Eugenia Kalnay and Zoltan Toth[1]

Environmental Modeling Center
National Centers for Environmental Prediction
NOAA/National Weather Service
Washington DC, USA

Summary: Methods to create perturbations for ensemble forecasting can be classified into two classes. Within the first class, denoted generically as Monte Carlo forecasting (MCF), are methods in which perturbations are chosen randomly, without regard to the "dynamics of the day". In the second class, which includes lagged averaged forecasting (LAF), short–range forecast differences (SRFD), combinations of singular vectors (SV), and breeding, the perturbations contain growing "errors of the day". Perturbations that contain "errors of the day" are found to be more effective for ensemble forecasting than randomly chosen perturbations. Breeding, used operationally at NMC since 1992, is a simple nonlinear perturbation method with periodic rescaling of the difference between two nonlinear integrations. Since breeding is a nonlinear generalization of the method used to obtain Lyapunov exponents and vectors, the bred growing vectors (BGVs) are closely related to the local Lyapunov vectors (LLVs), and share their lack of dependence on the choice of norm or on the period of rescaling. Nonlinear saturation, present in breeding, is important since it filters out irrelevant growing modes, such as convective modes, which have very fast linear growth but very small amplitudes. Because of this, breeding can also be used to generate the slow but high energy pertubations associated with the ENSO oscillations in a coupled ocean–atmosphere model, while at the same time filtering the faster but lower energy atmospheric weather perturbations.

The analysis cycle, in which a short–range forecast is statistically combined with observations to generate the next analysis, can be considered as a breeding cycle: The basic trajectory is the evolution of the real atmosphere, the short range forecasts are perturbed integrations, and the use of observations in the analysis acts to scale down the growing errors from the first guess. The strong similarity between breeding and the analysis cycle indicates that analysis errors contain Lyapunov vectors, and not just white noise, as often assumed. This conclusion is supported by examples in which breeding is used to improve the analysis, as well as by other suggestive evidences.

In addition to ensemble forecasting and improvement of the analysis, other applications of breeding include the detection and diagnosis of model instabilities. The existence of two new types of instabilities that occur close to the top of the atmospheric models, a form of equatorial Kelvin waves, and "upside down" baroclinic waves in the winter upper stratosphere, were found through breeding. A companion paper (Toth and Kalnay, 1995, this volume), compares the breeding method with the singular vectors method, and addresses in detail the application of breeding in ensemble prediction.

1. INTRODUCTION

In 1990 we started experimenting at the National Metorological Center (now NCEP) with several methods to generate efficient perturbations suitable for operational ensemble forecasting. We tested three methods that were available at that time: 1) Monte Carlo Forecasting (Leith, 1974; Kalnay and Toth, 1992a), or MCF, with realistic random perturbations defined as balanced perturbations spanning the space of observed daily analysis anomalies; 2) Lagged Averaged Forecasts (Hoffman and Kalnay, 1983), or LAF, where the perturbations are given by the apparent short range forecast errors (Fig. 1); 3) Scaled LAF (Ebisuzaki and Kalnay, 1991), or SLAF, where the LAF perturbations are scaled by their "age", and are applied both with positive and negative signs. Since it was apparent from the estimated short range forecast errors that the analyses contain significant random errors, we also tried a new method, 4) Short–range Forecast Differences (Kalnay and Toth, 1992b), or SRFD, which cancel some of the random errors present in the initial analysis (Fig. 1).

1. General Sciences Corporation (Laurel, MD) at NCEP

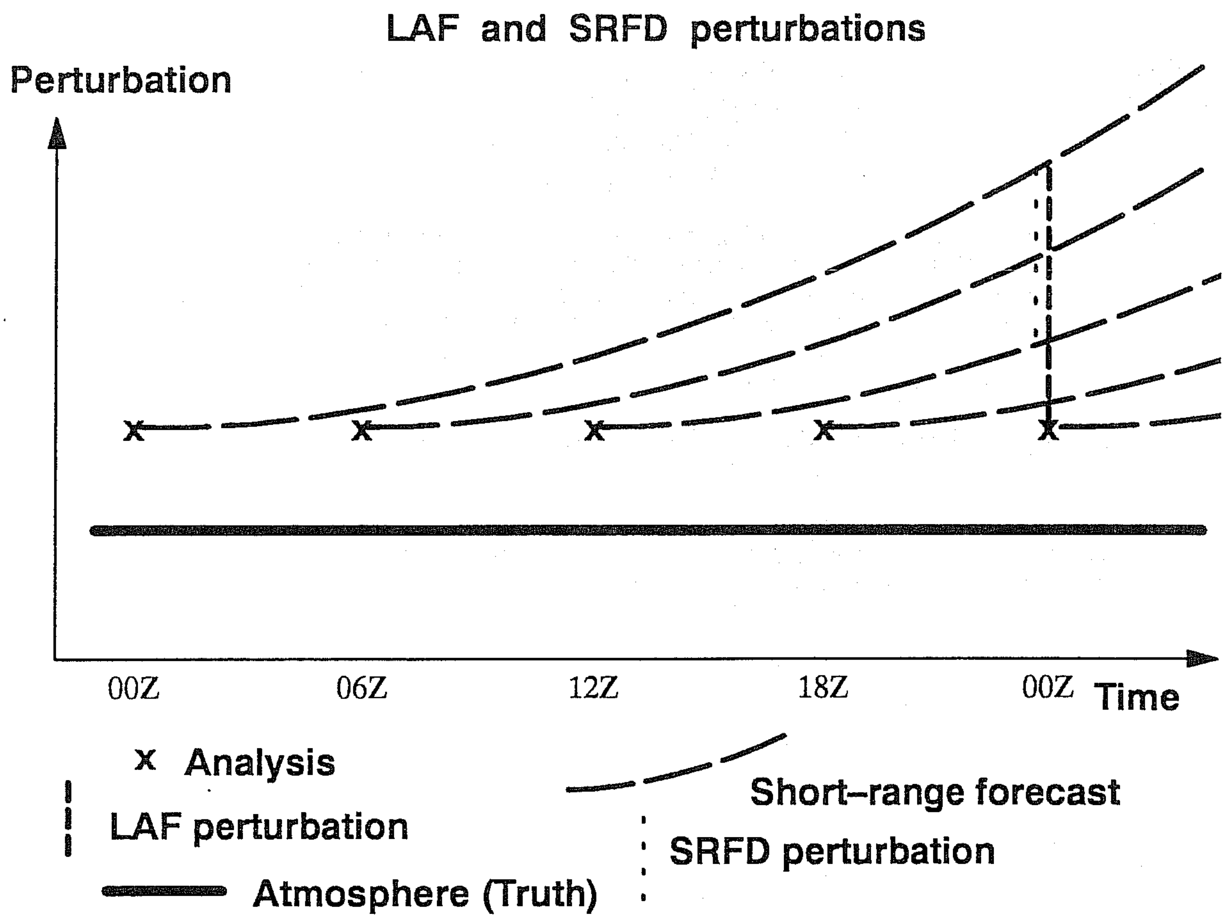

Fig. 1: Schematic of the creation of LAF and SRFD perturbations. Note that the LAF perturbation includes not only the short–range forecast errors but also the random errors of the latest analysis whereas the SRFD perturbation is not affected by the random error of the latest analysis. This results in a significant reduction of the random errors and therefore in a higher growth rate for the SRFD perturbations (from Toth and Kalnay, 1993).

