Potential Use of an Ensemble of Analyses in the ECMWF Ensemble Prediction System

Roberto Buizza, Martin Leutbecher and Lars Isaksen

European Centre for Medium-Range Weather Forecasts
Reading UK
www.ecmwf.int

Abstract

One of the crucial aspects of the design of an ensemble prediction system is the definition of the ensemble of initial states. This work investigates the use of singular vectors, an ensemble of analyses, and a combination of the two types of perturbations in the ECMWF operational ensemble prediction system. First, the similarity between perturbations generated using initial-time singular vectors (SVs) and analyses from an ensemble data assimilation (EDA) system is assessed. Results show that the EDA perturbations are less localized geographically and have a better coverage of the tropics. EDA perturbations have also smaller scales than SV-based perturbations, and have a less evident vertical tilt with height, which explains why they grow less with the forecast time. Then, the use of EDA-based perturbations in the ECMWF ensemble prediction system is studied. Results indicate that if used alone, EDA-based perturbations lead to an under-dispersive and less skilful ensemble then one based on initial-time SVs only. Combining the EDA and the initial-time SVs gives a system with a better agreement between ensemble spread and the error of the ensemble-mean, a smaller ensemble-mean error and more skilful probabilistic forecasts than in the current operational system based on initial-time and evolved SVs.

1. Introduction

One of the crucial aspects of the design of an ensemble prediction system is the definition of the ensemble of initial states. At the European Centre for Medium-Range Weather Forecasts (ECMWF, *Molteni et al*. 1996, *Buizza et al.* 2007), the initial time probability density function is represented by initial conditions generated by adding to and subtracting from the unperturbed analysis 25 perturbations defined by leading singular vectors (SVs) of the model tangent forward and adjoint version (*Buizza and Palmer* 1995), with the unperturbed analysis defined by the ECMWF 4-dimensional variational data assimilation system (*Rabier et al.* 2000). The 50 initial perturbations generated using the leading singular vectors, i.e. the phase-space directions of maximum growth measured by total energy, are scaled to have amplitude comparable to analysis error estimates and are designed to represent analysis error components along these directions of maximum growth. *Ehrendorfer and Tribbia* (1997) have shown that if the objective of an ensemble system is the optimal prediction of the forecast error covariance matrix (optimal in the sense of maximum possible fraction of forecast error variance), then the singular vectors constructed using covariance information at the initial time constitute the most efficient means for predicting the forecast error covariance matrix.

At the National Center for Environmental Prediction, initial perturbations are defined using an Ensemble Transform method (*Wei et al* 2008), and at the Chinese Meteorological Administration and the Korean Meteorological Administration initial perturbations are defined using bred-vectors instead (*Toth and Kalnay* 1993, 1997). At the Meteorological Service of Canada (MSC), a system simulation approach (*Houtekamer et al.* 1996) is used to generate the ensemble of initial perturbations, where a number of parallel data assimilation cycles are run randomly perturbing the observations and using different parameterization schemes for some physical processes in each run. The ensemble of initial states generated by the different data assimilation cycles defines the initial conditions of the Canadian ensemble system.

Several authors have compared the strengths and weaknesses of using a selective or a non-selective approach in ensemble prediction to simulate initial uncertainties. These studies can be classified in three groups depending on the experimental environment that has been followed. The first group of *simple-but-same-*

model environment studies uses simple, low-dimensional systems, often under 'perfect' model assumptions. These studies have the advantage of comparing ensembles obtained using different methods but the same forecast model, but suffer from the fact that these model are very simple, and under-represent the complexity of the atmospheric system. Furthermore, these studies often use synthetic data (i.e. data extracted from a model run rather than generated using real observations) as observation and for verification, a fact that limits the validity of their conclusions for real time applications. The second group of *complex-but-different-model environment* studies compares operational ensemble systems, using for the verification analyses from state-of-the-art data-assimilation systems fed with real observations. Although these studies give a realistic picture of the status of ensemble prediction, and of the relative strengths and weaknesses of the different systems, they cannot be used to draw any conclusion on the effectiveness of the different methods used to simulate the initial uncertainties, since they compare ensembles not only based on different approaches but also run with different models. The third group of *complex-and-same-model environment* studies tries to combine the merits of the two groups mentioned above, and test different perturbation strategies using only one state-of-the-art model, and for the verification an analysis obtained assimilating real observations with a state-of-the-art data assimilation system.

Among the first *simple-but-same-model environment* group, *Anderson* (1997) using the 3-parameter *Lorenz* (1963)'s system compared the use of dynamically constrained (SVs, bred vectors) or unconstrained perturbations, and concluded that unconstrained (random) perturbations produced more skilful systems. Using a more complex quasi-geostrophic model, *Houtekamer and Derome* (1995) found little differences between ensembles based on bred vectors, SVs or the Canadian observation perturbation method, while *Hamill et al* (2000) concluded that the observation perturbation method provides more accurate forecasts than the other two. More recently, *Bowler* (2006), using a low-order model developed by Lorenz (*Lorenz* 1996), concluded that using an ensemble Kalman filter approach gives a better ensemble than bred vectors or SVs. *Descamps and Talagrand* (2007), using the *Lorenz* (1996) model and the 3-level quasi-geostrophic model developed by *Marshall and Molteni* (1993) with synthetic data, compared the performance of ensembles based on SVs, bred vectors, Ensemble Kalman filter and Ensemble Transform Kalman filter, and concluded that the two latter outperform the two former ensembles.

Among the second *complex-but-different-model environment* group, *Buizza et al* (2005) compared the performance of the ECMWF, NCEP and Canadian ensemble systems for one season. They showed that the ECMWF ensemble system was performing better than the other two ensembles, but could not clearly estimate the relative importance of using the SV method instead of the bred vectors used at NCEP, or the Canadian perturbation observation approach. They concluded that "*the performance of an ensemble prediction system strongly depends on the quality of the data-assimilation system used to create the unperturbed (best) initial condition and the numerical model used to generate the forecasts"*. The reader is referred to *Park et al* (2008) for a similar, very recent work that compares the performance of operational ensemble systems sharing data within the TIGGE (the THORPEX interactive Grand Global Ensemble) project.

Magnusson et al (2008) and *Wei et al* (2008) are examples of the third *complex-and-same-model environment* group. *Wei et al* (2008) compare the performance of the NCEP ensemble system using bred vectors, ensemble transform and ensemble transform with re-scaling methods. Their work concluded that the ensemble transformed method with re-scaling outperformed the other two methods: this method replaced the bred vectors in the NCEP operational ensemble system on 30 May 2006. *Magnusson et al* (2008) compare the performance of three ensembles run with the ECMWF ensemble system at $T_L255L40$ resolution with initial perturbations defined using SVs and two types of bred vectors. They conclude that over the extra-tropics the SV method gives a slightly better performance, but over the tropics the breeding method performs better, due to the fact that the SVs sample only a limited region of the tropical band.

The work presented here is to be classified within this group of *complex-and-same-model environment* studies. It aims to investigate whether the performance of the ECMWF ensemble prediction system (EPS) can be improved by using alternative methods to generate the set of perturbed initial conditions. More specifically, it discusses the use of a set of perturbations generated using an Ensemble Data Assimilation (EDA, *Isaksen et al* 2007) system in the ECMWF ensemble prediction system. Following *Houtekamer et al* (1996, 2005), the ensemble of analyses has been generated by randomly perturbing the observations in a manner consistent with observation error statistics, and by using stochastic physics to simulate the effect of model uncertainties.

The three key questions that this communication is addressing are the following:

- How similar/different are EDA-based and SV-based initial perturbations?
- What is the difference in skill in an ensemble using EDA-perturbations only and an ensemble using SV-perturbations only?
- Can the skill of the EPS be improved by combining EDA- and SV-based perturbations?

These questions are addressed considering experiments performed for a 45-day period using a $T_L399L91$ version of the ECMWF 4-dimensional variational data assimilation system, and a 10-day $T_L399L62$ version of the EPS. In section 2, the methodology and experimental set-up are presented. In section 3, the characteristics of EDA-based initial perturbations are compared with the characteristics of SV-based perturbations. In section 4, the potential use of the ensemble of analyses to generate the EPS initial perturbations is discussed. Finally, some conclusions are drawn, and future work is discussed in section 5.

2. Methodology, ensemble definition and experimental set-up

A data assimilation system produces an optimal estimate of the true state of the atmosphere, called the analysis, given as input a first-guess and a set of observations. Schematically,

$$a(d,t_a) = D[fg(d,t_0 \le t \le t_1); o(d,t_0 \le t \le t_1)] \tag{1}$$

where $a(d,t_a)$ denotes the analysis of day d at time t_a, $fg(d,t)$ the first guess, $o(d,t)$ the set of observations at time t, and $D[..]$ the 4D-Var assimilation process. In a 4-dimensional variational (4D-Var) assimilation system, the input spans a time window $t_0 \le t \le t_1$. For each forecast time t, the first guess is given by the time integration of the model equations from the previous analysis t_a

$$fg(d,t) = a(d,t_a) + \int_{t_a}^{t} [A(d,\tau) + P(d,\tau) + \lambda_j(d,\tau)] d\tau \tag{2}$$

where $P(d,\tau)$ represents the tendency due to the parameterized physical processes (e.g. radiation, moist processes, turbulence), $A(d,\tau)$ is the tendency due to all the other processes, $\lambda_j(d,\tau)$ denotes the model perturbation added by the stochastic kinetic energy backscatter scheme (*Berner et al* 2008), and $a(d,t_a)$ denotes the initial condition. Ensembles of 4D-Var analyses have been generated by randomly perturbing the observations [term $o(d,t)$ in Eq. (1)] in each single analysis:

$$\begin{aligned} a_j(d,t_a) &= D[fg_j(d,t_0 \le t \le t_1); o(d,t_0 \le t \le t_1) + \varepsilon_j(d,t_0 \le t \le t_1)] \\ fg_j(d,t) &= a_j(d,t_a) + \int_{t_a}^{t} [A_j(d,\tau) + P_j(d,\tau) + \lambda_j(d,\tau)] d\tau \end{aligned} \tag{3}$$

Ensembles of analyses with 10 perturbed members have been generated with a $T_L399L91$ version, i.e. with spectral triangular truncation T399 with linear grid, and 91 vertical levels of the ECMWF model (*Isaksen et*

al 2007). Experiments have been performed with a 12-hour assimilation window, from the 14th of September to the 31st of October 2006. For each observation apart for cloud-track winds, perturbations have been defined by randomly sampling a Gaussian distribution with zero mean and standard deviation defined by the observation error standard deviation. For cloud-track winds observations, perturbations are horizontally correlated (*Bormann et al* 2003). Sea-surface temperature fields are also perturbed, with correlated patterns as currently done in the ECMWF seasonal ensemble system (*Vialard et al* 2005). At the first assimilation cycle, the randomly-perturbed observations are the only source of divergence between the perturbed analyses, while for the subsequent cycles differences in the first-guess fields also contribute to the analyses spread.

