# Comments on ''Short-Range Ensemble Forecasting of Explosive Australian East Coast Cyclogenesis''

THOMAS M. HAMILL

*Advanced Studies Program, National Center for Atmospheric Research, Boulder, Colorado*

21 April 1998 and 11 June 1998

Leslie and Speer (1998, henceforth LS98) examined a case of explosive Australian east coast cyclogenesis using a set of short-range ensemble forecasts. Their paper addresses a very interesting and relevant topic, namely, whether the predictability of explosive cyclogenesis and associated phenomena may be inferred from the spread produced by an ensemble of simulations. By varying the initial conditions in an interior nest, they produced 100 short-range ensemble forecasts (SREFs) of the event. Because the member forecasts were similar to each other and to a control forecast, they suggest that in this case the forecasters might have confidence in a forecast of potentially dangerous land gales.

Though I agree with the premise that short-range ensemble forecasting offers the *potential* to make such forecasts, and that this information would be very useful for this case, these conclusions must be regarded with skepticism. There are reasons other than greater than normal predictability which may explain the noted small dispersion. Specifically, this SREF experiment was conducted in a *small domain* using *identical lateral boundary conditions* for all members and with only one-way interaction between the interior nest and the coarse-resolution forecast. These may all suppress the realistic evolution of differences among ensemble members. Further, the method used to generate perturbations for the inner domain may also have contributed to limiting the differences among member forecasts.

SREFs are relatively new, and there is limited literature specific to the subject. The interested reader is referred to Brooks et al. (1992), Mullen and Baumhefner (1994), Stensrud and Fritsch (1994), Brooks et al. (1995), Hamill and Colucci (1997, 1998), Du et al. (1997), and Stensrud et al. (1998), and references therein. This literature is mixed, suggesting both the tremendous potential for improving probabilistic weather forecasts and notable deficiencies in early tests. Clearly, we all have much to learn.

The elusive goal of ensemble forecasting is to produce member forecasts that are realistic but that have uncorrelated (independent) errors. If member forecast errors are uncorrelated, the probability an event will happen can be inferred directly from the relative frequency of the ensemble. For example, if 70% of the member forecasts indicate land gales, then a 70% forecast of land gales is realistic. However, if member forecasts have correlated errors, then some or maybe all of the forecasts will have similar errors, and the dispersion among members will not realistically indicate the situational uncertainty. If one member is incorrect, most likely the others are incorrect.

Great care must be taken to ensure that the modeling approach is not constraining predictability error growth and contributing to correlated errors. To this end, SREF modelers must either use a sufficiently *large computational domain* and/or should supply *different lateral boundary conditions* (from an ensemble of coarse-mesh forecasts) to each member forecast in the innermost domain. When the domain is small and the coarse-mesh forecast supplying the lateral boundary conditions (LBCs) is imperfect, then both errors and reduced-resolution information transported across the LBCs can quickly ''sweep'' away the more accurate, higher-resolution simulation developing in the nest interior (Errico and Baumhefner 1987; Warner et al. 1997). Further, when the same LBCs are used for each member forecast as in LS98, whatever information propagates in from the LBC may be very similar among all members. Because of this, the dispersion among member forecasts was probably unrealistically small in LS98 and thus should not have been used as a predictor of the uncertainty in the forecast.

If LBCs do not differ realistically for each member forecast, then the domain must be enlarged enough so this sweeping effect does not seriously limit predictability error growth. A large domain is important for another reason, too: the forecast domain should be large enough to resolve the important scales of motion. A

*Corresponding author address:* Dr. Thomas M. Hamill, NCAR/MMM, P.O. Box 3000, Boulder, CO 80307-3000.
E-mail: hamill@ucar.edu

phenomenon like explosive cyclogenesis likely involves the interaction of energy at the planetary scale, the synoptic scale, and the mesoscale. Hence, the forecast model must faithfully represent interactions among these scales. This requires a sufficiently large computational domain and careful treatment of interactions with the coarse-mesh forecast (Warner et al. 1997). When a large domain is used, the synoptic-scale wavelengths are resolved in the fine-mesh simulation. Conversely, with a small domain, these wavelengths may not be resolved. Unless the LBCs are interactive, the nested simulation cannot interact with the larger-scale flow. In the LS98 simulation, the interior nested domain size was approximately $(2000\ \mathrm{km})^2$, centered at 33°S, and the nested forecast did not interact with the solution on the coarse-mesh grid. The largest wave that may thus be resolved in this computational domain is thus approximately wavenumber 16. The great majority of eddy kinetic energy is typically contained in scales larger than wavenumber 16 (Lee and Held 1993), with a peak typically around wavenumbers 5 to 6. Hence, the nested domain solution in LS98 excluded interactions with the most energy-containing scales of motion.

The methodology used by LS98 to generate perturbed initial conditions also may have contributed to producing an ensemble forecast with unrealistically small dispersion. There were two potential problems here: first, their technique *may* have produced unrealistically small perturbations. Second, the technique to generate these perturbations implicitly assumed that errors in the initial condition were related to the observations but unrelated to the atmospheric dynamics. Considering the first of these two points, perturbations in LS98 may have been unrealistically small because random, uncorrelated observational errors were added to each grid point for each member *before the objective analysis.* This typical error magnitude was consistent with the analysis error. However, the magnitudes of the perturbations may have been considerably less after the objective analysis and initialization, which smooths the initial, uncorrelated errors that were added to each grid point. Hence, the resulting perturbation magnitudes from the initialization process may have been much smaller than the analysis error, and the low dispersion noted in their subsequent forecasts a consequence of this initially small dispersion among initial conditions. The evidence provided in LS98 was not sufficient to determine whether their perturbation magnitudes for the initial condition (not the observations) were consistent with analysis uncertainty; the true initial condition may actually have laid outside the swarm of perturbed initial conditions.