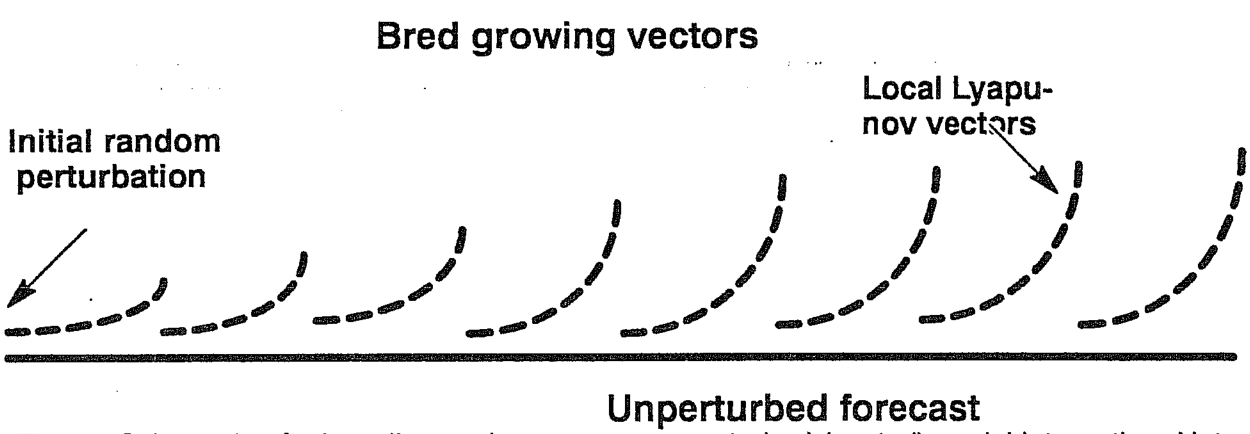

Fig. 2: Schematic of a breeding cycle run on an unperturbed (control) model integration. Note that after introducing an arbitrary initial perturbation, the growth rate of the perturbation increases. This is because decaying and neutral components in the perturbation field will loose their amplitude relative to the fast growing components since the whole perturbation field is scaled back periodically by the same factor. In an operational NWP environment, the unperturbed model integration can be substituted by short range control forecasts started from consecutive atmospheric analysis fields.

We tested the efficiency of the perturbations by keeping their initial RMS amplitude constant and measuring their growth rate, which is related to the improvement that the ensemble average yields when compared to the control forecast.

The results of these experiments were that SRFD was somewhat better than SLAF, which in turn was somewhat better than LAF. The most important difference, however, was that these three types of perturbations had a larger growth rate (and in turn a better ensemble average) than the MCF perturbations. It was clear that what made MCF the worst method was the fact that the MC perturbations were the only ones chosen randomly, without regard to the "dynamics of the day". By contrast, the SRFD, SLAF and LAF initial perturbations contain, by construction, "growing errors of the day", i.e., perturbations that will grow given the present evolving synoptic and large–scale flow.

At that point we started considering whether it was possible to design a special operational cycle, like the analysis cycle, with the purpose of "breeding" fast growing perturbations (Toth and Kalnay, 1992, 1993). Once this problem was posed, it was simple to design such a "breeding cycle" (Fig. 2): Given an evolving atmospheric flow (either a series of atmospheric analyses or a long model run), a breeding cycle is started by introducing a random initial perturbation with a given initial amplitude. It should be noted that the random seed is introduced only once. The same nonlinear model is integrated from the control and the perturbed initial conditions. From then on, at fixed time intervals (e.g., every 24 hours), the control forecast is subtracted from the perturbed forecast. The difference is then scaled down so that it has the same amplitude as the initial perturbation, and added to the corresponding new analysis or model state. We found that after an initial transient period of 3–4 days, the perturbations generated in the breeding cycle (denoted bred growing vectors, or BGVs) acquired a large growth rate, faster than the growth rate of any of the other dynamical perturbations discussed above.

We did extensive experiments with the BGVs and found that a) BGV perturbations performed better than SRFD, SLAF or MC perturbations; b) the characteristics of the BGVs were essentially independent of the choice of norm used to measure the amplitude of the perturbations, and also independent of the time interval (ranging from 6 hours to 2–3 days) after which the perturbations were scaled down.

The breeding cycle was introduced to generate one pair of bred vectors every day for ensemble forecasting at NMC in December 1992 (Toth and Kalnay, 1993; Tracton and Kalnay, 1993). Starting in March 1994, the configuration was changed to 7 independent breeding cycles run at T62/28 level resolution every day (Toth and Kalnay, 1995a; Tracton, 1994), leading to an ensemble of 17 forecasts integrated to 16 days each (Fig. 3). Note that in this configuration each pair of forecasts maintains its own breeding cycle, so that the initial perturbations are generated at *no cost* beyond that of running the forecast ensemble.

In the rest of the paper we first discuss the relationship between bred growing vectors (BGVs) and local Lyapunov vectors (LLVs), with an emphasis on the role of the nonlinear saturation in the generation of bred perturbations (section 2). We then turn to the relationship between the analysis cycle and the breeding cycle, and give examples of applications of BGVs to improve the analysis (section 3). Other applications of breeding such as diagnosis of model instabilities follow in section 4 while the conclusions are found in section 5. The properties of the breeding method and the BGVs are also summarized in Table 1.

2. LOCAL LYAPUNOV VECTORS AND THE BREEDING CYCLE

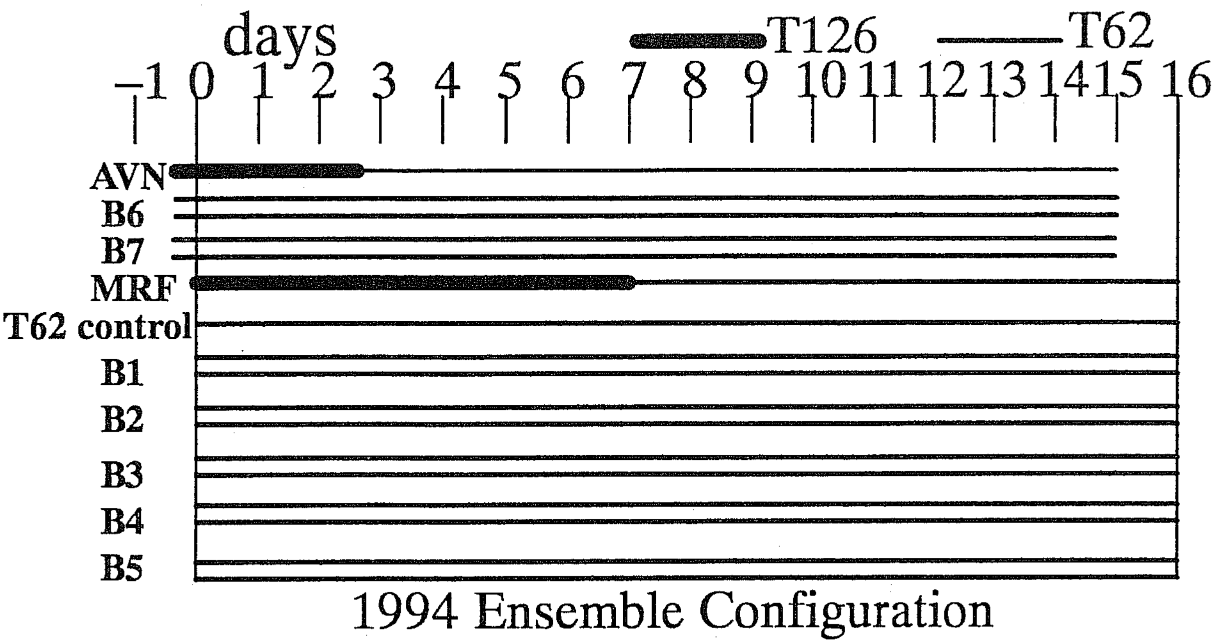

Fig. 3: Schematic of the configuration of the operational ensemble forecasting system at NCEP. Each horizontal line represents a numerical forecast. High resolution, T126 forecasts are marked with heavy lines while the other forecasts are run at T62 resolution. Note that at 00Z there are two control forecasts, one started at T126 resolution and then truncated to T62 at day 7, while the other started at a T62 truncated resolution. At 12Z, the high resolution control is truncated after 3 days of integration. Pairs of perturbed forecasts based on the breeding method are marked as B1–B7. (From Toth and Kalnay, 1996.)