Ensemble forecasts have been run every other day from the 22nd of September to the 30th of October, with forecasts generated with the same model (cycle 31r2, with $T_L399L62$ resolution), and with stochastic perturbations added to the physical tendencies to simulate model error uncertainties. Schematically, each ensemble forecast is defined by the time integration of the (stochastically perturbed) equations

$$\begin{aligned} e_j(d,t) &= e_j(d,t_a) + \int_{t_a}^{t} [A_j(d,\tau) + P_j(d,\tau) + \delta_j P_j(d,\tau)] d\tau \\ \delta_j &= r_j(\lambda,\phi) \end{aligned} \tag{4}$$

where δ_j is a random number that depends on the forecast time, the grid-point latitude λ and longitude φ (see *Buizza et al* 1999a for a detailed description of the scheme).

In the ensembles, initial uncertainties have been simulated using 48-hour total energy singular vectors, vectors that identify phase-space directions with maximum growth over a 48-hour optimisation time interval measured using a total energy norm. Since June 2004 (*Leutbecher* 2005), the operational day *d* ensemble has been using 'forecast' SVs, i.e. SVs computed between the +6h and the +54h trajectory, instead of between the initial and +48h trajectory. In the operational ensemble, for each initial date *d*, two sets of singular vectors are used:

- the initial singular vectors, defined as the SVs at initial-time growing for 48h in the future *SV(d,0)* (more precisely, these are the initial-time SVs growing between the +6h and the +54h trajectory started at day d-6h), and
- the evolved singular vectors, defined as the SVs at forecast time +48h growing from date *(d-48h)* for 48-hour in the future, *SV(d-48h,48h)* (more precisely, these are the SVs at +48h growing between the +6h and the +54h trajectory started at day *d-54h*)

Figure 1 is a schematic illustration of the configuration used in the ECMWF operational ensemble system at the time of writing (February 2008). The initial-time singular vectors were introduced at the time of implementation of the ECMWF EPS to simulate the effect of perturbations growing during the forecast time (*Buizza and Palmer* 1995, *Molteni et al* 1996). Then, in March 1998, the evolved singular vectors were added to improve the simulation of uncertainties growing during the data assimilation cycle (*Barkmeijer et al* 1999). Up to January 2002, the initial and evolved SVs covered only the extra-tropics (more precisely, their final-time total energy norm was maximized north of 30° N and south of 30° S). In January 2002, initial-time targeted tropical SVs were added to improve the prediction of tropical storms (*Barkmeijer et al* 2001, *Puri et al* 2001): these tropical SVs are computed only in some selected regions and do not cover the whole tropical band. The reader is referred to *Palmer et al* (2007) for a review of the development of the ECMWF ensemble prediction system from 1992 to 2007.

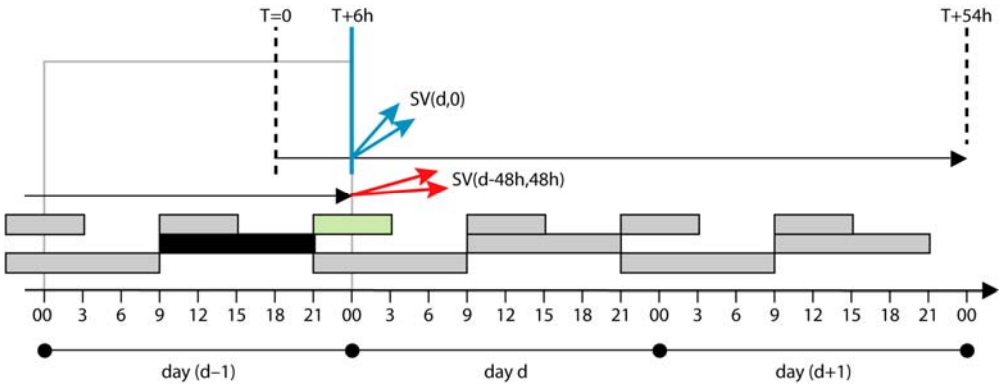

Figure 1: Schematic of the configuration used to generate the initial perturbations of the ECMWF ensemble at 00UTC of day d*. The 12h-long black and grey boxes between 9-21UTC and 21-9UTC mark the times during which the 12-hour 4D-Var runs; the 6h-long green and grey boxes between 21-03UTC and 9-15UTC mark the times during which the early-delivery 6h 4D-Var analysis runs. The unperturbed analysis at 00UTC of day* d *is defined by the 6h 4D-Var analysis generated by the early-delivery suite (green box). The initial perturbations are generated combining the evolved singular vectors* SV(d-48h,48h) *(red vectors) computed at 00UTC of day* (d-2)*, and initial-time singular vectors* SV(d,0) *(blue vectors) computed at 00UTC of day* d*. Note that (see text for more details) the trajectory along which the* SV(d,0) *grow is between the T+6h and the T+54h forecasts started at 18UTC of day* (d-1)*, and the trajectory along which* SV(d-48h,48h) *grow is between the T+6h and the T+54h forecasts started at 18UTC of day* (d-3)*.*

Each ensemble initial perturbation is defined by a combination of the initial-time and the evolved SVs computed over different regions: extra-tropical Northern Hemisphere (north of 30°N), extra-tropical Southern Hemisphere (south of 30°S), and between 1 and 6 selected regions targeted on tropical cyclones. Once computed they are combined and re-scaled to have initial amplitude comparable to the analyses error estimate given by the high-resolution data-assimilation system. The reader is referred to *Leutbecher and Palmer* (2008) and *Ehrendorfer and Beck* (2003) for more details on the Gaussian sampling method used to combine and re-scale the SVs in the current operational ensemble system.

Ensemble members have been integrated starting from perturbed initial condition $e_j(d,t_a)$, defined (a) using initial T42L62 singular vectors only, (b) as (a) plus evolved singular vectors (as it is currently done in the ECMWF operational ensemble system), (c) an ensemble of analyses only, or a (d) combination of singular vectors and perturbed analyses:

$$e_j(d,t_a) = e_0(d,t_a) + \sum_{s=1}^{S} \sum_{k=1}^{N_{SV,s}} \alpha_{j,k,s} SV_{k,s}(d,t_a) \tag{5.a}$$

$$e_j(d,t_a) = e_0(d,t_a) + \sum_{s=1}^{S} \sum_{k=1}^{N_{SV,s}} [\beta_{j,k,s} SV_{k,s}(d-48h,48h) + \alpha_{j,k,s} SV_{k,s}(d,t_a)] \tag{5.b}$$

$$e_j(d,t_a) = PA_j(d,t_a) \tag{5.c}$$

$$e_j(d,t_a) = PA_{m(j)}(d,t_a) + \sum_{s=1}^{S} \sum_{k=1}^{N_{SV,S}} [\alpha_{j,k,s} SV_{k,s}(d,t_a)] \tag{5.d}$$

where $s=1,S$ denotes the different areas for which SVs are computed.

In the experiments discussed in this work, the unperturbed analysis $e_0(d,t_a)$ has been defined as the T_L799L91 analysis interpolated to the T_L399L62 resolution. Furthermore, each T_L399L62 perturbed analysis $PA_m(d,t_a)$ has been generated by adding to the unperturbed analysis $e_0(d,t_a)$ the difference between the +6h forecast

started from the T_L399L62 j-th analysis valid for *(d-6h)*, a_j*(d-6h,6h)* and the +6h forecast from unperturbed analysis e_0*(d-6h,6h)*:

$$PA_m(d,t_a) = e_0(d,t_a) + [a_m(d-6h,6h) - a_0(d-6h,6h)] \quad (6)$$

Each ensemble of analyses includes 10 perturbed analyses and an unperturbed one (due to limited computer resources, the number of perturbed analyses could not have been extended to more than 10). For the first 5 perturbed forecasts, perturbed analysis number 1 is used, $m(j)=1$ for $j=1,5$, and in general $m(j) = 1 + \left[\frac{j-1}{5}\right]$.

To understand the effect on ensemble prediction of perturbations defined using SVs or perturbed analyses, and to investigate the potential use of perturbed analyses in ensemble prediction, the following four types of ensembles are compared:

- SVINI, with initial perturbations defined using only initial SVs, see Eq. (5.a)
- EDA, with initial perturbations defined using only perturbed analyses, see Eq. (5.c)
- EDA-SVINI, with initial perturbations defined using perturbed analyses and initial-time SVs, see Eq. (5.d)
- SVINI-EVO, with initial perturbations defined using initial and evolved SVs, see Eq. (5.b)

Figures 1 and 2 are schematic illustrations of the configuration used in the SVINI-EVO (which corresponds to the ECMWF operational ensemble system at the time of writing) and the EDA ensembles. The SVINI configuration is as the SVINI-EVO but without the evolved SVs, while the EDA-SVINI configuration is a combination of the SVINI and the EDA.

