DTC Annual Operating Plan 2011

1. Introduction

The Development Testbed Center (DTC) is a distributed facility with components residing in the Joint Numerical Testbed (JNT) of the National Center for Atmospheric Research's (NCAR) Research Application Laboratory (RAL) and the Global Systems Division (GSD) of the National Oceanic and Atmospheric Administration's (NOAA) Earth System Research Laboratory (ESRL). The objectives of the DTC are to: (i) advance science research by providing the research community an environment that is functionally similar to that used in operations to test and evaluate the components of state-of-the-art Numerical Weather Prediction (NWP) systems supported by the DTC, without interfering with actual day-to-day operations; (ii) reduce the average time required to implement promising codes emerging from the research community by performing the early steps of testing to demonstrate the potential of new science and technologies for possible use in operations; (iii) sustain scientific interoperability of the community modeling system; (iv) manage, provide and support the common baseline of end-to-end community software to users, including dynamic cores, physics and data assimilation codes, pre- and post-processors and codes that support ensemble forecasting systems; and (v) establish, maintain and support a community statistical verification and validation system for use by the broad NWP community. The fundamental purpose of the DTC is to serve as a coordinating mechanism that acts as a bridge between research and operations thereby facilitating the activities of both in pursuit of their own objectives. For the DTC's AOP 2011, the National Weather Service (NWS), the Office of Oceanic and Atmospheric Research (OAR), the Air Force Weather Agency (AFWA), NCAR, and the National Science Foundation (NSF) collectively provided approximately $5.3M to support the activities of the DTC. The breakdown of funding by sponsor is summarized in Table 1. This document provides a description of the tasks that will be carried out by the DTC during this performance period, without differentiating the funding sources for each specific task. The breakdown of the funding by task is summarized in Table 2. Most tasks are carried out jointly by JNT and GSD staff.

Table 1: Breakdown of DTC funding by sponsor

Source	Funding (in K)	Source	Funding (in K)
OAR	2,994	NSF	100
AFWA	718	HFIP	708
NCAR	250	Allocated NCAR carryover	68
USWRP	281	Allocated GSD carryover	61
GSD	250		

Table 2: Breakdown of DTC funding by task area

Task	Funding (in K)

Director's Office	755
Visitor Program	200
Mesoscale Modeling	1,037
Hurricanes	1,197
Data Assimilation	772
Verification	623
Ensembles	846

2. DTC Director's Office

The DTC Director's Office provides administrative and management support for all DTC activities. This support includes: (i) managing and coordinating all DTC tasks, (ii) conducting DTC workshops and tutorials, (iii) interacting with DTC partners (e.g., NOAA's National Centers for Environmental Prediction (NCEP) Environmental Modeling Center (EMC), the Mesoscale and Microscale Meteorology (MMM) division of the NCAR Earth System Laboratory (NESL), and AFWA) on collaborative efforts, (iv) creating and maintaining the DTC web site, and (v) providing administrative support for DTC management meetings. In addition to these basic support activities, the DTC Director's Office hosts the DTC Visitor Program.

Managing and Coordinating DTC Tasks

Due to the distributed nature of the DTC, the coordination and oversight of DTC tasks requires more effort on the part of the management team. On the other hand, the benefits of having a distributed facility (e.g., ability to maintain stronger connections to a broader intersection of the research community through daily interactions at two institutions rather than one, expanded access to computing resources, etc.) far exceed this slight increase in management overhead. This increase in overhead has also been reduced by establishing a structure where activities are managed in terms of tasks. The DTC currently has five major task areas: Mesoscale Modeling, Hurricanes, Data Assimilation, Ensembles, and Verification. Each task has an assigned lead responsible for developing the task work plan and ensuring the successful execution of the work plan. Hence, management's responsibilities are mainly related to coordinating the allocation of resources and staffing at the two nodes, collecting information from the leads to produce planning and reporting documents, resolving staffing issues when conflicting deadlines arise, identifying commonalities among the tasks that need broader coordination, and longer term planning. In addition to the five task areas, the DTC also collaborates with the Hazardous Weather Testbed (HWT) and the Hydrometeorological Testbed (HMT) to test and evaluate current and future mesoscale NWP systems. These testbed collaboration projects contribute to several of the DTC task areas.

DTC Workshops and Tutorials

The DTC conducts or assists with several workshops and tutorials each year. These events provide venues for strengthening the DTC connection with the broader community. For AOP 2011, the DTC will be involved in the annual WRF Users Workshop, a physics workshop and the 5th Ensemble User Workshop. Tutorials the DTC will be involved in are the bi-annual basic

WRF and MET Tutorials and annual WRF for Hurricanes and Gridpoint Statistical Interpolation (GSI) Data Assimilation (DA) System Tutorials. More information on these events is provided in the relevant sections below. Management involvement in the planning and conducting of the on-going workshops and tutorials is generally fairly limited because the members of the relevant task team are responsible for these activities. On the other hand, the administrative staff is a critical part of this process.

Interacting with DTC Partners on collaborative efforts

Maintaining a strong connection with the DTC's collaborative partners is critical to the success of the DTC given its role of providing a bridge between the research and operational communities. Since the DTC is not involved in the development of new NWP capabilities and techniques, the DTC needs to maintain a close working relationship with the developers of the packages it is tasked with supporting to the community. The success of community code activities relies on a strong connection with these development groups. Testing and evaluation activities also need to be planned such that their results will be relevant to both the research and operational communities.

DTC Website

The DTC website is an important communications tool for sharing information about the DTC and its accomplishments. Keeping information current in a quickly changing environment is challenging. During AOP 2011, the DTC plans to work towards filling the remaining gaps on its website and implementing procedures geared toward keeping the information on the site current. Achieving this goal will require participation from all staff members with some direction from management.

DTC Visitor Program

The DTC Visitor Program supports visitors to work with the DTC to test new forecasting and verification techniques, models, model components, and data assimilation approaches for NWP. It also offers an opportunity for visitors to introduce new NWP and verification techniques into the community codes supported by the DTC that would be of particular interest to the research community. The goal of this program is to provide the operational weather prediction centers (i.e., NCEP and AFWA) with options for near term advances in operational weather forecasting and to provide researchers with NWP codes that represent the latest advances in technology. The DTC has strived to release an annual Announcement of Opportunity (AO) for scientists to submit proposals for projects to be funded through this program. In the past, the AO has provided broad guidance on what types of projects would be of interest to the DTC, as well as a more detailed and quite lengthy list of suggested areas. The DTC Management Board (MB) is tasked with providing recommendations on the overall direction and areas of emphasis for this program; whereas the DTC Science Advisory Board (SAB) is tasked with reviewing the submitted proposals.

