P4.13

DYNAMIC SELECTION FROM AMONG AN ENSEMBLE OF LATERAL BOUNDARY CONDITIONS FOR LIMITED-AREA MODELS

Paul A. Nutter *
University of Oklahoma, Norman, Oklahoma

1. Introduction

Life cycles of certain mesoscale phenomena are intimately connected to the larger-scale environment in which they are embedded. For example, vertical wind shear and thermodynamic profiles influence the character and motion of severe convection. Skillful predictions of mesoscale events using a limited-area model (LAM) requires accurate prediction of the larger scale flow.

It is well known that predictive skill is limited due to the natural growth of errors resulting from imperfect analyses. Specifically, predictability theories ascribe an inverse relationship between wave number and time limits of predictive skill (Lorenz, 1969). In contradiction, several predictability experiments using LAMs report little or no error growth resulting from perturbed initial conditions (Paegle et al., 1997, references therein).

Attempts to explain these optimistic results have focused either on enhanced local forcing (e.g., topography or surface inhomogeneities) or on the artificial errors introduced by the use of "one-way" lateral boundary conditions (LBCs). The latter effect is favored in the literature and is the subject of the current paper.

Vukicevic and Errico (1990) found that error growth in a mesoscale model occurs at fairly large scales. They reported that the range of scales that can interact with growing errors on the domain interior is constrained by the size of the LAM domain. Therefore, uncertainties possessing the largest spatial scale appear as LBC uncertainties in a LAM (Paegle et al., 1997). The evidence suggests that initial errors in the LAM would grow more freely if the LBC constraint was weakened (Vukicevic and Errico, 1990, pg 1467).

The goal of this work is to explore methods to weaken the LBC constraint. This will be accomplished by allowing the LAM model to dynamically select its LBC from among the range of equally likely possibilities provided by an external (global) model ensemble.

* *Corresponding author address*: Paul A. Nutter, School of Meteorology, EC 1310, 100 E. Boyd St., Norman, OK 73019; *e-mail: pnutter@ou.edu.*

2. Recent experience

In addition to the predictability issues outlined above, the research being proposed in this paper has implications for short-range ensemble forecasting. Furthermore, the methods to be investigated rely on a comparison of LAM and external model solutions across the entire LAM domain. Recent experience in these two areas are briefly highlighted below.

a. SREF

Recent short-range ensemble forecasting (SREF) experiments have shown that the ensembles often are under dispersive. That is, the verifying analysis does not fall within the range of possibilities forecast by the ensemble. Du and Tracton (1999) found that a regional ensemble with a larger domain produces greater spread than does an ensemble with a smaller domain, especially for those variables that were perturbed in the ICs. Furthermore, they found that the contribution to ensemble spread increases with time from LBC perturbations and decreases with time from the IC perturbations. These and other similar results (Hamill and Colucci, 1997; Hou et al., 2001; Stensrud et al., 2000) demonstrate that, with time, the spread of the LAM forecast ensemble becomes increasingly determined by the spread in the global ensemble as high frequency components are "swept" from the LAM domain (Vukicevic and Paegle, 1989).

b. LBC nudging and other schemes

Waldron et al. (1996) introduced a technique to reduce LBC errors by nudging the spectrally filtered long wavelength components of the LAM forecast towards the external model solution. Unlike more traditional Davies (Davies, 1983) nudging at the domain boundary, Waldron's scheme modifies the LAM solution across the entire domain. von Storch et al. (2000) later applied the method to dynamical downscaling for extended regional climate simulations. The underlying premise for this approach is that the single external (global) model forecast will predict the larger scale motions more accurately than would the LAM. However, it does nothing to account for uncertainty in the external model forecast.

Waldron's research, combined with the perturbation model configuration used by NCEP's Regional Spectral Model (Juang, 2000), highlights the importance of evaluating long-wavelength structures across the extent of the LAM domain when considering the LBC problem.

