21.5

# ENSEMBLE METHODS FOR SEASONAL LIMITED AREA FORECASTS

Raymond W. Arritt[1], Christopher J. Anderson[1], Eugene S.Takle[1,2], Zaitao Pan[1,16], William J. Gutowski, Jr.[1,2], Francis O. Otieno[2], Renato da Silva[15], Daniel Caya[8], Jens H.Christensen[4], Daniel Lüthi[12], Miguel A. Gaertner[13], Clemente Gallardo[13], Song-You Hong[10], Colin Jones[11], H.-M.H. Juang[5], J. J. Katzfey[3], William M. Lapenta[7], René Laprise[10], Jay W. Larson[14], Glen E. Liston[9], John L. McGregor[3], Roger A. Pielke, Sr.[9], John O. Roads[6], John A. Taylor[14]

[1]Department of Agronomy, Iowa State University, Ames, Iowa
[2]Department of Geological and Atmospheric Sciences, Iowa State University, Ames, Iowa
[3]Commonwealth Scientific and Industrial Research Organisation, Aspendale, Australia
[4]Danish Meteorological Institute, Copenhagen, Denmark
[5]National Centers for Environmental Prediction, Camp Springs, Maryland
[6]Scripps Institution of Oceanography, La Jolla, California
[7]Marshall Space Flight Center, Huntsville, Alabama
[8]Université du Quebec à Montréal, Canada
[9]Department of Atmospheric Science, Colorado State University, Ft. Collins, Colorado
[10]Department of Atmospheric Sciences, Yonsei University, Seoul, Korea
[11]Rossby Center at the Swedish Meteorological and Hydrological Institute, Norrköping, Sweden
[12]Swiss Federal Institute of Technology (ETH), Zurich, Switzerland
[13]Environmental Sciences Faculty, Universidad de Castilla-La Mancha, Toledo, Spain
[14]Mathematics and Computer Science Division, Argonne National Laboratory, Chicago, Illinois
[15]Department of Civil and Environmental Engineering, Duke University, Durham, North Carolina
[16]St. Louis University, St. Louis, Missouri

## 1 Introduction

Limits on computational resources usually have constrained nested regional climate model (RCM) studies to a single realization using a single model. In contrast, experience with both short-range limited-area predictions and medium to longer-term global modeling has shown the utility of ensemble methods (e.g., Brooks et al. 1995; Buizza et al. 1999; Atger 1999). Thus, it is appropriate to explore the utility of ensemble simulations within the context of RCM applications.

Appropriate methods for constructing RCM ensembles have not been established. The problem is made more complicated because the tendency for solutions to diverge in seasonal limited-area forecast models or nested regional climate models differs both from short-term limited-area models and from global seasonal models. Short-term applications of limited-area models are strongly dependent on initial conditions, but "memory" for initial conditions tends to decay for seasonal and longer simulations using limited-area models because of the continual input of data at the lateral boundaries (Giorgi and Bi 2001). Likewise, while global model solutions are free to diverge based on small differences in initial conditions or model physics, the divergence of solutions in limited-area models may be constrained by specified lateral boundary conditions. Thus, experience with short-range limited-area models and global models does not necessarily provide useful guidance for producing ensembles of seasonal forecasts using limited-area models.

In the present study we make an initial attempt to address ensemble prediction using nested regional climate models by comparing methods for generating ensemble simulations of seasonal precipitation. We use the summer 1993 flood (1 June - 31 July) over the north-central U.S. as a test case. This period corresponds to the PIRCS 1-B experiment, so that the PIRCS 1-B simulation suite provides a multi-model ensemble for comparison to other ensemble methods.

## 2 Methodology

Four methods are examined for creating the limited-area forecast:

- *Lagged-average ensemble:* The lagged-average ensemble technique appears to have first been discussed by Hoffman and Kalnay (1983) and has been used in various applications since its development (e.g., Molteni et al. 1986; Brankovic et al. 1990). In this technique each ensemble member is started at a different initial time with all simulations overlapping for the period of interest. The predicted fields at the start of the period of interest are viewed as differing but physically plausible initial conditions on this date. Results for this overlapping period are used as members of the ensemble with each model's preceding "spinup" portion discarded. Our lagged-average ensemble consists of eight instances of the MM5 mesoscale model executed with the same model configuration but with different starting dates. Simulations were begun at 00 UTC 15 May 2003 and at preceding times separated by 12-hour intervals.

