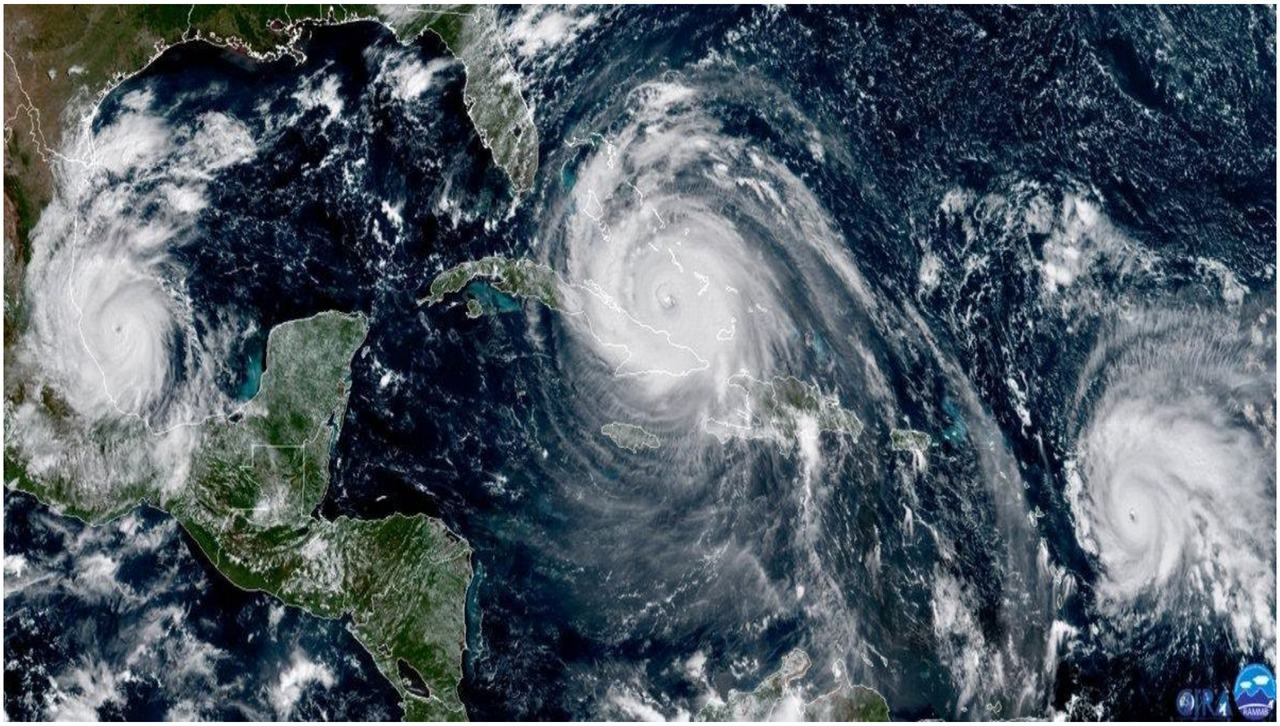
NATIONAL OCEANIC AND
ATMOSPHERIC ADMINISTRATION
United States Department of Commerce

Hurricane Forecast Improvement Program
Five-Year Plan: 2019-2024

Proposed Framework for Addressing Section 104 of the Weather Research Forecasting Innovation Act of 2017

22 June 2018
Updated 25 June 2019



Cover: GOES 16 image of hurricanes Katia (left), Irma (center), and Jose (right) captured on 8 September 2017. (NOAA)

National Oceanic and Atmospheric Administration
Hurricane Forecast Improvement Project
Years Ten to Fifteen Strategic Plan

Frank Marks[1], Nicole Kurkowski[2], Mark DeMaria[3], Michael Brennan[3]

22 June 2018
Updated 25 June 2019

Strategic Plan Writing Team

Mark DeMaria (NWS/NHC)	Michael Brennan (NWS/NHC)	Jamie Rhome (NWS/NHC)
Sundararaman Gopalakrishnan (OAR/AOML)	Rob Rogers (OAR/AOML)	Jason Sippel (OAR/AOML)
Vijay Tallapragada (NWS/EMC)	Avichal Mehra (NWS/EMC)	Jennifer Sprague (NWS)
Morris Bender (OAR/GFDL)	Tim Marchok (OAR/GFDL)	Alan Gerard (OAR/NSSL)
Pablo Santos (NWS/Miami)	James Nelson (NWS/WPC)	Cody Fritz (NWS/NHC)
Ed Mifflin (NWS/STI)	**Nicole Kurkowski** (NWS/STI)	**Frank Marks** (OAR/AOML)
James Franklin (NWS/NHC)	Nysheema Lett (NWS/STI)	

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

Table of Contents

Executive Summary

The Hurricane Forecast Improvement Project (HFIP) was established within the National Oceanic and Atmospheric Administration (NOAA) in June 2007 in response to a number of deadly and damaging hurricanes in 2004-2005 (i.e., Charley, Frances, Ivan, Jeanne, 2004; Katrina, Rita, Wilma, 2005). HFIP's efforts to reduce track and intensity forecast errors by 20% within 5 years and 50% within 10 years, and to extend forecasts out to 7 days, resulted in significant improvements in forecast performance over historical values (e.g., the improvement rate of intensity forecast error prior to HFIP was <1% per year). Those advancements required major investments in enhanced observational strategies to make best use of observations (oceanographic, aircraft, and satellite), improved data assimilation, numerical modeling systems, and expanded forecast applications based on high-resolution and ensemble-based numerical prediction systems and an improved computational infrastructure.

Despite these investments, and the advancements in hurricane forecasting resulting from them, hurricanes continue to cause significant impacts to the United States (U.S.) coastline as both the population near the coast (50% of the U.S. population lives within 50 miles of the coast) and coastal infrastructure and economic activity (estimated at $3 trillion) continue to rise. Hurricanes Harvey, Irma, and Maria in 2017 resulted in $265 billion in damage. Over the 38-year period from 1980-2017 Hurricane Harvey was the second costliest hurricane in damage (>$125 billion), while Hurricane Maria was the third costliest (~$90 billion), and Hurricane Irma the fifth costliest (~$50 billion)[1]. The large loss of human life from hurricanes like Katrina, Sandy and Maria indicates that hundreds or even thousands of deaths are still possible from the direct and/or indirect effects of tropical cyclones.

Impacts such as those discussed above can be significantly reduced in the future by providing improved forecasts and warnings and more actionable environmental intelligence[2], i.e., increasing forecast accuracy of tropical cyclone (TC) track, rapid intensification (RI) and extreme weather events (e.g., total water prediction). These are the goals of the next phase of HFIP. Further enhancements, such as providing a more accurate and reliable guidance for hurricane and storm surge watches and warnings at longer lead times, will allow greater lead times for pre-storm mitigation efforts. Such efforts include the protection of life and property and an overall reduction in the economic impact of damaging storms, potentially saving hundreds of lives and billions of dollars[3].

---

[1] Table 3a NOAA Technical Memorandum NWS NHC-6: The Deadliest, Costliest, and Most Intense United States Tropical Cyclones from 1851 to 2010 (2017 update https://www.nhc.noaa.gov/news/UpdatedCostliest.pdf)

[2] NWS Weather Ready Nation policy, 2016 (http://www.legislative.noaa.gov/policybriefs/NWS%20101%20Feb%202016.pdf).

[3] https://ascelibrary.org/doi/10.1061/%28ASCE%291527-6988%282007%298%3A3%2845%29

HFIP's approach is designed to accelerate the development and implementation of promising technologies and techniques from the research community into operations. That approach resulted in a 20% reduction in errors from numerical guidance for both storm track and intensity within the first 5 years of the project. In addition, to better communicate risks to the public and the emergency management community, HFIP invested in social science research that contributed to the development and implementation of operational storm surge products such as the Potential Storm Surge Flooding Map.