During AOP 2011, the DTC will be supporting six DTC visitor projects that were selected in September 2010. The DTC also plans to distribute an AO for the next round of visitor projects in September 2011. Before distributing this AO, the DTC will review its procedure for developing the AO and make adjustments that will increase the likelihood funded projects will have an impact on the NWP community that contributes to the DTC mission. Feedback from the SAB indicated an AO with more focused guidance on topics of interest to the DTC and requiring

proposals contain more detailed information on the proposed projects would improve the likelihood visitor projects would benefit the DTC mission. The DTC will work with the MB to adjust the AO based on these recommendations.

3. Support for Community Systems

Community code is a free and shared resource with distributed development and centralized support. Ongoing development of community codes is maintained under version control with periodic public releases that include the latest in developments of new capabilities and techniques. To serve as a bridge between operations and research, the DTC provides a framework for the two communities to collaborate in order to accelerate the transition of new technology into operational weather forecasting. The DTC currently supports the following software packages to the community:

- Weather Research and Forecasting (WRF) – NWP model + pre- and post-processors
- WRF for Hurricanes (set of tools for tropical storm forecasting, including a coupled atmosphere and ocean system)
- Gridpoint Statistical Interpolation (GSI) Data Assimilation (DA) System
- Model Evaluation Tools (MET) – Verification package

With the exception of MET, the DTC does not contribute to the development of these software packages. Rather, the DTC contributes to the software management and user support for these community codes. The main developers of these packages are affiliated with EMC, GSD, MMM, Global Modeling and Assimilation Office (GMAO) of the National Aeronautics and Space Administration (NASA), Geophysical Fluid Dynamics Laboratory (GFDL), University of Rhode Island (URI) and the Hurricane Research Division (HRD) of NOAA's Atlantic Oceanographic and Meteorological Laboratory (AOML).

3.1 Weather Research and Forecasting (WRF)

The DTC works in collaboration with MMM to maintain the WRF repository, including running tests on the repository code. Given the continued growth in the complexity of the WRF code and the number of available configuration options, the DTC will work with MMM to maintain and enhance the current weekly regression test as necessary in order to ensure meaningful cases are being exercised for both the ARW and NMM dynamic cores. In addition, the DTC will perform testing on a bi-monthly time frame that will exercise longer forecasts, restarts, data assimilation and other capabilities not included in the weekly regression test.

Major WRF releases occur annually in the March/April timeframe, with the possibility of a minor release (mainly to address bug fixes) later in the year. Prior to any public release, extensive pre-release testing on several platforms using multiple configurations of the code is conducted. Any bug fixes found during internal testing or brought to our attention by community users through *wrfhelp* correspondence are incorporated back into the WRF repository to ensure they are released with the next version of the code. Assistance is also provided to members of the DTC Visitor Program for inclusion of their new developments into the WRF repository.

The DTC works in collaboration with EMC to make a post processing tool available for WRF output that provides the capability to compute a variety of diagnostic fields, interpolate to

pressure levels, destagger grids and interpolate to specified grids. These grid manipulations produce GRIB output files that can be used directly by a number of plotting packages and MET. The current publicly available version of this tool is the WRF Post Processor (WPP). During the first quarter of AOP 2011, the final work towards transitioning WPP to the Unified Post Processor (UPP) will be completed. The UPP software package is based on WPP but has enhanced capabilities to post-process output from a variety of NWP models. The UPP is currently used in operations at NCEP for the Global Forecast System (GFS) and will be utilized for several other operational models, including the WRF Rapid Refresh (WRF-RR) and North American Mesoscale (NAM) models, in the coming year.

The first official release of the community UPP software will be in the spring of 2011. Prior to the release of UPP version 1.0, extensive testing will be performed to ensure that a broad range of model output formats are compatible with the UPP software. Both netCDF and binary output files from the WRF system will be tested in a serial and distributed memory environment and, similar to the WRF pre-release testing, the full suite of tests will be run on all of the computing platforms available to the DTC using a variety of compilers. Full documentation will be written and provided on the web with updates provided with each subsequent release. Future releases of UPP will typically coincide with the annual WRF releases.

During AOP 2011, the DTC will support the WRF Nonhydrostatic Mesoscale Model (WRF-NMM), WPP and UPP software packages to the community and answer email correspondence related to each of the aforementioned codes through *wrfhelp*. The DTC envisions the transition to supporting a single post processing package (i.e., UPP) will be final prior to AOP 2012. Requests for DTC staff assistance via *wrfhelp* average about 40-50 inquiries per month. Relevant bug fixes are made available throughout the year on the WRF-NMM User's Page maintained by the DTC.

Several community outreach events are co-hosted by MMM and the DTC in Boulder, CO each year, including the annual WRF User's Workshop and bi-annual WRF Tutorials. At the WRF User's Workshop, presentations are given on the use of WRF in both research and operational settings. The WRF Tutorials are an opportunity to teach participants how to run the WRF system (NWP model plus pre- and post-processing software). During these tutorials, participants benefit from lectures given by system experts and one-on-one assistance during the practical sessions. System documentation is available on the web and is updated prior to each tutorial and WRF public release.

Anticipated major accomplishments for AOP 2011:

- WRF version 3.3 release
- Community-UPP version 1.0 release
- Bi-annual WRF Tutorials (January and July 2011)
- Annual WRF User's Workshop (June 2011)

3.2 WRF for Hurricanes

The DTC, in collaboration with MMM and EMC, will continue to support WRF for Hurricanes to the community. MMM is responsible for supporting the Advanced Hurricane WRF (AHW) capabilities, whereas the DTC is responsible for supporting the Hurricane WRF (HWRF) system developed by EMC, URI and GFDL. For AOP 2010, the community HWRF system was

composed of WRF (modified version of the WRF-NMM dynamic core with tropics physics suite + pre- and post-processors), a vortex initialization package, the Princeton Ocean Model (POM), a features-based ocean initialization package, the NCEP coupler, and the GFDL vortex tracker. For AOP 2011, the community HWRF system will be enhanced with the addition of the GSI DA package and UPP replacing WPP.

