Towards Improving High-Resolution Numerical Hurricane Forecasting: Influence of Model Horizontal Grid Resolution, Initialization and Physics

*by*

Sundararaman G. Gopalakrishnan[[1]](#footnote-1), Stanley Goldenberg1, Thiago Quirino1, Frank Marks, Jr,1, Xuejin Zhang2, Kao-San Yeh2, Robert Atlas3 and Vijay Tallapragada4.

NOAA/Hurricane Research Division, Atlantic Oceanographic Meteorological Laboratory, Miami, FL, USA;

2 University of Miami, Cooperative Institute for Marine and Atmospheric Studies,

Miami, FL, USA;

3 NOAA/Atlantic Oceanographic Meteorological Laboratory

Miami, FL, USA;

4 NOAA/Environmental Modeling Center, National Centers for Environmental Predictions, Washington, DC, USA

(To be submitted to Weather and Forecasting, April 24, 2011)

Abstract

This paper provides an account of the performance of an experimental version Hurricane Weather Research and Forecasting system (HWRFX) for 87 cases of Atlantic tropical cyclones during the 2005, 2007 and 2009 seasons. The HWRFX system was used to study the influence of model grid resolution, initial conditions and physics. The model was run to produce 126 hours of forecast with two versions of horizontal resolution, namely, (i) a parent domain at a resolution of about 27 km with a 9 km moving nest (27:9) and (ii) a parent domain at a resolution of 9 km with a 3 km moving nest (9:3). The former was selected to be consistent with the current operational resolution, while the latter is the first step in testing the impact of finer resolutions for future versions of the operational model. The two configurations were run with initial conditions for the tropical cyclone obtained from the operational GFDL and HWRF models. In addition, some sensitivity experiments were carried out with the physical parameterization scheme. The study shows that the 9:3 HWRFX system using the GFDL initial conditions and a system of physics similar to the operational version (HWRF) provides the best results in terms of both track and intensity prediction. Use of the HWRF initial condition in the HWRFX model provides reasonable skill particularly when used in cases with initially strong storms (hurricane strength). However, initially weak storms (below hurricane strength) posed special challenges for the models. For the weaker storm cases, none of the predictions from HWRFX runs or the GFDL forecasts provided any consistent improvement when compared to the statistical-dynamical intensity model (DSHIP).

1. Introduction

The past decades have been marked by significant advances in numerical weather prediction models, such as the National Oceanic and Atmospheric Administration’s (NOAA) Global Forecast System (GFS), NOAA’s Geophysical Fluid Dynamics Laboratory (GFDL) regional hurricane forecasting system, U.S. Navy Operational Global Atmospheric Prediction System (NOGAPS),the National Aeronautics and Space Administration (NASA) model, the European Centre for Medium-Range Weather Forecasts (ECMWF) model and the U.K. Met Office model (UKMET). Such advances have contributed greatly to a steady improvement to the official tropical cyclone (TC) track forecasts as issued by NOAA’s National Weather Service’s (NWS) National Hurricane Center (NHC) resulting in substantial reduction in track forecast errors. This in turn has reduced warning and evacuation areas, thereby saving lives and resources. Forecasting intensity changes is also extremely important, especially in the case of storms that rapidly intensify or weaken just prior to landfall (e.g. TC’s Charley, 2004; Katrina and Wilma, 2005; Humberto, 2007; Karl 2010). However, forecasting intensity changes in TCs is a complex and challenging multi-scale problem. While cloud-resolving numerical models using a horizontal grid resolution of 1-3 km are starting to show some skill in predicting the intensity changes in individual cases, it is not clear at this time if such high-resolution models may be able to provide intensity guidance with fidelity. The goal is to produce consistently reliable results in the real-time (operational) venue. Lack of skill in numerical intensity forecasting is often associated with inaccurate initial conditions or limitations in modeling the physical processes within and around the hurricanes. Also, it is not even clear at this time what may be a reasonable horizontal resolution to forecast hurricane intensity changes on a day-to-day basis.

The Weather Research and Forecast (WRF) system for hurricane prediction (HWRF), became operational at the National Centers for Environmental Prediction (NCEP) in 2007. This advanced hurricane prediction system was developed at the Environmental Modeling Center (EMC) to address the Nation's next generation hurricane forecast problems. In this study an experimental version of HWRF (HWRFX) is used to explore all three factors (i.e., model grid resolution, initial conditions and model physics) that may influence the accuracy of track and intensity forecasts, and provide guidance for improvements of the operational HWRF system.

2. Background

The GFDL regional hurricane prediction system originated as research model in the 1970’s (Kurihara and Tuelya 1974; Kurihara 1975; Tuleya and Kurihara 1975, 1978, 1982). In the mid-1980s, the hurricane dynamics group at GFDL began a 10-year effort to transform their research hurricane model into an operational hurricane forecasting tool for NWS. As part of that effort, several years were spent developing a technique to insert a more realistic and model-consistent vortex into the global analysis. The initialization of hurricanes in the GFDL model, which is unique to this model and is relevant to our discussions later, uses a vortex replacement strategy, which consists of three major steps: 1) interpolate the global analysis fields from GFS onto the operational GFDL hurricane model grid; 2) remove the GFS vortex from the global analysis; and 3) add a high resolution, model-consistent vortex (Kurihara, et al. 1995). Since 1995, the GFDL hydrostatic hurricane prediction system has been used operationally by NHC and is regarded among the most skillful and reliable dynamical models by NHC for track forecasts. Since 2005, the model has demonstrated improved skill in intensity predictions, especially after the atmosphere-ocean coupled system (Ginis et al. 1999) and increases in vertical and horizontal resolution were implemented (Bender et al. 2007). A thorough summary of the modeling system and its performance is provided in Bender et al. (2007).

WRF is a general purpose, multi-institutional mesoscale modeling system. NCEP’s Non-hydrostatic Meso Model (NMM; Janjic et al. 2001; Janjic 2003) is a dynamical core option within the WRF model initiative. A moving nest capability was created within the WRF-NMM system to address the hurricane problem (Gopalakrishnan et al. 2006). This dynamical capability is the backbone of the HWRF system. The high-resolution nest is capable of capturing non-hydrostatic scales of motion within the hurricane inner core and has the potential to provide improved intensity guidance. Another advancement of the HWRF system over the GFDL model is the option to use the previous cycle of the model vortex (when available) as the first guess and to assimilate some inner-core observations. This capability was introduced to provide more realistic initial three-dimensional structure and is a critical ingredient towards advancing TC intensity/structure prediction (Liu et al. 2006). To build on the success of the GFDL model, the physics in the initial implementation of HWRF emulated the latest version of the GFDL system (Bender et al. 2007). In addition, the HWRF system, as in the GFDL model, is coupled to a three-dimensional version of the Princeton Ocean Model (POM) modified for hurricane applications over the Atlantic basin. The HWRF system, with 27 km and 9 km parent and movable mesh grid resolutions (27:9; Fig. 1a), respectively, has been used as an operational hurricane forecasting tool at the NWS since 2007. The documentation on the HWRF system (Gopalakrishnan et al. 2010) is now available at the Development Testbed center (DTC), National Center for Atmospheric Research (NCAR), Boulder, CO (<http://www.dtcenter.org/HurrWRF/users/docs/scientific_documents/HWRF_final_2-2_cm.pdf>).

In addition to the operational hurricane models, more recently, cloud-resolving models are being used as research tools to help understand the hurricane intensity prediction problem. For example, in a series of explicit simulations of Hurricane Andrew (1992) using the fifth generation mesoscale model (known as MM5) developed by Pennsylvania State University (PSU) and NCAR at a resolution of about 6 km, Da-Lin Zhang and his colleagues reported several interesting features of the modeled hurricane (Liu et al. 1997, 1999; Yau et al. 2004 and Zhang et al.,1999, 2000, 2001, 2002) that compared well to the observed structure of hurricanes (Marks and Houze 1987; Marks et al. 1992; Willoughby 1979, 1990a and b). As verified against various observations and the post-storm derived “best track” data (positions and intensities) provided by NHC, the PSU/NCAR MM5 simulations reasonably capture many of the features of the inner-core structure of the storm. In particular, the track, the explosive deepening rate, the minimum surface pressure preceding landfall, and the strong surface wind near the shoreline are well reproduced by these simulations (Yau et al. 2004). Also, other recent research suggests that forecasts of hurricane intensity would be improved with very high model resolution (grid spacing ≤ 1 km in the horizontal) that can adequately simulate the hurricane inner-core structures, such as the eyewall and rainbands (Davis et al. 2008, 2011 and Rotunno et al. 2009). For example, in an effort to demonstrate the value of high resolution hurricane forecast, Davis et al. (2011) used the Advanced Hurricane WRF (AHW) model with two sets of 69 simulations covering 10 Atlantic tropical cyclones, each using different horizontal resolution, 12 and 1.33 km. Their statistically significant results indicated that increasing resolution improves hurricane forecasts of both hurricane intensity and some of the hurricane's structure aspects. Using those results as the basis the authors used the AHW model to provide real-time forecasts during the 2009 season. The results were storm specific. While higher resolution improved forecasts of track and intensity especially on some of the weaker storms like Erica and Danny, the impacts were mixed on some other cases. This raises the question on what control the overall intensity and track predictions: resolution alone, physics, initial conditions, or all of these factors?

While the operational track forecasts show substantial skill (on the order of 30-40% at 72 h) versus the CLImatology PERsistence (CLIPER) models utilized by NHC (e.g., Gross 1999) the skill in predicting intensity changes using dynamical model has been marginal (DeMaria and Kaplan 1999 and Kaplan et al. 2010). To address this problem of lack of intensity prediction skill [with an emphasis on rapid intensification (RI) events], in 2007 NOAA established the Hurricane Forecast Improvement Project (HFIP) [http://www.nrc.noaa.gov/plans\_docs/HFIP\_Plan\_ 073108.pdf](http://www.nrc.noaa.gov/plans_docs/HFIP_Plan_%20073108.pdf)).

The HWRFX is a version of the HWRF system specifically adopted and developed at the Hurricane Research Division (HRD) of the Atlantic Oceanographic and Meteorological Laboratory (AOML) to study the intensity change problem at the cloud-resolving scales (about 1-3 km). This modeling system is supported by HFIP and compliments the operational HWRF system. The data acquisitioncomponent, designed with Java and shell scripts, runs automatically and in parallel with the triggering mechanism. It continuously retrieves meteorological data sets required to perform the HWRFX forecasts. These data sets include the initial conditions and the 126-hour forecast data of the operational Global Forecast System (GFS), the initial conditions of the operational GFDL hurricane model, and the initial conditions of the operational HWRF model. The HWRFX model can be run with at least two suites of physics options. The first suite mimics the operational HWRF suite of physics (Gopalakrishnan, et al. 2011) and the second is a suite of physics that was developed as part of high resolution modeling effort at NCAR. The physics options used in this work are discussed in detail later in the text. As in the case of the GFDL model developments, HWRFX can be run in both an idealized research framework as well as in real-time mode. As a first step, HWRFX was used in an idealized framework to gain a fundamental understanding of the influence of horizontal grid resolution on the dynamics of hurricane vortex intensification in three dimensions (Gopalakrishnan, et al. 2011). Based on a series of numerical experiments at the current operational resolution of about 9 km for the movable fine mesh and at a finer “research” resolution of about 3 km (Fig.1a and b), it was found that improved resolution had very little impact on the initial spin up of the vortex. However, the mature phase of the storm’s evolution exhibited significantly different behavior at 9 and 3 km with the finer resolution significantly improving the predictions of the storm’s structure.

