NOAA

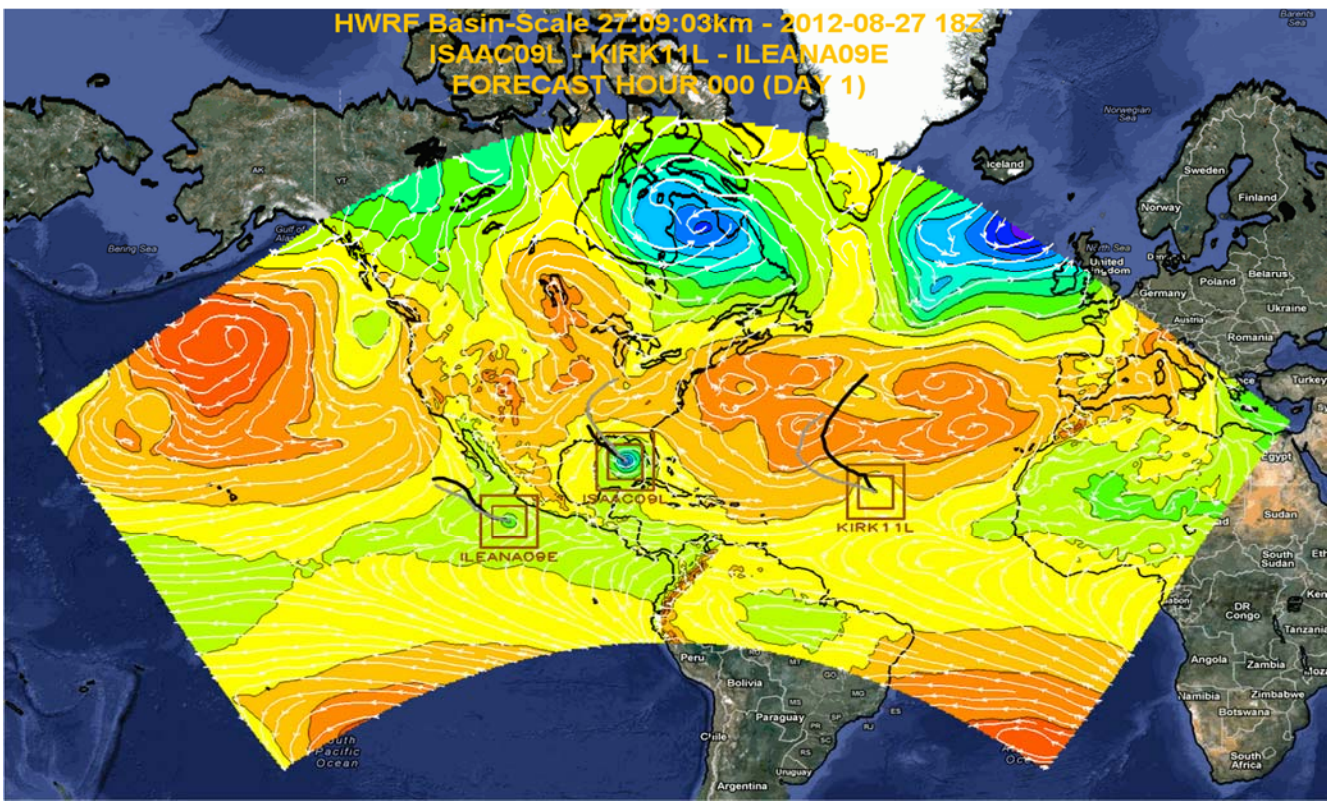
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
United States Department of Commerce

HURRICANE FORECAST IMPROVEMENT PROJECT

Hurricane Forecast Improvement Project
Years Five to Ten Strategic Plan

14 November 2013
Updated 5 November 2014

HFIP Technical Report: HFIP2014-1.1a



National Oceanic and Atmospheric Administration

Hurricane Forecast Improvement Project

Years Five to Ten Strategic Plan

R. Gall[3], F. Toepfer[3], F. Marks[1], E. Rappaport[2]

12 November 2013
Updated 5 November 2014

[1]Atlantic Oceanographic and Meteorological Laboratory OAR/NOAA
[2]National Hurricane Center NWS/NOAA
[3]Office of Science and Technology, NWS/NOAA

Table of Contents

Tables

Figures

Executive Summary

This document represents the second Five-Year Strategic Plan of the ten year Hurricane Forecast Improvement Project (HFIP). It is a continuation of the original 5-year plan published in December 2010. The overall goal of the project is to achieve a 50% improvement in hurricane numerical forecast guidance provided by the National Centers for Environmental Prediction (NCEP) to the National Hurricane Center (NHC). This improvement in guidance is for both track and intensity. HFIP also includes goals for predicting rapid intensification and for extending forecast guidance out to seven days.

Future configuration of the Hurricane Forecast System

Based on the first 4 years of results from HFIP, we project that the future operational hurricane forecast guidance system would be as described in the table below.

Component	**Specifications**
Global model ensemble with multiple interactive moving nests using Hybrid Data Assimilation	1) 20 members at 18 km, global model, 6 and 2 km for inner moving nests 2) Multi Model (at least two – FIM, GFS, Navy?) 3) Using all available aircraft and satellite data in core and near environment of hurricane 4) Run out to 7 days or more
Statistical Post Processing	1) LGEM, SHIPS, others

1. Introduction

This document represents the second Five-Year Strategic Plan of the ten-year Hurricane Forecast Improvement Project (HFIP). It is a continuation of the original 5-year plan published in December 2010. The overall goal of the project is to achieve a 50% improvement in hurricane numerical forecast guidance provided by the National Centers for Environmental Prediction (NCEP) to the National Hurricane Center (NHC). This improvement in guidance is for both track and intensity. HFIP also includes goals for predicting rapid intensification and for extending forecast guidance out to seven days.

a. Background on HFIP

Tropical cyclone activity in the Atlantic hurricane basin broke records for numbers and impacts during the first decade of the new millennium. A total of 13 hurricanes crossed the contiguous U.S. coastline from 2000-2010, including such now infamous storms as Charley (2004), Katrina (2005), Rita (2005), Wilma (2005) and Ike (2008). In 2005 alone, 27 Atlantic systems reached tropical storm status, far surpassing the previous record of 21. The heightened activity brought an increased awareness of the dangers from tropical cyclones and led to a number of studies concerning NOAA's ability to forecast hurricanes. The additional attention on the nation's hurricane warning program provided opportunities to give visibility to and initiate actions on intensity forecasting, a critical area where no appreciable improvement had been made over the preceding two decades (e.g., Cangialosi and Franklin 2011). To address this issue, NOAA, through its Science Advisory Board (SAB), established a Hurricane Intensity Research Working Group (HIRWG), which documented its recommendations to improve forecasts of hurricane intensity in October 2006 (NOAA SAB 2006). In addition, the National Science Foundation (NSF) National Science Board issued a report in January 2007 on the need for a National Hurricane Research Initiative (NSF 2007) and the Office of the Federal Coordinator of Meteorological Services (OFCM) issued a report in February 2007 calling for a federal investment of $70-85 million annually over the next 10 years for tropical cyclone research and development, transition of research to operations, and operational high performance computing (OFCM 2007).

NOAA's response was the establishment of HFIP, as noted in this November 2007 statement: "In response to the HIRWG report, NOAA convened a corporate hurricane summit developing unified strategy to address hurricane forecast improvements. On May 10, 2009, the NOAA Executive Council established the NOAA Hurricane Forecast Improvement Project, a 10-year effort to accelerate improvements in one- to five-day forecasts for hurricane track, intensity, and storm surge and to reduce forecast uncertainty, with an emphasis on rapid intensity change" (NOAA SAB 2008). In July 2008-July 2009 the President's proposed budget was amended to include $13M for HFIP, and this increment became part of NOAA's base budget.

This report describes the HFIP project, its goals, proposed methods for achieving those goals, and recent results from the program with an emphasis on recent advances in the skill of the operational hurricane forecast guidance.

b. The Hurricane Forecast Improvement Project

HFIP provides the unifying organizational infrastructure and funding for NOAA and other agencies to coordinate the hurricane research needed to significantly improve guidance for hurricane track, intensity, and storm surge forecasts. HFIP's 5-year (for 2014) and 10-year goals (for 2019) are:

- Reduce average track errors by 20% in 5 years, 50% in 10 years for days 1 through 5.
- Reduce average intensity errors by 20% in 5 years, 50% in 10 years for days 1 through 5.
- Increase the probability of detection (POD)[1] for rapid intensification (RI)[2] to 90% at day 1 decreasing linearly to 60% at day 5, and decrease the false alarm ratio (FAR) for rapid intensity change to 10% for day 1 increasing linearly to 30% at day 5. The focus on rapid intensity change is the highest-priority forecast challenge identified by NHC.
- Extend the lead-time for hurricane forecasts out to day 7 (with accuracy equivalent to that of the day 5 forecasts in 2006, approximately 260 nautical miles).

Forecasts of higher accuracy and greater reliability are expected to lead to higher user confidence and improved public response, resulting in savings of life and property. Reaching these goals, however, requires major investments in enhanced observational strategies, improved data assimilation, numerical model systems, expanded forecast applications based on the high-resolution and ensemble-based numerical prediction systems and improved computational infrastructure. NOAA also recognizes that addressing the challenges associated with improving hurricane forecasts requires interaction with, and the support of, the larger research and academic communities.

It is hypothesized that the ambitious HFIP goals could be met with high-resolution (~10-15 km) global atmospheric numerical forecast models run as an ensemble in combination with, and as a background for, regional models at even higher resolution (~1-5 km). In order to support the significant computational demands of such an approach, HFIP has developed a high-performance computational system in Boulder, Colorado. Demonstrating the value of advanced science, new

---

[1] POD, Probability of Detection, is equal to the total number of correct RI forecasts divided by the total number of forecasts that should have indicated RI: number of correctly forecasted ÷ (correctly forecasted RI+ did not, but should, have forecasted RI). FAR, False Alarm Ratio, is equal to the total number of incorrect forecasts of RI divided by the total number of RI forecasts: forecasted RI that did not occur ÷ (forecasted RI that did occur + forecasted RI that did not occur).

[2] Rapid Intensification (RI) for hurricanes is defined as an increase in wind speed of at least 30 knots in 24 hours. This goal for HFIP also applies to rapid weakening of a decrease of 25 knots in 24 hours.

observations, higher-resolution models, and post-processing applications is necessary to justify obtaining the commensurate resources required for robust real-time use in an operational environment.

For FY2012, HFIP program funding was approximately $7M, with an additional $4M dedicated to enhancing computer capacity available to the Program. This level was approximately half that of previous years. The funding for computing was used to enhance the HFIP system established in Boulder, Colorado in FY2009, and resulted in machines called t-jet and u-jet with a total of 23,000 processors. The $7M was distributed to: 1) various NOAA laboratories and centers, including the Environmental Modeling Center (EMC), Geophysical Fluid Dynamics Laboratory (GFDL), National Environment Satellite Data and Information Service (NESDIS), the Earth System Research Laboratory (ESRL) the Atlantic Oceanographic and Meteorological Laboratory (AOML), and NHC; 2) the National Center for Atmospheric Research (NCAR); 3) the Naval Research Laboratory in Monterey (NRL), and 4) several universities: University of Wisconsin (UW), The Pennsylvania State University (PSU), Colorado State University (CSU), Florida State University (FSU), University of Rhode Island (URI), University of Miami (UM), University of Colorado (UC), University of Maryland (UMD), University of California, Los Angeles (UCLA), University of Oklahoma (OU), and the State University of New York (SUNY), Albany (awarded through a NOAA Announcement of Opportunity) and the National Oceanographic Partnership Program (NOPP). Specifically, $1M was contributed each year for three years to the NOPP, and through an Announcement of Opportunity, for competed proposals related to improving understanding and prediction of hurricanes. The funding to NOPP from HFIP was matched by funding from the Office of Naval Research.

Distribution of the $7M was based on recommendations from nine teams focused on various components of the hurricane forecast problem. The current teams, made up of over 50 members drawn from the hurricane research, development and operational communities, are listed in Tables 1 and 2 along with the team co-leaders.

Table 1. Strategic Teams

FY 2014 Teams	**FY 2014 Team Leads**
HFIP Model/Physics Strategy	Vijay Tallapragada (EMC), Jian-Wen Bao (ESRL)
Data Assimilation / Initialization / Ensemble Development	Jeff Whitaker (ESRL), Xuguang.Wang (OU)
Post Processing and Verification Development	Mark DeMaria (NHC), David Zelinsky (NHC), Tim Marchok (GFDL)
Societal Impacts	Jennifer Sprague (NWS), Rick Knabb(NHC)

Table 2. Tiger Teams

<u>**FY 2014 Teams**</u>	<u>**FY2014 Team Leads**</u>
Web Page Design	Paula McCaslin (ESRL), Laurie Carson (ESRL)
Regional Hybrid DA System / Use of Satellite Data	Jeff Whitaker (ESRL), Xiaolei Zou (FSU)
Stream 1.5 and Demonstration System Implementation	James Franklin (NHC), Barb Brown (NCAR)
Reconnaissance Data Impact	James Franklin (NHC), Vijay Tallapragada (EMC)
Ocean Model Impact	Hyun-Sook Kim (EMC), Rich Yablonsky (URI)

HFIP's focus and long-term goal is to improve the numerical model guidance that is provided by NCEP operations to NHC as part of the hurricane forecast process. To accomplish this goal, the program is structured along three somewhat parallel development paths, known as "streams". Stream 1 is directed toward developments that can be accomplished using operational computing resources (either existing or planned). This stream covers development work planned, budgeted and executed over the near term (mostly one to two years) by EMC with HFIP augmenting support to enable participation by the broader modeling community. Since Stream 1 enhancements are implemented into operational forecast systems, these advances are automatically available to the Hurricane Specialists at NHC in the preparation of official forecast and warning products.

While Stream 1 works within presumed operational computing resource limitations, Stream 2 activities assume that resources will be found to greatly increase available computer power in operations above that planned for the next five years. The purpose of Stream 2 is to demonstrate that the application of advanced science, technology, and increased computing will lead to the desired increase in accuracy and other aspects of forecast performance. Because the level of computing necessary to perform such a demonstration is large, HFIP developed its own computing system at NOAA/ESRL in Boulder, Colorado.

A major component of Stream 2 is an Experimental Forecast System (EFS) that HFIP runs each hurricane season. The purpose of the EFS (also known as the Demonstration Project) is to evaluate the strengths and weaknesses of promising new approaches that are testable only with enhanced computing capabilities. The progress of Stream 2 work is evaluated each off-season to identify techniques that appear particularly promising to operational forecasters and/or modelers. These potential advances can be blended into the operational implementation plans through subsequent Stream 1 activities, or developed further outside of operations within Stream 2. Stream 2 models represent cutting-edge approaches that have little or no track record; consequently NHC forecasters do not use these models to prepare their operational forecasts or warnings.

HFIP was originally structured around this two-stream approach. However, it quickly became apparent that some Stream 2 research models were producing forecast guidance that was potentially useful to forecasters. Because these models could not be implemented at NCEP due to insufficient operational computing resources, a third activity, known as Stream 1.5, was initiated to expedite the testing and availability of promising new models to forecasters. Stream 1.5 is an approach that accelerates the transfer of successful research from Stream 2 into real-time forecasting, by following a path that temporarily bypasses the budgetary and technical bottlenecks associated with traditional operational implementations.