HFIP provides the unifying organizational infrastructure and funding for NOAA and other agencies to coordinate the TC research needed to significantly improve guidance for TC track, intensity, storm surge and precipitation forecasts and to accelerate the transition of research to operations. HFIP focuses multi-organizational activities to research, develop, demonstrate, and implement enhanced operational modeling capabilities, dramatically improving the numerical forecast guidance made available to the National Hurricane Center (NHC), which is the agency that issues the "official" U.S. TC forecast. Furthermore, through public-private partnerships, HFIP facilitates the development of the next generation of TC researchers for NOAA. These benefits will continue under the next generation of HFIP.

Tracking the progress of numerical TC prediction guidance during HFIP identified development areas capable of delivering the most significant results (e.g., data assimilation and improvements to physics parameterization schemes), areas into which HFIP's focus should be extended (e.g., storm size[4], maximum intensity, genesis, storm surge, rainfall), and the need for more meaningful ways to measure progress for high-impact hazards (e.g., storm surge, flooding, and tornadoes). Forecast challenges during the 2017 hurricane season underscored the need to address issues such as RI, genesis, storm size, and rainfall, as demonstrated by Hurricane Harvey. Hurricane Irma highlighted a need for improvements in forecasting track, storm size, and storm surge as even relatively small track uncertainties can have significant impacts on preparedness actions (e.g., FL east coast vs west coast evacuation effects). Hurricane Maria, which devastated Dominica and Puerto Rico, again pointed to a need to better predict RI, as did Hurricanes Matthew (2016), Joaquin (2015), and Patricia (2015).

To improve TC forecasting with the goal of developing and extending accurate TC forecasts in order to reduce loss of life, injury, and economic damage, the next generation of HFIP will focus on:

1. improving the prediction of RI and track of TCs;
2. improving the forecast and communication of surges from TCs; and
3. incorporating risk communication research to create more effective watch and warning products.

---

[4] defined as radii of maximum wind and 34-knot winds.

This plan details the specific research, development, and technology-transfer activities necessary to address HFIP's science and Research to Operations (R2O) challenges, which are articulated through the three primary focus areas as stated in Section 104 of the *Weather Research and Forecasting Innovation Act of 2017.*

In order to address the three primary focus areas outlined above, HFIP developed a set of specific goals and metrics to improve the accuracy and reliability of TC forecasts and warnings and increase the confidence in those forecasts to enhance mitigation and preparedness decisions by emergency management officials at all levels of government and by individuals.

To address all three areas, improved model guidance for TC formation, track, intensity and size will be essential. Basic TC forecast parameters will be improved, including the position, maximum wind (i.e., intensity), and storm size. Estimates of the uncertainty of those parameters will also be enhanced, enabling better risk communication to end users through accurate probabilistic information. RI remains an especially important and challenging forecast problem. Specific goals and metrics are defined for the prediction of the basic TC forecast parameters, new extended range forecasts, RI, and TC formation.

The next phase of HFIP will build upon the original goals of the project[5] through the following specific goals and metrics in order to address HFIP's science and R2O challenges, as well as Section 104 of *Weather Research and Forecasting Innovation Act of 2017*:

1. Reduce forecast guidance errors[6], including during RI, by 50% from 2017.
2. Produce seven-day forecast guidance as good as the 2017 five-day forecast guidance;
3. Improve guidance on pre-formation disturbances, including genesis timing, and track and intensity forecasts by 20% from 2017;
4. Improve hazard guidance and risk communication, based on social and behavioral science, to modernize the TC product suite (products, information, and services) for actionable lead-times for storm surge and all other threats.

Six key strategies are described that will continue improving track and intensity guidance, particularly for RI out to actionable lead times up to seven days, and extend the focus to improving guidance on storm size, storm surge, and all other TC hazards at actionable lead times up to three days. Improved hazard guidance will result from development of a TC analysis and forecast system producing ensemble guidance enabling probabilistic hazard products and improved track, intensity, and storm size predictions before formation and throughout the storm's life cycle. Using social and behavioral science research, HFIP will design a more effective TC product suite to better communicate risk and apply it to all current TC hazards products. Success requires increased resources for development, R2O, and infrastructure

---

[5] http://www.hfip.org/

[6] Percent improvement is determined by evaluating track, intensity, storm size, and RI error relative to those over the 3-year period 2015-2017.

(including high-performance computing). Additional resources will also enable re-engagement with the larger academic community in addressing the challenges and goals stated herein.

1. Introduction

The Hurricane Forecast Improvement Project (HFIP) was established within the National Oceanic and Atmospheric Administration (NOAA) in 2007 in response to sixteen named tropical storms--including ten hurricanes (e.g., Charley, Ivan, Frances, Jeanne, 2004; Wilma, Dennis, Katrina, Rita, Wilma, 2005) crossing US coastlines from 2004-2005. HFIP has been focused specifically on improving guidance for hurricane track, intensity, and storm surge forecasts. Efforts to reduce average track and intensity errors by 20% within five years and 50% in ten years for days one to five and to extend forecasts out to seven days resulted in significant improvements above the historical 1% annual improvement of forecast storm intensities. These advancements required major investments in enhanced observational strategies, improved data assimilation, numerical modeling systems, expanded forecast applications based on high-resolution and ensemble-based numerical prediction systems, and improved computational infrastructure.

HFIP provides the unifying organizational infrastructure and resources for the NOAA and other agencies to coordinate the research needed to significantly improve guidance for hurricane track, intensity, and storm surge forecasts, and to accelerate the transition of this research to operations. HFIP focused multi-organizational activities to research, develop, demonstrate, and implement enhanced operational modeling capabilities, dramatically improving the numerical forecast guidance made available to the National Hurricane Center (NHC). Furthermore, through public-private partnerships, HFIP developed and facilitated the next generation of hurricane researchers for NOAA.

HFIP’s approach is designed to accelerate the transition of promising technologies and techniques from the research community into operations. To that end, HFIP successfully:

- improved data assimilation and modeling technologies (global- and hurricane-scale);
- incorporated aerial reconnaissance, including Doppler, flight-level, and dropsonde data in model initialization;
- aligned focused research efforts within NOAA and with interagency and academic partners;
- established a high-performance computing infrastructure and attendant protocols to support research-to-operations activities;
- defined and implemented a solution (real-time experimental forecast system) to accelerate research into operational products;
- established a process to take advantage of outside research capabilities in support of project objectives (Federally-funded grantees working with a community code repository);

- expanded use of the Hurricane Weather Research and Forecast (HWRF) model in all tropical cyclone (TC) basins to meet the needs of Department of Defense (DOD) partners; and
- integrated social science methodologies in the development of products conveying risk associated with storm surge.

This approach resulted in several important advances, including a 20% reduction in numerical model errors for TC track and intensity forecasts.

The first phase of HFIP shed light on development areas yielding the most significant results to improved track and intensity guidance, such as data assimilation and improvements to physics parameterization schemes. These areas will be further improved during the next phase of HFIP, while additional forecast guidance metrics, such as storm size and rainfall, will be explored. The following issues must also be addressed in order to achieve the original goals of HFIP:

- reduction of short-term (<36 hour) forecast error growth in HWRF, to include rapid intensification (RI) episodes;
- investigation of long-term (> 4 days) track forecast error growth in global models, limiting the ability to extend forecasts beyond 5 days;
- better use of high-resolution ensembles for intensity to improve representation of forecast uncertainty in forecast products;
- better use of high-resolution ensemble information to improve model initialization;
- better use of satellite observations in cloudy/rainy regions surrounding TCs; and,
- identification of existing and new observations required to improve intensity forecasts at different lead times through a systematic evaluation of observation impacts using observing system experiments and observing system simulation experiments.

Tracking the progress of numerical TC prediction guidance during HFIP identified the development areas capable of delivering the most significant results (e.g., data assimilation and improvements to physics parameterization schemes), areas into which HFIP's focus should be extended (e.g., storm size[7], genesis, storm surge, rainfall), and the need for more meaningful ways to measure progress for high-impact events (e.g., those that include RI). Forecast challenges during the 2017 hurricane season underscored the need to address issues such as RI, genesis, storm size, and rainfall, as demonstrated by Hurricane Harvey. Hurricane Irma highlighted a need for improvements in forecasting track, storm size, and storm surge. Hurricane Maria, which devastated Dominica and Puerto Rico, again pointed to a need to better predict genesis and RI, as did Hurricanes Matthew (2016), Joaquin (2015), and Patricia (2015).

