An Ensemble-Based Four-dimensional Variational Data Assimilation Scheme

Chengsi Liu[1,2], Qingnong Xiao*[1], and Bin Wang[2]

1. ESSL/MMM, National Center for Atmospheric Research, Boulder, CO

2. LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

1. Introduction

The incremental approach of four-dimensional variational (4D-Var) data assimilation (Courtier et al. 1994) and Ensemble Kalman filters (EnKF, Evensen 1994) are well known as two advanced data assimilation approaches. As a retrospective assimilation algorithm, 4D-Var can gain the optimal trajectory and can effectively assimilate non-synoptic data (Xiao et al. 2002; Simmons and Hollingsworth 2002). EnKF, on the other hand, can use flow-dependent background error covariance (B matrix) calculated from ensemble forecast and can be easily implemented without tangent linear and adjoint models. In recent years, approaches of coupling the two data assimilation algorithms have been proposed, e.g. Ensemble Kalman Smoother (Evensen et al. 2000), maximum likelihood ensemble filter (Zupanski 2005), and 4DEnKF (Hunt et al. 2004; Fertig et al. 2007). These approaches use the flow-dependent B matrix based on statistics from ensemble forecasts while maintain retrospective assimilation character. Liu et al. (2008a) presented an ensemble-based four-dimensional variational (En4D-Var) algorithm, which uses the flow-dependent B matrix constructed by ensemble forecasts and performs 4D-Var optimization. This approach (En4D-Var) adopts the incremental and preconditioning idea in variational algorithm so that it can be easily incorporated in many operational and research communities' variational assimilation system. In addition, En4D-Var avoids tangent linear model and its adjoint, the two components that are difficult to develop and that make 4D-Var minimization computationally expensive.

2. Theoretical background of En4D-Var

The idea of En4D-Var is that the preconditioning matrix in the incremental approach of 4D-Var (Courtier et al 1994; Gilbert et al 1989) is replaced with a perturbation matrix. The column of perturbation matrix, $\mathbf{X}'_b$, is the normalized deviations from the ensemble mean and estimated by N ensemble members, *viz.*

$$\mathbf{X}'_b = \frac{1}{\sqrt{N-1}}(\mathbf{x}_{b1} - \overline{\mathbf{x}_b}, \mathbf{x}_{b2} - \overline{\mathbf{x}_b}, \ldots, \mathbf{x}_{bN} - \overline{\mathbf{x}_b}) . \quad (1)$$

The background error covariance **B** is approximately calculated by

$$\mathbf{B} \approx \mathbf{X}'_b \mathbf{X}'^{\mathrm{T}}_b . \quad (2)$$

Assuming the innovations at different times (with subscript *i*) are calculated by

$$\mathbf{d}_i = HM(\mathbf{x}_b) - \mathbf{y}_i , \quad (3)$$

where $\mathbf{x}_b$ is background state vector, *H* is observation operator, *M* is forecast model and **y** is observation state vector. The En4D-Var cost function in control variable space is defined by

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} + \frac{1}{2}\sum_{i=0}^{I}(\mathbf{HMX}'_b\mathbf{w} + \mathbf{d}_i)^{\mathrm{T}}\mathbf{O}^{-1}(\mathbf{HMX}'_b\mathbf{w} + \mathbf{d}_i), \quad (4)$$

where **w** is control variable, *I* is the total number of time levels on which observations are available, **H** is tangent linear observation operator, **M** is tangent linear forecast model and **O** is observation error covariance.

To avoid tangent linear and adjoint models in calculating the gradient of cost function, we transform the perturbation matrix to observation space via

* *Corresponding author: Dr. Qingnong Xiao, NCAR/MMM, PO Box 3000, Boulder, CO 80307-3000. Emai: hsiao@ucar.edu*

$$\mathbf{HMX}'_b \approx \frac{1}{\sqrt{N-1}}(HM\mathbf{x}_{b1} - HM\overline{\mathbf{x}}_b, HM\mathbf{x}_{b2} - HM\overline{\mathbf{x}}_b, ..., HM\mathbf{x}_{bN} - HM\overline{\mathbf{x}}_b). \tag{5}$$

