

Ensemble forecasting and data assimilation: two problems with the same solution?

Eugenia Kalnay^(1,2), Brian Hunt⁽²⁾, Edward Ott⁽²⁾ and Istvan Szunyogh^(1,2)
⁽¹⁾Department of Meteorology and ⁽²⁾Chaos Group
University of Maryland, College Park, MD, 20742

1. Introduction

Until 1991, operational NWP centers used to run a single computer forecast started from initial conditions given by the analysis, which is the best available estimate of the state of the atmosphere at the initial time. In December 1992, both NCEP and ECMWF started running *ensembles* of forecasts from slightly perturbed initial conditions (Molteni and Palmer, 1993, Buizza et al, 1998, [Buizza, 2005](#), Toth and Kalnay, 1993, Tracton and Kalnay, 1993, Toth and Kalnay, 1997).

Ensemble forecasting provides human forecasters with a range of possible solutions, whose average is generally more accurate than the single deterministic forecast (e.g., Fig. 4), and whose spread gives information about the forecast errors. It also provides a quantitative basis for probabilistic forecasting.

Schematic Fig. 1 shows the essential components of an ensemble: a control forecast started from the analysis, two additional forecasts started from two perturbations to the analysis (in this example the same perturbation is added and subtracted from the analysis so that the ensemble mean perturbation is zero), the ensemble average, and the “truth”, or forecast verification, which becomes available later. The first schematic shows an example of a “good ensemble” in which “truth” looks like a member of the ensemble. In this case, the ensemble average is closer to the truth than the control due to nonlinear filtering of errors, and the ensemble spread is related to the forecast error. The second schematic is an example of a “bad ensemble”: due to a poor choice of initial perturbations and/or to model deficiencies, the forecasts are not able to track the verifying truth, and remain relatively close to each other compared to the truth. In this case the ensemble is not helpful to the forecasters at all, since the lack of ensemble spread would give them unjustified overconfidence in the erroneous forecast. Nevertheless, for NWP development, the “bad” ensemble is still very useful: after verification time, the poor performance of the ensemble clearly indicates that there is a deficiency in the *forecasting system*. A single failed “deterministic” forecast, by contrast, would not be able to distinguish between a deficiency in the system and growth of errors in the initial conditions as the cause of failure.

Ideally, the initial perturbations should sample well the analysis “errors of the day” and the spread among the ensemble members should be similar to that of the forecast errors. The two essential problems in the design of an ensemble forecasting system are how to create effective initial perturbations, and how to

handle model deficiencies, which, unless included in the ensemble, result in the forecast error being larger than the ensemble spread.

In this paper we give a brief historic review of ensemble forecasting, current methods to create perturbations, and discuss Ensemble Kalman Filtering methods designed to perform efficient data assimilation, but which can, at the same time, provide optimal initial ensemble perturbations and estimate the model errors. We compare the properties of Ensemble Kalman Filters with those of 4D-Var, the only operational data assimilation method that currently includes the effect of the “errors of the day”.

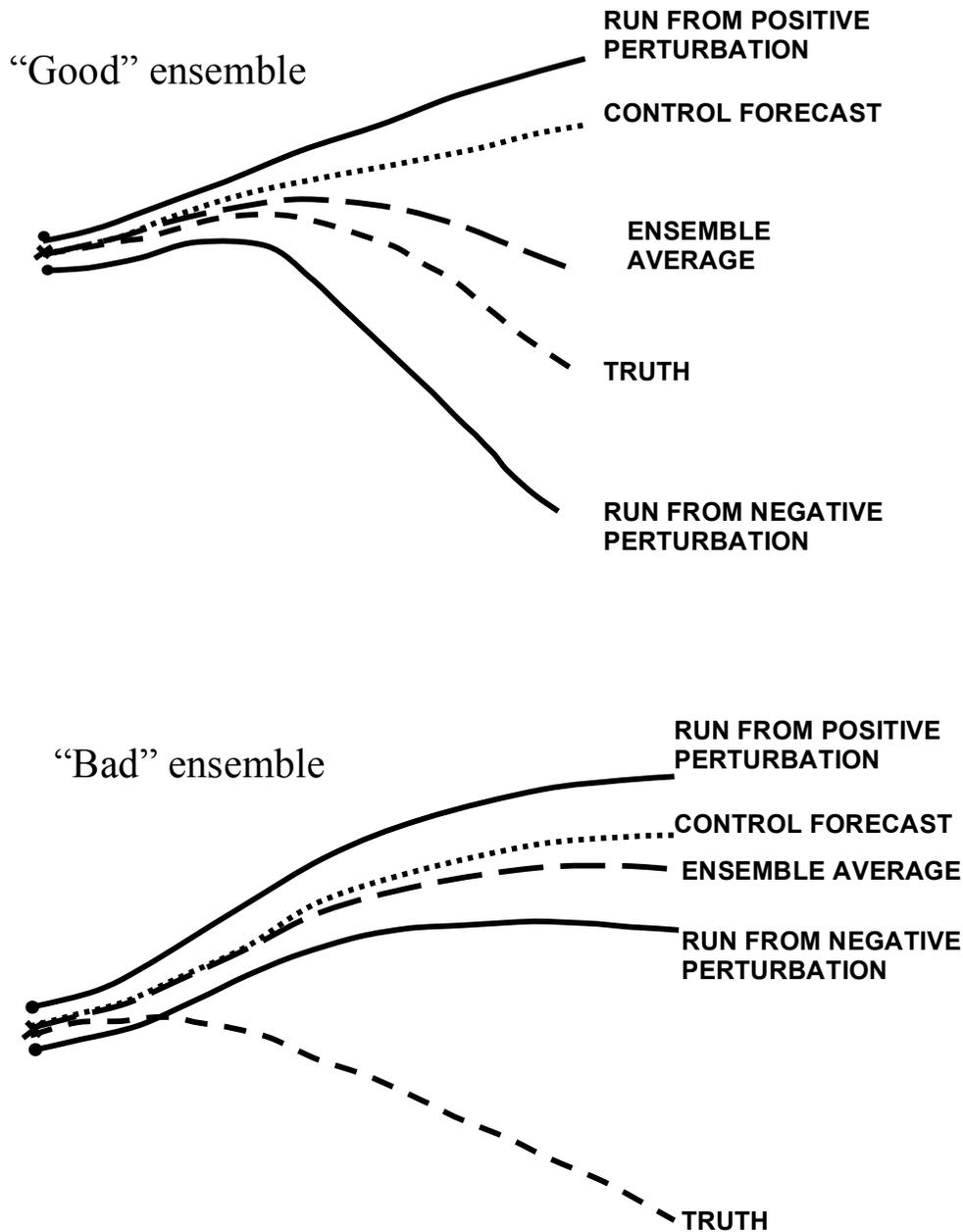


Figure 1: Schematic of the essential components of an ensemble of forecasts: The analysis (denoted by a cross) constitutes the initial condition for the control forecast (dotted); two initial perturbations (dots around the analysis), chosen in this case to be equal and opposite; the perturbed forecasts (full line); the ensemble average (long dashes); and the verifying analysis or truth (dashed). The first schematic is a “good ensemble” in which the truth is a plausible member of the ensemble. The second is an example of a bad ensemble, quite different from the truth, pointing to the presence of deficiencies in *the forecasting system* (in the analysis, in the ensemble perturbations and/or in the model).

2. Ensemble forecasting methods

Human forecasters have always performed *subjective ensemble forecasting* by either checking forecasts from previous days, and/or comparing forecasts from different centers, approaches similar to lagged forecasting and multiple systems forecasting. The consistency among these forecasts at a given verification time provided a level of confidence in the forecasts, confidence that changed from day to day and from region to region.

2.1 Early methods

Epstein (1969), introduced the idea of Stochastic-Dynamic forecasting (SDF), and pointed out that it could be also used in the analysis cycle to provide the forecast error covariance. Epstein designed SDF as a shortcut to estimate the true probability distribution of the forecast uncertainty, given by the Liouville equation (Ehrendorfer, 2005), which Epstein approximated running a huge (500) number of perturbed (Monte Carlo) integrations for the 3-variable Lorenz (1963) model. However, since SDF involves the integration of forecast equations for each element of the covariance matrix, this method is still not computationally feasible for models with large number of degrees of freedom.

Leith (1974) suggested the direct use of a Monte Carlo Forecasting approach (MCF), where random perturbations sampling the *estimated* analysis error covariance are added to the initial conditions. He indicated that a relatively small number of integrations (of the order of 8) is enough to approximate an important property of an infinite ensemble. In a large ensemble the average forecast error variance at long time leads converges to the climatological error variance, whereas individual forecast errors have an average error variance that is *twice* as large. Since the estimation of the analysis error covariance was constant in time, the MCF method did not include the effects of “errors of the day”. Errico and Baumhefner (1987) applied this method to realistic global models, using perturbations that represented a realistic (but constant) estimation of the error statistics in the initial conditions. Hollingsworth (1980) showed that for atmospheric models, random errors in the initial conditions took too long to spin-up into growing “errors of the day”, making MCF an inefficient approach for ensemble forecasting.

Hoffman and Kalnay (1983) suggested as an alternative to MCF, the Lagged Averaged Forecasting (LAF) method, in which forecasts from earlier analyses were included in the ensemble. Since the ensemble members are forecasts of different length, they should be weighted with weights estimated from the average forecast error for each lead-time. Hoffman and Kalnay found that compared to MCF, LAF resulted in a better prediction of skill (a stronger relationship between ensemble spread and error), presumably because LAF includes effects from “errors of the day”. The main disadvantage of LAF, namely

that “older” forecasts are less accurate and should have less weight, was addressed by the Scaled LAF (SLAF) approach of Ebisuzaki and Kalnay (1991), in which the LAF perturbations (difference between the forecast and the current analysis) are scaled by their “age”, so that all the initial SLAF perturbations have errors of similar magnitude. They also suggested that the scaled perturbations should be both added and subtracted from the analysis, thus increasing the ensemble size and the probability of “encompassing” the true solution within the ensemble. SLAF can be easily implemented in both global and regional models, including the impact of perturbed boundary conditions (Hou et al, 2001).