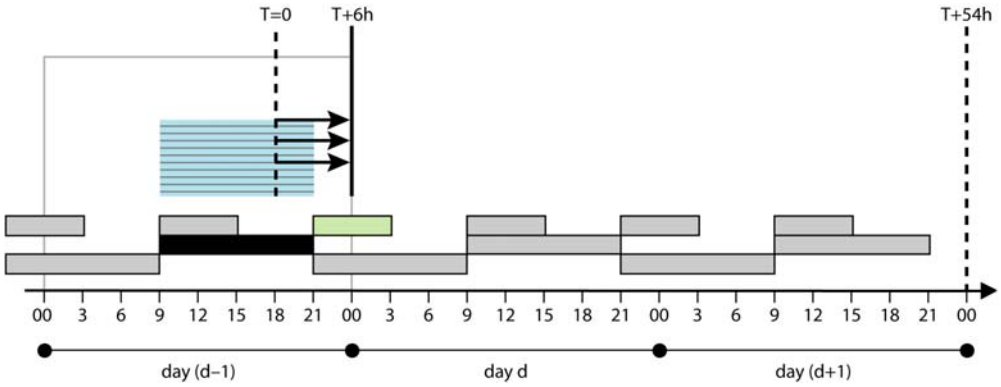

Figure 2: Schematic of the configuration used to generate the ensemble of perturbed analyses at 00UTC of day d*. The 12h-long black and grey boxes between 9-21UTC and 21-9UTC mark the times during which the 12-hour 4D-Var runs; the 6h-long green and grey boxes between 21-03UTC and 9-15UTC mark the times during which the early-delivery 6h 4D-Var analysis runs. The unperturbed analysis at 00UTC of day* d *is defined by the 6h 4D-Var analysis generated by the early-delivery suite (green box). The members of the ensemble of analyses used at 00UTC of day* d *(blue box with black lines) are generated by 12h 4D-Var cycles running between 9UTC and 21UTC of day* (d-1)*.*

All ensemble forecasts have been run for 10 days and with 50 perturbed and one unperturbed (the control) member. Most of the results presented in this work refer to the 500 hPa geopotential height and the 850 hPa temperature over the Northern Hemisphere extra-tropics (North of 20ºN), and the 850 hPa temperature and longitudinal wind component over the tropics (from 20ºS to 20ºN).

3. Characteristics of EDA- and SV-based perturbations

Figure 3 shows two perturbations at initial time, one from the SVINI and one from the EDA ensembles, for the forecasts started at 12 UTC on the 22nd of September 2006, in terms of temperature at 700 hPa. The SVINI and EDA perturbations have similar local maxima in temperature, but the EDA perturbations have larger local maxima in wind (not shown). The SVINI perturbations are more localized in the areas identified by the SVs, and have only a very small component in the tropical band (by construction). The EDA perturbation has smaller scales, and extends also in the tropics. The localization of the SVINI perturbation can be seen also in the vertical cross sections shown in Fig. 4; note that from this figure the vertical tilt typical of SV perturbations growing via baroclinic processes can be detected. Figure 4 confirms that the EDA perturbations are less localized also vertically, and shows that they are characterized by a less evident tilt.

These different characteristics of SV- and EDA-based perturbations are confirmed by the comparison of average values (computed over 20 cases) of the squared ensemble spread measured around the control forecast (defined as the average squared-difference between each perturbation and the control forecast). As shown in the top panels of Fig. 5, at initial time SVINI perturbations are confined to total wave-numbers smaller than 42 (by construction), with a similar amplitude to the EDA perturbations in terms of 500 hPa

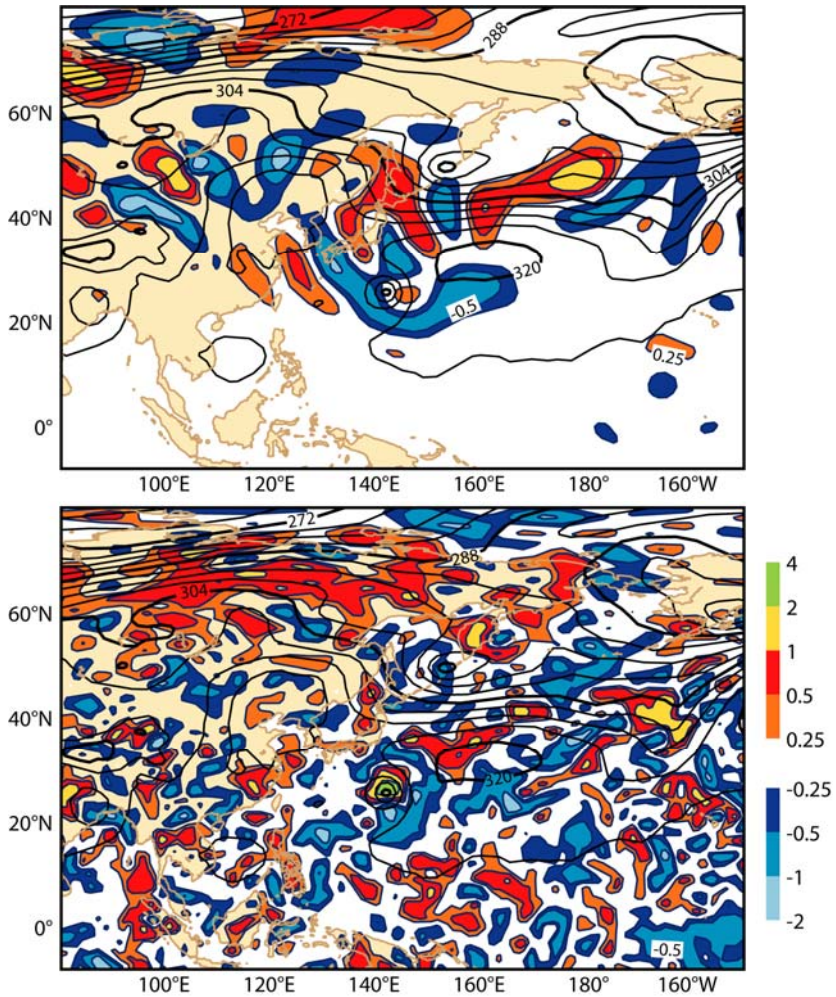

Figure 3: Initial-time perturbations of the forecast started at 12 UTC on the 22nd of September 2006 for the SVINI (top panel) and the EDA (bottom panel) ensembles. The solid lines show the control (con) initial state in terms of geopotential height, while shadings show the member-5 perturbations, i.e. the difference (mem5-con)*, in terms of temperature at 700 hPa. Contour interval is 4 dam for geopotential height and -2/-1/-0.5/-0.25/0.25/0.5/1/2/4 degrees for perturbation temperature.*

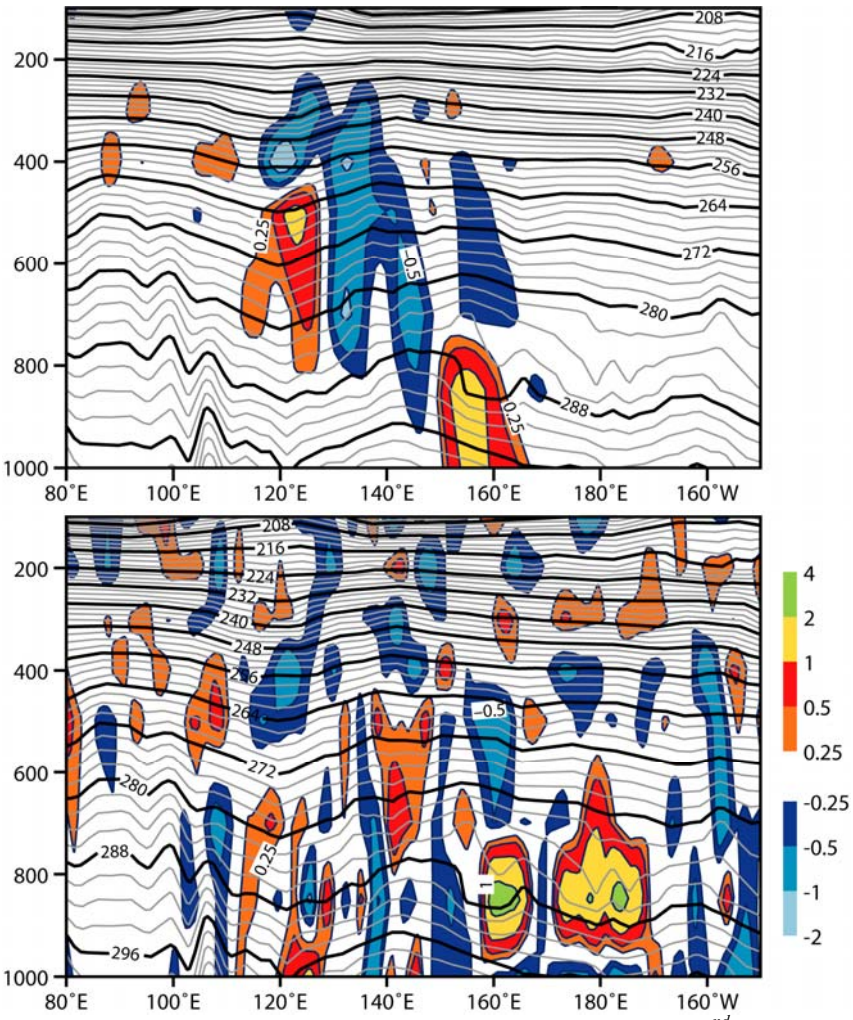

Figure 4: Initial-time perturbations of the forecast started at 12 UTC on the 22nd of September 2006 for the SVINI (top panel) and the EDA (bottom panel) ensembles. The solid lines show the control (con) initial state in terms of temperature, and the shadings show the member-5 perturbations, i.e. the difference (mem5-con), in terms of temperature for a vertical cross section at 30°N latitude. Contour interval is 2 degrees for the control temperature and -2/-1/-0.5/-0.25/0.25/0.5/1/2/4 degrees for perturbation temperature.

