P6.11

ESTIMATION OF UNCERTAINTIES IN ATMOSPHERIC DATA ASSIMILATION USING SINGULAR VECTORS

Hyun Mee Kim* and Michael C. Morgan

University of Wisconsin-Madison
Madison, Wisconsin

Rebecca E. Morss

National Center for Atmospheric Research
Boulder, Colorado

1 Introduction

The estimation of forecast error covariance is a principal task of tuning an atmospheric data assimilation system. The main sources of forecast error can be divided into inherent model error and the error associated with uncertainties in the initial data. Both uncertainties evolve in a numerical weather prediction model during the assimilation cycle so that the associated forecast error covariance is not constant but varies with respect to time and space depending on the flow.

There are several ways to estimate forecast error covariance in an atmospheric data assimilation system. Singular vectors (SVs), the most rapidly growing perturbations over a specified time period for a prescribed norm in a given model, can be used to construct a time and space-dependent forecast error covariance matrix. Statistical methods like the maximum likelihood method or generalized cross validation (Wahba et al., 1995) may also be used to estimate the forecast error covariance. As a data assimilation algorithm, the Kalman filter can explicitly calculate the forecast error covariance as part of the data assimilation algorithm.

In this presentation the flow dependent forecast error covariance calculated using singular vectors will be compared with the forecast error covariance calculated directly in an idealized framework. Since singular vectors are norm (metric) dependent (e.g., Palmer et al., 1998), the most appropriate norm to determine singular vectors in constructing the forecast error covariance will also be investigated.

2 Experimental framework

2.1 Model

The model used is a zonally periodic, quasi-geostrophic (QG) gridpoint channel model on a beta plane. The model was developed at NCAR and has been used in several studies including Rotunno and Bao (1996), Morss (1999), and Hamill et al. (1999). The model variables are potential vorticity in the interior and potential temperature at the upper and lower boundaries. The main forcing is a relaxation to a specific zonal mean state. There is no orography or seasonal cycle and it has fourth order horizontal diffusion and Ekman pumping at the lower boundary. Stratification is constant and the tropopause is fixed.

The domain of the model is 16000 km in circumference, 8000 km in channel width and 9 km in depth. The resolutions are horizontally 250 km, vertically 5 levels. More specific description of the model can be found in Morss (1999).

2.2 Data assimilation algorithm

A three-dimensional variational data assimilation algorithm (3DVAR) developed for the above QG channel model by Morss (1999) is used. Analysis

* *Corresponding author address:* Hyun Mee Kim, University of Wisconsin-Madison, Department of Atmospheric and Oceanic Sciences, 1225 W. Dayton Street, Madison, Wisconsin 53706; e-mail: khm@mapmaker.meteor.wisc.edu

in 3DVAR can be produced by minimizing the cost function which is a combination of forecast and observation deviations from the desired analysis, weighted by the inverses of the corresponding forecast and observation error covariance matrices.

$$J = \frac{1}{2}[\mathbf{x}^T B^{-1}\mathbf{x} + (L\mathbf{x}-\mathbf{y})^T (F+O)^{-1}(L\mathbf{x}-\mathbf{y})] \tag{1}$$

where $\mathbf{x}$ is an N component vector of analysis increments, which is the distance between the analysis and the background field. B is the $N\times N$ forecast error covariance matrix. O is the $M \times M$ observational error covariance matrix. F is the $M\times M$ representativeness error covariance matrix. L is a linear transformation operator that converts the analysis variables to the observation type and location. $\mathbf{y}$ is an M component vector of observational residuals. N is the number of degrees of freedom in the analysis. M is the number of observations. By setting the derivative of J equal to 0, and rearranging :

$$[I + BL^T(F+O)^{-1}L]\mathbf{x} = BL^T(F+O)^{-1}\mathbf{y} \tag{2}$$

The 3DVAR solves Eq.(2) at each assimilation time to obtain the analysis increments $\mathbf{x}$.