2.2 Operational Ensemble Forecasting methods

In December 1992 two methods to create perturbations became operational at NCEP and at ECMWF. They are based on bred vectors and singular vectors respectively, and like LAF, they include “errors of the day”. These and other methods that have since become operational or are under consideration in operational centers are briefly discussed. More details are given in Kalnay (2003).

a. Singular Vectors (SVs)

Singular vectors are the linear perturbations of a control forecast that grow fastest within a certain time interval (Lorenz, 1965), known as “optimization period”, using a specific norm to measure their size. SVs are strongly sensitive to the length of the interval and to the choice of norm (Ahlquist, 2000). Ehrendorfer and Tribbia (1997) showed that *if the initial norm used to derive the singular vectors is the analysis error covariance norm, then the initial singular vectors evolve into the eigenvectors of the forecast error covariance at the end of the optimization period*. This indicates that if the analysis error covariance is known, then singular vectors based on this specific norm are ideal perturbations.

ECMWF implemented an ensemble system with initial perturbations based on singular vectors using a total energy norm (Molteni and Palmer, 1993, Molteni et al, 1996, Buizza et al, 1997, Palmer et al 1998, [Buizza, 2005](#)).

b. Bred Vectors (BVs)

Breeding is a nonlinear generalization of the method to obtain leading Lyapunov vectors, which are the sustained fastest growing perturbations (Toth and Kalnay, 1993, 1997). Bred Vectors (like leading Lyapunov Vectors) are independent of the norm and represent the shapes of the instabilities growing upon the evolving flow. In areas where the evolving flow is very unstable (and where forecast errors grow fast), the BVs tend to align themselves along very low dimensional subspaces (the locally most unstable perturbations). An example of such situation is shown in Fig. 2, where the forecast uncertainty in a 2.5-day forecast

of a storm is very large, but the subspace of the ensemble uncertainty lies *within a one-dimensional space*. In this extreme (but not uncommon) case, a single observation at 500hPa would be able to identify the best solution. The differences between the forecasts are the bred vectors. The nonlinear nature of BVs allows for the saturation of fast growing instabilities such as convection, or, in the case of ENSO coupled instabilities, the weather noise (Peña and Kalnay, 2004).

In unstable areas of fast growth, BVs tend to have shapes that are independent of the forecast length or the norm, and depend only on the verification time. This suggests that forecast errors, to the extent that they reflect instabilities of the background flow, should have shapes similar to bred vectors, and this has been confirmed with model simulations (Corazza et al, 2003).

NCEP implemented an ensemble system based on breeding in 1992, and the US Navy, the National Centre for Medium Range Weather Forecasting in India, and the South African Meteorological Weather Service implemented similar systems. The Japanese Meteorological Agency implemented an ensemble system based on breeding, but they impose a partial global orthogonalization among the bred vectors, reducing the tendency of the bred vectors to converge towards a low dimensional space of the most unstable directions (Kyouda and Kusunoki, 2002).

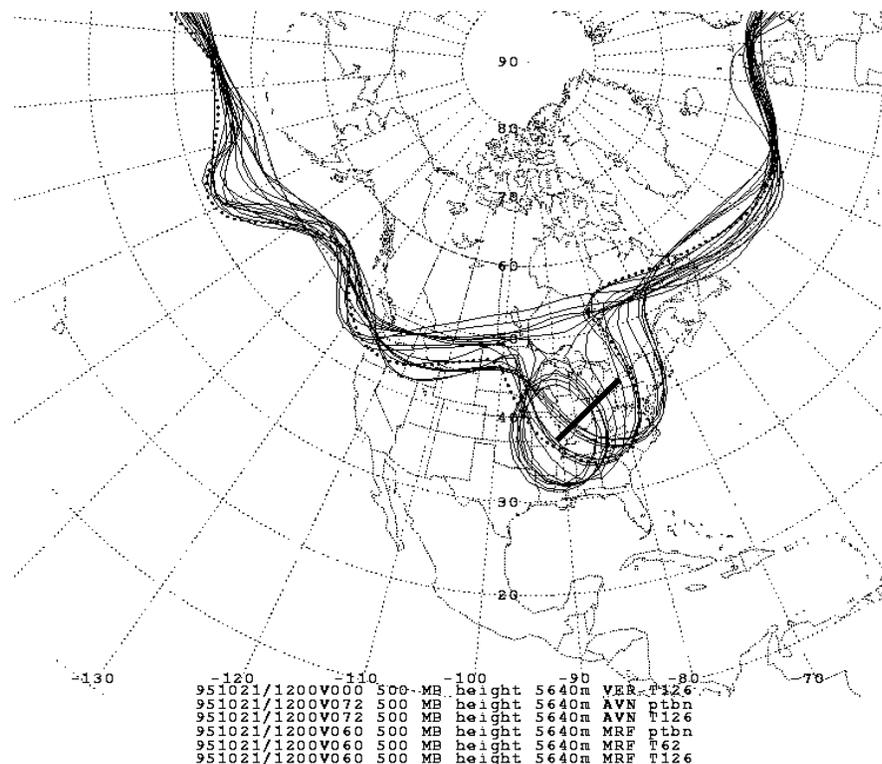


Fig. 2: “Spaghetti plots” showing a 2.5 day ensemble forecast verifying on 95/10/21. Each 5640gpm contour at 500 hPa corresponds to one ensemble forecast, and the dotted line is the verifying analysis. Note that the uncertainty in the location of the center of the predicted storm in the Midwest of the US is 6

c. Multiple data assimilation ensembles

Houtekamer et al (1996) developed a system based on running an ensemble of data assimilation systems using perturbed observations, implemented in the Canadian Weather Service. Hamill et al (2000) showed that in a quasi-geostrophic system, a multiple data assimilation system performs better than the singular vectors and the breeding approaches. With respect to the computational cost, the multiple data assimilation system and the singular vector approach are comparable, whereas breeding is essentially cost-free.

d. Perturbed physical parameterizations

The methods discussed above only include perturbations in the initial conditions, assuming that the error growth due to model deficiencies is small compared to that due to unstable growth of initial errors. Several groups have also introduced changes in the physical parameterizations to allow for the inclusion of uncertainties in the model formulation (Houtekamer et al, 1996, Stensrud et al, 2000). Buizza et al (1999) developed a perturbation approach that introduces a stochastic perturbation of the impact of subgrid scale physical parameterizations by multiplying the time derivative of the “physics” by a random number normally distributed with mean 1 and standard deviation 0.2. This simple approach resulted in a clear improvement of the performance of the ensemble system.

Both the perturbations of the initial conditions and of the subgrid scale physical parameterizations have been shown to be successful towards achieving the goals of ensemble forecasting. However, since they both introduce changes to the best estimate of the initial conditions and the model, which are used for the control forecast, it is not surprising that on the average, the individual forecasts are worse than the unperturbed control (see example in Figure 3). Nevertheless, the ensemble average is an improvement over the control, especially after the perturbations grow into a nonlinear regime that tends to filter out some of the errors because the ensemble solutions tend to diverge in the most uncertain components of the forecast.

e. Multiple system ensembles

An alternative to the introduction of perturbations is the use of multiple systems. Different operational or research centers, each aiming to be the best, choose different competitive approaches to data assimilation and forecasting systems. In principle, a combination of these different systems should sample well the uncertainty in both the models and the initial conditions. It has been known that the ensemble average of multiple center forecasts is significantly better than even the very best individual forecasting system (e.g., Kalnay and Ham, 1989, Fritsch et al, 2000, Arribas et al, 2004). This has also been shown to be true for regional models (Hou et al, 2001). Krishnamurti et al (2000), [Krishnamurti \(2005\)](#)

introduced the concept of “superensemble”, a multiple system ensemble where linear regression is used to correct the bias of each of the operational systems from past performance, and the predictors are combined to minimize the ensemble average prediction errors. This results in remarkable forecast improvements (e.g., Fig. 4). This method is also called “poor person” ensemble approach to reflect that it does not require running a forecasting system (Arribas et al, 2004).

Anomaly Correlation for the winter of 1997/98 (from 00Z) Controls (T126 and T62) and 10 perturbed ensemble forecasts

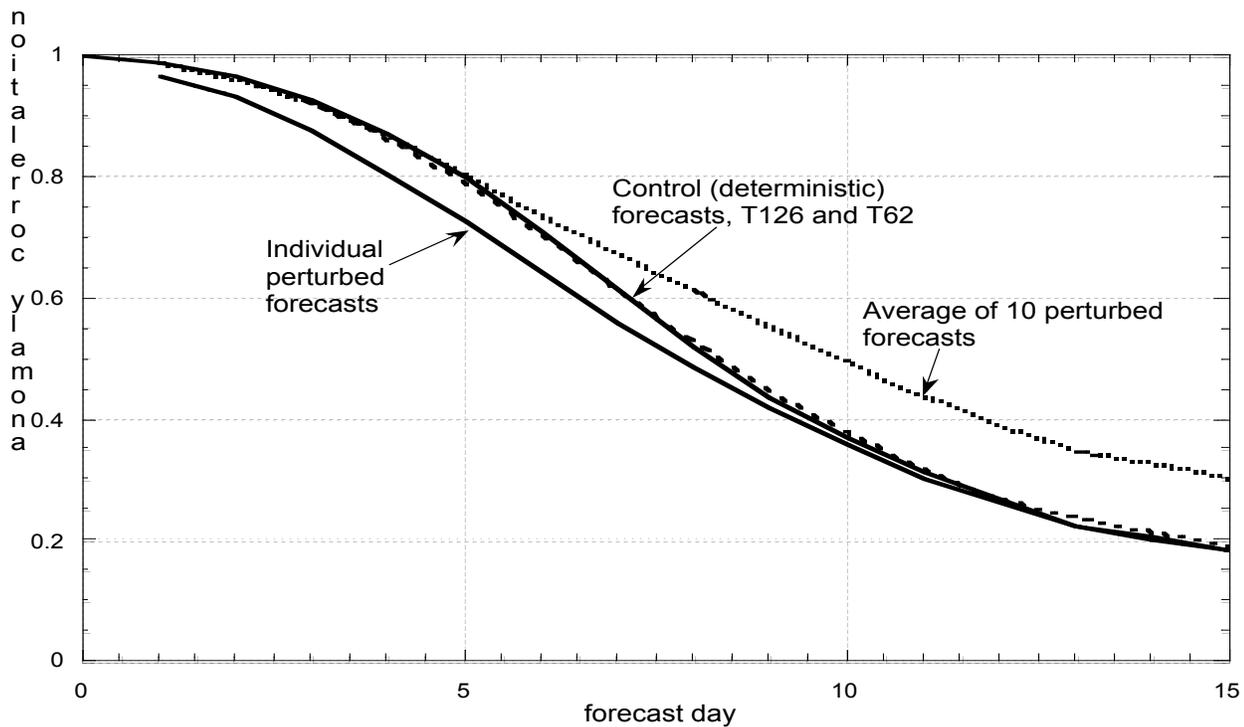


Fig. 3: Average anomaly correlation of the ensembles during the winter of 1997/1998 (data courtesy Jae Schemm, of NCEP). Note that, on the average, the individual perturbed forecasts are worse than the control, but their ensemble average is better.

