Background error statistics in 3DVAR derived from WRF ensembles

Mi-Seon Lee[1] and Dale M. Barker[2]

[1]*Korea Meteorological Administration, Sindaebang-Dong, Dongzak-Gu, Seoul, 156-720*

[2]*National Center for Atmospheric Research, P. O. Box 3000, Boulder, CO, 80307*

1. Introduction

A key avenue to improving data assimilation is an accurate specification of the error statistics for the background forecast, since it determines the filtering and propagation of observation information. Many current and past operational data assimilation method use long time series of previous forecasts to develop spatially homogeneous and temporally invariant approximations to background error statistics. The so-called "NMC method" (Parrish and Derber 1992) is widely used to calculate background error estimates in 3DVAR and 4DVAR systems. However a main shortcoming of the NMC method is likely to be evolution of the statistics of forecast error in the 12-24 (24-48 h) forecast ranges. For example, background error variances for winds tend to be over-estimated, and spatial correlation scales are too large (Fisher, 1999).

In the present work, we investigate the possibilities of using ensemble statistics to represent the background error in a limited-area 3DVAR application.

2. Experiments setup

WRF forecasts from the 1st to 10th of January 2002 are used over a continental US domain with 100km resolution. Sound, Synop, Airep, Pilot, Ship, Profiler, SATOB and METAR observations have been used in WRF-3DVAR with AVN background error statistics based on Wu *et al.* (2002). The differences between 12 and 24-h forecasts valid at the same time are used in the NMC method. In contrast, ensemble background error statistics use ensemble deviations (difference between forecasts and ensemble mean) to estimate the model error. The 50 ensemble members are arbitrarily numbered. Two different methods are used to generate ENS. The first set of statistics is based on perturbed observations (ENSo experiment). A second set uses perturbed observation as well as boundary perturbations (ENSb experiment) given by random control variable projections using 3DVAR's covariance model (Barker 2005). Boundary perturbations were designed to consider the boundaries effects in limited area mode and to maximize the ensemble spread.

The WRF 3DVAR control variable transform x'=Uv is in practice composed of a series of operations x'=UpUvUhv (Barker *et al.*, 2003) which should satisfy $B=UU^T$. Vertical (Uv) and horizontal (Uh) error covariance are represented by empirical orthogonal function (EOFs) and recursive filters respectively. Because the control variables in WRF 3DVAR are stream function, unbalanced velocity potential, unbalanced temperature, unbalanced surface pressure and pseudo RH (Q divided by background Qs), we need to estimate the regression coefficients to calculate the balanced part. These regression coefficients are defined using stream function as a predictor (Wu *et al.*, 2002). Therefore, Up means the variable transform processes from the control variables to physical variables (u, v, t etc.) using balanced coefficients and dynamic relationship. This new implementation (CV=5) in WRF 3DVAR are different to CV=3 in 2004 WRF 3DVAR version which uses the recursive filter both in horizontal and vertical.

3. Ensemble background error statistics

Fig. 1 shows the standard deviation of u and T forecast errors as a function of vertical level. These are calculated from 24-12 h forecast for NMC and ensemble deviation for ensemble statistics. The largest errors of winds (not shown here for v wind) are located at the jet level (sigma 16: around 300 hPa). The ENSo method estimates smaller errors than the NMC approach below the jet level, but similar errors in the upper troposphere. In this study, ENSb shows relatively larger errors in the whole troposphere especially in winds, due to the additional ensemble spread created by the perturbations to the lateral boundaries. For temperature, NMC and ENSb statistics indicate a very similar magnitude of errors except above 150 hPa. Two maxima of temperature background error are located around the top of the boundary layer (7 sigma – 850 hPa) and above the jet level (20 sigma : around 200 hPa).

The major reason of smaller variance in ENSo is that every members use same boundary condition, AVN analysis in this study, to generate ensemble forecasts except for initial boundary values and tendency.

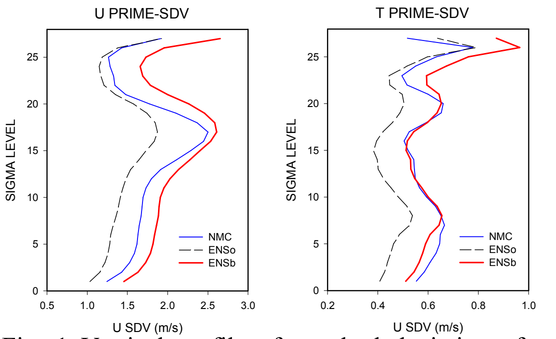

Fig. 1 Vertical profile of standard deviation of u wind (left) and temperature (right) background error.

Figure 2 indicates the derived correlations of unbalanced surface pressure in latitudinal index retrieved from NMC and ensemble statistics (ENSb). ENSb shows that the correlation of the unbalanced component correlated with stream function are increased in middle latitude. In other words, the regression analysis using ensemble data is less able to predict the surface pressure errors, than the regression using NMC-method data. This situation is very similar in velocity potential and temperature.