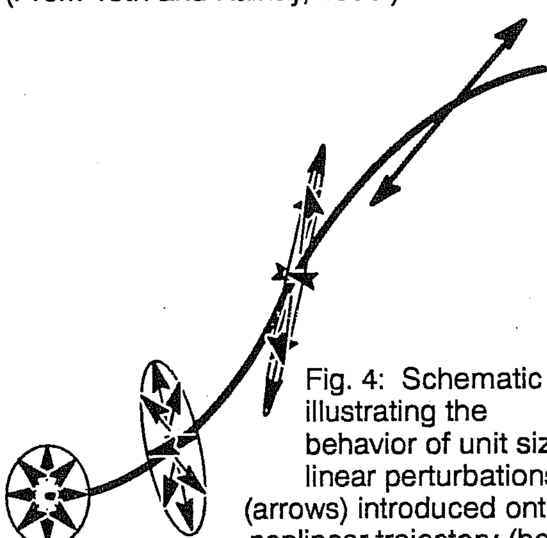

Fig. 4: Schematic illustrating the behavior of unit size linear perturbations (arrows) introduced onto a nonlinear trajectory (heavy line). Since in a general case growth is different in each direction in the phase space, the hypersphere representing the size of perturbations at initial time will transform into a hyperellipsoid.After some time (which is a characteristic of the system studied) all perturbations will practically collapse into one direction, which defines the first local Lyapunov vector (LLV) of the system. Note that the LLV changes with the location in the phase space.

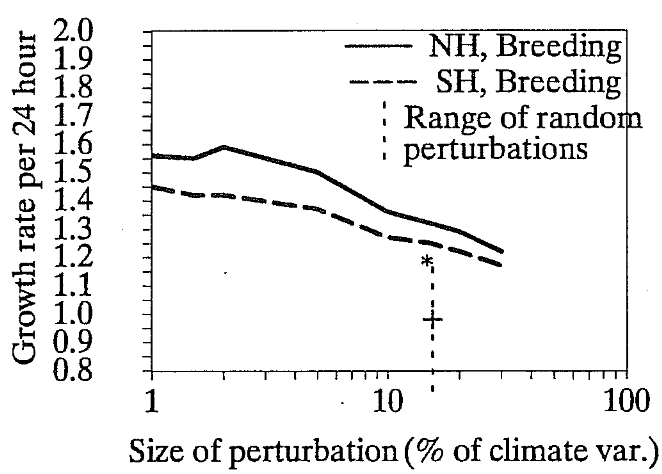

Fig. 5: Daily amplification of bred perturbations with different initial perturbation sizes over the Northern (solid) and Southern (dashed) hemisphere extratropics, computed for the period February 23–27, 1992. The range of amplification factors for different random (Monte Carlo) balanced perturbations is shown as a vertical dotted line. Average amplification factors for difference fields between different long short–range forecasts verifying at the initial time of perturbed forecast integrations are also shown with a star (NH) and a plus sign (SH). From Toth and Kalnay (1993).

Local Lyapunov vectors (LLVs, Trevisan and Legnani, 1995) are the perturbations that have the fastest *sustainable* growth in a dynamical system. The leading Lyapunov exponent (Tsonis, 1992) and along with it, the leading LLV is obtained by integrating random initial perturbations with a linear tangent model (Fig. 4). Further LLVs can be obtained by successive orthogonalization (e. g., Wolf et al.,, 1985). Lorenz (1965) was the first one to point out the fundamental role that LLVs play in atmospheric predictability when he observed in a low order atmospheric model that all linear perturbations tended to converge after a few days.

It is clear that there is a very close relationship between the method used to generate LLVs and the breeding method. In fact, if we choose infinitesimally small amplitudes for the bred perturbations, the two methods become identical. The fact that the BGVs are largely insensitive to the choice of norms, or to the time interval between scalings, is then naturally explained by the fact that the first LLV itself does not depend on these choices.

In breeding, therefore, there is essentially only one free parameter, the amplitude of the perturbations. This parameter is very important: we observed in our experiments with the NMC global forecast model (Toth and Kalnay, 1992; 1995a) that when we used perturbations with amplitudes between 1 and 5% of the RMS natural variability of the atmosphere (about 1 to 5 m in the 500 hPa geopotential height), the global growth rate of the BGVs was rather constant, about 1.5–1.6 per day; and when we increased these amplitudes to 10% or larger, nonlinear saturation slowly reduced the growth rate (Fig. 5). The bred vectors obtained using these amplitudes were clearly associated with mid–latitude baroclinic instabilities. However, when we reduced the amplitude below 0.1% of the natural variability, we observed a huge increase in the growth rate, from 1.6 per day to about 5–6 per day. The growing modes associated with this very large growth rate appeared on small scales and were intensive in the tropics, indicating that with these very small amplitudes, the dominant growing modes were associated with convection. This is not surprising, since convective instabilities are much faster than baroclinic instabilities, but they saturate at a much lower energy level than baroclinic instabilities (Fig. 6). Since the estimated analysis errors are of the order of 5–10% of the atmospheric variability, convective instabilities, which saturate at a level of much less than 1%, are not a dominant component of the analysis errors. Baroclinic and possibly barotropic instabilities are the relevant mechanisms for the creation of growing errors in the analysis cycle (Toth and Kalnay, 1993).

In summary, the BGVs are obtained by a nonlinear generalization of the method to construct LLVs. This constitutes a major advantage of the breeding method, because it allows the simple use of nonlinear saturation to eliminate those fast growing perturbations that would dominate in a linear tangent model approach, but which due to their low level of total energy, would be irrelevant for ensemble forecasting. It should be noted that convection is not the only example of fast but irrelevant perturbations that can be filtered out with the breeding method. Another important example is the use of breeding to create perturbations in a coupled ocean–atmosphere system for ensemble seasonal and interannual forecasting. In this case the relevant modes that we want to capture for ensemble forecasting are the coupled ocean–atmosphere modes responsible for the El Niño/Southern Oscillation (ENSO) phenomena (Zebiak and Cane, 1987; Schopf and Suarez, 1988), which have periods of recurrence of the order of several years, and a very high level of energy. At the same time, the coupled system also contains atmospheric and oceanic disturbances like baroclinic storms, that take place with much shorter time scales, of the order of a few days, and which compared to the ENSO oscillations, have much less energy. For this type of process, a judicious use of breeding using coupled ocean–atmosphere models, would also allow the generation of relevant ENSO perturbations, while filtering out the faster but lower–energy atmospheric storms (Toth and Kalnay, 1995b).