3. HFIP High-Resolution Test Plan

The first step towards the HFIP effort was to quantify the impact on intensity forecasting of increased horizontal grid resolution in numerical models. In early 2009 HFIP established the High Resolution Hurricane (HRH) tests. The plan for these tests was developed jointly by several segments of the community including specialists in hurricanes, numerical modeling, and forecast verification. Hurricane specialists at NHC selected 69 cases from 10 different hurricanes during the 2005 (Hurricanes Emily, Katrina, Ophelia, Philippe, Rita, and Wilma) and 2007 (Felix, Humberto, Ingrid, Karen) seasons for the HRH tests. The 69 HRH cases were designed to examine a variety of aspects of the model’s performance such as (1) the capability to simulate RI events (e.g. Katrina, Rita, Wilma, Humberto and Karen; see Table 1 and Fig. 2), (2) the impact of vertical wind shear on intensity evolution (e.g. Katrina, Philippe, Rita, Ingrid and Karen) (3) the effects of eyewall replacement (e.g. Emily, Rita, Wilma and Felix), (4) the effects of oceanic heating and cooling (e.g. Ophelia), and (5) the impact of terrain interaction during landfall on track and intensity predictions (e.g., Katrina, Rita and Wilma).

Zhang et al. (2011) used the HWRFX system to study the impact of resolution on the 69 cases. The GFDL initial conditions (of the storm vortex and immediate environment) and physics that emulate the operational system were used for this testing. The high-resolution forecasts, using a uniform resolution of 9 km with a moving nest at 3 km, demonstrate a reduction in the absolute track errors vs. the low-resolution forecasts at all forecast time intervals. Forecasts at both high- and low-resolutions show a westward and southward bias with the exception of the low-resolution forecast for 108 and 120 hours. However, results for the frequency of superior performance (FSP), a measure (in percent) of how often one model produces a better forecast than the other, are mixed. In contrast, use of the high-resolution model reduces the absolute intensity errors in comparison to the low-resolution model, with the exception of the 84-hour forecast. The intensity biases appear to be much better balanced over the forecast period with the high-resolution model. The high-resolution model has a smaller negative bias than the low-resolution model prior to 60 hours, but that changes to a positive bias at later forecast times. This demonstrates that the high-resolution model tends to intensify storms more than the low-resolution model. FSP results for intensity errors show that high-resolution forecasts also demonstrate general improvement over the low-resolution forecasts. However these results do not include a significant number of forecasts of initially (at start of forecast) weak (i.e., below hurricane strength) storms and of sheared systems.

The 2009 season was a challenging one for forecasts in part due to several weak and sheared storms (e.g., Ana, Erica and Danny). As a part of the HFIP real-time Demo-system, HWRFX was used for the 2009 season. For this study, cases from five 2009 storms (Ana, Bill, Danny and Erica) were added to the 2005 and 2007[[2]](#footnote-2) HRH storm sample to test the sensitivity of the HWRFX system to resolution, initial conditions and physics. In all there were 87 cases (Table 1; Fig. 2). Although this is a modest number of cases compared to verification samples for operational purposes, the sample size should be sufficient for the sensitivity tests presented here. These results will hopefully provide insights that will later be translated into improvements to operational forecasting. The 87 cases comprise a diverse sample of initial intensities, intensifying (including a number of RI events) and weakening situations and track types (see Table 1 and Fig. 2) to help understand the impacts of various factors that may influence the model performance. Also, this is the first time that the sensitivity of a version of the HWRF model to initial conditions and physics in addition to resolution has been examined for a comprehensive list of cases.

4. Resolution/Configuration, Vortex Initialization and Model Physics

Figure 1 shows the two model configurations (resolutions) used in this study. The domain of simulation along the horizontal direction was set to about 50 x 50 degrees with a moving nest of about 7 x 7 degrees. There were 40 hybrid levels along the vertical with the top level set to 50 hPa. The simulations reported here were performed with two kinds of horizontal resolution, namely, (i) a parent domain and moving nest with resolutions of about 27 km and 9 km, respectively (27:9 henceforth; Fig. 1a) and (ii) a parent domain and moving nest with resolutions of 9 km and 3 km, respectively (9:3 henceforth; Fig. 1b). While the former was picked to be consistent with the current operational resolution, the latter could become the possible operational resolution in the near future. All the results reported in this study use moving nests at 9-km (i.e., 27:9) or 3-km (i.e., 9:3) resolution. The model was run for 126 hours (more than 5 days). As mentioned earlier, the HWRFX system has an option to be initialized using either the operational HWRF grib or the GFDL grib products. The WRF post-processor (WPS) was used to initialize the model with either the HWRF or the GFDL vortex initializations. The GFS forecasts were used to produce boundary conditions for all the cases reported here.

A description of the operational HWRF suite of physics has been recently published as a technical document at the DTC, NCAR, Boulder (Gopalakrishnan et al. 2011). The HWRFX physics options used in this study are configured as close as possible to the operational HWRF system. The so-called GFDL surface (Bender et al. 2007) and GFS boundary layer formulations (Hong and Pan 1996) are used to parameterize the flux transport and the subsequent mixing in the atmosphere. The Ferrier scheme (Ferrier et al. 2002) is used to provide latent heating due to the microphysical processes in the atmosphere, and the Simplified Arakawa and Schubert scheme (SAS) is used to parameterize subgrid cumulus-cloud activity. The scheme includes a term for momentum mixing (Hong and Pan 1998) that is parameterized as a drag term in the model. Cumulus parameterization in combination with the Ferrier microphysical scheme have been found to have some value in operational NMM in scales down to about 5 km. However, the contribution to heating from SAS has diminishing returns, i.e., an increase in resolution and the grid volume in the inner core region quickly becoming saturated with the use of an explicit microphysics scheme alone and consequently we switched off the SAS convection scheme at 3 km resolution in this work consistent with Zhang et al (2011). Finally, to keep the radiation option simple, the effect of radiation on the atmosphere is approximated with the NCAR long- and short-wave radiation scheme available in the WRF framework. Detailed descriptions of the physics options in HWFX with appropriate references are reported in Yeh et al. (2011). Although the physics options in HWRFX were kept as close as possible to the operational HWRF physics, apart from the difference in the radiation scheme, there were several subtle but significant changes adopted in this study. We believe that it may be worthwhile to contrast these differences at least for the 27:9 version. In Table 2 we specify some main differences in physics between the 27:9 forecasts and the operational HWRF physics.

The WRF model provides an option for several combinations of physics packages (<http://www.mmm.ucar.edu/wrf/users/tutorial/200707/WRF_Physics_Dudhia.pdf>). In the recent past, a number of studies have reported inter comparisons between these packages for the specific case of a hurricane (example: Li and Pu 2009 and Fierro et al. 2009) or for a range of scale spanning problems such as warm season regional forecast (example: Gallus and Bresch 2006) and hurricane forecasting (Davis et al. 2011). The so called “Kain–Fritsch (KF; Kain and Fritsch 1992) convective scheme, the Yonsei University (YSU; Noh et al. 2003) PBL scheme and WRF-Single-Moment-Microphysics (WSM5)” is a well-documented combination for the hurricane applications (Davis et al. 2011). In this study we have used the YSU-WSM5-KF combination as the second option for testing. The effect of solar radiation on the atmosphere is approximated with the NCAR long- and short-wave radiation scheme available in the WRF framework in all the runs. Both physics packages in this study used the NOAH land surface scheme (Ek et al. 2003). Finally, in all the runs the sea surface temperatures (SSTs) were obtained from the GFS and were held fixed during the forecast. The summary of sensitivity experiments performed using the 87 cases as the basis are described in Table 3.

The track and intensity of a storm were determined on the basis of the position of the nest within the parent domain (Fig. 1), which in turn is based on the concept of dynamic pressure (Gopalakrishnan et al*.* 2002). At the end of every time step of the nested domain, the centroid of the dynamic pressure within this moving domain is determined. The minimum dynamic pressure determines the storm center. If the storm center is advected beyond one grid point of the parent domain (three grid points from the center of the nested domain due to the 3:1 parent-to-nest grid ratio), the nested domain is moved to a new position within the parent domain to maintain the storm near the center of the nested domain. The post processing software reads the history output file from the moving nest, which carries the information of the minimum sea level pressure, maximum wind, radius of maximum wind and the storm location corresponding to the minimum sea level pressure of the nested domain. As with any numerical model, at times tracking problems occurred where the algorithm (or nested grid) lost the storm center, was affected by a storm’s proximity to the boundaries or the nested grid algorithm jumped to a neighboring low pressure or other storm. These cases were removed typically by screening for situations when the storm center as determined from the parent domain and the center determined from the nested grid differed by more than ~200km. Once a center position was determined suspect, the rest of that forecast (position and intensity) was removed from the sample. Improvements to the automated tracking algorithm are now being developed but were not available when the data for this study were compiled.

5. Results and Discussions

Given the large amount of high-resolution data, it may not be feasible here to discuss our results in terms of individual cases. So for this manuscript, only statistical verification of the runs will be presented. We report the results in terms of standard metrics such as (i) absolute track errors, and (ii) intensity errors (absolute and bias). These measures provide only a part of the picture of the potential of the high-resolution model to improve track and intensity forecasts. In addition to these standard metrics we have examined the radius of maximum wind as a measure for structure predictions. The cumulative distribution function (CDF) of the radius of maximum wind at 10 meters above the ground for all the HWRFX forecasts were compared with the Hurricane Research Division’s HWIND analysis.