---

[7] defined as radii of maximum wind and 34-knot winds.

To improve TC forecasting with the goal of developing and extending accurate TC forecasts and warnings in order to reduce loss of life, injury, and damage to the economy, the next generation of HFIP will focus on:

1. improving the prediction of RI and track of TCs;
2. improving the forecast and communication of surges from TCs; and
3. incorporating risk communication research to create more effective watch and warning products.

This plan details the specific research, development, and technology transfer activities necessary to address HFIP's science and R2O challenges, which are articulated through the three primary focus areas as stated in Section 104 of the *Weather Research and Forecasting Innovation Act of 2017.*

2. Goals

In order to address the three primary focus areas outlined above, HFIP has developed a set of specific goals and metrics to improve the accuracy and reliability of TC forecasts and warnings and increase the confidence in those forecasts to enhance mitigation and preparedness decisions by emergency management officials at all levels of government and by individuals.

Improved model guidance for TC formation, track, intensity and size will be essential to address all three areas. Basic TC forecast parameters will be improved, including the formation time and location, position, maximum wind speed (i.e., intensity), and storm size. Estimates of the uncertainty of those parameters will also be enhanced, enabling better risk communication to end users through accurate probabilistic information (i.e., information that considers the likelihood, or probability, that an event will occur). Rapid intensification remains an especially important and challenging forecast problem. Specific goals and metrics are defined for the prediction of the basic TC forecast parameters, new extended range forecasts, rapid intensification, and TC formation.

The next generation of HFIP will build upon the original goals of the project[8] through the following specific goals and metrics:

---

[8] http://www.hfip.org/

1. Reduce track, intensity, and structure forecast guidance errors by 50% relative to a 2017 baseline[9]. Reduce intensity forecast guidance errors by 50%, relative to a 2017 baseline, for RI events.
2. Produce seven-day track and intensity forecast guidance as accurate as a 2017 five-day baseline.
3. Improve forecast guidance on pre-genesis disturbances, for track, intensity, and the timing of genesis, by 20% relative to a 2017 baseline.
4. Improve hazard guidance and risk communication for all of the TC hazards (wind, surge, rainfall, and tornadoes) at actionable lead times through the application of social and behavioral sciences, resulting in a modernized suite of TC products, information, and services.

Six key strategies are described that will continue improving track and intensity guidance, particularly for rapid intensification, out to actionable lead times up to seven days, and extend the focus to improving guidance on storm size, storm surge, and all other TC hazards at actionable lead times up to three days. As a framework for success in these efforts, HFIP intends to apply the "Forecasting a Continuum of Environmental Threats" (FACETs)[10] approach as an overarching strategy to guide some of this transition and modernization to NOAA TC hazard guidance. In the FACETs paradigm, a nearly continuous stream of high-resolution probabilistic hazard information, extending from days to within minutes of an event, will be driven by cutting-edge scientific tools and be optimized for user-specific decision making through the integrated application of social and behavioral sciences. While there are several components to FACETs, the generation and application of probabilistic hazard guidance are at the heart of the paradigm.

Improved hazard guidance will result from the development of a hurricane analysis and forecast system producing ensemble guidance, enabling the generation of probabilistic hazard products and improved track, intensity, and storm size predictions before formation and throughout the storm's life cycle. Using social and behavioral science research, HFIP will design a more effective TC product suite to better communicate risk and apply it to all current TC hazards products. Success requires increased resources for development, R2O, and infrastructure (including high-performance computing). Additional resources will also enable re-engagement with the larger academic community in addressing the challenges and goals stated herein.

3. Key Strategies

---

[9] Percent improvement is determined by evaluating track, intensity, storm size, and rapid intensification error relative to those over the 3-year period, 2015-2017.

[10] https://www.nssl.noaa.gov/projects/facets/

3.1 Advance the operational hurricane analysis and forecast system (HAFS)

The Hurricane Analysis and Forecasting System (HAFS) is NOAA's next-generation multi-scale numerical model and data assimilation package which will provide an operational analysis and forecast out to seven days, with reliable and skillful guidance on TC track and intensity (including rapid intensification), storm size, genesis, storm surge, rainfall and tornadoes associated with TCs. A key aspect of FACETs is the generation of calibrated probabilistic hazard guidance from numerical models and statistical analyses. The development of HAFS is the key strategy to generate this calibrated hazard guidance for the application of FACETs to all of the TC hazards.

One of the biggest successes achieved during the first phase of HFIP was the creation of the high-resolution HWRF system. The HWRF, storm following, nested grid modeling system was designed to operate at a horizontal resolution of 2 km or less desired for capturing tropical cyclone inner core processes as well as the interactions with the large scale processes, proven to be critical for improving track, intensity, rainfall and size predictions. It may be noted that HWRF has been providing only track and intensity guidance to forecasters until now. This system, with further advancements in ensembles, data assimilation techniques and better use of hurricane observations, will be the starting point for the first version of HAFS. As HWRF transitions to the Finite-Volume Cube Sphere (FV3) dynamical core, HAFS will evolve into advanced analysis and forecast system for cutting-edge research on modeling, physics, data assimilation, and coupling to earth-system components for high-resolution TC predictions within the Next Generation Global Prediction System (NGGPS) consistent with the Strategic Implementation Plan (SIP) objectives of creating the Unified Forecast System (UFS[11]).

3.1.1 Rationale for evolving to an operational HAFS:

- HAFS will provide avenues for HFIP goals to be met in the next 5-10 years
- HAFS (and other similar TC analysis and forecasts from other operational centers) will contribute to multi-model ensemble forecasts for improved probabilistic hazard guidance products.
- HAFS will evolve as the next-generation community-based hurricane modeling system for research and operations, taking advantage of resources and common community infrastructure, with support provided by the UFS, expanding into a unique global-to-local scale TC prediction system for all global basins.

HAFS comprises three major components: (1) modeling, (2) observations, and (3) data assimilation.

---

[11] https://www.weather.gov/sti/stimodeling_nggps_implementation

3.1.2 Modeling

Hurricane track forecast improvements beyond four days will require the use of high-resolution global models with at least some executed as ensembles. However, global model ensembles are likely to be limited by computing capability for at least the next five years to a horizontal 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 (NOAA Science Advisory Board, 2006). Maximizing improvements in hurricane intensity forecasts will therefore require high-resolution regional models or global models with moveable high-resolution nests, either of which should also be run as an ensemble. In order to evolve HWRF into HAFS, HFIP will maintain dedicated computing, develop high resolution convective-allowing regional ensembles and products for probabilistic predictions, develop nesting capabilities within the FV3[12] model, incorporate HWRF capabilities into FV3-Convection Allowing Model (FV3-CAM), increase spatial and temporal resolutions and number of vertical levels, carefully design and implement ensembles at optimal resolutions using state-of-the-art perturbation schemes, employ scale-aware physics to accommodate global-regional-convective allowing schemes, and implement full earth system modeling by way of coupled atmosphere, ocean, wave, and land modeling components.

3.1.2.1 *Current Modeling Capabilities*

3.1.2.1.1 Operational HWRF System

One of the major accomplishments of HFIP has been the development and operational implementation of the triple-nested, high-resolution HWRF model, which is now one of the top-performing track prediction models. Improvements to model resolution (3 km in 2012, 2 km in 2015 and 1.5 km in 2018), physics, and initial conditions, and the addition of aircraft observations have led to steady progress in improved numerical guidance. Figure 1 illustrates the progress of HWRF in forecasting track (Fig. 1a) and intensity (Fig. 1b). In fact, in 2017 the skill of HWRF was better than the global model at most forecast times. Figure 1b portrays the progress of HWRF in forecasting intensity. There is a steady decrease in track as well as intensity errors. More importantly, there is a steady decrease of intensity error from 2011 to the present by 15% to 20% per year. HWRF was the best intensity forecast guidance model in 2015 for the North Atlantic Basin and in East Pacific in 2016. In the 2017 season, the operational HWRF was again the best dynamical guidance hurricane model in Atlantic basin. The model is consistently performing better than statistical-dynamical models. The model has met the 5-year intensity improvement goals set-forth by HFIP. However, intensity forecast error improvements seem to be leveling off in the Atlantic basin (Fig. 1b). Sustained HFIP research and developments, especially focused on forecast failures and RI events, is recommended for further improvements in intensity predictions. It is also expected that the development of the Hurricane

---

[12] https://www.gfdl.noaa.gov/fv3/

Analysis and Forecasting System (HAFS) may be able to provide some accelerated progress in reaching the HFIP 10-year goal.

