J9.4 ERRORS OF THE DAY, BRED VECTORS AND SINGULAR VECTORS: IMPLICATIONS FOR ENSEMBLE FORECASTING AND DATA ASSIMILATION

Shu-Chih Yang[1], Matteo Corazza[2] and Eugenia Kalnay[1]

[1]University of Maryland, College Park, MD, and [2]ARPAL UO3 Centro Meteo Idrologico Regione Liguria

1. INTRODUCTION

In this project we plan to use the quasi-geostrophic 3D-Var data assimilation simulation system developed by Morss (1999) and Morss et al (2001) to compare several methods for data assimilation and ensemble forecasting. It is based on the a QG model of Rotunno and Bao (1996), and the 3D-Var system is similar to the NCEP operational Spectral Statistical Interpolation (SSI) method of Parrish and Derber (1992). Hamill et al (2000) have also used this system for testing comparing several ensemble forecasting methods, and Hamill and Snyder (2002) for a hybrid Ensemble Kalman Filter.

Since we want to compare 3D-Var with 4D-Var and with the recently developed Local Ensemble Kalman Filtering (LEKF) of Ott et al (2003), it was necessary to develop the adjoint of this model. As a preliminary step, we compare the analysis and forecast "errors of the day" obtained the 3D-Var with bred vectors and with singular vectors. Because of difficulties associated with the adjoint model, only preliminary results are presented in this paper. Further comparisons with Lyapunov vectors, and singular vectors with different norms will be presented at the conference. This comparison should provide guidance on the optimal generation of initial perturbations for ensemble forecasting and the relative advantages of 4D-Var and Ensemble Kalman Filtering.

2. EXPERIMENTAL SET-UP

The quasi-geostrophic model of Rotunno and Bao (1996) is a mid-latitude, beta plane, finite difference, channel model that is periodic in x and has impermeable walls at the north and south boundaries, and rigid lids at the top and bottom. Pseudo-potential vorticity is conserved except for Ekman pumping at the surface, ∇^4 horizontal diffusion and forcing by relaxation to a zonal mean state. The model is written in nondimensional form and has 64 grid points in the zonal direction, 32 grid points in meridional direction and 7 levels in vertical direction. The model variables are potential vorticity q defined at the interior levels (from levels 1 to 5) and potential temperature T defined on the bottom and top levels (levels 0 and 6).

As in Morss (1999) and Hamill et al. (2000) and others, we use a single model integration as the true or "nature" run. "Rawinsonde observations" are generated every 12 hours by randomly perturbed the true state at fixed observation locations, which were randomly chosen at initialization. The simulated data assimilation is performed with a 3D-Var data assimilation scheme, constructed by Morss (1999). In our experiments, the same model is used to generate the truth and forecasts, assuming a perfect model scenario.

(A) BRED VECTORS

The bred vectors (BV) are obtained through a breeding cycle (Toth and Kalnay, 1993, 1997), which starts by (1) adding random perturbation to the analysis; (2) integrating for 12 hours the initial conditions from both the breeding run and the analysis; (3) normalizing the differences between these two nonlinear runs; (4) adding the difference to the new analysis, and repeating steps (2) to (4). The bred vectors are defined as the normalized differences between breeding and 12 hour forecast runs.

(B) SINGULAR VECTORS

Singular vectors are the orthogonal sets describing the maximally growing perturbations. They are obtained by assuming the perturbation (δx) behaves linearly within the chosen "optimization" time interval *t* so that the linearization of the original QG model (tangent linear model, L) can represent its evolution this interval, expressed as $\delta x_t = \mathbf{L}\delta x_0$. The growth rate for a given norm at some optimization time can be defined as the perturbation norm after applying the tangent linear operator for this optimization time compared with to the initial perturbation norm defined with this norm, $<\mathbf{L}\delta x_0, \mathbf{L}\delta x_0> / <\delta x_0, \delta x_0>$, which can also be rewritten as $<\delta x_0, \mathbf{L}^T\mathbf{L}\delta x_0> / <\delta x_0, \delta x_0>$, where $\mathbf{L}^T$ is the adjoint of the tangent linear model. The leading initial singular vector is defined as the perturbation that maximizes the growth at the end of the optimization interval and can be obtained by finding the eigenvector of the matrix $\mathbf{L}^T\mathbf{L}$ with the largest singular value, as originally done by Lorenz (1965). But, in practice, the number of variables in numerical weather models is very large and because of its large dimension, it is difficult to write in matrix form. The tangent linear model and its adjoint codes, referred to as TLM and ADM have been developed for the QG model of Rotunno and Bao (1996) for this project. A first version of these codes for the TLM and ADM were originally generated by the widely used tangent linear adjoint model compiler (TMAC, Giering, 1996). This compiler provides automatic differentiation in forward (TLM) and reverse mode (ADM) for programs

written in Fortran. In our present experiments, SVs are defined with the potential enstrophy norm and with different optimization times: 12hours, 24hours and 48hours. We used both the power method and, as a verification, the Lanczos algorithm to calculate the leading SV.