Models based solely on climatology and persistence are created from statistical relationships between storm-specific information, such as location, time of year, and the behavior of historical storms. For track forecasts, NHC’s operational CLImatology and PERsistence model is CLIPER5 (Neumann 1972). For intensity forecasts, it is DSHIFOR5[[3]](#footnote-3). DSHIFOR5 (Decay-SHIFOR) is the SHIFOR5 (Statical Hurricane Intensity FORecast) model (Jarvinen and Neumann 1979, Knaff et al. 2003) adjusted for the inland decay rate of Demaria et al. (2006). NHC and others (Aberson 1998, Franklin 2010) routinely use these models for establishing the “skill” of research and operational models, where the CLIPER5 and DSHIFOR5 are regarded as a type of “no-skill” baseline since these models provide a simple measure of whether forecast situations are “easy” or “difficult”. Track forecast results in this study are shown in terms of “skill” versus CLIPER5. For intensity forecast verifications, NHC plots a comparable “DSHIFOR5 diagram”. However, because the ultimate focus of this project is to improve the intensity forecasts for dynamical models, a higher standard for the baseline, DSHIPS, was used. SHIPS (Statistical Hurricane Intensity Predicton Scheme) is a sophisticated statistical-dynamical model which predicts storm intensity utilizing multiple regression relationships with climatological, persistence and numerical model predictors (DeMaria et al. 2005). DSHIPS (Decay-SHIPS) is SHIPS adjusted for decay of the storms when they move inland according to Demaria et al. (2006) and is regarded as one of the most reliable intensity forecast models by NHC (Franklin 2010). In this study we use CLIPER5 and DSHIPS as a basis for comparison of tracks and intensity predictions, respectively from HWRFX.[[4]](#footnote-4)

5.1 Track Errors

Figure 3 provides an overview of the track error statistics. The mean track errors for all runs for the two resolutions (i.e. 27:9 and 9:3) and two initial conditions (i.e., GFDL and HWRF) are presented here, where H9hwrf and H9gfdl refer to results from 27:9 with HWRF and GFDL initial conditions, respectively, and H3hwrf and H3gfdl refer to results from 9:3 with HWRF and GFDL initial conditions, respectively (Table 3). The results are also compared with the operational GFDL model. As expected, the number of cases decreases for longer forecast intervals and consequently while there are 87 entries for the 12-h forecast, there are only 26 entries for the 120-h forecast. The average track errors for the various models increase almost linearly from near 60 km at 12 h to as low as 436 km (H3gfdl) and as high as 529 km (H9gfdl) at 120 h (Fig. 3a). In general, the large-scale variations that are expected to have a major influence on the TC tracks (Riehl 1954) are well captured by the HWRFX system in all the four model versions. The impacts of resolution and initial conditions are better seen when the results are normalized with reference to CLIPER5 (Fig. 3b). The operational GFDL model provides the best overall skill. The H3gfdl version of HWRFX has nearly the same skill level overall as the GFDL model. Consistent with the results shown in Zhang et al. (2011), the GFDL initialization with the 9:3 HWRFX version (H3gfdl) is more skilful with the 27:9 version (H9gfdl). On the contrary, the HWRFX prediction at 27:9 resolution with the corresponding HWRF initialization (H9hwrf) is more skillful in the mid-forecast intervals compared to the 9:3 version with the HWRF initialization (H3hwrf).

5.2 Intensity Errors

Figure 4 depicts the intensity error statistics from HWRFX. The intensity forecasts show a large diversity in behavior (Fig. 4a), with H3hwrf as the outlier with the largest errors. At 12 h the intensity error for this initialization is close to 9 ms-1 whereas the same initialization with the 27:9 resolution (i.e., H9hwrf) produces an intensity error of about 7.4 ms-1. When the average intensity forecast errors are shown as skill relative to DSHIPS (Fig. 3b), the dynamical models are better from at least about 36-96 h, with the exception of H3hwrf. These are encouraging results considering that DSHIPS is regarded as one of the most reliable intensity forecast models. [DSHIPS shows consistent skill relative to DSHIFOR5 (Franklin 2010).] And for this sample, the intensity forecast error results from the various versions of HWRFX (again with the exception of H3hwrf) are comparable or sometimes better than the results from the operational GFDL model from 36 to 120 h.

Looking at the impact of resolution and initial conditions, for this sample there is only a marginal improvement at 12, 24, 60 and 72 h using the increased resolution (9:3) with GFDL initial conditions (i.e., H3gfdl, vs. H9gfdl). However, using the GFDL initial conditions, Zhang et al. (2011) found a more consistent improvement in forecast with improved resolution. The change in results here are likely due to the addition of the sheared and weak cases from 2009, demonstrating the lack of stationarity in verification statistics with these modest sample sizes. On the other hand, runs with the HWRF initial conditions (H9hwrf vs. H3hwrf) perform much better for the lower resolution (27:9 i.e., H9hwrf). This was likely caused by the superfluous use of the bogus vortex and the overly restricted storm structure in the operational HWRF initialization (Liu et al. 2006) resulting in the artificial effects being amplified by the increase of resolution. Summarizing, the two resolutions give more or less similar results when using the GFDL initial conditions at least during 36-96 hours of forecasting, but the HWRF initial conditions appear to be much better suited to the coarser grid.

In order to further understand the model behavior the sample was stratified based on the initial intensity of each storm. Figures 5a and 5b show the results in terms of skill (vs. DSHIPS) for average absolute intensity forecast errors for storms with initial intensity (maximum sustained surface wind speed) ≥ 33.4 ms-1 (i.e., hurricane strength) and for storms with initial intensity < 33.4 ms-1 (i.e., less than hurricane intensity), respectively (see Table 1). Of course this further reduces the already modest sample size at the 12 h forecast interval to 38 cases for initially stronger storms and 49 cases for initially weaker storms and the sample sizes are even smaller for the other forecast intervals. Since the results with DSHIPS for the stratified samples (not shown) demonstrate very little differences for absolute errors between the runs with initially weaker and stronger storms, the baseline used for the skill diagrams is fairly stable. Therefore the skill differences discussed here for the stratification are the result primarily of differences in model results rather than the baseline.A comparison between Figs. 5a and b demonstrates that the skill in predicting intensity for initially stronger storms far exceed those of the initially weaker storms for the given sample. In fact, with the exception of the 12- and 24-h forecast intervals for the HWRFX with GFDL initial conditions (i.e., H9gfdl and H3gfdl), the skill is much better for every dynamical model represented here with the skill level generally higher than 20% from 24 to at least 96 h. In contrast, for the weaker storms, the skill was negative for every model at every forecast interval with the exception of a few models showing slightly positive skill from 60 to 84 h. For initially strong and weak storms the results for runs using the HWRF initialization (H9hwrf and H3hwrf) are fairly close except for 12 and 24 h, and also for 36 h with weak storms only. For runs with GFDL initial conditions (H9gfdl and H3gfdl), the finer resolution has better skill until 72 h. However, the sample size was limited to 17 at that time and goes down to only 9 cases at 120 h. Nevertheless, for the stronger storms, in general the HWRF initialization for the 27:9 km HWRFX modeling system performs better than the GFDL initialization at almost all lead times.

It should be noted that the extremely poor performance for the overall sample (Figs. 4a and b) of the fine grid version of HWRFX using HWRF initial conditions (H3hwrf) is mostly due to the large errors from the weak storm sample (see Fig. 5a vs. 5b). In fact, H3hwrf performs well out to 72-84 h for the strong intensity cases, but extremely poor relative to all of the other models depicted for the weaker cases. One would have to conclude that overall, the GFDL initial conditions (H3gfdl) are better suited (for the cases in this study) when using the 9:3 km version. And overall, the HWRF initial conditions (H9hwrf) are better suited when using the 27:9 km version.

With the exception of H3hwrf, in general the results from the various versions of HWRFX performed comparably to the operational GFDL model (Fig. 5a). As opposed to initially strong storms, an initially weak vortex has a very dramatic influence on prediction skills (Fig. 5b). Clearly there is complete lack in skill at almost all lead times. Results for intensity bias statistics discussed in the subsequent section may be able to shed light on these skill levels, at least for the initial period of forecast. Even with the relatively small sample sizes, these results are striking and make physical sense in that the models would be expected to have a more difficult time initializing and maintaining an initially weaker vortex. The results for the actual absolute intensity errors for initially strong and weak storms, and the full sample of cases (not shown) are also very consistent, with the lowest errors associated with the initially strong storms at all forecast intervals showing that these stratified results are not biased by scaling with the DSHIPS predictions. Stratification of the track forecast errors (not shown) also yields very similar results, with skill much higher for initially stronger storms.

5.3 Intensity Bias

Figure 6 compares the intensity bias for the models (four versions of HWRFX, GFDL and DSHIPS). Figures 6a, b and c show the biases for all the runs (87 cases at 12 h), initially strong storms (38 cases at 12 h) and initially weak storms (49 cases at 12 h), respectively. Several features are worth noting:

1. For the complete sample (Fig. 6a), the 27:9 km runs with GFDL initialization (H9gfdl) produce noticeable negative bias at lead times through 84 h. This is primarily from the strong storm cases (Fig. 6b) through about 60 h (see 5.3(ii)) and mainly from the weaker cases (Fig. 6c) at 84 h. The results also show a large negative bias for the 9:3 runs with GFDL initialization (H3gfdl) for the initially stronger storms (Fig. 6b) through 24 h. Yeh et al. (2011) used the GFDL initialization for HWRFX at 27:9 and studied the performance of the modeling system for the 2008 season. The study concluded that the early negative bias of the HWRFX results from the dynamical inconsistency due to the use of GFDL initial conditions, but the continuing negative bias is probably related to inadequate physical forcing for the given configuration (27:9). The study also suggested the need for model-consistent initial conditions for the HWRFX system. The use of initial conditions from the operational HWRF system is hence justified. In the present study the use of HWRF initial conditions for the HWRFX runs at the same resolution (i.e., 27:9) produces an improved intensity bias and a better intensity forecast (Fig. 5b). “Downscaling” (i.e., using the operational HWRF initial conditions at from the 27:9 version to drive the experimental 9:3 version: H3hwrf) produces noticeable positive bias (Fig. 6a). The positive bias, however, is almost entirely due to the errors from the weak cases (Fig. 6c).
2. The bias for H9gfdl for strong storms is negative through at least 84 h and especially large through at least 36 h (Fig. 6b) suggesting that a strong initial vortex invariably spins down in the first 12 hours of forecast. It takes 24-36 hours for the vortex to recover and this is partly reflected in the poor model skill for the same period (Fig. 5a). The intensity bias (Fig. 6b) is strongly associated (or correlated) with the model performance for the strong storms (Fig. 5a) for the other models as well. The GFDL initialization for the 9:3 km runs (H3gfdl) shows a noticeable spin down but this lasts for less than 24 hours (Fig. 6b) and consequently the model recovers from the initial shock to produce more skillful guidance than DSHIPS by 24 h (Fig. 5a). There is a systematic improvement (i.e., reduction) in the initial negative bias with resolution for stronger storms (fig. 6b)
3. Weak cases present special challenges in storm prediction (Fig. 6c). The HWRFX runs with HWRF initialization (H9hwrf and H3hwrf) show substantial positive bias even at 12 hours and the positive biases persist through 120 h. The worst biases are seen for H3hwrf and this is reflected in the very poor intensity forecasts (Fig. 5b). H9hwrf and even the GFDL model predictions also show significant initial positive bias suggesting that these models are cyclogenic in behavior. On the contrary, DSHIPS, which shows negligible bias for the overall and strong case samples (Figs. 6a and b, respectively) shows negligible bias for the weak cases through 36 hours but then maintains a negative bias throughout the rest of the forecast, especially from 60 through 96 h. While improved initial conditions for weak and sheared storms may be of pivotal importance in improving high-resolution model forecasting, it appears that improved physics especially for longer lead time of forecasting may also be an important issue. We explore some preliminary sensitivity experiments in the Section 6.

