The DTC Ensemble Testbed: A New Testing and Evaluation Facility for Mesoscale Ensembles

Zoltan Toth[1, 2], Barbara G. Brown[1, 3], Isidora Jankov[1, 2, 4], Tara L. Jensen[1, 3], Huiling Yuan[5], Edward I. Tollerud[1, 2], Linda S. Wharton[1, 2], Paula T. McCaslin[1, 2], Bill Kuo[1, 3], Louisa Nance[1, 3], and Steven E. Koch[1, 2]

[1]*Developmental Testbed Center, Boulder, Colorado*
[2]*NOAA Earth System Research Laboratory (ESRL), Boulder, Colorado*
[3]*Research Application Laboratory, National Center for Atmospheric Research (NCAR), Boulder, Colorado*
[4]*Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado*
[5]*School of Atmospheric Science, Nanjing University, Nanjing, Jiangsu, P.R. China*

1. Introduction

As operational centers move towards ensemble-based probabilistic forecasting, the Developmental Testbed Center (DTC) has been asked to expand its efforts by providing a testbed platform to serve as a bridge between research and operations. Hence, the DTC has established the DTC Ensemble Testbed (DET). The goal of DET is to *provide an environment in which extensive testing and evaluation of ensemble-related techniques can be conducted such that the results are immediately relevant to the operational centers (e.g., NCEP and AFWA)*. DET activities will involve supporting and maintaining community codes, as well as conducting extensive testing and evaluation of promising new capabilities and techniques that have been incorporated into these community codes. Community codes already supported by the DTC will serve as building blocks for the end-to-end ensemble testing and evaluation system to be assembled. Inclusion of research techniques targeted for upcoming operational implementation will ensure the DET ensemble system does not lag behind the operational capabilities and will allow the DTC to contribute to future operational decisions.

To truly act as a bridge between research and operations, the DET has defined a process that allows the community to provide input into the process, and has engaged a WRF ensemble modeling working group (EMWG) to provide input to the DET planning activities. The initial outcome of this collaboration is an initial plan for the structure and function of the DET. This paper will discuss this conceptual structure, the principal priorities of the DET, its linkages with other DTC activities, and some preliminary results and anticipated accomplishments.

2. DET Infrastructure

Given the DET requirement to facilitate testing and evaluation of competing techniques and capabilities for specific components of the ensemble system, the DET infrastructure is designed to be modular. In order to keep the testing and evaluation results of new ensemble capabilities developed by the research community relevant to operational upgrade decisions, the DET modules will be configured such that they are able to replicate the algorithms used by the operational centers. Modules to be included in the infrastructure are:

- *Ensemble configuration*: Defines membership and horizontal/vertical resolution of members, such that different models and/or different configurations of the same model can be included.

- *Initial perturbations*: Provides the ability to represent uncertainty in initial conditions based on a variety of techniques.

- *Model perturbations*: Provides the ability to represent model-related uncertainty based on a variety of techniques.

- *Statistical post-processing*: Provides ability to specify techniques for fusing information from ensemble and high resolution control forecasts, climatology, and other sources such as the latest set of observations; Bias

correct / calibrate forecast distribution; Statistically downscale information to user relevant variables.

- *Product generation*: Provides ability to specify technique for deriving information from the ensemble, generating probabilistic products, providing decision support services, etc.

- *Verification*: Provides the ability to specify techniques to be used to evaluate ensemble and derived probabilistic forecasts.

Implementing the modules that are the key pieces of the DET infrastructure involves two important steps:

- Establish an initial basic capability: This step refers to establishing the capability to implement a promising new technique appropriate to the module.
- Establish a benchmark: This step refers to establishing the capability to functionally reproduce a current operational technique appropriate to the module.

Figure 1 presents a schematic illustration of the interrelationships between these modules.

Once a new capability and its respective benchmark have been established, the DET has the basic infrastructure necessary to conduct testing and evaluation directed at assessing whether the new capability shows promise for improving performance over the operational benchmark.

The DET will be starting AOP 2011 with detailed plans for the overall DET infrastructure (including all modules) and a test and evaluation protocol, as well as basic capabilities for the following modules: initial perturbations and verification modules. The infrastructure plans relied heavily on input from the ensemble community obtained through the DET Workshop held in August 2010 and interactions with the reconstituted WRF Ensemble Working Group, as well as regular meetings with colleagues at EMC directed at ensuring DET design, software development, and testing is well coordinated with a clear path to influence NCEP operations, starting with NCEP's next upgrade of the operational Short Range Ensemble Forecast (SREF) system.