The Stream 1.5 process for the each hurricane season involves extensive evaluation, by the Tropical Cyclone Modeling Team (TCMT) at NCAR, of the previous season's most promising Stream 2 models or techniques. This testing involves rerunning the models or techniques over storms from the demonstration period (August 1 to October 31) for the three previous seasons involving several hundred cases. If operational computing resources are not available for immediate implementation, those enhancements that meet certain pre-defined standards for improvement over existing techniques can be run on HFIP computing resources and be provided to NHC forecasters in real-time during the upcoming hurricane season as part of the EFS. This process moves forward the availability of real-time advances to forecasters one or more years. It also serves as a proof of concept for both the developmental work (Stream 2) and augmented computational capabilities.

2. The HFIP Model Systems

HFIP believes that the best approach to improving hurricane track forecasts, particularly beyond four days, involves the use of high-resolution global models, with at least some run as an ensemble. However, global model ensembles are likely to be limited by computing capability for at least the next five years to a resolution no finer than about 15-20 km, which is inadequate to resolve the inner core of a hurricane. It is generally assumed that the inner core must be resolved to see consistently accurate hurricane intensity forecasts (e.g., HIRWG Report). Maximizing improvements in hurricane intensity forecasts will therefore likely require high-resolution regional models, perhaps also run as an ensemble. Below we outline the modeling systems currently in use by HFIP.

a. The Global Models

Global models provide the foundation for all of HFIP's modeling effort. They provide hurricane forecasts of their own, and are top-tier performers for hurricane track. They also provide background data and/or boundary conditions for regional and statistical models, and can be used to construct single-model ensembles, or be members of multi-model ensembles. HFIP EFS involves two global models: the Flow-following finite-volume Icosahedral Model (FIM) and the Global Forecast System (GFS).

Built by ESRL, FIM is an experimental global model that can be run at various resolutions and uses initial conditions from a number of sources (Benjamin et. al. 2004, NOAA ESRL 2011). It is currently using a constant sea surface temperature underneath.

Two versions of GFS, the NWS's global model, are currently in use by the HFIP EFS. One of these is the current operational model run at NOAA NCEP. The second is an experimental version developed at ESRL, which differs from the operational GFS by featuring a fixed ocean and, at least in 2014, using semi-Lagrangian time differencing, allowing higher resolution. The operational GFS is expected to transition to semi-Lagrangian in 2015. The experimental version of GFS is maintained at ESRL to allow testing before new technologies become operational.

Some specifics of the global models are shown in Table 3.

Table 3. Specifications for the HFIP Global Models

Models	Horizontal resolution	Vertical levels	Cumulus Parameterization	Microphysics	Planetary Boundary Layer (PBL)	Land Surface Model (LSM)	Radiation	Initialization
FIM	15km	64	From 2011 GFS – Simplified Arakawa Schubert	Zhao-Carr	GFS Non-local PBL	Noah LSM	Rapid Radiative Transfer Model (RRTM)	GFS GSI operational hybrid-ensemble variational
GFS/Experimental Semi-Lagrangian	18 km	64	Simplified Arakawa Schubert	Ferrier	GFS Non-Local PBL	Noah LSM	GFDL scheme	GSI/hybrid
GFS/Operational	27 km	64	Simplified Arakawa Schubert	Ferrier	GFS Non-Local PBL	Noah LSM	GFDL	GSI/hybrid

b. The Regional Models

Specifics of the regional models are shown in Table 4. Note that GFDL (OPS) and HWRF (OPS) refer to the current operational regional models. The Weather Research and Forecasting (WRF) modeling system in use by HFIP contains two options for its dynamic core, and several options for physics as well as initialization and post processing systems (for a list of model publications, see DTC 2012a). The two dynamic core configurations are the Advanced Research WRF (ARW), built by NCAR, and the Non-hydrostatic Mesoscale Model (NMM), built by EMC.

Table 4. Specifications for the HFIP Regional Models

Models	Domains / Horizontal Resolution (km)	Vertical Levels core	Cumulus Parameterization	Microphysics	PBL	Land Surface	Radiation	Initial and Boundary Conditions	Initialization
HWRF (OPS)	3 27/9/3	42 NMM	Simplified Arakawa Schubert (27/9 km only)	Ferrier	GFS Non-Local PBL	GFDL Slab Model	GFDL Scheme	GDAS and GFS	One-way hybrid GSI-EnKF with vortex initialization
GFDL (OPS)	3 55/18/9	42 GFDL	Simplified Arakawa Schubert	Ferrier	GFS Non-Local PBL	GFDL Slab Model	Schwarzkopf-Fels (longwave) / Lacis-Hansen (shortwave)	GFS	GFDL Synthetic Bogus Vortex
HWRF-HRD/EMC Basin Scale	3 27/9/3	61 NMM	Simplified Arakawa Schubert	Ferrier	GFS Non-Local PBL	GFDL Slab Model	GFDL Scheme	GFS	Vortex initialization
HWRF-HRD (HEDAS)	2 9/3	42 NMM	Simplified Arakawa Schubert	Ferrier	GFS Non-Local PBL	GFDL Slab Model	GFDL Scheme	GFS	EnKF with aircraft and satellite data
ARW (NCAR)	3 36/12/4	36 ARW	Tiedtke (36/12 km only)	WSM6	Yonsei University (YSU)	NOAH LSM	RRTM Longwave and Shortwave)	GFS (BC only)	EnKF method in a 6-hour cycling mode
COAMPS-TC	3 45/15/5 (15/5 km follow the storm)	40 COAMPS	Kain Fritsch on 45 and 15 km meshes	Explicit microphysics (5 class bulk scheme)	Navy 1.5 Order Closure	Slab with the NOAH LSM as an option	Fu-Liou	GFS in WATL and EPAC	3D-Var with synthetic observations
Wisconsin NMS	2 45/4.1	42 UW-NMS	Modified Emanuel	Tripoli-Flatau (1 liquid, 2 ice categories)	1.5 Order Closure	NOAH LSM	RRTM Longwave and Shortwave	GFS	Bogus vortex with 12-hour dynamic initialization
Penn State AHW-EnKF	3 27/9/3	43 AHW	Grell-Devenyi ensemble scheme (27 km only)	WSM 6-class graupel scheme	YSU	5-layer thermal diffusion scheme	RRTM (longwave) / Dudhia (shortwave)	GFS	Cycling EnKF with all Recon data

The operational NCEP Hurricane WRF (HWRF) derives from the NMM dynamic core and has a movable, triply nested grid capability for one 3-km innermost nest and one 9-km intermediate nest. A coarser outer domain covers an 80° x 80° region at 27-km resolution. The model has 42 vertical layers. Advanced physics include atmosphere/ocean fluxes, coupling with the Princeton Ocean Model, and the NCEP GFS and GFDL physics. HFIP also supports the WRF ARW system, which NCAR runs using a simplified one-dimensional model of the ocean. It features three interactive nests with an inner-nest resolution of 4 km.

The PSU Regional Ensemble constitutes another version of the WRF ARW system, with similarities to the NCAR WRF ARW. It uses a static interactive inner nest of 3 km but no interactive ocean (PSU 2011; Zhang et al 2011; Weng and Zhang 2012; Snyder and Zhang 2003).

The Coupled Ocean/Atmosphere Mesoscale Prediction System – Tropical Cyclone (COAMPS-TC©) (Doyle et al. 2012) and the Wisconsin Model are detailed in the table and have been members of the stream 1.5 suite of models each year. Note that COAMPS-TC© features an interactive ocean (Chen et al. 2010 and Doyle et al. 2012).

c. Initialization and Data Assimilation Systems

A number of approaches are used to create the initial state for the global and regional models in the HFIP EFS:

1. GFS: The initial state created for the current operational global model, GFS, is interpolated to the grids used by HFIP global models. The GFS in 2012 used the new Hybrid Ensemble-Variational DA System (HEVDAS—see below) that is a combination of the GSI system formerly used and an ensemble based system to define the background error matrix. The GSI initialization system that has been run operationally since 2006 is a three-dimensional variational approach (3D-Var) (DTC 2012; Wu et al. 2002; Parrish and Derber 1992; Cohn and Parrish 1991).

2. HWRF: The 2012 operational HWRF used an advanced vortex initialization and assimilation cycle consisting of four major steps: 1) interpolation of the global analysis fields from the Global Data Assimilation System (GDAS) onto the operational HWRF model grid; 2) removal of the GFS vortex from the global analysis; 3) addition of the HWRF vortex modified from the previous cycle's six-hour forecast based on observed location and strength (or use of a corrected GDAS or bogus vortex for a cold start); and 4) addition of data observed outside of the hurricane area using GSI. The one-way hybrid ensemble-variational DA system, designed for regional hurricane applications and implemented for operations in the 2013 HWRF, is based on HEVDAS ensembles and is capable of ingesting inner core data to optimize the vortex initialization.

3. NRL Atmospheric Variational Data Assimilation System (NAVDAS): This is the system used to provide the initial conditions to the Navy Global Environmental Model (NAVGEM). Previously a 3D-Var system, it was upgraded in September 2009 to NAVDAS-AR, a four-dimensional variational (4D-Var) approach (NRL 2001, Daley and Barker 2001). The 3D-Var version of NAVDAS is used to initialize COAMPS-TC.

4. EnKF: This is an advanced assimilation approach, somewhat like 4D-Var, that uses an ensemble to create background error statistics for a Kalman filter (Tippett et al 2003, Keppenne 2000, Evensen 1994, Houtekamer et al 1998). This approach has shown considerable promise (Hamill et al 2011). For example, the Hurricane Ensemble Data Assimilation System (HEDAS), developed at AOML, is an EnKF system applied to the HWRF (Aksoy et al 2012).

5. Hybrid: This system combines aspects of the EnKF and 3D- or 4D-Var, such as using the ensemble of forecasts to estimate the covariances at the start of the variational component of the DA system. This technology was developed at EMC, Oceanic and Atmospheric Research (OAR)/ESRL and AOML/Hurricane Research Division (HRD) and was used in operations for the 2012 season.

6. Vortex Initialization: The initial vortex for some of the regional models is produced by a vortex initialization procedure. First, the vortex circulation is filtered from the first guess

fields interpolated from global model; then a new vortex modified by the observed intensity is inserted back in the filtered environment. The new vortex is the model balanced vortex cycled from previous six-hour forecast or defined based on a synthetic vortex profile. On the first initialization for a particular storm, the size and intensity of the vortex are modified based on real-time observations. In the HWRF system, the tropical cyclone vortex is cycled from the previous six hour forecast and the vortex is relocated based on the observed position. The one-way hybrid GSI-EnKF DA system assimilates the modified vortex and ambient fields to generate initial conditions for the HWRF system. Vortex relocation is also utilized by the current operational GFS and Global Ensemble Forecast System (GEFS) in NCEP.

d. The HWRF Community Code Repository

During 2009-2011 EMC and the Developmental Testbed Center (DTC) unified the operational and research versions of HWRF and created a code management protocol, making the operational model completely compatible with the codes in the central DTC repository. In 2012, researchers had direct access to the repository, encouraging code-sharing and allowing development using the latest code, making improvements in HWRF easily transferable into operations. This was one of the initial goals of the WRF program. DTC maintains a portal (www.dtcenter.org/HurrWRF/users) providing access to HWRF and AHW documentation, datasets, and tutorials. That portal includes a link to the GFDL vortex tracker for users with models other than HWRF.

3. The HFIP Goals

To measure progress toward meeting the HFIP goals outlined in the introduction, a baseline level of accuracy was established to represent the state of the science at the beginning of the program. Results from HFIP model guidance could then be compared with the baseline to assess progress. HFIP accepted a set of baseline track and intensity errors developed by NHC, in which the baseline was the consensus (average) from an ensemble of top-performing operational models, evaluated over the period 2006-2008. For track, the ensemble members were the operational aids GFSI, GFDI, UKMI, NGPI, HWFI, GFNI, and EMXI, while for intensity the members were GHMI, HWFI, DSHP, and LGEM (Cangialosi and Franklin 2011). Figure 1 shows the mean errors of the consensus over the period 2006-2008 for the Atlantic basin. A separate set of baseline errors (not shown) was computed for the eastern North Pacific basin.

The baseline errors in Figure 1 are also compared to the errors for the same cases for the Climatology and Persistence model (CLIPER5) for track and the Decay Statistical Hurricane Intensity Forecast (Decay-SHIFOR5) model for intensity (NHC 2009). Errors from these two models are large when a storm behaves in an unusual or rapidly changing way, and therefore are useful in assessing the inherent difficulty in a set of forecasts. When a track or intensity model error is normalized by the CLIPER5 or Decay-SHIFOR5 error, the normalization yields a measure of the model’s skill.

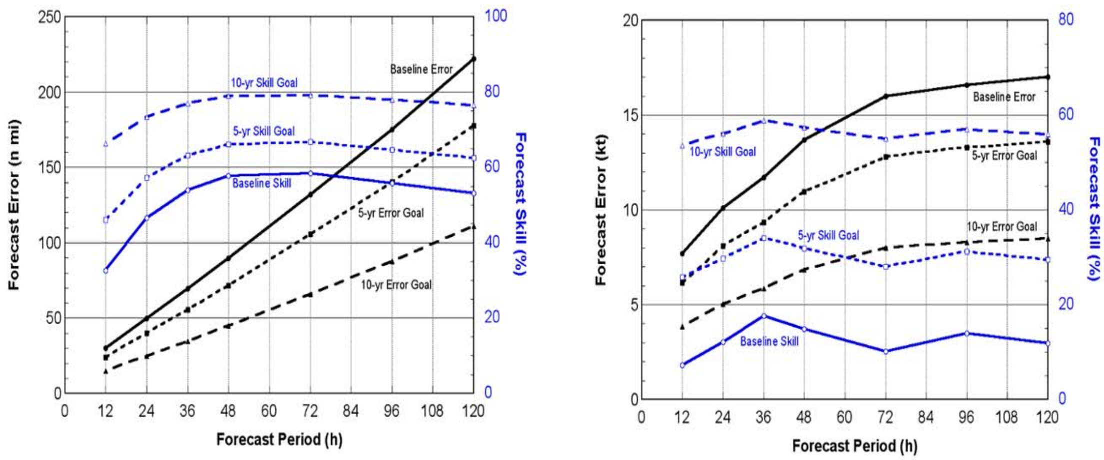


Figure 1. Track and Intensity Error Baselines
HFIP baseline track (left panel) and intensity errors (right panel). The baseline errors (black lines) were determined from an average of the top-flight operational models during the period 2006-2008. The HFIP expressed goals (dashed lines) are to reduce this error by 20% in 5 years and 50% within 10 years. Comparisons of forecasts over non-homogenous samples, however, are best done in terms of skill. To obtain the 5-year and 10-year HFIP goal in terms of skill (blue lines—baseline skill in solid, HFIP goals dashed), the goals are expressed as the percentage improvement over the CLIPER5 errors (track) and Decay-SHIFOR5 (intensity) of the baseline sample (see text).

Because a sample of cases from, say, the 2013 season might have a different inherent level of difficulty from the baseline sample of 2006-2008 (for example, because it had an unusually high or low number of rapidly intensifying storms), evaluating the progress of the HFIP models in terms of forecast skill provides a longer-term perspective. Figure 1 shows the baseline errors and the 5- and 10-year goals as skill, represented in blue and labeled on the right side of the graph. Skill in the figure is the percentage improvement over the Decay-SHIFOR5 and CLIPER5 forecasts for the same cases. Note the skill baseline and goals for intensity at all lead times is roughly constant with the baseline representing a 10% improvement over Decay-SHIFOR5 and the 5- and 10-year goals, representing 30% and 55% improvements, respectively.

