*. Instead, the experiment without cycling (*OBS-OBS-EDA*) counts only for 40% of the spread of *OBS-EDA* for all modes (not shown).

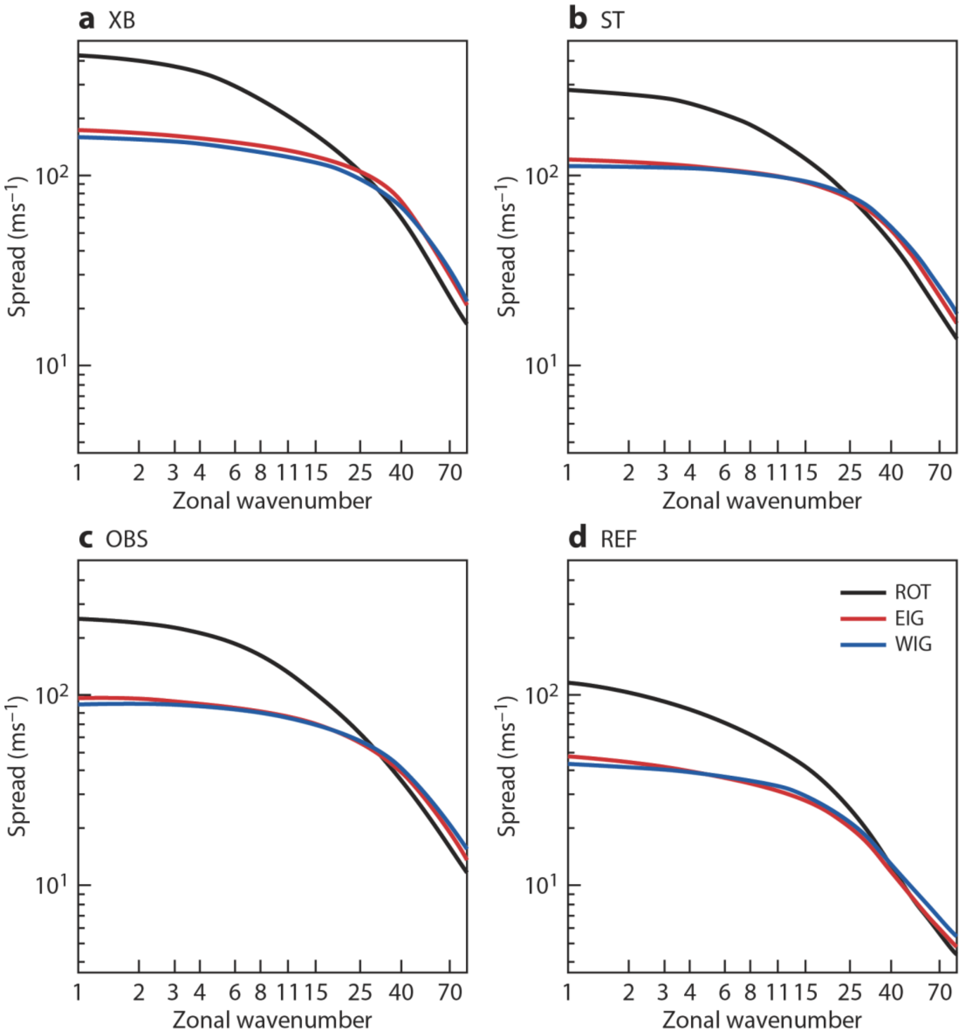

Figure 7. Time-averaged ensemble spread in the balanced and inertio-gravity modes for a) XB-EDA, b) ST-EDA, c) OBS- and d) OBS-OBS EDA. Black curves correspond to the spread associated with balanced modes (ROT), red curves to spread due to easterly-propagating inertia-gravity (EIG) modes whereas blue curves represent spread due to westerly-propagating inertia-gravity modes (WIG).

When the spread is summed up across all scales, the percentages of ROT, EIG and WIG spread in the four experiments is 43, 28, 29 % for *OBS-OBS-EDA*, 41, 29, 30 % for *OBS-EDA*, 39, 30, 31% for *ST-EDA* and 40, 31, 29 % for *XB-EDA*, respectively. Overall, the *XB-EDA* is the only experiment with the total EIG spread greater than the WIG spread. As can be seen in Fig. 8, which presents ratios between the balanced, EIG and WIG spread with the total spread as a function of the zonal scale, this applies for every zonal wavenumber. The dominance of the EIG spread in the *XB-EDA* is most likely associated with the larger tropical spread in this experiment (see also Fig. 3) and it is also in agreement with Žagar et al. (2013) who presented the same conclusion for the 3-hr and 12-hr forecast-error variances in an earlier model cycle. As discussed there, easterly propagating tropical modes represent the most important variability and largest forecast error source in the tropics.