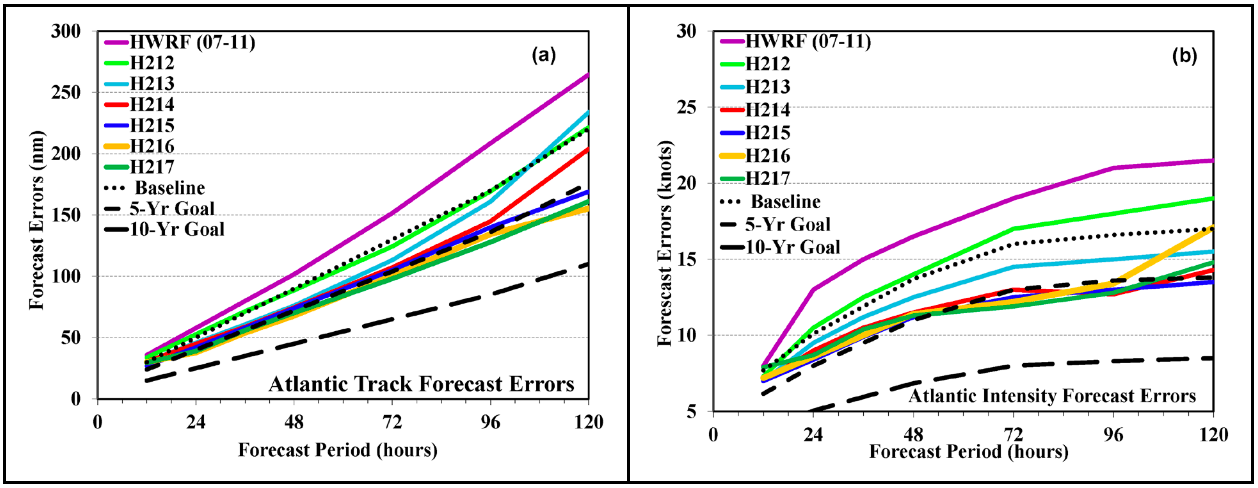


Figure 1: HWRF improvements in forecast track (a) and intensity (b) error for the Atlantic Basin from 2007-2011 to 2017 under HFIP are shown. Seasons for which models were run are depicted on each line. Note that some samples (years) are not homogeneous between models.

3.1.2.1.2 Operational HMON System

Until its retirement in 2016, the Geophysical Fluid Dynamics Laboratory (GFDL) hurricane model was a second operational dynamical model. The model contributed to the diversity required for a robust multi-model ensemble. But, its capabilities and future were limited by its hydrostatic framework. The need for a better, high-resolution non-hydrostatic hurricane model to complement HWRF in the ensembles led to the development of the Hurricanes in a Multi-scale Ocean-coupled Non-hydrostatic (HMON) model. HMON became operational in 2017. It was built using a common-modeling infrastructure shared with unified model development within the NOAA Environmental Modeling System (NEMS[13]), and is capable of being coupled to other models (e.g., ocean, wave, land, surge, inundation, etc.) within that infrastructure. Use of NEMS also paves the way for use of the Common Community Physics Package (CCPP[14]). HMON has different characteristics than HWRF for both the North Atlantic and North Eastern Pacific basins, hence providing the diversity for operating multi-model ensembles for intensity predictions. HMON is not yet available to the community.

3.1.2.1.3 "Basin-Scale HWRF"

Although the current operational HWRF system shows improved skill in intensity forecasting over its earlier versions and predecessors, it should be noted that the operational HWRF employs a storm-centric (SC) configuration with a single embedded nest of higher spatial resolution near the storm center. These attributes are not ideal for representing multi-scale interactions or for TC

---

[13] https://www.weather.gov/sti/stimodeling_nggps_implementation

[14] https://dtcenter.org/community-code/ipd-ccpp

genesis forecast applications and greatly limited the model's potential for improving forecast skill beyond five days. Also, the current operational configuration poses many challenges for advancing data assimilation techniques and downstream applications at and after landfall. Thus, an experimental version of HWRF, called "basin-scale HWRF", was created with a large outer domain covering approximately one-fourth of the globe. It comprises multiple moving nests at 1-3 km horizontal resolution for each TC in its domain, enabling it to produce simultaneous forecasts for multiple TCs. This has great potential because tropical cyclones interact both with the large-scale environment and with one another (see Fig. 2).

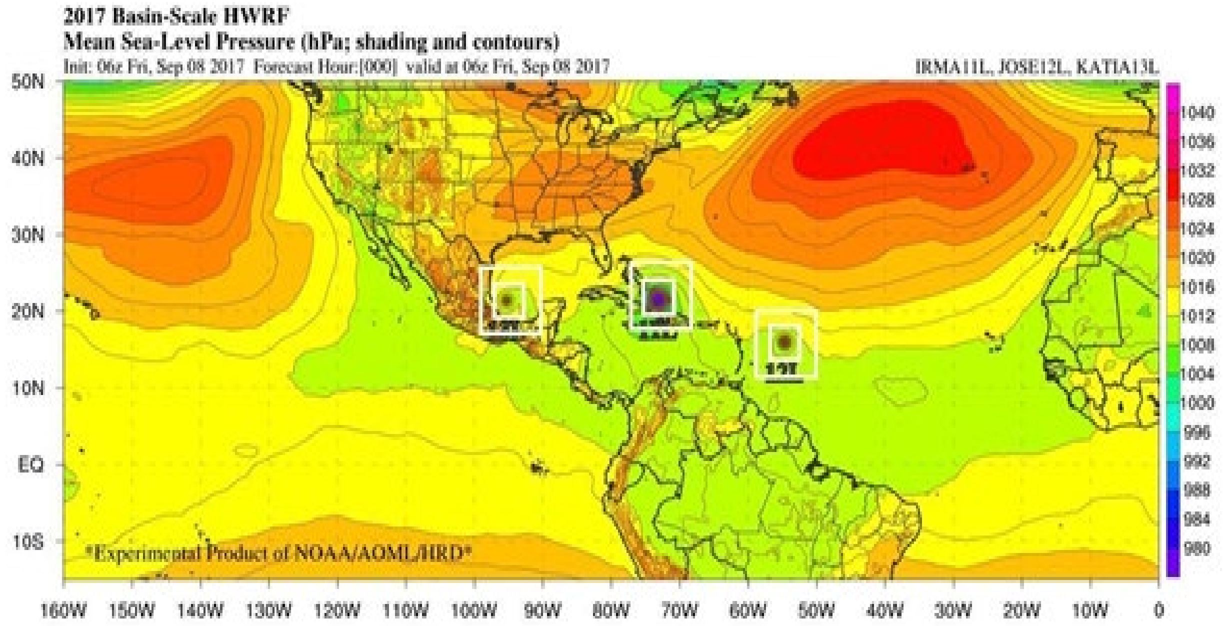


Figure 2: Basin-Scale configuration of 2017 operational HWRF with multiple moving nests at 2-km resolution covering Hurricanes Irma, Jose and Katia. This configuration was run under stream 2 for the entire 2017 season. The above picture shows one cycle initialized on 06Z on 08 September 2017. This capacity was developed by HFIP.