It is also important to note that the HFIP performance baselines were determined from a class of operational aids known as “early” models. Early models are those that are available to forecasters early enough to meet forecast deadlines for the synoptic cycle. Nearly all the dynamical models currently in use at tropical cyclone forecast centers, however, (such as the GFS or the GFDL model, referred to as GFDL, for short) are considered “late” models because their results arrive too late to be used in the forecast for the current synoptic cycle. For example, the 1200 Coordinated Universal Time (UTC) GFDL run does not become available to forecasters until around 1600 UTC, whereas the NHC official forecast based on the 1200 UTC initialization must be issued by 1500 UTC, one hour before the GFDL forecast can be viewed. It’s actually the older, 0600 UTC run of the GFDL that would be used as input for the 1500 UTC official NHC forecast, through a procedure developed to adjust the 0600 UTC model run to match the actual storm location and intensity at 1200 UTC. This adjustment, or interpolation, procedure creates the 1200 UTC "early" aid GFDI that can be used for the 1500 UTC NHC forecast. Model results so adjusted are denoted with an "I" (e.g., GFDI). The distinction between early and late models is important to assessing model performance, since late models have an advantage of more recent observations/analysis than their early counterparts.

Strategies for meeting these goals are noted in sections 6 and 7.

4. Results from the first 5 years of HFIP

In the next sections we will outline results from the first 4 years of the project. It is from these results that the final 5-year program for the project is built. We refer the reader to the various annual reports available on line at http://hfip.org. In fact, much of this section was taken from the most recent Annual reports, 2012 and 2013. The previous strategic plan is also available at that site. In going back over the milestones listed in the original 5-year plan, most of the milestones were reached except for some of the operational implementation milestones that were delayed because of insufficient operational computing power. That is currently being remedied by NWS.

a. Operational Hurricane Guidance Improvements

The HFIP goals described in section 4 are only met when the model guidance provided to NHC by NCEP reaches those goals. Since 2013 represents the fourth year of the project we would expect to see progress toward meeting the five-year goals in the operational models and not just in the experimental models. In this section the emphasis will be on improvements in the hurricane forecasts from the models operational in 2012 and 2013. This includes the global GFS model, and the HWRF and GFDL operational regional models.

i. Global Model (GFS)

In May of 2012 the GSI data assimilation system in the GFS was replaced by the hybrid data assimilation system. The hybrid system uses an ensemble to generate a flow dependent background error covariance matrix which is then used in the GSI for the analysis. The reader may note that in previous annual reports starting with the first one in 2010 we have described the impact of changing the data assimilation system in the global models, particularly the GFS from the 3D-Var GSI to an ensemble based system, called Ensemble Kalman Filter (EnKF). The hybrid system is basically a combination of the EnKF and the GSI and has been shown to provide somewhat better results than EnKF alone. HFIP regards the implementation of the hybrid system, which it has promoted, a component of transferring HFIP results into operations.

Figure 2 shows track errors as a percent improvement over CLIPER5 (i.e., skill). Here we just emphasize the three operational numerical prediction guidance systems that NCEP runs related to hurricanes; the global model GFS, the HWRF and the GFDL. HWRF and GFDL will be discussed further in the next sections. The baseline and the 5-year goal of a 20% improvement in track guidance are also shown on the plot as skill.Note that all three models are showing track skills at or near the 5-year goal and the GFS shows skill considerably above the 5-year goal, approaching the 10-year goal of a 50% decrease in track error (at 72 hours the 10 year goal is an 80% improvement over CLIPER5). Since comparing relative to CLIPER5 errors removes most of the year to year variation from the results, what is displayed in Figure 2 represents a true improvement in skill and is not simply due to 2012 being an easy year for track (if indeed it was).

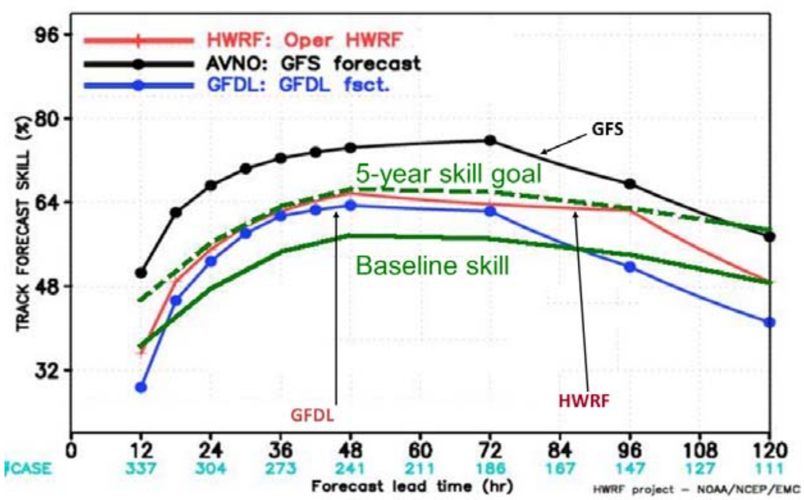


Figure 2. NCEP Operational Models Track Skill
Track Skill (error relative to CLIPER5) for the NCEP Operational models including the global and regional models. Models are indicated on the panel. The green dashed line is the 5-year HFIP goal shown in terms of skill

Figure 3 shows intensity skills for the 2012 operational systemsrelative to the skill of Decay-SHIFOR5. Here, when compared to the improvements in track forecasts, the results are not nearly as impressive. None of the models even come up to the baseline skill let alone the 5-year goal. Surprisingly, the intensity forecasts by the GFS actually exceeded, in skill, the intensity forecasts of the two operational regional models beyond 48 hours. The HWRF did better than either the GFS or the GFDL in the first 36 hours though not at the skill level of the Decay-SHIFOR5.

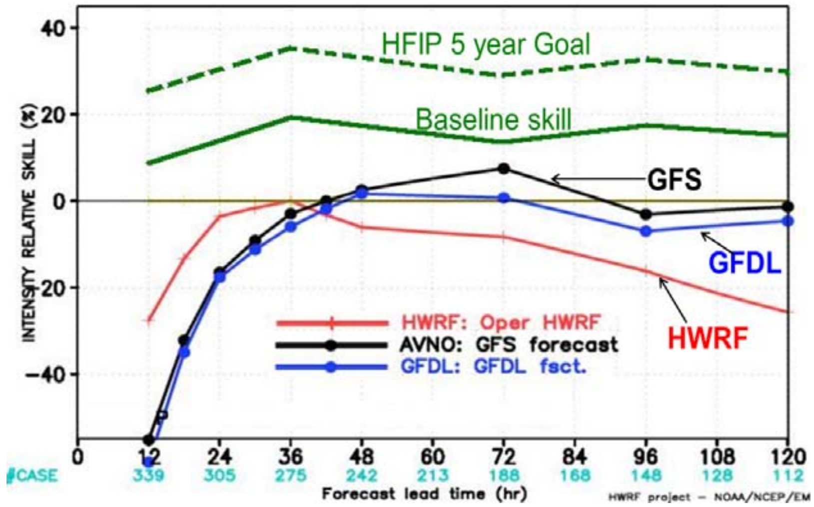


Figure 3. NCEP Operational Models Intensity Skill
Intensity Skill (error relative to Decay-SHIFOR5) for the NCEP Operational models including the global and regional models. Models are indicated on the panel. The green dashed line is the 5-year HFIP goal shown in terms of skill

Further discussion of the intensity skill of the most recent HWRF (2013) follows.

ii. Hurricane WRF (HWRF)

Even though the results for HWRF in 2012 fell short of the 5-year HFIP goal for intensity, Figure 3, the HWRF model has undergone a vast improvement over previous versions of the operational HWRF. There were two major changes in HWRF for the 2012 hurricane season: the introduction of a third nest allowing an inner core resolution of 3 km and a stream 2.0 demonstration of using aircraft data in HWRF. There were other changes including physics package upgrades, bug fixes etc. that also led to improvements in the model but the introduction of the third nest was a game changer.

Up until the 2012 hurricane season, HWRF had been run with two nests; an outer domain with a 27-km grid spacing and an inner nest of 9 km. The inner nest moves with the hurricane and interacts with the outer domain. HFIP results from previous years and described in annual reports indicated that a higher resolution inner or third nest provided superior results. Hence the HWRF team at EMC undertook a project to include an interactive third nest with a resolution of 3 km that would fit in the operational time slot allotted to the HWRF system. This was a major effort requiring, among other things, making the code more efficient so that it would fit in the time slot which was only modestly expanded for this improvement.

Figure 4 shows the errors for intensity for the HWRF using the 2011 operational model, the 2012 operational model and the 2013 operational model. Errors averaged over three years are shown (the 2011 results are for 2008-2010; the 2012 and 2013 results are for 2010-2012). Between

2011 and 2013 improvements included adding the third (3-km) nest in 2012 as noted above and adding a one way hybrid (using the global model ensemble to define the background errors in the DA scheme rather than the HWRF model itself) in 2013. In 2013 the frequency of calls to the physics packages was increased as well. The improvement in the intensity scores has been impressive, dropping about 15% between 2011 and 2012 and again between 2012 and 2013. By 2013, the scores were not yet up to the HFIP 5-year goal but if this rate of improvement continues (which is expected) we should reach the 5-year goal on time. Figure 4 also shows the results of retrospective tests of the 2014 version of HWRF run over the 2008-2013 seasons. Beyond 60 hours, the new model, at least for the retrospective runs, exceeds the HFIP 5-year intensity goal.

This level of improvement from one annual version to the next has never before been observed for the operational hurricane models. We attribute this improvement largely to the introduction of the third nest with its higher resolution near the hurricane core and the aforementioned changes to the physics packages.

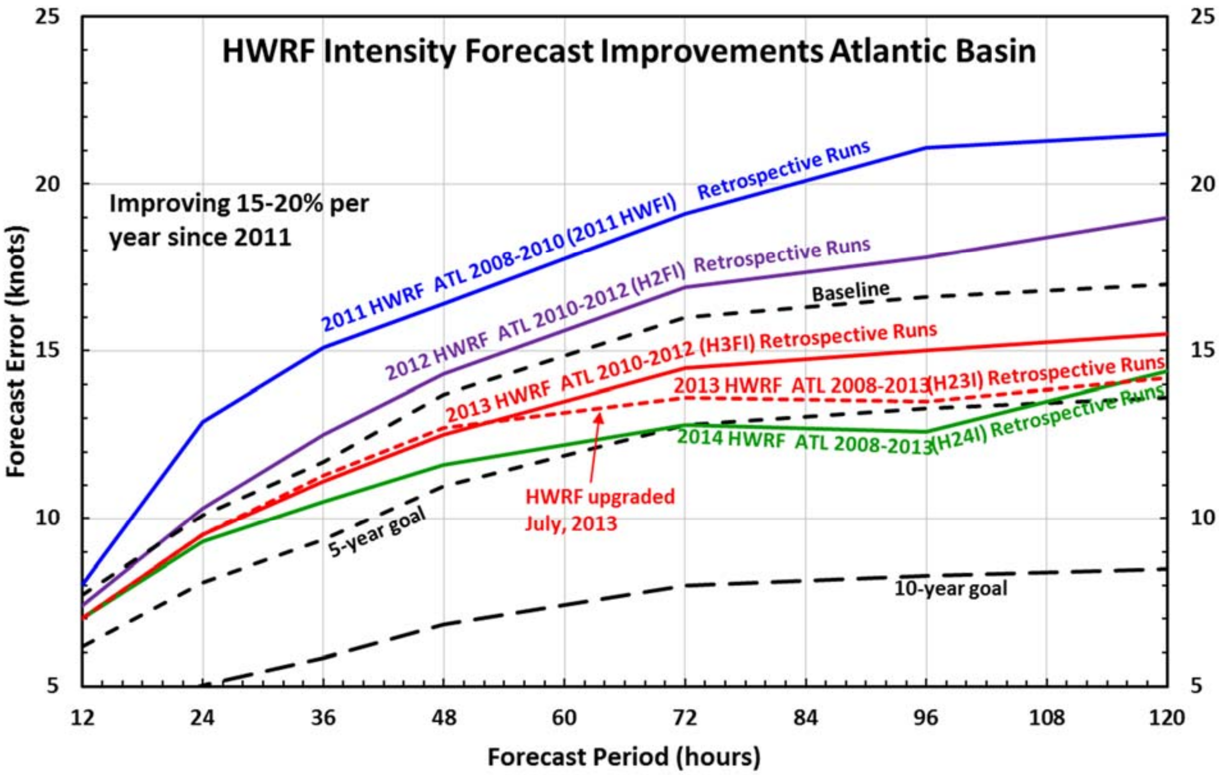


Figure 4. HWRF Intensity Forecast Improvements 2011-2014
Intensity errors for HWRF Operational models for the Atlantic. Dashed lines indicate the baseline and 5- and 10-year intensity error goals.

Figure 5 shows official NHC intensity forecast errors from 1990.   Notable in this figure is the rapid decrease in official intensity errors since 2009. While this decrease could be due in part to improvements in numerical guidance such as that shown in figure 4 it is still a bit early to draw such a conclusion. There are other years with similar dramatic changes in skill (e.g., 2003 and 2008) which are related to peculiarities of a particular season.   The only way to be sure that the 2009-2013 drop is not also related to such a season to season change in forecast difficulty is to wait and see if the trend continues. Still, the figure is encouraging.

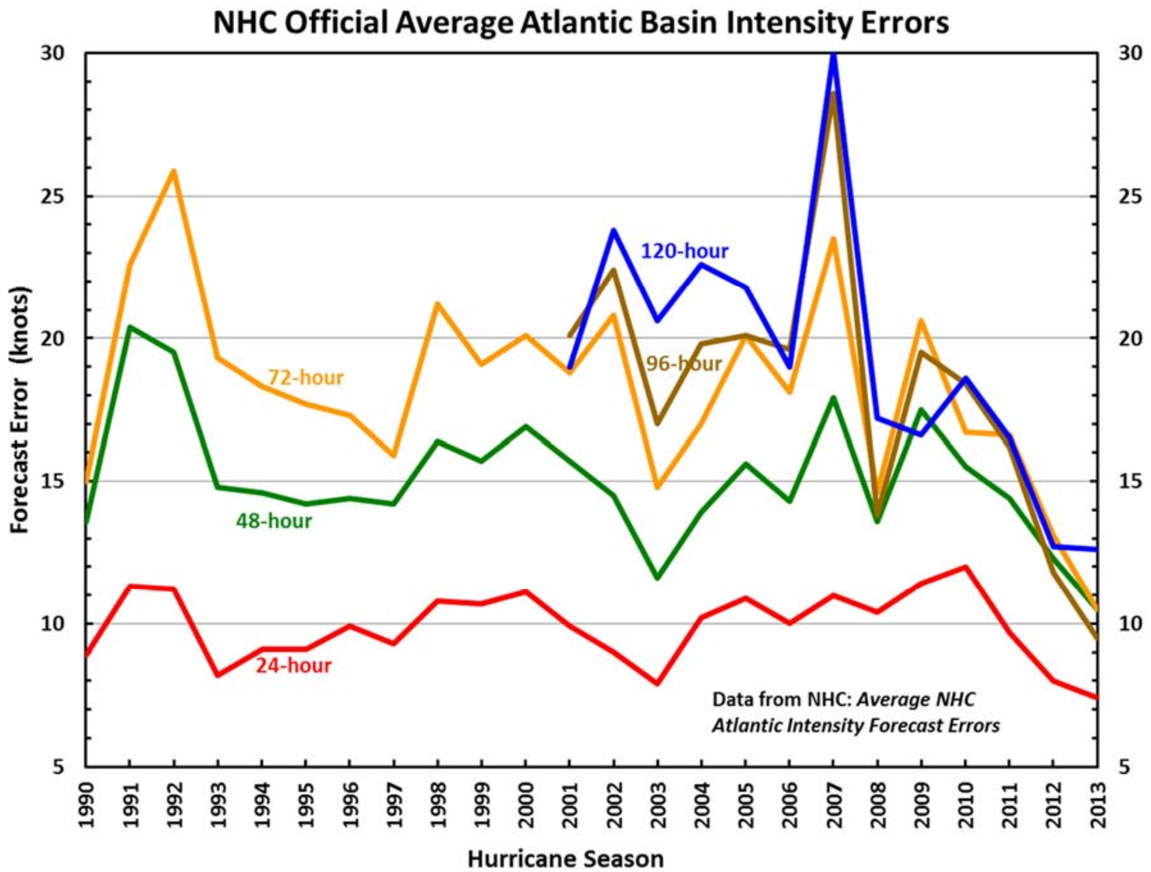


Figure 5. NHC Official Intensity Forecast Errors (1990-2013)

iii. HWRF Ensemble

For the first time HFIP and EMC conducted a real-time experiment on the HFIP machines in Boulder where the HWRF system was run as an ensemble. The basic model used in the ensemble was identical to the operational HWRF for 2013. A 20 member ensemble was run where the initial- and boundary-condition perturbations were from the GEFS. The members were created using the ensemble transform with a rescaling scheme. Additionally, the model physics were perturbed by adding a stochastic component to the convective trigger function in the operational HWRF. In particular, sub-grid scale convection is triggered only when the difference between partial pressure at the convection starting level and that at the level of free convection is less than the (large-scale-vertical-velocity-dependent) trigger function (120 to 180 hPa) plus a random component between -50 and 50 hPa. Random components are generated separately for each member at each cycle thereby avoiding spatial or temporal correlations. Thus the large scale perturbations of the ensemble came from the GEFS to initially define each member and then a stochastic convective trigger was added within each member.