5.4 Preliminary evaluation of structure

The radius of maximum wind (RMW) is defined to be the distance between the center of a TC and its band of strongest wind. It is considered an important parameter in the TC dynamics and forecasting. In a recent study using idealized initial conditions Gopalakrishnan et al. (2011) demonstrated that by changing the initial radius of maximum wind the structure and intensity of modeled storms are significantly altered. Yet RMW is one of the parameters that has large measurement uncertainties except where aircraft reconnaissance data are available. As progress is made towards improvements leading to structure and intensity predictions with high-resolution models, there is a need for an evaluation of RMW from the model outputs. Recently, Zhang et al. (2011) used the H\*WIND analysis produced at the HRD (Powell et al. 1998) to evaluate the predictions from HWRFX. The same analysis procedure is followed here. Figure 7 provides the Cumulative Distribution Function (CDF) of the RMW[[5]](#footnote-5) from the HWRFX forecasts with the probability counted in 1-km bins of RMW, compared with those from the H\*WIND analysis. Excluding the samples in which the storm centers were located over land, totals of 230 H\*WIND analyses, and 2,653 GFDL and 2966 HWRF initialized and 2,704 GFDL and 2865 HWRF initialized low- and high-resolution model samples, respectively, are used to calculate the distribution functions. Figure 7a shows that, according to the H\*WIND analysis, 60% (from the 20th to 80th percentile) of the observed RMW are distributed over 20-64 km. For the high- and low-resolution HWRFX configurations using GFDL initial conditions (H3gfdl .vs. H9gfdl; Fig. 7a), 60% of the simulated RMW are distributed over 34-93 km and over 43-104 km, respectively. This indicates that the storm sizes are better predicted with the high-resolution forecasts than with the low-resolution forecasts. Consistent with earlier results, there is a small degradation with the use of HWRF initial conditions (Fig. 7b). However, when compared to the intensity errors produced at higher resolution using the HWRF initial conditions (H3hwrf .vs. H9hwrf), there is a noticeable improvement in predictions of RMW with resolution even with the HWRF initial conditions (H3hwrf).

6. Additional Sensitivity Experiments

The operational HWRF system is likely to be run at about 3 km resolution starting from the 2012 season using the advanced, model consistent initialization procedure described in Section 4. In order to compliment the HWRF system and provide some inputs for further improvements to that system at 3 km resolution, some sensitivity experiments were performed to its high-resolution proto type (HWRFX). Specifically the model sensitivity to (1) model physics suite and (2) lateral diffusion coefficient were examined.

6.1 Sensitivity to model physics

As mentioned in Section 4, while the “GFS-FERRIER-SAS” combination for surface and boundary layer, microphysics and cumulus convection parameterization scheme is used in the operational HWRF model, “YSU-WSM5-KF” is another well-researched combination and is used in the Advanced Research WRF (ARW) model. However, it still remains a great challenge for the research community to reach a consensus on whether the current physics parameterizations in operational NWP models are suitable for horizontal grid spacing of 3 km or smaller. Using physics packages reported in Table 2, Gopalakrishnan et al. (2011) performed a series of idealized experiments with HWRFX at current operational resolution of 9 km and at 3 km, a resolution proposed to be used in NOAA operations. They found that the “GFS-FERRIER-SAS” combination may be extended to higher resolution and produced some realistic intensification of an isolated hurricane vortex at 3-km-resolution, as well. Nevertheless, more recently Bao et al. (2011; personal communication) performed a series of idealized experiments with HWRFX and found significant sensitivity of the vortex intensification process to the boundary-layer parameterization, convective scheme followed by the microphysics scheme. They also found that the GFS boundary-layer scheme produces deeper inflow and larger storms. Those studies are complimented here by examining the sensitivity of the TC tracks and intensity statistics for the real cases (Table 1) to the model physics. In this case the 9:3 configuration of HWRFX was run with HWRF initial conditions but with the physics options set to the “YSU-WSM5-KF” combination.

Figure 8 provides an overview of track and intensity prediction skills of HWRFX. Figure 8a provides the track statistics obtained from using the modified physics combination (H3Nhwrf). There is only a slight difference between the new combination (H3Nhwrf) and “GFS-FERRIER-SAS” package (H3hwrf) except at 120 h when there is a 5-7% improvement in skill. However the sample size are limited at the larger forecast intervals. Nevertheless, “YSU-WSM5-KF” (H3Nhwrf) produces very noticeable improvements on the overall intensity statistics (Fig. 8b) at all forecast intervals except for 12 and 84 h. The skill score significantly improves as when compared to the 3-km-run with HWRF initial conditions and “GFS-FERRIER-SAS” physics (H3hwrf), and the H3Nhwrf results are comparable with the 3-km-run with GFDL initial conditions or the 9-km-run with GFDL or HWRF initial conditions with “GFS-FERRIER-SAS” physics (see Fig. 4b). It should be noted that the GFDL model still provides the best skills in terms of commonly used skill measures for this sample. When stratifying the sample by initially weak and strong storms (Fig. 8c and d), the “YSU-WSM5-KF” combination in HWRFX (H3Nhwrf) produces modest improvements for the initially strong vortex cases (Fig. 8c) from 24 to 60. There is even about 5% improvement over the GFDL predictions for those times. Between 96 and 120 hours there is a significant improvement using the modified physics scheme, even showing improvement compared with the GFDL model results for 108 and 120 h. However the sample sizes are very limited at the later forecast intervals. In the case of initially weak storms the “YSU-WSM5-KF” combination produces intensity skills much better than the 3 km HWRF runs with “GFS-FERRIER-SAS” combination at all forecast intervals except for 84 h (with skill improvements on the order of 30% between 48 and 72 h) and the results are even equivalent to the GFDL forecasts from 48 to 72 h. However, even with the improvement, H3Nhwrf still demonstrates no skill when compared to DSHIPS predictions (except slight skill at 72 h) for the initially weak storms. All this suggests that improving initialization techniques (e.g., Ensemble Kalman Filter technique) has the potential to provide improved predictions. A careful re-examination of these results in terms of structure predictions needs to be done in future to better understand the source and nature of the improvements.

6.2 Sensitivity to lateral diffusion

Apart from sensitivity to modeled physical process discussed above, recent work by Rotunno et al (2009) illustrates the importance of horizontal diffusion especially in the eye wall region of the hurricanes. Nevertheless the application of lateral diffusion in atmospheric models in general and hurricane models in particular has always been a subject of debate since it is not clear to what extent one can model the horizontal diffusion. At this time the diffusion along the horizontal direction is reduced to a minimum value in order to alleviate the numerical dicretization issue (Janjic 1990). In the HWRF/HWRFX system that uses the NMM core as the basis, the lateral diffusion is formulated following the Smagorinsky type non-linear approach (Janjic 1990), where the horizontal diffusion coefficient is defined as K\_h=(Smagorinsky constant\*minimum grid length\*diffusion strength). The Smagorinsky constant is usually set as a tunable parameter. Apart from horizontal deformation Janjic (1990) included the diffusion strength to be a function of the turbulence kinetic energy (TKE). Also the namelist parameter, known as "COAC", is an input parameter and is defined as 10\* (Smagorinsky constant)\*\*2. This coefficient was set to 1.6 in the parent domain and 0.7 in the nested domain in the operational HWRF and in all of the earlier experiments as well (Table 3). However, unlike the NMM system, the HWRF and HWRFX systems are operated with the GFS physics that does not include an additional equation for the TKE. Consequently one can expect an effective reduction in lateral diffusion with a different system of physics that does not include a TKE equation for the NMM core. The COAC was increased to 5 in the inner domain but retained at 1.6 in the outer domain (i.e., increased the lateral diffusion by a factor of roughly 3 in comparison with the parent domain which is a third of the resolution of the nested domain) and the 9:3 model configuration was re-ran, still using the with HWRF initial conditions. The standard option of “GFS-FERRIER-SAS” physics package was used in this case.

Figure 8 also provides an overview of the sensitivity of the track and intensity statistics to changes in the lateral diffusion coefficient in the 3-km-domain. Increase in lateral diffusion in the moving nest (H3Chwrf) produces only small improvements to track predictions at some forecast intervals (Fig. 8a), however, it produces significant improvements to the intensity predictions at all forecast intervals (Fig. 8b). From 24 through 108 h the intensity forecasts with enhanced diffusion are close to the GFDL model predictions. Figures 8c and d show the skill results (vs. DSHIPS) stratified by initial storm intensity. A comparison between the 3 km HWRFX predictions using HWRF initial conditions (H3hwrf) and the current set of results (H3Chwrf) illustrates that the effect of increasing lateral diffusion produces a slight improvement at most forecast times for the initially strong (hurricane strength) storms. However, the improvement for the initially weak storms is dramatic at all forecast times, producing results comparable to the GFDL model forecasts at some forecast intervals. It is very likely that increasing lateral diffusion eases the strong gradients in the eye-wall region especially in the case of initially weak storms and, yet, does not deter the skills on strong initial vortex (Fig. 8c). All this leads to improved skills in the overall intensity predictions with HWRF initial conditions at 3 km resolution.

7. Summary and Conclusions

In this study HWRFX, the research version of the operational HWRF system, was run for 5 days at two different horizontal resolutions, namely, (i) a parent domain with a resolution of about 27 km with a 9 km moving nest (i.e., 27:9) and (ii) a parent domain at a resolution of 9 km with a moving nest at 3 km (i.e., 9:3) each using two different initial conditions namely, (a) GFDL and (b) operational HWRF initializations. As a part of sensitivity experiments the NCAR suite of physics was used with the operational HWRF initial conditions for the fine grid model version (9:3 km). Also, sensitivity experiments were done using the 9:3 km version with HWRF initial conditions by adjusting the diffusion coefficient. This study has shown that:

1. The 9:3 km runs with the HWRFX system using the GFDL initial conditions and a system of physics close to the operational version of the HWRF model (i.e, H3gfdl) provided the best overall skill in terms of both track and intensity predictions. This configuration also produced improved structure prediction, as measured in terms of probability estimates of the radius of maximum wind.
2. The 27:9 km runs with the HWRFX system using the HWRF initial conditions and a system of physics close to the operational version of the HWRF model (H9hwrf) provided reasonable skill in terms of both track and intensity predictions. However, the 9:3 km version with the same configuration produced very poor results showing that the HWRF initial condition may need to be adjusted to be consistent with the finer scale version of the model. Work is currently proceeding in this area.
3. In cases of initially strong storms (hurricane strength), initial spin up of the hurricane vortex in HWRFX appears to have a major influence on the intensity bias and errors at least during the first 36 h of the model run. In general the model appears to spin down the initial vortex for strong storms when using the GFDL initial conditions. Use of HWRF initial conditions for the stronger storms reduces the model bias and improves skill between 0-24 hours. Nevertheless the operational GFDL model performs better than any other models discussed here for the 12 h forecasts for initially strong storms.
4. Forecasting initially weak storms poses major issues and needs some immediate attention. Other than some marginal skill from 72-84 h, none of the models including GFDL forecasts had any skill relative to DSHIPS for these cases. Also, the 9:3 km runs using HWRF initialization (H3hwrf) illustrated a substantial positive bias indicating overly cyclogenic behavior of HWRFX for the initially weaker storms. Increasing the lateral diffusion on the 3 km domain for runs initialized with the HWRF initial conditions substantially improves the intensity skill scores leading to scores that equal the GFDL forecasts. But those improvements are marginal compared to the DSHIPS predictions.
5. Increasing the lateral diffusion on the 3-km-domain runs with operational “GFS-FERRIER-SAS” combination for surface and boundary layer, microphysics and cumulus convection parameterization scheme and HWRF initial conditions has nearly the same effect on track and intensity skill as using the NCAR “YSU-WSM5-KF” physics combination. A further analysis beyond the standard metrics adopted here will be the subject of a future work.