- *Perturbed physics ensemble:* This method uses MM5 with a single set of physics options, but internal parameters within the convective parameterization are assigned different values in order to create different realizations for the ensemble.

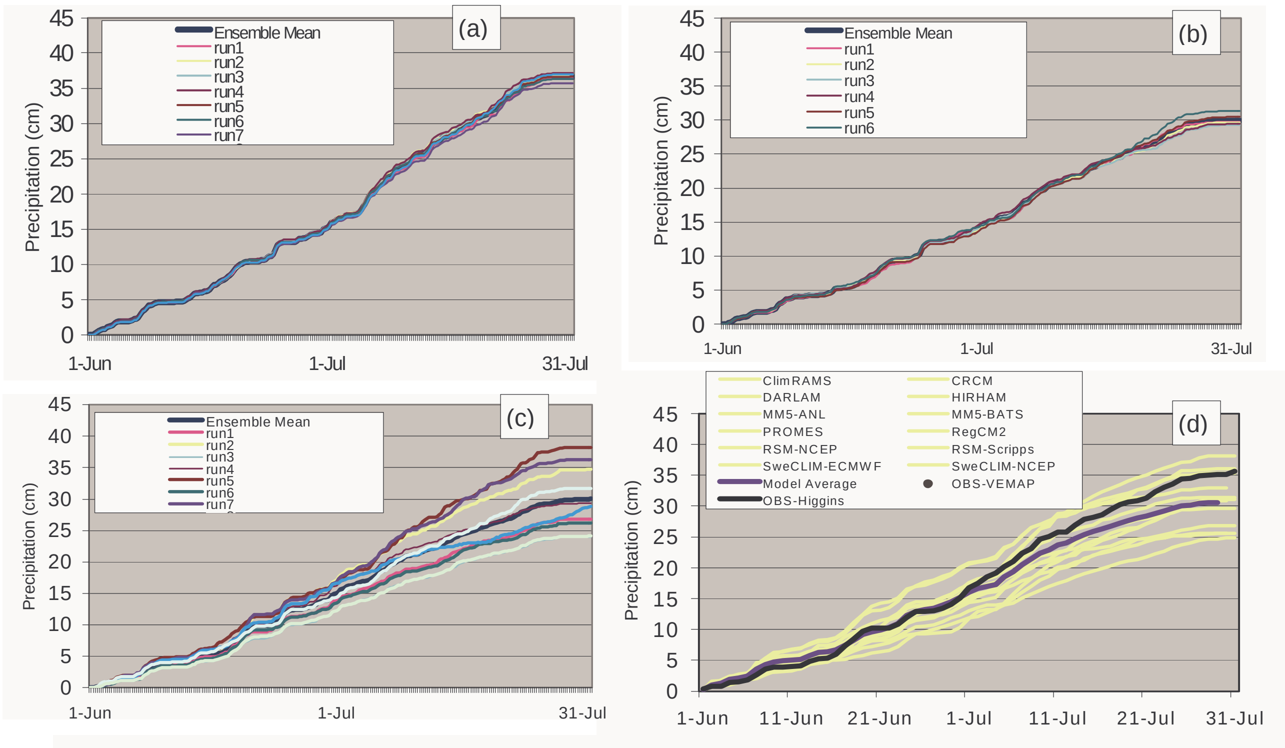


**Figure 1:** Cumulative simulated precipitation averaged over a region of the upper Mississippi River basin (37-47 $^{o}$N, 99-89 $^{o}$W).

All simulations use the same initial conditions. Here, our approach follows Yang and Arritt (2002) in which two parameters are varied within the Grell (1993) convective parameterization.

- *Mixed physics ensemble:* This method uses the same numerical model and initial conditions but differing schemes for physical parameterizations. We performed simulations with MM5 using a variety of moist physics options but the same initial conditions. The moist physics options included choices of either the Grell (1993) or Kain-Fritsch (Kain and Fritsch 1990) convective parameterizations, and choices of different explicit moisture schemes.
- *Multi-model ensemble:* In this approach distinct model simulations are used as individual realizations of an ensemble. Here we used models participating in the PIRCS 1-B experiment (Anderson et al. 2003).