Support of WRF for Hurricanes to the community falls into two basic categories: code management and user support. User support for HWRF includes an email helpdesk, availability of code and dataset downloads, and up-to-date Users' Guide and scientific documentation. The DTC hosts a WRF for Hurricanes website through which users can access the code and obtain support. For AOP 2011, the user support will be enhanced with the addition of an online tutorial for HWRF.

In addition to the services listed above, user support during AOP 2011 will include a 4-day resident tutorial. This tutorial, which will be limited to 40 participants, will consist of lectures given by system experts from a variety of institutions and one-on-one assistance with running test cases during the practical sessions. Additionally, the tutorial will contain a science component, in which students will be given an overview of tropical cyclone NWP and its main challenges. Preparations necessary for the tutorial include: lecture materials, instructions for the practical sessions, and test cases for the students to practice running the end-to-end system.

The DTC, in collaboration with EMC, made significant progress toward adding the operational HWRF capabilities to the DTC community codes during AOP 2010. As a result of this effort, as well as significant testing and evaluation of the community codes, the adoption of the community codes for HWRF operations is anticipated for the 2011 Hurricane season. Having a single code base for operations and the research community is a DTC paradigm to facilitate transition of new research to operations. During AOP 2011, the work of maintaining a single code base will continue with the incorporation of new developments for HWRF (coming from EMC and elsewhere) into the community code repositories and the ongoing transition of those codes to use at EMC. New capabilities slated to be added to the community codes during AOP 2011 include a coupled third nest and ability to run idealized simulations using HWRF. This transition will only be possible with the continued testing of the code. An important component of the current code management for HWRF is regular testing of the HWRF configuration to ascertain that no unintended changes have occurred in the HWRF forecasts due to introduction of code for other WRF configurations.

During AOP 2011, the DTC will also transition the Hybrid Coordinate Ocean Model (HYCOM) from EMC to DTC. A timeline for the release of this package to the community will be determined after the DTC has transitioned the code. This timeline will depend on the timing of operational adoption of HYCOM for HWRF, complexities of the porting and debugging of the code, along with the effort needed to create scientific documentation and a users' guide.

Anticipated major accomplishment for AOP 2011:

- WRF for Hurricanes Tutorial (April 2011).
- Full support for users of WRF for Hurricanes (code availability, helpdesk, documentation, website).
- Maturity of the HWRF configuration in the general WRF code repository, such that it matches the 2011 operational HWRF.

- Two new packages added to the community HWRF system: GSI and UPP.
- Code management document.

3.3 Data Assimilation Systems

For AOP 2011, the DTC will continue to support the GSI DA system to the user community through a collaborative effort with EMC, GMAO, ESRL, AFWA and MMM. The administrative structure for coordinating the distributed development for the GSI system is now the responsibility of the GSI Review Committee that was established during AOP 2010. The DTC is responsible for maintaining the community GSI repository through oversight of the processes for source code modification, source code testing, and repository upkeep. The DTC responsibilities also include managing the preparation and release of GSI code, providing GSI users' support, and hosting the GSI annual resident tutorial.

Supporting and maintaining the community GSI system requires continuous coordination of the scientific and software development from all the participants. The DTC will use the GSI Review Committee to orchestrate the integration of these contributions into the community repository. The function of the GSI Review Committee is twofold. The committee evaluates the scientific relevance of contributions from the GSI partners and the research community, as well as reviews the actual code revisions prior to being added to the software repository. In this way, the GSI Review Committee provides full oversight for the community development. The DTC is responsible for coordinating GSI Review Committee meetings and subsequent communication with the community.

One important function of the DTC is to provide a pathway for the research community to contribute to the GSI development. A draft procedure for managing research to operation (R2O) contributions was established during AOP 2010. For AOP 2011, the DTC will test this procedure and adjust as issues arise with the guidance of the GSI Review Committee. Following this procedure, the DTC will continue to test and subsequently merge community GSI development into the GSI repository upon approval by the GSI Review Committee.

The DTC will continue to maintain the community GSI repository (Boulder). This repository stores the community version of the GSI system, under subversion control to accommodate community contributions, as well as the operational capability. The community GSI system includes the flexible DTC build system, the required external libraries, and run scripts, in addition to the common shared GSI source code. The DTC will continue to sync the common GSI source code, on a weekly basis, with the EMC GSI trunk code, in order to maintain a single version of GSI. The DTC will work closely with the GSI Review Committee on code merging and commitment, and hosting Boulder GSI developer meetings. The DTC will assist the developers, as appropriate, with incorporating new development, more generalized utilities, and bug fixes into the GSI repository. The DTC will also continue to work towards enhancing the stability and portability of the GSI system by focusing on the improvement of support functionality, such as compilation utilities, run scripts, and diagnostic tools.

The DTC will continue to improve its capability to test the integrity of the GSI code in the community repository. The DTC will work towards adding new standard test routines to the regression test suite, including regression test configurations from EMC, as well as creating a GSI cycling test suite for conducting regular short-term cycling experiments.

GSI support and outreach continues to be an important aspect of the coming year's effort. A new version of GSI will be released in spring 2011 using the source code contained within the community repository. A new release is expected to occur on an annual basis. Prior to any public release of the GSI code, the DTC conducts pre-release testing on several computer platforms, and updates the GSI User's Guide, the DTC community GSI webpage, and other GSI documentation. The DTC will also coordinate with the other GSI partners to host the annual Community GSI Tutorial. The organization and content of the tutorial will be updated based on past experience and further discussion. The DTC will host a one-day workshop held, in conjunction with the tutorial, for the first time. This workshop will serve as a forum for GSI developers and members of the GSI user community to share ideas and discuss the future direction of GSI development. In addition, GSI will be publicized at appropriate conferences and workshops. GSI user support will continue to be provided through the GSI helpdesk (gsi_help@ucar.edu). The DTC will also enhance the community support on BUFR/PrepBUFR data processing and conversion. BUFR format is a WMO standard data format and is the only observation format accepted by GSI. "PrepBUFR" is a NCEP term for "prepared" or QC'd conventional observations in BUFR format. To meet an increasing demand from the community, the DTC will build a separate community website for BUFR/PrepBUFR support and document the BUFR/PrepBUFR processing procedures and conversion codes. The DTC will also start developing an online tutorial to provide user self-training courses on BUFR/PrepBUFR format conversion.