The research version of the system is not identical to the operational system. Minor differences in physics, domain size, methods in which the initial conditions are generated and ocean coupling in the operational version all contribute differences in the versions (i.e., HWRFX .vs. HWRF). As the research and operational communities continue to work together under the auspices of HFIP to address the improvement of tropical storm intensity forecasts, it remains a great challenge for the community to reach a consensus unified version of the modeling system that will be useful for providing input for operations. HFIP continues working towards this goal. The HWRFV3.2 is an ocean coupled and high resolution operational and research system that is currently under development. This system is a merger between HWRFX and the operational HWRF system. The current work will provide some basis for further evaluations of this system.

Acknowledgments:

The authors acknowledge funding from NOAA’s Hurricane Forecast Improvement Project that supported this work. The model was primarily developed at the NCEP. The first author wishes to acknowledge several of the scientists at NCEP for their help during the development of this modeling system. The authors wish to thank Drs. Robert Rogers and Tomi Vukicevic for providing useful comments on improving the original version of this manuscript.

References

Aberson, S. D., 1998: Five-day tropical cyclone track forecasts in the North Atlantic basin. *Wea. Forecasting,* **13**, 1005-1015.

Bender, M.A., I. Ginis, R.E. Tuleya, B. Thomas, and T. Marchok, 2007: The operational GFDL coupled hurricane-ocean prediction system and a summary of its performance. *Mon. Wea. Rev.,* **135**, 3965-3965

Davis, C., W. Wang, S. Chen, Y. Chen, K. Corbosiero, M. DeMaria, J. Dudhia, G. Holland, J. Klemp, J. Michalakes, H. Reeves, R. Rotunno, Q. Xiao, S. Chen, K. Corbosiero, M. DeMaria, and H. Reeves, 2008: Prediction of landfalling hurricanes with the advanced hurricane WRF model. *Mon. Wea. Rev*., vol. **136**, p. 1990-2005.

Davis, C.A., Wei Wang Cavallo, S. Done, J.M. Dudhia, J. Fredrick, S. Michalakes, J. Caldwell, G. Engel, T. Torn, 2011: High-Resolution  Hurricane Forecasts. *Computing in Science and Engineering,* January/February, 2011, 22-30.

DeMaria, M., and J. Kaplan, 1999: An updated statistical hurricane intensity prediction scheme for the Atlantic and eastern North Pacific basins. *Wea. Forecasting*, **14**, 326-337.

DeMaria, M., Kaplan, J., Knaff, J. A., Mainelli, M., Shay, L. K., 2005: Further improvements to the Statistical Hurricane Intensity Prediction Scheme (SHIPS), *Wea. Forecasting*, vol. **20**, 531-543.

DeMaria, M., J. A. Knaff, and J. Kaplan, 2006: On the decay of tropical cyclone winds crossing narrow landmasses, *J. Appl. Meteor.*, **45**, 491-499.

Ek, M. B., K. E. Mitchell, Y. Lin, P. Grunmann, E. Rogers, G. Gayno, and V. Koren, 2003: Implementation of the upgraded Noah land-surface model in the NCEP operational

mesoscale Eta model. *J. Geophys. Res.,* **108,** 8851, doi:10.1029/2002JD003296.

Ferrier, B. S., Y. Lin, T. Black, E. Rogers, and G. DiMego, 2002: Implementation of a new grid-scale cloud and precipitation scheme in the NCEP Eta model. Preprints, *15th Conference on Numerical Weather Prediction,* San Antonio, TX, Amer. Meteor. Soc., p. 280-283.

Fierro, A.O., R.F. Rogers, F.D. Marks, and D.S. Nolan, 2009: The impact of horizontal grid spacing on the microphysical and kinematic structures of strong tropical cyclones simulated with the WRF-ARW model. Mon. Wea. Rev., **137**, 3717-3743.

Franklin, J. L., 2010: 2009 National Hurricane Center verification Report. <http://www.nhc.noaa.gov/verification/pdfs/Verification_2009.pdf>

Gallus, W. A., Jr., and J. F. Bresch, 2006:  Comparison of impacts of WRF dynamic core, physics package, and initial conditions on warm season rainfall forecasts.  *Mon. Wea. Rev*., **134**, 2632-2641.

Ginis, I, W. Shen, and M. A. Bender, 1999: Performance evaluation of the GFDL coupled hurricane ocean prediction system in the Atlantic basin. Preprints, *23d Conf. on Hurricanes and Tropical Meteorology,* Dallas, TX, Amer. Meteor. Soc., 607–610.

Gopalakrishnan, S.G., D.P. Bacon, N. Ahmad, Z. Boybeyi, T.J. Dunn, M.S. Hall, Y. Jin, P.C.S. Lee, R.V. Madala, R.A. Sarma, M.D. Turner, and T. Wait, 2002: An operational multi-scale atmospheric model with grid adaptivity for hurricane rorecasting. *Mon. Wea. Rev.*, **130** (7), 1830-1847.

Gopalakrishnan, S.G., N. Surgi, R. Tuleya, and Z. Janjic, 2006 : "NCEP's Two-way-Interactive-Moving-Nest NMM-WRF modeling system for Hurricane Forecasting", *27th Conference on Hurricanes and Tropical Meteorology*, 24-28 April 2006, Monterey, California.

Gopalakrishnan, S., Q. Liu, T. Marchok, D. Sheinin, N. Surgi, R. Tuleya, R. Yablonsky, and X.

Zhang, 2010: Hurricane Weather and Research and Forecasting (HWRF) Model scientific documentation. L. Bernardet, Ed., 75 pp.

Gopalakrishnan, S. G., F. D. Marks, X. Zhang, J.-W. Bao, K.-S. Yeh, and R. Atlas, 2011: The Experimental HWRF System: A Study on the Influence of Horizontal Resolution on the Structure and Intensity Changes in Tropical Cyclones using an Idealized Framework. *Mon. Wea. Rev*. (in press)

Gross, J. M., 1999: 1998 National Hurricane Center forecast verification. Minutes *of the 53d Interdepartmental Hurricane Conf.,* Biloxi, MS, Office of the Federal Coordinator for Meteorological Services and Supporting Research, B24–B63. [Available from Office of the Federal Coordinator for Meteorological Services and Supporting Research, 8455 Colesville Rd., Suite 1500, Silver Spring, MD 20910.]

Hong, S.-Y., and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.*, vol. **124**, p. 2322-2339.

Hong, S.-Y., and H.-L. Pan, 1998: Convective Trigger Function for a Mass-Flux Cumulus Parameterization Scheme. *Mon. Wea. Rev.*, vol. **129**, p. 1164-1178.

Janjic, Z. I., 1990: The step–mountain coordinates: physical package. *Mon. Wea. Rev*, **118**, 1429–1443.

Janjic, Z.I., J.P. Gerrity Jr., and S. Nickovic, 2001: An alternative approach to non-hydrostatic modeling. *Mon. Wea. Rev*., **129** (5), 1164-1178.

Janjic, Z.I., 2003: A non-hydrostatic model based on a new approach. *Meteor. Atmos. Physics*, **82**, 271-285.

Jarvinen, B. R., and C. J. Neumann, 1979: Statistical forecasts of tropical cyclone intensity for the North Atlantic basin. NOAA Tech. Memo. NWS NHC-10, 22 pp.

Kain, J. S., and J. M. Fritsch, 1992: The role of convective “trigger function” in numerical forecasts of mesoscale convective systems. *Meteor. Atmos. Phys.,* **49,** 93–106.

Kaplan, J., M. DeMaria, and J.A. Knaff, 2010: A revised tropical cyclone rapid intensification index for the Atlantic and eastern North Pacific basins. Wea. Forecast., **25**, 220-241.

Knaff, J.A., M. DeMaria, B. Sampson, and J.M. Gross, 2003: Statistical, five-day tropical cyclone intensity forecasts derived from climatology and persistence. *Wea. Forecasting*, **18**, 80-92.

Kurihara, Y., and [R.E. Tuleya](http://www.gfdl.noaa.gov/bibliography/resultstest.php?author=1101), 1974: S**tructure of a tropical cyclone developed in a three-dimensional numerical simulation model**.*J. Atmos. Sci.*, **31 (4)**, 893-919.

Kurihara, Y., 1975: Budget analysis of a tropical cyclone simulated in an axisymmetric numerical model. *J. Atmos. Sci*., **32**, 25-59.

Kurihara, Y., M. A. Bender, R.E. Tuleya and R. J. Ross, 1995: Improvements in the GFDL hurricane prediction system. *Mon. Wea. Rev*., **123**, 2791-2801.

Li, X. and Z. Pu, 2009: Sensitivity of numerical simulation of early rapid intensification of Hurricane Emily (2005) to cumulus parameterization schemes in different model horizontal resolutions, *J. Meteor. Soc. Japan.*, **87**, 403-421.

Liu, Q., N. Surgi , S. Lord, W.-S. Wu, S. Parrish, S. Gopalakrishnan, J. Waldrop and J. Gamache, 2006: Hurricane Initialization in HWRF Model. Preprints, *27th Conference on Hurricanes and Tropical Meteorology,* Monterey, CA*.*

Liu, Y., D.-L. Zhang and M.K. Yau, 1997: A multiscale numerical study of Hurricane Andrew (1992). Part I: Explicit simulation and verification. *Mon. Wea. Rev*., **125**, 3073-3093.

Liu, Y., D.-L. Zhang and M.K. Yau, 1999: A multiscale numerical study of Hurricane Andrew (1992). Part II: Kinematics and inner-core structures. *Mon. Wea. Rev*., **127**, 2597-2616.

Marks, F.D., and R.A. Houze, 1987: Inner core structure of Hurricane Alicia from airborne Doppler radar observations. *J. Atmos. Sci.,* **44**, 1296-1317.

Marks, F.D., R.A. Houze, and J.F. Gamache, 1992: Dual-aircraft investigation of the inner core of the Hurricane Norbert. Part I: Kinematic structure. *J. Atmos. Sci.,* **49**,919–942.

Neumann, C. B., 1972: An alternate to the HURRAN (hurricane analog) tropical cyclone forecast system. NOAA Tech. Memo. NWS SR-62, 24 pp.

Noh, Y., W. G. Cheon, S.-Y. Hong, and S. Raasch, 2003: Improvement of the K-profile model for the planetary boundary layer based on large eddy simulation data. *Bound.-Layer Meteor.,***107,** 401–427.

Powell, M. D., S. H. Houston, L. R. Amat, and N Morisseau-Leroy, 1998: The HRD real-time hurricane wind analysis system. *J. Wind Engineer. and Indust. Aerodyn.*, vol. **77-78**, p. 53-64.

Riehl, H. 1954. *Tropical meteorology*. McGraw-Hill Book Company. New York. 392 p.

Rotunno, R., T. Chen, W. Wang, C. Davis, J. Dudhia and G.J. Holland, 2009: Large-eddy Simulation of an Idealized tropical Cyclone. *Bull. Amer. Meteor. Soc.*, vol. **90**, p. 1783-1788.

Tuleya, R.E., and Y. Kurihara, 1975: The energy and angular momentum budgets of a three-dimensional tropical cyclone model. J. Atmos. Sci., **32**, 287-301.

Tuleya, R.E., and [Y. Kurihara](http://www.gfdl.noaa.gov/bibliography/resultstest.php?author=1061), 1978: **A numerical simulation of the landfall of tropical cyclones**. J. Atmos. Sci., **35**, 242-257.