The second potential problem is that perturbations were unrelated to the atmospheric dynamics. An objective analysis involves corrections of a first guess toward a set of observations. Hence, it may be important to understand the structure of errors in the first guess; for a chaotic dynamical system, it is likely that these structures will project onto growing modes, and that even after the assimilation of observations, there can be a significant projection of the analysis error onto these modes. Admittedly, the wisdom of using a perturbation scheme connected to the dynamics is an area of active debate (Anderson 1997). Mullen and Baumhefner (1989, 1994) suggest the use of ‘‘intelligent random perturbations,’’ which are unconnected to the dynamics, while two operational schemes (Molteni et al. 1996; Toth and Kalnay 1993, 1997) produce perturbations that are related to the atmospheric dynamics. A theoretically appealing approach is to run separate forecast and analysis cycles for all members, adding realistic, independent perturbations to the observations used in each member's analysis cycle (e.g., Houtekamer et al. 1996; Houtekamer and Mitchell 1998). Thus, the LS98 perturbation methodology may be a reasonable one, but the lack of a connection with the atmospheric dynamics and the possible use of unrealistically small perturbations gives reason for concern.

Other SREF experiments have been over larger domains and with different perturbation methodologies. Using 13 separate cases, this author examined the ability of a set of ensemble forecasts to ‘‘forecast the forecast skill’’ of precipitation (Hamill and Colucci 1998). We found that the magnitude of the deviation of the ensemble about its mean provided no useful evidence of the situational unpredictability. Instances with low dispersion typically had no lower error than instances with high dispersion. Using a larger set of cases, Stensrud et al. (1998) found no correlation between the spread of cyclone positions and the uncertainty of the forecast cyclone position. These studies do not invalidate the concept of SREF; rather, they illustrate that even in more carefully designed experiments a user cannot automatically expect reliable probability forecasts to be generated from uncorrected ensemble forecasts. We still need to address many theoretical and technical questions.

The temptation to run SREFs at the highest possible resolution is understandable, given the ample evidence of the benefits of increased resolution. However, as was summarized in Warner et al. (1997), the presumed benefits of high resolution are often not realized when small domains are used. This warning must be considered even more urgently when running ensembles of forecasts. So, what possible alternatives might LS98 have considered? Given that the authors conducted a 100-member simulation at 15-km resolution over a limited area, by reducing the horizontal resolution to 30 km, the domain area could have been enlarged by a factor of 8, and there is much evidence to suggest that 30-km resolution is adequate to predict explosive cyclogenesis. If the number of members were also reduced from 100 to 20, the domain area could have been 40 times as large. This simulation would have been on so large a computational domain that lateral boundary conditions would have been largely irrelevant and all important wavelengths could have been resolved.

REFERENCES

Anderson, J. L., 1997: The impact of dynamical constraints on the selection of initial conditions for ensemble predictions: Low-order perfect model results. *Mon. Wea. Rev.,* **125,** 2969–2983.

Brooks, H. E., C. A. Doswell, and R. A. Maddox, 1992: On the use of mesoscale and cloud-scale models in operational forecasting. *Wea. Forecasting,* **7,** 120–132.

——, M. S. Tracton, D. J. Stensrud, G. DiMego, and Z. Toth, 1995: Short-range ensemble forecasting: Report from a workshop (25–27 July 1994). *Bull. Amer. Meteor. Soc.,* **76,** 1617–1624.

Du, J., S. L. Mullen, and F. Sanders, 1997: Short-range ensemble forecasting of quantitative precipitation. *Mon. Wea. Rev.,* **125,** 2427–2459.

Errico, R., and D. P. Baumhefner, 1987: Predictability experiments using a high-resolution limited-area model. *Mon. Wea. Rev.,* **115,** 488–504.

Hamill, T. M., and S. J. Colucci, 1997: Verification of Eta–RSM short-range ensemble forecasts. *Mon. Wea. Rev.,* **125,** 1312–1327.

——, and ——, 1998: Evaluation of Eta–RSM ensemble probabilistic precipitation forecasts. *Mon. Wea. Rev.,* **126,** 711–724.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.,* **126,** 796–811.

——, L. LeFaivre, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.,* **124,** 1225–1242.

Lee, S., and I. M. Held, 1993: Baroclinic wave packets in models and observations. *J. Atmos. Sci.,* **50,** 1413–1428.

Leslie, L. M., and M. S. Speer, 1998: Short-range ensemble forecasting of explosive Australian east coast cyclogenesis. *Wea. Forecasting,* **13,** 822–832.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.,* **122,** 73–119.

Mullen, S. L., and D. P. Baumhefner, 1989: The impact of initial condition uncertainty on numerical simulations of large-scale explosive cyclogenesis. *Mon. Wea. Rev.,* **117,** 2800–2821.

——, and ——, 1994: Monte-Carlo simulations of explosive cyclogenesis. *Mon. Wea. Rev.,* **122,** 1548–1567.

Stensrud, D. J., and J. M. Fritsch, 1994: Mesoscale convective systems in weakly forced large-scale environments. Part III: Numerical simulations and implications for operational weather forecasting. *Mon. Wea. Rev.,* **122,** 2084–2104.

——, H. E. Brooks, J. Du, S. M. Tracton, and E. Rogers, 1999: Using ensembles for short-range forecasting. *Mon. Wea. Rev.,* in press.

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

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

Warner, T. T., R. A. Peterson, and R. E. Treadon, 1997: A tutorial on lateral boundary conditions as a basic and potentially serious limitation to numerical weather prediction. *Bull. Amer. Meteor. Soc.,* **78,** 2599–2617.