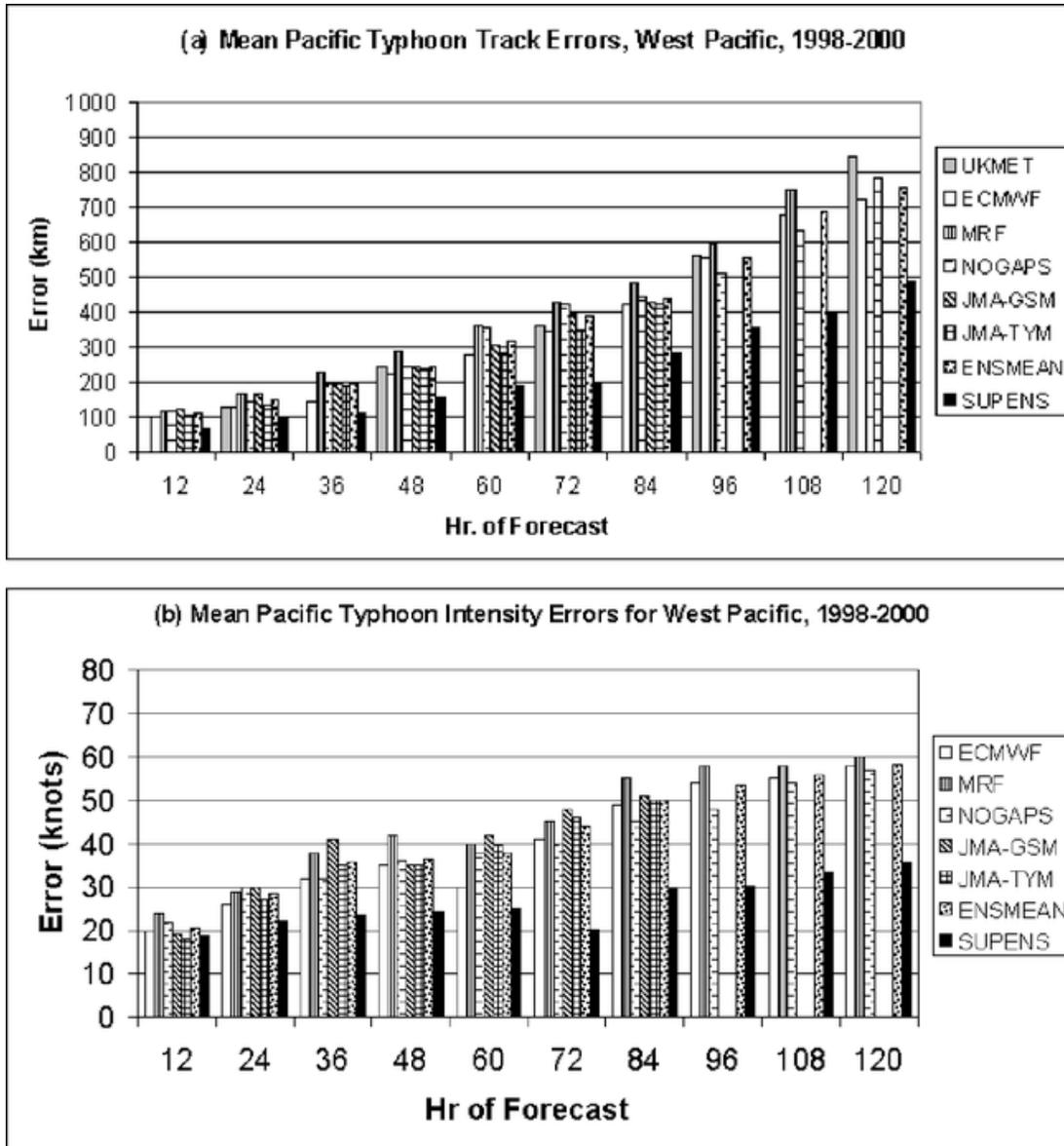


Fig. 4: Mean typhoon track and intensity errors for the west Pacific, 1998-2000 for several global models (ECMWF, UK MetOffice, US MRF, US Navy NOGAPS, Japan JMA global and typhoon models, the ensemble mean of all these models), and a superensemble obtained by linear regression trained on the first half of the same season. From Kumar et al. (2003).

f. Other methods

This field is changing quickly, and improvements and changes to the operational systems are under development. For example, ECMWF has implemented changes in the length of the optimization period for the SVs, a combination of initial and final or evolved SVs (which are more similar to BVs), and the introduction of a stochastic element in the physical parameterizations, all of which contributed to improvements in the ensemble performance (Buizza et al, 2000). NCEP is considering the implementation of the Ensemble Transform Kalman Filter (Bishop et al, 2001) to replace breeding since the BVs have a tendency to converge to the leading Lyapunov Vectors, providing insufficient spread. A recent comparison of the ensemble performance of the Canadian, US and ECMWF systems (Buizza et al, 2004) suggests that the ECMWF SV perturbations behave well beyond the 2-day optimization period, at which time the model advantages of the ECMWF system are also paramount, giving the best performance. The BV perturbations of NCEP are somewhat better at short ranges, and the multiple analysis method performs well at all ranges. See [Hagedorn et al, 2005](#), [Lalurette and van der Grijn, 2005](#), [Mylne, 2005](#) [Palmer, 2005](#), [Shukla and Kinter, 2005](#), [Tibaldi et al, 2005](#), [Timmermann and Jin, 2005](#), [Latif et al, 2005](#), [Toth et al, 2005](#), [Waliser, 2005](#) and [Webster et al, 2005](#) for further discussions on applications of ensemble forecasting at all ranges.

3. Ensemble Kalman Filtering for data assimilation and ensemble forecasting.

As indicated before (e.g., Ehrendorfer and Tribbia, 1997), “perfect” initial perturbations for ensemble forecasting should sample well the analysis errors. Thus, ideal initial perturbations $\delta\mathbf{x}_i$ in an ensemble with K members should have a covariance that spans well the analysis error covariance \mathbf{A} :

$$\frac{1}{K-1} \sum_{i=1}^K \delta\mathbf{x}_i \delta\mathbf{x}_i^T \approx \mathbf{A} \quad (1.1).$$

Until recently, the problem has been the lack of knowledge of \mathbf{A} , which changes substantially from day to day and from region to region due to instabilities of the background flow. These instabilities, associated with the “errors of the day”, are not taken into account in data assimilation systems, except for 4D-Var and Kalman Filtering, methods that are computationally very expensive. 4D-Var has been implemented at ECMWF, MeteoFrance, the Canadian Meteorological Service (CMS) (Andersson et al, 2004, Gauthier, 2004, Desroziers et al, 2003) [and Japan](#). The implementation of 4D-Var at ECMWF required some cost-saving simplifications such as reducing the resolution of the analysis from ~40km in the forecast model to ~120km in the assimilation model. Even at the lower resolution,

the 4D-Var clearly outperformed the previous operational 3D-Var. Versions of 4D-Var are also under development in other centers.

The original formulations of Kalman Filtering and Extended Kalman Filtering are prohibitive because they would require the equivalent of N model integrations, where N is the number of degrees of freedom (dof) of the model, of the order of a 10^6 or more. Considerable work has been done on finding simplifying assumptions to reduce the cost of KF (e.g., Fisher et al, 2003), but so far they have been successful only under special circumstances.

An alternative approach to Kalman Filter which is much less expensive is Ensemble Kalman Filter (EnKF) suggested first by Evensen (1994) and now under development by several groups (Hamill, 2005). In the original Extended KF the background error covariance is updated by using the linear tangent model and its adjoint (Ghil et al, 1981, Cohn, 1997), equivalent to running the model about N times, where N is the number of degrees of freedom. By contrast, EnKF attempts to estimate the evolution of the background error covariance from an ensemble of K forecasts, with $K \ll N$. In the formulation of EnKF of Evensen (1994) and Houtekamer and Mitchell (1998), ensembles of data assimilation are driven by *perturbed observations* and used to derive the background error covariance from the ensemble of forecasts. A more recent class of EnKF is known as *square-root filters* (Tippett et al, 2002, Bishop et al., 2001, Anderson, 2001, Whitaker and Hamill, 2002, Ott et al, 2002, Ott et al, 2004). The ensemble forecasts are used to obtain a most likely forecast (the ensemble mean

$\bar{\mathbf{x}}_b = \frac{1}{K} \sum_{i=1}^K \mathbf{x}_i$) and the background error covariance (obtained from the

perturbations of the ensemble with respect to the

mean $\mathbf{B} = \frac{1}{K-1} \sum_{i=1}^K (\mathbf{x}_i - \bar{\mathbf{x}}_b)(\mathbf{x}_i - \bar{\mathbf{x}}_b)^T$) at the time of the analysis. The full Kalman

Filter equations and the new observations are then used to obtain the most likely analysis \mathbf{x}_a , with the analysis increment (difference between the single analysis and the forecast) lying within the subspace of the forecast ensemble

perturbations. After this is completed, the new initial analysis perturbations $\delta \mathbf{x}_i$ for the next analysis cycle are obtained by solving the square root problem

$\frac{1}{K-1} \sum_{i=1}^K \delta \mathbf{x}_i \delta \mathbf{x}_i^T = \mathbf{A}$, where \mathbf{A} is the analysis error covariance estimated by the

Kalman Filter, and K is the number of ensemble members. The solution of this problem is not unique and different square-root filters have adopted different solutions to the square-root problem (Tippett et al, 2002, Ott et al, 2002). An advantage of the square-root filter approach is that there is no need to add perturbations to the observations, which reduces the sampling error (Whitaker and Hamill, 2002).