2.1 Overarching design

For AOP 2011, the DET will start development and testing of "DET Portal", an interactive interface that in its advanced form can be used for setting up, running, monitoring and evaluating DET ensemble experiments. The ensemble configuration and initial perturbation modules will be the first modules to be incorporated into the DET Portal. With the gradual development of the DET Portal, the use of the DET infrastructure by DET and visiting scientists will become much easier.

2.2 Ensemble Configuration Module

For AOP 2011, the DET plans to establish a basic capability for the ensemble configuration module. Building on the HMT ensemble experience during AOP 2010, the DET will incorporate the capability of running various models within the NEMS and WRF frameworks into the DET infrastructure. These models will include the ARW and NMM in the WRF infrastructure and the NMMB in the NEMS framework. The DET will balance requirements regarding NCEP implementations (need to work in NEMS framework) and the community's continued interest to work within the WRF framework. Specific arrangements for the testing of stochastic perturbation / physics schemes will be taken into consideration when establishing this module.

2.3 Initial Perturbations Module

For AOP 2011, the DET plans to contribute to the next upgrade of NCEP's SREF system through testing and evaluation of the "cycling of perturbations" method that NCEP is considering for its next upgrade. A simplified version of this technique was introduced as an initial basic capability implemented for this module during AOP 2010. The technique is equivalent to dynamical downscaling of the operational global ensemble that currently uses Ensemble Transform with Rescaling (ETR) perturbations. The cycling perturbations method, along with the simple interpolation of the operational global initial ensemble perturbations currently used in the operational SREF system will be fully implemented into the DET infrastructure. Once both capabilities are incorporated into the infrastructure, head-to-head tests between these two initial perturbation techniques will be run over an extended time period and a report on the evaluation of these results will be prepared and delivered to NCEP. In addition to the report, the

DET will make the tested software available to NCEP for further testing and operational implementation. If this new initial perturbation method is implemented by EMC, this approach will become the benchmark for the initial perturbation module. The report from this test will also be posted on the DTC website so the NWP community will have access to the results.

2.4 Product Generation Module

At the request of HWT participants, and through the ingenuity of CAPS research, many ensemble products were generated during the 2010 SE. Some of these products proved useful subjectively, others did not. The DTC evaluated a subset of these products for a particular forecast variable, but did not evaluate all products for a variable. Additionally, only products generated in real-time were evaluated. Objective evaluation of single-value products (i.e., simple ensemble mean and probability matched mean), as well as the probabilistic products (simple probability and neighborhood probability) for composite reflectivity, QPF, and radar echo top height, showed that each product provided a different level of skill for each field. This result suggests an ensemble, and its products, cannot be evaluated by one method and variable alone.

In support of the Ensemble task area, the DTC plans to accelerate its implementation of the DET product generation module through its collaboration with HWT. The approach will be twofold:

i) Identification of Promising Techniques

The DTC will evaluate in real-time and prepare a report on an evaluation of all HWT ensemble products (both single-value and probabilistic) for composite reflectivity, accumulated precipitation, and, if available, a selected synthetic satellite product or explicit lightning product. Ensemble mean, spread, and probabilistic guidance will be compared to the SREF (Short Range Ensemble Forecast) and the new HREF (High-res Short Range Ensemble Forecast) products provided by NCEP. In brief, the HREF combines output from the NMM and ARW runs of the Hires Window with SREF output to produce a 44 member "ensemble". The DTC will work with EMC and HWT to determine whether any other ensemble products from NCEP should be included in this evaluation, (i.e., the VSREF or Very Short Range Ensemble Forecast).

ii) Begin incorporation of promising techniques into Product Generation Module

The DTC will work with EMC and CAPS to acquire some of their ensemble product algorithms for inclusion in the DET product generation module, thus, establishing an initial basic capability for this module that can undergo further testing by the DET. Resources permitting, results may also be produced and made available to the community related to the optimal choice for the number of ensemble members.