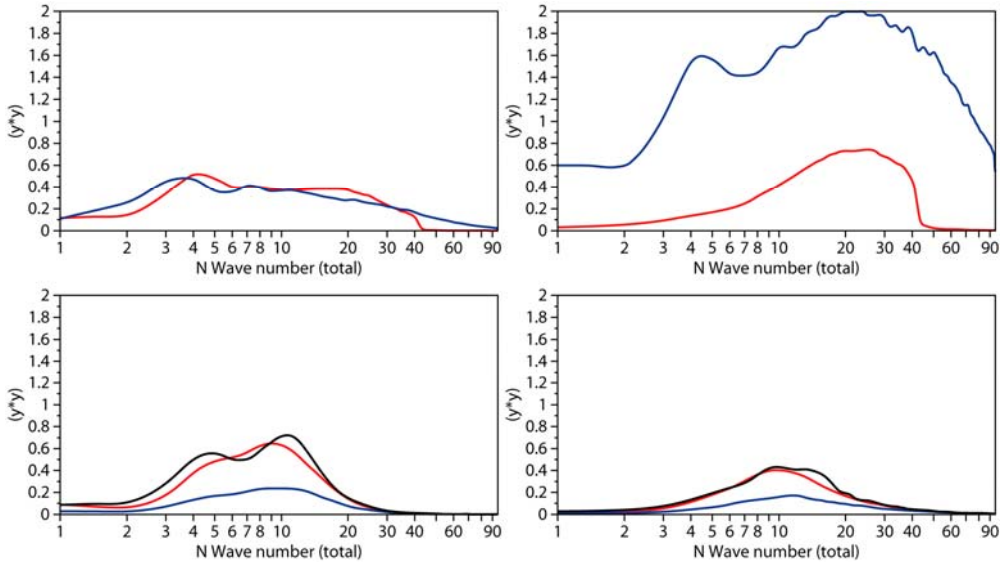

Figure 5: Average (20 cases) variance of the SVINI ensemble (red lines) and the EDA ensemble (blue lines) at initial time (top panels, multiplied by 500) and at T+120h (bottom panels) as a function of total wave number, computed over the Northern Hemisphere extra-tropics (north of 30°N) for the 500 hPa geopotential height (left panels) and the 850 hPa temperature (right panels). As reference for forecast times, the bottom panels show the average spectra of the control forecast error (black lines).

geopotential height but with a much smaller amplitude in terms of 850 hPa temperature. At forecast step T+120h (Fig. 5, bottom panels) the spread of the EDA ensemble is about half the spread of the SVINI ensemble. The comparison of the ensemble spread with the control error (black lines in the bottom panels of Fig. 5) indicates that at T+120h the SVINI spread is very close to the control error, while the EDA spread is much smaller. More comments on the level of ensemble spread and on perturbation growth in the different ensembles are reported in section 4.

Following *Buizza* (1998), to quantify the differences between the initial perturbations of the four types of ensembles, the similarity index and the forecast-error projection index have been computed. First, for each ensemble, a set of 50 ortho-normal vectors has been generated using the 50 initial perturbations defined by the difference between the 50 perturbed and the unperturbed (control) initial states applying a Gram-Schmidt orthogonalization procedure. Then, for any two ensembles X and Y, the similarity index between the two ortho-normal bases at initial time $v_{X,j}(d,0)$ and $v_{Y,j}(d,t)$ has been computed to measure the degree of similarity between the two sets of initial-time perturbations:

$$SI_{X,Y}(d,t=0) = \frac{1}{50}\sum_{j,k=1}^{50} < v_{X,j}(d,t=0); v_{y,k}(d,t=0) >^2 \qquad (7)$$

where <..;..> denotes the Euclidean inner product.

The similarity index is 1 for ensembles that span exactly the same sub-space of the forecast phase-space, while it is 0 for orthogonal sub-spaces.

Finally, the forecast error projection index, defined as the percentage of control forecast error explained by each ortho-normal basis, has been computed:

$$\begin{aligned} err(d,t) &= e_0(d,t) - a(d+t) \\ proj_X(d,t) &= \sum_{j=1}^{50} < err(d+t); v_{X,j}(d,t) > v_{X,j}(d,t) \\ PI_X(d,t) &= \frac{|proj_X(d,t)|}{|err(d,t)|} \end{aligned} \qquad (8)$$

where $a(d+t)$ is the T_L799L91 high-resolution analysis interpolated to the ensemble resolution (T_L399L62). The projection index is 1 if the control forecast error is confined to the subspace of the forecast phase-space spanned by the X basis, while it is 0 if it is confined to an orthogonal one.

Similarity indices between the different sub-spaces have been computed up to the forecast time for which SVs are optimized, i.e. 48 hours. Figure 6 shows the 20-case average similarity indices computed for all possible couples of experiments for three variables: 500 hPa geopotential heights and 850 hPa temperatures over the Northern Hemisphere, and the 850 hPa temperatures over the tropics. Results show that in the early forecast range SVINI and EDA, and SVINI-EVO and EDA perturbations are the most different. This confirms the differences detected in the comparison of single perturbations of the SVINI and the EDA ensembles discussed above. Note also that the similarity between SVINI and EDA-SVINI is close to the similarity between SVINI and SVINI-EVO: this indicates that adding to SVINI perturbations a set of perturbations defined either by the evolved singular vectors or the perturbed analysis has a similar impact on the dimensionality of the sub-space spanned by the initial perturbations. Note that over the Northern Hemisphere the EDA-SVINI perturbations are more similar to SVINI perturbation than to EDA perturbations: this is not surprising given the fact that there are only 10 different perturbed analyses. But note that the reverse is true over the tropics: again, this is not surprising given the fact that only a limited region of the tropics is sampled by the tropical targeted SVs.

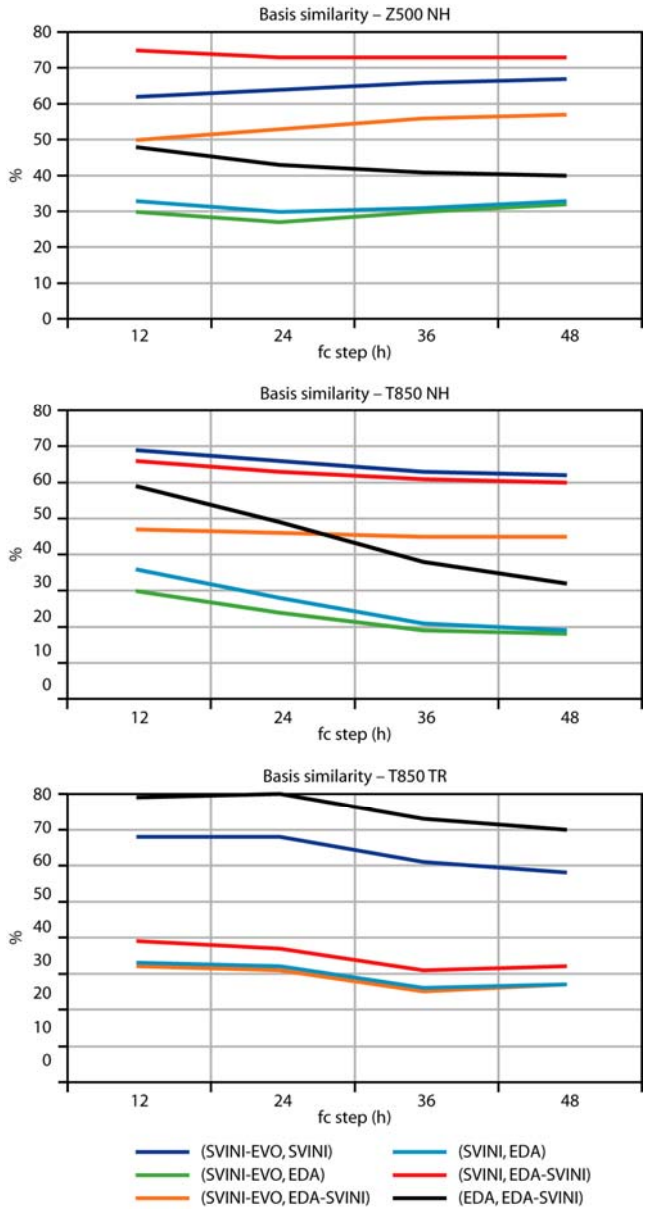

Figure 6: Average (20 cases) similarity index between different sets of ortho-normal bases defined using SVINI, SVINI-EVO, EDA and EDA-SVINI perturbations, computed for the 500 hPa geopotential height over the Northern Hemisphere (top panel), the 850 hPa temperature over the Northern Hemisphere (middle panel), and the 850 hPa temperature over the tropics (bottom panel).