Tuleya, R.E., and [Y. Kurihara](http://www.gfdl.noaa.gov/bibliography/resultstest.php?author=1061), 1982: **A note on the sea surface temperature sensitivity of a numerical model of tropical storm genesis**. Mon. Wea. Rev, **110**, 2063-2069.

Willoughby, H.E., 1979: Forced secondary circulations in hurricanes. *J. Geophys. Res*., **84**, 3173-3183.

Willoughby, H.E., 1990a: Temporal changes of the primary circulation in tropical cyclones. *J. Atmos. Sci.*, **47**, 242-264.

Willoughby, H.E., 1990b: Gradient balance in tropical cyclones. *J. Atmos. Sci.,* **47**, 265-274.

Yau, M.K., Y. Liu, D.-L. Zhang, and Yongsheng Chen, 2004: A multiscale numerical study of Hurricane Andrew (1992). Part VI: Small-scale inner-core structures and wind streaks. *Mon. Wea. Rev*., **132**, 1410-1433.

Yeh, K.-S., X. Zhang, S. Gopalakrishnan, S. Aberson, R. Rogers, F.D. Marks and R. Atlas, 2011: Performance of the experimental HWRF in the 2008 hurricane season. J. Natural Hazards (accepted for publication.)

Zhang, D.-L., Y. Liu and M.K. Yau, 2002: A multiscale numerical study of Hurricane Andrew ( 1992). Part V: Inner-core thermodynamics. *Mon. Wea. Rev*., **130,** 2745-2763.

Zhang, D.-L., Y. Liu and M.K. Yau, 1999: Surface winds at landfall of Hurricane Andrew (1992) - A reply. *Mon. Wea. Rev*., **127**, 1711-1721.

Zhang, D.-L., Y. Liu and M.K. Yau, 2000: A multiscale numerical study of Hurricane Andrew (1992). Part III: Dynamically-induced vertical motion. *Mon. Wea. Rev*., **128**, 3772-3788.

Zhang, D.-L., Y. Liu, and M.K. Yau, 2001: A multiscale numerical study of Hurricane Andrew (1992). Part IV: Unbalanced flows. Mon. Wea. Rev., **129,** 92-107.

Zhang, X, T.S. Quirino, S. Gopalakrishnan, K.S. Yeh, F.D. Marks, Jr.,and S.B. Goldenberg, 2011: HWRFX: Improving Hurricane Forecast with High Resolution  Modeling. *Computing in Science and Engineering,* January/February, 13-21.

27 km

9 km

9 km

3 km

GFDL/

HWRF

IC &

GFS BCs

GFDL/

HWRF

IC &

GFS BCs

500

500

500

Figure 1: Versions of the two domain configurations used in this study: (a) Parent domain at a resolution of 27 km with a moving nest at 9 km and (b) parent domain at a resolution of 9 km with a moving nest at 3 km. The approximate size of the domain is shown as they appear in a rotated latitude-longitude grid system (i.e., without any projection to the actual latitude longitude grid)

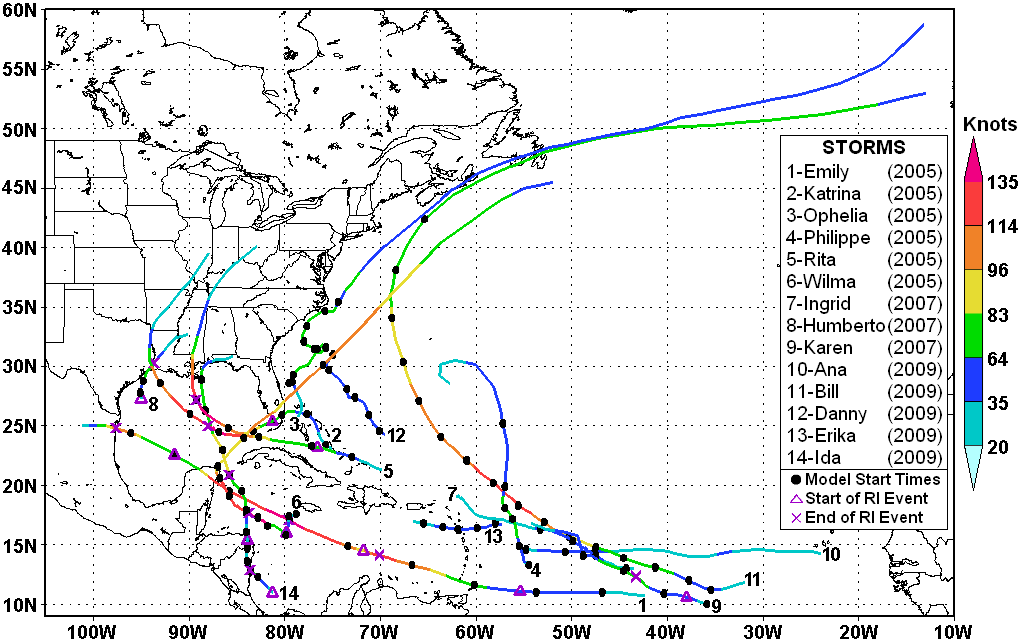


Figure 2: Tracks of the 14 tropical cyclones used in this study. Model initialization times (see Table 2) for the 87 cases are indicated by solid dots. Maximum surface (10 m) wind speeds are color coded. (Speeds on legend are in knots.) Beginning and ending points of rapid intensification (RI) events are indicated by triangles and X’s, respectively.

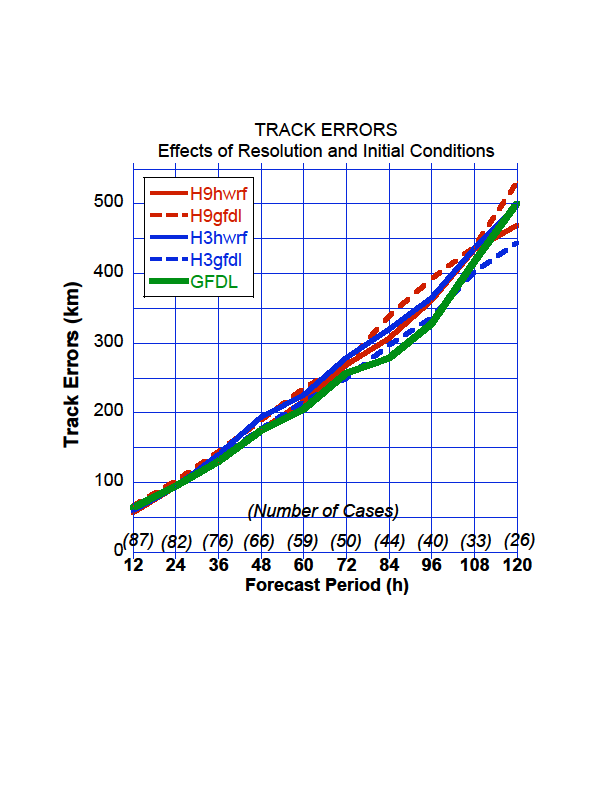
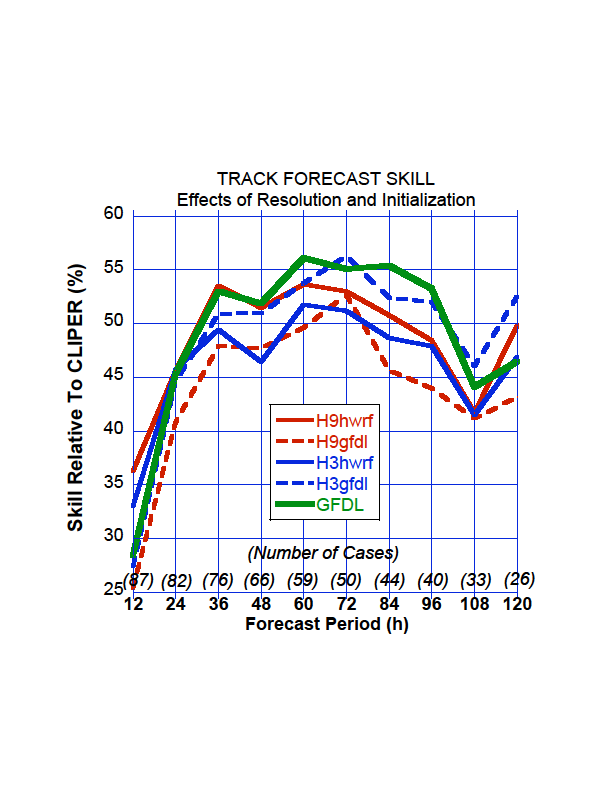
 

Figure 3: Verification of HWRFX track errors: (a) Absolute track errors and (b) Skill relative to CLIPER. H9hwrf refer to results from nested domain at 09-km-resolution with HWRF initial conditions, H9gfdl refer to results from nested domain at 09-km-resolution with GFDL initial conditions, H3hwrf refer to results from nested domain at 03-km-resolution with HWRF initial conditions, H3gfdl refer to results from nested domain at 03-km-resolution with GFDL initial conditions.

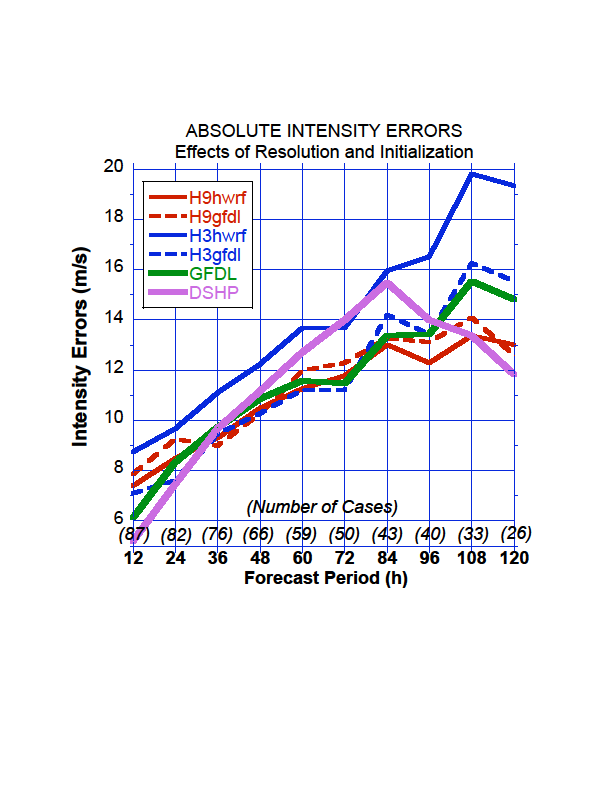
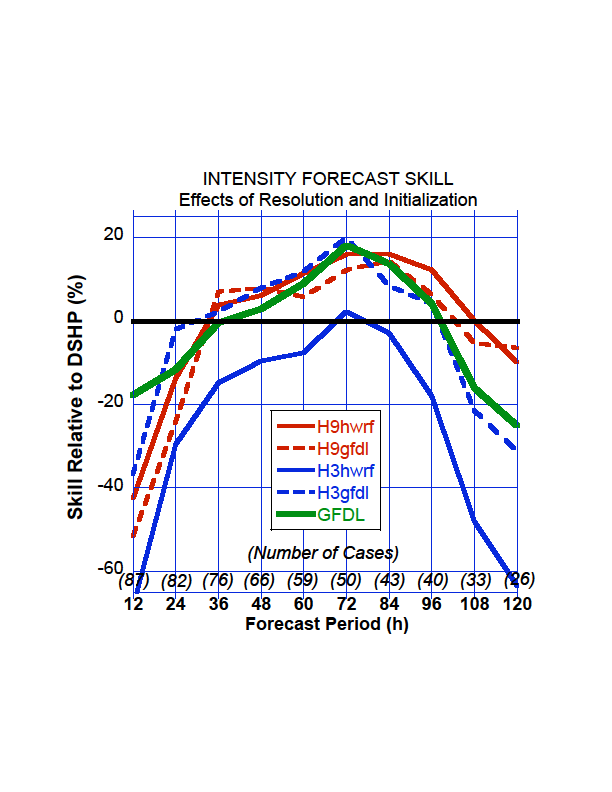
 

Figure 4: Verification of HWRFX intensity errors: (a) Absolute intensity errors and (b) Skill relative to DSHIPS. H9hwrf refer to results from nested domain at 09-km-resolution with HWRF initial conditions, H9gfdl refer to results from nested domain at 09-km-resolution with GFDL initial conditions, H3hwrf refer to results from nested domain at 03-km-resolution with HWRF initial conditions, H3gfdl refer to results from nested domain at 03-km-resolution with GFDL initial conditions.