2.5 Verification Module

The DET will apply the DET verification module, which is comprised of ensemble verification capabilities developed during AOP 2010 through the HMT collaboration, to its testing and evaluation activities. In addition, verification metric packages, which consist of a set of forecast variables, forecast levels, and statistical metrics to be used for several forecast and evaluation challenges (i.e. operational model evaluation, aviation weather forecasts, etc…) will be established. These verification metric packages can be considered parallel to DTC RCs and will be defined through discussions with major operational users (EMC, FAA, OHD, etc). Experimental DET runs for the initial perturbation module will be evaluated using a defined EMC verification metric package. Additionally, HWT and HMT ensembles may be evaluated beyond the testbed collaboration efforts to further understanding and interactions with these organizations. Resources permitting, the DET verification module will also be used for the evaluation of some HFIP regional ensemble experiments.

2.6 Computing Resources

From the start, the DET has been faced with a major challenge in terms of identifying computational resources for its efforts to construct and test its infrastructure and test and evaluate new methods developed by the community. During AOP 2010, the DET prepared a request and was granted a startup allocation on the Teragrid computational network supported by NSF. This allocation will allow the DET to test the portability of its infrastructure and run a limited amount of tests in support of its AOP 2011 work. The DET also prepared a proposal to HFIP during AOP 2010 requesting computational resources in support of testing the next version of the SREF system. For

this testing, the SREF domain would be extended to encompass more of the Atlantic Ocean. The proposal was favorably received by HFIP management and the request for computing resources was granted. This work would benefit HFIP by providing potentially useful guidance on tracks and boundary conditions for nested hurricane ensemble forecasts. Through this allocation, the DET will be able to perform extensive testing during AOP 2011 in support of NCEP's next SREF implementation.

2.7 Community Interactions

The DET plans to maintain its critical connections with the ensemble developer and user community by collaborating with NCEP and the National Unified Operational Prediction Capability (NUOPC) on the organization of the 5th Ensemble User Workshop to be held in the Washington, DC area in spring 2011. This workshop will provide important input to the planning and requirements for the statistical post-processing, product generation, and verification modules of DET infrastructure and will further strengthen the DET's interactions with NCEP.

Note that additional funding, if available, would allow DET to accelerate benchmark development and testing and evaluation associated with the model perturbations module and basic capabilities for the statistical post-processing module, planned at current level of funding for AOP 2012, and to carry out real-time testing of benchmarks. The DET will benefit from ensemble verification capabilities added to MET through work reflected under the verification

task and capabilities implemented through the DTC's collaborations with HWT and HMT. The DTC also anticipates the ensemble modeling efforts for 2011-2012 HMT field exercise will make use of the DET infrastructure for its real-time forecast system.
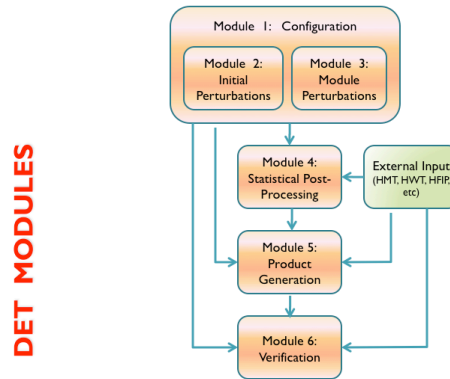

Fig. 1. Schematic Depiction of DET Infrastructure.

5. Future Plans

Table 1 provides an initial set of deliverables scheduled the fiscal years indicated. Note that a schedule for achievement of benchmarks for each functional module is shown, starting with modules 3 and 2 for the AOP for 2011. Many of these benchmarks will focus on NCEP requirements and utilities.

- AOP10 deliverables – Planning, basic capabilities
  - Detailed plan for infrastructure (6 modules)
  - Basic capability for modules 1-2 & 6
  - Protocol for T&E
  - Plan for & commence collaboration with HMT
- AOP11 deliverables – Basic capabilities, benchmarks
  - Basic capability for module 3
  - Benchmark for modules 2
  - T&E for module 2
  - Contribute to NCEP implementation for modules 2 (FY11-12)
  - Contribute to design and evaluation of HMT, HFIP, & HWT ensembles
- AOP12 deliverables – Benchmarks, T&E
  - Benchmark for module 3
  - T&E for module 3
  - Basic capability for modules 4-5
  - TET of community methods for module 2
- AOP13 deliverables
  - Benchmark for modules 4 & 6
  - Contribute to NCEP implementation for module 4 (FY13-14)
  - TET of community methods for all modules
  - Update benchmarks

Table 1. Deliverables