Several of the EnKF approaches reduce the computational cost by assimilating the observations one at a time, for the whole physical domain, a method known as sequential assimilation of observations. This is done using a localization of the error covariance in the horizontal and in the vertical, to avoid spurious long-distance correlations due to sampling. Although the sequential assimilation of observations is very efficient with limited observations (such as those available before 1979), it becomes less practical with the abundant satellite observations currently available (about 3.5 million data per assimilation cycle, about 40 times more than a decade ago) and planned for the next decade.

A different approach, also within the class of the square-root filters is the Local Ensemble Kalman Filter (LEKF) method (Ott et al, 2002, 2004, Szunyogh et al, 2004). In the EnKF problem is solved *locally in physical space*, not in observation space. For each grid point a local 3D volume of the order of 700km by 700km by a few vertical layers is used to perform the analysis. The Kalman Filter equations are solved exactly in the subspace locally spanned by the global ensemble members, using *all the observations* available within the volume. This localization in space results in a further reduction of the number K of ensemble members needed to obtain an accurate solution, so that matrix operations are done in a very low dimensional space. The analysis is carried out independently at each grid point, leading to a completely parallel algorithm.

In schematic Fig 6 we compare Ensemble Kalman Filtering (EnKF) with the 3D-Var approach, in which the background error covariance is estimated as an average over many cases. Fig. 6a shows how the 3D-Var analysis maximizes the joint probability defined by the observations error covariance and the background error covariance, both of which are high dimensional. Since the 3D-Var background error covariance is a statistical average, the 3D-Var analysis does not know about “errors of the day”. Fig. 6b shows that in the EnKF, the ensemble perturbations define a very low-dimensional subspace within which the forecast errors lie, and the KF analysis maximizes the joint probability within that subspace. Because the computations are performed within this subspace, the rank of the matrices involved is low, and the Kalman Filter equations providing the analysis and analysis error covariance can be solved directly, not iteratively.

Figure 7 shows the background errors and the analysis corrections based on a given set of noisy observations in a quasi-geostrophic data assimilation system (Morss et al, 2000, Hamill and Snyder, 2000, Corazza et al, 2004). The top panel corresponds to 3D-Var, with a background error covariance constant in time. Because the system does not know about the dynamical stretching produced by the “errors of the day”, the analysis increments introduced by the new observations tend to be isotropic. The bottom panel shows that the Local Ensemble Kalman Filter (Ott et al, 2004, Szunyogh et al, 2004) is much more efficient in correcting the background errors. The large improvements made on the analysis are also apparent in forecasts (not shown).

Performing the EnKF locally in space substantially reduces the number of ensemble members required for the analysis, as shown in Figure 8, from Ott et al, 2004, obtained using the Lorenz (1996) model. When performed globally, the number of ensemble members required for the EnKF to converge to the optimal value is proportional to the size of the model. When done locally, the number of ensemble members is reduced from 27 to 8, and it does not increase with the size of the model. In addition, the analysis for different grid points can be carried out in parallel, since they are independent from each other. The efficiency of localization is schematically shown in Figure 8, where three independent unstable regions A, B, C, can each have wavenumber 1 and wavenumber 2 instabilities, in the same way that different areas of the world can develop baroclinic instabilities that evolve independently from each other. Three of the possible ensemble perturbations are depicted. From a local point of view, the first two perturbations are enough to represent all possible combinations of wavenumbers 1 and 2 instabilities, whereas from a global point of view, the third perturbation and many others are linearly independent from the first two.

The LEKF has been tested in a “perfect model” mode, using the operational NCEP Global Forecasting System at a resolution of T62/28 levels, with excellent results (Szunyogh et al, 2004). Fig. 9 shows the evolution of the analysis of surface pressure, when about 11% of the points (separated by about 200km) have “rawinsonde” observations. The analysis errors for temperature and winds show similar quick convergence to values much smaller than the observational errors. Fig. 10 shows the vertical rms analysis errors for temperature and zonal wind for several levels of observational density, including 11% of the grid points, a density similar to that of the rawinsondes in the NH extratropics, and 2%, a density similar to rawinsondes in the SH. The ability of the LEKF to extract information through the knowledge of the “errors of the day” is very encouraging. Fig. 10 shows that despite the local nature of the analysis, in the perfect model simulation, the LEKF, with just 2% observations, is able to reach “superbalance”, being able to reproduce the evolution of not only slow synoptic waves but also that of a gravity wave present in the “nature” run.

4. Prospects for operational implementation

At the time of this writing (December 2004), tests of the LEKF with real observations have not yet been carried out. Tests with other square-root filters have been performed only with rawinsonde upper air observations and cloud tracked winds, and yield results comparable in the Northern Hemisphere with the 3D-Var used in the NCEP Reanalysis, which also used satellite retrievals (Jeff Whitaker, pers. comm., 2004). The perturbed observations approach has been tested at CMC in an operational environment and also yields results comparable to those of the very mature operational 3D-Var in the CMS (Houtekamer et al, 2004). By contrast, 4D-Var has been shown to be superior to 3D-Var in several centers (ECMWF, MeteoFrance, CMC). This relatively disappointing performance of the EnKF approach, although it is to be expected in the still early

stages of testing, is in contrast with the excellent results obtained with perfect model simulations.

From this experience we cannot say at this point in time whether EnKF will be able to compete with or replace 4D-Var in operational centers as the next data assimilation system of choice. However, EnKF has a number of very attractive advantages that hold promise once the new systems are tested and tuned. Because of familiarity we will describe the advantages of the LEKF system that has been developed at the University of Maryland (Ott et al, 2002, 2004, Hunt et al, 2004, Szunyogh et al, 2004), comparing with the corresponding characteristics of 4D-Var, but some of the advantages are generic to other EnKFs.

a) The LEKF is very efficient, due to its complete parallelism and relatively few required ensemble members, and can use data simultaneously. In the original formulation it takes only 15 minutes to assimilate 1.5 million observations using 40 ensemble members on a cluster of 25 dual processor 2.8 GH PCs. This includes the 6 hour 40-member ensemble forecast, which takes about 6 minutes (and which would be free within an operational center that performs ensemble forecasting). Hunt (pers. comm., 2003) has developed an alternative algorithm (Local Ensemble Transform Kalman Filter, [LETKF](#)) not based on singular value decomposition, which is about 3 times faster than the LEKF while yielding essentially identical results (I. Szunyogh, and E. Kostelich, pers. comm., 2004). The efficiency of the LEKF ensures that it can be used operationally with a resolution at least the same as that of the operational ensemble forecasting system.

b) Like other EnKF methods, the LEKF does not require the development and maintenance of the model's linear tangent or adjoint models, saving the large effort needed for these models, and avoiding the need for linear approximations, since the full nonlinear model is used for every operation.

c) Similarly, like other EnKF, the LEKF does not require the Jacobian or the adjoint of the observation operator H , another important advantage (Houtekamer and Mitchell, 2001, Miyoshi, 2004, Szunyogh et al, 2004). Basically, if we define a matrix \mathbf{E} of perturbation vectors, so that $\mathbf{B} = \mathbf{E}\mathbf{E}^T$; $\mathbf{H}\mathbf{B}\mathbf{H}^T = (\mathbf{H}\mathbf{E})(\mathbf{H}\mathbf{E})^T$, one can replace the linear $\mathbf{H}\mathbf{E}$ with its fully nonlinear expression,

$\mathbf{H}\mathbf{E} = \mathbf{H}[\delta\mathbf{x}_1 | \dots | \delta\mathbf{x}_k] \approx [H(\mathbf{x}_1) - H(\bar{\mathbf{x}}) | \dots | H(\mathbf{x}_k) - H(\bar{\mathbf{x}})]$, which is simpler and more accurate.

d) It can be easily extended to 4 dimensions (4DLEKF), so that observations can be assimilated at their time of observation, in between analysis times (Hunt et al, 2004). This is performed at a relatively low computational cost, by using the ensemble to "transport" observational increments from the time of observation to the analysis time. This is important since experience suggests that the ability to assimilate observations at the right time is the main advantage of 4D-Var (rather than the evolution of the covariance).

e) The 4DLEKF can be used as a smoother, taking advantage of future observations, as would be possible in a Reanalysis mode, at the cost of doubling the computations. At a given time t_0 a preliminary (operational) analysis and analysis error covariance is performed using only past observations, then ensemble forecasts between this time and the next analysis time t_1 are used to bring the observations in that interval back to t_0 , and a final, more accurate analysis and analysis error covariance can be obtained at t_0 . The problem of using the data twice (which would erroneously reduce the analysis error covariance) can be handled by increasing the observation error covariance (e.g., doubling it for observations used both at t_0 and t_1).

f) The LEKF provides diagnostic tools that can be used to tune the system in the future. One of them is the local effective dimension ED (Patil et al, 2000)

obtained from the singular values σ_i of the ensemble: $ED = \frac{\left(\sum_{i=1}^k \sigma_i\right)^2}{\sum_{i=1}^k \sigma_i^2}$. Szunyogh

et al (2004) showed that there is a strong relationship between the projection of the forecast error on the ensemble perturbations (not available in real data assimilation experiments) and the effective dimension. Their Fig. 7 shows that in regions where the projection of the error on the ensemble is close to 100%, as in most of the mid-latitudes, the effective dimension (about 10) is much smaller than the actual ensemble dimension (40). In the tropics, where the projection is only about 40%, the effective dimension is larger, of the order of 30 or more. Other diagnostics that can be used for tuning are comparisons of observations minus forecast and observations minus analysis, which can be compared with their predicted values.

g) EnKF can be used to obtain best estimates of model bias as part of the data assimilation (e.g., Anderson, 2001).

h) Finally, the LEKF provides estimates of the background and the analysis error covariances, (and thus ideal initial perturbations) although they are based on a space with limited dimension given by the ensemble size. This information has other important applications such as adaptive observations, a new area of great growth in recent years (Thorpe and Petersen, 2005). Fig. 12 suggests that EnKF makes possible to have interactive targeted observations, with remote sensing instruments dwelling only on regions identified by the ensemble as having low accuracy forecasts.