Figures 6 and 7 compare the intensity and track forecasts of the operational model (red line) and the ensemble mean (blue line) from the ensemble experiment in the Atlantic in 2013. Because there were few long-lived storms in the Atlantic this season, there are not enough cases beyond 72 hours to draw reliable conclusions. Figure 15 shows significant improvements in the track forecast out to 72 hours.

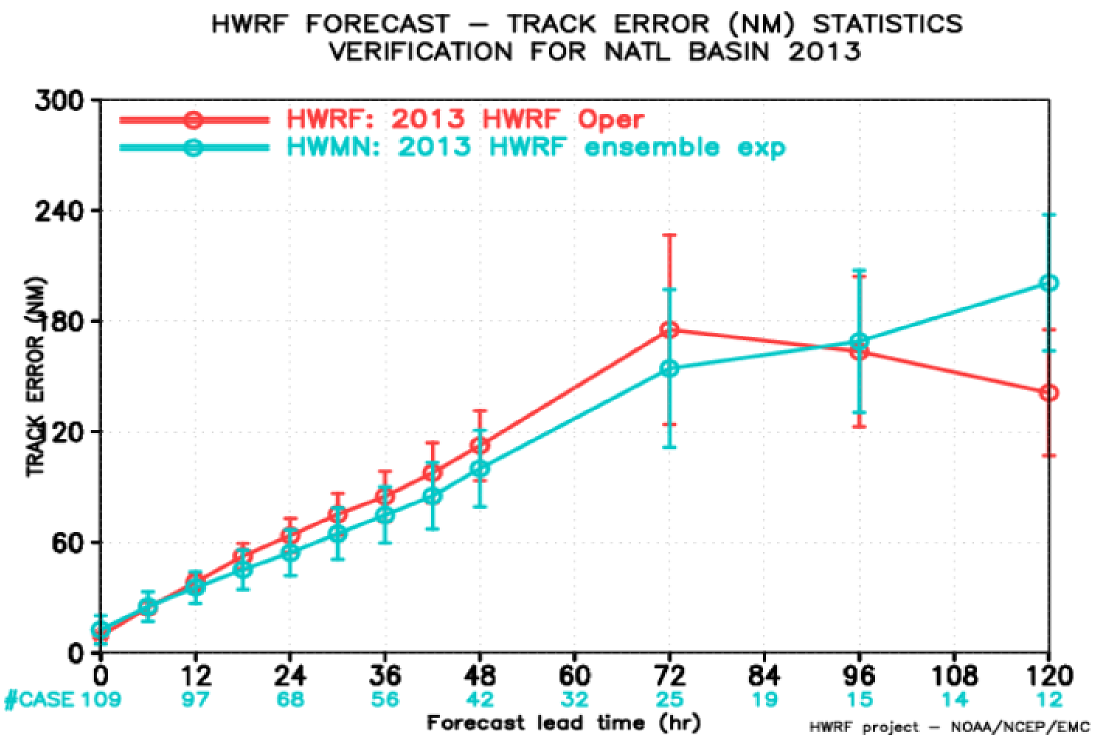


Figure 6. HWRF 2013 Experimental Ensemble Track Performance, Atlantic Basin
Red shows the operational HWRF and blue the ensemble mean. The ensemble model was identical to the operational HWRF, it used an inner grid resolution of 3 km.

Figure 7 shows even more improvement in the intensity forecast, approaching 50% at some lead times. This effect alone (from this type of ensemble) might give the additional improvement overall in intensity guidance from the HWRF system to meet the HFIP goal at 5 years (Figure 4).

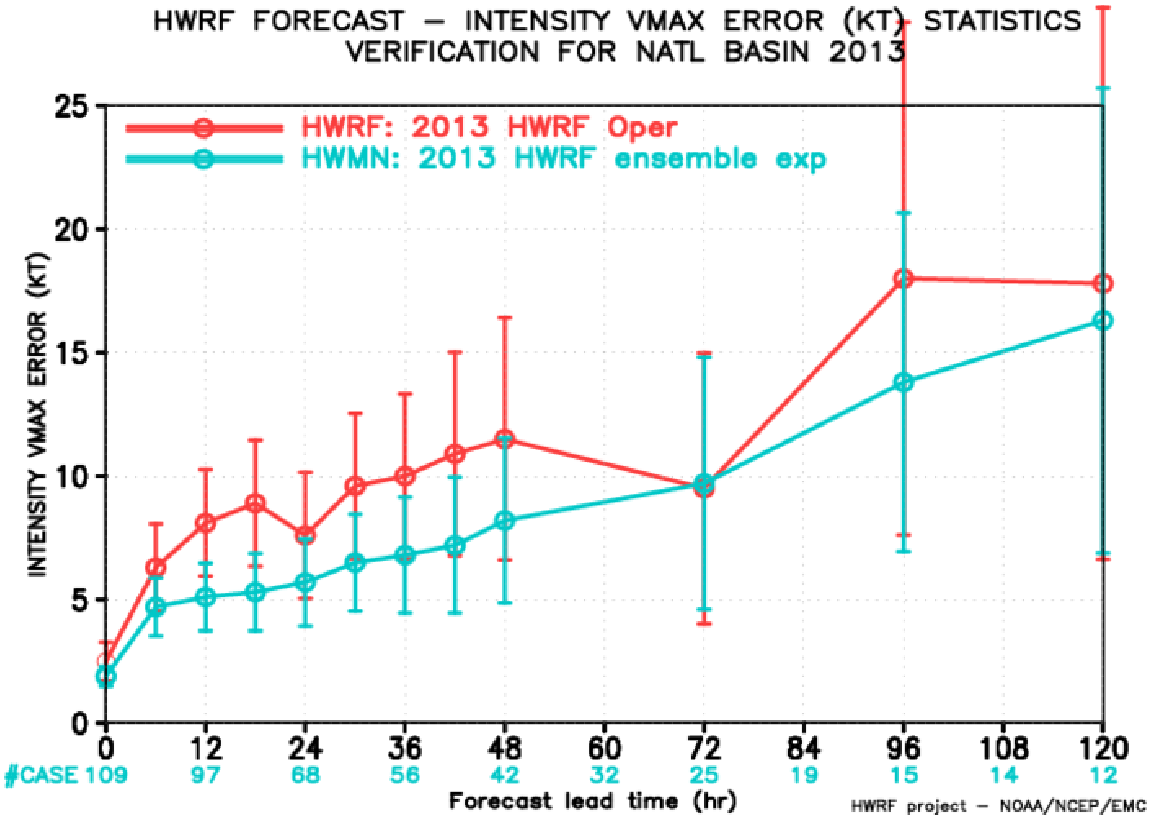


Figure 7. HWRF 2013 Experimental Ensemble Intensity Performance, Atlantic Basin
Again, red shows the operational HWRF and blue the ensemble mean. The ensemble model was identical to the operational HWRF, it used an inner grid resolution of 3 km.

iv. Rapid Intensification

One of the goals for HFIP is to increase the probability of detection (POD) for rapid intensification (RI) to 90% at Day 1 decreasing linearly to 60% at day 5, and decreasing the false alarm ratio (FAR) for RI to 10% for day 1 increasing linearly to 30% at day 5.

We were not able to test the ability of the 2013 HWRF to forecast rapid intensification in the Atlantic this year because of too few cases. There were enough cases for such a test in the WPAC. Figure 8 shows a comparison of the observed 24-hour intensity change in the West Pacific plotted against the HWRF model forecast, left panel, and observed compared to the JTWC official forecast in the right panel. The numbers in the lower right quadrant of each panel show the POD and FAR for each set of forecasts. The red squares show the points that were correctly forecast as rapid intensifiers. While the POD was still considerably below the 10 year goal for HFIP, the result shown in figure 8 is promising: note that the HWRF RI forecasts are better than the official forecasts.

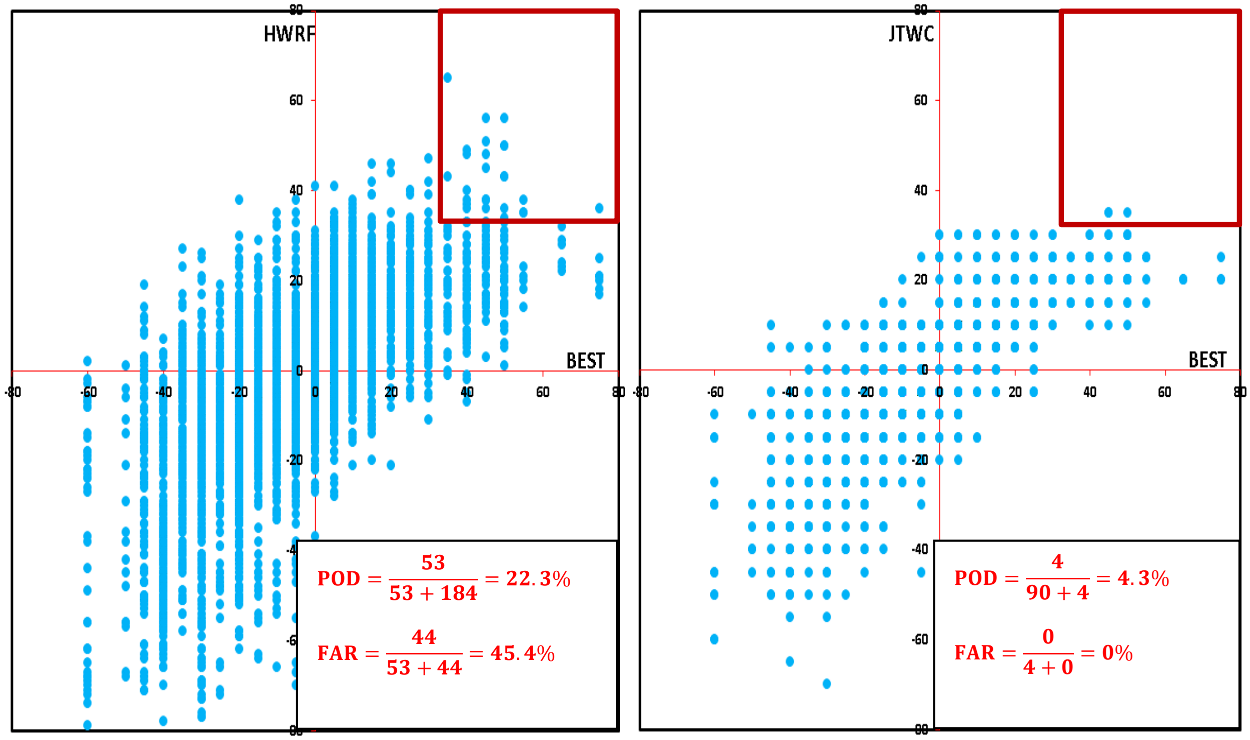


Figure 8. Forecasted 24-hour Intensity Changes in WPAC, 2013
Forecast intensity change along the x-axis and best track intensity change along the y-axes. The red square denotes the region of increased intensity greater than 30 knots in 24 hours—the region regarded as rapid intensification. The Probability Of Detection (POD) and the False Alarm Rates (FAR) are shown. The left panel is for HWRF and the right panel is for the official JTWC forecast. Note that multiple data points with the same forecast-best track value appear as a single blue dot in the plots

b. Seven Day Forecast

One of the goals of HFIP is to "extend the lead-time for hurricane forecasts out to Day 7 (with accuracy equivalent to that of the Day 5 forecasts in 2006 ~260 nm)". Figure 9 shows the operational GFS and ECMWF track errors for the Atlantic averaged for the years 2006-2008 and plotted against the baseline as a gain over the baseline. For days 6 and 7 a linear interpolation from days 4 and 5 was used to define the baseline. Note that over the three years the actual operational models used did change somewhat.

During that period, GFS performed significantly worse than ECMWF and plots below (worse than) the baseline for all forecast lead times.

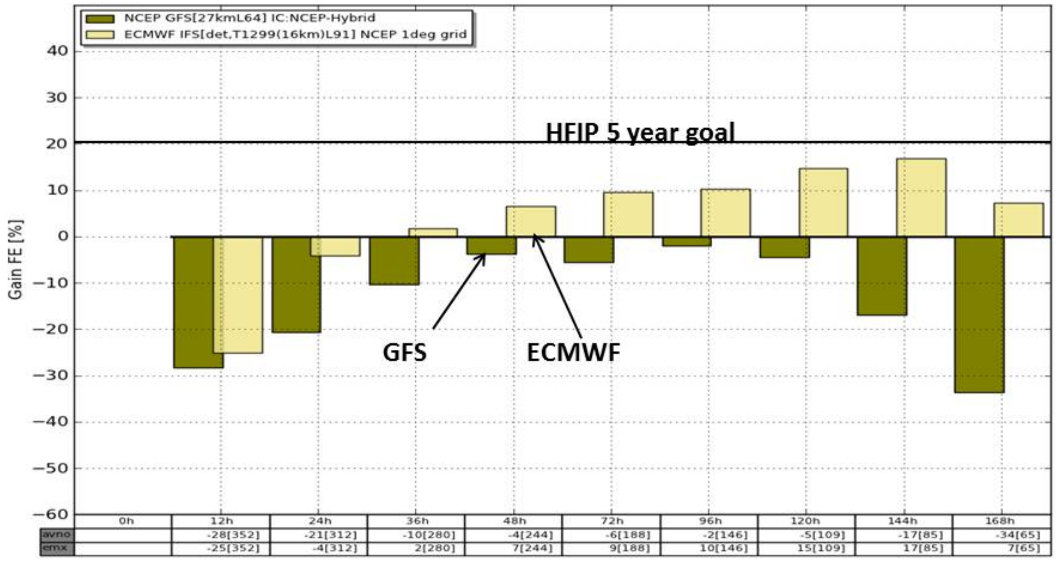


Figure 9. GFS & ECMWF Track Errors (2006-2008)
Track forecast errors of the GFS and ECMWF models over the years 2006-2008 in the Atlantic. Data are plotted against the HFIP baseline and the 5-year goal is noted on the figure. Numbers of cases are shown in parentheses along the bottom.