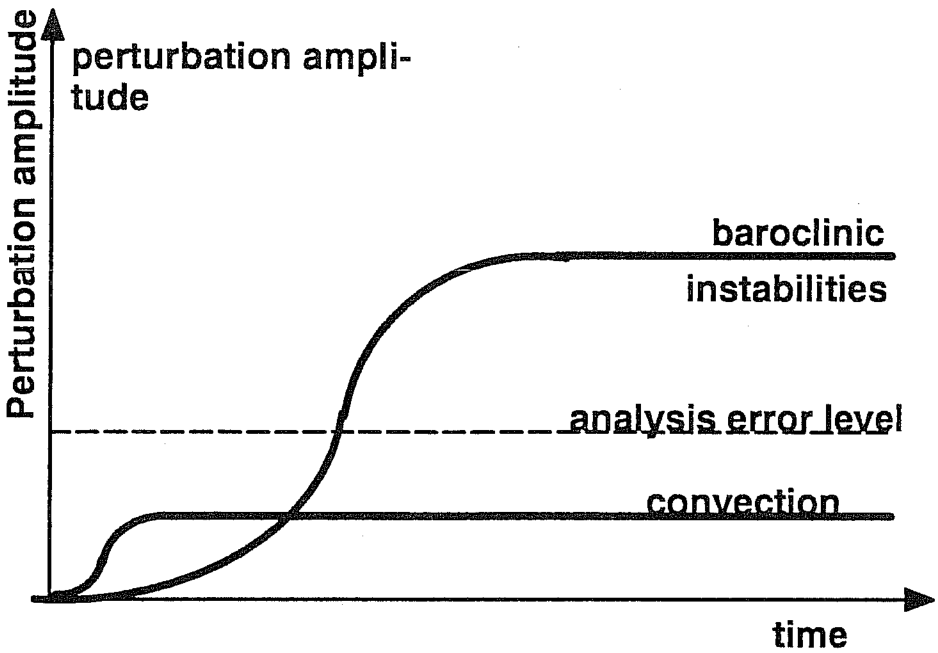

Fig. 6: Schematic of the time evolution of the rms amplitude of high–energy baroclinic modes and low–energy convective modes. Note that although initially growing much faster than the baroclinic modes, convective modes saturate at a substantially lower level. These modes are therefore insignificant for the analysis/ensemble perturbation problem since the errors in the analysis (dashed line) are much larger than the convective saturation level (from Toth and Kalnay, 1993).

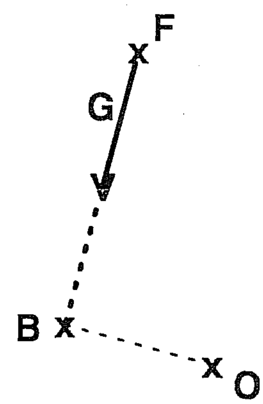

Fig. 8: Schematic for the adjustment of the first guess F towards the observations **O** in the direction of the growing vector **G**.

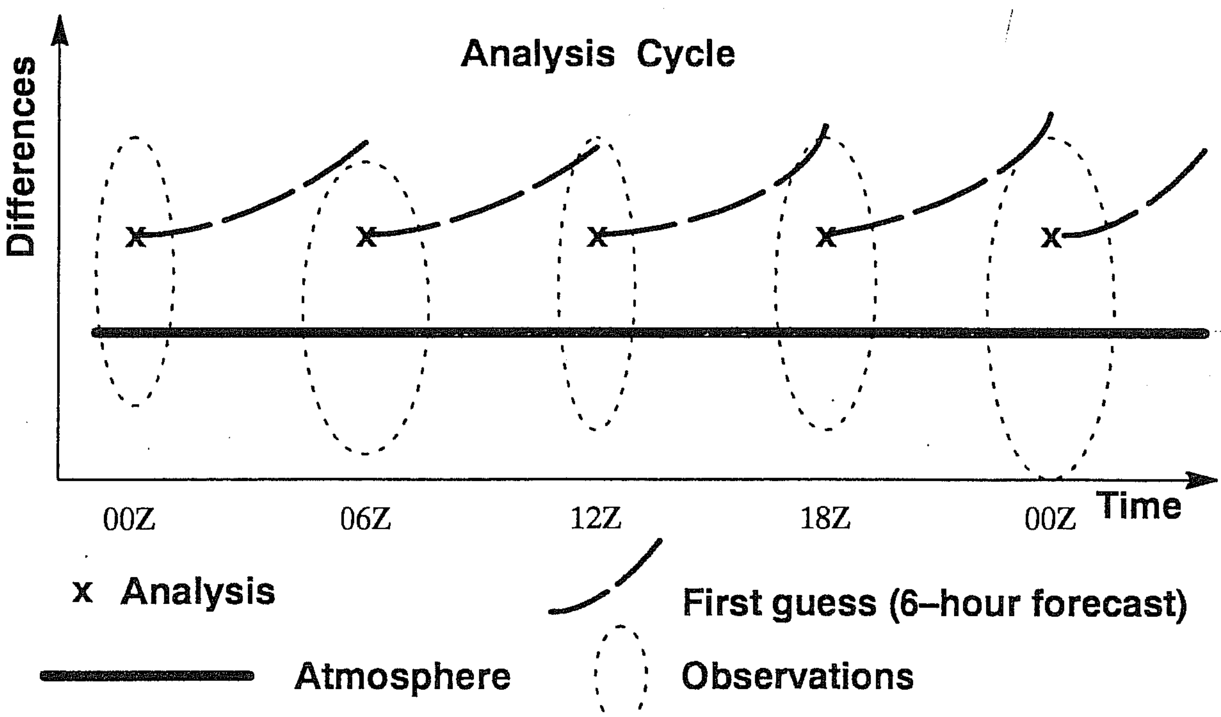

Fig. 7: Schematic of the 6–hour analysis cycle. Indicated on the vertical axis are differences between the true state of the atmosphere (or its observational measurements, burdened with random errors) and the analysis or forecasts of it. Note that the difference between a forecast and the true state of the atmosphere (or the observations) increases with time, due to the growing type of errors in the initial analysis (from Toth and Kalnay, 1993).

Lorenz (1995, this volume) has also presented in this meeting a very clear example of a simple model with coupled "synoptic" and "convective" modes, and has shown that if the linear tangent model contains the "convective" physics, the LLVs are dominated by these irrelevant but fast growing modes. His results agree with ours in that the use of nonlinear perturbations in breeding has the advantage of filtering out these irrelevant modes by nonlinear saturation, without having to make any explicit modification to the model. It should be noted that in the method of singular vectors, the irrelevant convective modes are also filtered out by the simple fact that the linear tangent model and its adjoint does not yet contain the physics of convection (Buizza, 1994). If the full model physical parameterizations were included, the singular vectors would also be dominated by convective instabilities.

3. BREEDING AND THE ANALYSIS CYCLE

A fundamental question for ensemble forecasting is which type of growing perturbations are present in the analysis errors. If the analysis errors contained only random noise, then the dominant forecast errors for each forecast lead time would be given by the singular vectors optimized for that particular time interval (Lorenz, 1965; Molteni and Palmer, 1993; Buizza et al., 1993). This may be a good assumption under the following conditions: a) if the analysis is performed with a "cold start", using climatology or persistence as the first guess, or b) if we had a perfect data assimilation system. In current data assimilation systems, we do not perform "cold start analyses", and instead use analysis cycles, and they are certainly not optimal. In the analysis cycle, the first guess is provided by a short–range forecast (e.g., 6 hours) from the previous analysis, and this first guess is statistically combined with the data gathered in the following 6 hour interval (Fig. 7) to create the new analysis. This procedure is clearly similar to a breeding cycle with respect to the true evolution of the atmosphere: The analysis errors act as initial perturbations, the 6–hour forecast corresponds to the perturbed integration, and the statistical combination of the first guess and the data corresponds to the periodic scaling down (cf. Figs. 7 and 2). This implies that the analysis cycle errors will inevitably contain LLVs, since any random errors generated by observations or deficiencies in the analysis procedure will quickly evolve into LLVs.