2.3 Forecast error covariance

Given a linear model, the final stage of the model, ψ_t^f, can be obtained by the linear integration of the initial state, ψ_0^a :

$$\psi_t^f = P\psi_0^a \tag{3}$$

where P is the tangent linear model of the nonlinear QG model. The forecast error covariance matrix can be defined as,

$$B = E[\psi'\overline{\psi'}^T] = E[(\psi - E[\psi])\overline{(\psi - E[\psi])}^T] \tag{4}$$

where ψ' is error or perturbation and $E[\cdot]$ denotes the expectation operator. The time evolution of the total error is the sum of the dynamical growth of the initial error and the model error.

$$\psi' = P{\psi_0}' + \psi'' \tag{5}$$

where ψ'' is model error. Substitution of Eq.(5) into Eq.(4) yields :

$$\begin{aligned} B &= E[(P\psi_0' + \psi'')(P\overline{\psi_0'} + \overline{\psi''})^T] \\ &= PE[\psi_0'\overline{\psi_0'}^T]P^T + E[\psi''\overline{\psi''}^T] \\ &= PAP^T + Q \end{aligned} \tag{6}$$

where $A = E[\psi_0'\overline{\psi_0'}^T]$ is analysis error covariance matrix and $Q = E[\psi''\overline{\psi''}^T]$ is model error covariance matrix.

By perfect model assumption Q is zero and the forecast error covariance B becomes :

$$B = PAP^T \tag{7}$$

2.4 Singular Vectors

Because of the rapidly growing property of SVs, they have been used to :

- study the instability properties of the atmosphere-ocean system,
- construct the initial ensembles to produce ensemble prediction of ECMWF,
- adaptive (or targeting) observations to detect regions of large sensitivity to small perturbations, and
- construct the eigenvectors of the forecast error covariance matrix for the end of the optimization interval.

SVs can be calculated by maximizing the final amplitude of the perturbation subject to the constraint that the initial perturbation be of unit amplitude for a specified metric. The final amplitude may be represented by the cost function J as follows :

$$J(\psi_0') = (P\psi_0')^T(P\psi_0') \tag{8}$$

The constraint is ${\psi_0'}^T(A)^{-1}\psi_0' = 1$ and where A is the analysis error covariance at the initial time. By the Lagrange multiplier method,

$$L = (P\psi_0')^T(P\psi_0') - \lambda({\psi_0'}^T(A)^{-1}\psi_0' - 1) \tag{9}$$

Differentiating Eq.(9) with respect to ψ_0 and equating that derivative to zero yields :

$$P^TP\psi_0' = \lambda(A)^{-1}\psi_0' \tag{10}$$

The formulation of Eq.(10) can be related to that of Eq.(7) by calculating eigenvectors of Eq.(7).

$$PAP^T y = \lambda y \tag{11}$$

By rearranging and using $y = P\psi_0'$, Eq.(11) becomes the exactly same form of Eq.(10). Therefore the time evolved SVs can be used to efficiently construct that part of the forecast error covariance associated with the uncertainty of initial data (Ehrendorfer and Tribbia, 1996).

The Lanczos algorithm was used to calculate SVs for the QG channel model and adjoint of the tangent linear version of the QG channel model was developed.

3 SVs for different norms

While SVs consistent with the forecast error covariance may be calculated based on the initial analysis error covariance metric (Barkmeijer et al., 1998), several other norms including the potential enstrophy, L^2, kinetic energy, and total energy have been considered in the predictability studies since the actual analysis error covariance is not known.

In this study the potential enstrophy, L^2 and total energy norms are used to approximate the analysis error covariance norm to estimate forecast error covariance. The SVs are calculated for a time varying basic state.

3.1 Potential enstrophy norm

The potential enstrophy (square of the disturbance QG potential vorticity) of QG model in discrete form is :

$$\begin{aligned} Q &= \frac{1}{2}\sum_{l=1}^{L+1}\sum_{m=1}^{M+1}\sum_{n=1}^{N} {q'}^2 \\ &+ \frac{1}{2S}\sum_{l=1}^{L+1}\sum_{m=1}^{M+1}({\theta'}^2_{n=0} + {\theta'}^2_{n=N+1}) \end{aligned} \tag{12}$$

where S is static stability and l, m, n are indexes of x, y, z grid points respectively. L, M, N are corresponding numbers of grid points for x, y, z.

The most rapidly amplifying SV in the potential enstrophy norm for 6 hour optimization time is shown in Fig.1. The structure of SV is quite zonal.