Figure 10 shows results from 2012. We only show one year in this figure because there were major changes in the GFS that reflect what HFIP feels is required to make progress toward its goals; introduction of the hybrid DA system and running at higher resolution (T574). Note that the number of cases is relatively small. As was noted earlier, in Figure 2 where the data were plotted in terms of skill, GFS was meeting or exceeding the HFIP 5-year goal through 4 days and beats the ECMWF from forecast lead times of 48 to 96 hours.

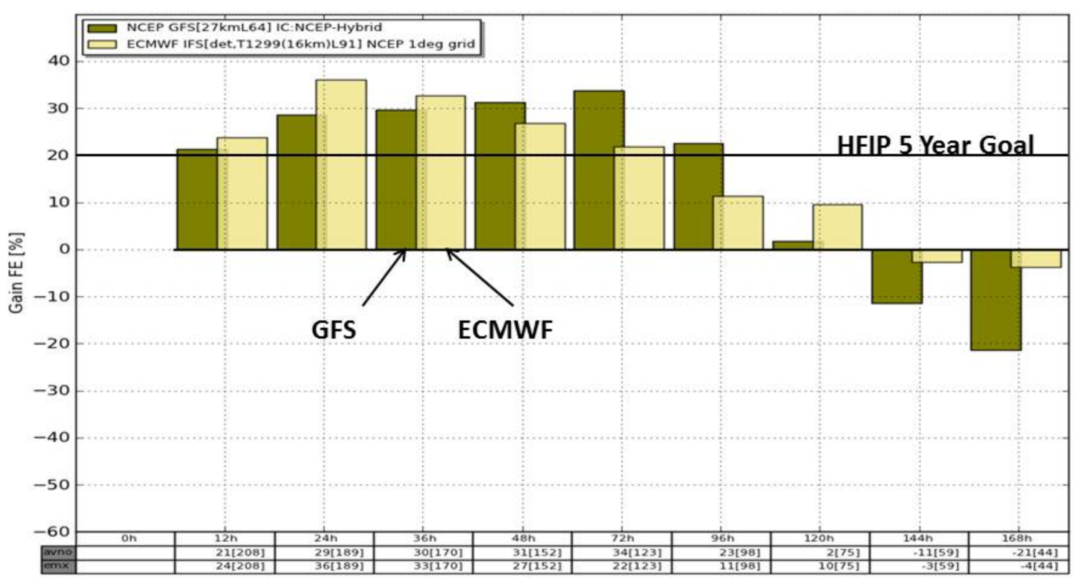


Figure 10. GFS & ECMWF Track Errors (2012 Operational)
Same as Figure 9 except for the operational models in 2012

Around day 5 the improvements in both models drop off dramatically and in fact the GFS shows almost no improvement at 7 days. Because of the small sample size at this forecast range and the fact that there is only one season shown, care must be taken interpreting the changes. This is most noticeable for ECMWF where comparison of the two figures suggests that model has gotten worse at 7 days. Our assessment is that there has been essentially no change in the 7-day forecast accuracy for both models. Figure 10 shows that we have made very substantial improvement in the forecasts out to 4-5 days but after that the improvements that have been put into the GFS have had no impact. The question here is why.

The improvements that have been added include a major change in the data assimilation system and it has been our opinion over the years (as noted in the annual reports) that this change has had the largest impact. We note that when we ran the GFS at T256 but using an EnKF DA system we got results similar to those shown in Figure 10. The higher resolution has led to improvements over the lower resolution versions of the GFS (not shown here) but the DA seems to be having the largest impact. This being said then the question arises as to why the model error rises much more rapidly after day 4-5 as compared to the earlier forecast periods. One suggestion has been that 4 days is about what it takes for errors from the analysis in the Pacific to reach the Atlantic. The presumption here is of course that the errors in the analyses over the Pacific are greater than those over the Atlantic and North America. As a rule of thumb, one might expect this to be the case. Regardless, the reasons for the slow to no improvement in the GFS (and ECMWF) forecasts at day 7 remains a mystery and a challenge for HFIP in its second 5 years.

The above discussion is suggesting that the problem may be related to the data used in the analysis. This will be considered in the next 5 years. For example there may be data that will be coming online like Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) data that will help. However another possibility is ocean interaction. At 4 days large scale atmosphere ocean interactions may be becoming important and lead to a rapid growth in error. For example, the GFS uses a constant sea surface temperature for the duration of the 7 day forecast (based on a current analysis). So another possible line of investigation might be to look, on a global scale, at the impact of surface temperature changes on the long term (out to 7 days) error growth. If that is important in the 7 day forecast then coupled ocean atmospheric models will need to be explored.

c. Genesis

While not a specific goal of HFIP, hurricane genesis is an implied goal since hurricane forecasts in general extend out to at least 5 days and being able to forecast genesis is important for the longer lead times. Many hurricanes that exist at the end of the 5 -7 day forecast won't exist at the initial time. HFIP has been encouraging the development of methods to diagnose genesis in our experimental and the operational models. This involves the development of a tracker that can identify genesis in an ensemble and then an effort to determine how well the various ensembles perform. Figure 12 shows the genesis prediction for four of the operational models in 2012 (ECMWF, GFS, CMC, NOGAPS) and shows genesis in the first 48 hours of the forecast. Since these are ensembles, if 50% of an individual model's forecasts showed genesis in the first

48 hours then the probability of genesis is 50% for that ensemble forecast. Thus if genesis occurred 50% of the time in cases where 50% of the members forecast genesis, then the forecast was considered perfect (the thin diagonal line), and so forth.

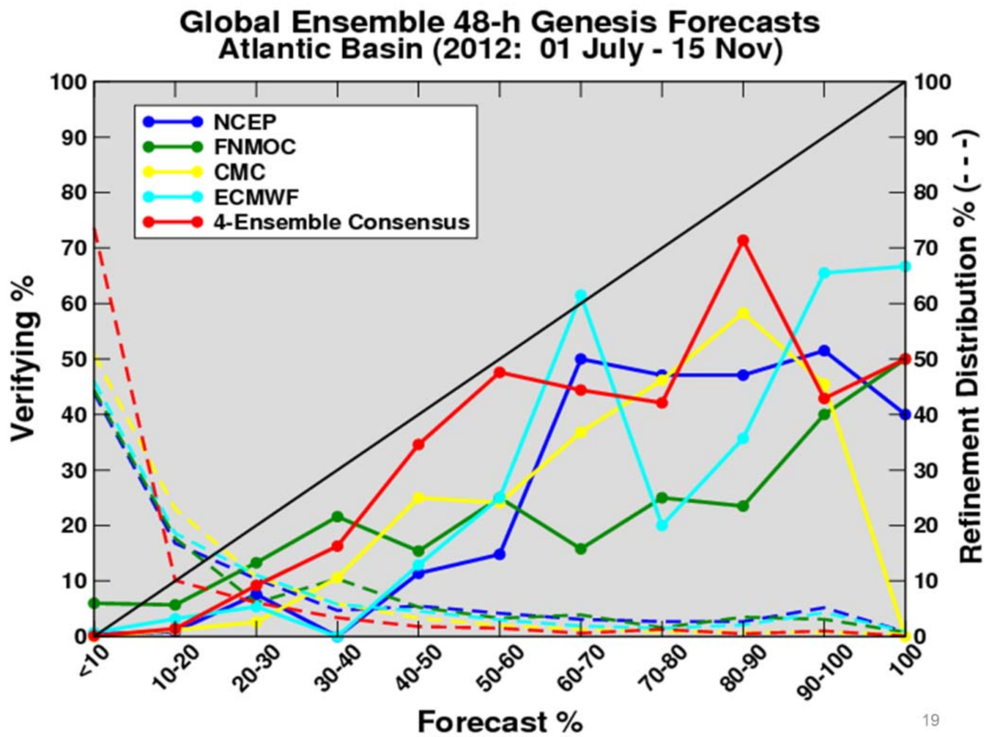


Figure 11. 2012 Genesis Modeling Statistics
Genesis statistics from 4 operational global models: NCEP GFS model, FNMOC NOGAPS model, Canadian Meteorological Centre (CMC) model, and the ECMWF model ensembles. The 4 member ensemble consensus statistics are shown in red. Solid lines show the percent verifying, the dashed lines are called the refinement distribution and each indicates the percentage of times the forecast probability falls in a particular category. So, for example, for the ensemble mean (red line) forecasts of 10-20% probability of formation occur 10% of the time.

All four models shown in Figure 11 over predict genesis and there is a slight tendency to over predict more when the number of members in the ensemble showing genesis is larger. In general the Fleet Numerical Meteorology and Oceanography Center (FNMOC) model tends to over predict more than the other three but all vary widely with increasing forecast probability. The four-ensemble consensus shows the best skill in forecasting genesis at least up to a forecast probability of 60%. The dashed lines (the refinement distribution) indicate the percentage of times the predicted genesis probability fell within the various categories. Thus 75% of the time the four-ensemble consensus predicted less than a 10% probability of genesis. This work was done by GFDL.

Figure 12 is a similar figure for the NHC operational genesis forecast. That forecast is rather good for the lower probabilities of genesis, greatly under predicts around a 50% probability of genesis and slightly over predicts above 70%.

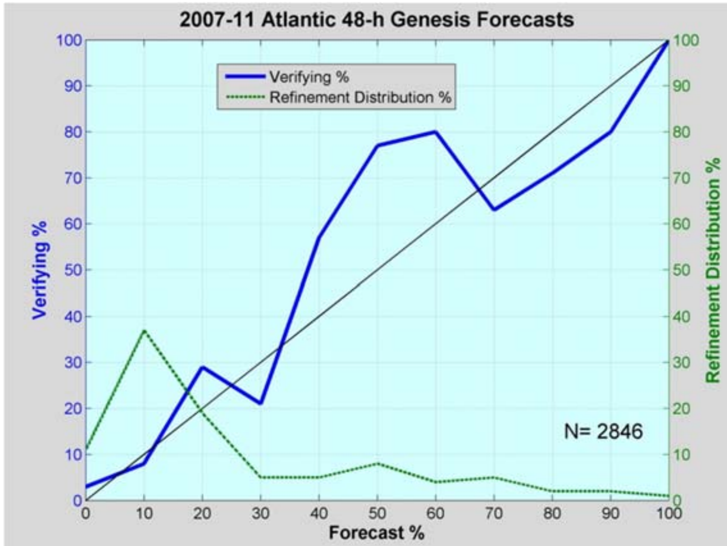
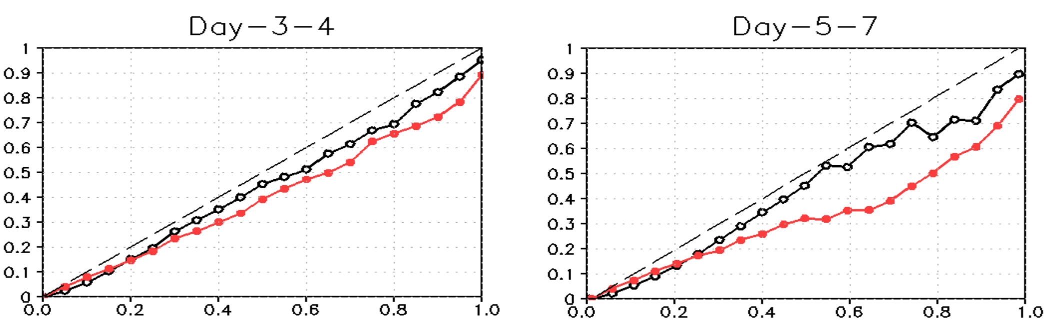


Figure 12. NHC Genesis Prediction Statistics
Statistics for the NHC genesis prediction, otherwise same as Figure 11. Data for years 2007-2011 are shown.

Figure 13 shows the verification of probability calculated from the experimental HFIP GFS ensemble of winds greater than 34 knots at a point during the time period indicated on each panel.



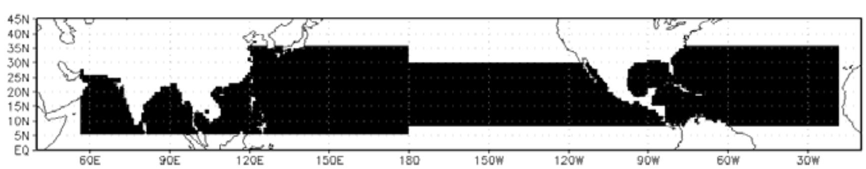


Figure 13. Verification of Tropical Storm Force Wind Probabilities
Verification statistics for probability of tropical storm force winds (greater than 34 knots) at points shown in the black area in the lower panel. This is from the HFIP GFS T574 ensemble and 100% or 1 in the figures would imply that all ensemble members forecast a 34 knot wind at that point during day 3-4 (left panel) and days 5-7 (right panel), black and the operational GEFS (T256), red. Reliability Scores are aggregated over the domain shown below. Observed probabilities are shown on the y-axis and forecast probabilities on the x-axis. A perfect score would be along the dashed line. Forecast probabilities are over forecast below the line and under forecast above.

Figure 14 shows the probability forecast for winds 34 knots or greater over a 168-hour period starting 12Z September 11, 2013. The storm in the eastern Atlantic is Humberto which was a hurricane at that time. Humberto later weakened rapidly as indicated in the figure and then regained tropical storm force winds as it turned northeast. The storm in the Gulf formed on the 12th and became hurricane Ingrid on the 14th. The storm in the EPAC is Manuel which formed on the 13th.

