JP1.25 **ESTIMATION OF UNCERTAINTIES IN ATMOSPHERIC DATA ASSIMILATION USING SINGULAR VECTORS**

Hyun Mee Kim* and Michael C. Morgan

University of Wisconsin-Madison
Madison, Wisconsin

1 Introduction

The estimation of the 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 (Ehrendorfer and Tribbia, 1997). 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 SVs will be compared with the true forecast error covariance calculated directly in an idealized framework under the perfect model assumption. The potential enstrophy norm constrained by the initial analysis error covariance is used to calculate SVs to construct the forecast error covariance.

Section 2 contains a brief description of model, data assimilation algorithm, and the relationship of the forecast error covariance and SVs. The initial and evolved structures of SVs for the potential enstrophy norm are compared in section 3.

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 in 3DVAR can be produced by minimizing the cost

* *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

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.

$$J = \frac{1}{2}[\mathbf{x}^T B^{-1}\mathbf{x} + (L\mathbf{x} - \mathbf{y})^T (F+O)^{-1}(L\mathbf{x} - \mathbf{y})] \tag{2.1}$$

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 :

$$[I + BL^T(F+O)^{-1}L]\mathbf{x} = BL^T(F+O)^{-1}\mathbf{y} \tag{2.2}$$

The 3DVAR solves (2.2) at each assimilation time to obtain the analysis increments $\mathbf{x}$.

2.3 Forecast error covariance and SVs

Given a linear model, the final stage of the model, q_t^f, can be obtained by the linear integration of the initial state, q_0^a :

$$q_t^f = Pq_0^a \tag{2.3}$$

where q is the statevector of the QG channel model, which is potential vorticity in the interior and potential temperature at the upper and lower boundaries. P is the tangent linear model of the nonlinear QG model. The forecast error covariance matrix can be defined as,

$$B = E[q'\overline{q'}^T] = E[(q - E[q])\overline{(q - E[q])}^T] \tag{2.4}$$

where q' is the PV perturbation for the whole domain including the boundary potential temperature 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:

$$q' = Pq_0' + q'' \tag{2.5}$$

where q'' is model error. Substitution of (2.5) into (2.4) yields :

$$\begin{aligned} B &= E[(Pq_0' + q'')(P\overline{q_0'} + \overline{q''})^T] \\ &= PE[q_0'\overline{q_0'}^T]P^T + E[q''\overline{q''}^T] \\ &= PAP^T + Q \end{aligned} \tag{2.6}$$

where $A = E[q_0'\overline{q_0'}^T]$ is analysis error covariance matrix and $Q = E[q''\overline{q''}^T]$ is model error covariance matrix.

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

$$B = PAP^T \tag{2.7}$$

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 for the potential enstrophy norm in the QG channel model may be represented as follows :

$$\begin{aligned} \|q_t'\|_Q^2 = \langle q_t', q_t'\rangle_Q &= \langle Pq_0', QPq_0'\rangle_{L_2} \\ &= \langle q_0', P^TQPq_0'\rangle_{L^2} \end{aligned} \tag{2.8}$$

where Q is the norm matrix corresponding to the potential enstrophy norm and determined as,

$$\begin{aligned} Q &= \frac{1}{2}\sum_{l=1}^{L+1}\sum_{m=1}^{M+1}\sum_{n=1}^{N} {q'}^2 \\ &+ \frac{1}{2S}\sum_{l=1}^{L+1}\sum_{m=1}^{M+1}({\theta'}^2_{n=0} + {\theta'}^2_{n=N+1}) \end{aligned} \tag{2.9}$$

