High Resolution Hurricane Forecast Test Plan

2008-09 effort

06/27/2008

1 Introduction

Tropical cyclones are a serious concern for the nation, causing significant risk to life, property and economic vitality. The NOAA National Weather Service has a mission of issuing tropical cyclone forecasts and warnings, aimed at protecting life and property and enhancing the national economy. In the last 10 years, the errors in hurricane track forecasts have been reduced by about 50% through improved model guidance, enhanced observations, and forecaster expertise. However, little progress has been made during this period toward reducing forecasted intensity errors.

To address this shortcoming, NOAA established the Hurricane Forecast Improvement Project (HFIP) in 2007. HFIP is a 10-year plan to improve one to five day tropical cyclone forecasts, with a focus on rapid intensity change (for more details, see http://www.nrc.noaa.gov/HFIP%20Draft%20Plan.html). The HFIP plan details a variety of approaches to improve hurricane forecasting, including research and development to provide: (1) an observing strategy analysis capability for hurricanes, (2) an improved understanding of hurricane intensity change, and (3) an advanced hurricane numerical modeling system. Recent research suggests that prediction models with grid spacing less than 1 km in the inner core of the hurricane may provide a substantial improvement in intensity forecasts. The 2008-09 staging of the high resolution test aims at quantifying the impact of increased horizontal resolution in numerical models on hurricane intensity forecasts. The primary goal of this test is an evaluation of the effect of increasing horizontal resolution within a given model across a variety of storms with different intensity, location and structure. A secondary goal is to provide a data set that can be used to explore the potential value of a multi-model ensemble for improving hurricane forecasts.

The Developmental Testbed Center (DTC) and the HFIP Team hosted a workshop at the National Hurricane Center in Miami, FL, 11-12 March 2008. Experts on hurricanes, numerical modeling and model evaluation met for two days to discuss the strategy for this test. Table 1 contains a list of the workshop participants. The plan presented in this document reflects the consensus reached on the framework for this testing effort. The participants, timeline, selection of test cases, agreed model configuration, evaluation metrics, and plans for distribution of results are outlined below.

2 Participants

2.1 DTC

In addition to hosting the planning workshop and preparing the test plan document, the DTC will facilitate the test by:

- Creating an archive containing the following information that will be made available to the modeling groups:

- *GFS gridded forecasts for the selected tropical cyclones*. The forecasts in this archive will be for the 2007 implementation of the GFS. The archive will include GFS analyses and forecasts out to 126 h (6-h intervals) for all initialization times starting with the earliest selected initialization time through the last selected initialization time for a particular storm. For the selected cases where hours with track is less than 126 h, only the appropriate portion of the GFS forecast will be included in the archive. The archive will also include the GFS analyses and forecasts for the periods leading up to the earliest selected initialization time that modeling groups need to spin up their data assimilation/cycling systems, as well as any intermediate files needed for these systems.
- *GFDL initial fields.* These fields consist of the GFS initialization with the embedded GFDL vortex and will be used to determine the outer domain for each run. The archive will include GFDL initial fields at 6-h intervals starting 24 hours before the earliest selected initialization time through the last selected initialization time for a particular storm. Each modeling group can decide whether or not they want to incorporate the GFDL initial fields into their model initialization.
- *Operational storm data from the TCM at 6-h intervals during the life time of a particular storm.* This information includes the latitude and longitude coordinates of the storm center, the intensity, wind radii (64, 50 and 34 kt) in each quadrant and the storm motion.
- *Observations/Analyses for computing the selected verification metrics at 6 -h intervals during the life time of a particular storm.* The NHC ATCF B-decks contain all but one of the observations or analyses needed to compute the selected verification metrics. The radius of maximum wind will be based on a review of the NHC operational estimates that will be performed by Jack Beven of NHC.

- Specifying the requirements for computing the selected verification variables. The DTC will provide the modeling groups with a list of output fields required for computing the verification variables, including specifications on the frequency of output and the level(s) at which the fields should be made available. In addition to providing this list, the DTC will also provide each modeling group with a common output routine for post-processing their model output. This output routine will produce files compatible with the GFDL Vortex Tracker program. Each modeling group will be responsible for developing the front-end portion of the post-processing routine necessary to produce the specified output fields on a non-staggered latitude-longitude grid for the specified levels.
- Coordinating the transfer of post-processed model files from the various modeling groups to the DTC and computing the selected verification metrics for all available model forecasts. The DTC will create an archive containing the post-processed model files and verification metrics for each experimental forecast. For model configurations that utilize feedback between coarser and finer domains, verification metrics will only be computed for the innermost domain of any configuration.

- Collecting and archiving information from each modeling group specifying the model configuration used to generate their experimental forecasts, including information related to domain configuration, physics packages, data assimilation approaches, cycling, special input data, and special initialization techniques.
- Creating selected plots for the retrospective forecasts and displaying them on a website.
- Providing the group with access to the verification results and all archived data.
- Preparing a report summarizing the verification results for the test.