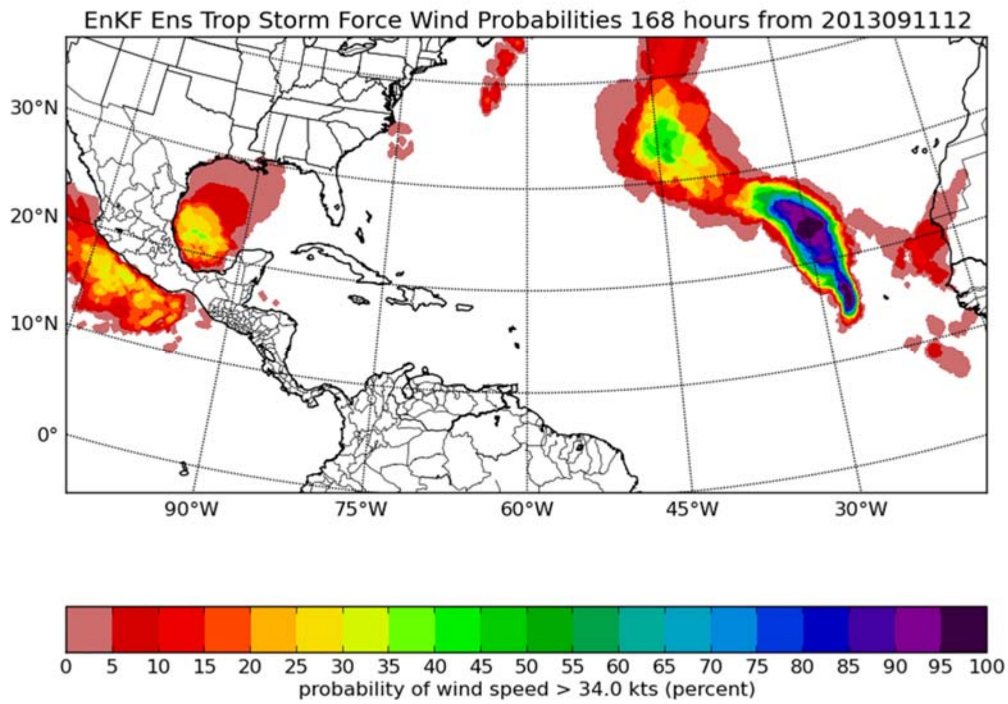


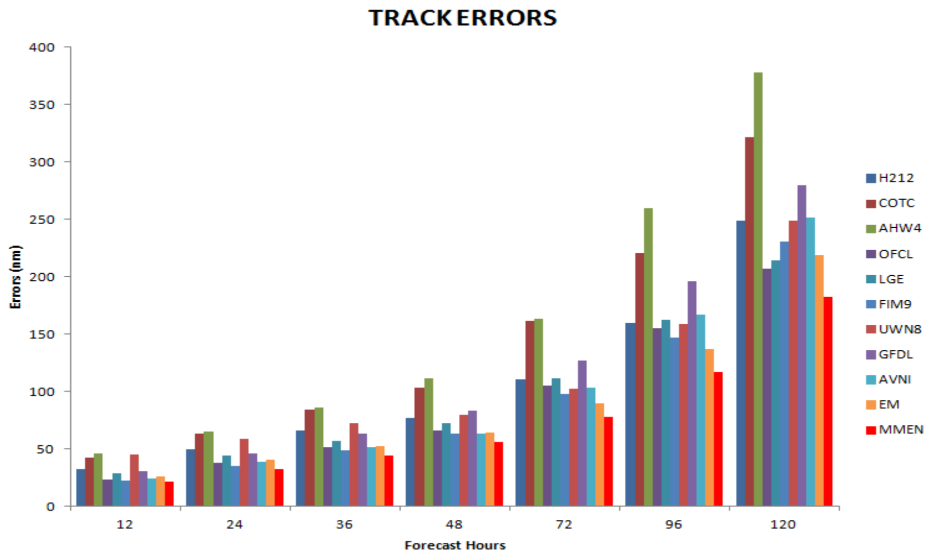
Figure 14. Tropical Storm Force Wind Probabilities within 168 Hours

An example of a probability of winds greater than 34 knots at the various points during a 7 day period starting at 12 Z September 11 2013. A 100% probability implies that all 20 members of the ensemble forecast a 34 knot wind at the point at some time during the 7 days of the forecast. Forecast is from September 11 at 12Z 2013. Humberto is the storm in the east Atlantic that formed on September 8, 2013, Ingrid in the Gulf of Mexico that formed on September 12, 2013 and Manual in the East Pacific that formed on September 13, 2013.

d. Statistical Post Processing of Model Output

Much of the discussion above focused on using numerical model improvements to achieve the HFIP goals. Typically, statistical models (for example DSHP and LGEM) are among the best predictors of hurricane intensity. A statistical model is one where a limited number of variables (measured in single to double digits) are weighted, through correlation with past data, and combined. The variables are generally selected from parameters describing the current state of the hurricane or various environmental data. Values of environmental data can be specified using current observations or model forecasts. Perhaps the simplest statistical model for intensity is SHIFOR5 (also called OCD5) where the variables are current position and intensity, position and intensity 12 hours earlier, and date. CLIPER5 is a similar model for track.

Another class of statistical model takes particular predictions (say track or intensity) from several dynamical models in a multi-model ensemble and combines those predictions as a weighted average. The weights are determined by comparing the performance of the various models over a period of years. The FSU Multi-Model Ensemble is of that class. As in past years, the FSU Multi-Model Ensemble was among or the best performers of the statistical models for 2013. Figure 15 shows the track and intensity errors in 2012 for the various models that went into the FSU Multi-Model Ensemble. The orange and red bars on the right side of the groups for each forecast lead time are the equally weighted ensemble mean (EM) and for the variably weighted ensemble mean (MMEN). At all lead times the weighted mean was better than the equally weighted mean and all the other models. At some lead times the official forecast was better.



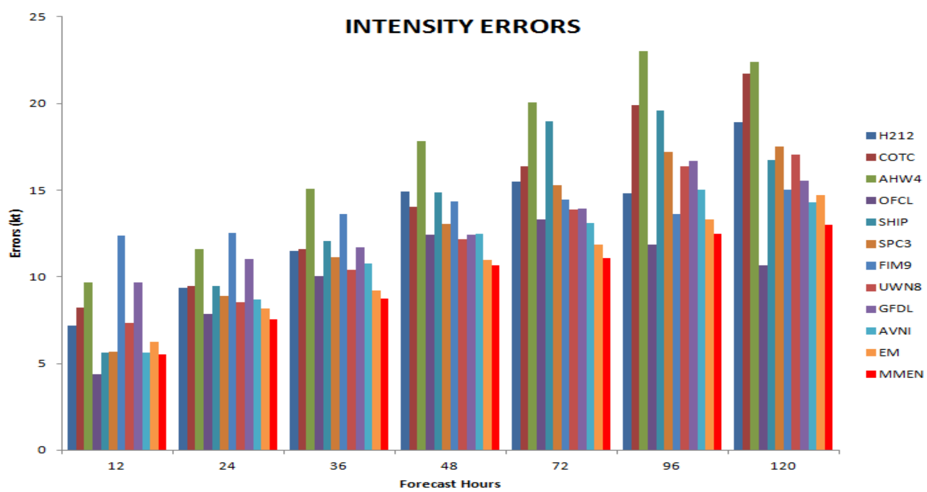


Figure 15. FSU Multi-model Ensemble: Members and Mean Performance (2012)
Track and intensity errors for all components of the FSU Multi-model Ensemble for various forecast lead times. The acronyms are shown on the right (see Appendix for details) and the number of cases is shown across the bottom. The ensemble mean (EM) of the models shown is in orange and the weighted ensemble mean (MMEN) is shown by the red bar. The official forecast is dark purple.

More complex statistical models used operationally for intensity are SHIPS, LGEM and SPICE (SPC3). SPC3, in recent years, showed improvement compared to the operational statistical and dynamical models, by using multiple operational numerical models as input for the environmental predictors of DSHIP and LGEM. SPC3 derives input for both DSHIP and LGEM from the operational GFS, HWRF and GFDL models. This gives six variations which are then averaged as an ensemble.

Both SPC3 and the FSU Multi-Model ensemble system are examples of statistical post processing being pursued by HFIP.

5. HFIP Challenges

HFIP has been working on the hurricane forecast improvement problem for about four years. From this work, several challenges have been identified that HFIP must resolve before meeting its goals.

- Most model initialization times will not have the advantage of aircraft data. Exceptions will occur in the Western Atlantic but this challenge will hold true for almost all Pacific storms. Better use of available satellite data near the storm center is needed in these cases, particularly in the regional models. Satellite data need to be fully utilized in characterizing the storm environment, whether or not aircraft data are available.

- Improvement in the initial states of the regional models is absolutely essential to achieving progress in the rapid intensification goals. Emphasis is on rapid intensification in the first 24 hours and current models take that long to settle down from the shock of initialization. HFIP is developing better data assimilation systems for the regional models to alleviate the initialization startup problem.

- Development and tuning of physics packages for hurricane models at high resolution is critical. This effort is ongoing within HFIP and many of the improvements in skill of HWRF such as noted in Figure 4 are a result of gradual improvements in the physics package. Each year, improvements are added to the physics packages and, in 2013, increased computer time allowed the physics to be called more frequently. That is likely responsible for much of the improvement.

- Advanced data assimilation systems in both regional and global models appear to lead to substantial forecast improvement. These advanced assimilation systems should be entered into operations as soon as possible. The global model was transferred to the hybrid DA system in 2012 and likely accounts for the improvements noted in Figure 2. As noted above, introducing the hybrid system into the regional models remains a challenge.

- Statistical post-processing of model and ensemble output and improved model data assimilation (such as addition of supplemental observations) have shown considerable promise and may be the key to meeting the intensity forecast improvement goals. HFIP continues developing existing and new statistical post processing systems.

- Better products are needed to effectively convey ensemble information to forecasters.

- Error growth, especially as manifested in track error becomes much larger after 4-5 days compared to the earlier forecast leads times. Reasons for this are unclear but it could be due to errors propagating in from data poorer areas (like Pacific); or some model deficiency (ocean coupling); or model biases becoming prevalent. It will be a goal of the project to better define the source of this error even if there is not time in 5 years to make changes that will eliminate it.

- The current hurricane models are regional in that they involve a large (generally covering all the Atlantic and North America) outer grid, or domain, with two or three high resolution grids nested around the hurricane center to capture the processes near the core. They consider one storm at a time and, so, when more than one storm is modeled, the same computations are performed for the outer domain (often at global model scale resolutions) more than once. The outer domain takes boundary conditions from a global model.

- Recently HFIP has been working with basin scale models with a somewhat larger outer domain but with more than one set of nested grids within the domain. This allows for interaction between storms but the outer domain still depends on a global model which receives no feedback from the hurricanes. One solution to this is to make the global model the outer domain of the ”regional” model. A challenge of HFIP will be to move the HWRF in that direction by moving the current model from the NMME grid onto the Non-hydrostatic Multi-scale Model on B grid (NMMB) and using the NOAA Environmental Modeling System (NEMS) framework. This will also move the hurricane model toward a unified modeling system within EMC.

The plan outlined below addresses the challenges to be met in achieving the operational system for 2019 proposed in Section 7.

6. Overall HFIP Strategic Plan

This section describes the overall plan for the project with further information on how HFIP goals will be met.

a. Track

Improved global modeling is needed to improve hurricane track forecasting. This is especially true for the longer lead times, since a forecast at a particular location (after about 48 hours) is affected by systems distributed globally. The track in a regional model is very much affected by the track in the global model, which the regional model uses for boundary conditions. Therefore, primary emphasis on improving track forecasts will continue to be placed on improvements in the global models. A key here is that the models must also accurately forecast genesis since many storms can go through their lifecycles in 7 days or less. A storm not forming for another 2 days and then making landfall 5 days latter will not be predicted 7 days in advance unless genesis is correctly predicted.
.

b. Intensity

Intensity is controlled by processes both external and internal to the hurricane. The external processes are controlled by the large scale flow surrounding the hurricane. Improving forecasts of the external factors controlling intensity will require improved global models. The internal processes include convection occurring within the hurricane, particularly in or near the eyewall. Improvements in forecasting the internal factors controlling intensity must be tackled with very high-resolution definition of the hurricane in regional models.
.

c. Global Models

Hurricanes alter the flow in which they are embedded, which in turn controls their track and intensity. To capture this environmental feedback, the global model must be run at a resolution high enough to adequately resolve the hurricane. It is generally thought that the hurricane is reasonably resolved at a horizontal resolution of 5 km. High-resolution modeling requires a large number of computer processors for real-time run operations. This degree of resolution for a global model running in real-time will not be possible for another 3-5 years, even with currently planned research supercomputers. Hence, the strategy is to use the finest resolution possible to begin to demonstrate the ability of global models to forecast hurricanes at long lead times. The goal is to develop global models to run at 10-15 km by the end of the project.

Specific steps:

- Leverage Sandy Supplemental supported Global Model Development activities
- High Impact Weather Prediction Project (HIWPP) selection of a preferred non-hydrostatic global model for future OAR development for data assimilation and physics would be a likely focus for HFIP global modeling for the 2017-2019 period
- Evaluate reasons for increased growth of track forecast errors beyond Day 4
- Address 7-day track forecast goals and improved TC genesis prediction goals

d. Regional Models

Much higher resolution is possible today with regional models. A single-model deterministic run at 3 km in real-time is possible now and a run at 1 km may be possible operationally within a few years. This resolution should be adequate to capture the convective processes occurring in or near the center of the storm, including processes coupled to the ocean. This, however, needs to be demonstrated and the appropriate physical processes need to be included in the physics package appropriate for a 1-km model. Our strategy is to use the regional models, coupled to high-resolution global models, to address the problem of improving forecasts of intensity.

Specific Steps:

- Advancements to model infrastructure: High-resolution nests with focus on multi-scale interactions
- Regional high-resolution multi-model ensembles (Stream 1.5 and Demonstration System Implementation Tiger Team)
- Ocean Model impact on hurricane forecasts (Ocean Model Impact Tiger Team)
- Evaluation of aircraft recon data impact (Reconnaissance Data Impact Tiger Team)
- Assessment of benefits of 2-way nesting versus 1-way nesting
- Advanced air-sea-wave-surge-land-hydrology coupled models

e. Physics Packages for regional and global models

It has been shown many times (see for example the annual reports of HFIP) that inaccurate physics is a primary driver behind model error. Therefore there will continue to be an emphasis in HFIP on developing and refining the physics packages in both the global and regional models. A goal is to develop a "scale aware physics package" that can be used interchangeably in global models as well as regional models or the counterpart -- high resolution nests within global models. Overall the program will focus on a few critical areas:

- Surface physics: Stochastic nature of drag-enthalpy coefficient ratios (C_d/C_k); Sea-state dependence (waves and spray)
- Microphysics: Role of individual hydrometeors
- PBL: Role of rolls, vertical/horizontal diffusion
- Convection: Magnitude and impact of downdrafts and connection to Microphysics
- LSM: Connection to surface physics
- Radiation: Cloud Radiative Forcing (CRF); impact on air-sea fluxes

Specific steps:

- Advanced scale-aware physics for multi-scale applications including vortex-shear interactions
- Aerosol aware physics for radiation and microphysics
- Development/adaptation of 3D physics schemes

- Advanced physics for air-sea interface including impacts of waves, sea spray and ocean currents
- Physics development closely tied with observations and model diagnostics
- Focus on RI/Rapid Weakening processes to accomplish HFIP goals for increased POD and reduced FAR
- Stochastic physics to assist DA and ensemble development

f. *Unified Hurricane Model*

As noted in Section 5, HFIP will move toward unifying its hurricane model, consisting of multiple nests each over one of multiple hurricanes and contained within the global model, with other NCEP models. The hurricane nests will be fully interactive with the global model. Converting to this system will only occur after it is shown that forecasts using this model equal those provided by the HWRF model which will continue to be improved.

Specific Steps:

- Develop a self-consistent 3D, moving to 4D, Ensemble Variational Data Assimilation (Ens Var DA) system using cycled HWRF backgrounds with moving nests, that can be used by ARW/NMME/NMMB
- Adopt NMMB/NEMS framework for HWRF ---- FY15/16
- Accelerate transition of HWRF components to NMMB/NEMS, further increase of model resolution to ~1-2 km near the storm region --- FY16/17
- Configure and test multiple moveable nests within the NMMB/NEMS framework (basin-scale HWRF) using advanced computationally efficient procedures --- FY16/17
- Efficient coupling between various components within NEMS including post-processing and product generation --- FY17/18
- Adopt the high-resolution nesting strategies to develop global-to-local scale modeling for hurricane forecast applications --- FY18/19
- Design efficient high-resolution ensemble strategies to provide probabilistic guidance on track, intensity (including RI), size and structure forecasts --- FY18/19

g. *Ensembles*

A single deterministic run of a model is essentially one member of an ensemble. If the forecasts of tracks from all members of an ensemble are plotted from the initial position of the hurricane, the result is a diverging fan of tracks. A deterministic run could be any one of these members. At lead times beyond two days when the tracks have generally diverged significantly from the ensemble mean (or median or mode), the ensemble provides the best information to forecast the most probable track. The same holds true for forecasting intensity. In five years, global ensembles with resolutions of 10-20 km should be available in real-time with enough members to give a good statistical distribution of forecast probability. High-resolution regional ensembles from multiple HFIP models running at 3 km resolution will be used to provide probabilistic forecasts of intensity, as well. Inclusion of multiple models will provide a more diverse ensemble and a better representation of model errors.