Figure 7 shows the 20-case average (control error) projection indices for all four ensembles for up to T+48h. Over the Northern Hemisphere, the EDA perturbations have the second largest projection indices up to T+36h, but thereafter the two SVINI and SVINI-EVO perturbations have the second largest values. The fact that the SV-based perturbations are less optimal than EDA-based perturbations up to T+36h is not surprising, given the fact that the SVs are optimized for a 48-hour time interval. Over the tropics, the EDA perturbations have the second largest projection indices for the whole forecast range. The fact that for this region the SV-based perturbations are less optimal than the EDA-based perturbations is a consequence of the fact that by construction tropical SVs are computed only for some selected regions. Overall, the EDA-SVINI perturbations have the largest index over both the Northern Hemisphere and the tropics, suggesting that combining these two sets of perturbations could lead to an ensemble system spanning a subspace that includes a larger component of the control forecast error.

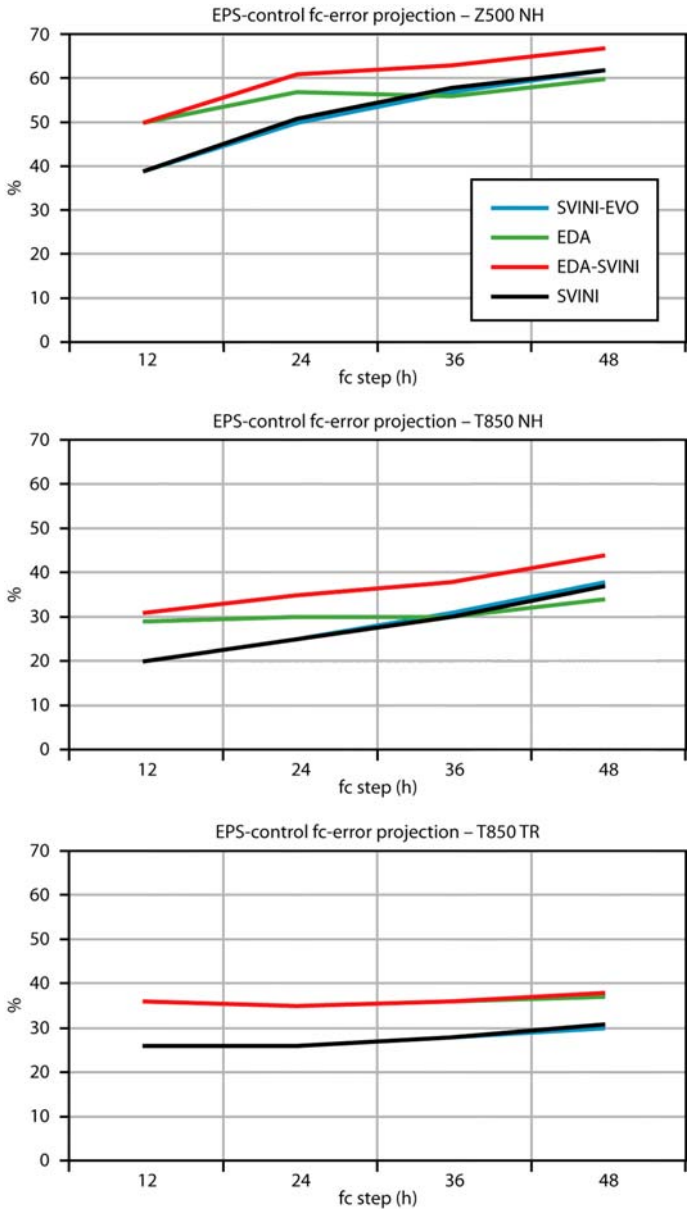

Figure 7: Average (20 cases) Percentage of control forecast error explained by ortho-normal bases defined using SVINI, SVINI-EVO, EDA and EDA-SVINI perturbations, computed for the 500 hPa geopotential height over the Northern Hemisphere (top panel), the 850 hPa temperature over the Northern Hemisphere (middle panel), and the 850 hPa temperature over the tropics (bottom panel).

In conclusion, the results discussed in this section indicates first of all that SV- and EDA-based perturbations have different characteristics (coverage, scale, tilt, amplitude) and span different regions of the phase-space of the system, and second that the subspace spanned by combined SVINI-EDA perturbations explains a larger percentage of control forecast error than if used individually. In the next section, the characteristics of the four ensembles (SVINI, EDA, SVINI-EVO and EDA-SVINI) are discussed.

4. Use of SV- and EDA-based perturbations in the ECMWF EPS

The performances of the SVINI, EDA, EDA-SVINI and SVINI-EVO ensembles have been compared in terms of ensemble spread measured by the ensemble standard deviation (STD, the average distance of an ensemble member from the ensemble-mean), root-mean-square-error (RMSE) for the ensemble-mean and the ranked probability skill score (RPSS) for probabilistic predictions. The RPSS has been computed using ten climatologically equally likely categories, with skill computed with respect to a climatological distribution (*Jung and Leutbecher* 2008). Attention has been focused on the 500 hPa geopotential height and the 850 hPa temperature over Northern Hemisphere, and the 850 hPa temperature and zonal-wind component over the tropics. When the average performance of two ensemble systems is compared, the [0.05,0.95] confidence interval of the difference between the scores of the two systems has been estimated using a boot-strapping technique (in figures that show the average scores of two systems, the confidence intervals are shown by bars drawn around one of the two score curves).

Figures 8 and 9 show the average RMSE of the ensemble-mean forecast and the ensemble spread (measured by the ensemble STD) for the 850 hPa temperature over the Northern Hemisphere and the tropics. Over the Northern Hemisphere, results show that after forecast step T+48h the STD of the EDA ensemble is ~50% smaller than the RMSE of the ensemble-mean, while for the other three ensembles the STD and the RMSE of the ensemble-mean are very close. Over the tropics, by contrast, the spread of the EDA and the EDA-SVINI ensembles is only ~25% smaller than the RMSE of the ensemble-mean, while the STD of the SVINI and the SVINI-EVO ensembles is ~50% smaller than the RMSE of the ensemble mean. These average results are in agreement with the results presented in section 3 for one case. In terms of RMSE of the ensemble-mean, over the Northern Hemisphere the four ensembles are equivalent up to forecast step T+120h, but thereafter the EDA ensemble-mean has the largest values (with differences statistically significant after forecast step T+108h). Over the tropics, the ensemble-mean of the EDA and the EDA-SVINI ensembles is the smallest between forecast steps T+48h and T+120h. These results are confirmed by Figures 10-11, which show the STD and the RMSE of the ensemble-mean for the 500 hPa geopotential height over the Northern Hemisphere and the 850 hPa zonal-wind component over the tropics.

Figure 12 shows the average RPSS of probabilistic forecasts of the 850 hPa temperature over the Northern Hemisphere and the tropics, and Fig. 13 the average RPSS of probabilistic forecasts of the 500 hPa geopotential height over the Northern Hemisphere and the 850 hPa zonal-wind component over the tropics. Over the Northern Hemisphere, the EDA ensemble has a statistically significant lower skill than the SVINI ensemble, and the EDA-SVINI and the SVINI-EVO ensembles have very similar skill (with difference not statistically significant). Over the tropics, the EDA ensemble has a statistically significant better skill than the SVINI ensemble, and the EDA-SVINI ensemble has a statistically significantly better skill than the SVINI-EVO ensemble

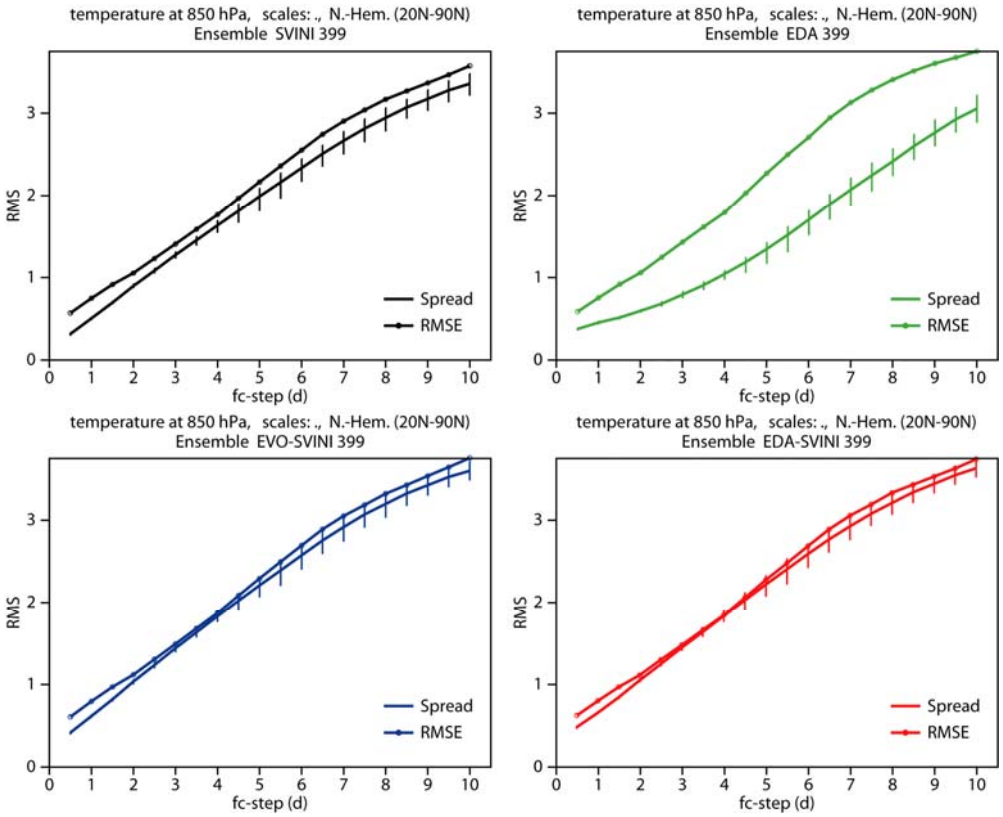

Figure 8: Average (20 cases) ensemble-mean RMSE (solid lines) and ensemble spread (solid lines with full circles) of SVINI (black lines, top-left panel), EDA (green lines, top-right panel), SVINI-EVO (blue lines, bottom-left panel) and EDA-SVINI (red lines, bottom-right panel) ensembles, computed for the 850 hPa temperature over the Northern Hemisphere extra-tropics. In each panel, bars around the RMSE lines the [0.05,0.95] confidence interval for the difference (spread-RMSE) estimated using a bootstrapping technique (see text for more details).