Since the analysis statistics (forecast error covariance) do not currently include information about the "growing errors of the day", it is impossible for the analysis scheme to zero out the LLVs that have been generated in previous analyses, and propagated forward through the first guess. New random errors introduced by observations will thus continuously generate new LLVs, which continue to propagate through the first guess, and cannot be completely eliminated by the new observations. The LLVs, in fact, act as a "magnet" for the random errors. It should be noted that, by contrast, singular vectors act as "repellent" for random errors, since any random perturbation that happens to project strongly on a singular vector will quickly move away from the singular vector and towards an LLV.

There are other indications that analyses contain well organized LLVs within the analysis errors, and not just random white noise. For example, we have observed that the differences between analyses from two different operational centers frequently contain patterns closely resembling the BGVs obtained from one or more breeding cycles. In addition, the pattern of the perceived forecast errors (differences between a forecast and the verifying analysis) usually contain very similar patterns that survive for several days. In other words, the 2, 3,...5–day forecasts verifying at the same time have errors that frequently are remarkably similar in shape, but with amplitudes which increase with the length of the forecast. This indicates that the same growing error pattern has at least partially survived the analysis cycle for many analyses (performed 4 times per day), and are carried forward and evolved by the six hour forecast. Such long lasting error patterns are, by definition, LLVs.

It seems natural to take advantage of the similarity between the analysis cycle and a breeding cycle, and of the fact that the same LLVs that are generated in the breeding cycle must be present as well in the analysis errors. In this section we present two examples of applications of breeding to data assimilation, and discuss a third approach. The first two examples (Kalnay and Toth, 1994) are based on the idea that, once we know the shape of the bred growing modes (LLVs), we can improve the first guess **F** in the analysis cycle by removing the projection of the difference between the first guess and the data along the growing errors. The advantage of this method is that the minimization is done along a single degree of freedom which is known (the growing mode **G**), and therefore it is simple to perform and computationally inexpensive. From the schematic Fig. 8, we can see that the improved first guess **B** is given by $\mathbf{B}=\mathbf{F}+\mu\mathbf{G}$, and that in order to minimize the distance **B–D** between the best first guess and the data, we simply need to require that the distance vector be normal to **G**. From this we obtain the amplitude $\mu=[\mathbf{D}-\mathbf{F},\mathbf{G}]/[\mathbf{G},\mathbf{G}]$. If there are a few BGVs available, as is the case at NCEP with 7 breeding cycles, the formulation can be easily generalized (Purser et al., 1994).

In the first example, we performed a simulated analysis cycle in a 3–variable model (Lorenz, 1963), using the same set up as in Miller et al (1994), except that we used "observations" obtained from the "nature run" for 2 variables (X and Z) rather than for all three variables as in Miller et al. We first tested an extended Kalman filtering (EKF) data assimilation. As was the case in Miller et al, the EKF failed to reproduce the evolution of the "nature" run because during periods of strong growth the linear model became a poor approximation of the nonlinear evolution (Fig. 9a). By contrast, a much simpler and less expensive correction to the first guess using the BGVs as described above, followed by a simple Optimal Interpolation analysis resulted in an analysis very close to nature (Fig. 9b).

In a second example, we used the full NCEP analysis/forecast system at T62 horizontal resolution, and performed a correction to the first guess within sliding windows of 10^o by 10^o, to allow for the fact that the BGVs are regional in character. We first computed the projection amplitudes μ using rawinsonde winds as data, and the distance of the data to both the control first guess and analysis. We then performed a first guess improvement as indicated above, minimizing the distance between the first guess and the rawinsonde winds in the Northern Hemisphere, interpolating μ linearly in between the 10^o by 10^o windows. As in the 3–variable model, we followed this correction of the first guess by a regular analysis (SSI, a spectral 3–D VAR, Parrish and Derber, 1992) using the improved first guess. The experiments were performed for a 44 days long period in April–June, 1992, and the results indicated that the experiment was better than the control (the regular SSI analysis) in both the first guess, which was closer to the observations (even before the last adjustment correction was introduced) in 71 cases and worse in only 7 cases, and in the 5–day forecasts, which had higher anomaly correlations in 24 cases and worse anomaly correlations in 10 cases. Purser et al. (1994) generalized this simple approach to improving the first guess to allow for multiple BGVs, include different observational errors and use a much better space filter than the sliding window. We hope to test the general approach in the near future.

A third approach to use breeding to improve the analysis cycle is to include the outer product of the BGVs within the forecast error covariance matrix (Parrish, 1992, pers. comm.). This approach would take into account the growing "errors of the day" within the statistical interpolation, instead of assuming constant statistics, as is the current operational practice in all centers. Because of the low number of degrees of freedom involved in this approach, it would also be quite efficient,