It is necessary to test the correctness of both the TLM and ADM codes. For the TLM, we use the gradient check ratio is used as in Navon et all (1992), defined as $d = \frac{\left\| \mathrm{M}^{\tau}[\mathrm{x}(t) + \alpha\delta\mathrm{x}(t)] - \mathrm{M}^{\tau}[\mathrm{x}(t)] \right\|}{\left\| \alpha \mathrm{L}^{\tau}\delta\mathrm{x}(t) \right\|}$, where **M** represents the nonlinear model, α is the size of the perturbation and the ratio, d, represents the degree of linearity. With TAMC, we are able to try the expensive but accurate approach of letting the background flow evolve within the different steps of the TLM. Using single precision, the ratios are good (close to 1) within one-day time integrations but start to degrade for the potential temperatures at the bottom and top levels. When testing with different size of the perturbations (Fig. 1), the ratios remain close to 1 only if α is within the range 0.001 to 2, suggesting that round-off errors dominate, especially for potential temperatures, for amplitudes less than 0.001. If we assume (as usually done in operations) that the background is constant within one time step, the gradient check becomes worse. The ratios with different perturbation sizes are listed in table 1. Given these results, we modified TLM and adjoint models using double precision in order to avoid the dominance of round-off errors. Fig. 2 shows that the gradient check is greatly improved and can be extended to $1.e^{-10}$ when including the background flow evolution within each time step. The ratios for the TLM with constant background stay close to a value of 1.00 for integrations as long as 3 days.

For testing the correctness of the adjoint code, it is customary to take advantage of the identity $\frac{< \mathrm{L}\delta x_0, \mathrm{L}\delta x_0 >}{< \delta x_0, \delta x_0^* >} = 1$ where the star represents an adjoint variable. Even with single precision, when the tangent model accuracy is poor, the adjoint verification remains extremely good as shown in Fig 3, which can only suggest that the adjoint model is correctly coded.

3. COMPARISONS BETWEEN BVS AND SVS

Corazza et al (2002) showed that bred vectors (BVs) have the ability to depict the structures of the background error in a data assimilation system like 3D-Var, which does not include errors of the day in the background error covariance. The fact that BVs carry information on the shape of the errors of the day, is apparent in Fig. 4 adapted from Corazza et al. (2002). It shows that the subspace spanned by a set of 10 bred vectors explains 96-98% of the variance of the background error. Fig 5 (a)-(d) are examples showing the resemblance between one randomly chosen BV and the background error from day Jan31 to Feb02, even when the background error grows into smaller scales like the area in the northeastern portion in Fig.5(d)

The same background flow is used to calculate the initial and final SVs corresponding to the periods in Fig. 5, and is shown in Fig 6(a)-(f). Unlike the BVs, the initial singular vectors with different optimization times at day 31 show a more zonal pattern but evolve into final singular vectors, which also share a similarity with background errors and bred vectors. In this QG mode, it seems that the singular vectors with the optimization time of 12 hour provide the best description of the background error. This is because large growth rates of the temperature perturbations at top and bottom are created when applying the adjoint model. This pattern persists in the initial perturbations when they are normalized with the analysis error. This problem will be discussed in more detail in the next section.

In order to compare the extent to which the background error locally projects on either the BVs or on the initial and final SVs, we compute the cosine of the angle between two chosen fields. We use local regions of 5 by 5 grid points centered at each grid point and compute the cosine between the vectors of 25 values for the two fields. For comparison, we also use a flat field to project to the background error. As shown in Fig. 7, the results from projecting Initial SVs to the background error are similar to those using the flat fields. They project more to the flat part of the background error at initial time and miss the portion with fast growth, like the northeast part in Fig 7(c), thought the final SV do project much better onto some of the structures of the background error. On the other hand, the BVs project more onto fast growing background errors, and maintain a shape closer to the background error throughout the integration.

4. PRACTICAL PROBLEMS WITH THE ADJOINT MODEL

In this section we discuss two of several problems we have found in developing the adjoint of this model starting with the TAMC compiler, which is otherwise very helpful, also associated with the way the original nonlinear QG model is coded.

(A) ADJOINT OF THE POISSON SOLVER

In this QG model, the Poisson equation relating the potential vorticity and potential temperature to the streamfunction are solved in spectral coordinates.
Because of the boundary conditions, the spectral transform is written as a complex exponential in x direction, and in the y direction it is written in terms of cosine y for the zonal mean component, and sine y for the remainder of the solution. The adjoint code for the solver of the Poisson equation was originally created by Morss (1998).