3. Idealized model experiments

It is clear that random and systematic components of both IC and model error growth are unavoidable in common LAM configurations. However, the specific effects of model error due to the LBC forcing alone can be isolated by generating a model simulated truth and running global and regional experiments with identical numerics.

An appropriate model choice is the barotropic vorticity grid point model for a mid-latitude beta channel. Emphasis lies with large scale mid-tropospheric flow since these are the patterns that are important for accurate placement of developing mesoscale and smaller features (Paegle et al., 1997).

At the time of writing, this model had just been coded, and results are not yet available. Thus, the emphasis of the current paper is on a discussion of planned experiments.

4. LAM ensemble configurations

Figure 1 is introduced as a schematic illustration of the parameter space defined by the range of available IC and LBC solutions. It is argued that traditional LAM configurations do not fully utilize this parameter space, and that the available information is more completely used by interfacing with a range of variable LBCs. Emphasis is placed on an ensemble configuration, but possible configurations for a single LAM forecast are revealed by considering individual ensemble members.

a. Many ICs, one LBC

Suppose an ensemble is constructed by providing a range of m equally likely ICs, yet only one LBC (Fig. 1a). Evidence suggests that in the absence of differences in model physical and numerical schemes (including grid spacing), every member of the ensemble might converge towards the single external solution. This is the essence of the

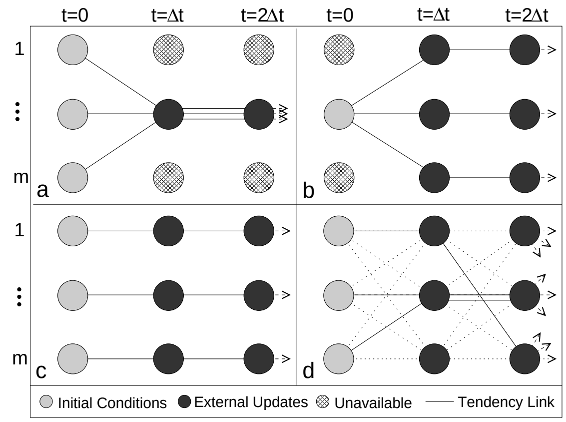

Figure 1: Schematic illustration of the IC/LBC parameter space available while configuring a LAM forecast ensemble system. Each ensemble simulation begins with m equally plausible ICs (light grey circles). LBCs are updated by regular data dumps from one or more external forecasts (dark grey circles) at time intervals Δt. At every time step of the LAM integration, LBCs are determined by computing linear tendencies between any pair of subsequent external updates (solid lines).

error growth constraint imposed by the one-way LBC forcing (Vukicevic and Paegle, 1989) which limits SREF dispersion (Du and Tracton, 1999).

b. One IC, many LBCs

Another possible LAM ensemble configuration is to provide a range of n unique boundary conditions to obtain solutions starting from a single initial state (Fig. 1b). Ensemble variance develops quickly since the LBC forcing is unique for each forecast. Given sufficient time, the boundary "sweeping" effect combined with nonlinear wave interactions within the LAM domain replaces the single solution with n different solutions. An ensemble configured as in Fig. 1b will attain greater dispersion than the previous case (Fig. 1a). However, the contribution to variance at smaller scales still is not represented well, a characteristic which largely motivates the use of LAM ensembles in the first place.

c. Parallel runs

The ongoing discussion (and common experience) suggests that an optimal ensemble should consist of both IC and LBC perturbations. The contribution to ensemble spread at the initial time reflects analysis uncertainty, and the growing variance in the LBCs reflects the uncertainty in larger scale external solutions. Furthermore, natural IC error growth and artificial LBC error growth interact nonlinearly to enhance the variance with time.

Existing SREF systems have been configured by setting $m = n$ and requiring that the initial assignment of a single LBC to a given analysis remains fixed for the duration of the forecast as shown in Fig. 1c.