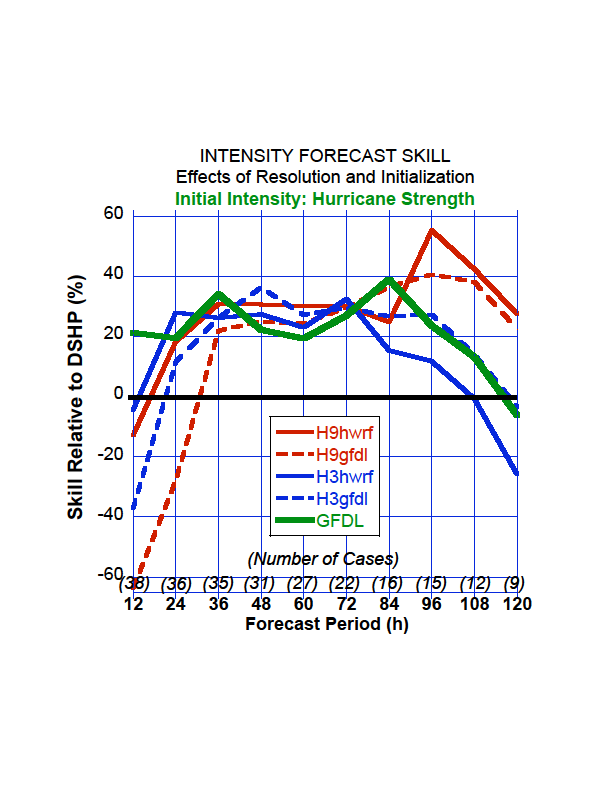
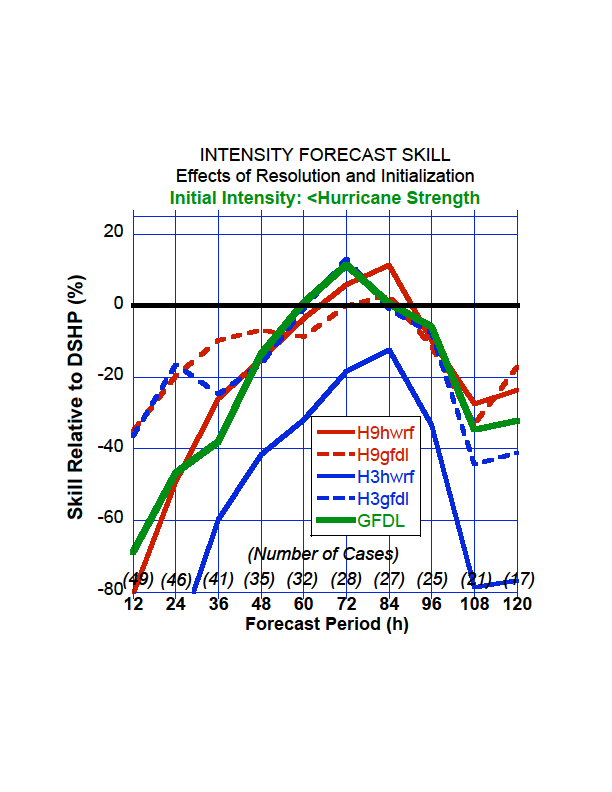
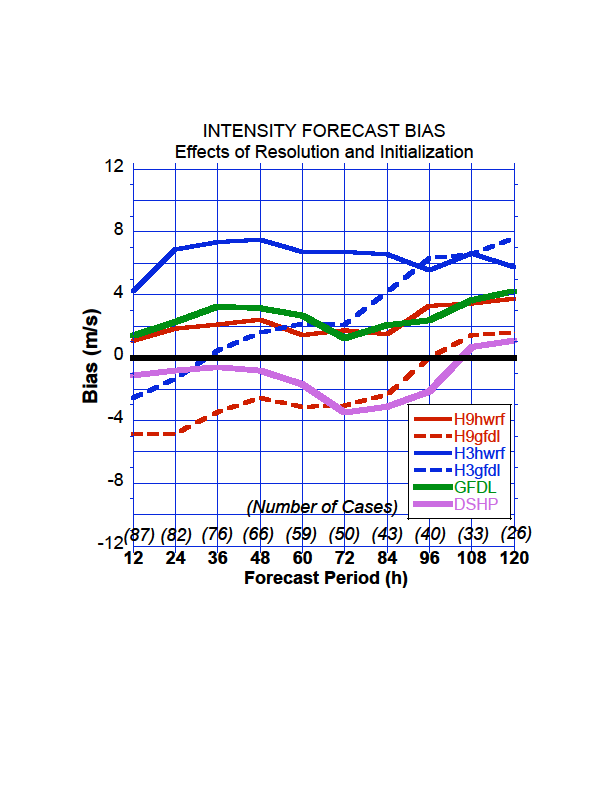
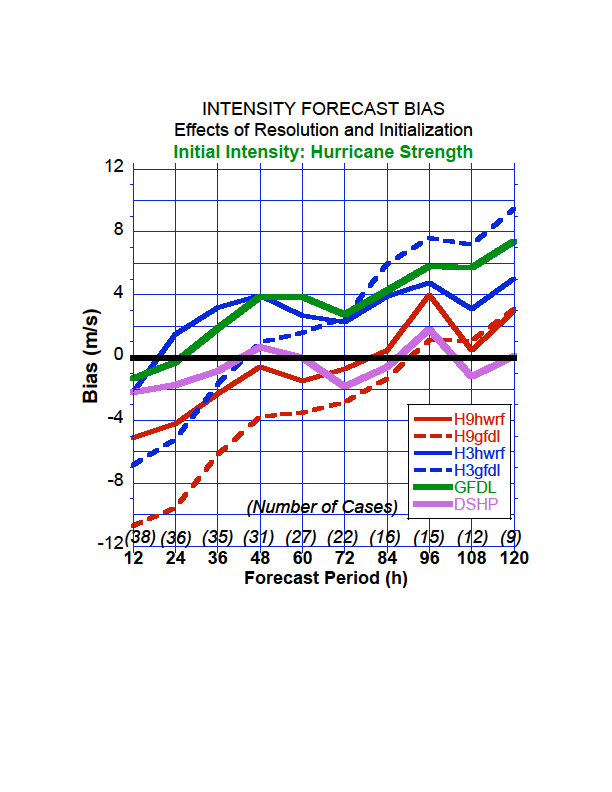
 

Figure 5: Skill relative to DSHIPS for (a) storms with initial intensity (maximum sustained surface wind speed) ≥ 33.4 ms-1 (i.e., hurricane strength) and (b) storms with initial intensity < 33.4 ms-1 (i.e., less than hurricane intensity). See Table 2 for (a) and (b) cases.





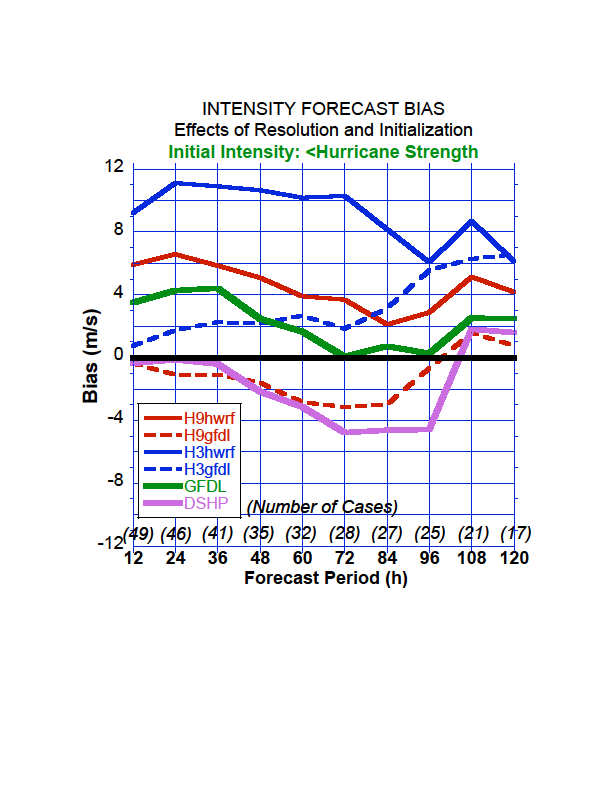


Figure 6: intensity bias for the models (four versions of HWRFX, GFDL and DSHIPS). Figures 5a, b and c show the biases for all the runs (87 cases at 12 h), initially strong storms (38 cases at 12 h) and initially weak storms (49 cases at 12 h), respectively