Mathematically, the Poisson solver multiplies each

spectral component of the field by the inverse of the wave number squared and thus the amplification becomes proportional to this factor, causing long waves to be naturally amplified. The model solves the continuous Poisson equation

$$q = \nabla^2_{x,y}\psi + \frac{\partial}{\partial z}\left(\frac{\partial \psi}{N \partial z}\right)$$

where $\nabla^2_{x,y}$ is the horizontal Laplace operator and N is the stratification.

The boundary conditions at bottom and top levels are cast in terms of potential temperature, the vertical derivative of the streamfunction.

$$\theta = \left.\frac{\partial \psi}{\partial z}\right|_{\text{bottom and top}}$$

After converting into spectral coordinates, the Poisson equation can be written as a simple tri-diagonal matrix equation, where p is the zonal wave number, q is the meridional wave number and k is the vertical level, which can solved with a fast LU decomposition.

$\theta_{pq} = \sum_{pq} \frac{\psi_{pqk} - \psi_{pqk-1}}{\Delta z}$ for the bottom and top levels

$Q_{pqk} = \sum_{pqk} b_k \psi_{pqk} + a_{k-1}\psi_{pqk-1} + c_k \psi_{pqk+1}$ for the interior levels

where

$$b_{pqk} = -4\left(\frac{\sin^2 p/2}{\Delta x^2} + \frac{\sin^2 q/2}{\Delta y^2}\right) - \frac{2}{N_k \Delta z^2}$$

$$a_k = \frac{1}{N_{k-1}\Delta z^2}$$

$$c_k \frac{1}{N_{k+1}\Delta z^2})$$

The diagonal elements include the coefficients of both horizontal wave numbers and vertical stratification while the off diagonal elements contains the vertical terms only. For low wave numbers, this matrix is less diagonal dominant (need $|b_k| > |a_k| + |c_k|$) and has large values of Q_{pqk}.

For simplicity, we write the streamfunction as $\psi = (\nabla^2)^{-1} q$ in the forward model and assume it is self-adjoint. In the adjoint model, the adjoint of the Poisson solver will be expressed as $q^* = (\nabla^2)^{-T}\psi^*$ and therefore the adjoint of the potential vorticity will be dominated by low wave numbers. Since potential temperature is a first order vertical derivative of the streamfunction, the temperature perturbation becomes larger than potential vorticity after the application of the adjoint of the Poisson solver. Therefore, the long wave patterns in the temperature take over the growth of the total perturbation after several time steps and show a strongly zonal pattern in all adjoint fields. Fig. 8 is an example of the applying the Poisson solver for potential vorticity and temperature and taking the resulting streamfunction as the input for the adjoint of the Poisson solver. In this example, because of the presence of (small amplitude) long wave components in the initial potential vorticity and temperature, the long waves in the adjoint of vorticity and temperature fields become 10 times larger than the initial fields. Thus, this long wave pattern has large growth rate in this QG model and dominates after a short time. Note also that in Fig. 8, extremely large values appear close to the walls after the adjoint of the Poisson solver is applied. This problem is created when applying a slow cosine Fourier transform for the zonal mean component, because the problem is numerically ill-posed, creating a kink on the streamfunction for zonal wave number zero (Fig.9). We were able to correct for this problem by making the input array into a real odd function and solving with the Fast Fourier cosine transform developed for the ECMWF model (Fig. 9).

(B) PROBLEMS WITH THE BOUNDARY CONDITIONS IN THE ZONAL DIRECTION

The codes for TLM and adjoint models were automatically generated by TAMC based on the original nonlinear model. This is done by linearizing each line in the code and casting it into the equivalent of a matrix and its transpose form without the knowledge of the physical meaning of each term. In the original QG model, an extra grid point is used to impose the periodic boundary conditions, and the boundary points are considered separately at the beginning and last step in a horizontal direction loop. Code for a simple nonlinear advection equation is shown as an example of the problem that arises from this approach. Here j is index for the time integration, i is the index for the loop in the x direction, nx is the total number of grid points and nt the number of time steps. The original code with a leap-frog scheme has the form

```
   do j=3,nt
       u(1,j)=u(1,j-2)+(dt/dx)*u(1,j-1)*(
u(2,j-1)-u(nx-1,j-1))
      do i=2,nx-1
          ....
      enddo
      u(nx,j)=u(1,j)
   enddo
```

The tangent linear code created by TAMC follows, also maintaining correctly the periodic boundary condition:

```
   do j=3,nt
       g_u(1,j)=g_u(1,j-2)+(dt/dx)*
(g_u(1,j-1)*(u(2,j-1)-u(nx-1,j-1))
+u(1,j-1)*(g_(2,j-1)-g_u(nx-1,j-1)))
       do i=2,nx-1
           ....
       enddo
       g_u(nx,j)=g_u(1,j)
   enddo
```