The ensemble dispersion for the configuration shown in Fig. 1a can be calculated algebraicaly and compared with that for the configuration shown in Fig. 1c. Let

$$\bar{\mathbf{u}}^{(c)} = \frac{1}{n}\sum_{k=1}^{n} \mathbf{u}_k^{(c)} \tag{1}$$

be the ensemble mean defined for any time at each point on the nested domain, where the superscript (c) indicates that a single control forecast from the global model provides the LBC for all n LAM ensemble members. The ensemble dispersion is defined as (Stephenson and Doblas-Reyes, 2000)

$$\mathbf{D}^{(c)} = \frac{1}{n}\sum_{k=1}^{n} \left\| \mathbf{u}_k^{(c)} - \bar{\mathbf{u}}^{(c)} \right\|^2 \tag{2}$$

When each LAM forecast in the ensemble is forced by a unique LBC from a global ensemble, a difference vector $\mathbf{z}_k$ is introduced for every ensemble member $1 \leq k \leq n$ such that

$$\mathbf{z}_k = \mathbf{u}_k^{(e)} - \mathbf{u}_k^{(c)} \tag{3}$$

The superscript (e) denotes the LAM ensemble forced using an ensemble of LBCs. After taking the ensemble average of $\mathbf{z}$ and making appropriate substitutions, the dispersion for the LAM ensemble becomes

$$\begin{aligned} \mathbf{D}^{(e)} &= \frac{1}{n}\sum_{k=1}^{n} \left\| \mathbf{u}_k^{(e)} - \bar{\mathbf{u}}^{(e)} \right\|^2 \\ &= \frac{1}{n}\sum_{k=1}^{n} \left\| \mathbf{u}_k^{(c)} + \mathbf{z}_k - \bar{\mathbf{u}}^{(c)} - \bar{\mathbf{z}} \right\|^2 \qquad (4) \\ &= \frac{1}{n}\sum_{k=1}^{n} \left\| (\mathbf{u}_k^{(c)} - \bar{\mathbf{u}}^{(c)}) + (\mathbf{z}_k - \bar{\mathbf{z}}) \right\|^2 \end{aligned}$$

The square norm of a vector is the inner product of that vector with itself. Thus,

$$\mathbf{D}^{(e)} = \mathbf{D}^{(c)} + \frac{1}{n}\sum_{k=1}^{n} \|\mathbf{z}_k - \bar{\mathbf{z}}_k\|^2 + \frac{2}{n}\sum_{k=1}^{n} \left\langle \mathbf{u}_k^{(c)} - \bar{\mathbf{u}}^{(c)},\ \mathbf{z}_k - \bar{\mathbf{z}} \right\rangle \tag{5}$$

where the angle brackets denote the inner product. If the inner product term is small, then $\mathbf{D}^{(c)}$ is enhanced largely by the dispersion of the external model ensemble. If the inner product term is large, then the interaction between perturbatation growth on both domains enhances or dimishes the dispersion depending on the sign of the product. Since both models respond to nearly the same dynamics, the interaction terms should be positively correlated. If true, this result suggests that SREF systems must be configured using both IC perturbations and LBCs from the global ensemble to maximize the ensemble variance.

However, experience indicates that LAM ensembles configured this way remain under dispersive and scale deficient due to the suppression of error growth at the smallest scales (Du and Tracton, 1999).

d. Interdependent runs

Consider again the full ensemble configuration as in Fig. 1c, but relax the existing requirement that correspondence between the elements of each vector be fixed in time. Now, LBC tendencies may be computed between any pair of external updates as shown by dotted lines in Fig. 1d.

Suppose m equally likely ICs are provided for the LAM ensemble. Then, given a set of n choices for LBCs, a set of $m \times n$ forecasts are generated by interpolating in time between every IC and every LBC. Upon completing the first time interval, each of the $m \times n$ ensemble members could then be forced by tendencies obtained from any one of the n LBCs given at the second time interval. Thus, given T external model updates, $m \times n^T$ individual LAM ensemble members are generated. Of course, such an approach is impossible to implement in practice, but let us proceed in thought to uncover some (un)desirable properties.