The gradient of the cost function is then calculated by

$$\nabla_{\mathbf{w}} J = \mathbf{w} + \sum_{i=0}^{I}(\mathbf{HMX}'_b)^{\mathrm{T}}\mathbf{O}^{-1}(\mathbf{HMX}'_b\mathbf{w} + \mathbf{d}_i). \tag{6}$$

After minimization iteration, the optimal analysis $\mathbf{x}_a$ can be obtained from

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{X}'_b\mathbf{w}. \tag{7}$$

3. Horizontal and vertical localization in En4D-Var

In order to reduce sampling errors due to finite ensemble numbers, Houtekamer and Mitchell (2001) employed Schur operator (Gaspari and Cohn 1999) in EnKF, and Lorenc (2003) and Buehner (2005) used Schur operator in variational scheme. In En4D-Var, we introduce an EOF decomposed correlation function operator to modify the perturbation matrix, $\mathbf{X}'_b$. This is an approach similar to spatial localization of Buehner (2005). The modified perturbation P'_b is defined by

$$\mathbf{P}'_b = [\mathbf{E}_v\lambda_v^{1/2}\cdot(\mathbf{E}_{h1}\lambda_{h1}^{1/2}\cdot\mathbf{X}'_{b1}, ..., \mathbf{E}_{h1}\lambda_{h1}^{1/2}\cdot\mathbf{X}'_{bN}), ..., \mathbf{E}_v\lambda_v^{1/2}\cdot(\mathbf{E}_{hn}\lambda_{hn}^{1/2}\cdot\mathbf{X}'_{b1}, ..., \mathbf{E}_{hn}\lambda_{hn}^{1/2}\cdot\mathbf{X}'_{bN})]. \tag{8}$$

In Eq. (8), the subscripts *h* and *v* are horizontal and vertical indices. **E** contains all of eigenvectors and λ is a diagonal matrix that the diagonal elements are eigenvalues. The eigenvalue is obtained from EOF decomposed correlation function **C**,

$$\mathbf{C} = \mathbf{E}\lambda\mathbf{E}^{\mathrm{T}}. \tag{9}$$

The new perturbation matrix (8) used in En4D-Var is equivalent to the modified **B** matrix by correlation function in EnKF (Appendix A of Liu et al. 2008b), e.g. Schur product operator,

$$\mathbf{P}_b = \mathbf{C}_v\cdot(\mathbf{C}_h\cdot\mathbf{B}). \tag{10}$$

We adopt the compactly-supported second-order auto-regressive function as horizontal correlation model (Liu and Rabier 2002), i.e.

$$\rho(s) = \begin{cases} (1+\frac{s}{s_0})e^{-\frac{s}{s_0}}(1-\frac{s}{s_1}) & s \le s_1 \\ 0 & s > s_1 \end{cases}, \tag{11}$$

where *s* stands for the spherical separation in degree between two data points, s_0 and s_1 are the correlation scale and the cut-off distance beyond which the correlation become zero. To perform vertical localization, we use a correlation function following Zhang et al. (2004),

$$\rho(\Delta\log p) = \frac{1}{1 + 5\times(\Delta\log p)^2}. \tag{12}$$

In (12), $\Delta\log p$ is the distance between two vertical levels in $\log p$, and *p* is pressure.

4. Observing system simulation experiments with the blizzard of 2000

The WRF-ARW (Skamarock et al. 2005) was used for this study. All experiments are conducted over a grid mesh of 94×94 with grid-spacing of 27 km. the domain covers the eastern half of United States. In the vertical, there are 27 η layers. The control run (CTRL) is conducted by integrating the first-guess at 1800 UTC 23 Janunary 2000 for 42 hours. The random pertubations are added to the first-guess of CTRL to produce.37 ensemble members. One member, which compares most favorably to observations in terms of the location and strength of the cyclone is chosen as true simulation and the other members are treated as reference forecast ensemble. The random pertubations are derived from the background error covarince of the WRF 3D-Var data assimilation system, with the similar approach to the initial ensemble generating method in Houtekamer et al (2005). The pertubations are therefore consistent with the background error covarince defined by the WRF 3D-Var data assimilation. We also perturbed boundary conditions using the Data Assimilation Research Testbed (DART) system (Anderson 2001; 2003).

a. Test of the En4D-Var localization technique

To perform localization, we introduce the correlation function operator in the B matrix. However, when the correlation function operator is applied to En4D-Var, the control vectors are enlarged and the computation cost becomes very expensive. In order to reduce the computation cost, EOF is used to decompose the correlation function and limited truncation modes are selected.