In addition to managing and supporting GSI, the DTC will start expanding current GSI efforts to include a GSI-based variational-ensemble hybrid data assimilation ("GSI-hybrid") system. The data assimilation team at the DTC will collaborate with the Hurricane and Ensemble tasks and the appropriate operational data assimilation groups (i.e., NCEP/EMC and AFWA) to develop a work plan to build and test a GSI-hybrid benchmark system, which will eventually be adopted by operational community. The DA team at the DTC will work with HFIP and NCEP/EMC to organize an HFIP workshop directed at developing plans for a unified framework for a GSI-hybrid system that is suitable for research and operations.

The DTC will keep its current staff member resident at EMC in Maryland to encourage direct interactions with the data assimilation system developers at EMC and GMAO, while also staying closely connected with the data assimilation groups at ESRL, AFWA and MMM. This arrangement has and will continue to facilitate efficient communication and coordination among the data assimilation partners and minimize the divergence of the software baselines among distributed groups.

Anticipated major accomplishments for AOP 2011:

- Annual Community GSI Tutorial/Workshop
- GSI version 3.0 release
- Enhanced GSI regression test suite
- GSI community contribution (R2O) procedure and implementation
- BUFR/PrepBUFR community website/documentation
- GSI-hybrid Workshop
- Set up a GSI-hybrid framework

3.4 Verification

The main goal of the verification task is to provide advanced verification tools and support to the community. The primary effort for this task is currently the MET software development and support. Past verification tool development has focused on general forecast verification. Now that the DTC has a substantial effort dedicated to tropical cyclone forecasting, the DTC's verification efforts will also include the determination of appropriate methods for hurricane forecast evaluation and the development of tools for hurricane forecast evaluation. Community outreach in the form of tutorials and collaborations with NOAA testbeds are ongoing. Current NOAA Testbed collaborations include the Hydrometeorology Testbed (HMT) and the Hazardous Weather Testbed (HWT). The MET software is being used and demonstrated during the annual experiments for both of these testbeds, with special attention being paid to evaluation of ensemble forecasts. The Ensemble task plans include a verification component, and thus collaboration with the Ensemble task on methods and tools will be of increasing importance for the verification group.

3.4.1 MET Development

It is widely acknowledged in the NWP community that the traditional verification metrics that are applied to assess model skill (such as bias, root-mean-squared error, and mean absolute error) provide a useful but limited view of how well a model is performing. A number of tools (including object-based, scale separation, neighborhood, and field deformation approaches) are under development by the international verification community to cope with this dilemma, and to provide more meaningful information about the performance of high-resolution models. As the premier, independent, organization responsible for evaluating new NWP techniques, the DTC requires state-of-the-art verification tools. To truly serve as a bridge between research and operations, the DTC needs to make traditional as well as state-of-the-art verification tools available to the NWP community. Hence, the DTC began assembling its Model Evaluation Tools (MET) package in the fall of 2006. This package includes standard verification approaches, as well as a number of more advanced approaches recently developed by the verification community. To maintain truly state-of-the-art verification tools, the DTC must engage the international verification community and continue to add new capabilities to MET.

The MET software package is on an annual development/release schedule with the next major release (MET version 4.0) scheduled for fall of 2011. This release will contain several enhancements, including enhanced cloud forecast and ensemble verification methods. Following the release of MET version 4.0, further enhancements to the cloud and ensemble methods will be implemented. Some of the planned cloud verification enhancements will stem from utilities developed for a separately funded CloudSat project. MET development during AOP 2011 will also continue to focus on better diagnostic methods for evaluating wind forecasts and the capability to examine forecast statistics through time.

The DTC currently uses the National Hurricane Center (NHC) verification package to evaluate its retrospective tests for HWRF. While this package provides a good starting point for this work, the package has a number of limitations. During AOP 2011, the DTC will work with NHC and EMC, as well as HFIP collaborators and verification experts outside of this community, to develop a plan for a more flexible set of hurricane forecast evaluation tools and establish an initial basic capability. These tools will also include methods for performing statistical significance testing. In addition to providing robust, flexible tools for evaluating the

basic storm properties contained in the Best Track analysis, the verification team will explore the application of advanced general forecast verification tools to the hurricane forecast problem. One area ripe for exploration is object-oriented techniques (i.e., MODE) already available in MET. New tools will be made available to the hurricane modeling and forecasting community as they are implemented, with an initial version constructed and made available during AOP 2011.

In addition to advancing the core MET software, the DTC will continue the development of its database and display system for MET output, referred to as METviewer, albeit at a reduced pace due to budget constraints. Components of METviewer will be applied to various DTC testing and evaluation activities during the development stage (e.g., Reference Configuration testing, HWT Spring Experiment (SE), HMT-west).

The DTC engages the international verification community in its development planning process through its Verification Advisory Group (VAG). The VAG is a group of verification experts from both sponsoring agencies and the community that provides guidance and vision for MET. In particular, the VAG provides advice on what new verification methods should be incorporated into MET and thus made available to the modeling and verification communities. The VAG will meet several times during AOP 2011.

Anticipated major accomplishments for AOP 2011:

- MET v4.0 release

3.4.2 MET User Support

Each new release of MET requires activities that fall under the user support category: pre-release testing and updating the MET User's Guide to include information on new verification methods and tools included in the software. Periodic bug fixes are also made available, as necessary, throughout the year on the MET Users Page. A new SVN code repository was implemented in 2009 to streamline the development process and provide consistency with other DTC supported software. Enhancements and additions to the repository will continue through AOP 2011.

Other on-going support tasks include providing assistance via the *met_help* email and bi-annual tutorials held in Boulder, CO. The MET helpdesk receives about 25 emails each month from users seeking support for the MET software, particularly with compilation issues. User assistance is often provided on the same day. The tutorials teach participants how to run the MET software and understand verification methods. The DTC staff designs and gives all presentations and hands-on sessions. The practical session information from these tutorials is incorporated into the online tutorial as well. Thus, the online tutorial will be enhanced to include the cloud verification procedures and ensemble verification methods.

Anticipated major accomplishments for AOP 2011:

- Bi-annual MET tutorials
- Updated MET Users Guide and online tutorial
- Timely support via *met_help*

4. Testing and Evaluation

Testing and evaluation undertaken by the developers of new NWP techniques from the research community is generally focused on case studies. However, to adequately assess these

techniques, extensive testing and evaluation must be performed to ensure that these new techniques are indeed ready for operational consideration. The DTC undertakes two types of rigorous evaluation activities: 1) studies for which the DTC generates forecasts and evaluates the skill of these forecasts (end-to-end internal testing), and 2) studies for which the DTC evaluates the skill of forecasts provided by external modeling groups.