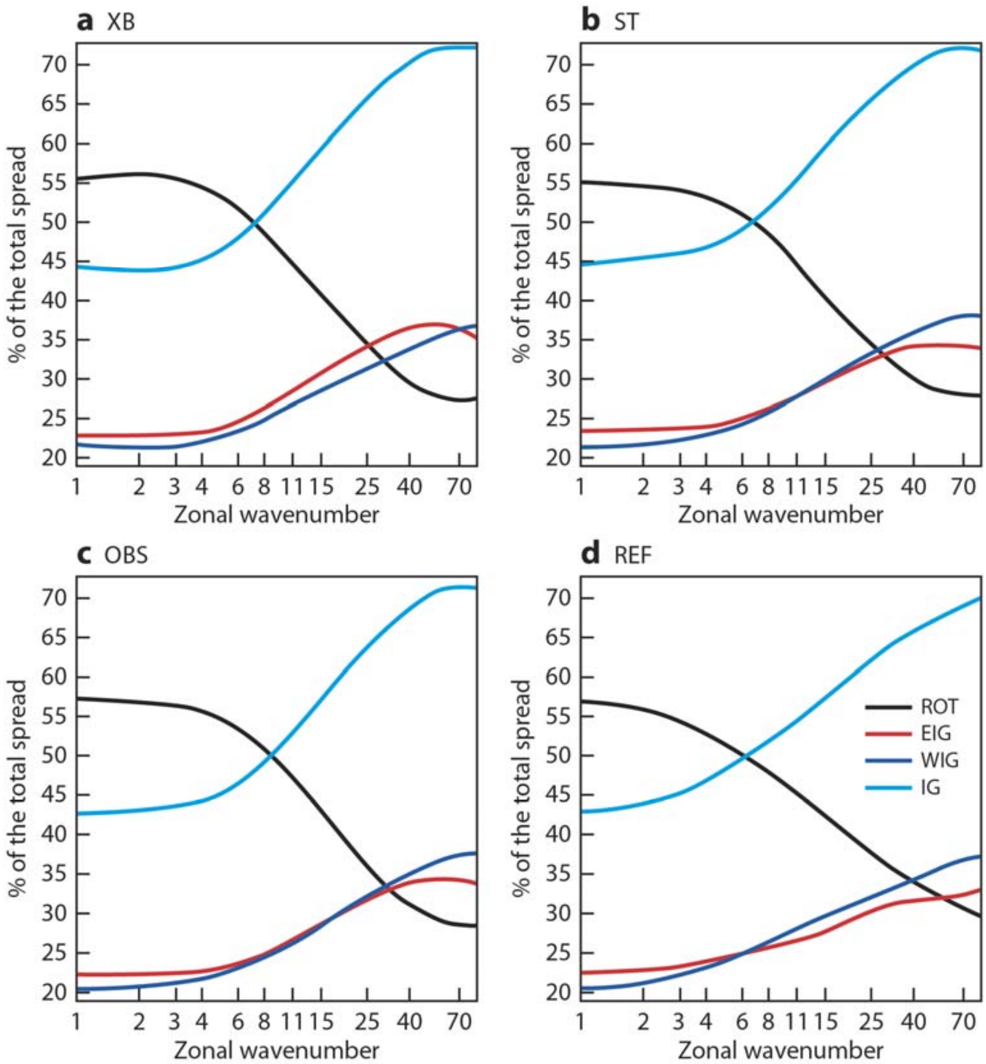

Figure 8. Ratios of the balanced (ROT), EIG, WIG and IG (EIG+WIG) spread to the total spread in each zonal wavenumber for (a) XB-EDA, (b) ST-EDA, (c) OBS-EDA and (d) OBS-OBS-EDA. The ratios are multiplied by 100. Black curves correspond to balanced modes (ROT), red to easterly-propagating inertia-gravity modes (EIG), dark blue to westerly-propagating inertia-gravity (WIG) while light blue curves correspond to all inertia-gravity modes (IG=EIG+WIG).

Contrary to *XB-EDA*, the *OBS-OBS-EDA* and *ST-EDA* experiments contain more WIG spread than EIG spread for zonal wavenumbers greater than 6 and 11, respectively (Fig. 8). If the spread in each experiment is normalized by *OBS-EDA*, it is found that *XB-EDA* is characterized by an increased EIG spread on all scales with respect to *OBS-EDA* while *ST-EDA* increases more the WIG spread than the EIG spread with respect to *OBS-EDA* (not shown). This result may be a consequence of the variance inflation as found in Žagar et al. (2011) for an ensemble Kalman filter system DART/CAM.

Figure 8 also shows that in all experiments the IG spread is dominant over the balanced spread for the zonal wavenumbers greater than 10. At the shortest analyzed scales, the IG spread makes about 70% of the total spread. Although differences in the EIG and WIG spread may seem small when expressed in percentages, they illustrate a sensitive balance affected either dynamically (larger growth of forecast errors in tropical easterly modes) or artificially (variance inflation) with potentially major impacts on subsequent forecasts.

3.4 *Using EDAs to define background error covariance matrices for assimilation*

Background error statistics, the static B covariance matrix, computed from the *OBS*-, *ST*- and *XB-EDAs* are provided to the assimilation system and three $T_L399L91T_L255$ resolution analyses are computed for the period 20081005-20081115. Because at the time these experiments were performed, the operational configuration was still using the static B matrix, the use of a 'flow dependent' B matrix was not possible. Description of the computation of the static B matrix from an ensemble analysis can be found in Fisher (2003); see also Derber and Bouttier, 1999 for covariance modeling.The background error covariance matrix is modelled using coordinate transformations and spherical wavelet techniques (Fisher 2003). In addition, a non-linear, analytical balance is included in the covariance model (Fisher 2003). The OBS, ST and XB analyses use, respectively, *OBS*-, *ST*- and *XB-EDAs* estimated B matrices. Diagnostics have been performed to assess the background error covariance impact on the assimilation system. The analysis experiments have the same name of the EDA experiments but bold fonts are used instead.

The first diagnostic presented is based on the observation influence (OI) (Cardinali *et al.*, 2004, Cardinali 2013) which quantifies the observational leverage in the analysis. The mean Observation Influence is the degree of freedom for signal, DFS, or total observation influence (Tukey, 1972; Velleman and Welsch, 1981; Wahba et al., 1995; Purser and Huang, 1993) divided by the total number of observation N

$$OI = \frac{DFS}{N} = \frac{tr(\mathbf{HK})^T}{N} \tag{3.3}$$

H is the linear observation operator and K is the gain matrix. OI and DFS depend on the assigned accuracy of the observations and background as well as the model itself which is a space and time propagator. The DFS quantifies the number of statistically independent directions constrained by each observation. Differences in the OI or DFS in the three assimilation experiments reflect differences on the B-matrices. The OI is proved to be bounded between 0 and 1; 0 influence indicates that an observation has had not influence on the estimate but only the background counted whilst OI=1 means that an entire degree of freedom has been devoted to fit that observation point. The OI can be gathered e.g. by observation type; in Fig. 9 the OI in OBS, ST and XB analyses is shown for different satellite and conventional observation types. Results indicate that XB shows the largest OI.