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

The constraint is that the initial perturbation is of unit amplitude for the inverse of analysis error covariance at the initial time and may be represented as :

$$\|q_0'\|_{A^{-1}}^2 = \langle q_0', q_0'\rangle_{A^{-1}} = \langle q_0', A^{-1}q_0'\rangle_{L^2} \tag{2.10}$$

where A is the analysis error covariance at the initial time. The ratio of the final and initial amplitude of perturbation becomes :

$$\lambda = \frac{{q_0'}^T P^T QPq_0'}{{q_0'}^T A^{-1}q_0'} \tag{2.11}$$

Therefore the problem maximizing the ratio λ becomes eigenvalue problem as :

$$AP^TQPq_0' = \lambda q_0' \tag{2.12}$$

After multiplying $A^{-1/2}$ to both sides of (2.12), the matrix AP^TQP becomes symmetric so that we can use the Lanczos algorithm to solve this eigenvalue problem. The (2.12) becomes :

$$A^{1/2}P^TQPA^{1/2}\widetilde{q_0'} = \lambda\widetilde{q_0'} \tag{2.13}$$

The formulation of (2.12) can be related to that of (2.7) by calculating eigenvectors of (2.7).

$$PAP^Ty = \lambda y \tag{2.14}$$

By rearranging, (2.14) becomes the equivalent form of (2.12). 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.

The basic state is time-variant and the adjoint of the tangent linear version of the QG channel model was developed.

3 Results

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.

The most rapidly amplifying SV in the potential enstrophy norm for 24 hour optimization time for the arbitrary basic state is shown in Fig.1. The initial structure of SVs in the potential enstrophy norm is of large scale and it is similar to the results found in the Eady model (Kim and Morgan, 1999) and in the general circulation model (*e.g.*, Palmer et al. 1997). The evolved SV structure, however, is of smaller scale and localized structure.

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

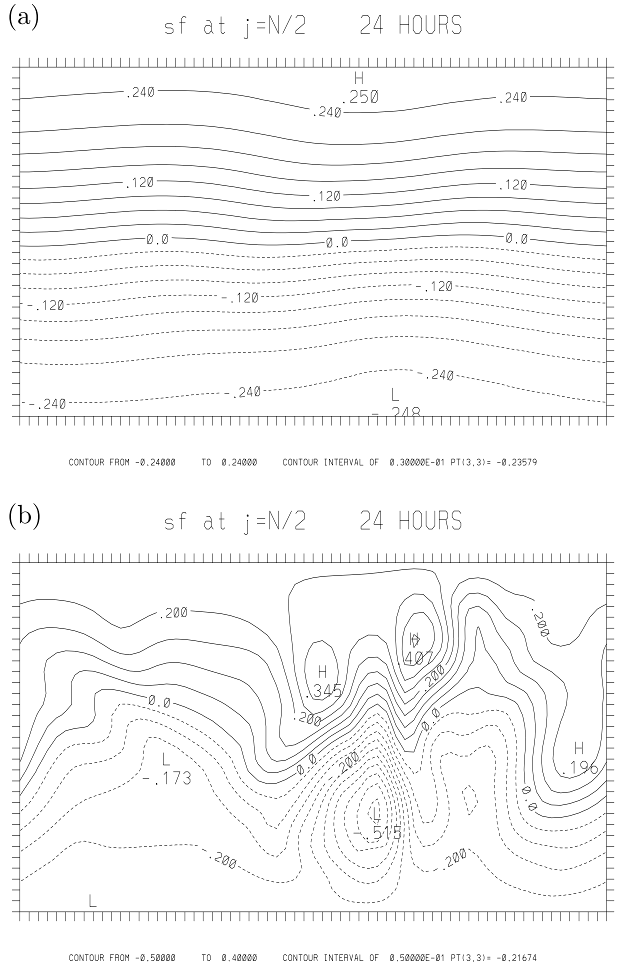

Figure 1: Horizontal cross-section of the (a) initial structure and (b) evolved structure of the leading SV streamfunction in potential enstrophy norm at middle (N=3) of the domain for 24 hour optimization time

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. We also thank to Dr. Rebecca Morss for providing us the data assimilation system.

4 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, 1997: 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.