5. Final comments

Although the EnKF approach in general and the LEKF/LETkf in particular, seem to have many advantages, so far there has been no direct comparison between LEKF and 4D-Var in a realistic system. The EnKF is based on a reduced dimension, albeit computed with full nonlinearity, and it provides an estimate of the background and analysis covariances. 4D-Var requires an estimate of the background error covariance \mathbf{B} at the initial time of the assimilation window, and it is common to start with a 3D-Var \mathbf{B} . Obtaining a

better evolving estimate of \mathbf{B} with a reduced Kalman Filter has been difficult (Fisher et al, 2003). Mike Fisher showed in a recent presentation that 4D-Var is equivalent to a full Kalman Smoother (which at the end of the interval yields the same estimate as the Kalman Filter) if the assimilation window is sufficiently long, because the lack of an initial \mathbf{B} is “forgotten”. Fisher proposed that a 4D-Var with a 3-10 day window and including model errors (weak constraint) will provide a solution equivalent to a *full rank* Kalman Filter, and it could perhaps be made computationally affordable. This is an attractive alternative to EnKF and only experience will tell whether one is better than the other.

An important remaining problem is that of model deficiencies, leading to model systematic errors and to problems such as those suggested in figure 1b. A successful approach to start addressing this problem for ensemble forecasting applications is using multiple models (Krishnamurti et al, 2000, Hou et al, 2001). Other approaches (DelSole and Hou, 1999, Kaas et al, 1999) rely on empirical methods to reduce the errors using past observations. Yet another method, known as “dressing” adds random perturbations to the ensemble forecasts in order to reproduce the observed error covariance with the ensemble (Roulston and Smith, 2003, Wang and Bishop, 2004). It is possible that the Ensemble Kalman Filtering approach will also be able to handle efficiently model errors by augmenting the model variables with a relatively small number of parameters associated with model errors, and using the observations to estimate the optimal value of their time-varying coefficients.

If current research in Ensemble Kalman Filtering methods achieves their promise, then the problems of data assimilation and ensemble forecasting may indeed have a unified solution.

5. References

Ahlquist, Jon, 2000: Almost Anything can be a Singular Vector.
<http://www.met.fsu.edu/ftp/ahlquist/singvect.ps>

Anderson, J. L. 2001. An ensemble adjustment filter for data assimilation.
Mon. Wea. Rev. **129**, 2884–2903.

Anderson, J. L. and Anderson, S. L. 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts.
Mon. Wea. Rev. **127** 2741–2758.

Andersson, Erik, C. Cardinali, M. Fisher, E. Hólm, L. Isaksen, Y. Trémolet, and A. Hollingsworth, 2004: Developments in ECMWF's 4D-Var system. Paper J1.4, American Meteorological Society, Symposium on Forecasting the Weather and Climate of the Atmosphere and Ocean, 20th Conference on Weather Analysis and Forecasting, 16th Conference on Numerical Weather Prediction, Seattle, Washington, January 2004.

Arribas, A.; Robertson, KB; Mylne, KR Test of a poorman's ensemble prediction system for short-range probability forecasting. Submitted

Bishop, C. H., B. J. Etherton and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, 129, 420-436.

Buizza, R., M. Miller and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quart. Journal of the Royal Meteorolog. Soc.*, 125, 2887-1908.

Buizza, R., Barkmeijer, J., Palmer, T.N., & Richardson, D.S., 2000: Current status and future developments of the ECMWF Ensemble Prediction System. *Meteorol. Appl.*, 7, 163-175.

Buizza, R., P.L. Houtekamer, Z. Toth, G. Pellerin, M. Wei, Y. Zhu, 2004: Assessment of the Status of Global Ensemble Prediction. (Submitted to MWR).

Buizza, R. **2005**. The ECMWF Ensemble Prediction System. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Cohn, S. 1997: An introduction to estimation theory. *J. Meteor. Soc. Of Japan*, 75 (1B), 257-288.

Corazza, M., E. Kalnay, D. J. Patil, E. Ott, J. Yorke, I. Szunyogh, and M. Cai, 2002: Use of the breeding technique in the estimation of the background error covariance matrix for a quasi-geostrophic model. *AMS Symposium on Observations, Data Assimilation and Probabilistic Prediction*, pp 154-157, Orlando, Florida, 2002.

Corazza, M., E. Kalnay, D. J. Patil, R. Morss, I. Szunyogh, B. R. Hunt, E. Ott, and M. Cai, 2003: Use of the breeding technique to estimate the structure of the analysis "errors of the day". *Nonlinear Processes in Geophysics*, 10, 233-243.

Daley, R. 1991. *Atmospheric data analysis*. Cambridge University Press, Cambridge.

Dee D. P., and A. DaSilva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Royal Meteorol. Soc.*, 124, 269-295.

DelSole, T., and A. Y. Hou, 1999: Empirical Stochastic Models for the Dominant Climate Statistics of a General Circulation Model. *J. Atmos. Sci.*, **56**, 3436-3456.

Desroziers, G., G. Hello, and J.N. Thépaut, 2003 : A 4D-VAR re-analysis of the FASTEX experiment. *Quart. J. Roy. Meteor. Soc.*, 129, 1301-1316.

Dowell, D. C., Zhang, F., Wicker, L. J., Snyder, C. and Crook, N. A. 2004. Wind and thermodynamic retrievals in the 17 May 1981 Arcadia, OK supercell: Ensemble Kalman filter experiments. *Mon. Wea. Rev.*, **132**, 1982-2005

Ehrendorfer, M. and J. J. Tribbia, 1997: Optimal prediction of forecast error covariances through singular vectors. *J. Atmos. Sci.*, **54**, 286-313.

Ehrendorfer, M. **2005**. The Liouville Equation and Atmospheric Predictability. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Epstein, E. S., 1967: Stochastic-dynamic prediction. *Tellus*, 21, 739-759.

Evensen, G. 2003. The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics* **53** 343–367.

Fisher, M., L. Isaksen, M. Ehrendorfer, A. Beck and E. Andersson, 2003: A critical evaluation of the reduced-rank Kalman filter (RRKF) approach to flow-dependent cycling of background error covariances. ECMWF Tech. Memo.

Fritsch, J. M., J. Hilliker, J. Ross, and R. L. Vislocky, 2000: Model consensus. *Weather and Forecasting*, 15, 571-582.

Gauthier, P., 2004: THORPEX activities in Canada. Presented at THORPEX workshop, Boulder, CO, June 2004.
(http://box.mmm.ucar.edu/uswrp/recent_meetings/ThorpeXWorkshop/Canada_Gauthier.pdf)

Ghil, M., S. Cohn, J. Tavantzis, K. Bube, and E. Isaacson, 1981: Applications of estimation theory to numerical weather prediction. In *Dynamic Meteorology: Data Assimilation Methods*, L. Bengtsson, M. Ghil and E. Kallen, Editors. Springer-Verlag, New York.

Hagedorn, R., Doblas-Reyes, F.J., and Palmer, T.N. **2005**. DEMETER and the application of seasonal forecasts. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Hamill, Thomas M., Chris Snyder., 2000: A Hybrid Ensemble Kalman Filter-3D Variational Analysis Scheme. *Mon. Wea. Rev.*, **128**, 2905-2919.

Hamill, T.M. **2005**. Ensemble-based atmospheric data assimilation. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

- Hou, D., E. Kalnay, and K.K. Droegemeier, 2001: Objective verification of the SAMEX '98 ensemble forecasts. *Mon. Wea. Rev.* , 129, 73-91.
- Houtekamer, P.L., L. Lefavre and J. Derome, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, 124, 1225-1242.
- Houtekamer, P. L. and Mitchell, H. L. 2001. A sequential ensemble Kalman Filter for atmospheric data assimilation. *Mon. Wea. Rev.* **129**, 123–137.
- Houtekamer, P. L., Mitchell, H. L., Pellerin, G., Buehner, M., Charron, M. et al. 2004. Atmospheric data assimilation with the ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.* **132**, in press.
- Hunt, B. R., Kalnay, E., Kostelich, E. J., Ott, E. and Patil, D. J. et al. 2004. Four-dimensional Ensemble Kalman filtering *Tellus* **56A**, 273–277.
- Kaas, E., A. Guldberg, W. May and M. Decque., 1999: Using tendency errors to tune the parameterization of unresolved dynamical scale interactions in atmospheric general circulation models. *Tellus*, **51A**, 612-629.
- Kalnay, Eugenia and M. Ham., 1989: Forecasting forecast skill in the Southern Hemisphere. Preprints of the 3rd International Conference on Southern Hemisphere Meteorology and Oceanography, Buenos Aires, 13017 November 1989. Boston, MA: Amer. Meteor. Soc.
- Kalnay, E., 2003. *Atmospheric modeling, data assimilation, and predictability*. Cambridge University Press, Cambridge, 341 pages.
- Keppenne, C. and Rienecker, H. 1961. Initial testing of a massively parallel ensemble Kalman Filter with the Poseidon Isopycnal Ocean General Circulation Model. *Mon. Wea. Rev.* **130**, 2951–2965.
- Krishnamurti, TN, CM Kishtawal, Timothy E LaRow, David R Bachiochi, Zhan Zhang, Eric C Williford, Sulochana Gadgil, and Sajani Surendran (1999): Improved Weather and Seasonal Climate Forecasts from Multimodel Superensemble. *Science* 285 (5433), 1548-1550.
- Krishnamurti, T.N., Kumar, T.S.V.V., Mitra, A.K., Yun, W.T., Stefanova, L., Mackey, B.P., O'Shay, A.J., and Dewar, W.K. **2005**. Weather and Seasonal Climate Forecasts Using the Superensemble Approach. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.
- Kumar, T.S. V., Krishnamurti, T.N., M. Fiorino and M. Nagata, 2003. Numerical Prediction of Typhoon Tracks and Intensity using a Multimodel Superensemble. *Monthly Weather Review*, 131, 574-583.

Kyouda, M. and S. Kusunoki, 2002: Ensemble Prediction System: Outline of the Operational Numerical Weather Prediction at the Japan Meteorological Society. Appendix to the WMO Numerical Weather Prediction Progress Report, pp 59-63.

Lalurette, F., and van der Grijn, G. **2005**. Ensemble forecasts: can they provide useful early warnings? In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Latif, M., Pohlmann, H., and Park, W. **2005**. Predictability of the North Atlantic Thermohaline Circulation. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Lorenz, E. N., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**, 321-333.