In particular, the largest OI increase is noticed for wind reporting observations (0.3 OBS , 0.5 ST and 0.7 XB), GPS-RO (0.2 OBS, 0.3 ST and 0.7 XB), AMSU-B radiances (0.2 OBS, 0.3 ST and 0.4 XB) and All-Sky SSMI radiances (0.1 OBS, 0.2 ST and 0.3 XB). The OI diagnostic indicates that when model errors are under-represented in the ensemble analysis, the background error statistics are also under-estimated and the observations have smaller leverage in the assimilation procedure. XB analysis provides better observations fit (not shown) in agreement with the higher OI. XB DFS is 50% larger than OBS and 30% larger than ST DFS.

A measure of the consistency of the assimilation system is provided by the diagnostics on the background-error statistics computed in observation space (Talagrand 2002; Desroziers *et al* 2005). If the K gain matrix is consistent with the 'true' covariances for background and observation errors, the innovation d and the analysis errors should be de-correlated from a statistical point of view. It can be simply shown (Desroziers *et al* 2005) that the covariance between the analysis increment in observation space (Hx_a - Hx_b) and the innovation vector (d), quantities archived during the asimilation procedure, should satisfy

$$\mathbf{HBH}^T \approx E[(\mathbf{Hx}_a - \mathbf{Hx}_b)\mathbf{d}^T] \quad (3.4)$$

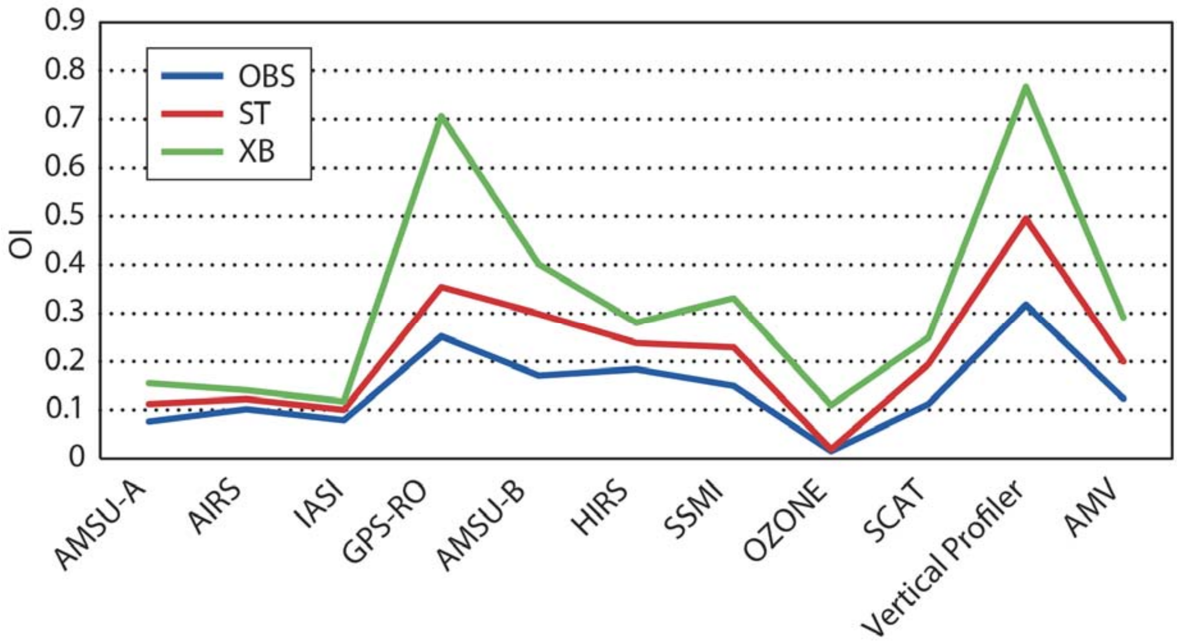

Figure 9 Global average observation influence (OI) for the different observation types assimilated in the XB (green line), ST (red line), OBS- (blue line) 4DVar analyses. AMSU-A and -B are microwave radiances, AIRS and IASI and HIRS infrared radiances, SSMI microwave imager radiance, GPS-RO satellite GPS radio occultation, OZONE retrieval, SCAT retrieved wind information from microwave scatterometer, Atmospheric Motion Vector (AMV) from geostationary cloud imagery and Vertical profiler consists of wind from radiosonde, pilot, aircraft and wind profiler observations.

The assigned background error variances, HBH^T (in observation space), are also archived; therefore the difference between the assigned and estimated background variances can be computed and averaged over the period of interest. In the context of linear estimation theory, a consistent unbiased analysis should result in no difference between the estimated and assigned background error variance. The following *Variance Consistency Check* (VCC)

$$VarianceConsistencyCheck = \frac{(\mathbf{HBH}^T)_{estimated} - (\mathbf{HBH}^T)_{assigned}}{(\mathbf{HBH}^T)_{estimated}} \tag{3.5}$$

measures the difference between the background error variances estimated from the analysis residuals (Desroziers *et al* 2005) and the background error variances assigned from the ensemble analysis normalized with respect to the estimated ones.

The *VCC* computed for the period 20081005-20081115 for OBS, ST and XB analyses shows small but non-zero values. . Figure 10 shows the *VCC* for AMSU-A and -B, HIRS, SSMI, SCAT, Vertical profilers and AMV: XB *VCC* is smaller than OBS and smaller or similar than ST.