Note that in the limit of continuous LBC updates ($\lim \Delta T \to 0$), the forecast trajectory evolves continuously without necessarily converging towards any particular global solution. Indeed, the trajectory through the $m \times n$ parameter space need not be linear.

Many of the integration paths through this parameter space may involve LBCs that are completely inappropriate. For example, trajectories through the parameter space in which the LBCs are completely out of phase with the LAM solution must be avoided.

Application of an interdependent modeling strategy requires careful selection from among the many available paths of integration. Such paths might be indicated, for example, by the solid lines in Fig. 1d. This is the focus of the current research, and options are considered in the following section.

5. Planned model configurations

Having established the scope of the problem, it is now appropriate to offer possible solutions that minimize the impact of LBC constraints while using methods that are practical in application.

a. Benchmark

A good starting point is to generate LAM ensembles using each of the configurations shown in Figs 1a-c. The results of such basic experiments would establish a benchmark for comparison and validate the expected outcomes summarized above. Of course, solutions using a two-way nesting configuration also will be generated for additional benchmark simulations.

b. Ensemble mean LBC

The mean forecast from global ensemble systems possesses greater skill than individual forecasts run at finer resolution (Toth and Kalnay, 1997). The reason for this is that the unpredictable components of the solution are averaged out, leaving only the structure of the predictable part of the flow. This result suggests that the skill of SREF might be improved by simply applying the global ensemble mean as the LBC.

However, such improvement is likely not possible since the effect would be the same as applying a single LBC to a set of different ICs as in Fig. 1a. Furthermore, smaller scale features would be averaged out and the LAM solution would become scale deficient as the larger scale features 'swept' the domain. Although not useful in practice, this method could be computed as another benchmark for comparison, especially for comparison against the control LBC.

c. Random selection

An unintelligent method of selection is to simply pick the tendency path randomly at each juncture in Fig. 1d. Such a method limits the number of ensemble members and produces another benchmark configuration, but is likely not useful in application.

d. Dynamic (active) selection

Now consider approaches that attempt to actively select from the range of possible LBCs as shown above in Fig. 1d. The idea is not to select the "best" LBC, but rather the one that is most consistent with growing errors across the nested domain at any given time. Indeed, it is not possible in practice to select the "best" LBC from among the global ensemble members because truth is unknown and each LBC is by design equally likely.

An objective measure is needed to determine which choice of LBC from the global ensemble is most consistent with the nested domain solution as the forecast proceeds. The distance measure needed to select consistent LBCs from the global ensemble should apply to the entire nested domain rather than just along a limited zone near the lateral boundary (Waldron et al., 1996).

Multivariate methods as suggested by (Stephenson and Doblas-Reyes, 2000) provide useful measures of distance. The covariance between the kth and lth members of an n member ensemble is given by

$$B_{kl} = \frac{1}{p}\sum_{i=1}^{p}[(X_{ki} - \overline{X}_i)(X_{li} - \overline{X}_i)] \qquad (6)$$

where X is the $n \times p$ data matrix and p is the number of points in the limited domain. The ensemble covariance matrix indicates which ensemble members are most similar. The diagonal elements measure the variance (in space) of each ensemble member about the ensemble mean. In application, the ensemble covariance matrix is expensive to calculate. For the purposes of choosing the global ensemble member(s) that provides the most consistent boundary conditions to the LAM forecast at any given time, it is simpler to compute only the distance between the current LAM forecast and each ensemble member. Moreover, this would be the approach for generating any individual LAM forecast. The appropriate distance metric may be computed as the mean-square difference between the kth and lth ensemble members:

$$D_{kl}^2 = B_{kk} + B_{ll} - 2B_{kl} = \frac{1}{p}\sum_{i=1}^{p}(X_{ki} - X_{li})^2 \qquad (7)$$