Specific Tasks:

Ensemble forecast priorities

- Evaluate HWRF ensembles initialized from Ens Var system versus those initialized from GEFS
- Evaluate estimations of forecast uncertainty (at different spatial and temporal scales) to inform development of stochastic physics
- Extract more information/develop new products from days 5-7 in global ensembles (particularly relating to genesis/decay)
- Identify/correct deficiencies in ensemble system for days 5-7 (including storm population biases, track biases - feedback on model physics development)
- Evaluate using multiple models (HWRF, TC-COAMPS, WRF-ARW) within the ensemble
- Include ocean uncertainty in ensembles
- Evaluate using coupled models (longer term) and atmosphere-only models with perturbed sea surface temperatures (shorter term)

Predictability issues

- Use ensemble re-forecasts (global and regional) to answer predictability questions such as:
    - What kinds of systems are more or less predictable?
    - What situations have large model error (small spread/large error in the ensemble)?
- What observations are needed to improve forecasts at different time scales (particularly intensity)?
    - Leverage Sandy Supplemental Observing System Simulation Experiments (OSSE) effort
    - Investigate use of ensemble forecast sensitivity to observation algorithms (e.g., Ota et al 2013)

h. Post-Processing

There are statistical techniques that extract certain predictors from the models (such as shear in the vicinity of the hurricane) as well as from various observations to make specific forecasts such as maximum wind speed. Statistical models using those techniques currently provide the best intensity guidance to forecasters when compared with dynamical models. A major part of HFIP will include developing these statistical methods.

HFIP is also promoting new techniques of presenting model data to forecasters, techniques for diagnosing model errors and their causes and model verification.

Specific steps include:

Development of new products and statistically post-processed guidance

- Transition the NHC-HFIP corrected consensus to operations
- Continue to expand the use of SHIPS, LGEM, and SPICE with input from global ensembles
- Develop and implement techniques that provide guidance on guidance (in coordination with Joint Hurricane Testbed efforts on this topic)
- Develop a web-based interface that can display ensemble forecast tracks, stratified by parameters such as intensity, moisture, or shear
- Coordinate with NRL to develop and implement a method for itemizing NHC requirements for Automated Tropical Cyclone Forecasting (ATCF) support and upgrades, including budgets

Development of new diagnostic and verification techniques

- Include new DTC/TCMT verification techniques into tropical cyclone model evaluation tools (MET-TC), and promote the expanded use of MET-TC by modelers and forecasters
- Survey the HFIP groups to develop a list of available verification and diagnostic tools, and make the list available on the HFIP website so other groups can benefit.
- Develop new verification techniques that evaluate storm size and structure forecasts
- Continue advanced diagnostics work for model improvement. Examples include, but are not limited to, diagnostics of model radius to maximum wind, vortex tilt, eyewall slope, convective bursts, PBL, surface layer, and microphysics
- Use the HWRF (or other models) to develop tools that distinguish tropical cyclones that undergo RI from those that do not

Real-time display, diagnostics and verification

- Continue to support upgrades to operational diagnostic code, such as the GFDL Vortex Tracker
- Develop and refine tools for making real-time verification statistics available to forecasters at NHC
- Expand the HFIP products webpage, and integrate more interactive displays, such as the new ensemble track display system

Community tool development

- Expand the Cooperative Institute for Research in the Atmosphere (CIRA) model (SHIPS) diagnostic code to include additional output parameters and update the documentation
- Upgrade the CIRA diagnostic verification code and update the documentation

i. Observations and Data Assimilation

Advanced data assimilations systems using 4D-Var, EnKF or hybrids with EnKF and either 3D-Var or 4D-Var are currently in the planning process. These systems will need to take advantage of additional data as they become available from current or planned satellite systems, various aircraft observations, and both land-based and airborne radar. To solve the initialization problem in regional models, satellite data assimilation methods already used and planned for global applications need to be applied at higher resolution in the vicinity of the hurricane. Many satellite data assimilation components, including data thinning, bias correction, cloud detection, and quality control, currently used in global forecast systems need to be revised for higher resolution applications. Clouds affect satellite data presenting a major challenge to assimilating a large amount of satellite data within and near tropical cyclones and, therefore, we need to develop new and enhance existing cloud detection and cloudy-radiance assimilation methods. All aircraft-derived data should be utilized and observing strategies must be synchronized to the availability of planes to maximize the value to model initialization. There may still be data gaps for both model initialization and direct forecaster use that will need to be addressed perhaps by use of unmanned aircraft.

Specific tasks:

DA System Development

- Develop a self-consistent 3D, moving to 4D, Ens Var DA system using cycled HWRF backgrounds. This system will:
  - Use an HWRF-based EnKF to define background-error covariances
  - Have a moving nest DA cycling capability
  - Be able to use WRF/ARW, WRF/NMME and NMMB dynamic cores
  - Assimilate all operational observations (including radiances and airborne TDR data)
  - Have a flexible cycling interval (down to 1 hour)

Systematic Evaluation of DA/Observations impacts

- Use self-consistent hourly cycling HWRF EnKF Var system to evaluate impact of high-frequency (temporal and spatial) observations at the vortex-scale. Observations to be evaluated include:
  - Airborne acquired data (including those data already tested in RDITT); NHC is particularly interested in the impact of data from dropsondes, and increasing the number of drops from Air Force reconnaissance flights
  - Ground-based radar data
  - Wind derived from from geostationary satellite data
  - Global Position System Radio Occultation (GPSRO)
  - Data from hyperspectral sounders using statistical retrieval techniques
  - Radiance data from geostationary platforms

DA research priorities

- Dealing with displacement errors via
  - field alignment
  - rapid cycling
  - storm-relative DA
- Increase the use of radiances in clear and cloudy regions; A key aspect of this will be developing bias correction methods that work in limited domains and/or better leveraging the bias corrections derived from the global assimilation
- Dealing with multi-scale sampling error (vortex-scale and environment)
- Representation of model uncertainty (through the development of stochastic physics in HWRF, emphasis on microphysics/surface layer) – also applicable to ensemble forecasts

j. Ocean/Wave Models

Some of the models used in HFIP interact with active ocean models both one-dimensionally and three-dimensionally. Both the Princeton Ocean Model (POM) and Hybrid Coordinate Ocean Model (HyCOM) models are used. During the next five years, most atmosphere models with an interactive ocean will likely shift to use of the more complex HyCOM. Use of coupled ocean with a global prediction model also needs to be at least explored to determine if some of the rapid track error growth after Day 5 of a forecast is related to a lack of a coupled model. HFIP also supports development of advanced air-sea-wave coupled models where the explicit interaction with waves is modeled using the operational Wave Watch III model, developed at NCEP.

7. The Configuration of a Numerical Model Hurricane Forecast Guidance System to Meet the HFIP Goals

While it appears that use of aircraft data will likely help HFIP meet its intensity goals for storms for which such data is available, these data will not be available for storms for a large majority of model initializations. For those storms we will need to rely on better use of satellite data taken in the near vicinity of the hurricane. A longer-term major focus for HFIP is to improve satellite data assimilation in regional model initialization systems.

We have not addressed the goal of HFIP to improve the forecasting of rapid changes in tropical cyclone intensity because, at this juncture, none of the HFIP dynamical models are capable of providing reliable forecasts of RI. The global models are not able to resolve the inner core processes that are likely to be very important in the RI process and all the regional models have serious spin up (and spin down) problems initially. This was noted in earlier Annual Reports and is not likely to be resolved until cycled DA with a hybrid DA system is available for the regional models (currently projected to be 2014). Except for the RI issue, we can now say with considerable confidence what a final end-state operational configuration of the hurricane numerical prediction system should look like in 2019, the end of the project.

The longer range predictions, out to one week, of both track and intensity will be accomplished by global models run as an ensembles and initialized with a hybrid data assimilation system and post-processed with various statistical models. Resolution of these global models needs to be no coarser than about 20 km and the results will be improved if more than one global model is used in the ensemble.

The intensity goals for forecast periods out to 48-72 hours will be accomplished with regional models run with resolution at least as fine as 3 km as a multi-model ensemble. All regional models will use all available aircraft and satellite data. These will also be post-processed with statistical models. With the high resolution, the RI goals may be met with the regional models.

Ultimately, the end system might include a global model ensemble with hybrid data assimilation, a regional model ensemble with hybrid data assimilation and statistical post processing (Table 5). The ability to run this system with its high-resolution ensembles, however, will require at least a ten-fold increase in computer resources in operations.

Table 5. Numerical Model Hurricane Forecast Guidance System

Component	**Specifications**
Global model ensemble with multiple interactive moving nests using hybrid data assimilation	• 20 member ensemble at 18 km, global model, 6 and 2 km for inner moving nests • Multi Model (at least two – FIM, GFS, Navy?) • Use all available aircraft and satellite data in core and near environment of hurricane • Run out to 7 days or more
Statistical Post Processing	LGEM, SHIPS, others

We also note that this nested concept could also be extended to mid-latitude systems like squall lines (for example Derechos). In other words, the basin scale concept being tested in HFIP could also likely improve forecasts of other types of weather systems, not just hurricanes.

8. Milestones and Deliverables

This section provides a broad overview of the main deliverables and milestones for HFIP. They follow on from the milestones outlined in the first Strategic Plan. Note that operational implementation is not determined by HFIP. That is done by EMC. In the following, “operational implementation” denotes our estimate of when the models will be ready, assuming adequate operational computer time.

FY15

- Continued development of the HWRF/COAMPS-TC multi-model ensemble system -- real-time test in summer of 2015
- Develop a self-consistent 3D, moving to 4D, EnsVar DA system using cycled HWRF backgrounds with moving nests, which can be used by ARW/NMME/NMMB
- Operational implementation of high-resolution global GFS(15 km) using semi-implicit time differencing
- Development of new model products for forecasters (at least one per year)
- Test new high resolution physics package in HWRF Demonstration system
- Test assimilation of various new sources of satellite data into regional hurricane models
- HWRF operational model upgrade (June)
    - Increase horizontal resolution to 2km and vertical resolution to 64 levels
    - Replace POM with HyCOM
    - Coupling to wave model
    - Coupling to NOAH (for NCEP, Oregon State University, Air Force, Hydrologic Research Lab/NWS) land surface model (LSM)
    - Advanced Microphysics
    - Assimilation of microwave/infrared (IR) cloudy radiance and the Geostationary Operational Environmental Satellite (GOES) Atmospheric Motion Vectors (AMVs)
- Experimental downstream applications for Wave, Surge and Hydrological models
- Transition HWRF to NMMB/NEMS framework and conduct real-time testing (HFIP Stream 1.5/2.0)
- Improve physics and further improve HWRF data assimilation/initialization system—hybrid DA
- Retrospective testing of regional model components of multi-model ensemble for Stream 1.5
- Conduct demonstration experiment (Aug 1-Oct 30)
- Continued development of global models, regional models and their ensembles
- Continued development of physics packages for both global and regional models
- Experiment with coupled ocean-atmosphere global models and new global data sets (such as COSMIC) to reduce track error growth rate after Day 5 thereby improving 7-day forecast
- Implement the NHC-HFIP Corrected Consensus in operations
- Develop techniques that provide guidance on guidance
- Upgrade the CIRA model diagnostic code
- Enter CIRA model diagnostic code into DTC as a community code
- Develop a set of standard metrics and verification techniques that assess surface wind structure and storm size
- Continue HRD vortex- and convective-scale diagnostics -- radii to maximum winds (RMW), vortex tilt, eyewall slope, convective burst radial and azimuthal locations, contoured frequency by altitude diagrams (CFADs) of vertical velocity, etc.

- Develop air-sea interaction diagnostics -- continue HRD PBL & surface layer diagnostics, surface fluxes, ocean sea surface temperature (SST), mixed layer depth (MLD), ocean heat content (OHC), currents, upwelling, etc.
- Include new DTC/TCMT verification techniques into MET-TC (e.g., forecast consistency verification, RI, genesis verification)
- Release latest version of GFDL vortex tracker to the community
- Evaluate HFIP website for content and revisions
- Create a link on the HFIP web page of the list of available verification and diagnostic tools from all the various groups
- Develop a procedure for evaluating web-based HFIP products and transferring high-priority items to NHC
- Coordinate with NRL to develop and implement a list of NHC ATCF support and upgrade requirements, including budgets

FY16

- Operational Implementation HWRF/ COAMPS-TC multi-model ensemble system.
- Operational Implementation of the hybrid GFS semi-implicit global high resolution (T574) ensemble
- Development of new model products for forecasters (at least one per year)
- Test new high resolution physics package in HWRF demonstration system
- Test assimilation of various new sources of satellite data into regional hurricane models
- HWRF operational model upgrade (June)
    - Implementation of basin-scale NMMB HWRF with multiple nests
    - Atmosphere-ocean-wave-surge-land-hydrology coupled modeling system in NMMB/NEMS framework
    - High-resolution physics upgrades
    - DA upgrades
    - Increase resolution to 1-2 km
- Experimental 7-day forecasts of tropical cyclogenesis
- Experimental HWRF ensembles in NEMS framework (HFIP Stream 1.5/2.0)
- Configure and test multiple moveable nests within the NMMB/NEMS framework (basin-scale HWRF) using advanced computationally efficient procedures
- Retrospective testing of regional models of multi-model ensemble (HWRF/ COAMPS-TC) for Stream 1.5
- Conduct demonstration experiment (Aug 1-Oct 30
- Continued development of global models, regional models and their ensembles
- Continued development of physics packages for both global and regional models
- Experiment with coupled ocean-atmosphere models and new global data sets to reduce rate of growth of track error growth after Day 5 in global models thereby improving 7-day forecast
- Test versions of LGEM, SHIPS and SPICE that use input from global ensembles
- Implement graphical displays that distinguish tracks and intensities based on various thresholds of environmental parameters

- Continue entering CIRA model diagnostic verification code into DTC as community code
- Continue HRD vortex- and convective-scale diagnostics (e.g., RMW, vortex tilt, eyewall slope, convective burst radial and azimuthal locations, CFADs of vertical velocity)
- Develop air-sea interaction diagnostics (e.g., continue HRD PBL and surface layer diagnostics, surface fluxes, ocean SST, MLD, OHC, currents, upwelling)
- Release latest version of GFDL vortex tracker to the community
- Release latest version of CIRA diagnostic code to the community
- Evaluate HFIP website for content and revisions
- Update and release the MET-TC software
- Begin to transfer selected experimental web-based HFIP products to NHC for implementation
- Update NHC's ATCF requirements list for NRL and implement changes

FY17

- Development of new model products for forecasters (at least one per year)
- Initial retrospective testing of high resolution (10-km) FIM
- HWRF operational model upgrade (June),
    - Improved physics and advanced data assimilation system
    - Testing of global-scale NMMB HWRF with multiple nests
- Testing of new aircraft observation strategies for improved model initialization
- Conduct demonstration experiment (Aug 1-Oct 30)
- Continued development of physics packages for both global and regional models
- Pre-implementation testing of inner core data assimilation based vortex initialization
- Implement versions of LGEM, SHIPS and SPICE that use input from global ensembles
- Continue HRD vortex- and convective-scale diagnostics (e.g., RMW, vortex tilt, eyewall slope, convective burst radial and azimuthal locations, CFADs of vertical velocity)
- Continue developing air-sea interaction diagnostics (e.g., continue HRD PBL & surface layer diagnostics, surface fluxes, ocean SST, MLD, OHC, currents, upwelling)
- Release latest version of GFDL vortex tracker to the community
- Release latest version of CIRA diagnostic and verification codes to the community
- Evaluate HFIP website for content and revisions
- Update and release the MET-TC software

FY18

- Efficient coupling between various components within NEMS including post-processing and product generation
- Adopt high-resolution nesting strategies to develop global-to-local scale modeling for hurricane forecast applications
- Design efficient high-resolution ensemble strategies to provide probabilistic guidance on track, intensity (including RI), size and structure forecasts

- HWRF Upgrades
    - Global NMMB/NEMS with multiple high-resolution nests
    - Dynamic and adaptive nesting for genesis cases
    - High-resolution ensembles of nests within the global framework
    - Fully cycled EnKF/3D- 4D-VAR DA with advanced satellite DA
    - Continuous advancement of physics for multi-scale interactions
    - Forecasts for tropical cyclones of all ocean basins
- Continue HRD vortex- and convective-scale diagnostics (e.g., RMW, vortex tilt, eyewall slope, convective burst radial and azimuthal locations, CFADs of vertical velocity)
- Continue developing air-sea interaction diagnostics (e.g., continue HRD PBL & surface layer diagnostics, surface fluxes, ocean SST, MLD, OHC, currents, upwelling)
- Release latest version of GFDL vortex tracker to the community
- Release latest version of the CIRA diagnostic and verification codes to the community
- Evaluate HFIP website for content and revisions
- Update and release the MET-TC software
- Update NHC's ATCF requirements list for NRL, and implement changes.