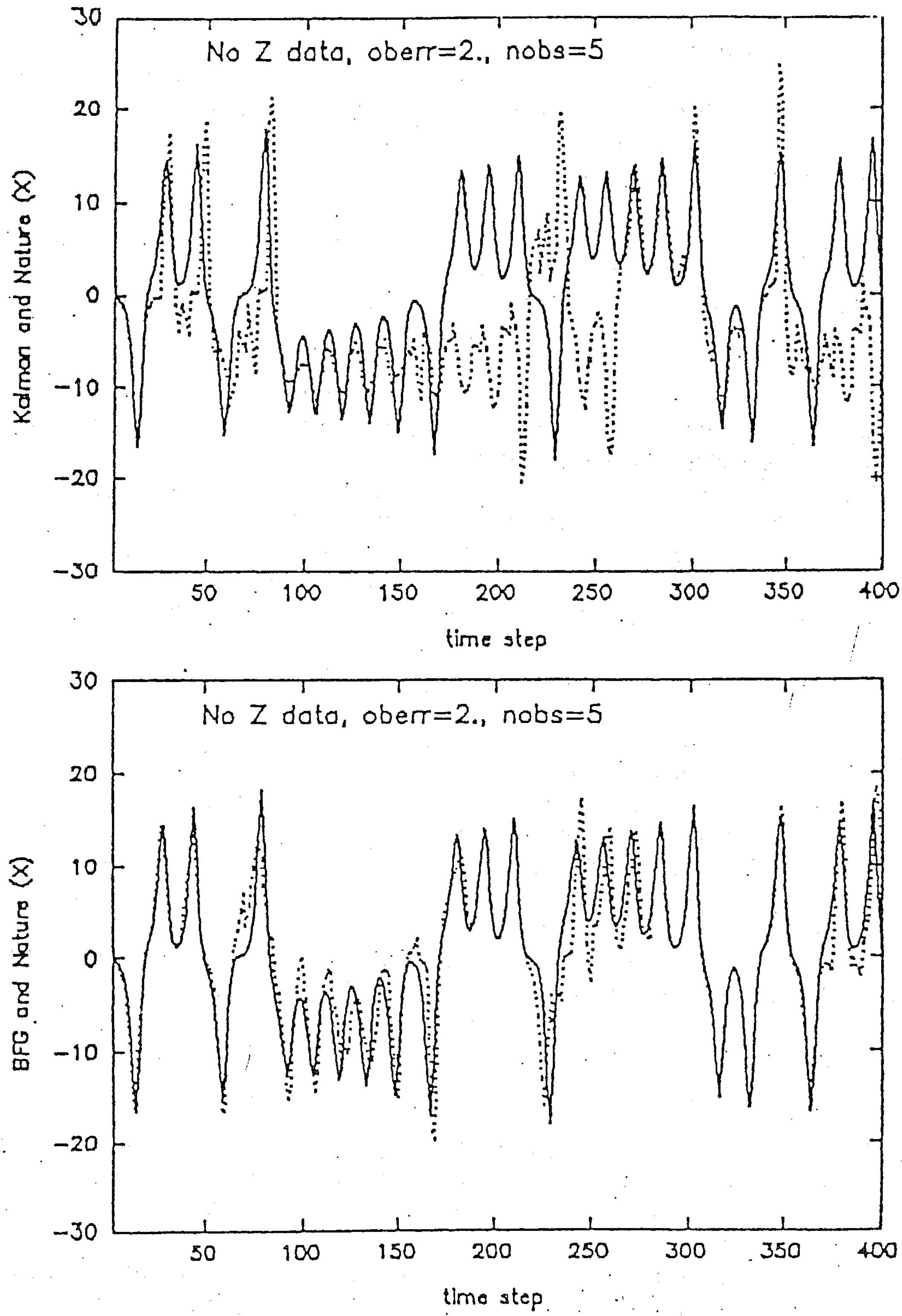

Fig. 9: Comparison between the true evolution of Y (solid line) and that estimated by an analysis (dotted line). Top: Kalman filtering. Bottom: Adjusted First Guess followed by a simple OI.

requiring very little additional computational cost in the analysis cycle.

4. APPLICATIONS TO ATMOSPHERIC DYNAMICS

Because the BGVs are obtained from a simple difference between nonlinear integrations with periodic rescaling to avoid excessively large perturbations, the resulting growing vectors are those that grow within the unstable basic flow, without any approximation or *a priori* assumption. We can take advantage of this set to observe or diagnose the characteristics of atmospheric instabilities (or more precisely model instabilities). We give here some examples.

When we started growing vectors through a breeding cycle, we noticed a maximum of kinetic energy in the upper winter stratosphere, (Toth and Kalnay, 1993, last paragraph), which we mistakenly attributed to planetary waves transmitted through the Charney–Drazin mechanism. However, when in 1993 the vertical resolution of the NMC model was increased from 18 to 28 levels, and the model top raised from 50hPa to 3hPa, the bred perturbations in the winter stratosphere were observed to move up with the top of the model, and their growth rate increased considerably. They are apparent in the winter stratosphere near the top of the model (Fig. 10a), have horizontal wavelengths of wavenumber typically 4–6 (Fig. 10b), and seem to be due to "baroclinic instability upside down" induced by the presence of the model's rigid lid boundary condition. Similar instabilities were found afterwards by Hartmann and Palmer (1995) in the ECMWF model using singular vectors. Hartmann and Palmer, however, do not attribute these instabilities mainly to the presence of the model top boundary condition but rather interpret them as genuine instabilities of the atmosphere. In addition to the stratospheric winter perturbations, we also observe during the transition seasons equatorial growing modes in the stratosphere, which may be unstable Kelvin waves (Fig. 10c and 10d).

It would seem advantageous to use bred vectors to diagnose sensitivity studies within general circulation model experiments such as runs with double CO_2. For example, at the cost of running the model a second time, it is simple to generate a breeding cycle alongside each long integration, and determine the change in the basic stability characteristics (such as growth rate of baroclinic instabilities) of the new atmospheric climate. Similarly, breeding can be easily used to diagnose the evolution of the atmosphere in specific case studies, such as in the great East Coast storm of March 1993, where the origin of the disturbance can be clearly traced back to the Gulf of Alaska through the bred modes. The characteristics of down–stream development of baroclinic modes pointed out by Orlanski and Sheldon (1993) are easily observable in the BGVs.

5. CONCLUSIONS

The breeding method is a nonlinear extension of the technique to estimate the local Lyapunov vectors (LLVs) of the atmosphere. It follows that the bred growing vectors (BGVs) practically do not depend on the choice of norm or renormalization period used in the breeding cycle. We emphasized the nonlinear aspect of breeding. Since the atmosphere is a complex system with motions of many different scales and types it is important to represent perturbation development with realistic amplitudes. Indeed, the breeding method has only one free parameter which is perturbation amplitude. This makes it possible to estimate the fastest sustainable growth connected to different physical processes, such as convection or baroclinicity by adjusting perturbation size appropriately. Since processes with faster perturbation growth typically have smaller scales and saturation levels, by choosing a higher perturbation amplitude we can

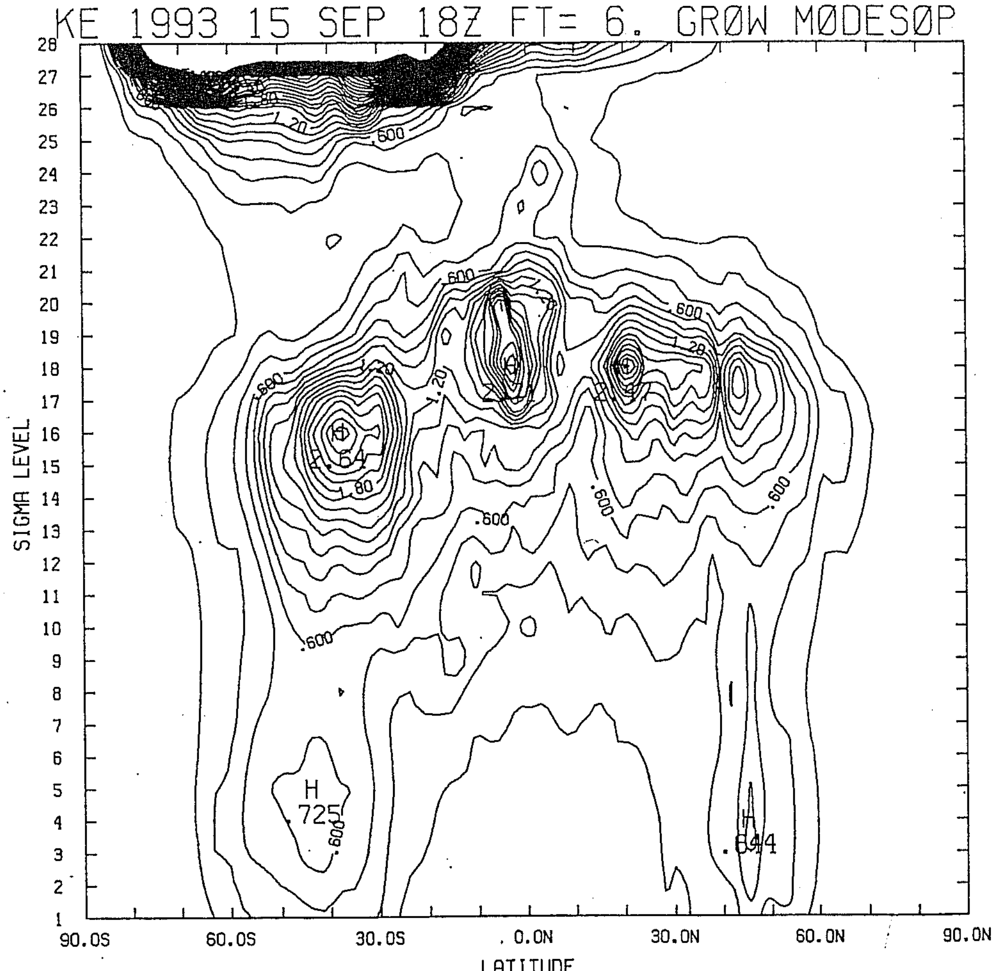

Fig. 10a: The vertical distribution of the zonally averaged kinetic energy for the operational bred vector at NCEP, valid at 18Z on September 15, 1993.