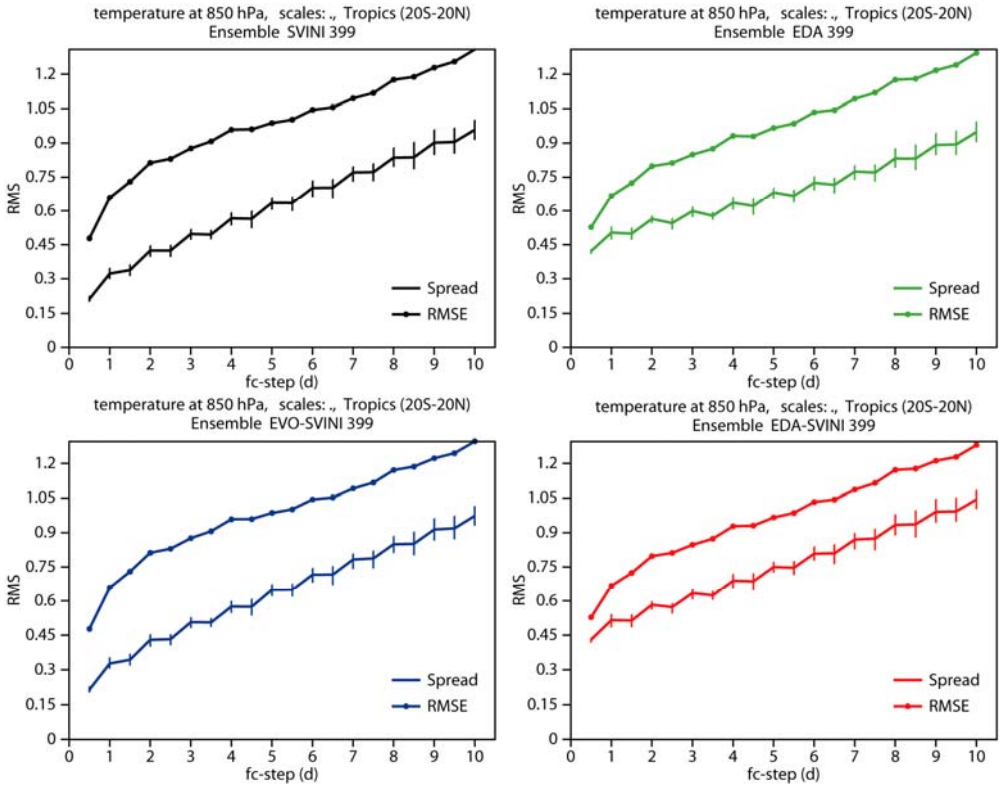

Figure 9: As Fig. 8 but for the 850 hPa temperature over the tropics.

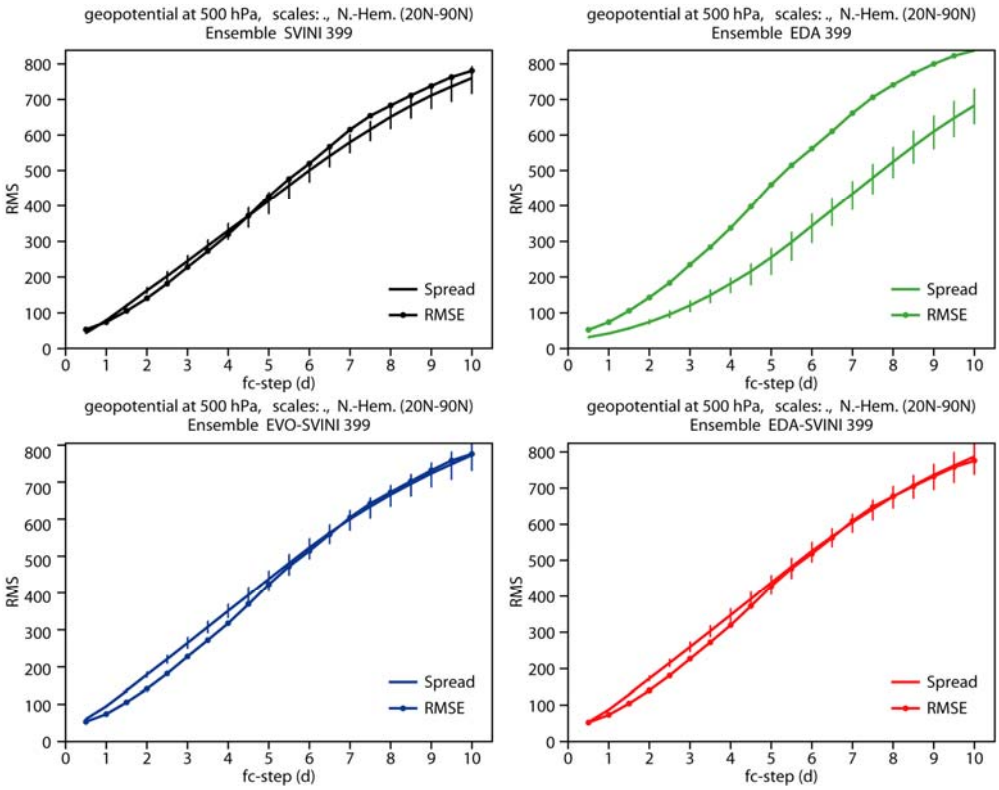

Figure 10: As Fig. 8 but for the 500 hPa geopotential height over the Northern Hemisphere extra-tropics.

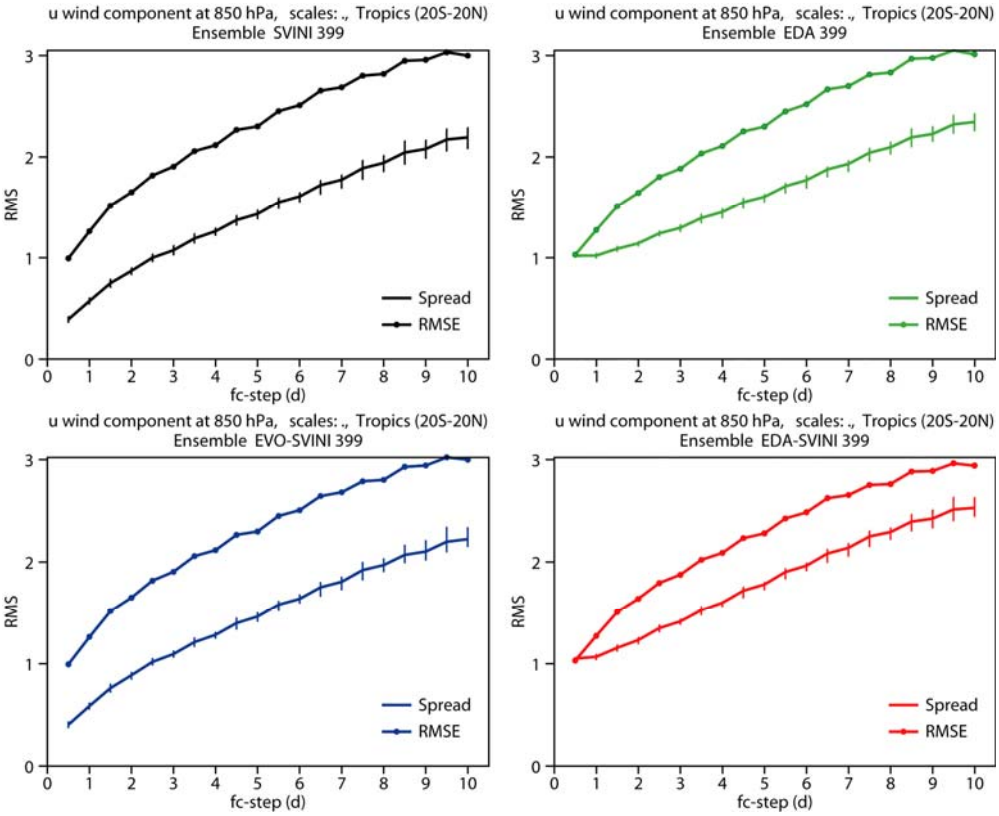

Figure 11: As Fig. 8 but for the zonal-wind component at 850 hPa over the tropics.