Internal DTC testing and evaluation generally focuses on extended retrospective time periods. The cases selected encapsulate a broad range of weather regimes ranging from null, to weak and strong events. The exact periods chosen vary based on the phenomenon of focus for the test. Before testing begins, a design document is created in consultation with the developers, relevant area experts, and verification experts, detailing the specifications of the testing to be done. The technique to be tested must be part of the repository of community codes supported by the DTC to ensure that the code has reached a certain level of maturity. The DTC's evaluation of these retrospective forecasts includes standard verification techniques, as well as new verification techniques when appropriate. All verification statistics also undergo a statistical significance assessment when appropriate. By conducting carefully controlled, rigorous testing, including the generation of objective verification statistics, the DTC is able to provide the operational community with guidance for selecting new NWP technologies with potential value for operational implementation. DTC testing also provides the research community with baselines against which the impacts of new techniques can be evaluated. The statistical results may also aid researchers in selecting model configurations to use for their projects.

In addition to the internal testing conducted, the DTC participates in a number of testing activities in which the model output is produced by a modeling group or groups. These model evaluations can consist of demonstrations of model evaluation capabilities in real-time forecast projects or long-term retrospective evaluations in which the model setup (e.g. initial conditions, resolution, other model system attributes) is well-defined in advance and the model output is provided to the DTC by the modeling group. The DTC's evaluation of these externally generated forecasts follows the same principles applied to the evaluation of the forecasts generated internally by the DTC; objective verification statistics are computed, including standard verification metrics as well as new verification techniques when appropriate. The verification statistics also undergo a statistical significance assessment when appropriate.

One important area that allows the DTC to expand its testing and evaluation activities is its collaboration with NOAA testbeds. For AOP 2011, these fruitful collaborations will include the Hazardous Weather Testbed (HWT) and the Hydrometeorological Testbed (HMT). The DTC's involvement in these testbeds provides the DTC with access to extensive experimental model forecast data sets that can either be used to supplement its internal testing and evaluation activities for its Mesoscale Modeling, Data Assimilation, Hurricanes, or Ensembles task areas, or provide fertile ground for demonstrating advanced verification techniques. Activities associated with these testbed collaborations will be described under the appropriate task area in order to clearly delineate how these activities contribute to the DTC mission in these five key areas. Note that testbed activities may fall all within one task area or be divided across a few tasks areas, depending on the focus of the particular activity.

4.1 Mesoscale Modeling

Since its inception, the DTC has undertaken testing and evaluation activities directed at assessing the skill of WRF model configurations for mesoscale mid-latitude applications. While the DTC

will continue its WRF testing and evaluation activities for the foreseeable future, broadening the scope of its internal testing and evaluation activities to include new software frameworks utilized by the relevant operational centers (i.e., NOAA and AFWA) will be necessary in order to continue to serve as a bridge between research and operations. Hence, the WRF task area has been renamed Mesoscale Modeling to reflect this broader scope. The testing and evaluation described in this section of the AOP includes the evaluation of forecasts generated by DTC staff, as well as the evaluation of operational forecasts or experimental forecasts that are part of NOAA testbed activities.

Communication is key to serving as a bridge between research and operations. While the DTC website is one important means of making information on DTC activities available to the community, this method alone is not sufficient. Hence, the DTC needs to also dedicate time to producing journals articles on its key activities. For AOP 2011, manuscripts describing the DTC testing and evaluation methodology and the GFS-NAM inter-comparison completed during AOP 2010 will be prepared, as well as any additional manuscripts that may emerge from new test activities.

4.1.1 Reference Configurations (RCs)

The WRF model is a state-of-the-art NWP system that is highly configurable and suitable for a broad range of weather applications. With the numerous options available in this end-to-end system, it is extremely difficult to test all option combinations. Thus, there is a need within the WRF community for widely publicized verification results from a variety of configurations that have been extensively tested and evaluated by either the DTC or a member of the WRF community. These evaluations will ultimately support both the research and operational communities. To fill this need, the DTC conducts Reference Configurations (RCs) testing and evaluation. Early in AOP 2011, a DTC RC Implementation Plan will be written describing the general experiment design for each RC testing and evaluation activity. Feedback on this plan will be solicited from the MB and SAB to ensure the DTC RC testing meets the needs of the community at large.

Several DTC RCs were designated during AOP 2010. To expand the number of baselines for operational or soon-to-be operational models and build a historical record of performance, the DTC will perform testing and evaluation during AOP 2011 of configurations based on the NAM and WRF-RR model physics suites. Upon completion, both configurations will be designated DTC RCs and the baseline data and statistical verification results will be made available to the community on the DTC RC webpage. In addition to testing new WRF configurations, previously designated RCs will be considered for retesting with the latest WRF release. At a minimum, retesting will include the AFWA Operational Configuration. A prioritized list of additional configurations to either be retested or tested for the first time will be established and the tests conducted as resources permit. Since RCs are associated with a particular version of the code, retesting will lead to new RC designations. Retesting configurations for new versions of WRF provides the community with an assessment of the code change impacts within the latest version of the code for particular configurations and serves to monitor progress of forecast improvement as the WRF system evolves.

Widely publicizing verification results is imperative to the RC process. Information on specific RC testing will be presented at appropriate scientific conferences and workshops. The RC

webpage on the DTC website will also continue to be updated throughout AOP 2011 with the latest statistical verification results from RC testing.

The DTC also recognizes that extensive data sets are being produced by the WRF user community and these data sets may provide additional resources to leverage. Members of the university community have been encouraged to collaborate with the DTC to establish Community Contributed Reference Configurations (CCRCs) in order to take advantage of these data and potentially create a more extensive collection of model verification results across a broader range of forecast applications. During AOP 2011, the DTC will continue to encourage and support the research community to share their results from significant testing activities in a manner that will benefit the community as a whole.

4.1.2 Functionally Similar Operational Environment

A fundamental purpose of the DTC is to facilitate operational and research collaborations with the goal of accelerating the transfer of new science and technology from research into operations. To facilitate these collaborations, the DTC must have the capability to run a functionally similar operational system. This framework will ultimately help assess the overall performance a new model configuration will have in an operational setting. One key aspect in this end-to-end system testing, that is not necessarily included in RC testing, is the inclusion of data assimilation/cycling.