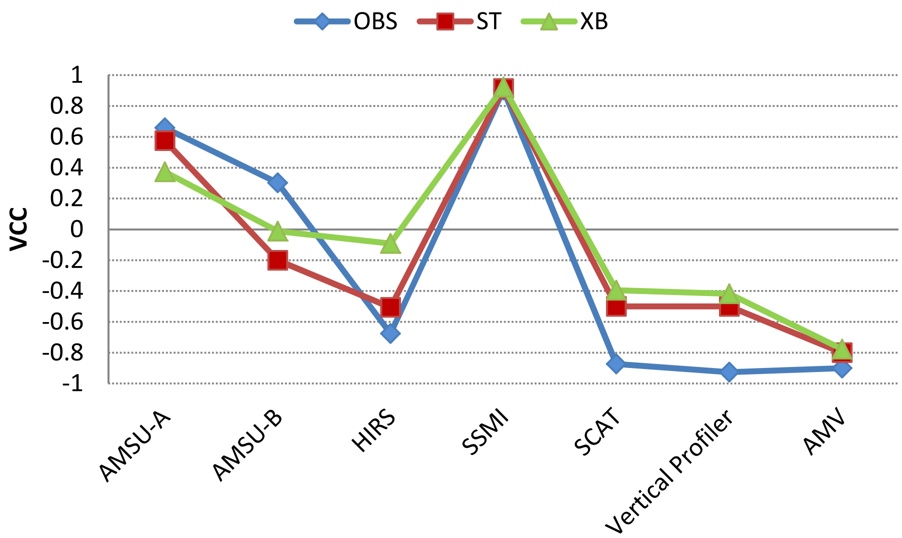

Figure 10. VCC for AMSU-A and -B, HIRS, SSMI, SCAT, Vertical profilers and AMV from XB (green line), ST (red line), OBS- (blue line) analyses.

3.5 *EDA-only based Ensemble Prediction System (EPS) forecasts*

Buizza *et al* (2008) proposed to use EDA-based perturbations in the ECMWF operational Ensemble Prediction System (EPS), and in June 2010 EDA-based perturbations have been introduced in the EPS to improve the simulation of initial uncertainties (Isaksen et al 2010). The replacement in the EPS of the evolved singular vectors with EDA-based perturbations improved substantially the EPS spread over the tropics, with a detectable impact in the early forecast range also over the extra-tropics. A positive impact was also detected on the EPS skill.

The three EDAs discussed in this work can be used to assess the sensitivity of EPS forecasts to the EDA configurations. Two types of ensembles were run: the first type included EDA-only perturbations, and the second type also included singular vectors (this is the configuration of the operational EPS). Since the impact of the EDA is more evident in the EDA-only type, attention will be focused mainly on them. EDA-only EPS have been run with the perturbations defined using the *OBS-*, *ST* and *XB-EDAs*. All three EPS configurations included 50 perturbed and 1 unperturbed members, with variable resolution T_L399L62 between forecast day 0 and 10, and T_L255L62 between forecast day 10 and 15 (in uncoupled mode). All forecast have been run with both the SPPT and the SPBS stochastic scheme as in the operational EPS (in other words, even the *OBS-* EPS that starts from the *OBS-EDA*-based perturbations that did not use any stochastic model, included stochastic perturbations in each of the 50-perturbed members). Forecasts have been run for 18 cases, with initial conditions from the 12th of October to the 14th of November 2008 every other day (with 12 UTC as initial time).

In all ensemble configurations, following the methodology used in the ECMWF EPS operational at the time when the experiments were conducted (for more details, please see Isaksen *et al* (2010) and references therein), the EDA-based component of the 50 EPS initial perturbations have been constructed by (a) defining 10 EDA-based perturbations by computing the difference between each of the 10 EDA perturbed members and the unperturbed (control) member and by (b) adding and subtracting these EDA-based perturbations from the un-perturbed analysis, defined by the ECMWF operational high-resolution 4D-Var system. Since this procedure provides only 20 perturbations, EPS members 21-40 have the same initial EDA-based perturbations as members 1-20, and EPS members 41-50 have the same as members 1-10. The fact that up to 3 EPS members can use the same EDA-based perturbations is not a problem in the ensembles run with initial perturbations generated using both EDA-based perturbations and singular vectors (as it is the case of the operational EPS), since 25 different SV-based initial perturbations are also used to generate the 50 positive and negative SV-based perturbations. For the EDA-only ensembles, the EPS members starting with the same EDA-based perturbation diverge, albeit in a slower way than the ensembles initialized by blending EDA- and SV-based perturbations, since each EPS member is integrated with different model error perturbations generated by the stochastic physic schemes.

The performance of an ensemble prediction system is usually measured by a range of metrics that compare, in a statistical sense, the forecast probability distribution function with the verification (either the analysis, or observations). Skill metrics that are routinely used include the ranked probability score and skill score, the Brier score (Brier 1950), the area under a relative operating characteristic and the ignorance skill score (Roulston and Smith 2002). The reader is referred to Wilks (1995) for a general overview, and e.g. to Palmer et al (2007) for a review of metrics used to assess the skill of the ECMWF Ensemble Prediction System.

In this work, attention has been focused on three aspects: firstly the ensemble reliability, i.e. the consistency between forecast probabilities and observed frequencies of occurrence, measured by the agreement between ensemble spread and ensemble-mean error, secondly the error of the ensemble-mean forecast, and thirdly the skill of probabilistic forecasts measured by the continuous rank probability skill score (CRPSS) and the ignorance score. Considering the first aspect, in a reliable ensemble, on average the spread of the system measured by the standard deviation should be equal to the average error of the ensemble-mean. This property follows from the fact that in such a system, one ensemble-member can be considered as the verification (Buizza et al 2005, Palmer et al 2006). Figure 11 shows the EPS spread (measured by the standard deviation) and the root-mean-square-error of the ensemble-mean forecast for the zonal wind at 850-hPa (verified against the operational high resolution analysis), computed over the Northern Hemisphere (20°-80°N, Fig 10a), over the Southern Hemisphere (20°-80°S, Fig 10b) and over the Tropics (20°S-20°N, Fig 10c).