## 3 Results

### *3.1 Area average precipitation*

Precipitation averaged over a portion of the upper Mississippi River basin is shown in Figure 1 for each ensemble member. We show results for individual members in order to illustrate the spread created by each ensemble method. It is immediately apparent that the lagged-average ensemble had very little spread. We interpret this result to reflect the strong control of lateral boundary conditions on the regional model solution, so that there was little sensitivity to initial conditions. The multi-model and mixed-physics ensembles had the largest spread. Notably, the spread obtained by using different moist physics parameterizations was about as large as the spread obtained by using completely different models.

### *3.2 Equitable threat score*

The usual definition of the equitable threat score (ETS) was employed, i.e.,

$$ETS = (H - C) / (F + O - H - C)$$

where H is the number of hits (correctly forecasted occurrences), F is the number of forecasts of the event in question that were forecast to occur (whether correct or incorrect), O is the number of occurrences of the event in question, and C is the number of correct forecasts that would be expected by chance. Here the event is taken as accumulated precipitation over the period 1 June - 31 July exceeding a specified threshold, and the number of occurrences is the number of gridpoints at which the threshold was exceeded. We considered thresholds from 200 to 600 mm at intervals of 25 mm. In evaluating each ensemble we took the simple arithmetic mean of all members to create an ensemble forecast, and evaluated categorical exceedence for this ensemble mean.

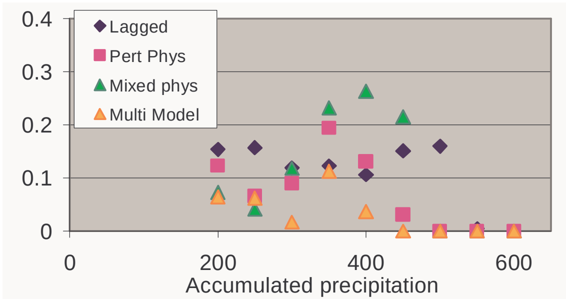


**Figure 2:** Equitable threat score (ETS) for each ensemble method at various precipitation thresholds.

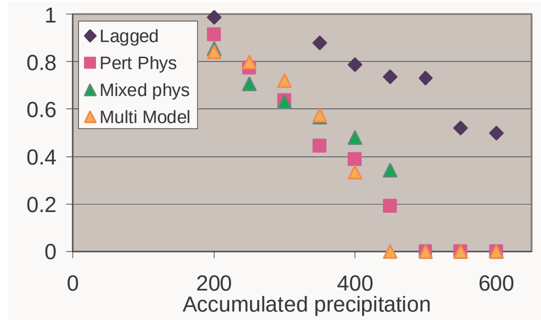


**Figure 3:** Bias for each ensemble method at various precipitation thresholds.

The method producing the highest ETS depended on the threshold (Figure 2). For thresholds below 250 mm the lagged ensemble method had the highest ETS, while for thresholds above 250 mm the mixed physics approach tended to produce the highest ETS. The multi-model ensemble tended to have low ETS. Inspection of the spatial pattern of precipitation (not shown) suggests that the low ETS for the multi-model ensemble can be attributed to the widely differing spatial patterns of the individual models, so that precipitation maxima were strongly smoothed in the ensemble mean. All of the methods yielded ETS of essentially zero for thresholds above 500 mm, reflecting the fact that there were very few hits at the highest thresholds.

### *3.3 Bias*

Using the same terminology as for ETS, the bias (B) is defined as

$$B = F / O$$

B ranges from 0 to infinity. While a perfect forecast system has B = 1, it does not follow that B = 1 implies that forecasts are accurate; e.g., in the context of ETS we could have F = O so that B = 1 and yet have no "hits" and zero ETS.

All of the methods produced low bias (i.e., B < 1) for all categories (Figure 3). The highest bias was produced by the lagged average ensemble. This is consistent with the small spread of the lagged average method; that is, since all the members were similar, averaging them together did not produce as great smoothing as for methods with larger spread. Therefore the lagged average method produced the greatest *number* of forecasts of category exceedence even though these forecasts were not necessarily *accurate*.