3.1.2.1.4 Regional Multi-Model Ensembles

HFIP has been testing a 41-member multi-model regional ensemble system comprising HWRF (20 members), the Navy's Coupled Ocean/Atmosphere Mesoscale Prediction System-TC (COAMPS-TC, 10 members) and the GFDL model (now HMON, 11 members). Regional ensemble model progress has generally lagged behind single-member "deterministic" models, in part, due to a lack of computational resources to run ensembles. Spread, or more often a lack of realistic spread, among ensemble tracks is a problem. In addition, often the ensemble mean track diverges significantly from the verifying track and/or the parent deterministic forecast of either global or regional models. Regarding the latter, for instance, HWRF ensembles lack the DA and (ocean) coupling advances that are part of the operational HWRF. In 2017, HWRF ensembles run under the HFIP "Stream 2" (real-time experimental) system provided useful information to researchers to improve the utility of the ensemble for producing probabilistic hazard guidance. In fact, the operational HWRF outperformed the ensemble-mean forecasts for track and intensity predictions. Nevertheless, there were a few interesting cycles that provided some useful information to researchers. The HWRF-HMON-COAMPS multi-model ensembles for Hurricane Irma (12Z cycle 2017/09/08; Fig. 3a) and HWRF ensemble for Hurricane Harvey (12Z cycle 2017/08/22; Fig. 3b) are worth noting: both these forecasts provided reasonable spread around the mean which was close to observations.

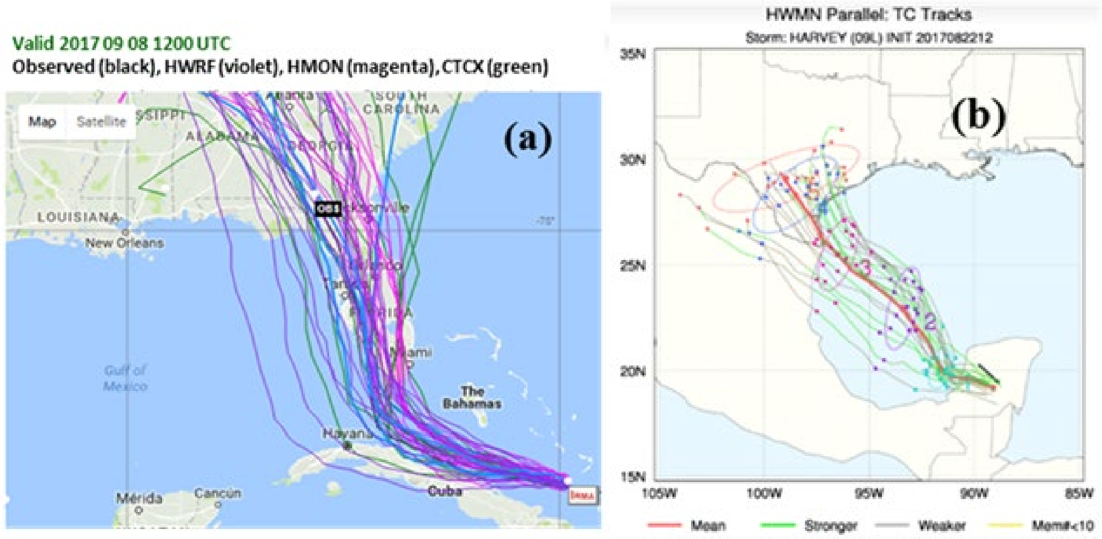


Figure 3: The HWRF-HMON-COAMPS multi-model ensembles for Hurricane Irma (12Z cycle 2017/09/08) and HWRF ensemble for Harvey (12Z cycle 2017/08/22).

3.1.2.1.5 Unified Modeling System for Hurricane Forecasting

In 2016, GFDL's FV3 model was selected to be the next dynamical core for NOAA's Global Forecasting System. FV3 is a fully non-hydrostatic model and will replace the operational Global Spectral Core at National Centers for Environmental Predictions (NCEP) in 2019. The first version will be run at a horizontal resolution of about 13 km. Nevertheless, two kinds of downscaling are possible with this model: grid stretching and telescopic nesting[15]. The GFDL FV3 dynamical core with Global Forecast System (GFS) physics (fvGFS) was used to perform near-real-time forecasts of tropical cyclone track, structure, and intensity out to 132 h during the 2017 Atlantic hurricane season under the HFIP stream 2 experimental efforts. The model domain covered the entire Atlantic basin with a horizontal resolution of 3 km and forecasts were run from early August through late October.

---

[15] https://journals.ametsoc.org/doi/pdf/10.1175/MWR-D-11-00201.1

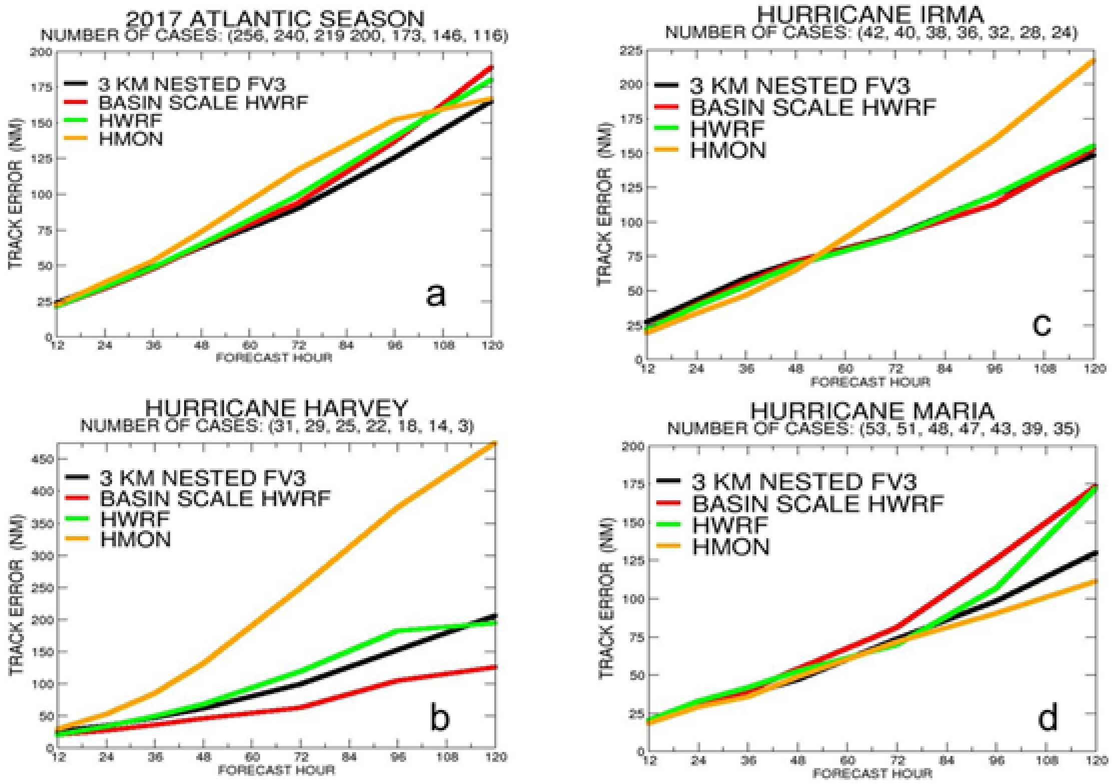


Figure 4: Stream-2 models (Basin-scale HWRF and 3km-FV3) compared to Stream-1 operational models (HWRF & HMON) track error for (a) 2017 AL season and Hurricanes (b) Harvey, (c) Irma and (d) Maria.

Figure 4 illustrates track errors from the Stream-2 models compared to the operational models. Overall (Fig 20a), both Stream-2 models were competitive with the operational model. Significant storm-to-storm variability between high resolution models, but nested fvGFS was the best track performer for entire season (Fig. 4a). Basin-Scale HWRF performed extremely well for Harvey (Fig. 20b). This HFIP effort demonstrates that the FV3 global model has reached a stage of maturity for further advancement of the HAFS system.

3.1.2.1.6 Modeling Community Involvement

R2O and Operations to Research (O2R) is supported by HFIP through the development of a repository for a community-based hurricane modeling system, which ensures the same code base can be used for research and in operations. From 2009-2016, Environmental Modeling Center (EMC) and the Developmental Testbed Center (DTC) worked to update the operational version of HWRF from version 2.0 to the current community version of HWRF (version 3.9a). The 3.9a version makes the code in the community repositories completely compatible with operational model codes, allowing researchers access to operational codes, thereby enabling improvements made by the research community to be easily transferable into operations. In 2017, there were more than 1,300 registered HWRF researchers worldwide. A similar testbed activity is recommended for conducting R20 activities with the proposed FV3-based hurricane forecasting system.