Thus, the domain mean-square error vector includes information about how each ensemble member varies about the ensemble mean and how each ensemble member fluctuates about the mean in correspondence with the regional forecast. If, in particular, l denotes the LAM forecast and $1 \leq k \leq n$ represents the members of the global ensemble, a p-element difference vector is easily computed as

$$\mathbf{d}_{kl}^2 = \frac{1}{p}(\mathbf{x}_k - \mathbf{x}_l)^T(\mathbf{x}_k - \mathbf{x}_l) \quad (8)$$

Other measures to be investigated include multivariate regression or eigenvector and multidimensional scaling. The latter consider issues of rank deficiency as described by Stephenson and Doblas-Reyes (2000).

Waldron et al. (1996) illustrates that spectral analysis should be considered for selective comparisons of LAM and external model solutions at longer wavelengths across the domain. Boer (1993) decomposed the vorticity equation into equations to diagnose systematic and random error components in an extended-range forecast. These methods likely will prove useful in the currently proposed research.

6. Summary

The goal of this work is to examine the interaction between IC error growth and artificial error growth caused by inaccurate specification of the larger scale environment through the LBCs. A properly constructed LAM ensemble should not restrict the error growth resulting from interactions between these sources of uncertainty. Once such interactions are quantified, methods will be explored that do not inhibit the error growth across the model interfaces by allowing solution trajectories to more completely utilize the parameter space defined by the set of ICs and LBCs. It is hoped that dynamic selection of uncertain LBCs will improve the accuracy of individual LAM forecasts which, in turn, should help enhance the dispersion of SREF.

References

Boer, G. J., 1993: Systematic and random error in an extended-range forecasting experiment. *Mon. Wea. Rev.*, **121**, 173–188.

Davies, H. C., 1983: Limitations of some common lateral boundary schemes used in regional nwp models. *Mon. Wea. Rev.*, **111**, 1002–1012.

Du, J. and M. S. Tracton, 1999: Impact of lateral boundary conditions on regional-model ensemble prediction. *Research activities in atmospheric and oceanic modeling, Report 28*, H. Ritchie, ed., WMO, TD-942, 6.7–6.8.

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

Hou, D., E. Kalnay, and K. K. Drogemeier, 2001: Objective verification of the samex '98 ensemble forecasts. *Mon. Wea. Rev.*, **129**, 73–91.

Juang, H. H., 2000: The ncep mesoscale spectral model: A revised version of the nonhydrostatic regional spectral model. *Mon. Wea. Rev.*, **128**, 2329–2362.

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289–307.

Paegle, J., Q. Yang, and M. Wang, 1997: Predictability in limited area and global models. *Meteorol. Atmos. Phys.*, **63**, 53–69.

Stensrud, D. J., J.-W. Bao, and T. T. Warner, 2000: Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 2077–2107.

Stephenson, D. B. and F. J. Doblas-Reyes, 2000: Statistical methods for interpreting monte carlo ensemble forecasts. *Tellus*, **52A**, 300–322.

Toth, Z. and E. Kalnay, 1997: Ensemble forecasting at ncep and the breeding method. *Mon. Wea. Rev.*, **125**, 3297–3319.

von Storch, H., H. Langenberg, and F. Feser, 2000: A spectral nudging technique for dynamical downscaling purposes. *Mon. Wea. Rev.*, **128**, 3664–3673.

Vukicevic, T. and R. M. Errico, 1990: The influence of artifical and physical factors upon predictability estimates using a complex limited-area model. *Mon. Wea. Rev.*, **118**, 1460–11482.

Vukicevic, T. and J. Paegle, 1989: The influence of one-way interacting lateral boundary conditions on predictability of flow in bounded numerical models. *Mon. Wea. Rev.*, **117**, 340–350.

Waldron, K. M., J. Paegle, and J. D. Horel, 1996: Sensitivity of a spectrally filtered and nudged limited-area model to outer model options. *Mon. Wea. Rev.*, **124**, 529–547.