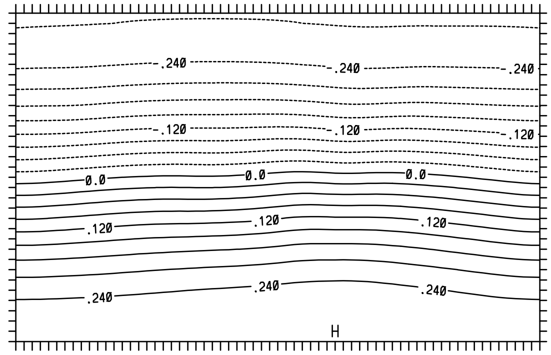

Figure 1: Horizontal cross-section of the leading SV streamfunction in potential enstrophy norm at middle (N=3) of the domain for 6 hour optimization time

3.2 L^2 norm

The square of the disturbance of streamfunction of QG model in discrete form is :

$$L^2 = \sum_{l=1}^{L+1}\sum_{m=1}^{M+1}\sum_{n=0}^{N+1} {\psi'}^2 \tag{13}$$

The leading SV in L^2 norm of 6 hour optimization time for the same starting time with SV in potential enstrophy norm is shown in Fig.2. The leading SV in the L^2 norm has more horizontal structure compared with SV calculated for the potential enstrophy norm in Fig.1.

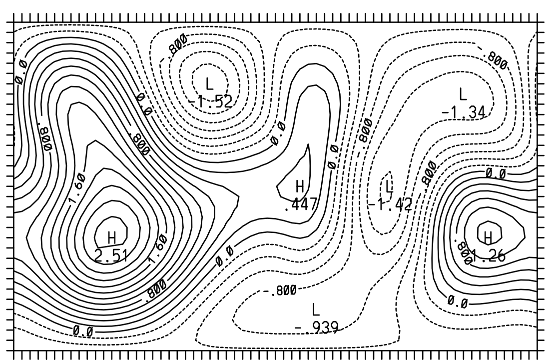

Figure 2: Horizontal cross-section of the leading SV streamfunction in L^2 norm at middle (N=3) of the domain for 6 hour optimization time

3.3 Energy norm

The energy of disturbance in QG model can be represented in discrete form as follows :

$$E = \frac{1}{2}\sum_{l=1}^{L+1}\sum_{m=1}^{M+1}\sum_{n=0}^{N+1}({\psi'_x}^2 + {\psi'_y}^2 + \frac{1}{S}{\psi'_z}^2) \qquad (14)$$

SVs have not yet been calculated for the energy norm which is the most frequently used norm for predictability studies.

4 Conclusion and future work

SVs calculated for QG channel model show different structures with different metrics. SVs in the potential enstrophy norm show large scale and zonal structure while those in the L^2 norm show smaller and localized structure. These results are similar to results found in the Eady model (Kim and Morgan, 1999).

Until now we have written the adjoint code and SVs calculation routine for QG channel model and looked at the characteristics of SVs based on different norms. In the presentation the flow dependent forecast error covariance calculated using singular vectors based on several metrics will be compared with the forecast error covariance calculated directly in 3DVAR since we know what the truth is in this idealized framework.

Acknowledgment This work was supported in part by the National Science Foundation grant ATM-9810916. We are indebted to Dr. Chris Snyder of NCAR for numerous discussions on the adjoint coding and providing us the quasi-geostrophic channel model.

5 References

Barkmeijer, J., M. van Gijzen, and F. Bouttier, 1998: Singular vectors and estimates of the analysis error covariance metric. *Quart. J. Roy. Meteor. Soc.*, **124**, 1695-1713.

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

Hamill, T. M., C. Snyder, and R. E. Morss, 1999: Error characteristics of bred, singular vector, and perturbed observation ensemble forecasts. *Mon. Wea. Rev.*, **128**, 1835-1851.

Kim, H. M., and M. Morgan, 1999: Dependence of singular vector structure and evolution on the choice of norm. Preprints, 12^{th} *Conference on Atmospheric and Oceanic Fluid Dynamics,* New York, NY.

Morss, R., E., 1999: Adaptive observations: Idealized sampling strategies for improving numerical weather prediction. Ph.D. thesis, M.I.T.

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

Rotunno, R., and J. W. Bao, 1996: A case study of cyclogenesis using a model hierarchy. *Mon. Wea. Rev.*, **124**, 1051-1066.

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