We performed single temperature observation tests with a temperature at a level of 850 hPa (11 sigma level) to see the detailed response from the new background error statistics and new background error model option (CV=5) in WRF-3DVAR with the recursive filter length scale, EOF, and balance coefficient. In this figure, NMC, ENSo and ENSb experiments use the new

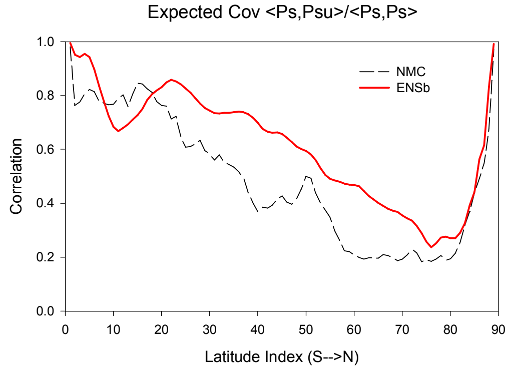

Fig. 2 The correlation of the "unbalanced" and total components of surface pressure according to the NMC and ENSb experiments.

background error statistics and CV=5 described in Section 2. Default CV3 has strong vertical autocorrelation than CV5 below sigma level 5 (940 hPa). ENSo shows smaller autocorrelation than NMC and CV3 which is corresponded by smaller scale length. ENSb shows similar response range to NMC, but vertical correlation was decreased. Actually we expected that ENSb's scale length to represent horizontal correlation will be smaller than NMC such as Fisher (1999). However, we found a similar or somewhat larger scale length than that using the NMC-method due to the use of 12 h forecasted ensembles. The true state of 12 h forecasted ensemble seems to be similar to 24-12 h forecast differences in NMC in this study.

4. Forecast Impacts

We generated 12 h forecast errors (against sonde observations) using WRF forecast and different option in WRF 3DVAR (CV=3 and CV5). The CV2 and CV3 error statistics are base on global (MM5 and AVN) forecasts, and hence require interpolation to the regional domain. Fig. 4 shows the performance from local and interpolated B (upper panel) and different B statistics (lower panel). First we found that the interpolated statistics included in WRF 3DVAR does not give better forecasted errors regardless of CV=2 (MM5) CV=3 (from AVN). Local background error statistics (CV=5) from WRF forecasts gives better forecast errors in both NMC and ENSb. However ENSo shows the increased forecast errors, especially in 12 hours cycling. In this study, ENSb gives the very similar errors to NMC.

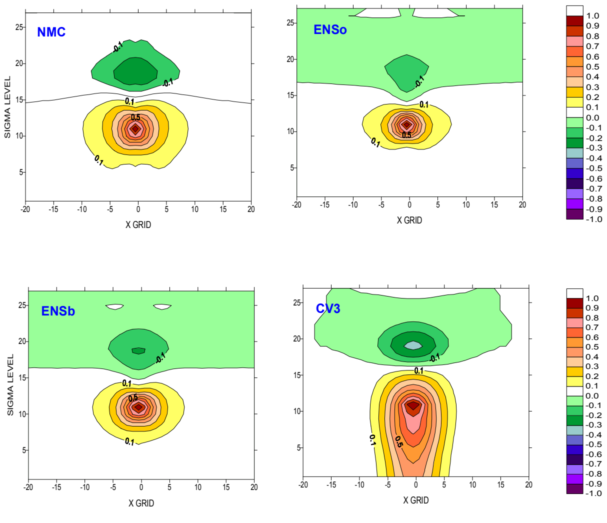

Fig. 3 Single observation test: Temperature response by a temperature around 850 hPa.

4. Summary

This study shows the possibilities of using ensemble forecasts to estimate background statistics of a limited area model. At this moment, ENSb shows similar skill to NMC. However, in short-range forecast ensembles (ex. 6hr), the impact of ensembles will be increased when we compare NMC using 12-24 (or 36-12) hour forecasts. Furthermore, the perturbation of lateral boundary conditions is effective in generating increased spread to depict forecast errors in a limited area model. However, suitable ensemble size and generation method of ensembles require further research.

Acknowledgments

This research was supported by the Post-doctoral Fellowship Program (2002) of Korea Science & Engineering Foundation (KOSEF), MMM Division Visitor Program, and the Air Force Weather Agency.

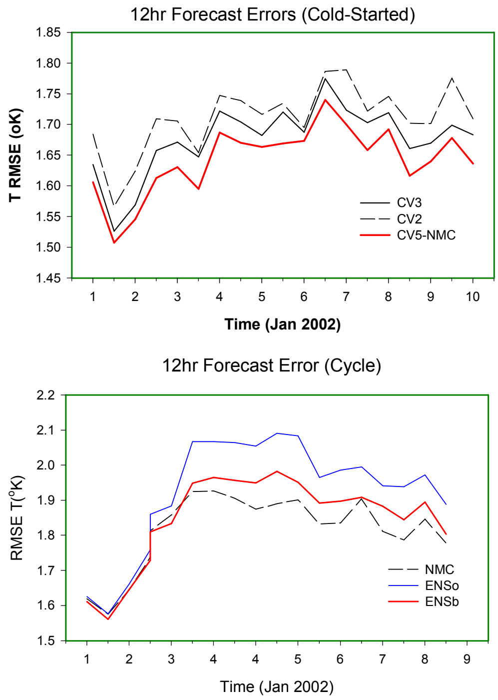

Fig. 4 12 hour forecast errors against sounding temperature. Upper panel shows the comparison between interpolated statistics (CV3 and CV2) from other models and WRF model statistics (CV5-NMC). Lower panel shows the forecast errors from NMC and ENS statistics.

References

Barker, D. M., Huang, W., Guo. Y-R., and Bourgeois, Al, 2003: A three-dimensional Variational (3DVAR) Data Assimilation System for use with MM5. NCAR Technical Note, NCAR/TN-453+STR, pp68.

Barker, D. M., 2005: High Southern-Latitude Ensemble Data Assimilation in the Antarctic Mesoscale Prediction System. *Mon. Wea. Rev.*, In Press.

Fisher, M., 1999: Background Error Statistics derived from an ensemble of analyses. ECMWF Research Department Technical Memorandum, 79, 12pp.

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

Wu W.-S, R. J. Purser, and D. F., Parrish, 2003: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev* **130**. 2905-2916.