However, based on the previous code, the TAMC converts the tangent linear code into the adjoint form as,

```
   do j=nt,3,-1
      adu(1,j)= adu(1,j)+ adu(nx,j)
```

```
      adu(nx,j)=0.
      do i=nx-1,2,-1
           ....
      enddo
      adu(1,j-2)=adu(1,j-2)+adu(1,j)
      adu(nx-1,j-1)=adu(nx-1,j-1)

      -adu(1,j)*(dt/dx)*u(nx,j-1)
      adu(2,j-1)=adu(2,j-1)

      +adu(1,j)*(dt/dx)*u(nx,j-1)
      adu(1,j-1)=adu(1,j-1)
                 +adu(1,j)*(dt/dx)*(u(2,j-
                 1)-u(nx-1,j-1))
       adu(1,j)=0.
  enddo
```

With the statement at the beginning, the value of the adjoint with be doubled at x=1, and terms related to the left end points will increase artificially their value. On the other hand, the terms related with x=nx will become smaller. Without checking the code carefully, this will create an unrealistic gradient near the both end points. To avoid this, all these wrong boundary conditions need to be commented out or considered separately.

5. SUMMARY

In this paper, the tangent linear and adjoint models for the QG model are constructed for obtaining singular vectors. The gradient check and adjoint relation check support the conclusion that these model codes have no errors. But there are some practical difficulties showing that the TLM and adjoint models obtained from TAMC have to be examined carefully in order to be consistent with the physics of the problem. Although the gradient check shows most accuracy when including the background flow evolving under different physical processes within each time step, this is too expensive, and even more expensive in the adjoint model. This will limit the optimization time for calculating accurate singular vectors. Also, some procedures like the slow cosine transforms, that are apparently harmless in the nonlinear model, are not purely invertible and cause serious boundary problems at walls if the adjoint of the cosine transform is applied repeatedly. Bred vectors are constructed using only the nonlinear model, and are not affected by these problems. They naturally lie on the attractor of the evolving background flow since they are the differences between slightly perturbed ensembles of nonlinear runs of the model.

As expected from theory, the final SVs have a similar structure as the BVs and both of them have the ability to depict the shape of the background error. The initial SVs are dominated by zonal patterns with a strong meridional gradient but the detailed structures are somewhat different with different optimization times.

The SVs obtained with the streamfunction and energy norms will be used in order to compare the initial perturbations with BVs and with Lyapunov vectors, and quantify the extent to which these vectors represent the initial and evolving background errors.

We found that the enstrophy initial SVs, dominated by zonal scales (Fig. 6, left panels), evolve towards the shape of the more realistic final SVs (Fig. 6, right panels) after only one time step. A wave number filter for SV may be needed for a better understanding of the initial evolution of the synoptic scales and their representation of the errors of the day.

We plan to use the experience acquired with these experiments when comparing 3D-Var with 4D-Var and Ensemble Kalman Filtering.

6. ACKNOWLEDGEMENTS

We are very grateful to Rebecca Morss for providing the initial code for this system and for discussions. We had helpful discussions with Istvan Szunyogh, Salil Das, Joe Tribbia and Mike Navon. This research was partially supported by the Taiwan Ministry of Education, the NPOESS Integrated Program Office and by a James McDonald 21st Century Award.

REFERENCES

Corazza, M, E. Kalnay, D. J. Patil, S-C Yang, R. Morss, M. Cai, I. Szunyogh, B.R. Hunt and J. A. Yorke, 2003: Use of the breeding technique to estimate the structure of the analysis “errors of the day”. *Nonlinear Processes in Geophysics* **10**: 1–11

Hamill, T. and C. Snyder, 2000: A comparison of probabilistic forecasts from bred, singular-vector, and perturbed observation ensembles. *Amert. Meteor. Soc.*,**128**, 1835-1851.

Hamill, T. and C. Snyder, 2000: A hybrid ensemble Kalmanfilter-3D variational analysis scheme. *Mon. Wea. Rev.* **128**, 2905-2919.

Lorenz, E. N., 1965: A study of the predictability of a 280variable atmospheric mode. *Tellus*, **17**, 321-333.

Morss, R. E., 1999: Adaptive observation: Idealized sampling satrtegies for improving numerical weather prediction. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Navon, I. M., X. Zou, J. Derber, and J. Sela, 1992: Variational data assimilation with an adiabatic version of NMC spectral model. *Mon. Wea. Rev*. **120**, 1433-1446.

Ott, E, B.R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D.J. Patil and J. A. York, 2002: A Local Ensemble Kalman Filter for Atmospheric Data Assimilation. Summit to Mon. Wea. Rev.,

Parish, D. F., and J. C. Derber, 1992: The national meteorological center’s spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747-1763.

Rotunno, R., and J.-W. Bao, 1996: A case study of

cyclogenesis using a model hierarchy. *Mon. Wea. Rev.*, **124**, 1051-1066.