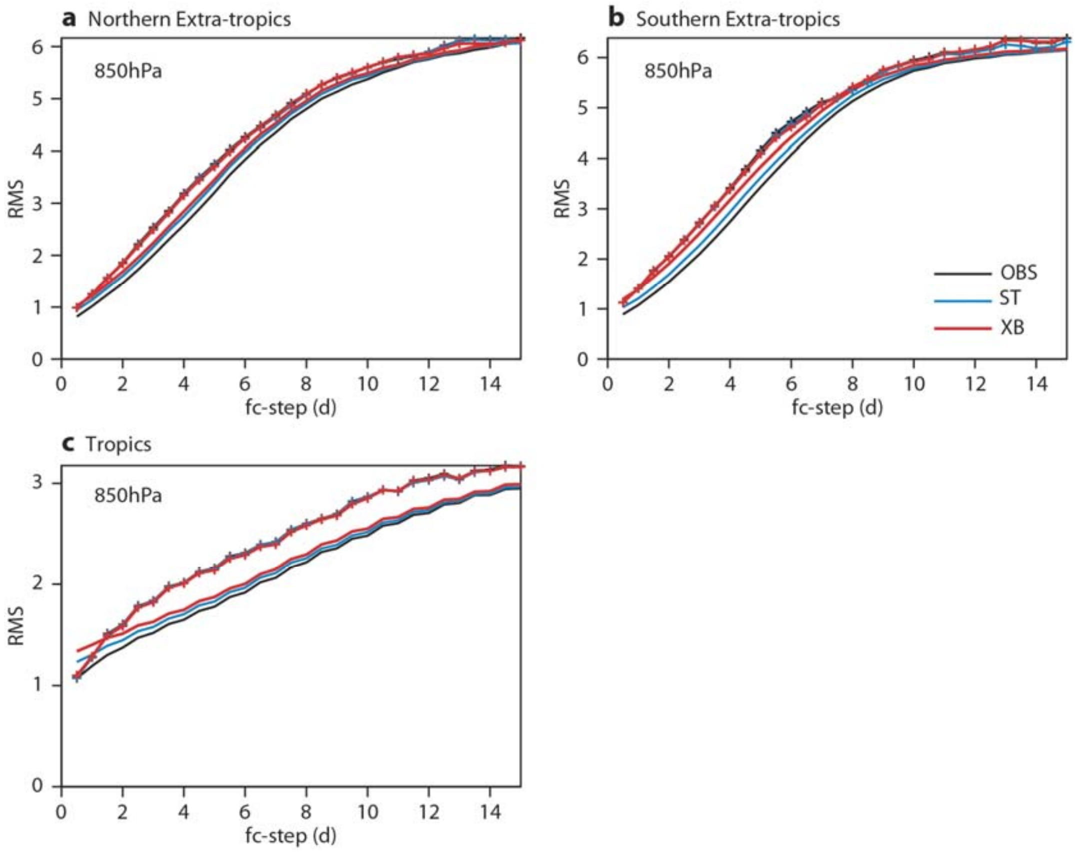

Figure 11 Averaged spread measured by the standard deviation (solid lines) and error of the ensemble mean (lines with symbols) of the EPS run with EDA-only perturbations generated from OBS-EDA (black), ST-EDA (blue) and XB-EDA (red). Results refer to the zonal wind component at 850 hPa over the a) Northern Hemisphere (20°-80°N), b) Southern Hemisphere (20°-80°S) and c) Tropics(20°S-20°N). The average has been computed considering 18 cases, each with 51-Members EPS forecast with initial condition from 12 to 14 November 2008, every other day (12 UTC only).

Figure 11 shows that for all configurations the ensembles are underdispersive, especially over the -tropics, indicating that EDA-only perturbations, if used as generated by the EDA and not re-scaled, are not sufficient to produce reliable ensemble forecasts. Among the EDA configurations, the *XB*- EPS has the largest spread, with differences evident up to about forecast day-10. Considering the error of the ensemble-mean, Fig. 11 shows that the ensemble-mean forecasts are very similar, almost undistinguishable for most of the forecast times, with the *XB*-EPS showing the smallest error for the forecast times when the spread is closer to the ensemble-mean root-mean-square error level (e.g. between forecast day 5 and 8 over the SH and between forecast day 5 and 10 over the tropics).

