Development and Evaluation of a New Initialization Technique for Weather Forecasting Ensembles

Principal Investigator: Dr. Isidora Jankov, Cooperative Institute for Research of the Atmosphere (CIRA), and Developmental Testbed Center (DTC)

BACKGROUND

The Developmental Testbed Center (DTC) is an interagency organization with a mission to serve as a bridge between the atmospheric science research and operational numerical weather prediction (NWP) communities. There is a clear need to test and evaluate the computing and scientific innovations developed within the NWP research community (i.e., in academia and at national laboratories), to identify the most promising methods for transfer into the operational NWP environment. A recent report from the National Research Council (NRC, 2006), and an ensuing initiative from the National Weather Service (NWS, Eckel et al., 2010), emphasizes the critical need for improved ensemble numerical weather prediction capabilities in the United States. Ensemble forecasting is a branch of NWP where not only one but a number of numerical forecasts are carried out with slightly modified initial conditions and model configurations allowing for an assessment of the uncertainty associated with a given forecast, facilitating much improved decision making. NWP, and especially ensemble forecasting, is a very computer intensive discipline and is the foundation of most weather forecasts. The DTC Ensemble Testbed (DET) was formed in 2010 in response to the ensemble forecasting initiative with the intent to thoroughly test and evaluate ensemble techniques developed at universities for transfer into operations. With this proposal, we seek to augment the very limited computational resources DET has at this time for its development of the infrastructure necessary for testing and evaluation of methods proposed by the academic community. The DET infrastructure will be comprised of the following modules (see Figure 1):

1. Ensemble configuration
2. Initial Condition and Lateral Boundary Condition (IC/LBC) Perturbations
3. Model Physics Perturbations
4. Statistical Post-processing
5. Products and Display
6. Verification and Evaluation

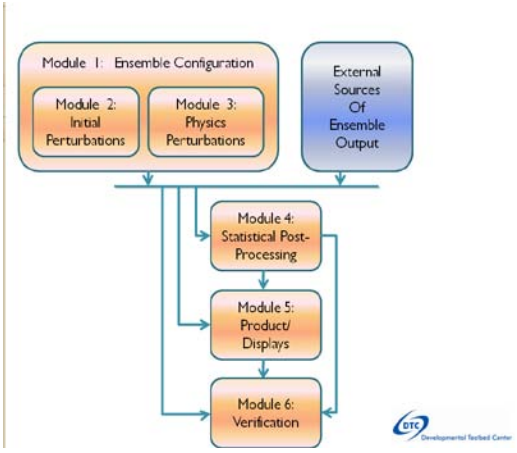

Figure 1. Overall design of DTC Ensemble Testbed testing and evaluation system. Arrows indicate process flow.

PROPOSED WORK

DET has partnered with National Oceanic and Atmospheric Administration (NOAA)/ National Center for Environmental Prediction (NCEP)/Environmental Modeling Center (EMC) to test the next generation Short-Range Ensemble Forecast system (SREF). DET's initial focus is on the Ensemble Configuration (Module 1) and IC/LBC Perturbations (Module 2). The definition of a perturbation for this application is a mathematical alteration of the best initial condition based on the uncertainties in the latest observations and short range NWP forecast. A multitude of perturbations are prepared and added to the best analysis to arrive at an ensemble of initial conditions consistent with plausible scenarios based on past and current observations. The perturbation technique under preliminary investigation is called "cycling of perturbations" and involves running a limited area ensemble initialized and driven by global ensemble forecasts. The perturbations of interest will be derived from the difference between the limited area ensemble mean and the individual ensemble members. Part of the initial investigation will be exploring the necessity of rescaling the perturbations and the optimal way of performing that operation. In other words, the perturbations will be centered over the chosen fine scale initial condition (or analysis) (e.g. NAM-GSI, LAPS/STMAS…) and these finer scale adjustments of the perturbations will be performed by using the fine scale analysis of choice. Rigorous testing and evaluation of the "cycling of perturbations" technique for transfer into operations requires a large number of simulations performed in a highly parallelized computing environment. The TeraGrid would be the ideal environment to perform these simulations.