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

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

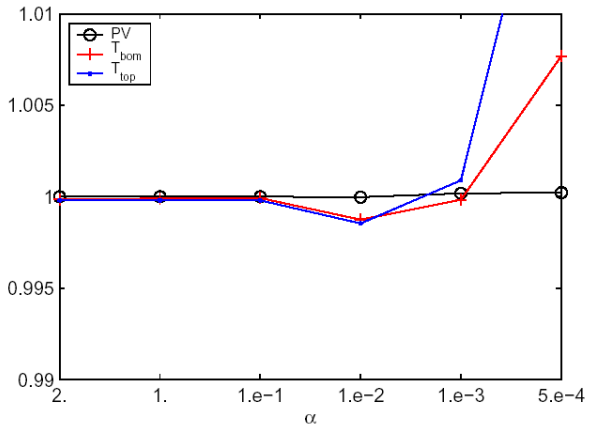
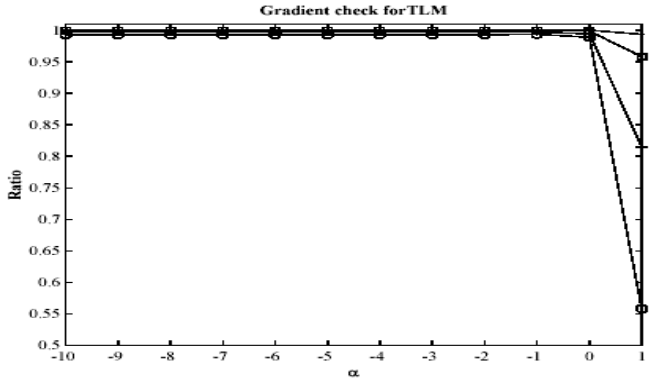


Fig. 1: Gradient check for different size perturbations using single precision

Table 1: Gradient check for different sizes of perturbations, using single precision and assuming a constant background flow within one time step

α	PV mid-level	θ_{bom}	θ_{top}
1.0	1.001716	0.9693545	0.9768672
$1.e^{-1}$	1.001742	0.9693025	0.9770426
$1.e^{-2}$	1.001708	0.9676584	0.9771303
1.e-3	1.002060	1.021203	1.029764
$1.e^{-4}$	1.001810	1.370200	1.708439

Fig. 2: Gradient check for different sizes of perturbation (10^{α}), using double precision



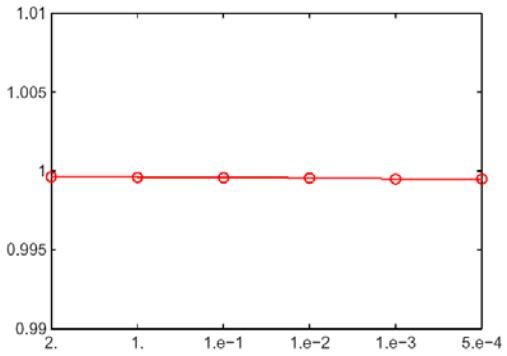


Fig. 3: Adjoint check $\frac{< \mathbf{L}\delta x_0, \mathbf{L}\delta x_0 >}{< \delta x_0, \delta x_0^* >}$ **for different size of perturbation with single precision**

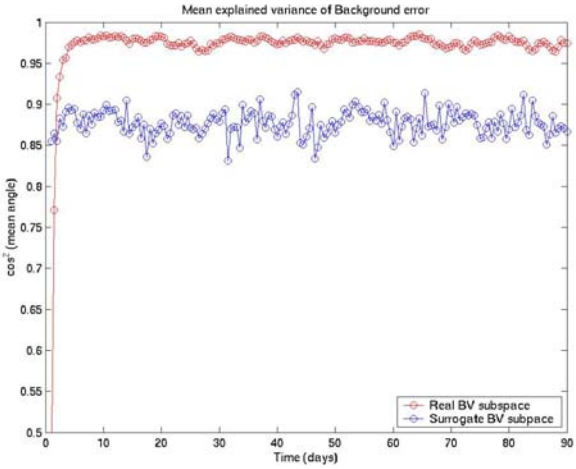


Fig. 4: Evolution of the average explained variance of the 12 hour forecast "errors of the day" as a function of time, using a set of 10 bred vectors started from initial random perturbations (red dots), and using instead surrogates (randomly chosen bred vectors, blue line). Adapted from Corazza (2003).