2.2 Modeling groups

During the March 2008 workshop, eight modeling groups expressed interest in participating in the high resolution test activity planned for the coming year. The range of model resolutions and number of test cases for which forecasts will be run varies depending on the resources each group has available to commit to this project. The prospective participants and their proposed level of participation are summarized in Table 2. The participation of the NASA-Goddard modeling team is contingent on the approval of a proposal they will be submitting to NASA this summer. NCEP/EMC has run the current operational configuration of HWRF (grid spacing of 9 km) for the cases included in this study, and does not plan to run any higher resolution forecasts. GFDL will run the current operational configuration of its hurricane model (grid spacing of 1/12°) for the selected cases using the 2007 version of GFS. The initial condition from these runs will be made available by the DTC. In the fall, GFDL/URI plans to run an additional higher resolution configuration (grid spacing of 1/18°) of the GFDL hurricane model.

2.3 Verification Team

A verification team was appointed during the workshop: Barb Brown (NCAR), James Franklin (NHC), Mike Fiorino (NHC), Mark DeMaria (NOAA/CIRA), and Tim Marchok (GFDL). This team was tasked with determining the sample size for the cases, the criteria for case selection, the verification metrics that will be used to assess the difference in skill due to changes in model resolution, what tools are available for computing the selected metrics, and finally what observations, analyses, and model output fields will be needed to compute the selected metrics. The membership of this team includes a good mixture of research and operational scientists, but does not include any members of the modeling groups.

2.4 Case Selection Team

A case selection team was also appointed during the workshop: Mark DeMaria (NOAA/CIRA) and Jack Beven (NHC). This team was tasked with selecting a diverse set of storms and time periods from each of these storms that will meet the criteria set forth by the verification team. This team was also tasked with determining the order in which the retrospective forecasts for the selected cases should be run. The membership of this team also includes a good mixture of research and operational scientists.

2.5 HFIP Team

During the March 2008 workshop, the HFIP Team provided the participants with guidance on the goals for the testing activity, as well as information on short-term and long-term time lines for HFIP activities. The HFIP Team will also be responsible for evaluating the results of the planned testing activity and providing direction for future testing activities.

3 Time Line

The driving force behind the time line for this test plan is a need for a preliminary assessment of the impact of high resolution on hurricane forecasts by the end of the second quarter of 2009 (31 March 2009). The time line put forth in this document was obtained by dividing the tasks that need to be completed into five phases: pre-test preparations, retrospective forecast runs, verification computations, report preparation and distribution, and workshop. Given the limited time to arrive at the preliminary assessment, timely transfer of data from the modeling groups to the DTC for the computation of verification metrics will be critical. Hence, the modeling groups are strongly encouraged to provide files to the DTC for post-processing as soon as they are available, rather than waiting until they have completed all of the retrospective forecasts. This need is reflected in the overlapping time windows for these complementary tasks.

3.1 Pre-Test Preparations (April – August 2008)

The following tasks need to be addressed prior to the start of the retrospective testing:

- Finalize prioritized list of cases (Test Case Team).
- Create archive of input and verification data sets (DTC).
- Obtain necessary existing verification tools, test on sample high resolution grids, and develop any necessary new verification tools (DTC and Verification Team).
- Define guidelines for how model fields provided for verification will be computed. For instance, a maximum instantaneous wind from a high resolution model may not be the appropriate field for comparing with Best Track assessments of intensity (DTC in consultation with test participants).
- Finalize list of model output fields needed for computing verification metrics (DTC and Verification Team).
- Develop common output routine for post-processing native model grids and distribute to each modeling group (DTC).
- Determine method for transferring input data/post-processed model output files to/from the various computing platforms used to generate the retrospective forecasts from/to the DTC (DTC).
- Test sample post-processed output files from each model group to evaluate procedure (DTC).

Given the tight time line for this testing activity, the modeling groups are encouraged to pursue whatever tasks they deem necessary during this pre-test time period to arrive at their final model configuration for the retrospective testing. The DTC will strive to provide the modeling groups with the finalized list of output fields and common output routine no later than 1 August 2008. The DTC is also committed to testing sample output files from each modeling group as quickly as possible to avoid delaying the generation of the retrospective forecasts. Once a modeling group has provided a set of sample output files and these files have been tested by the DTC, the model group will be deemed ready to start generating retrospective forecasts.

3.2 Retrospective Testing (September 2008 – early January 2009)

The modeling groups will have a four-month window to generate retrospective forecasts for as many of the selected test cases as possible. Given some modeling groups will not be able to generate retrospective forecasts for all of the selected cases, each group should run the forecast cycles in the order listed in Table 3. Following this guideline will assure a complete data set for a subset of the cases. Each modeling group is welcome to continue generating retrospective forecasts during the time window leading up to the workshop. The DTC will strive to include as many of the available retrospective forecasts in their analysis as possible, but will only guarantee inclusion of the data received by 9 January 2009 in its preliminary report. If additional data are available by the time of the workshop, a second report can be prepared following the workshop that includes this additional information. Any group may modify their model-system configuration during the test, but all cases submitted for evaluation must be run with the same model configuration, apart from modifying the location (not the size) of the domains.