3.1.2.2 *Proposed Modeling Strategy for Implementation of HAFS*

As noted, the FV3 dynamic core will replace the Global Spectral core within GFS in 2019. This lays the groundwork for adoption of the FV3 core for all regional/convective allowing model (CAM) forecast system developments at National Weather Service (NWS), including the HAFS within the broader framework of UFS. The hurricane intensity and structure forecasting problem requires both a large domain, as well as a very high-resolution (1-2 km) domain (1-2 km) to resolve convective-scale motions in the eyewall region. FV3-based CAM experiments have shown the model's capability to operate at high-resolution, cloud-resolving scales.

Some of the needed steps for various elements of the proposed HAFS include:

- Develop storm-following (moving), telescopic, two-way interactive nests operating at about 1-2 km resolution which can be located anywhere on the globe and capable of following TCs for several days;
- Add hurricane-specific physics (from HWRF and HMON) to the CCPP for use with FV3-based HAFS. In addition, seek opportunities for unification of physics between various UFS applications in consultation with UFS Physics Working Group;
- Build high resolution initialization and pre-processing capabilities for HAFS-nested domains, including high-resolution terrain, and land-sea masks;
- Adopt HWRF's vortex initialization and storm relocation capability for HAFS;
- Build inner-core data assimilation capability for HAFS aligned with Joint Effort for Data Assimilation (JEDI[16]) developments;
- Adopt National Unified Operational Prediction Capability (NUOPC[17]) based mediators for coupling HAFS to ocean, wave, surge and inundation models;
- Extend HAFS to 7-10 days for tropical cyclogenesis and potential TCs;

Earlier efforts supported by NGGPS allowed scientists at Atlantic Oceanographic and Meteorological Laboratory (AOML), GFDL, and EMC to design a prototype technique for nest motion shown in Fig. 5. Consequently, a plan was developed to implement moving nest techniques into FV3 during 2019-2021 along with transition of further scientific and technical enhancements from HWRF/HMON to HAFS.

HFIP will also continue to build upon multi-model regional ensemble capabilities by designing and implementing ensembles at optimal resolutions using state-of-the-art perturbation schemes, while ensuring progress is tested and implemented within the HAFS.

---

[16] https://www.da.ucar.edu/sites/default/files/Auligne_JCSDA_WS.pdf

[17] https://www.earthsystemcog.org/projects/nuopc/

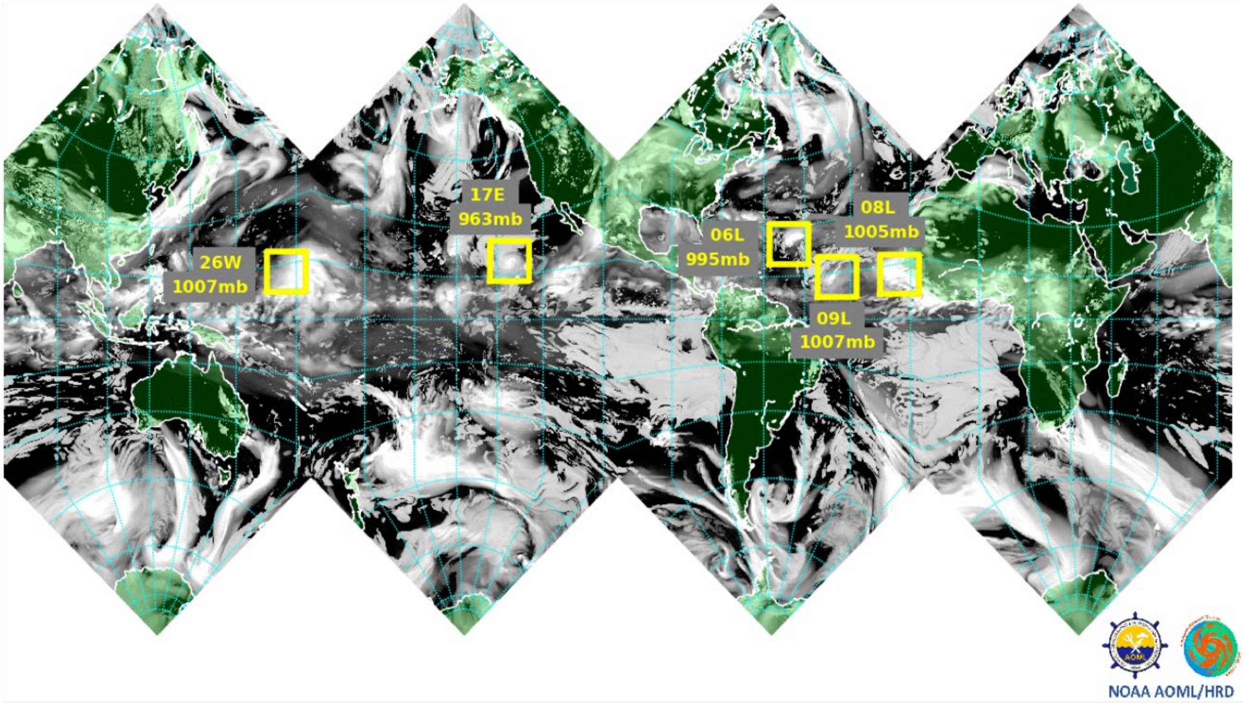


Figure 5: How high-resolution nests may be moved seamlessly within the six faces of the FV3 cube sphere grid. High-resolution TC nests are shown for 5 storms: 26W: Mangkhut; 17E: Olivia; 06L: Florence; 08L: Helene; and 09L: Isaac. For example, if the nest for 26W crosses the edge of one face to another the nest will stay on one projection. The feedback and downscale at the leading edge of the moving nest will be on the interchangeable equivalent projections between faces.

3.1.3 Data Assimilation

Proper initialization of the TC vortex and its environment with advanced data assimilation methods, together with proper treatment of physics, is key to achieving the HFIP goals for track, intensity, and structure. The data assimilation (DA) system is also necessary to produce a high-resolution analysis using all available observations to evaluate model guidance and provide a historical record of events for storm attributes not analyzed by the NHC. Though great strides have recently been made in HWRF DA, more work remains to be done In particular, there are a number of known problems in the current hurricane DA system that will require varying degrees of effort to resolve. These include:

1. Vortex initialization procedures need to work more seamlessly with the data assimilation system. The current procedure, while helpful in some ways, destructively interferes with the data assimilation system when inner-core observations are available. A possible alternative that needs to be explored is to assimilate synthetic observations to supplement inner-core observations.
2. All state variables need to be carried from one cycle to the next, which is not currently the case in HWRF. Most crucially, HWRF currently does not cycle condensate or vertical motion, which is known to impact the analysis. The current self-cycled three-dimensional

hybrid ensemble-variational (3DEnVAR) HWRF DA system improves upon the old DA system, but more development is needed to improve dynamic balance, particularly for intense hurricanes where inner core gradients are extremely large. Among necessary improvements are an upgrade to four-dimensional hybrid ensemble-variational data assimilation (4DEnVAR) from 3DEnVAR and also to cycle DA more frequently (e.g., every hour instead of 6 hours).

3. The current HWRF DA makes suboptimal use of observations. For example, though all reconnaissance data are now assimilated into HWRF, much of this data has had no assumed observation error tuning. Also, though the HWRF system assimilates satellite radiances, it currently uses bias correction from the global model, which is problematic since HWRF and the global model does not have the same biases.

3.1.3.1 *Current Data Assimilation Capabilities*

The HWRF is the focus of this section since it is presently the only operational hurricane model to include data assimilation. While all DA developments described below are in reference to HWRF, they also have applicability to HMON. Much of the description regarding new data is also applicable to the development of the next-generation global model (FV3-GFS). It is anticipated that HWRF will remain the most advanced hurricane modeling system through the next 5 years (2023). Sometime thereafter, FV3-GFS will likely supplant HWRF.