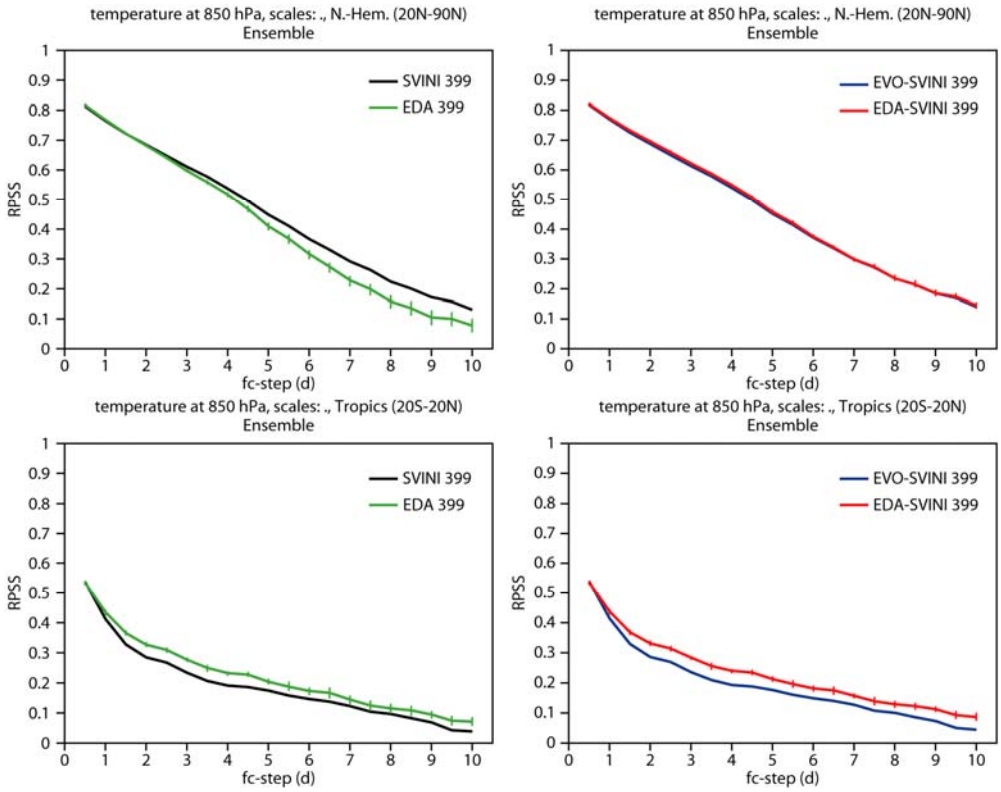

Figure 12: Average (20 cases) ranked probability skill score (solid lines) of SVINI-EVO (right panels, blue lines), EDA-SVINI (right panels, red lines), SVINI (left panels, black lines) and EDA (left panels, green lines) ensembles, computed for the 850 hPa temperature over the Northern Hemisphere extra-tropics (top panels) and the tropics (bottom panels). In each panel, bars around one of the two lines denote the [0.05,0.95] confidence interval for the difference of the scores of the two ensembles estimated using a bootstrapping technique (see text for more details).

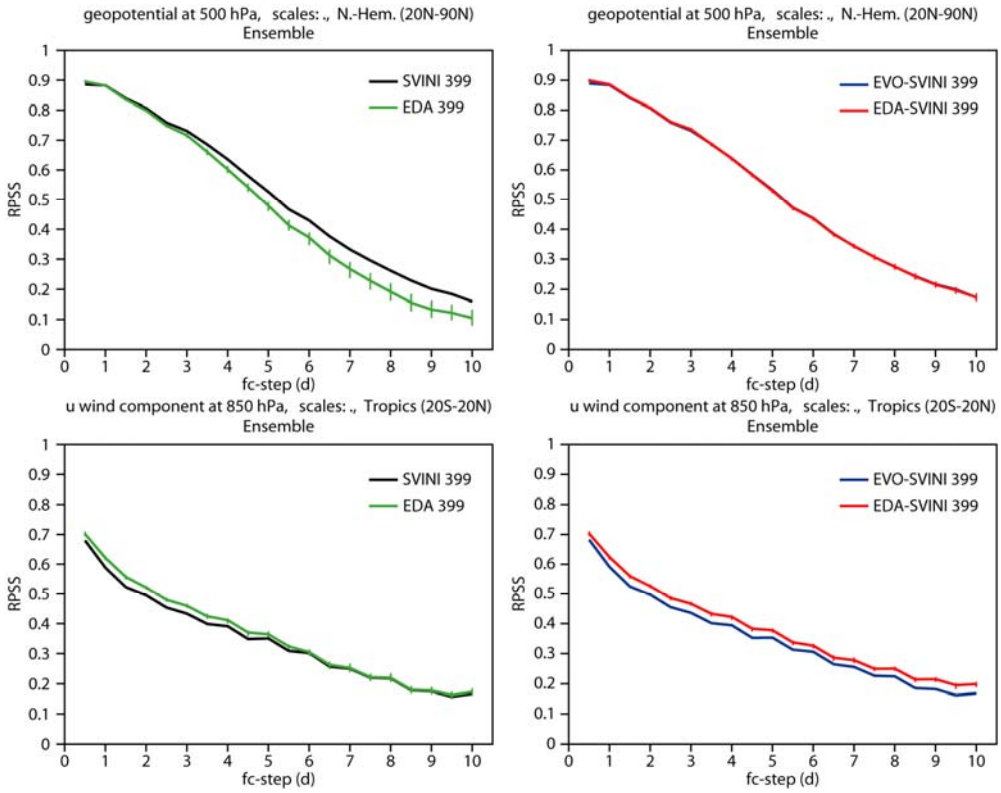

Figure 13: As Fig. 12 but for the 500 hPa geopotential height over the Northern Hemisphere extra-tropics (top panels) and for the zonal-wind component at 850 hPa over the tropics (bottom panels).

Overall, the EDA-SVINI ensemble shows the best agreement between STD and RMSE of the ensemble-mean, smallest ensemble-mean RMSE and the best RPSS for probabilistic forecasts. Having, on average, a good agreement between STD and ensemble-mean RMSE is considered as a necessary condition for a skilful ensemble system: in fact, in the ideal situation that one perturbed member is a perfect forecast, the average STD should coincide with the average ensemble-mean RMSE, see e.g. the discussion in *Buizza and Palmer* (1998). The EDA-SVINI ensemble benefits from the strengths of the SVs over the extra-tropics and of the EDA perturbations over the tropics.

5. Discussion and conclusions

The simulation of initial uncertainties is one of the key problems in ensemble prediction. At ECMWF, these uncertainties have been simulated with singular vectors (SVs), perturbations characterized by the fastest growth, measured using a total energy norm (*Palmer et al* 1998), over a finite time interval. In the current operational system, different sets of singular vectors are used to better sample the initial uncertainties: to improve geographical coverage, SVs are computed separately over the Northern and the Southern Hemisphere extra-tropics, and for up to 6 local regions in the tropics. Furthermore, initial-time SVs growing into the first 48 hours of the forecast range are mixed with evolved SVs, computed to grow during the 48 hours leading to the analysis time: the former represent uncertainties growing during the forecast time, while the latter represent uncertainties that have been growing during the data-assimilation cycle. The initial-time and evolved SVs, computed for the different areas, are orthogonalized and scaled to have an amplitude comparable to the analysis error estimate provided by the ECMWF data assimilation system (see *Palmer et al* 2007 for a review of the evolution of the ECMWF ensemble system from its implementation in 1992 to date).

In this work, experiments have been performed for 20 cases spanning a 40-day period with a TL399L62 resolution, and results have been presented to address three key questions:

- How similar/different are EDA-based and SV-based initial perturbations?
- What is the difference in skill in an ensemble using EDA-perturbations only and an ensemble using SV-perturbations only?
- Can the skill of the EPS be improved by combining EDA- and SV-based perturbations?

Considering the first question, results presented in section 3 have shown that the EDA perturbations are less localized geographically and have a better coverage of the tropics. They also have smaller scales than SV-based perturbations, and have a less evident vertical tilt with height, which explains why they grow less with the forecast time. The comparison of EDA-based and SV-based ortho-normal bases has given a measure of their similarity: results indicate that they span different regions of the phase-space of the system. Considering the second question, results presented in section 4 have shown that the EDA ensemble has too little spread, due to a combination of too small initial amplitudes and too slow growth, and this has a negative impact on its performance. Considering the third question, combining the EDA and the initial-time SVs gives a superior ensemble system to the operational system, which uses a combination of initial-time and evolved SVs. In fact, the EDA-SVINI ensemble shows the best agreement between STD and RMSE of the ensemble-mean, the smallest ensemble-mean RMSE and the best RPSS for probabilistic forecasts. The EDA-SVINI ensemble benefits from the strengths of the SVs over the extra-tropics and of the EDA perturbations over the tropics. Results discussed in section 3 have also shown that the subspace spanned by combined SVINI-EDA perturbations explains a larger percentage of control forecast error than if used individually. It is worth to mention that similar conclusions were reached by earlier experiments performed with a T_L255L40 model version (*Leutbecher et al* 2007).

As briefly reviewed in the Introduction, several studies have compared the strengths and weaknesses of the different methods used in ensemble prediction to simulate initial uncertainties following three approaches: a simple-but-same-model environment, a complex-but-different-model environment or a complex-and-same-model environment approach. It is interesting to contrast the results discussed in this communication with the conclusions of three very recent studies obtained using the three different approaches.

Firstly, consider *Descamps and Talagrand* (2007), who followed a simple-but-same-model environment approach and compared the performance of ensembles based on SVs, bred vectors, ensemble Kalman filter and Ensemble Transformed Kalman filter, and concluded that the two latter outperform the two former ensembles. In our view, the key reason why their results do not agree with our results is the over-simplification of the approach that they followed. Their over-simplification, on the one hand due to the use of a perfect model assumption (they only included the effect of model errors due to a resolution truncation from T81 to T21) and to their use of simple, low resolution models (e.g. without moist processes) compared to the one used in this study, and on the other hand due to the use of synthetic data as verification, makes it difficult, if not impossible, to apply their conclusions to the real problem. This is, in our view, the main reason why they found that an SV-based system is sub-optimal, while results discussed in this work indicate that an SV-based approach still outperforms an EDA-based system.