During AOP 2011, the DTC will work toward establishing a functionally similar operational environment at the DTC for current operational mesoscale applications that utilize or will soon utilize the community codes currently supported by the DTC (e.g., WRF-RR and NCEP's High Resolution Windows). Software packages including pre-processing (e.g. WPS), data assimilation (e.g. GSI), the model itself (e.g. WRF), and post-processing (e.g. UPP) will be linked and run end-to-end for specified testing. Input datasets used in the operational environment will also be utilized in retrospective testing performed by the DTC.

As NCEP and NOAA transition their operations to the NOAA Environmental Modeling System (NEMS), the DTC will continue to strengthen its foundation of expertise with the NEMS software and the science associated with the NEMS-based models. To accomplish this, one DTC staff member will continue to work at EMC in Camp Springs to collaborate with EMC on the development of NMMB in the NEMS framework for select areas (e.g. portability, interoperability, I/O layer capabilities). Emphasis on community software will be a focus in future years. The DTC will also continue to participate in technical discussions and host a workshop focused on the status of NEMS development and options regarding the integration of new features or capabilities. Rather than providing full community support for this package, the DTC will establish an in-house capability that will provide the framework for new configuration settings from a research environment to be included in an operational environment and tested. Contributions from the research community that have shown significant potential to positively impact operations may be tested within this framework by the DTC.

4.1.3 Model Physics

Facilitating interactions between the research and operational communities in order to identify, examine and address key national NWP problems is at the core of the DTC mission. Physics parameterizations have been identified as a key area that needs more focused and coordinated attention. To facilitate a stronger connection between research and operations in this area, the

DTC, in collaboration with EMC, will host a physics workshop in the Washington D.C. area for invited participants. Discussions at this workshop will focus on both short-term and longer-term strategies for accelerating improvements to physics parameterizations through close collaborations between the research and operational communities.

In the short-term, this workshop will help identify new techniques or capabilities available in the research community with the potential to have a positive impact on operations that can be readily tested and evaluated by the DTC. Discussions focusing on longer-term strategies will identify which physical processes remain challenging to the community and identify approaches toward making substantial progress in these areas. In the end, multiple approaches to forecast improvement will continue to be pursued, addressing different needs and objectives; however, the DTC hopes this workshop will result in greater coordination and communication on model physics within the community of developers.

Observations beyond that available through the operational network are often times needed in order to perform detailed analysis of the strengths and weaknesses of physics parameterizations. HMT has, over the course of several years, put in place a number of research observational platforms that provide important information about the precipitation processes. These observations provide detailed information about the temporal evolution of precipitation at the surface and moisture processes within columns. The DTC, in collaboration with HMT, plans to use these research data sets to evaluate microphysical algorithms as applied in operational forecasts. In addition to standard verification scores, this analysis will utilize enhanced MET tools (e.g., MODE Time Dimension) to compare point and column observations with parallel representations of model microphysical forecasts. These comparisons will provide much needed feedback to the developers on the performance of microphysical parameterizations. Hence, the DTC's collaboration with HMT will contribute to the efforts of advancing model physics development.

Anticipated major accomplishments for AOP 2011:

- New RCs designated for WRF v3.3 with results posted on the DTC RC webpage and presented at relevant conferences
- White paper outlining short-term and longer-term approaches for making significant progress toward improving physics parameterizations
- Testing and Evaluation Plan for assessing microphysical parameterizations using HMT observations
- Functionally similar operational environment established at DTC for appropriate operational mesoscale applications.
- Manuscript(s) submitted to appropriate peer-reviewed journals

4.2 Hurricanes

During AOP 2011, the DTC will continue to update its HWRF Testing and Evaluation suite on the HFIP computer platform. This work will include adopting changes in operational procedures for running HWRF and the addition of HYCOM to the testing suite.

In addition to updating the test suite, the DTC will perform comprehensive tests for HWRF with the objectives of assessing the skill of 1) the operational HWRF configuration in the general

WRF repository and 2) alternative HWRF configurations based on promising advances for the 2012 EMC operational implementation and beyond. These tests will be conducted using the DTC HWRF Testing and Evaluation suite on the HFIP computer platform. Comprehensive tests of the operational HWRF configuration from the community repositories will consist of a large number of cases (approximately 400 cases), and some will be used to designate DTC RCs. Testing of promising new capabilities for operational implementation, which will give careful consideration to code management practices, will include three levels of testing:

- Preliminary testing and evaluation of proposed model upgrades, including diagnostics, to determine potential suitability for operations.
- Advanced testing and evaluation of proposed model upgrades, to assess cost/benefit ratio for operational implementation.
- Comprehensive testing and evaluation leading to recommendations for upgrades to be considered for transition to operational testing.

Most, if not all, of the new capabilities that will be tested by the DTC during AOP 2011, will likely stem from development partially funded by HFIP. The actual tests to be conducted are yet to be determined, but will include increased resolution, alternate physics configurations and revised initialization procedures. Diagnostic activities will help identify the strengths and weaknesses of HWRF and will provide information for improving the model. Diagnostics will be conducted to assess the impact of alternate physical parameterizations on the forecast and to analyze the representation of the large scale fields in the model. In addition, a multi-year DTC HWRF Test Plan will be written in collaboration with EMC and the HFIP Program Office to determine DTC activities that will help HFIP reach its milestones. This plan will contain details of the tests to be conducted during AOP 2011, along with protocols for tests to be conducted in upcoming years and the process for developers to deliver new capabilities to the DTC for testing.

Communication of DTC activities related to Hurricanes is also important. While the DTC website is one important means of making information on DTC activities available to the community, this method alone is not sufficient. Hence, the DTC needs to also dedicate time to producing journals articles on its key activities. For AOP 2011, a manuscript describing the DTC's evaluation of the High Resolution Hurricane Test that was completed in September 2009 will be prepared and submitted to the appropriate journal.