3.3 Verification Computations (September 2008 – late-February 2009)

Each modeling group is strongly encouraged to provide the DTC with sample post-processed files for each grid resolution they plan to run as soon as files are available. By providing sample files early in the process, the DTC will be able to determine whether any modifications need to be made to the post-processing procedure before a modeling group goes into production mode. Catching problems early in the process and quickly resolving these problems will be critical to achieving the goals of this test plan.

Each modeling group is also strongly encouraged to provide the DTC with access to their retrospective forecasts as soon as they are available. By computing the verification metrics as the data become available, the DTC will be able to provide timely feedback to the modeling groups about any problems with the processing they encounter. In addition, this arrangement will reduce the time required to complete the verification computations, compile overall verification statistics and prepare a report following the completion of the retrospective runs.

3.4 Report Preparations (mid-January – late-February 2009)

During January and February, the DTC will complete the verification computations for the retrospective forecasts completed by 9 January 2009. All forecasts completed by this deadline will be used to produce the overall statistics that will be included in the report the DTC prepares and distributes to the modeling groups and HFIP Team in late-February. The inclusion of data in the report that are provided after the 9 January 2009 deadline may be possible but cannot be guaranteed at this time.

3.5 Workshop (mid-March 2009)

The DTC will sponsor a second workshop in mid-March 2009. This workshop will provide a forum for discussing the experiences of the modeling groups who participated in the 2008-09 testing, the results presented in the DTC report, and plans for the next testing activity.

4 Test Cases

4.1 Prioritized list of test cases

Table 3 contains the list of the forecast cycles provided by the Case Selection Team. The tracks and intensities of the ten storms included in the list of cases are shown in Figure 1. Forecast cycles should be run in the order listed in Table 3. In other words, all modeling groups should start with the forecast cycle initialized at 00 UTC on 16 October 2005. By running the cycles in the order listed, modeling groups who do not have time to complete the full set of retrospective runs by the 9 January 2009 deadline will be able to cover a diverse subset of the selected cycles. Running the selected forecast cycles in the order listed will also assure a complete data set for a subset of the cases.

4.2 Brief description of selected Tropical Cyclones

Wilma 2005

Wilma was a large-envelope cyclone that had a slow initial development over the northwestern Caribbean Sea, followed by the greatest intensification rate ever seen in the Atlantic basin. The small central core evolved into a very large central core as the hurricane crossed the Yucatan Peninsula of Mexico and southern Florida. Wilma became extra-tropical over the western Atlantic. Five forecast periods for this system met the DeMaria-Kaplan RI criteria. Wilma represents perhaps the ultimate in tropical cyclone structure, with the eye contracting to 2-3 n mi in diameter at the time of peak intensity. The system then underwent a series of eyewall replacement cycles to become 50-60 n mi in diameter during its crossing of south Florida. Wilma is an excellent test case for RI, eyewall replacement, land interaction, and extra-tropical transition.

Philippe 2005

Philippe reached hurricane strength east of the Lesser Antilles, and then weakened due to the impact of shear and interaction with an upper-level low. This storm is an excellent test case for arrested development due to shear, which is good for testing false alarms. It should be noted that the GFDL model produced false alarms for RI in the real-time runs.

Rita 2005

Rita developed rapidly over the Florida Straits and the eastern Gulf of Mexico. This development was followed by weakening due to a series of eyewall replacement cycles and increasing vertical shear. Rita made landfall over western Louisiana and eastern Texas. Six forecast periods for this system met the DeMaria-Kaplan RI criteria. Rita is an excellent test case for RI and eyewall replacement. Data collected from the RAINEX experiment would allow a detailed verification of the model depiction of the cyclone structure.

Karen 2007

Karen was a large-envelope cyclone that formed over the eastern Atlantic Ocean. The system started to develop rapidly and reached hurricane strength just before encountering strong shear, then dissipated over water due to shear. Two forecast periods for this system met the DeMaria-Kaplan criteria RI. Karen represents a case where RI might have been forecast in NWP models, only to be abruptly halted by the increased shear.

Katrina 2005

Katrina formed over the Bahamas, becoming a hurricane just before landfall in southern Florida. This system then intensified over the Gulf of Mexico and underwent an eyewall replacement before reaching category 5 intensity. Katrina then weakened due to another eyewall replacement, and possibly shear and ocean effects before landfall along the northern Gulf coast. Five forecast periods for this storm met the DeMaria-Kaplan RI criteria. Katrina is a good test case for RI, which for this storm was interrupted by an eyewall replacement cycle. It is also a good case for improved track forecasts (particularly before the Florida landfall), land interaction, and oceanic effects (the Loop Current).

Humberto 2007

Humberto developed rapidly before landfall on the Texas coast, going from a tropical depression to a hurricane in less than 24 hours. Two forecast periods for this system met the DeMaria-Kaplan RI criteria. RI started unusually early in the development of this cyclone, making this storm a good candidate for testing how models will forecast RI in weak systems.