Lorenz, E. N., 1996. Predictability: A problem partly solved. In *Proc. Seminar on Predictability, Vol. 1*. European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, Berkshire, RG2 9AX, United Kingdom.

Molteni, F. and T. N. Palmer, 1993: Predictability and finite time instability of the northern winter circulation. *Quart. J. Roy. Meteor. Soc.*, **119**, 269-298.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliajgis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73-119.

Morss, R. E., Emanuel, K. A., and Snyder, C.: Idealized Adaptive Observation Strategies for Improving Numerical Weather Prediction, *J. Atmos. Sci.*, **58**, 210–234, 2001.

Mylne, K. **2005**. Predictability from a Forecast Provider's Perspective. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Ott, E., Hunt, B. H., Szunyogh, I., Corazza, M. Kalnay, E. et al., 2002: Exploiting local low dimensionality of the atmospheric dynamics for efficient Kalman filtering. arXiv:archive/paper **0203058**, <http://arxiv.org/abs/physics/0203058>.

Ott, E., Hunt, B. H., Szunyogh, I., Zimin, A. V., Kostelich, E. J. et al. 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus* **56A** 415–428.

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

Palmer, T.N. **2005**. Predictability of Weather and Climate: From Theory to Practice. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Patil, D. J., Hunt, B. R., Kalnay, E., Yorke, J. A. and Ott, E. 2001. Local low dimensionality of atmospheric dynamics.. *Phys. Rev. Lett.* **86**, 5878-5881.

Peña, M. and E. Kalnay, 2004: Separating fast and slow modes in coupled chaotic systems. *Nonlinear Processes in Geophysics*, vol 19, 319-327.

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.

Roulston, M. S. and L. A. Smith, 2003: Combining dynamical and statistical ensembles. *Tellus*, **55A**, 16-30.

Shukla, J. and Kinter, J.L. **2005**. Predictability of Seasonal Climate Variations: A Pedagogical Review. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Simmons, A. J. and A. Hollingsworth, 2002: Some aspects of the improvement in skill of numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **128**, 647—678.

Simmons, A.J. **2005**. Observations, assimilation and the improvement of global weather prediction - Some results from operational forecasting and ERA-40. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Snyder, C. and Zhang, F. 2003 Assimilation of simulated Doppler radar observations with an Ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663-1677.

Stensrud, D. J., J.-W. Bao, and T. T. Warner, 2000: Using initial condition and model physics perturbations in short-range ensembles. *Mon. Wea. Rev.*, **128**, 2077–2107.

Szunyogh, I., and Toth, Z., 2002 The effect of increased horizontal resolution on the NCEP global ensemble mean forecasts *Mon. Wea. Rev.*, **130**, 1125-1143.

Szunyogh, Istvan, Eric J. Kostelich, G. Gyarmati, D. J. Patil, Brian R. Hunt, Eugenia Kalnay, Edward Ott And James A. Yorke, 2004: Assessing a local ensemble Kalman filter: Perfect model experiments with the NCEP global model. *Tellus*, in press.

Thorpe, A. and Petersen, G.N. **2005**. Predictability and Targeted Observations. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Tibaldi, S., Paccagnella, T., Marsigli, C., Montani, A., and Nerozzi, F. **2005**. Limited-Area Ensemble Forecasting: the COSMO-LEPS system. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Timmermann, A. and Jin, F.-F. **2005**. Predictability of Coupled Processes. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M. and Whitaker, J. S. 2002. Ensemble square-root filters. *Mon. Wea. Rev.*, **131**, 1485–1490.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, 74, 2317-2330.

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

Toth, Z., Talagrand, O., and Zhu, Y. **2005**. The attributes of forecast systems: A framework for the evaluation and calibration of weather forecasts. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Waliser, D.E. **2005**. Predictability of Tropical Intraseasonal Variability. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Wang, X., and C. H. Bishop, 2003: A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.*, 60, 1140-1158.

Wang, Xuguang and Craig H. Bishop, 2004: Ensemble Augmentation With A New Dressing Kernel. AMS, Seattle, January 2004.

Webster, P.J., Hopson, T., Hoyos, C., Subbiah, A., Chang, H.-R., and Grossman, R. **2005**. A three-tier overlapping prediction scheme: Tools for strategic and tactical decisions in the developing world. In Palmer, T.N. and Hagedorn, R., editors, *Predictability of Weather and Climate*. Cambridge University Press.

Whitaker, J. S. and Hamill, T. H. 2002. Ensemble Data Assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.

Whitaker, J. S., Compo, G. P., Wei, X. and Hamill, T. H. Hamill 2004. Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.*, **132**, 1190-1200.

Zhang, F., Snyder, C. and Sun, J. Impacts of initial estimate and observation availability on convective-scale data assimilation with an Ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238-1253.

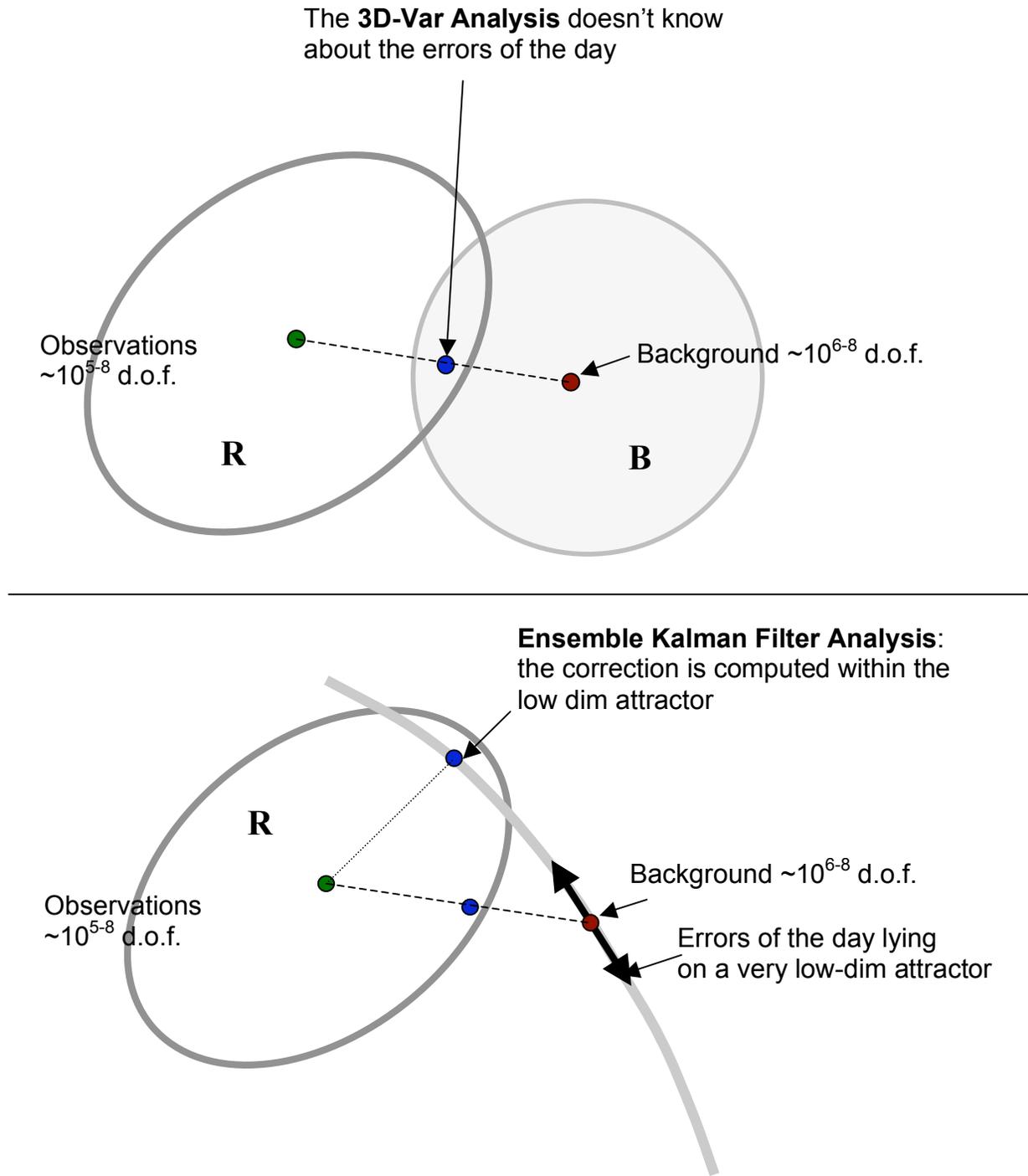
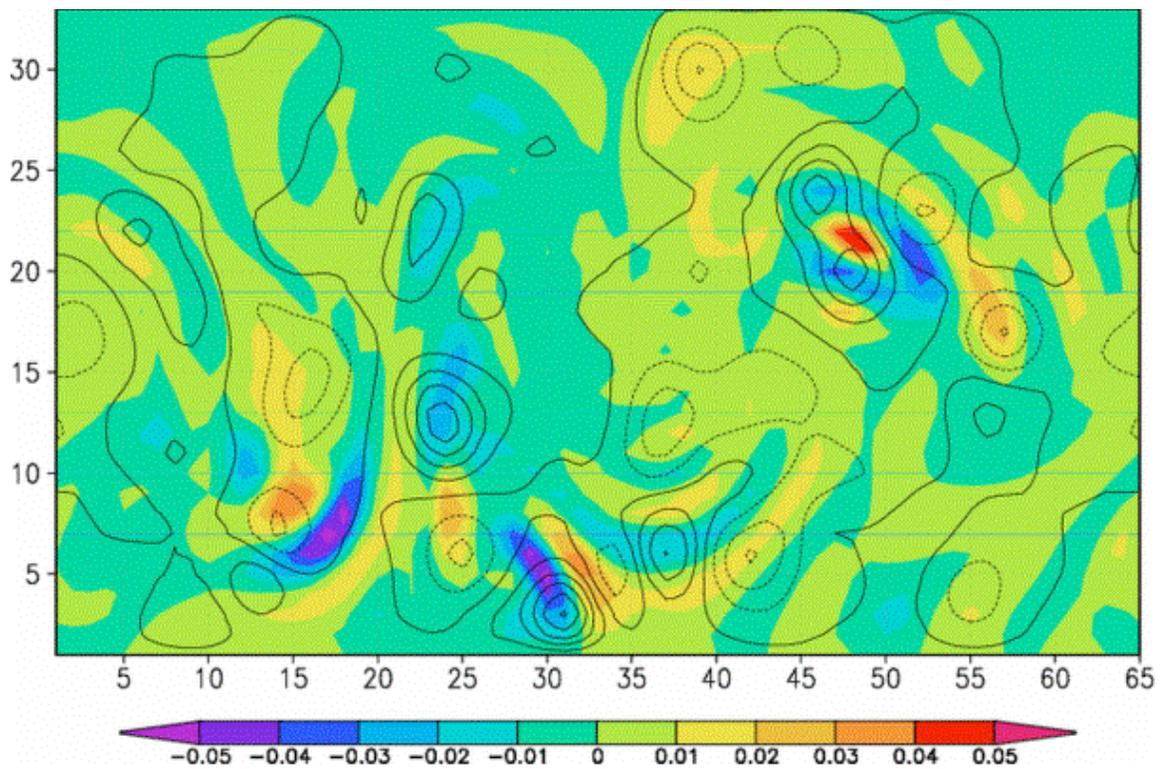


Figure 5: Schematic of the analysis given the background forecast in a very large dimensional space, the background error covariance **B** (which in the case of 3D-Var is isotropic and constant in time), the vector of observations, in a very large dimensional space, with observations error covariance **R**. The analysis estimate of the true state of the atmosphere maximizes the joint probability distribution. Top: 3DVar. Bottom: EnKF, in which the ensemble forecast members define a subspace within which the analysis lies.