We employ a single observation test to examine the En4D-Var localization technique in this study. The single temperature observation at 33.5^{o}N, 78.4^{o}W and 850-hPa are assimilated using WRF 3D-Var, En4D-Var without localization (En4D-Var-NL), and En4D-Var with localization (En4D-Var-L). The results of wind vector and temperature increment at 1000-hPa are shown in Figure 1. The temperature analysis increment in WRF 3D-Var (Fig.1a) indicates a homogeneous and isotropic structure that the B matrix of WRF 3D-Var has. Due to the quasi-geostrophy relationship in the variable transform of WRF 3D-Var, the wind analysis increments show quasi-geostrophy characteristics in the single temperature observation assimilation experiment. On the other hand, Figs. 1b and 1c demonstrate flow-dependent B matrix structure in En4D-Var. The temperature analysis increment in En4D-Var extends along the eastern coast, corresponding to the flow direction and isotherm (not shown). Different from the constraint of variable transform technique in WRF 3D-Var, the relations among different physical variables in En4D-Var depend on ensemble statistics. The increments of wind vectors in En4D-Var also show some quasi-geostrophy characteristics (Figs. 1b and 1c), which indicates quasi-geostrophy constraint between temperature and winds in En4D-Var is presented well by ensemble statistics.

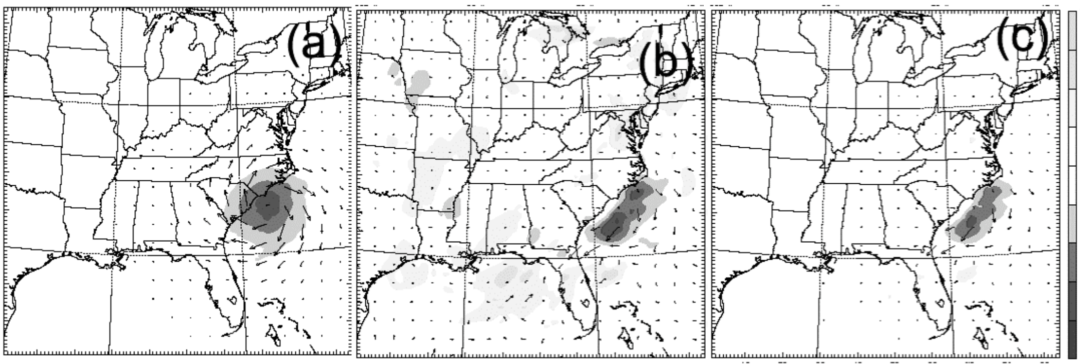

Figure 1: The response increments of wind vector and temperature increment (shadow) at 1000-hPa.from single observation test with (a) WRF 3D-Var, (b) En4D-Var without localization, and (c) En4D-Var with localization.

If no localization is performed in En4D-Var (Fig. 1b), a lot of increment noises are found due to sampling errors. This has been discussed in section 1. The amplitude of these noises can be comparable with the increment signal at observation location. Using localization by the EOF decomposed correlation function operator, the noises almost disappears but the increment signal from observation is still maintained (Fig. 1c). Figure 2 shows the vertical profile of temperature analysis increment at observation location by En4D-Var-NL and En4D-Var-L. At the high levels, there is obvious fake analysis increment if the localization isn't applied but the noise is filtered out after localization. Although the noises can be filtered out by localization technique, the analysis increment signal is also a little reduced compared to the one without localization because limited modes are used. However, the major analysis feature is retained since over ninety percent signal is still maintained.

b. Comparison of En4D-Var and En3D-Var cycling

En4D-Var is an ensemble-based retrospective algorithm. If no sampling errors are considered, the En4D-Var analysis fits to the optimal trajectory. However, it does not necessarily fit to each single optimal point as ensemble-based sequential algorithm provides. The temporal sampling errors can affect En4D-Var analysis as discussed in previous discussions of this paper. In order to compare the ensemble-based sequential and ensemble-based retrospective algorithms, we design the En3D-Var cycling experiment that uses the same configuration of En4D-Var but assimilates only one-time observation each time and cycles all the observations in the assimilation window.