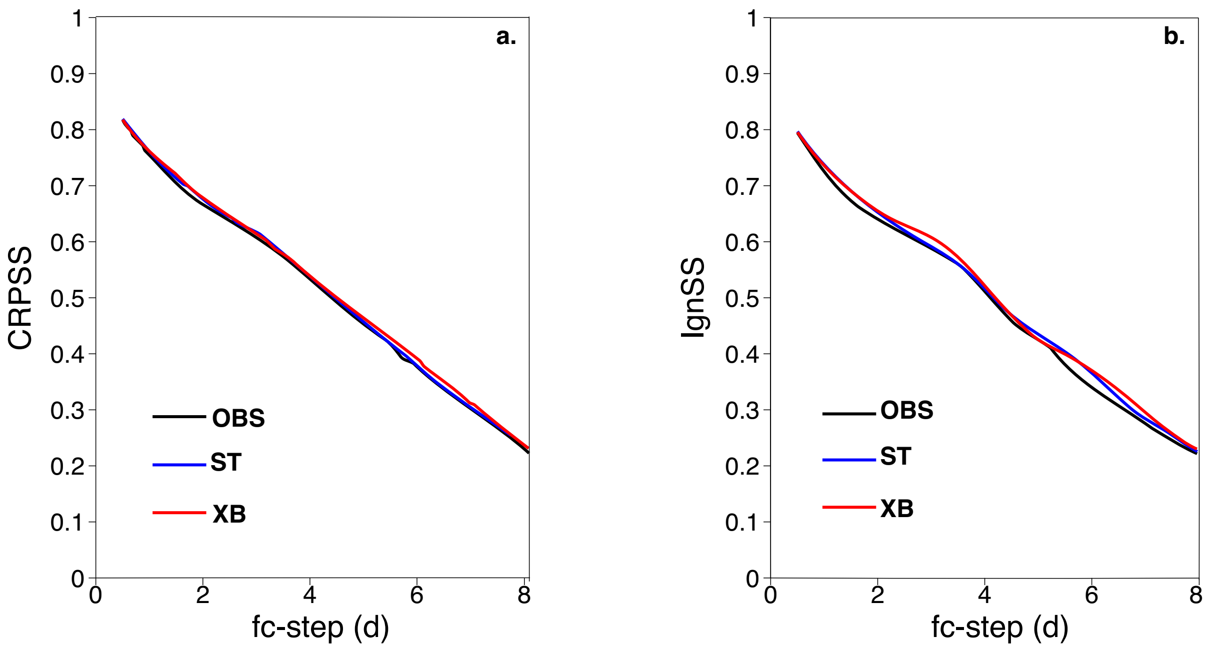

Figure 12 Average CRPSS (a) and Ignorance Score (b) of EPS run with EDA-only perturbations generated from OBS-EDA (black), ST-EDA (blue) and XB-EDA (red) for temperature at 850 hPa over the Southern Hemisphere. The average has been computed considering18 51-Members EPS forecast with initial condition from 12 to 14 November 2008, every other day (12 UTC only).

Considering the skill of probabilistic forecasts, all the metrics mentioned above have been considered, and since results are all consistent, only CRPSS and ignorance skill scores will be shown. The CRPS is the equivalent of the mean squared error for single forecasts, and give a measure of the average distance between the forecast and observed distributions; the corresponding skill score, the CRPSS have been computed using a climatological probabilistic forecast as reference (thus a perfect probabilistic forecast would score 1, and a forecast as skillful as climatology 0). The ignorance skill score is a logarithmic score defined using information theory (Roulston and Smith 2002), based on the information deficit (or ignorance) in the forecast. According to Benedetti (2010), for probabilistic forecast systems the ignorance score and skill scores are more fundamental scores than the Brier score and skill scores, given that these latter are second-order approximations of the former. A clear advantage of these two scores compared to the Brier score and skill score, or the area under a relative operating characteristic, is that both the CRPSS and the ignorance skill score consider the whole forecast probability distribution function of forecast states, and not simply some specific event (e.g. the probability of a variable exceeding a certain threshold). Thus, they provide a more complete assessment than these latter two (Wilks, 1999).

Figure 12a shows that in terms of CRPSS, the three EPSs have a similar performance. The same conclusions can be drawn from the ignorance skill score (Fig. 12b). However, it is worth to notice that the small but consistent improvement at day-2 and 3 and day-6 as shown in Fig 12 a) and b) is observed for all parameters at different levels. Similar conclusions could be drawn by considering different performance metrics for these variables e.g. the Brier skill score or the area under the relative operating characteristic (ROC) for the probabilistic prediction of dichotomous events such as 'wind anomalies above or below the climatological standard deviation' (not shown).

Concluding, these results based on EDA-only ensembles indicate that the use of the *XB-EDA* in the EPS would lead to a slightly better match between spread and error, i.e. a better reliability, and to similar forecast skill than the use of the *ST*- and *OBS-EDA*s The EPS experiments based on EDA-based perturbations and singular vector perturbation indicate smaller spread differences, detectable only up to forecast day 5 instead of day 10, and practically no differences in skill (not shown).

4 Conclusions

In this paper, the representation of model error in the ECMWF 4DVar Ensemble Data Assimilation (EDA) system by additively perturbing the background field is shown. The method follows the idea used in the Meteorological service of Canada (MSC) Ensemble Kalman Filter (Houtekamer et al 2005) and in the ARPEGE 4D-Var system (Raynaud et al 2012) to account for model uncertainties in the EDA by perturbing the model background field. In the MSC Ensemble Kalman Filter the perturbation magnitude is globally constant (but the representation of other errors is also considered) whilst at Météo-France the perturbation, as a function of latitude band and model level, is multiplied to the background field (multiplicative approach). Here, the additive approach is presented.

The idea behind it is that a large fraction of model error is represented by the short-range forecast error. Thus perturbing the 12 hour model forecast would include error sources introduced by dynamics, spatial and temporal discretization and, in the case of assimilation, linearized physical processes and the misspecification of the probability distribution of errors in the observations and the numerical background model. The proposed methodology, contrary to the stochastically perturbed parameterization tendency scheme, does not require routine diagnostic and tuning.

Ensembles of data assimilation with different representations of model error have been compared. In particular, two more EDAs are examined, all with the same methodology to represent the observation uncertainties but different techniques to account for model uncertainties. One model error representation technique is based on the assumption that random model errors due to the physical process parameterizations are the main model error source (*ST-EDA*). In the *ST-EDA*, stochastic perturbations are added to the physical model tendency at each model time-step. The other ensemble considered (*OBS-EDA*) only includes an observation error representation, which implicitly modifies the background fields in the assimilation cycling process. In the additive *XB*-method, the magnitude of the perturbation is calculated by comparing the variance of the innovation vector of the high resolution analysis system with the ensemble data assimilation variance in which the ensemble data assimilation is produced by only perturbing the observations. The perturbation is a function of latitude band, vertical model level and the model parameter.