Felix 2007

Felix was a small cyclone that moved westward through the Caribbean Sea and strengthened from a tropical storm to a Category 5 hurricane in 48 hours. The hurricane made landfall in Nicaragua and dissipated over Central America. Eight forecast periods for this system met the DeMaria-Kaplan RI criteria. Felix is an excellent test case for RI, as well as a cyclone that underwent a notable eyewall replacement cycle.

Ingrid 2007

Ingrid was a weak system over the tropical Atlantic whose development was limited by shear and dissipated over water. This system is a good test case for model false alarms of RI.

Emily 2005

Emily developed over the tropical Atlantic, became a hurricane near the Windward Islands and then, after an eyewall replacement cycle, became a category 5 hurricane over the Caribbean Sea. Emily made landfall over the Yucatan Peninsula as a major hurricane, weakened, and then re-intensified over the western Gulf of Mexico. At final landfall in Mexico, the hurricane featured a prominent concentric eyewall in both radar and aircraft data. Eleven forecast periods for this system met the DeMaria-Kaplan RI criteria. Emily is a good test case for RI and intensity fluctuations, with 2 eyewall replacement cycles and one over-water weakening phase with no obvious cause. The initial intensification to hurricane strength was accompanied by a position jump, suggesting the vortex re-formed or re-organized.

Ophelia 2005

Ophelia formed from a non-tropical trough and followed a meandering course off the southeastern coast of the U. S. for many days. While there were periods of RI and the peak intensity was only 75 kts, there were numerous fluctuations in intensity. Ophelia eventually passed Cape Hatteras and became extra-tropical as it approached Nova Scotia. This cyclone is the poster child for oceanic effects, with a slow meandering track that brought it across its own cold wake or allowed it to remain stationary long enough to upwell.

5 Model Configurations

The DTC made every effort to avoid putting undue constraints on the configuration of the numerical models. Each modeling group has extensive expertise in hurricane forecasting and is currently employing a broad spectrum of physics parameterizations, data assimilation techniques and coupling complexities, making it impossible to expect all groups to use the same configuration. The minimum guidelines outlined below will allow the modelers to employ the best features of each forecast system.

In order for the DTC to conduct a sound scientific investigation, one important constraint is that all groups must fully disclose their model configurations by the time production runs start, and all forecasts submitted for evaluation must be generated with the same code and configuration.

5.1 Domains and grid spacing definitions

All groups are encouraged to strive to configure three grids with horizontal grid spacing O(10 km), O(4 km) and O(1 km). Due to the hydrostatic nature of the GFDL model, its smallest grid spacing will be 1/18^{o} (6.2 km). The coarsest grid should cover an area of approximately 75^{o} x 75^{o}, roughly matching the GFDL coarse domain. The GFDL grid extends from 15^{o} S to 60^{o} N, but its east-west placement varies depending on the storm location and projected track. The DTC will provide the participants with the GFDL coarsest domain specifications for each of the selected cases.

No specific guidelines were put forth at the workshop for the intermediate resolution grids. On the other hand, a minimum size of 3^{o} x 3^{o} was agreed upon for the innermost nest. No constraints were placed on the number and distribution of vertical levels, or on the placement of model top.

Modelers are free to choose nesting configurations with or without feedback between domains. To isolate the impact of high-resolution, groups using nests with feedback must run each case for the O(10 km) grid only, the O(10 km) and O(4 km) grids only, and finally all three grids.

5.2 Length of Forecasts

The cycles will be run to 126 h or to the last point in the NHC best track, including the extra-tropical stage, whichever is shorter. The prescribed length of each run is specified in the “Hours with track” column of Table 3.

5.3 Initial and Boundary Conditions

No restrictions have been placed on the global model used to produce initial and boundary conditions for the retrospective forecasts. On the other hand, most of the modeling groups expressed an interest in using a common source of GFS forecasts. Hence, the DTC will make available to all modeling groups the GFS analyses and forecasts generated using the 2007 version of GFS.

The DTC will also make available the TCM operational data from the NHC for each storm. This information includes the latitude and longitude coordinates of the storm center, intensity, and storm motion. Additionally, the DTC will provide the GFDL initial state fields, consisting of the GFS initialization for the 2007 version of the code with the embedded GFDL vortex. Modelers may choose to use the TCM data and/or the GFDL initial fields for initialization, but are not required to use this information.

Data assimilation may be employed on the coarsest grid only or on all grids. Data assimilated on the high-resolution grid must also be assimilated for the coarse resolution grids. The DTC is not responsible for providing data sets for data assimilation.

5.4 Coupling and Physics

The use of coupled models for the representation of the ocean and/or waves is at the discretion of each modeling group. Likewise, the modeling groups are free to use any physics suite they find appropriate. The physics suite (except for cumulus parameterization) and the nature of the ocean and wave coupling should be fixed as the grid spacing is altered for a given model configuration.