COMPUTATIONAL NEEDS

Once optimization of the perturbation method has been achieved, thoroughly testing the impact of the new IC/LBC perturbations on operational forecast peformance will require at a minimum simulations from 21 ensemble members run for initialization times starting every 36 hours for an entire year of data. This represents just over 5,100 individual model runs extending out to 120 hours of simulated time. Additionally, five 6-hour runs will be performed at 6 hour increments during the intervening time, as part of the "cycling of perturbations" method, which is analogous to the "breeding method" introduced by Toth and Kalnay (1993).

The domain over which this test will be preformed covers the entire North American region at a 20 km grid spacing. The domain size is 427 by 401 gridpoints in the horizontal with 28 vertical levels. This represents 4,794,356 sets of weather-related calculations per time step. A timing run was performed on a DTC computational resource, the NOAA Jet System. Based on the timing information listed in Table 1, approximately 774,100 SU are needed to complete the entire project. Timing is broken out by steps and number of processors needed. As discussed below, the 200,000 SU start-up allocation requested here will help get the project started. Post-Processing and Verification are listed as serial processes to allow for reduction of gridded datasets and hence reduction of data transfer requirements. We are estimating the data transfer needs to be approximately 105 GB per 5-day run (and a fraction of that for a 6-hr run) for a total data transfer of approximately 255 TB for the 200,000 SU we will use. If needed, data-compression, such as converting the output files to Grib2 format, can be preformed to decrease required transfer rates. We will work the TeraGrid resource on optimal data storage procedures and do not anticipate needing any long-term storage at this time.

Table 1. Break-down of processes for individual model run for an individual ensemble member.

Step	Process	RunTime (sec)	# processors	SU (core hours)	Number of Model runs	Total SU per process
Preprocess	Ungrib	1885	1	0.52	5100	2,670
	Metgrid	345	1	0.10		510
	Real	570	1	0.16		807
5-day Model Run	WRF run	13300	32	118.22		602,933
Post-Processing	LFM Post	3600	3	3		15,300
Verification	MET	900	1	0.25		1,275
6-hr Model Runs for cycling	WRF run	665	32	5.91	25500	150,705
					Total SU needed	**774,100**

Key: Ungrib – reads Grib data and writes in intermediary NetCDF format; Metgrid – horizontally interpolates initial conditions to model grid; Real – reads data from namelist, allocates space for domain, generates initial condition file; WRF – non-hydrostatic numerical weather prediction model; LFM Post – post-processor to interpolate to constant height an pressure surfaces and derive additional variables; MET – verification package for numerical weather prediction models.

ALLOCATION REQUEST

Our proposal for a Start-Up allocation on the TeraGrid covers the initial set-up and preliminary tests of the ”cycling of perturbations” ensemble initialization technique. We anticipate using the knowledge gained through the Start-Up allocation will be used to prepare a one-year or multi-year Research Allocation proposal for submission in mid-2011. That proposal will include completion of the initial DET project as well as testing of newly developed methods on much finer spatial scales (4 km resolution). If granted, we would be ready to begin testing 15 December 2010. Our preferred platform is the TACC Ranger with post-processing and model verification on TACC Longhorn. These systems have the same batch system as the DTC resources and would expedite our initial investigation.

REFERENCES

NRC, 2006: COMPLETING THE FORECAST: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts, Committee on Estimating and Communicating Uncertainty in Weather and Climate Forecasts, Board on Atmospheric Sciences and Climate, Division on Earth and Life Studies, NATIONAL RESEARCH COUNCIL OF THE NATIONAL ACADEMIES, THE NATIONAL ACADEMIES PRESS, Washington, D.C., pp 112. www.nap.edu.

Eckel, F. A., Glahn, H.R., Hamill, T.M., Joslyn, S., Lapenta, W.M., Mass, C.F., 2010: National Mesoscale Probability Prediction: Status and the Way Forward. A White-Paper Report from the National Workshop on Mesoscale Probabilistic Prediction, 23-24 September 2009. NWS Office of Science and Technology, Washington D.C., May 28, 2010. pp. 28. http://www.weather.gov/ost/NMPP_white_paper_28May10_with%20sig%20page.pdf.

Toth, Z., and Kalnay, E., 1993: Ensemble Forecasting at the NMC: The generation of perturbations. Bull. Amer. Meteorol. Soc., 74, 2317-2330.