Results have shown that the EDA generated using *XB*-method accounts for the largest spread, followed by the *ST-EDA* with stochastically perturbed parameterization tendencies and the *OBS-EDA* with observation only perturbations. The increase of spread depends on the location: the stratosphere accounts for the largest increase of spread. In the tropics, the increase occurs mainly in the midtroposphere, while in the extra-tropics spread enhancement is detected at all model levels and in the Southern hemisphere in particular. The largest difference compared to other EDA configurations is found in the extra-tropics where all but the *XB-EDA* are not able to produce significant spread. However, the spread structure is quite similar among the different methods confirming that a large part of the model error sources are affecting the short-range forecast. Comparison with respect to the un-cycled ensemble analysis (*OBS-OBS EDA*) shows that the background perturbation adds spread from high to low levels in the troposphere and into the extra-tropics in both Northern and Southern hemispheres. Normal-mode diagnostics compare the percentage of balanced and unbalanced spread in the four ensembles *XB-* , *ST- OBS-OBS* and *OBS-EDA*. It was found that various experiments contain around 40% of their spread in the balanced modes with *ST-EDA* being the least balanced. Both *ST-* and *XB-EDA* add both balanced and IG spread homogeneously at all scales when compared with *OBS-EDA*. An interesting difference between *XB-* and *ST-EDA* is found in the distribution of the added easterly IG spread relative to the westerly IG spread. *XB-EDA* adds more EIG spread on all scales whereas the *ST-EDA* increases the WIG more than the EIG spread when compared with *OBS-EDA* spread. Overall, *XB-EDA* is the only experiment which contains more spread in the easterly propagating IG modes than in the westerly IG component. It is believed that this is due to the tropical flow properties which are strongly influenced by the easterly-propagating Kelvin waves. Although more investigation is needed to confirm this interpretation, the increase of spread in the westerly IG mode for the *ST* is likely to be related to the inflation of the model background error variances (as was found by Žagar et al. (2011) for the DART-CAM ensemble). In fact, it should be kept in mind that *OBS-* and *ST-EDAs* include a global inflation factor of the background error variances that penalizes the model background further with respect to the observations. The inflation is a spatial and temporal constant that does not depend on synoptic weather developments but only intends to introduce larger ensemble spread.

The covariance matrices produced by the three ensemble analysis have been provided to a higher-resolution data-assimilation system. The diagnostic performed on the three resulting analyses show that the largest Observation Influence (OI) is obtained from the assimilation system with *XB-EDA* background statistics. This is due to the larger background error variances. When the analysis residual diagnostic is applied to investigate the consistency of the assimilation systems considered, systematical smaller differences are found with XB showing closer agreement between the assigned and estimated background error variances.

Finally, the three EDAs have been used in ensemble prediction mode. Since June 2010, EDA-based perturbations have been used with singular vectors to simulate initial uncertainties in the ECMWF Ensemble Prediction System (Buizza et al 2008, Isaksen et al 2010). Results have indicated that the use of the *XB-EDA* in the EPS would lead to the largest spread, with differences evident up to about forecast day-10 and with the smallest error for forecast times from day-5 to 8 over Southern and from day-5 to 10 over the tropics when the spread is closer to the ensemble-mean root-mean-square error. In terms of EPS forecast skill, very small but consistent improvements up to day-7 have been detected when, in particular, the ignorance score is used.

In conclusion, the *XB*-method discussed in this work is shown to be a valuable alternative of the method used in the current ECMWF EDA to simulate model uncertainty in the ensemble analysis. It accounts for different sources of error coming from the dynamics, the parameterizations, the linearization and interpolation schemes and it is easier to tune and maintain. The tuning of the perturbation is performed automatically every 3 days from a three-day sample of the high resolution operational innovations. A possible extension of the work presented would be to combine the background perturbation method with the stochastic physical tendency perturbation method or other methodologies that are considered to simulate longer range random model error sources. The estimation of the background perturbations magnitude should, in this case, be achieved by comparing the innovations of the high resolution analysis system with the EDA in which not only the observations are perturbed but also e.g. stochastic perturbations are added to the physical model tendency, that is, the *ST-EDA*.

Acknowledgements

The authors thank Elias Holm who provided the B matrices computed from the ensemble of background states from the different EDAs considered in this study. The authors are very thankful to Gerald Desroziers and Loïk Berre for fundamental discussions on the methodology applied here. The authors are also very grateful to the thorough and constructive review provided by two reviewers, Antony Weaver and an anonymous one, which not only has improved the manuscript but has enriched our understanding of ensemble analysis methodologies.

References

Bormann, N., S. Saarinen, G. Kelly and J-N. Thépaut, 2003: The spatial structure of observation errors in atmospheric motion vectors from geostationary satellite data. *Mon. Wea. Rev.*, **131,** 706-718.

Berner J., G. Shutts, M. Leutbecher, and T.N. Palmer, 2008: A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system, J. Atmos. Sci., **66**, 603-626.

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. *ECMWF Workshop Proceedings on Flow Dependent Aspect of Data Assimilation*, 11-13 June 2007, 151-168.

Buehner, M., P.L. Houtekamer, C. Charette and H.L. Mitchell, 2010: Intercomparison of variational data assimilation and ensemble kalman filter for global deterministic NWP. Part II: One-Month experiments with real observations. *Mon. Wea. Rev.,* **138,** 1567-1586.

Benedetti, R, 2010: Scoring rules for forecast verification. *Mon. Wea. Rev.*, **138**, 203-211.

Buizza R, M. Miller and T.N. Palmer, 1999. Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Q. J. R. Meteorol. Soc.*, **125**: 2887-2908.

Buizza, R., J.R. Bidlot, N. Wedi, M. Fuentes, M. Hamrud, G Holt and F. Vitart, 2007: The new ECMWF VAREPS (Variable Resolution Ensemble Prediction System). *Q. J. R. Meteorol. Soc.* **133,** 681-695.

Buizza, R., Houtekamer, P. L., Toth, Z., Pellerin, G., Wei, M., and Zhu, Y., 2005: A comparison of the ECMWF, MSC and NCEP Global Ensemble Prediction Systems. *Mon. Wea. Rev.,* **133**, 5, 1076-1097.

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.