Data assimilation in HWRF has improved dramatically as a direct result of HFIP. Among the improvements was a major upgrade to its DA system in 2017. Past versions of HWRF used error covariance provided by the NCEP global model, which is suboptimal for TC applications in a GSI hybrid scheme. The 2017 version of HWRF introduced a fully-cycled EnKF to provide error covariance for GSI, which is more accurate for TCs. The workflow for this system, which was developed in collaboration with the University of Oklahoma with funding from HFIP, is illustrated in Fig. 6.

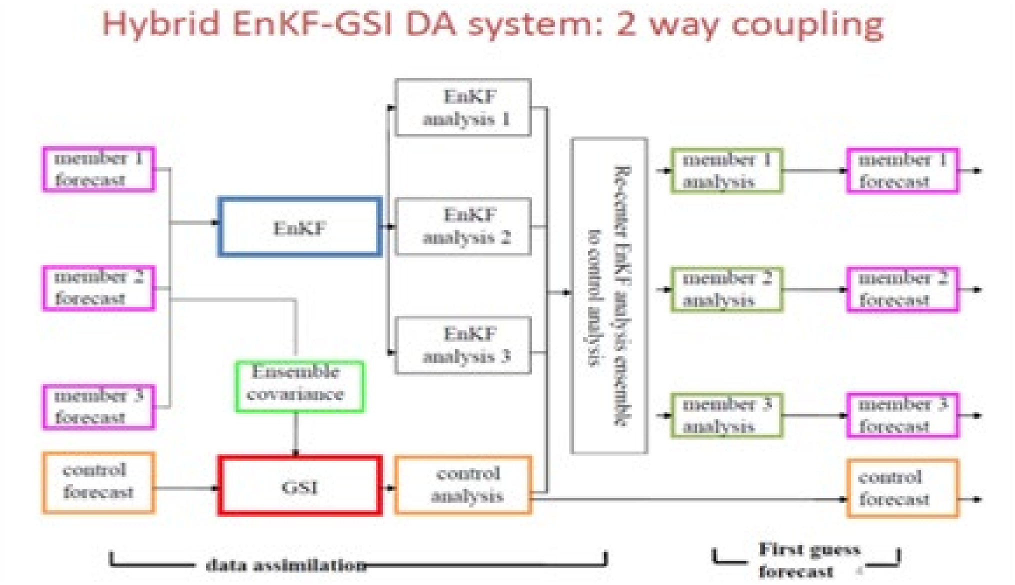


Figure 6: Diagram illustrating the new HWRF DA system with full covariance cycling provided by EnKF

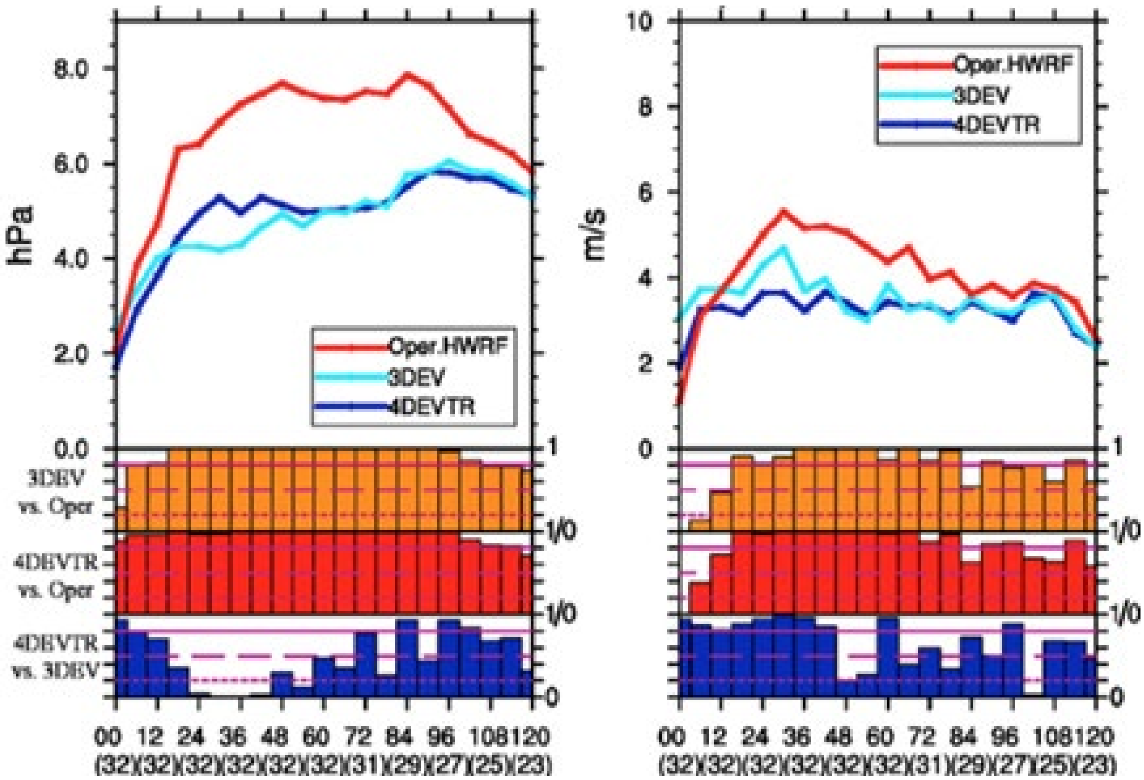


Figure 7: Performance of the research version of the HWRF DA system (adapted from Lu et al, 2017) for Hurricane Edouard in terms of (left) MSLP error and (right) V_{max} error. The 2014 operational HWRF (red line) is compared with the 3D ensemble-variational (light blue) and 4D ensemble variational (dark blue) versions of the research system. Statistical significance is assessed in the bar graphs beneath the error plots, where values near 1 indicate the research system is significantly better than the operational HWRF. Sample size for each forecast hour is indicated at the bottom of each figure panel.

Testing revealed that adopting the new DA strategy can significantly improve TC intensity forecasts. Lu et al. (2017) examined a research version of this DA system with the 2014 HWRF implementation and found major improvements for the forecasts of Hurricane Edouard (2014). These improvements were in part due to a more accurate initial vortex structure, which allowed the model to better capture Edouard's intensification phase. For Edouard, this resulted in a ~25% reduction in the maximum wind speed (V_{max}) error as well as a ~35% reduction in minimum sea-level pressure (MSLP) error (Fig. 7).

There have been a number of other HWRF DA system upgrades over the past few years. Among the more important upgrades is tuning of the new DA system that was implemented in 2017. For example, stochastic physics perturbations were added to the covariance ensemble in the 2018 upgrade. This includes a stochastic convective trigger for the HWRF cumulus parameterization scheme, stochastic boundary layer height perturbations, and stochastic perturbations to the drag coefficient. This change, which increases ensemble spread so more weight is given to observations, improves track and intensity across a large sample of storms.

Concurrent with the DA system upgrades, HWRF now makes much greater use of data than in the past, particularly within the inner core. Flight-level high-density reconnaissance data were added in the 2017 HWRF upgrade, and the 2018 upgrade improved the use of dropsonde data and added both Stepped-Frequency Microwave Radiometer (SFMR) 10-m wind speeds and tail-Doppler radar (TDR) data from the NOAA G-IV. The improvements in data usage, particularly for the inner core, mean that inner core reconnaissance flights are becoming increasingly important for improving intensity forecasts. Retrospective studies of 2017 storms revealed that reconnaissance aircraft data improved HWRF intensity forecasts by about 10% through 48 h (Fig. 8) for a sample of major hurricanes. Note that major hurricanes represent the most difficult cases for reconnaissance DA improvement in HWRF.