Background error (shaded) and 3D-Var analysis increments



Background error (shaded) and LEKF analysis increments

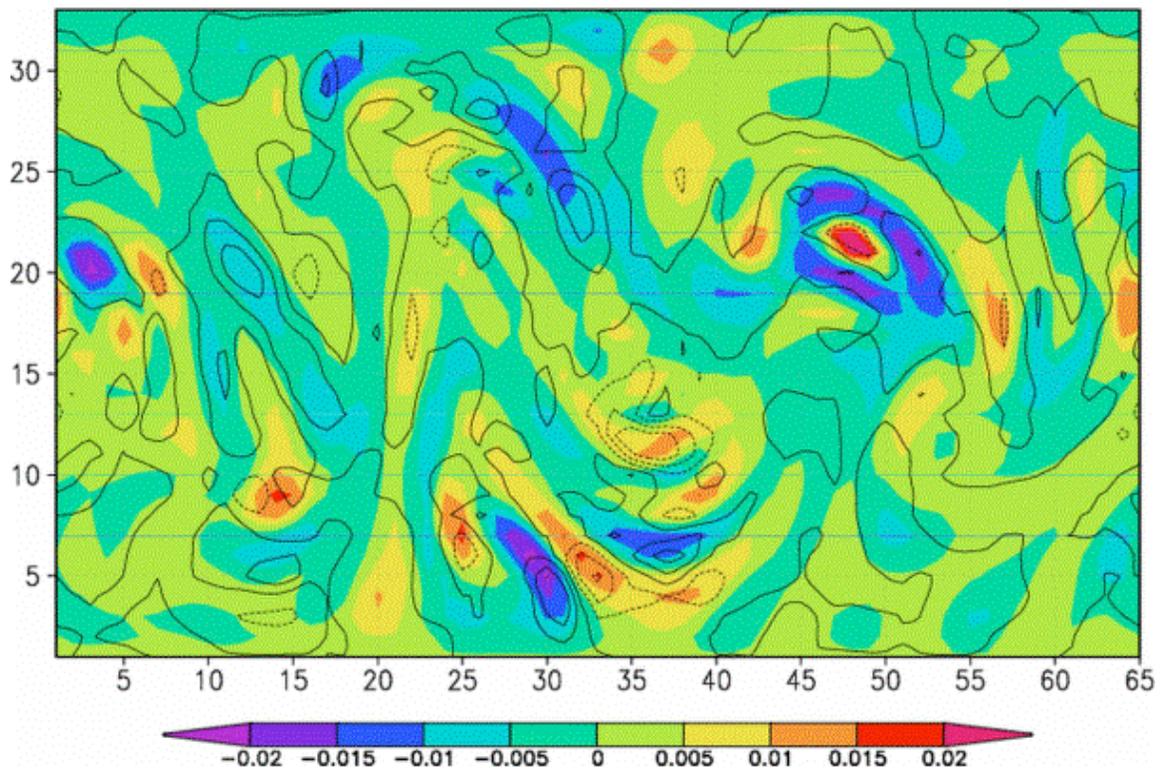
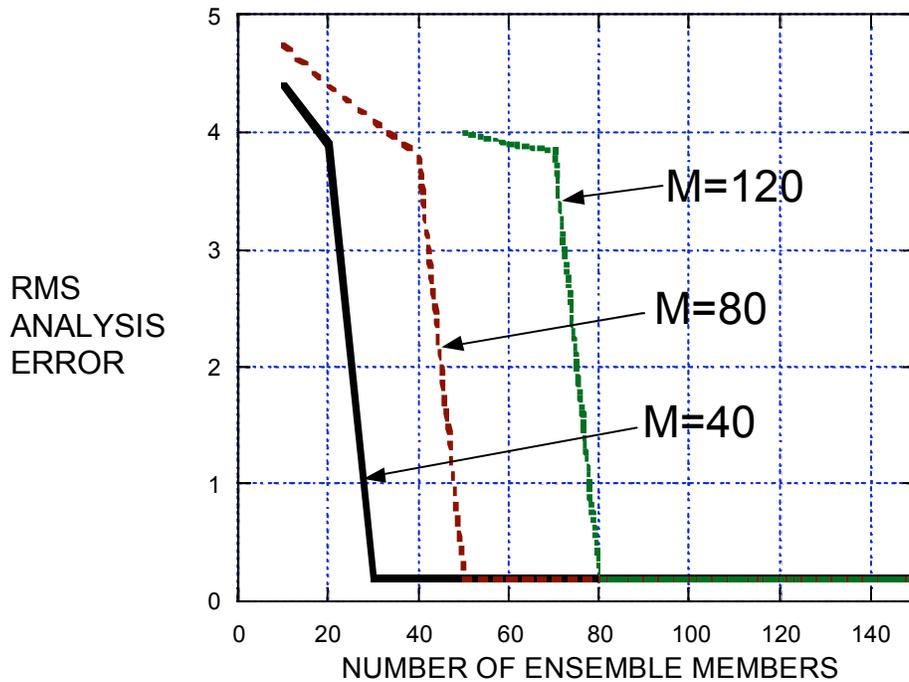


Fig. 6: Simulation of data assimilation in a quasi-geostrophic model, assimilating potential vorticity observations at a particular day (June 15). The shades represent the 12 hr forecast (background) error and the contours the analysis corrections. Top: 3D-Var. Bottom: Local Ensemble Kalman Filter. Figures courtesy of Matteo Corazza.

FULL ENSEMBLE KALMAN FILTER



LOCAL ENSEMBLE KALMAN FILTER

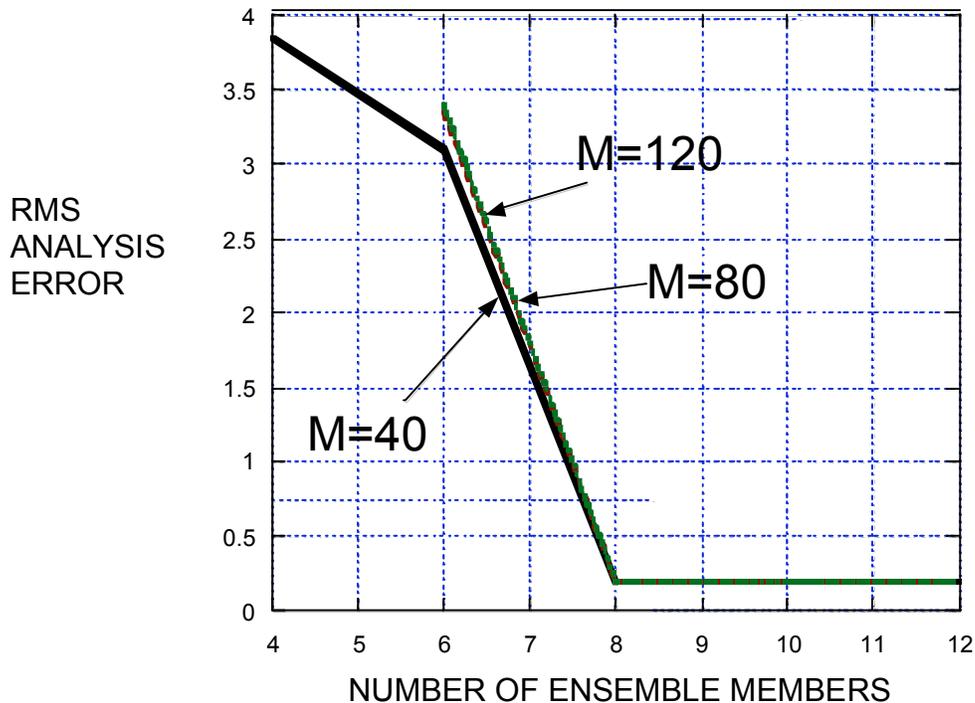


Fig. 7: Number of ensemble members required for convergence to the optimal solution in a Lorenz (1996) model. Top: using a full global Ensemble Kalman Filter. Bottom: using a Local Ensemble Kalman Filter. The size of the domain, M is either 40, 80 or 120. Note that the Kaplan-Yorke dimension is about 27 for the 40 variable model and increases linearly with size. Figure adapted from Ott et al (2004).