Cardinali, C., S. Pezzulli and E. Andersson, 2004: Influence matrix diagnostics of a data assimilation system. *Q. J. R. Meteorol. Soc.*, **130,** 2767-2786

Cardinali C. 2013. Observation influence diagnostic of a data assimilation system. *Data Assimilation for Atmospheric, Oceanic and Hydrology Applications*, II, Chapter 4. Springer-Verlag: Berlin and Heidelberg, Germany.

Derber, J. and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195-222

Desroziers, G. and S. Ivanov, 2001. Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. *Q. J. R. Meteorol. Soc.*, **127,** 1443-1452

Desroziers G, L. Berre and B. Chapnik and P. Poli, 2005: Diagnosis of observation, background and analysis error statistics in observation space. *Q. J. R. Meteorol. Soc.*, **131,** 3385-3396

Fisher, M., 2003: Background error covariance modeling. *Proceedings of the ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, ECMWF, pages 45–63 (available online at: http://www.ecmwf.int/publications/).

FujitaT., D.J. Stensrud and D.C. Dowell, 2007: Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physiss uncertainties. *Mon. Wea. Rev.*, **135,** 1846-1868.

Gauthier, P., M. Buehner and L. Fillion, 1999: Background-error statistics modelling in a 3D variational data assimilation scheme: Estimation and impact on the analyses. *Proceedings of ECMWF workshop on Diagnosis of Data Assimilation Systems*, 131-145. (available online at: http://www.ecmwf.int/publications/).

Hamill, T.M. and J.S. Whitaker, 2005. Accounting for the error due to unresolved scales in ensemble data assimilation: a comparison of different approaches. *Mon. Wea. Rev.*, **133,** 3132-3147.

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

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

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

Isaksen, L., J. Haseler, R. Buizza and M. Leutbecher, 2010: The new esemble of data assimilation. ECMWF Newsletter No. 123, p 17-21.

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

Lopez, P. and E. Moreau, 2005: A convection scheme for data assimilation: Description and initial tests. *Q.J.R.Meteorol.Soc.*, **131**, 409-436

Janiskova, M., J.-F. Mahfouf, J.-J. Morcrette and F. Chevallier, 2002: Linearized radiation and cloud schemes in the ECMWF model: Development and evaluation, *Q. J. R. Meteorol. Soc.*, **128**, 1403-1423.

Meng, Z. and F. Zhang 2007: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect model experiments. *Mon. Wea. Rev.*, **135,** 604-620.

Molteni, F., R. Buizza, T.N. Palmer and T. Petroliagis, 1996: The new ECMWF Ensemble Prediction System: Methodology and validation. *Q. J. R. Meteorol. Soc.*, **122,** 73-119.

Palmer, T., R. Buizza, R. Hagedorn, A. Lawrence, M. Leutbecher, and L. Smith, 2006: Ensemble prediction: a pedagogical perspective. ECMWF Newsletter No.106, p 10-17.

T.N. Palmer, R. Buizza, M. Leutbecher, R. Hagedorn, T. Jung, M. Rodwell, F. Vitart, J. Berner, E. Hagel, A. Lawrence, F. Pappenberger, Y-Y. Park, L von Bremen, I Gilmour, 2007: The ECMWF Ensemble Prediction System: recent and on-going developments. A paper presented at the 36th Session of the ECMWF Scientific Advisory Committee. ECMWF Research Department Technical Memorandum No. 540, 53 pp. (available online at: http://www.ecmwf.int/publications/)

Palmer, T.N., R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G.J. Shutts, M. Steinheimer and A Weisheimer, 2009: Stochastic parametrization and model uncertainty. ECMWF Research Department Technical Memorandum No. 598, 42 pp.
(available online at: http://www.ecmwf.int/publications/)

Purser, R. J. and H.-L. Huang, 1993: Estimating effective data density in a satellite retrieval or an objective analysis. *J. Appl. Meteorol.*, **32,** 1092—1107

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

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.

Raynaud, L., Berre, L. and Desroziers, G., 2012: Accounting for model error in the Météo-France ensemble data assimilation system. *Q.J.R. Meteorol. Soc.*, **138**, 249–262.

Roulston, M S, and L A. Smith, 2002: Evaluating probabilistic forecasts using information theory. *Mon. Wea. Rev*., **130**, 1653-1660.

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

Talagrand, O., 1997: Assimilation of observations, an Introduction. *J. Meteorol. Soc. Japan*, **75,** N.1B,191-209.

Talagrand, O., 2003: A posteriori validation of assimilation algorithms. Chap. 8 In *Data Assimilation for the Earth System*, edited by Richard Swinbank, Victor Shutyaev and WilliamAlbert Lahoz. Nato Science Series, 85-95: Springer Netherlands.

Tompkins, A.M. and M. Janiskova: 2004, A cloud scheme for data assimilation: Description and initial tests. *Q.J.R.Meteorol.Soc*, **130**, 2495-2517

Tukey, J. W. 1972: Data analysis, Computational and Mathematics. *Q. of Applied mathematics*, **30**, 51-65

Velleman, P. F., and R. E. Welsch, 1981: Efficient computing of regression diagnostics. *The American Statisticians,* **35**, 234-242.

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

Wahba, G., D.R. Johnson, F. Gao and J. Gong, 1995: Adaptive tuning of numerical weather prediction models: Randomized GCV in three- and four-dimensional data assimilation. *Mon. Wea. Rev*., **123,** 3358-3369.

Wilks, D.S. 1995: *Statistical methods in the atmospheric sciences. An introduction*. Geophysics Series, Vol. **59**.

Žagar, N. J. Tribbia, J. Anderson and K. Raeder, 2011: Balance of the background-error variances in the ensemble assimilation system DART/CAM. *Mon. Wea. Rev*., **139**, 2061-2079.

Žagar, N., D. Tan, L. Isaksen and J. Tribbia, 2013: Balance properties of the short-range forecast errors in the ECMWF 4D-Var ensemble. *Q. J. R. Meteorol. Soc*. **139,** 1229-1238.