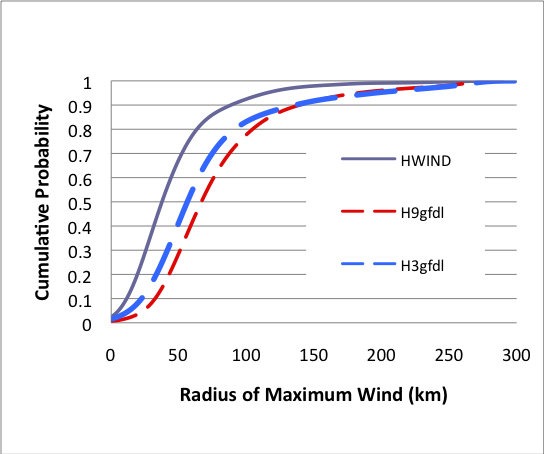
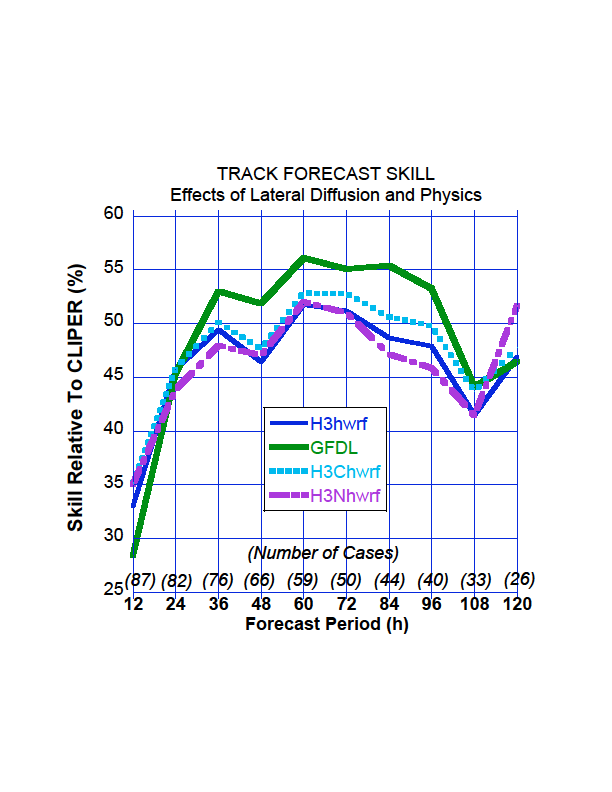
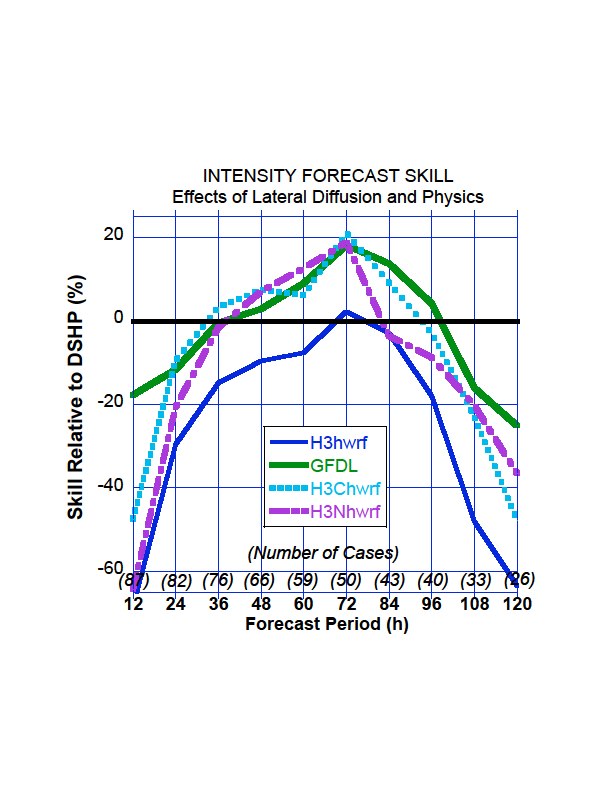


Figure 7: Cumulative probability distribution of the radius of maximum wind at 10 meters above the ground for the HWRFX forecasts, compared with the H\*WIND analysis using (a) GFDL initial conditions for the 3-km-run (H3gfdl) and the 9-km-run (H9gfdl) and (b) HWRF initial conditions for the 3-km-run (H3gfdl) and the 9-km-run (H9gfdl)

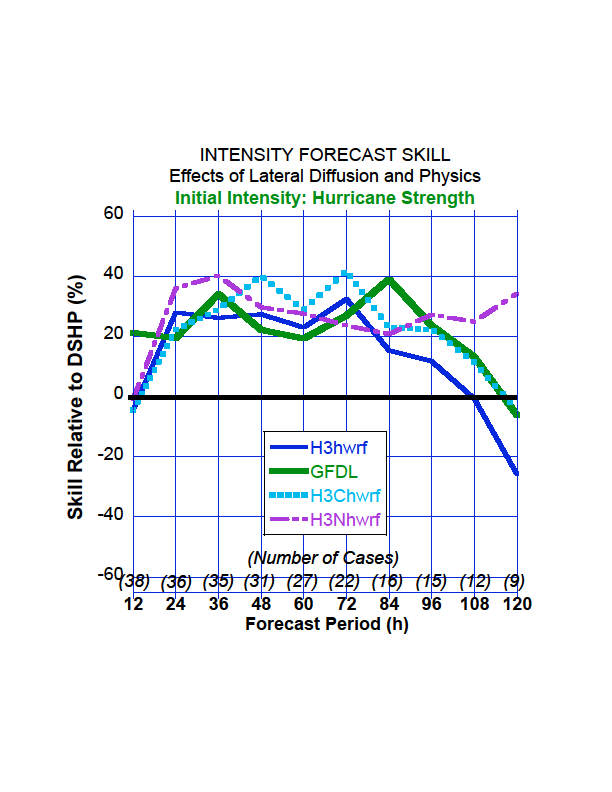
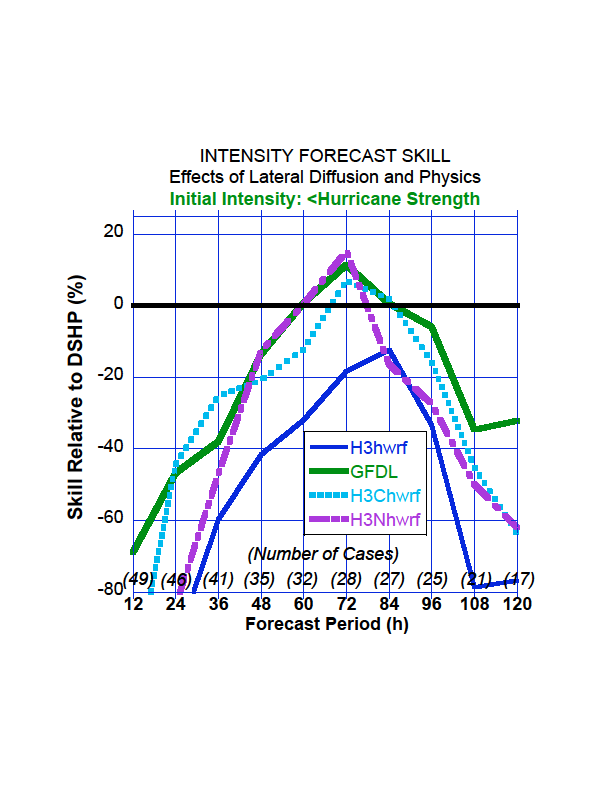
 

Figure 8: Verification of HWRFX (a) track errors: skill relative to CLIPER and intensity errors: skill relative to DSHIPS for (b) all storms, (c) storms with initial intensity (maximum sustained surface wind speed) ≥ 33.4 ms-1 (i.e., hurricane strength) and (d) storms with initial intensity < 33.4 ms-1 (i.e., less than hurricane intensity). Effect of “YSU-WSM5-KF” physics package (H3Nhwrf) is shown in pink and enhanced horizontal diffusion on the 3 km HWRFX runs with HWRF initial condition (H3Chwrf) is shown in cyan.

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **# Cases** | **Dates** | **RI Events** |
| Emily, 2005 | 8 | **071400, 071500, 071600, 071700, 071800, 071900, 072000** | 071306-071512, 071518-071700, 071900-072012 |
| Katrina, 2005 | 6 | 082400, 082500, **082600, 082700, 082800, 082900** | 082606-082900 |
| Ophelia, 2005 | 9 | 090712, 090812, 090912, **091012**, 091212, 091312, **091412, 091512**, 091612 |  |
| Philippe, 2005 | 6 | 091712, 091812, **091912**, 092012, 092112, 092212 |  |
| Rita, 2005 | 5 | 091900, 092000, **092100, 092300, 092400** | 092000-092206 |
| Wilma, 2005 | 7 | 101600, 101700, 101800, **101900, 102200, 102300, 102400** | 101718-101918 |
| Humberto, 2007 | 2 | 091212, 091300 | 091206-091312 |
| Ingrid, 2007 | 4 | 091212, 091312, 091412, 091512 |  |
| Karen, 2007 | 3 | 092500, 092700, 092800 | 092512-092618 |
| Ana, 2009 | 2 | 081512, 081600 |  |
| Bill, 2009 | 15 | 081600, 081612, 081700, **081712, 081800, 081812, 081900, 081912, 082000, 082012, 082100, 082112, 082200, 082212, 082300** |  |
| Danny, 2009 | 5 | 082612, 082700, 082712, 082800, 082812 |  |
| Erika, 2009 | 4 | 090200, 090212, 090300, 090312 |  |
| Ida, 2009 | 11 | 110500, **110512,** 110600, 110612, 110700, 110712, **110800, 110812, 110900,** 110912, 111000 | 110406-110512, 110618-110812 |

Table 1: Summary of the 87 TC cases used in this study. Cases in bold are for initial intensity (maximum surface wind speed) ≥33.4 ms-1 (i.e., hurricane strength). Rapid Intensification (RI) events are defined as ≥ 15.4 ms-1 (i.e., ≥30 kts) increase in maximum surface wind speed in 24 h (Kaplan et al. 2010). See Figure 2 to see the actual tracks and locations of the model initialization times and RI events.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Convection | Microphysics | Boundary and surface layer | Land Surface &  Ocean Coupling | Dissipative  heating | Radiation |
| Operational  HWRF | SAS with momentum mixing | Ferrier | GFS boundary layer and GFDL hurricane model surface layer scheme (uses modification  to surface  roughness over ocean) | GFDL Slab  &  Princeton Ocean model coupling | Switched on | GFDL  radiation scheme |
| HWRFX  (Low resolution forecasts) | SAS without momentum mixing | Ferrier | GFS boundary layer and  GFDL hurricane  model surface layer scheme  (uses the original Charnock’s approximation for surface roughness over ocean) | NOAH LSM  &  Uncoupled/Initially presecribed SST | Switched  off | NCAR |

Table 2: List of differences between operational HWRF and HWRFX physics for the configuration at 27:09

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| No. | Initial Conditions | Grid configuration | Physics | Lateral Diffusion Coefficient in the parent & nested domain | Specification |
| 1  2  3  4  5  6 | HWRF  GFDL  HWRF  GFDL  HWRF  GFDL | 27:09  27:09  09:03  09:03  09:03  09:03 | HWRFX  HWRFX  HWRFX  HWRFX  HWRFX  NCAR | 1.6 & 0.7  1.6 & 0.7  1.6 & 0.7  1.6 & 0.7  1.6 & 5.0  1.6 & 0.7 | H9hwrf  H9gfdl  H3hwrf  H3gfdl  H3Chwrf  H3Nhwrf |

Table 3: List of experiments performed, each containing 87 cases

1. Corresponding author address: S.G. Gopalakrishnan, AOML/HRD, 4301 Rickenbacker Causeway, Miami, Florida 33149 <E-mail: gopal@noaa.gov [↑](#footnote-ref-1)
2. For this study, cases from Felix (2007) could not be included in the homogeneous sample since all of the necessary initial conditions were not available for that storm. [↑](#footnote-ref-2)
3. The “5” refers to the fact that the original CLIPER (developed in 1972) and SHIFOR (developed in 1979) models only provided forecasts out to 72 h. These forecasts were extended to 120 h (five days) in 1998 and 2003, respectively. [↑](#footnote-ref-3)
4. These comparisons are not considered completely “fair” judges of the performance of the HWRFX (or GFDL) models in the actual operational environment since the HWRFX and GFDL versions used here are considered “late” models, i.e., they are available late in the operational forecast cycle, whereas the CLIPER5 and DSHIPS are “early” models, in that they are available to the forecasters much sooner (Franklin 2010). [↑](#footnote-ref-4)
5. The model RMW’s were obtained from the hourly model output of the azimuthally-averaged winds. [↑](#footnote-ref-5)