Secondly, consider *Park et al* (2008), who followed a *complex-but-different-model environment* approach and compared the performance of seven operational ensemble prediction systems. This work provides some explanations of the different performances detected by *Park et al* (2008). For example, the fact that the EDA ensemble is severely under-dispersive compared to the SVINI ensemble can help understanding why the Canadian ensemble system needs to start with much larger initial perturbations to have a spread comparable to the ECMWF ensemble. In fact, the EDA ensemble has been generated following the approach used by *Houtekamer et al* (1996) to set-up the Canadian ensemble. It is interesting to compare one result of *Descamp and Talagrand* (2007) with *Park et al* (2008): in *Descamp and Talagrand* (2007) the SV method produces an under-dispersive ensemble at all forecast ranges (see their Fig. 4). This is in contrast with the results of *Park et al* (2008), who show that SV-based ensembles have the best tuned ensemble spread, especially in the medium-range (see their Fig. 4). This disagreement between the two works confirms, in our view, the fact that it is difficult to apply *Descamp and Talagrand* (2007)'s conclusions to real systems.

Thirdly, consider *Magnusson et al* (2008), who followed a *complex-and-same-model environment* approach, and compared the performance of ensembles based on the ECMWF model, and with initial perturbations defined using SVs and two types of bred vectors. The key difference between this study and *Magnusson et al* (2008), is that this work uses an ensemble of analyses instead of bred vectors as an alternative to SVs. Compared to using bred vector, using an ensemble of analyses is, in our view, a better approach, since the ensemble of analyses is influenced by observation coverage and observation error statistics, while bred vectors only try to mimic the effect of an analysis cycle. The ensemble of analyses has larger spread in areas characterized by less coverage and/or larger observation errors, and is thus better capable to simulate the effect of observation errors on initial conditions. *Magnusson et al* (2008)'s conclusions are very similar to the ones drawn in this work: both studies indicate that over the extra-tropics SVs provides a better performance, and that the ECMWF operational ensemble system has a poorer performance over the tropics due to the fact that the SVs sample only a sub-region of the tropical band. Furthermore, this work goes beyond simply comparing the two methods, and proposes to combine initial-time SVs and an ensemble of analyses to benefit from the strengths of both of them.

The three questions posed above will be re-addressed in the near future with a new set of ensemble of analyses generated using a new model version that is producing a larger, more realistic spread, and possibly different schemes. The main differences between this new model cycle and the one used in the experiments

discussed in this work are that it has a different physics and a reduced initial perturbation amplitude (*Bechtold et al* 2008), and a revised version of the stochastic backscatter scheme used to simulate model uncertainties in the ensemble of data assimilation (*Shutts* 2005, *Berner et al* 2008). Furthermore, work is in progress to take account of correlations of radiance errors in the ensemble of analyses: this could lead to a larger initial spread among the EDA analyses, leading to a more realistic representation of analysis errors. Finally, the re-assessment will also consider the possibility to include an extra set of SVs computed for the entire tropical band.

Acknowledgements

Tim Palmer is acknowledged for very valuable discussions and comments on an earlier version of this manuscript. Anabel Bowen is thanked for substantially improving the quality of the figures.

References

Anderson, J. L., 1997: Impact of dynamical constrains on the selection of initial conditions for ensemble predictions: low-order perfect model results. *Mon. Wea. Rev.*, **125**, 2969-2983.

Barkmeijer, J., Buizza, R., and Palmer, T. N., 1999: 3D-Var Hessian singular vectors and their potential use in the ECMWF Ensemble Prediction System. *Q. J. R. Meteorol. Soc.*, **125**, 2333-2351.

Barkmeijer, J., Buizza, R., Palmer, T. N., Puri, K., and Mahfouf, J.-F., 2001: Tropical singular vectors computed with linearized diabatic physics. *Q. J. R. Meteorol. Soc.*, **127**, 685-708.

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

Bechtold, P., Kohler, M., Jung, T., Leutbecher, M., Doblas-Reyes, F., Rodwell, M., Vitart, F. and Balsamo G. 2008: Advances in simulating atmospheric variability with the ECMWF model: from synoptic to decadal time-scales. *Q. J. R. Meteorol. Soc.*, submitted.

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

Bowler, N., 2006: Comparison of error breeding, singular vectors, random perturbations and ensemble Kalman filter, Part I: theoretical aspects. *Tellus*, **58A**, 538-548.

Buizza, R., 1998: Impact of horizontal diffusion on T21, T42 and T63 singular vectors. *J. Atmos. Sci.*, **55**, 1069-1083.

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

Buizza, R., and Palmer, T. N., 1998: Impact of ensemble size on the skill and the potential skill of an ensemble prediction system. *Mon. Wea. Rev.*, **126**, 2503-2518.

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

Buizza, R., and Palmer, T. N., 1999b: Ensemble Data Assimilation. Pre-prints of the 17th *Conference on Weather Analysis and Forecasting*, 13-17 September 1999, Denver Colorado, 231-234.

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**, 1076-1097.

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

Descamps, L., and Talagrand, O., 2007: On some aspects of the definition of initial conditions for ensemble prediction. *Mon. Wea. Rev.*, **135**, 3260-3272.

Ehrendorfer, M. and Beck, A., 2003: Singular vector-based multivariate normal sampling in ensemble prediction. *ECMWF Technical Memorandum No. 416*, pp.23 (http://www.ecmwf.int/publications/)

Hamill, T. M., Snyder, C., and Morss, R. E., 2000: A comparison of probabilistic forecasts from bred, singular-vector, and perturbation observation ensembles. *Mon. Wea. Rev.*, **128**, 1835-1851.

Houtekamer, P. L., and Derome, J., 1995: Methods for ensemble prediction. *Mon. Wea. Rev*., **126**, 796-811.

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

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

Isaksen, L., Fisher, M., and Berner, J., 2007: Use of analysis ensembles in estimating flow-dependent background error variance. Proceedings of the ECMWF Workshop on *Flow-dependent aspects of data assimilation*, 11-13 June 2007, pp. 65-86. (http://www.ecmwf.int/publications)

Jung, T., and Leutbecher, M., 2008: Scale-dependent verification of ensemble forecasts. *Q. J. R. Meteorol. Soc.*, submitted. (Also available as ECMWF Tech Memo No. 551, http://www.ecmwf.int/publications)

Leutbecher, M, 2005: On Ensemble Prediction Using Singular Vectors Started from Forecasts. *Mon. Wea. Rev.*, **133**, 3038-3046.

Leutbecher, M. and T. N. Palmer, 2008: Ensemble forecasting. *J. Comp. Phys.*, **227**, 3515-3539. (Also available as ECMWF Tech. Memo. No. 514. http://www.ecmwf.int/publications)

Leutbecher, M., Buizza, R., and Isaksen, L., 2007: Ensemble forecasting and flow-dependent estimates of initial uncertainty. Proceedings of the ECMWF Workshop on *Flow-dependent aspects of data assimilation*, 11-13 June 2007, pp. 185-201. (http://www.ecmwf.int/publications)

Lorenz, E. N., 1963: Deterministic non-periodic flow. *J. Atmos. Sci.*, **20**, 130-141.

Lorenz, E. N., 1996: Predictability: a problem partly solved. Proceedings of the *ECMWF Seminar on Predictability,* 1-18. (http://www.ecmwf.int/publications)

Magnusson, L., Leutbecher, M., and Kallen, E., 2008: Comparison between singular vectors and breeding vectors as initial perturbations for the ECMWF ensemble prediction system. *Mon. Wea. Rev.*, under revision.

Marshall, J., and Molteni, F., 1993: Toward a dynamical understanding of planetary-scale flow regimes. *J. Atmos. Sci.*, **50**, 1792-1818.

Molteni, F., Buizza, R., Palmer, T. N., and Petroliagis, T., 1996: The new ECMWF ensemble prediction system: methodology and validation. *Q. J. R. Meteorol. Soc.*, **122**, 73-119.

Palmer, T. N., F. Molteni, R. Mureau, and R. Buizza, 1993: Ensemble prediction. ECMWF Proceedings of the ECMWF Seminar on *Validation of models over Europe: Vol. I*, (http://www.ecmwf.int/publications)

Palmer, T. N., Gelaro, R., Barkmeijer, J., and Buizza, R., 1998: Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.*, **55**, 633-653.

Palmer, T. N., Buizza, R., Leutbecher, M., Hagedorn, R., Jung, T., Rodwell, M, Vitart, F., Berner, J., Hagel, E., Lawrence, A., Pappenberger, F., Park, Y.-Y., van Bremen, L., Gilmour, I., and Smith, L., 2007: The ECMWF Ensemble Prediction System: recent and on-going developments. *ECMWF Technical Memorandum No. 540*, (http://www.ecmwf.int/publications).

Park, Y.-Y., Buizza, R., and Leutbecher, M., 2008: TIGGE: preliminary results on comparing and combining ensembles. *Q. J. R. Meteorol. Soc.*, submitted (also published as ECMWF Technical Memorandum No. 548, http://www.ecmwf.int/publications).

Puri, K. Barkmeijer, J., and Palmer, T. N., 2001: Ensemble prediction of tropical cyclones using targeted diabatic singular vectors. *Q. J. R. Meteorol. Soc.*, **127**, 709-731.

Rabier, F., Jäärvinen, H., Klinker, E., Mahfouf, J.-F., and Simmons, A., 2000: The ECMWF operational implementation of four dimensional variational assimilation. Part I: experimental results with simplified physics. *Q. J. R. Meteorol. Soc.*, **126**, 1148-1170.

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

Toth, Z., and Kalnay, E., 1997: Ensemble Forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297-3319.

Tracton, M. S., and Kalnay, E., 1993: Operational ensemble prediction at the National Meteorological Center: practical aspects. *Weather and Forecasting*, **8**, 379-398.

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

Wei, M., Toth, Z., Wobus, R., and Zhu, Y., 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system *Tellus A*, **60**, 62-79.

Wilks, D. S., 1995: *Statistical methods in the atmospheric sciences*. Academic Press, Inc., San Diego, pp. 467 (ISBN 0-12-751965-3).