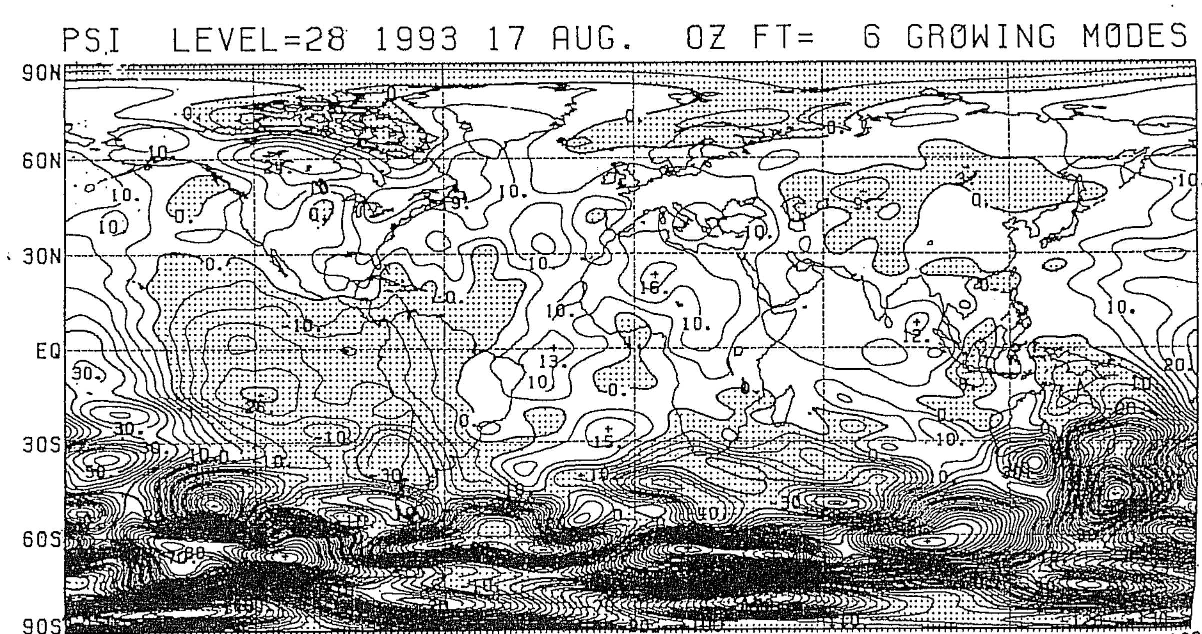

Fig. 10b: The horizontal distribution of the streamfunction perturbation (*10^{-5}) for the operational NCEP bred vector at sigma level 28 (top) for 00Z, August 17, 1993.

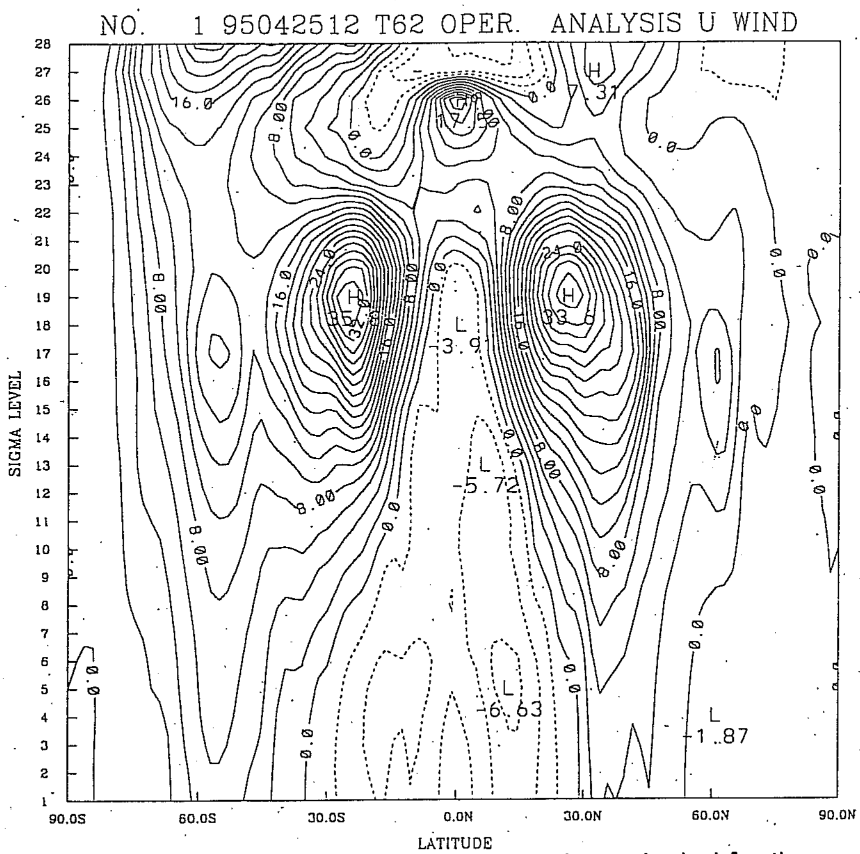

Fig. 10c: The vertical distribution of the zonally averaged zonal wind for the operational NCEP analysis, valid at 12Z on April 25, 1995.

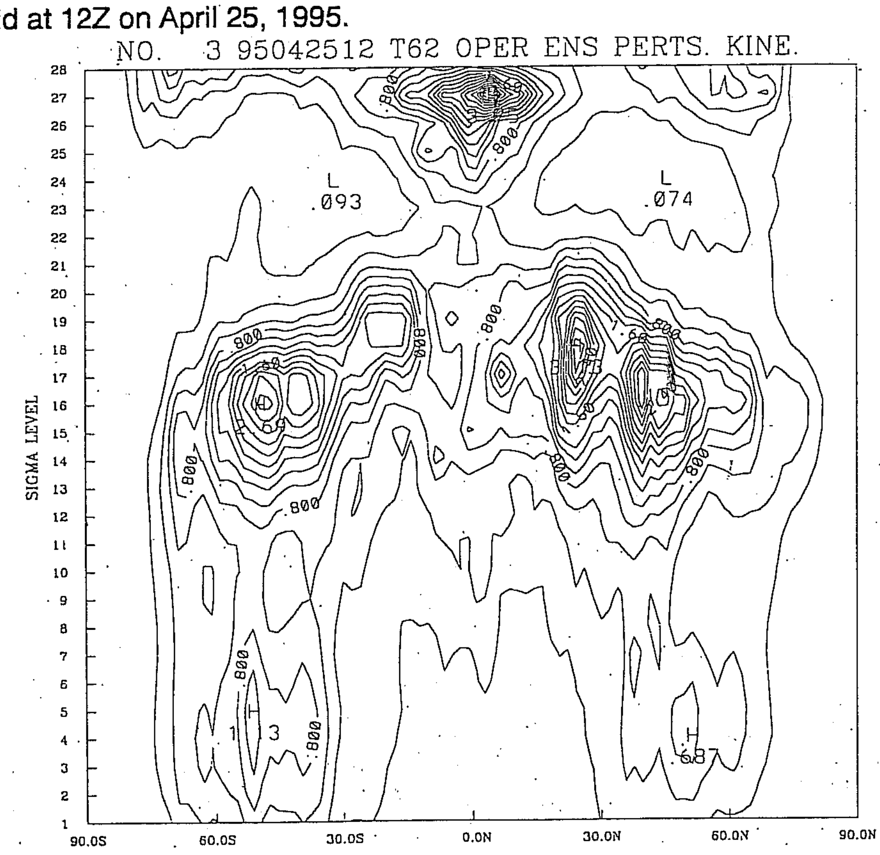

Fig. 10d: The vertical distribution of the zonally averaged kinetic energy for the operational bred vector (number three) at NCEP, valid at 12Z on April 25, 1995.

effectively filter out the smaller scale processes (e. g., convection) that would otherwise be very difficult to achieve in a strictly linear approach.