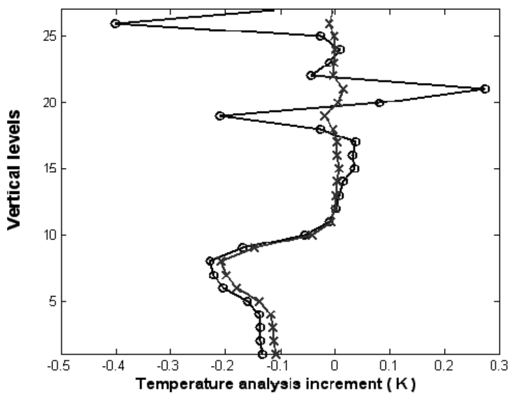

Figure 2: Corresponding to Figure 4, the vertical profile of temperature increment at the observation location by En4D-Var-NL (circle-line) and En4D-Var-L (cross-line)

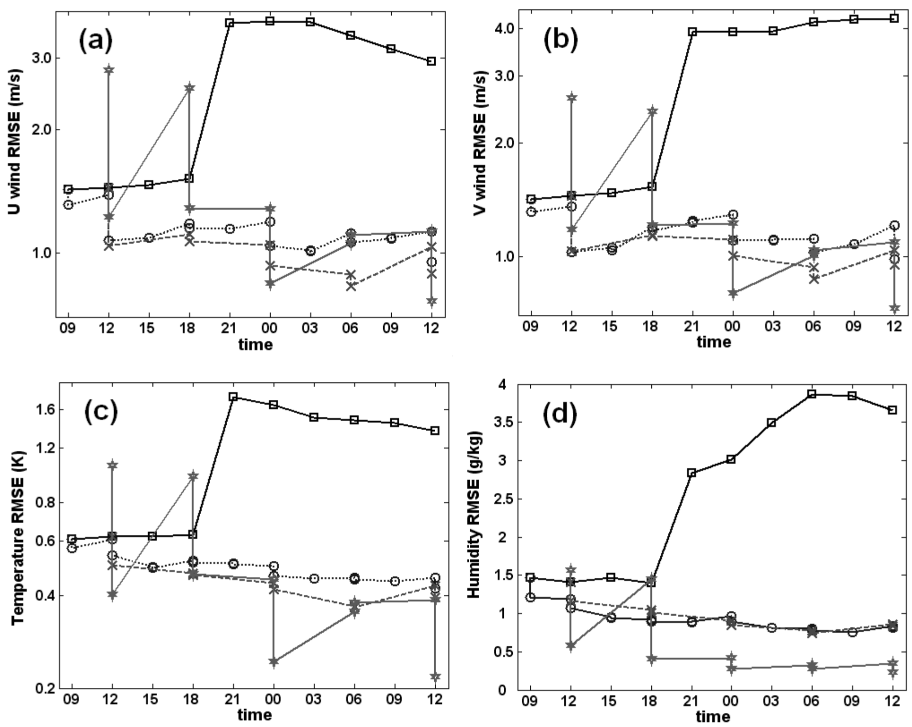

Figure 3: The variation of domain-averaged RMSE in CTRL (square-line), En3D-Var (circle-line) and En4D-Var (cross-line) with time for (a) U winds, (b) V winds, (c) temperature, and (d) humidity. The star-line is the variation of forecast-analysis spread statistics with time in En4D-Var.