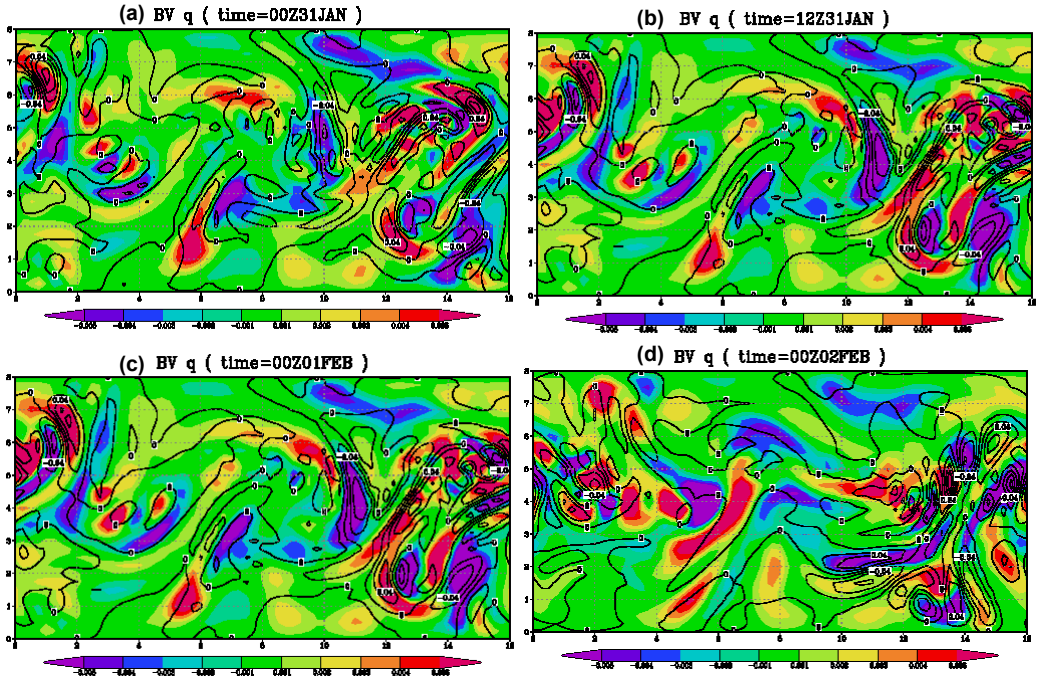


Fig. 5: Background error evolution (color shaded) compared with bred vectors (contours) for (a) at 00Z31Jan (b) at 12Z31Jan(c) 12Z31Jan (c) at 00Z01Feb and (d) 00Z02Feb.