5.5 Post-Processing

No restrictions have been placed on the post-processing approach that each modeling group chooses to apply to the fields submitted to the DTC for evaluation except that each group should use a consistent approach for all grid resolutions. In addition, each modeling group agrees to fully disclose all aspects of their post-processing, including the interpolation scheme(s) used to move fields from their native grid to the required destaggered latitude-longitude grid.

6 Evaluation Requirements

The DTC will need at least the following hourly fields from all participant models: 1-h accumulated precipitation, MSLP, winds at 10 m AGL, 850 hPa, 700 hPa, and 500 hPa, and geopotential height at 850 hPa and 700 hPa. Additional output requirements will be communicated to the modeling groups once plans for verification and images to be posted on the web have been finalized. This information will be provided to the modeling groups no later than 1 August 2008.

Modelers need to provide these fields for the grids with O(10 km), O(4 km), and O(1 km) grid spacing. For the models using nesting with feedback between the nested domains, the O(10 km) grid should be from the run that has O(10 km) grid as the finest domain, and the O(4 km) grid should be from the run that has O(4 km) as the finest domain.

7 Evaluation Metrics

Table 4 contains a list of verification variables and metrics, with assigned feasibility and priority ratings. For each variable, the necessary observation/analyses, model fields, and tools to compute the metrics are listed. The DTC is responsible for the development of the currently non-existent verification tools and for the computation of all the verification metrics.

Most of the variables listed in Table 4 are self-explanatory. One that requires a definition is “rapid intensification” and “rapid weakening.” For this study, we will utilize the definitions proposed by Dr. M. DeMaria: 30 kt change in 24 h and 25 kt change in 24 h. These two definitions will be applied to the observed and forecasted hurricane maximum wind values for all 24-h periods when data are available to determine the occurrence or non-occurrence of rapid intensification or weakening. Counts of the occurrences and non-occurrences will be used to construct 2x2 contingency tables from which standard statistics such as Probability of Detection (POD), False Alarm Ratio (FAR), and Critical Success Index (CSI) will be computed.

Some of the statistics included in Table 4 indicate “skill” will be computed. Because skill is a relative measure of performance, a standard of comparison is required. For this study, the obvious standard for the high-resolution forecasts is the low-resolution forecasts. Thus, the report prepared by the DTC will focus on skill scores where the verification statistics for the low-resolution forecasts are used to compute the skill scores for the high-resolution forecasts. The DTC will also compute skill scores with respect to CLIPER5 and Decay-SHIFOR5 for track and intensity, respectively. These computations will be made available to the group for their own assessment and included in the archive the DTC creates for this testing activity.

Error distributions will also be presented graphically for some of the variables (e.g., center position, intensity). These plots will consist of either the error values themselves (when the sample size is very small) or box plots showing the distributions of values. In cases where the sample size is small, the error distributions can be more informative than summary statistics such as MAE.

In addition to providing verification statistics for individual storms, the verification results will be aggregated across cases, where possible, according to lead time. In some cases, the sample sizes will still be quite small, but the aggregated values will be more representative of overall performance than the statistics for individual storms.

8 Plans for Distribution of Results

The DTC will be responsible for distributing the results of the retrospective testing put forth in this test plan. The plans for distributing results are described below.

8.1 Email list

The DTC has setup a mailing list that will be used to communicate with participants. The address for this mailing list is: hrh@rap.ucar.edu.

8.2 Website for common plots

The DTC will create and post a set of common plots from all models at the three proposed grid spacing: O(10 km), O(4 km), and O(1 km). Plots for the O(10 km) domain will be based on the run for the O(10 km) only grid. The plots for the O(4 km) domain will be based on the run with only the O(10 km) and O(4 km) grids. Plots will include at least: track, surface wind and pressure, and 3-h accumulated precipitation.

8.3 Website for forecast verification metrics

The DTC will maintain a website displaying the verification statistics.

8.4 Summary Report

A preliminary report summarizing the findings will be distributed to all workshop participants and the HFIP team in late-February 2009. This report will also be posted on the DTC website.

Table 1: List of workshop participants and their affiliation

Participant	Affiliation
Fred Toepfer	NOAA/HFIP
Frank Marks	NOAA/AOML
S. Gopalakrishnan	NOAA/AOML
Robert Rogers	NOAA/AOML
Xuejin Zhang	NOAA/AOML
Kevin Yeh	NOAA/AOML
Bill Read	NOAA/NHC
Mike Fiorino	NOAA/NHC
James Franklin	NOAA/NHC
Ahsha Tribble	NOAA/NHC
Jack Beven	NOAA/NHC
Naomi Surgi	NOAA/NCEP/EMC
Steve Lord	NOAA/NCEP/EMC
Mark DeMaria	NOAA/CIRA
Jian-Wen Bao	NOAA/ESRL
Ligia Bernardet	NOAA/ESRL/DTC
Morris Bender	NOAA/GFDL
Tim Marchok	NOAA/GFDL
Isaac Ginis	University of Rhode Island
Bob Gall	NCAR/RAL/DTC
Rich Wagoner	NCAR/RAL
Louisa Nance	NCAR/RAL/DTC
Barb Brown	NCAR/RAL/DTC
Greg Holland	NCAR/ESSL/MMM
Chris Davis	NCAR/ESSL/MMM
Bo-Wen Shen	NASA/GSFC
Melinda Peng	DOD/NRL
Shuyi Chen	University of Miami/RSMAS
Greg Tripoli	University of Wisconsin-Madison
Mike Montgomery	Navy Postgraduate School
Fuquing Zhang	Texas A & M University
Roger Smith	University of Munich
Daniel Melendez	NOAA/NWS

Table 2: Modeling groups who are planning to participate in 2008-09 testing activity.