Figure 3 shows the time variations of the domain-averaged RMSE in the analyses of CTRL, En3D-Var cycling and En4D-Var. It is clear that the error in CTRL suddenly rises after 1800 UTC 24 and become stable after 2100 UTC 24. But the humidity error still slowly increases until 0600 UTC 25. It is found that the overall performance of En4D-Var is better than that of En3D-Var, indicating that the ensemble-based retrospective algorithm possesses more robust ability than ensemble-based sequential algorithm. The time variation of forecast-analysis spread statistics in En4D-Var is also plotted in Figure 3. At first two cycles, it is obvious that the perturbation spread is larger than RMSE. After the third cycle, however, the perturbation spread is a little less than RMSE, which means the filter divergence also exists in En4D-Var. The filter divergence seems more serious in humidity analysis (Fig. 3d) because the humidity perturbations are controlled within small amplitude in our experiments in case of over-saturation or negative humidity. It suggests a better humidity perturbation method should be explored in the future. Another reason for the humidity analysis divergence is humidity spatial and temporal scales are less and it is more sensitive to sampling errors. Adding inflation factor (Anderson 1999) can be expected to relax the filter divergence problem.

Acknowledgements:

This research has been supported by NOAA grant under contract of 05111076.

References

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

Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background error covariances: evaluation in a quasi-operation NWP setting. *Quart. J. Roy. Meteor. Soc.,* **131,** 1013-1043.

Countier, P., Thepaut, J. N, and A, Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.,* **120,** 1367-1387.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Montre Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99.** 10, 143-10,162.

Evensen, G., and P. J. van Leeuwen, 2000: An ensemble Kalman smoother for nonlinear dynamics. *Mon. Wea. Rev.*, **128**, 1852-1867.

Fertig, E. J., J. Harlim, and B. R. Hunt, 2007: A comparative study of 4D-VAR and a 4D Ensemble Filter: perfect model simulations with Lorenz-96. *Tellus*, **59A,** 96-100.

Gilbert, J. C., and C. Lemarechal, 1989: Some numerical experiments with variable storage quasi-Newton algorithm. *Mathematical Programming*.

Houtekamer, P. L., and H. L. Mitchell, 2001: A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Mon. Wea. Rev.,* **129,** 123-137.

Houtekamer, P. L., and H. L. Mitchell, 2005: Ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.,* **131**, 3269-3289.

Hunt, B. R., E. Kalnay, E. J. Kostelich, E. Ott, D. J. Patil, T. Saur, I. Szunyogh, J. A. Yorke, and A. V. Zimin,, 2004: Four-dimensional ensemble Kalman filtering. *Tellus*, **56A**, 273-277.

Liu, C., Q. Xiao, and B. Wang, 2008a: An Ensemble-based Four-dimensional Variational Data Assimilation Scheme. Part I: Technical Formulation and Preliminary Test. *Mon. Wea. Rev.*, in press.

Liu, C., Q. Xiao, and B. Wang, 2008b: An Ensemble-based Four-dimensional Variational Data Assimilation Scheme. Part II: Observing System Simulation Experiments with the

Advanced Research WRF (ARW). *Mon. Wea. Rev*., submitted.

Liu, Z., and F. Rabier 2003: The potential of high-density observations for numerical weather prediction: A study with simulated observations. *Quart. J. Roy. Meteor. Soc.,* **129,** 3013-3035

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP: a comparison with 4D-VAR. *Quart. J. Roy. Meteor. Soc.,* **129,** 3183-3203.

Skamarock W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the Advanced Research WRF version 2. NCAR Tech. Note NCAR/TN-468+STR, 88pp.

Simmons, A. J., and A. Hollingsworth. 2002: Some aspects of the improvement in skill of numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **128,** 647-677

Xiao, Q., X. Zou, M. Pondeca, M. A. Shapiro, and C. S. Velden, 2002: Impact of GMS-5 and GOES-9 satellite-derived winds on the prediction of a NORPEX extratropical cyclone. *Mon. Wea. Rev*., **130,** 507-528.

Zhang, H., J. Xue, S. Zhuang, G. Zhu, and Z. Zhu, 2004: GRAPeS 3D-Var data assimilation system ideal experiments. *Acta Meteorologica Sinica.,* **62,** No.1 31-41.

Zupanski, M., 2005: Maximum likelihood ensemble filter: theoretical aspects. *Mon. Wea. Rev.,* **133,** 1710-1726.