Anticipated major accomplishment for FY 2011:

- DTC HWRF Test Plan
- Updated HWRF testing infrastructure on HFIP computing platform
- HWRF Reference Configuration
- Testing and diagnostics to contribute towards HWRF 2012 operational implementation

4.3 Data Assimilation

For AOP 2011, the DTC will continue testing DA systems and techniques in applications that are relevant to the GSI partners and DTC sponsors. Testing for AOP 2011 will consist of GSI baseline tests. The DTC will test the GSI code extensively in an end-to-end system, including data pre-processing, data assimilation, forecast, post-processing and verification. Prior to conducting the tests, the components of the end-to-end system, which include WPS, GSI, WRF-

ARW, and MET, will be updated to their latest version available in the spring of 2011. The configuration of the testing for GSI v3.0 will be based on recommendations from the DTC GSI version 2.0 testing and evaluation report. The reference for GSI v3.0 testing will be one extensive run of the WRF-ARW (standard no data assimilation or "NODA" run). The DTC will then conduct experiments comparing GSI v3.0 and WRFDA 3D-Var v3.3 in AFWA regional applications. A report summarizing the test results and recommendation will be provided to AFWA. Results will also be posted on the DTC website for the NWP community to access.

In addition to GSI baseline test, the DTC will start investigating Ensemble Kalman Filter (EnKF) techniques and conducting benchmark EnKF tests. The DTC will set up the AFWA testbed with a regional EnKF system and develop expertise in regional EnKF. Specific efforts include: a) develop a scripting suite for running the NCAR Data Assimilation Research Testbed EnKF package (DART-EnKF) in an end-to-end system, including WPS and WRF-ARW; b) conduct a benchmark run using DART-EnKF to assimilate conventional data in the end-to-end system. Appropriate scripts and running sequences will be tested and the robustness of the system will be evaluated.

For AOP 2011, the DTC will also start to build a framework for testing and evaluating a GSI-hybrid system for regional prediction models (e.g., HWRF). Testing and evaluation for the GSI-hybrid technique will entail two important pieces: 1) the investigation of different methods for ensemble generation, and 2) conducting benchmark tests using GSI to assimilate observations in hybrid mode using the information provided by a selected ensemble generation scheme. To assure results of the DTC's testing are useful to operations, the DTC will work closely with its operational partners to build this framework such that it includes a testing standard, which will emulate the operational environment. Establishing the testing and evaluation framework will be the focus of work undertaken during AOP 2011, with the focus turning to testing and evaluation using this framework during AOP 2012.

Anticipated major accomplishments for AOP 2011:

- A report summarizing the results of GSI system testing and comparison of GSI and WRFDA 3D-Var
- End-to-end system for benchmark testing of EnKF techniques
- GSI-hybrid testing and evaluation framework

4.4 Ensemble

In response to the operational centers' move toward ensemble-based probabilistic forecasting, the DTC allocated resources during AOP 2010 to establish an Ensemble task in order to meet the needs of this important area. The goal of this task is to *provide an environment in which extensive testing and evaluation of ensemble-related techniques developed by the NWP community can be conducted such that the results are immediately relevant to the operational centers (e.g., NCEP and AFWA).* Once all the pieces are assembled, Ensemble activities will involve maintaining the Ensemble testing and evaluation infrastructure and updating this infrastructure to include promising new capabilities and techniques developed by the community, as well as conducting extensive testing and evaluation of these new capabilities and techniques. The current DTC community codes will serve as building blocks for the end-to-end ensemble testing and evaluation system assembled by the DTC. These building blocks will be elements of

a DTC Ensemble code repository, which will also include code from the NCEP Short Range Ensemble Forecast (SREF) and other ensemble forecast systems.

Given the DTC needs to facilitate testing and evaluation of competing techniques and capabilities for specific components of the ensemble system, the Ensemble testing and evaluation 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 modules will be configured such that they are able to replicate the algorithms used by the operational centers. Modules 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 Ensemble 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.

Once a new capability and its respective benchmark have been established, the DTC 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 DTC will be starting AOP 2011 with detailed plans for the overall Ensemble 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 Mesoscale Ensembles 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 Ensemble testing and evaluation infrastructure 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 SREF system. Further input will be gathered at the NCEP 5th Ensemble Users Workshop, in May of 2011.

4.4.1 Overarching design

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

4.4.2 Ensemble Configuration Module

For AOP 2011, the DTC plans to establish a basic capability for the ensemble configuration module. Building on the HMT ensemble experience during AOP 2010, the DTC will incorporate the capability of running various models within the NEMS and WRF frameworks into the Ensemble infrastructure. These models will include the ARW and NMM in the WRF infrastructure and the NMMB in the NEMS framework. The DTC 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.

4.4.3 Initial Perturbations Module

For AOP 2011, the DTC plans to contribute to the next upgrade of NCEP's SREF system by working with historical SREF data to determine an appropriate vertically-varying rescaling factor in the blending perturbations method planned for the next SREF implementation. This work will be transitioned to EMC for incorporation into their SREF system for pre-implementation testing. A report on this work will be provided to EMC and posted on the DTC website soothe NWP community will have access to the results. In addition, the DTC will work closely with EMC to assure that all aspects of the initial perturbations method for the next SREF implementation are incorporated into the DTC Ensemble infrastructure. Through this work, the benchmark for the initial perturbations module will be established.

A simplified version of the "cycling of perturbations" technique was introduced as an initial basic capability for the initial perturbations 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 will be fully implemented into the Ensemble infrastructure during AOP 2011. Head-to-head tests between the two initial perturbation techniques available in this module will be run over an extended time period for the HMT ensemble and a report on the evaluation of these results will be prepared and delivered to NCEP. In addition to the report, the DTC will make the tested software available to NCEP for further testing and operational implementation. The report from this test will also be posted on the DTC website so the NWP community will have access to the results.

4.4.4 Product Generation Module

At the request of HWT participants, and through the ingenuity of research at the Center for Analysis and Prediction of Storms (CAPS), 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, quantitative precipitation forecast (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 product generation module through its collaboration with HWT. The approach will be twofold:

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).

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 product generation module, thus, establishing an initial basic capability for this module that can undergo further testing by the DTC. Resources permitting, results may also be produced and made available to the community related to the optimal choice for the number of ensemble members.

4.4.5 Verification Module

The DTC will apply the 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. MET will be one of the verification packages used. These verification metrics will be defined through discussions with major operational users (EMC, Federal Aviation Administration, Office of Hydrologic Development, etc). Experimental DTC runs for the initial perturbation module will be evaluated using MET. Additionally, HWT and HMT ensembles may be evaluated beyond the testbed collaboration efforts to further understanding and interactions with these organizations. Resources permitting, the verification module will also be used for the evaluation of some HFIP regional ensemble experiments.