Institution (Contact)	Model	Computing Platform	Grid Spacing	Tentative Start Date	# of Cases by 2 Jan 2009	Total # of cases (completion date)
URI (Isaac Ginis)	GFDL	Oak Ridge	1/12° - 1/18° (~9 – 6 km)	mid Sept '08	All	All (2 Jan '09)
NCAR (Chris Davis)	AHW	NCAR IBM	12-4-1.33 km	~1 Oct '08	~25	All (Mar '09)
AOML (S. Gopalakrishnan)	WRF-NMM	Oak Ridge & NOAA Linux cluster	9 – 3 – 1 km	Early Nov or Dec '08	~25	All (Mar '09)
NRL (Melinda Peng)	COAMPS-TC	DOD HPC	9 - 3 km	Sept or early Oct '08	~30	All (Mar '09)
U of Wisconsin-Madison (Greg Tripoli)	UWNMS	Oak Ridge	Down to 1 km	July	~3	All (End of '09)
NASA Goddard (Bo-Wen Shen)	WRF-ARW	NASA SGI Altix	9 - 3 -1 km	Unknown	unknown	~3 months from start date
Texas A&M (Fuqing Zhang)	WRF-ARW	Texas A&M cluster	3-5 km or 1-2 km	Case studies underway	2-5	Subset (not sure)
NCEP/EMC (Naomi Surgi)	HWRF	NCEP IBM	9 km	Runs complete	All	All (complete)

Table 3: Prioritized list of test cases. Forecasts should be run in the order listed to assure a complete data set for a diverse set of cases.

Storm	Forecast Date	Forecast Time	Hours w/ track	Hours as TC	RI periods	RW periods
Wilma					5	1
	10/16/2005	0000 UTC	126	126	5	0
	10/17/2005	0000 UTC	126	126	5	0
	10/18/2005	0000 UTC	126	126	4	0
	10/19/2005	0000 UTC	126	126	0	0
	10/19/2005	1200 UTC	126	126	0	0
	10/20/2005	0000 UTC	126	126	0	0
	10/21/2005	0000 UTC	126	114	0	1
	10/22/2005	0000 UTC	114	90	0	1
	10/23/2005	0000 UTC	90	66	0	1
	10/24/2005	0000 UTC	66	42	0	1
	10/25/2005	0000 UTC	42	18	0	0
Philippe					0	0
	9/17/2005	1200 UTC	126	126	0	0
	9/18/2005	1200 UTC	126	120	0	0
	9/19/2005	1200 UTC	126	96	0	0
	9/20/2005	1200 UTC	90	72	0	0
	9/21/2005	1200 UTC	66	48	0	0
	9/22/2005	1200 UTC	42	24	0	0
Felix					8	1
	8/31/2007	1200 UTC	126	114	8	1
	9/1/2007	1200 UTC	126	90	5	1
	9/2/2007	0000 UTC	114	78	3	1
	9/2/2007	0600 UTC	108	72	2	1
	9/2/2007	1200 UTC	102	66	1	1
	9/2/2007	1800 UTC	96	60	0	1
	9/3/2007	0000 UTC	90	54	0	1
	9/3/2007	1200 UTC	78	42	0	0
Rita					6	3
	9/18/2005	0000 UTC	126	126	6	2
	9/19/2005	0000 UTC	126	126	6	3
	9/20/2005	0000 UTC	126	126	6	3
	9/21/2005	0000 UTC	126	120	2	3