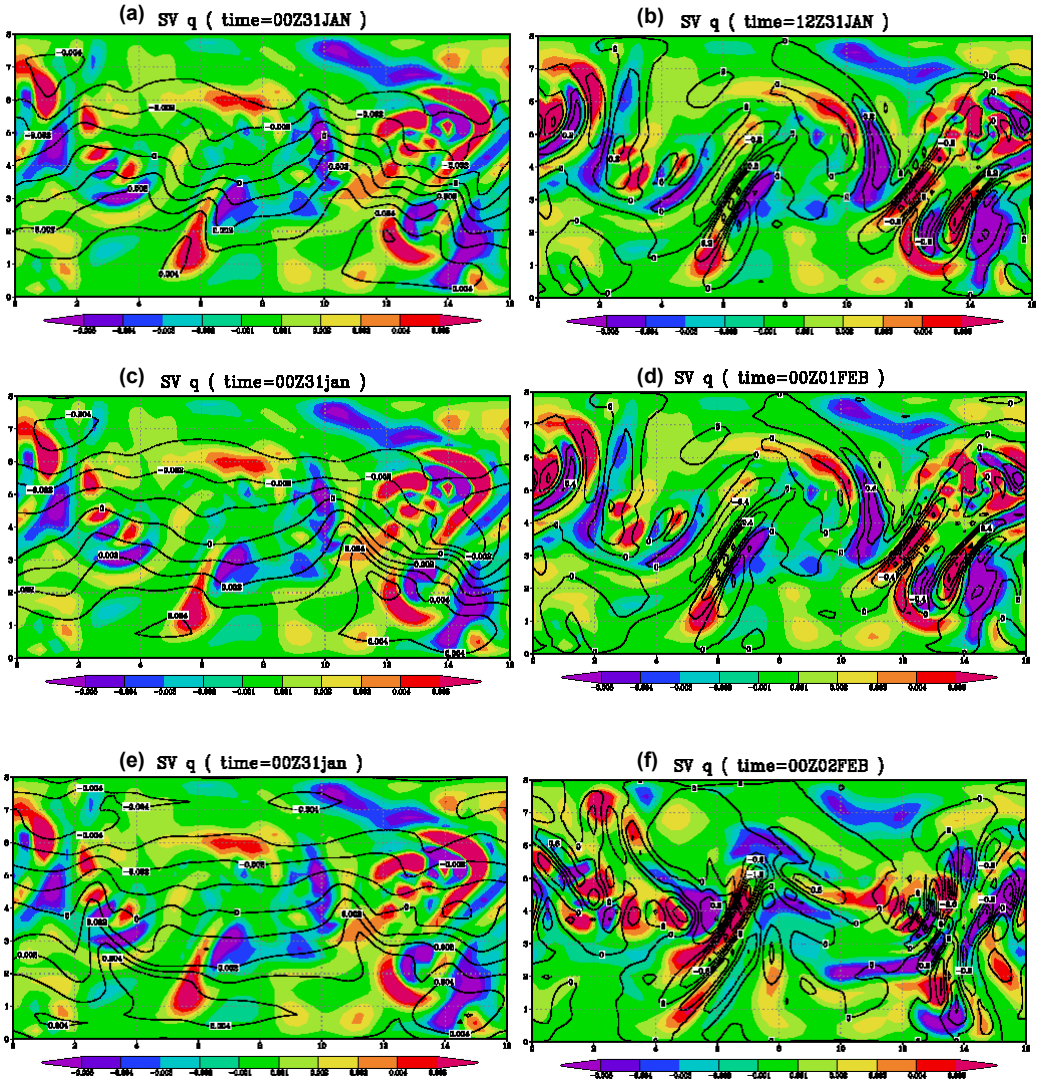


Fig. 6 Background error evolution (color shaded) compared with initial singular vectors (contours) at 00Z31Jan with an optimization time of (a) 12 hours(b) 24 hours (c) 48 hours and the corresponding final singular vectors (d) at 12Z31Jan (e) 00Z01Feb and (f) 00Z02Feb

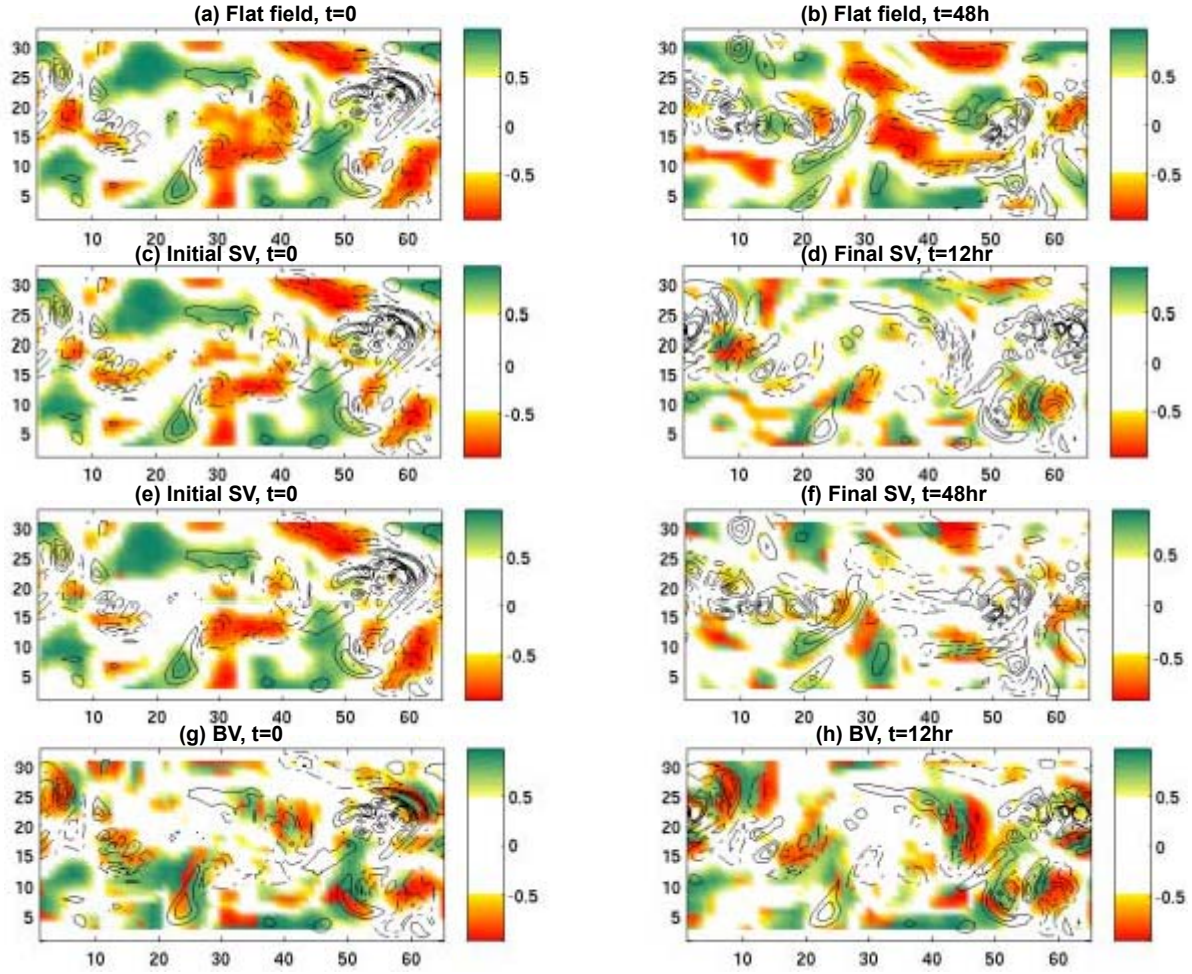


Fig. 7: Cosine of local (5x5) grid point projections between two fields. (a) Flat field projected on the background error (t=0); (b) Flat field projected on the background error (t=48hr); (c) Initial SV (optimization time=12hr) projected on the background error (t=0); (d) Final SV (optimization time=12hr) projected on the background error (t=12hr); (e) Initial SV (optimization time=48hr) projected on the background error(t=0); (f) Final SV (optimization time=48hr)) projected on the initial background error (t= 48hr); (g) BV projected on the initial background error (t=0); (h) BV projected on the final background error (t=48hr)