## 4 Conclusions

- The lagged ensemble method yielded very low spread. The implication of this result is that nested regional climate simulations have little sensitivity to initial conditions, at least for the case considered here, and are primarily boundary-value problems. The continual input of specified lateral boundary data causes the simulations to "forget" details of the initial conditions. This is in marked contrast both to global models, where the solutions are free to diverge, and to short-range limited-area forecasts, where the forecast time is short enough that the initial conditions are critical.
- Spread obtained by using different moist physics schemes within MM5 was about as great as the spread obtained by using completely different models.
- The method that produced the highest equitable threat score (ETS) depended on the specified threshold for accumulated precipitation. For the lower values the lagged ensemble produced the highest ETS although the mixed-physics ensemble was nearly as high. For higher precipitation values the mixed-physics ensemble

The similarity in spread between the mixed-physics and multi-model ensembles raises the possibility that a mixed-physics ensemble approach could be considered as a proxy for a multi-model ensemble. From a practical standpoint it is much more straightforward to produce a mixed-physics ensemble using a single model than to run multiple models each having its own data format, optimal computational platform and so forth. The mixed-physics ensemble also performed well in terms of equitable threat score, especially for higher precipitation amounts.

These results need to be evaluated further using more detailed statistical approaches (e.g., relative operating characteristic and decomposition of mean square error). We also need to develop better methods for creating ensembles than simply taking the mean of all realizations. Finally, it is essential to keep in mind that the results shown here are for only a single 60-day period, and a highly anomalous period at that. Despite this being a summertime case, the synoptic environment was quite active so that we might expect lateral boundary forcing to exert stronger control on the RCM solutions than in some other circumstances. Our findings need to be extended by considering multi-year simulations that capture different seasonal regimes and different interannual climate states (e.g., positive and negative phase of ENSO). Such simulations are presently underway as an extension of the PIRCS 1-C project.

*Acknowledgments.* This research was funded by the NOAA GAPP/PACS programs and by the National Science Foundation through grant ATM-9969650. Additional support was provided by Iowa Agriculture and Home Economics Experiment Station project 3803, supported by Hatch Act and State of Iowa funds.

## 5 References

Anderson, C.J., et al. 2003: Hydrological processes in regional climate model simulations of the central United States flood of June-July 1993. *J. Hydrometeorology*, in press.

Atger, F., 1999: The skill of ensemble prediction systems. *Mon. Wea. Rev.,* 127, 1941-1953.

Brankovic, C., T.N. Palmer, F. Molteni, S. Tibaldi and U. Cubasch, 1990: Extended-range predictions with ECMWF models: Time-lagged ensemble forecasting. Q.J.R. Meteorol. Soc., 116, 867-912.

Brooks, H.E., M.S. Tracton, D J. Stensrud, G.J. DiMego, and Z. Toth, 1995: Short-range ensemble forecasting (SREF): Report from a workshop. *Bull. Amer. Meteor. Soc.,* 76, 1617–1624.

Buizza, R., M. Miller and T. N. Palmer, 1999: Stochastic representation of model uncertainties in ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.,* 125, 2887-2908

Hoffman, R.N. and E. Kalnay, 1983: Lagged average forecasting, an alternative to Monte Carlo forecasting. Tellus, 35A, 100 - 118.

Giorgi, F. and X. Bi, 2000: A study of internal variability of a regional climate model. *J. Geophys. Res.* 105, 29503-29521.

Grell, G.A., 1993: Prognostic evaluations of assumptions used by a cumulus parameterization. *Mon. Wea. Rev.*, 121, 764-787.

Kain, J.S. and J.M. Fritsch, 1990: A one dimensional entraining/detraining plume model and its application in convective parameterization. J. Atmos. Sci. 47, 2784-2802.

Molteni, F., U. Cubasch, S. Tibaldi, 1986: Experimental monthly forecasts at ECMWF using the lagged-average forecasting technique. Proceedings of the first WMO Workshop on the diagnosis and prediction of monthly and seasonal atmospheric variations over the globe (College Park, USA, 29 Jul. - 2 Aug. 1985). WMO/TD no. 87, LRFRR no. 6, vol. 2, 598-607.

Yang, Z. and R.W. Arritt, 2002: Tests of a perturbed physics ensemble approach for regional climate modeling. *J. Climate*