We pointed out that the analysis cycle itself is a breeding cycle that runs on top of the true state of the atmosphere, where "rescaling" is done through the use of observational data. With a perturbation amplitude that is representative of the estimated size of error in the analysis, we can estimate the fastest growing analysis errors with a breeding cycle. We presented evidence that the bred vectors indeed are present in the analysis error and with a simple method their amplitude in the error field can be reduced. Breeding can also be used to diagnose instabilities. We have to be aware, though that some of these instabilities may not be real but rather indicate problems with our nonlinear models. A comparison of the breeding method with the singular vectors approach, and the application of breeding in ensemble forecasting are presented in a companion paper (Toth and Kalnay, 1995, this volume).

TABLE 1: Properties of the breeding method and of bred growing vectors

- Breeding is a nonlinear extension of the method used to generate local Lyapunov vectors (LLVs).
- The bred growing vectors (BGVs) are a superposition of the leading LLVs (perturbations that have the fastest sustainable growth).
- Just as LLVs, bred perturbations have time continuity.
- Breeding has one free parameter, the perturbation amplitude.
- Nonlinear saturation is important to filter out fast–growing but low energy modes, such as convective modes, atmospheric instabilities within a coupled ocean atmosphere system, etc.
- Breeding is as simple and as inexpensive as running a nonlinear model twice.
- The model is used at full resolution and with no approximations.
- Within an ensemble forecasting system, the generation of bred perturbations is cost–free.
- Bred perturbations are generated over the whole globe (including the Southern Hemisphere and the tropics) at no additional cost.
- Because the analysis cycle is similar to a breeding cycle, forecasts and analyses contain BGV "errors of the day".
- Breeding can be used to improve the analysis in several different ways: a) minimize the projection of the data minus the first guess along BGVs; b) this can be done within 3 or 4D data assimilation; c) use tha BGVs to define forecast error covariances; d) use BGVs to guide adaptive observing systems (e.g., pilotless aircraft).
- BGVs can be used to detect and diagnose instabilities, including model induced, spurious ones.

REFERENCES:

Buizza, R., 1994: Sensitivity of optimal unstable structures. *Q. J. R. Meteorol. Soc.*, **120**, 429–451.

Buizza, R., J. Tribbia, F. Molteni, and T. Palmer, 1993: Computation of optimal unstable structures for a numerical weather prediction model. *Tellus*, **45A**, 388–407.

Ebisuzaki,W., and E. Kalnay, 1991: Ensemble experiments with a new lagged analysis forecasting scheme. Research Activities in Atmospheric and Oceanic Modelling, Report No. 15, WMO. {Available from WMO, C. P. No. 2300, CH1211, Geneva, Switzerland.}

Hartmann, D. L., and T. N. Palmer, 1995: Finite–Time Instabilities of Lower Stratospheric Flow. J. Atmos. Sci., under review.

Hoffman, R. N., and E. Kalnay, 1983: Lagged average forecasting. *Tellus*, **35**, 110–118.

Kalnay, E., and Z. Toth, 1992a: Efficient selection of Monte Carlo forecast ensemble members. Research Highlights of the NMC Development Division: 1989–1991. U.S. Department of Commersce, NOAA/NWS, 157–159. {Available from: NCEP EMC, 5200 Auth Rd., Room 207, Camp Springs, MD 20746.

Kalnay, E., and Z. Toth, 1992b: Estimation of the growing modes from short range forecast errors. Research Highlights of the NMC Development Division: 1989–1991. U.S. Department of Commersce, NOAA/NWS, 160–165. {Available from: NCEP EMC, 5200 Auth Rd., Room 207, Camp Springs, MD 20746.

Kalnay, E., and Z. Toth, 1994: Removing growing errors in the analysis. Proceedings of the Tenth Conference on Numerical Weather Prediction, July 18–22, 1994, Portland, OR. AMS, p. 212–215.

Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**, 409–418.

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

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

Miller, R. N., M. Ghil, and F. Gauthiez, 1994: Advanced data assimilation in strongly nonlinear dynamical systems. *J. Atmos. Sci.*, **51**, 1037–1056.

Molteni, F., and T. N. Palmer, 1993: Predictability and finite–time instability of the nothern winter circulation. *Q. J. R. Meteorol. Soc.*, **119**. 269–298.

Orlanski, I., and J. Sheldon, 1993: A case of downstream baroclinic development over Western North America. *Mon. Wea. Rev.*, **121**, 2929–2950.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.

Purser, R. J., D. F. Parrish, Z. Toth, and E. Kalnay, 1994: Numerical filtering applied to the enhancement of pre–analysis corrections of background errors expressible as rapidly amplifying modes. Preprints of the Tenth Conference on Numerical Weather Prediction, July 18–22, 1994, Portland, Oregon, p. 243–244. Available from: AMS, 45 Beacon Str., Boston, MA 02108–3693.

Schopf, P. S., and M. J. Suarez, 1988: Oscillations in a coupled ocean–atmosphere model. *J. Atmos. Sci.*, **45**, 549–566.

Toth, Z, and E. Kalnay, 1992: Estimation of the growing modes of the atmosphere: The breeding method. Research Highlights of the NMC Development Division: 1989–1991. U.S. Department of Commersce, NOAA/NWS, 439–443. {Available from: NCEP EMC, 5200 Auth Rd., Room 207, Camp Springs, MD 20746.

Toth, Z., and E. Kalnay, 1993: Ensemble Forecasting at the NMC: The generation of perturbations. *Bull. Amer. Meteorol. Soc.*, **74**, 2317–2330.335–345.

Toth, Z., and Kalnay, E., 1995a: Ensemble forecasting at NCEP and the breeding method. Mon. Wea. Rev., under review.

Toth, Z., and Kalnay, E., 1995b: Climate ensemble forecasts: How to create them? *Idojaras*, in press.

Tracton, M. S., 1994: Operational ensemble prediction – The NMC experience. Preprints of the Tenth Conference on Numerical Weather Prediction, July 18–22, 1994, Portland, Oregon, p. 206–208. Available from: AMS, 45 Beacon Str., Boston, MA 02108–3693.

Tracton, M. S. and E. Kalnay, 1993: Ensemble forecasting at NMC: Operational implementation. *Wea. Forecasting*, **8**, 379–398.

Trevisan, A., and R. Legnani, 1995: Transient error growth and local predictability: a study in the Lorenz system. *Tellus*, **A47**, 103–117.

Tsonis, A. A., 1992: Chaos: From theory to applications. Plenum Press, New York, pp. 83–96.

Wolf, A., J. B. Swift, H. L. Swinney, and J. A. Vastano, 1985: Determining Lyapunov exponents from a time series. *Physica*, **16D**, 285–317.

Zebiak, S. E., and M. A. Cane, 1987: A model El Nino/Southern Oscillation. *Mon. Weather Rev.*, **115**, 2262–2278.