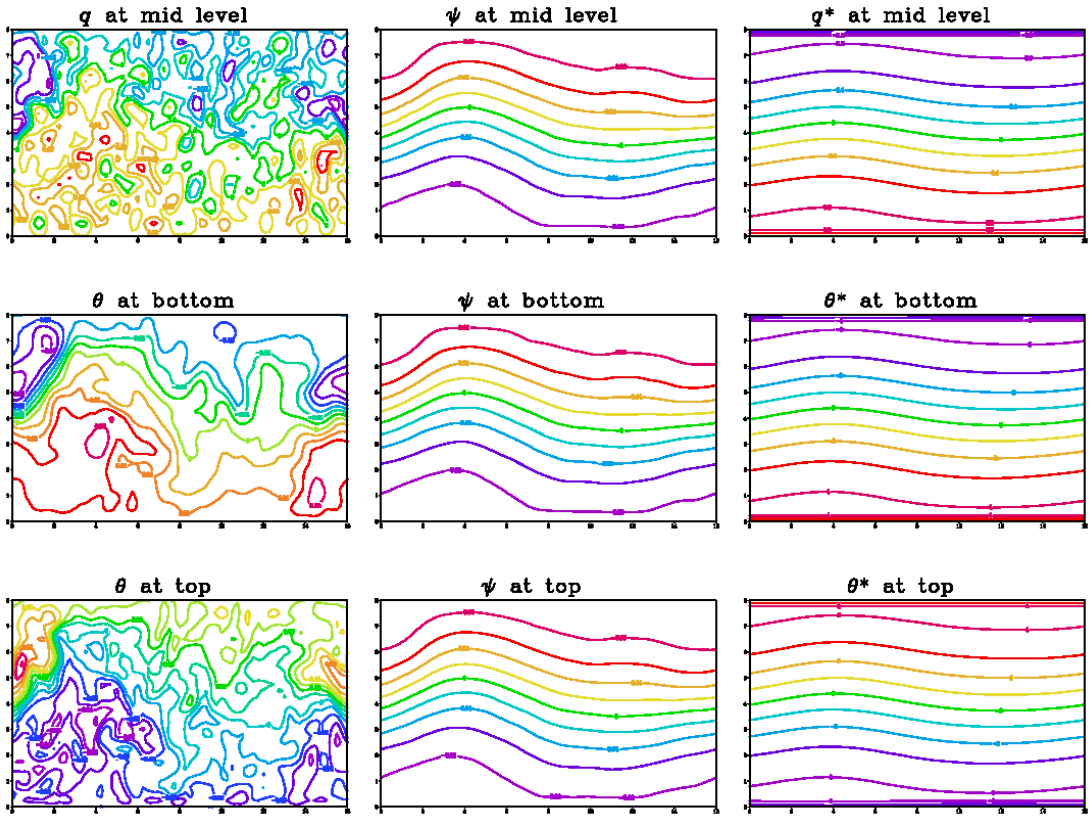


Fig. 8: Example of the application of the Poisson solver and adjoint of Poisson solver. Left panels are the potential vorticity, q, and potential temperature, θ, at bottom and top levels; central panel are the streamfunction, ψ, obtained after applying the Poisson solver; right panels are the adjoint of the potential vorticity and potential temperatures obtained by taking the stream function as the input for the adjoint of Poisson solver.

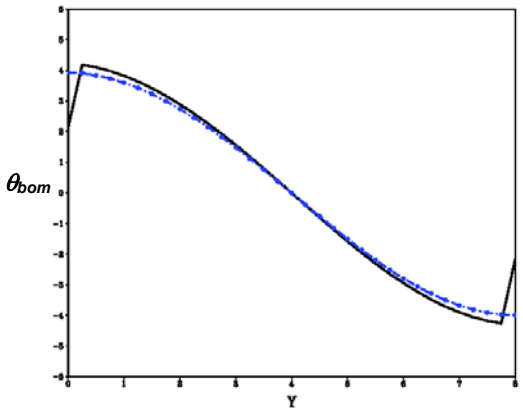


Fig. 9: Meridional cross-section for potential temperature at the bottom level using the original slow cosine Fourier transform (black) and after using the Fast complex Fourier transform (blue)