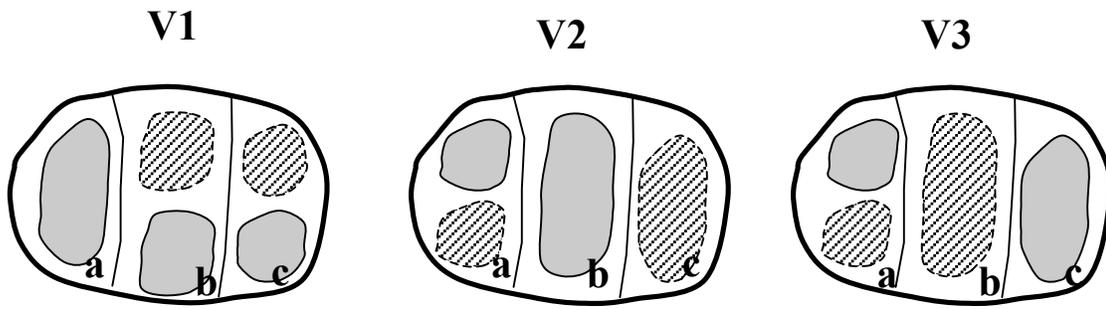


Fig. 8: Schematic showing the advantage of performing a local rather than a global analysis. The domain is composed of three regions, **a**, **b**, **c**, each of which has possible instabilities with wave numbers 1 and 2. From a local point of view, the ensemble perturbations **V1** and **V2** are sufficient to represent all possible unstable perturbations, whereas from a global point of view, **V3** is independent from **V2**, and there are many more independent perturbations.

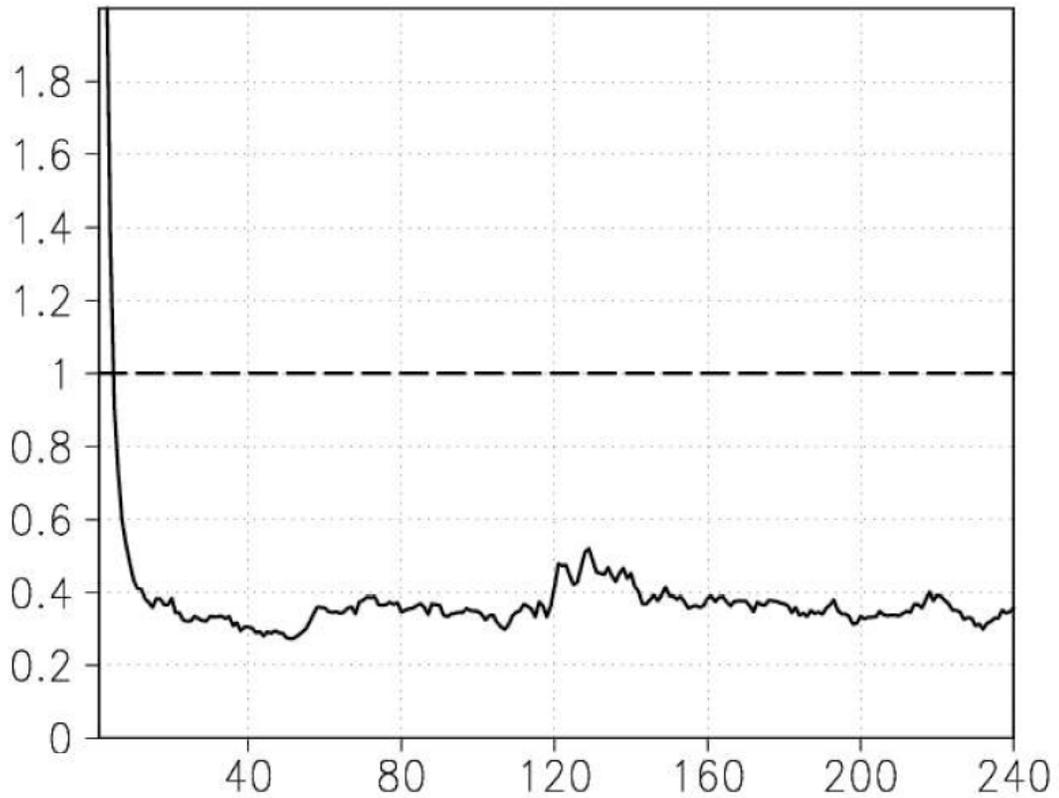


Fig. 9: Evolution of the LEKF analysis error in surface pressure in hPa as a function of assimilation step (in units of 6 hr). The rms error of the observations is shown by the dashed line. Observations are made at 11% of the grid points, and the model has T62 horizontal resolution (about 200km). (From Szunyogh et al., 2004).

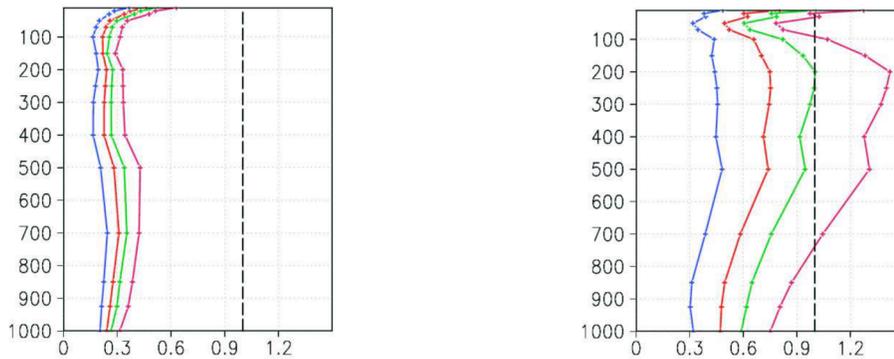


Fig. 10: RMS global analysis error for temperatures (left, °C) and tropical analysis error for zonal winds (right, m/sec). The dashed line is the rms of observations. From left to right, the following percentage of the grid points have “rawinsonde” data: 100% , 11%, 5%, 2%. Since the grid resolution is about 200km, the second is similar to the current rawinsonde density in the Northern Hemisphere, and the 4th to the Southern Hemisphere and tropics (From Szunyogh et al., 2004).

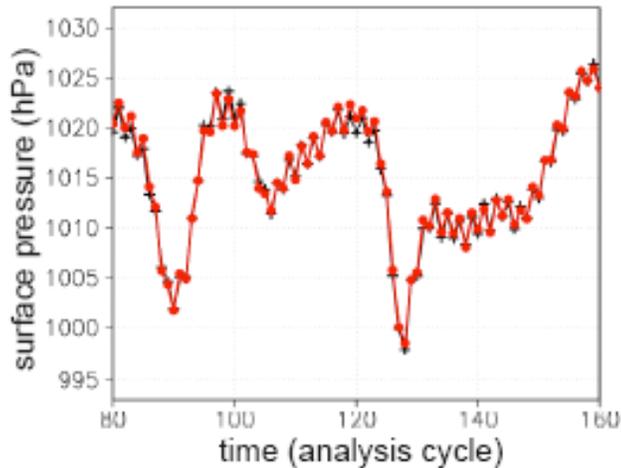


Fig. 11: Comparison of the true (crosses) and analyzed (circles) gravity wave observed at 30N 150W. The observing network has density similar to that of the SH. (From Szunyogh et al, 2004)

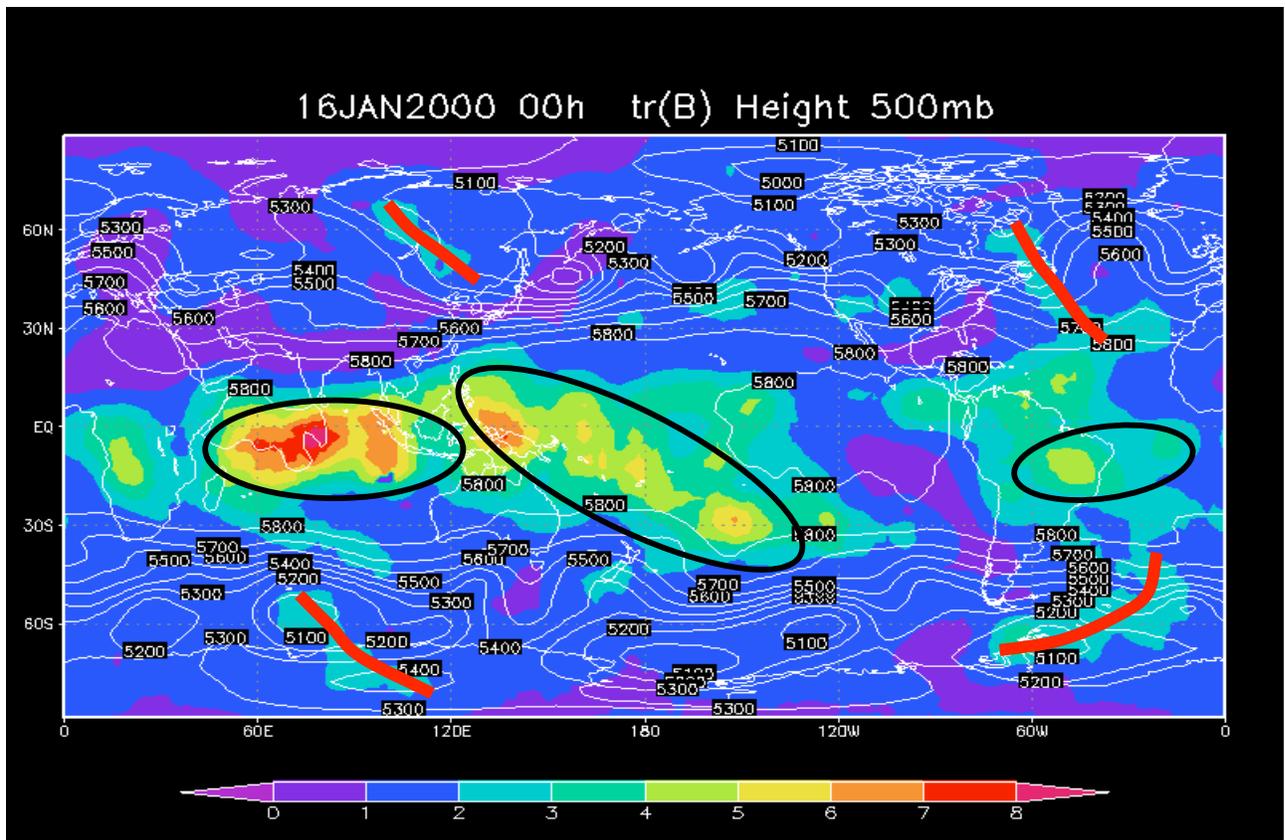


Fig. 12: Example of a 6 hr trace of the 500mb height forecast error covariance showing the potential use of LEKF for adaptive observations. Regions in blue and purple do not need immediate observations. Mid-latitude areas marked with red have large errors but a low effective ensemble dimension, so that they are prime areas for targeting. Tropical regions with large errors (ovals), on the other hand, have also large effective ensemble dimension presumably because the error growth is dominated by convection.