4.4.6 Computing Resources

From the start, the Ensemble task 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 Ensemble task prepared a

request and was granted a startup allocation on the Teragrid computational network supported by NSF. This allocation will allow the DTC to test the portability of its infrastructure and run a limited number of tests in support of its AOP 2011 work. The DTC 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. HFIP management favorably received the proposal and the request for computing resources was granted. This work will benefit HFIP by providing potentially useful guidance on tracks and boundary conditions for nested hurricane ensemble forecasts. Through this allocation, the DTC will be able to perform extensive testing during AOP 2011 in support of NCEP's next SREF implementation.

4.4.7 Community Interactions

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

The Ensemble task 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 DTC's Ensemble infrastructure for its real-time forecast system.

4.4.8 GSI-Hybrid for HWRF

The Hurricane Forecast Improvement Program (HFIP) has a strong interest in advancing a regional model version of the GSI-Hybrid, specifically for the HWRF model. The ensemble task will work with the DA task on the GSI-Hybrid project. It is common to initialize a regional scale model, such as HWRF, with initial and boundary conditions from a global ensemble, such as NCEP's Global Ensemble Forecast System (GEFS). Using a global model to initialize a local model results in analysis uncertainty at scales the global model does not resolve being absent. For example, the Hurricane Ensemble Data Assimilation System (HEDAS) uses the GFS ensemble to initialize its HWRF ensemble members. The HWRF members then 'spin up' the higher resolution structures. Given the inner core tropical cyclone structures that HWRF can represent, but are not present in GEFS based initial conditions, methods for including additional initial condition uncertainty in the ensemble, at the finer scales, will be investigated. Methods that will be investigated will include downscaling of GEFS, construction of high resolution perturbations (via blending/ cycling), and other methods suggested by the community.

Anticipated major accomplishments for AOP 2011:

- Requirements and plans for the statistical post-processing, product generation, and verification modules based on community input at 5th Ensemble User Workshop (spring 2011)
- Report on HMT & HWT ensemble product evaluation

- Benchmark for all modules based on quarter one FY2012 NCEP SREF
- DTC ensemble code repository
- Basic capability for the testing and evaluation of initial perturbation and product generation techniques
- Expanded capability for the verification module
- Report on vertical variation of rescaling initial perturbation technique testing
- Report on results of initial perturbation technique testing for HMT ensemble

4.5 Verification

Demonstrations of the advanced verification techniques that are part of MET through the DTC's collaborations with NOAA testbeds are critical to the DTC mission in a number of ways. First and foremost, such demonstrations provide an opportunity to educate the NWP community on advantages and limitations of these new techniques and the type of information these techniques can provide. Providing model developers with the information available through the application of advanced verification techniques and showing them how this information can be utilized to assess the strengths and weaknesses of their NWP techniques is critical to accelerating the rate at which operational numerical guidance improves. Secondly, the DTC's collaboration with NOAA testbeds leads to advances in the MET software package through refinements of current tools, expansion of the applicability of current tools or the addition of new capabilities critical to the collaborative effort. The application of these techniques during real-time experiments associated with the testbeds also provides beneficial feedback to the MET developers on the software package. And finally, this type of application provides feedback to the MET scientific team as to strengths and weaknesses of these new techniques, as well as the merits or appropriate approach for applying the methods to new fields.

4.5.1 Object-Oriented Verification Tools

For AOP 2010, the DTC's demonstration of MET capabilities for HMT focused on the application of the Method for Object-Based Diagnostic Evaluation (MODE), which provides an object-based verification for comparing gridded forecasts to gridded observation, to the spatial characteristics of Atmospheric Rivers (ARs). This application of MODE utilized satellite-derived integrated water vapor fields to generate a qualitative assessment of forecast models' ability to capture the spatial characteristics of ARs. For AOP 2011, MODE output will be used to produce more quantitative measures. In addition, the application of MODE for HMT will turn its focus to spatial forecasting of other field objects (e.g., moisture flux) and alternative representations (e.g., objects defined for time-height representations). In addition to providing valuable information about the QPF problem along the West Coast of the United States, the work undertaken to make these new types of analyses possible will result in enhanced capabilities for the MODE tool that will be incorporated into future MET releases.

The capability to extend MODE to include the time dimension (developed under separate funding) is in the process of being added to MET. While this tool appears to hold substantial promise for characterizing and evaluating timing and propagation errors associated with particular numerical models, testing and evaluation of the tool using additional datasets is needed. The DTC's collaboration with HMT provides an ideal setting for the application of the MODE Time-Dimension (MODE-TD) tool. For the evaluation of AR forecasts, the HMT-west

plans to extend spatial verification capabilities beyond observations and forecasts of integrated water vapor to the actual moisture flux. This application will utilize MODE-TD to identify and verify two-dimensional flux objects observed by satellite imagery and forecast by the operational GFS model over the Pacific Ocean and by the HMT-West WRF model ensemble near the California coast.

4.5.2 Ensemble Verification Tools

HMT has on-going QPF-related projects that provide excellent opportunities for leveraged model-assessment research that feeds directly back into the enhancement and extension of the DTC's verification tools. These opportunities include the direct verification of EMC-based numerical models using advanced verification techniques, and the provision of datasets useful for direct testing of verification methods for both deterministic and ensemble forecasts systems in a real-life operational setting. For AOP 2011, the DTC's collaboration with HMT will initially focus on gathering data and completing analysis of verification statistics produced for EMC operational and research models during both the 2009-2010 and 2010-2011 HMT-West winter experiments. This evaluation will include EMC's operational GFS and NAM, the updated version of SREF, and the developmental NMMB and High-Resolution Rapid Refresh (HRRR). Verification will focus on a U.S. West Coast regional domain corresponding to that of HMT's experimental WRF-based ensemble system, which will provide the opportunity to make fair comparisons between models and modeling methods. This evaluation, which will consider deterministic and ensemble-based forecast systems, as well as ensemble-based probabilistic methods, will provide an opportunity to apply and test MET's new Ensemble Stat tool to calculate ensemble mean, spread, and rank histograms. Additionally, it provides a dataset for development and testing of METviewer's ensemble verification capabilities, including plotting Receiver Operator Characteristic (ROC) diagrams, Rank-Histograms, and spread-skill relationships, which are all core capabilities that will be used by the Ensemble task in its verification module. Hence, the DTC's collaboration with HMT will result in enhanced, as well as new verification tools important to ensemble verification efforts by both operational entities and the research community.

Anticipated major accomplishments for AOP 2011:

- White paper report and draft journal manuscript describing 2-year aggregate HMT-West QPF verification for operational and research numerical forecasts