Storm	Forecast Date	Forecast Time	Hours w/ track	Hours as TC	RI periods	RW periods
	9/22/2005	0000 UTC	102	96	0	3
	9/23/2005	0000 UTC	78	72	0	0
	9/24/2005	0000 UTC	54	48	0	0
Karen					2	0
	9/25/2007	0000 UTC	108	102	2	0
	9/26/2007	0000 UTC	84	78	0	0
	9/27/2007	0000 UTC	60	64	0	0
	9/28/2007	0000 UTC	36	30	0	0
Katrina					5	1
	8/24/2005	0000 UTC	126	126	4	0
	8/25/2005	0000 UTC	126	126	5	1
	8/26/2005	0000 UTC	126	114	5	1
	8/27/2005	0000 UTC	102	90	4	1
	8/28/2005	0000 UTC	78	66	1	1
	8/29/2005	0000 UTC	54	42	0	0
Humberto					2	0
	9/12/2007	1200 UTC	48	36	1	0
	9/13/2007	0000 UTC	36	24	0	0
Ingrid					0	0
	9/12/2007	1200 UTC	126	108	0	0
	9/13/2007	1200 UTC	120	84	0	0
	9/14/2007	1200 UTC	96	60	0	0
	9/15/2007	1200 UTC	72	36	0	0
Emily					11	0
	7/11/2005	0000 UTC	126	126	8	0
	7/12/2005	0000 UTC	126	126	8	0
	7/13/2005	0000 UTC	126	126	5	0
	7/14/2005	0000 UTC	126	126	4	0
	7/15/2005	0000 UTC	126	126	4	0
	7/16/2005	0000 UTC	126	126	3	0
	7/17/2005	0000 UTC	108	108	3	0
	7/18/2005	0000 UTC	84	84	3	0
	7/19/2005	0000 UTC	60	60	3	0
	7/20/2005	0000 UTC	36	36	0	0

Storm	Forecast Date	Forecast Time	Hours w/ track	Hours as TC	RI periods	RW periods
Ophelia					0	0
	9/6/2005	1200 UTC	126	126	0	0
	9/7/2005	1200 UTC	126	126	0	0
	9/8/2005	1200 UTC	126	126	0	0
	9/9/2005	1200 UTC	126	126	0	0
	9/10/2005	1200 UTC	126	126	0	0
	9/11/2005	1200 UTC	126	126	0	0
	9/12/2005	1200 UTC	126	126	0	0
	9/13/2005	1200 UTC	126	102	0	0
	9/14/2005	1200 UTC	126	78	0	0
	9/15/2005	1200 UTC	126	54	0	0
	9/16/2005	1200 UTC	126	30	0	0

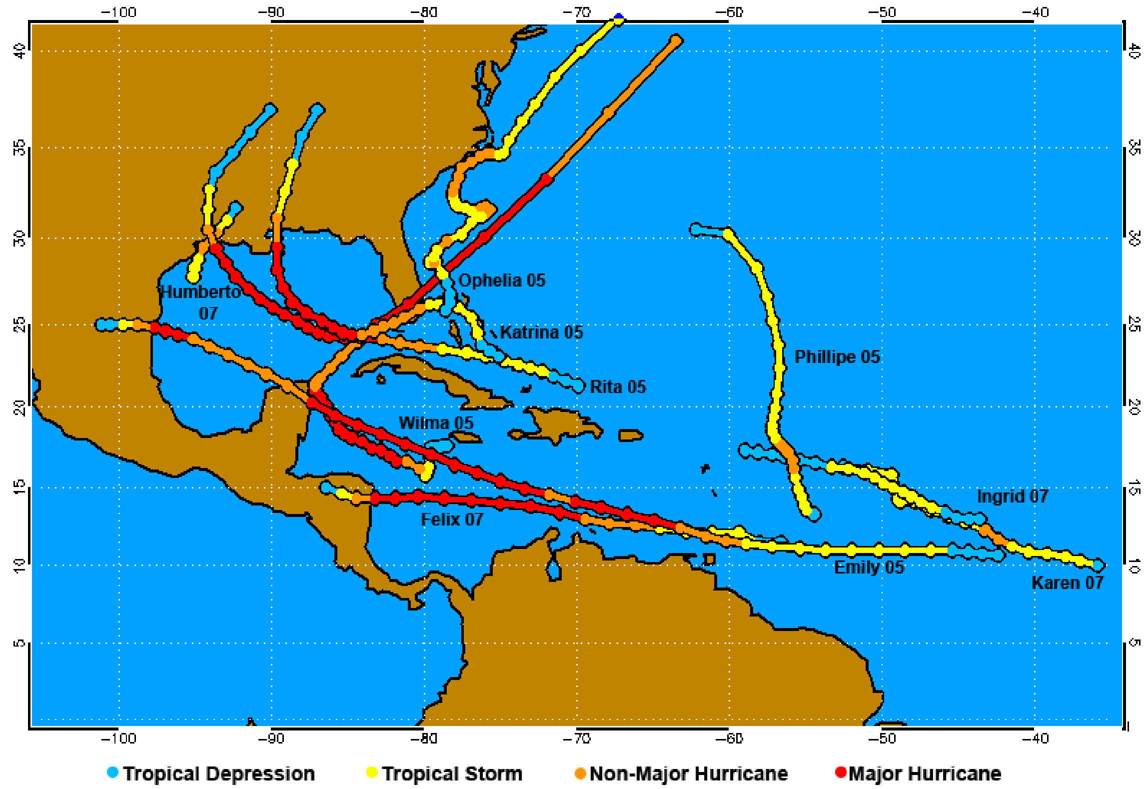

Figure 1: Track and intensity of ten storms selected for retrospective testing during 2008-09 phase of high resolution test.