FY19

- Release latest version of GFDL vortex tracker to the community
- Release latest version of the CIRA verification and diagnostic codes to the community
- Ensure that all relevant content from the HFIP website has been transferred to / implemented at NHC (the transition of useful products will begin sooner)
- Update and release the MET-TC software
- Update NHC's ATCF requirements list for NRL, and implement changes
- Full Operational implementation of the system described in 12 by the 2019 hurricane season

9. HFIP Priorities for the Sandy Supplemental funding

After Hurricane Sandy, Congress appropriated funding specifically to improve hurricane forecasts. Some of that funding supported much of the 2013 HFIP effort. This effort is included in the descriptions of the overall program outlined above. The Sandy Supplemental funded parts of the following efforts:

Components of the HWRF System funded

- Plan and conduct the Experimental Numerical Forecast System (real-time) during hurricane season on HFIP Boulder Jet System
- Add and evaluate use of standard reconnaissance data (flight level, dropsondes, SFMR) and the addition of Doppler Radar in HWRF
- Demonstrate Impact of Multi-model ensembles for Track and Intensity using HWRF and COAMPS-TC

Developmental Projects

- Increase use of Satellite Data, in hurricane DA
- Continue to improve HWRF physics
- Develop HWRF GSI hybrid DA--high resolution inner domain/vortex scale
- Other model Improvements
  - Vertical resolution increase to ~128 levels
  - Increase computational efficiency
  - Horizontal resolution increase to 10 km (if computational efficiencies are found)
  - North American Model (NAM) tropical cyclone relocation
- Development of the Global HWRF noted earlier

HWRF Model Infrastructure development

- Convert HWRF to NMM and optimize code for application of accelerator technologies

Global Ensemble upgrade

- Cloud and precipitation physics
- Improved ensemble initialization
- Stochastic surface physics
- Model error estimation (scaling for ensemble perturbations)
- Possible increased horizontal resolution

10. References

Aksoy, A., S. Lorsolo, T. Vukicevic, K. J. Sellwood, S. D. Aberson, F. Zhang, 2012: The HWRF Hurricane Ensemble Data Assimilation System (HEDAS) for High-resolution Data: The Impact of Airborne Doppler Radar Observations in an OSSE. *Monthly Weather Review*, 140(6):1843–1862, (doi:10.1175/MWR-D-11-00212.1).

Benjamin, S., G. Grell, J. Brown, T. Smirnova, and R. Bleck, 2004: Mesoscale Weather Prediction with the RUC Hybrid Isentropic-Terrain-Following Coordinate Model. *Monthly Weather Review*, 132(2):473–494, (doi:10.1175/1520-0493(2004)132<0473:MWPWTR>2.0.CO;2)

Cangialosi, J. P., and J. L. Franklin, 2011: 2010 National Hurricane Center Forecast Verification Report. NOAA, 77 pp. [Available online at http://www.nhc.noaa.gov/verification/pdfs/Verification_2010.pdf.]

Chen, S., T. J. Campbell, H. Jin, S. Gaberšek, R. M. Hodur, and P. Martin, 2010: Effect of Two-way Air–sea Coupling in High and Low Wind Speed Regimes. *Monthly Weather Review*, 138(9):3579–3602, (doi:10.1175/2009MWR3119.1)

Cohn, S. E., and D. F. Parrish, 1991: The Behavior of Forecast Error Covariances for a Kalman Filter in Two Dimensions. *Monthly Weather Review*, 119(8):1757–1785, (doi: 10.1175/1520-0493(1991)119<1757:TBOFEC>2.0.CO;2).

Daley, R. and E Barker, 2001: NAVDAS: Formulation and Diagnostics, *Monthly Weather Review*, 129(4): 869-883, (doi:10.1175/1520-0493(2001)129<0869:NFAD>2.0.CO;2)

Doyle, J.D., Y. Jin, R. Hodur, S. Chen. H. Jin, J. Moskaitis, A. Reinecke, P. Black, J. Cummings, E. Hendricks, T. Holt, C. Liou, M. Peng, C. Reynolds, K. Sashegyi, J. Schmidt, and S. Wang, 2012: Real Ttime Tropical Cyclone Prediction Using COAMPS-TC. *Advances in Geosciences*, 28, Eds. Chun-Chieh Wu and Jianping Gan, World Scientific Publishing Company, Singapore, 15-28.

DTC, cited 2012a: WRF for Hurricanes Documents and Publications. [Available online at http://www.dtcenter.org/HurrWRF/users/docs/index.php.]

DTC, cited 2012b: Community Grid point Statistical Interpolation GSI Documents and Publications. [Available online at http://www.dtcenter.org/com-GSI/users/docs/index.php.]

Evensen, G., 1994: Sequential Data Assimilation with a Nonlinear Quasi-geostrophic Model Using Monte Carlo Methods to Forecast Error Statistics. *Journal of Geophysical Research*, **99**(C5):10,143–10,162.

Hamill, T.M., J. S. Whitaker, M. Fiorino, and S. G. Benjamin, 2011: Global Ensemble Predictions of 2009's Tropical Cyclones Initialized with an Ensemble Kalman Filter. *Monthly Weather Review*, 139(2):668-688, (doi: 10.1175/2010MWR3456.1).

Houtekamer, P. L., and H. L. Mitchell, 1998: Data Assimilation Using an Ensemble Kalman Filter Technique. *Monthly Weather Review*, 126(3):796–811 (doi: 10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2)

Keppenne, C.L., 2000: Data Assimilation into a Primitive-Equation Model with a Parallel Ensemble Kalman Filter. *Monthly Weather Review*, 128(6):1971-1981, (doi: 10.1175/1520-0493(2000)128<1971:DAIAPE>2.0.CO;2).

NHC, cited 2009: NHC Track and Intensity Models. [Available online at http://www.nhc.noaa.gov/modelsummary.shtml.]

NOAA ESRL, cited 2011: FIM Documentation. [Available online at http://fim.noaa.gov/fimdocu_rb.pdf.]

NOAA SAB, cited 2006: Hurricane Intensity Research Working Group Majority Report. [Available online at http://www.sab.noaa.gov/Reports/HIRWG_final73.pdf.]

---------, cited 2008: Proposed Framework for Addressing the National Hurricane Research and Forecast Improvement Initiatives [Available online at http://www.nrc.noaa.gov/plans_docs/HFIP_Plan_073108.pdf]

NRL, cited 2012: NAVDAS Source Book 2001. [Available online at http://www.nrlmry.navy.mil/docs/_NAVDAS01.pdf.]

NSF, cited 2007: Hurricane Warning: The Critical Need for a National Hurricane Research Initiative. [Available online at http://www.nsf.gov/nsb/committees/archive/hurricane/initiative.pdf.]

OFCM, cited 2007: Interagency Strategic Research Plan for Tropical Cyclones - The Way Ahead, FCM-P36-2007 [Available online at http://www.ofcm.gov/p36-isrtc/fcm-p36.htm]

Ota, Y., J. C. Derber, E. Kalnay, and T. Miyoshi, 2013: Ensemble-based Observation Impact Estimates Using the NCEP GFS. Tellus, 2013(65): 20038, (doi: 10.3402/tellusa.v65i0.20038).

Parrish, D.F. and J. C. Derber, 1992: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Monthly Weather Review*, 120(8):1747-1763, (doi:10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2)

PSU, cited 2011: PSU WRF/EnKF Real-time Atlantic Hurricane Forecast. [Available online at ocean http://hfip.psu.edu/realtime/AL2011/forecast_track.html.]

Snyder, C., and F. Zhang, 2003: Assimilation of Simulated Doppler Radar Observations with an Ensemble Kalman Filter. *Monthly Weather Review*, 131(8):1663-1677, (doi: 10.1175//2555.1).

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble Square Root Filters. *Monthly Weather Review*, 131(7):1485-1490, (doi:10.1175/1520-0493(2003)131<1485:ESRF>2.0.CO;2).

Weng, Y. and F. Zhang, 2012: Assimilating Airborne Doppler Radar Observations with an Ensemble Kalman Filter for Convection-Permitting Hurricane Initialization and Prediction: Katrina (2005). *Monthly Weather Review*, 140(3):841-859, (doi: 10.1175/2011MWR3602.1)

Wu, W-S., J. Purser, and D. F. Parrish, 2002: Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances. *Monthly Weather Review*, 130(12):2905-2916 (doi: 10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2).

Zhang, F., Y. Weng, J. Gamache and F. Marks, 2011: Performance of Convection-permitting Hurricane Initialization and Prediction During 2008-2010 with Ensemble Data Assimilation of Inner-core Airborne Doppler Radar Observations. *Geophysics Research Letters*, 38(15), (doi:10.1029/2011GL048469)

APPENDIX: MODEL ACRONYMS

The following is a list of acronyms used to identify models in this document. Many of the acronyms follow the four-character naming convention in the Automated Tropical Cyclone Forecasting (ATCF) system. For example, 6-hour "earlier" forecasts from "late" models (see Section 3) are adjusted so that the previous 6-hour forecast matches the conditions at the beginning of the current forecast. This is simply known as an interpolated forecast. Forecasts of those future conditions are denoted with an "I" at the end, for "interpolated" (12-hour interpolations are denoted with a "2").

Other conventions (although not exclusively) in the model naming include using the acronym "A" to denote advanced version, "D" to denote the addition of inland decay, "E" to denote ensemble, "H" to denote hurricane, "R" to denote research, "S" to denote statistical, "T" to denote track, "V" to denote Variable (ensemble of at least 2, for example), and beginning with an "I" to denote intensity.

AEMI:	GEFS with 6-hour interpolation.
AVNI:	GFS with 6-hour interpolation.
AHW:	National Center for Atmospheric Research Advanced Hurricane WRF.
AHWI:	AHW with 6-hour interpolation.
APSI:	AWR with 6-hour interpolation
ARW:	Pennsylvania State University Advanced Research WRF
CMC:	Canadian Meteorological Centre model.
CMCI:	CMC with 6-hour interpolation.
CLIPER5:	Climate and Persistence model.
COAMPS-TC:	Fleet Numerical Meteorology and Oceanography Center Coupled Ocean/Atmosphere Mesoscale Prediction System-Tropical Cyclone.
COTI:	COAMPS-TC with 6-hour interpolation.
Decay-SHIFOR5:	Decay Statistical Hurricane Intensity Forecast model.
DSHP:	Decay SHIPS.
ECMWF:	European Centre for Medium-range Weather Forecasts model.
EGRI:	United Kingdom Meteorological Office model, subjective tracker, with 6-hour interpolation.

EMXI:	ECMWF with 6-hour interpolation.
FIM:	Flow-following finite-volume Icosahedral Model.
FM9I:	FIM with 6-hour interpolation
FSSE	Florida State University Super Ensemble
G01I:	GFDL ensemble member 01 with 6-hour interpolation (in general, G##I denotes GFDL ensemble member ## with 6-hour interpolation).
GEFS:	National Centers for Environmental Prediction Global Ensemble Forecast System.
GFDI:	GFDL with 6-hour interpolation.
GFDL:	Geophysical Fluid Dynamics Laboratory model.
GFNI:	Navy version of GFDL with 6-hour interpolation.
GFS:	Global Forecast System.
GFSI:	GFS with 6-hour interpolation.
GHMI:	GFDL adjusted using a variable intensity offset correction that is a function of forecast time, with 6-hour interpolation.
GPMN:	GFDL ensemble mean
GPMI:	GFDL ensemble mean (note all members of the ensemble include 6-hour interpolation).
HEDAS:	Hurricane Ensemble Data Assimilation System.
HWFI:	HWRF with 6-hour interpolation.
HWRF:	Hurricane WRF.
HyCOM:	The Hybrid Coordinate Ocean Model. A general circulation ocean model, describing the effects of the tides, winds, earth's rotation, and other factors on flow.
ICON:	National Hurricane Center Intensity Consensus.
IV15:	Intensity forecast ensemble including 2012 stream 1.5 forecasts.
LGEM:	Logistics Growth Equation Model.

NAVGEM: Fleet Numerical Meteorology and Oceanography Center Navy Global Environmental Model (replaced NOGAPS February, 2013).

NGPI: NOGAPS with 6-hour interpolation.

NGXI: Experimental NOGAPS with 6-hour interpolation.

NOGAPS: Fleet Numerical Meteorology and Oceanography Center Navy Operational Global Atmospheric Prediction System (replaced by NAVGEM February, 2013).

NMM: Environmental Modeling Center Nonhydrostatic Mesoscale Model.

POM: Princeton Ocean Model sigma coordinate (terrain-following), free surface ocean model with embedded turbulence and wave sub-models, and wet-dry capability.

SHIPS: Statistical Hurricane Intensity Prediction System.

SPC3: Six member weighted SPICE ensemble using output from GFS, HWRF, and GFDL as input for DSHP and LGEM. The ensemble weights vary with forecast lead time.

SPICE: Statistical Prediction of Intensity from a Consensus Ensemble.

TV15: Track forecast ensemble including 2012 stream 1.5 forecasts.

TVCA: Track Variable Consensus of at least two of AVNI, EGRI, EMXI, NGPI, GHMI, HWFI forecasts

TVCN: National Hurricane Center Track Variable Consensus

UKMI: United Kingdom Meteorological Office model, automated tracker, with 6-hour interpolation.

UWNI: UW-NMS with 6-hour interpolation.

UW-NMS: University of Wisconsin Nonhydrostatic Modeling System

WRF: Weather Research and Forecasting model. It is a regional system with options for the dynamic core, physics, initialization, post processing and verification. Variations include the Hurricane WRF (HWRF), PSU Advanced Research WRF (ARW), and NCAR Advanced Hurricane WRF (AHW)