Table 4: Verification Metrics for HFIP Model Forecast Evaluations

Variable		Observations/Analyses		Model		Verification		Status	
Name	**Definition**	**Type**	**Archive Location**	**Required fields**	**Tools for diagnosing variable from model fields**	**Statistical methods and metrics**	**Tool**	**Feasibility**	**Priority**
Center Position	Location of cyclone center (generally refers to the surface center)	NHC Best Track	NHC ATCF B-decks	MSLP; winds at 10-m, 850, 700, and 500 hPa; height at 850 and 700 hPa.	GFDL Vortex Tracker	MAE, error distribution, along- and cross-track MAE and bias; MAE skill scores	NHC Hurricane Forecast Verification System	High	High
Intensity	Primary measure is max 1-min wind at 10 m elevation (MSSW), secondary measure is minimum MSLP	NHC Best Track	NHC ATCF B-decks	MSLP, 10-m winds	GFDL Vortex Tracker	MAE, bias, error distribution, MAE skill score	NHC Hurricane Forecast Verification System	High	High
Occurrence of Rapid Intensifica-tion (RI)	Increase in MSSW of at least 30 kts in a 24 h period	NHC Best Track	NHC ATCF B-decks	10-m winds	GFDL Vortex Tracker	2-D contingency table, POD, FAR, CSI, skill scores	None presently exists	High	High
Occurrence of rapid weakening (RW)	Decrease in MSSW of at least 25 kts in a 24 h period for cyclones over water	NHC Best Track	NHC ATCF B-decks	10-m winds	GFDL Vortex Tracker	2-D contingency table, POD, FAR, CSI, skill scores	None presently exists	High	High
Onset of rapid intensifica-tion	Starting time of a 24 h period of RI that was not preceded by another RI event 6 h earlier (i.e., only the first in a series of consecutive and overlapping RI events)	NHC Best Track	NHC ATCF B-decks	10-m winds	GFDL Vortex Tracker	Timing error (restricted to model events that occur within 36 h of nearest actual event); MAE, bias, MSE, RMSE, RMSE skill score, error distribution	None presently exists	High	High

Variable		Observations/Analyses		Model		Verification		Status	
Name	**Definition**	**Type**	**Archive Location**	**Required fields**	**Tools for diagnosing variable from model fields**	**Statistical methods and metrics**	**Tool**	**Feasibility**	**Priority**
Onset of rapid weakening	Starting time of a 24 h period of RW that was not preceded by another RW event 6 h earlier (i.e., only the first in a series of consecutive and overlapping RW events)	NHC Best Track	NHC ATCF B-decks	10-m winds	GFDL Vortex Tracker	Timing error (restricted to model events that occur within 36 h of nearest actual event); MAE, bias, MSE, RMSE; MAE and RMSE skill scores, error distribution	None presently exists	High	High
Wind radii	Radius of max extent of winds of 34, 50, and 64 kt, by quadrant	NHC Best Track	NHC ATCF B-decks	10-m winds	GFDL Vortex Tracker	MAE, bias, RMSE, error distribution; MAE and RMSE skill scores	NHC Hurricane Forecast Verification System	High	Med
Radius of maximum wind (RMW)	Azimuthally averaged radial distance of maximum wind from center	Analysis to be provided by Jack Beven (NHC) based on a review of NHC operational estimates.	Does not exist at this time.	10-m winds	GFDL Vortex Tracker	MAE, bias, RMSE, error distributions; MAE and RMSE skill scores	None presently exists	Med	Med
Consistency	Similarity of forecast variables from one forecast time to the next	N/A	N/A	All fields listed above	GFDL Vortex Tracker	Correlation, Distributions of changes over time	None presently exists	Med	High

List of acronyms

AGL – Above Ground Level
AHW – Advanced Hurricane WRF
AOML - Atlantic Oceanographic and Meteorological Laboratory
ATCF – Automated Tropical Cyclone Forecast System
CIRA – Cooperative Institute for Research in the Atmosphere
CSI – Critical success index
DTC – Developmental Testbed Center
EMC – Environmental Modeling Center
FAR – False-alarm rate
GFDL – Geophysical Fluid Dynamics Laboratory
GFS – Global Forecasting System
HFIP - Hurricane Forecast Improvement Project
HWRF – Hurricane WRF
IBM – International Business Machines
MAE – Mean absolute error
MSE - Mean-squared error
MSLP – Mean Sea Level Pressure
MSSW – Maximum Sustained Surface Wind
NASA – National Aeronautics and Space Administration
NCAR – National Center for Atmospheric Research
NCEP – National Centers for Environmental Prediction
NHC – National Hurricane Center
NRL – National Research Laboratory
NOAA – National Oceanic and Atmospheric Administration
NOGAPS – Navy Operational Global Atmospheric Prediction System
NWP – Numerical Weather Prediction
POD – Probability of detection
QPF – Quantitative Precipitation Forecast
RI – Rapid Intensification
RW – Rapid Weakening
SST – Sea Surface Temperature
TC – Tropical Cyclone
TCM – Tropical Cyclone Forecast/Advisory Message
UWNMS – University of Wisconsin Non-hydrostatic Modeling System
WRF – Weather Research and Forecasting