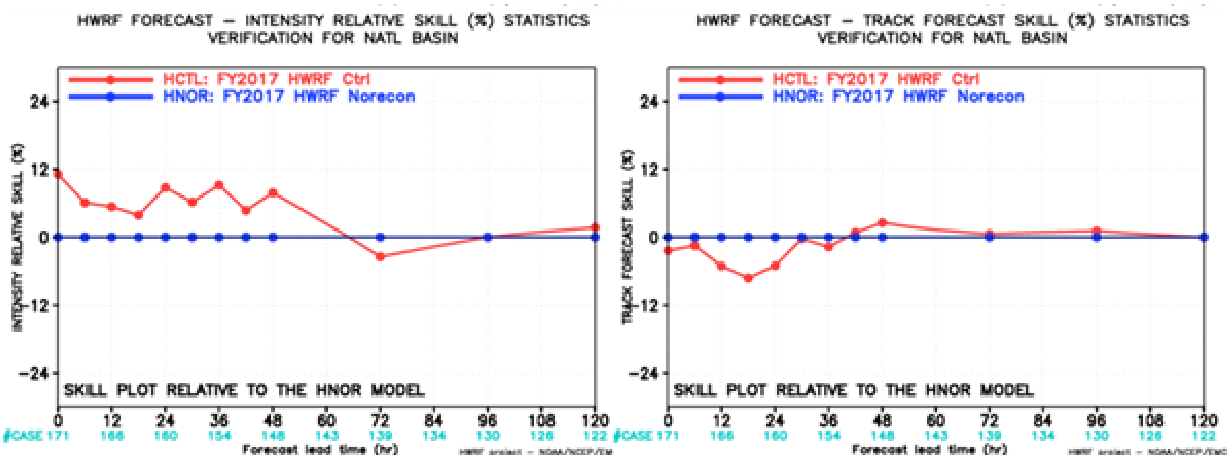


Figure 8: The impact on HWRF V_{max} (left) and track (right), in terms of skill, of reconnaissance on HWRF.

3.1.3.2 *Proposed Data Assimilation Strategy for Implementation of HAFS*

Though great strides have recently been observed in HWRF data assimilation, much more work remains to be done to make the system truly state-of-the-art. In particular, there are currently a number of known problems in the HWRF DA system that will require varying degrees of effort to resolve.

3.1.3.2.1 Modifications to vortex initialization procedures

One significant problem lies within the procedure used to provide the first-guess for each DA cycle. Typically, DA relies on a first guess provided solely by a previous forecast, but that approach does not work well for TC vortices when observations are sparse. Thus, HWRF relies upon a vortex initialization procedure that can be separated into a part that relocates (VR) and modifies (VM) the vortex (size and intensity) from the previous 6-h forecast. These procedures provide a first guess for HWRF DA.

Unfortunately, the vortex modification (VM) procedure in HWRF frequently produces vortices that are too large and strong, which results in GSI analysis increments (i.e., differences between the first guess given to GSI and the subsequent GSI analysis) that overwhelmingly tend to weaken the analysis of the TC vortex. This problem can clearly be seen by comparing the first guess to the analysis increments for a cycle of Hurricane Irma (Fig. 9). The increments strongly oppose the wind components through the entire depth of the TC vortex out to a radius of several degrees. This circumstance is suboptimal since DA systems do not perform as well when there are very large errors in the "first guess field".

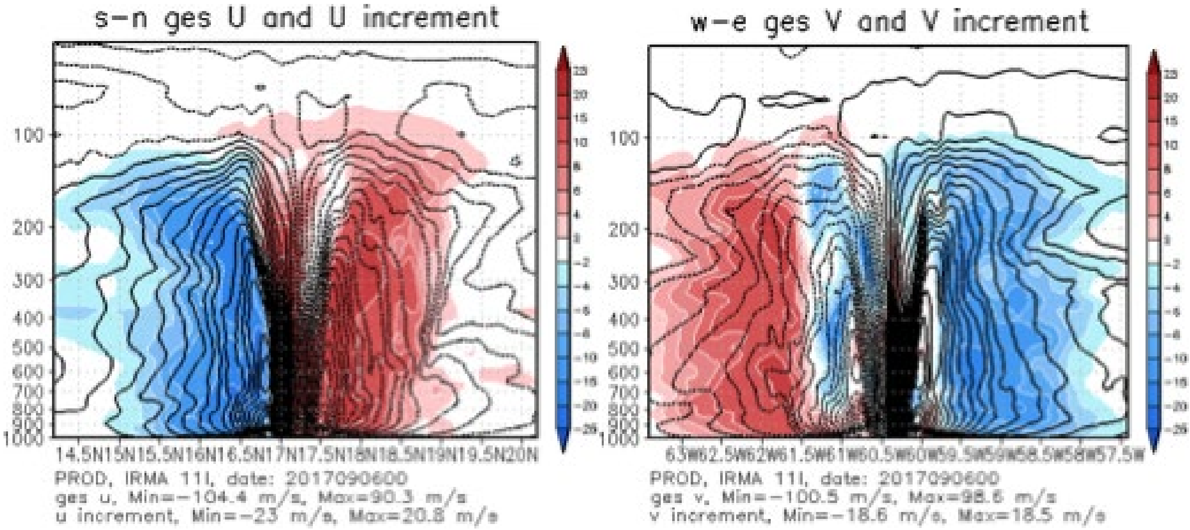


Figure 9. Cross-sections of GSI analysis increments (shaded) of (left) zonal and (right) meridional wind as compared with the first guess (contoured; dashed negative) for the 0000 UTC Sept. 6, 2017 cycle of Hurricane Irma.

The most immediate short-term solution to these issues is to discontinue use of VM when sufficient inner-core data are available. Indeed, this configuration was employed by Lu et. al.

(2017) in their research version of the HWRF DA system, and the operational HWRF system has this ability as well. The impact of implementing such a configuration has not yet been tested with the operational system, but it is a high priority. The long-term solution to this issue is likely to move completely away from VM. One possible solution is to employ self-consistent DA of synthetic wind observations. Such observations could be assimilated alongside real data, or they could be assimilated as a pre-processing step to provide a first guess in the same manner that VM currently does. HFIP will determine as a high priority the solution to this important deficiency in the current model system.

3.1.3.2.2 Improved use of GSI

In addition to improving the first guess for Gridpoint Statistical Interpolation (GSI), a number of improvements ranging in complexity are also needed for GSI. One relatively straightforward improvement that needs to be tested is the addition of one or more outer loops in GSI (for the current implementation, this is essentially the number of times GSI is run in a given cycle). Hsiao et al. (2012) argued that up to four outer loops would be appropriate, particularly for mesoscale situations where the first guess is poor. Given that a poor first guess occurs frequently in HWRF, additional outer loops could improve the analysis and the HWRF forecasts that depend on it. Another update that will require a greater investment is the extension of the control and state variables in GSI to include vertical motion and condensate. This improvement should work in concert with the items discussed above, to provide a more realistic analysis with a self-consistent secondary circulation. Such advancements are likely necessary to alleviate forecast spindown issues (e.g., erroneous weakening of 10 kt or greater in the first 12 h), which are known to plague both operational and research systems. This development is also necessary to support assimilation of all-sky radiances.

Perhaps the most significant development necessary to appropriately initialize the TC vortex is to upgrade the HWRF DA system to 4DEnVAR from 3DEnVAR. In the operational HWRF, observations are assimilated using fixed error covariances that do not evolve over the 6-h DA window. When storms are rapidly changing, such as going through RI and eyewall replacement, DA methods that account for the temporal evolution of the error covariances within the 6-h window are likely needed. Additionally, while the analysis is valid at the center of the 6-h DA window, the inner-core observations are usually not valid at the analysis time but rather distributed over the 6-h window depending on the aircraft flight times. In such cases, using a four-dimensional error covariance approach may more accurately update the analysis at the specified time.

An excellent example of the benefits of 4DEnVAR over 3DEnVAR is illustrated in Fig. 10. It compares analyses of Hurricane Edouard (2014) to radar observations. The 3DEnVAR analysis (Fig. 10 b,e) contains a very strong spurious wind maximum northeast of the circulation center, whereas the 4DEnVAR analysis (Fig. 10 c,f) exhibits a much better fit to the available observations. The resulting forecast from the 4DEnVAR analysis demonstrates a much better fit to the observed evolution (not shown) , and the overall TC intensity forecasts initialized from 4DEnVAR are superior to those initialized from 3DEnVAR (Fig. 7).

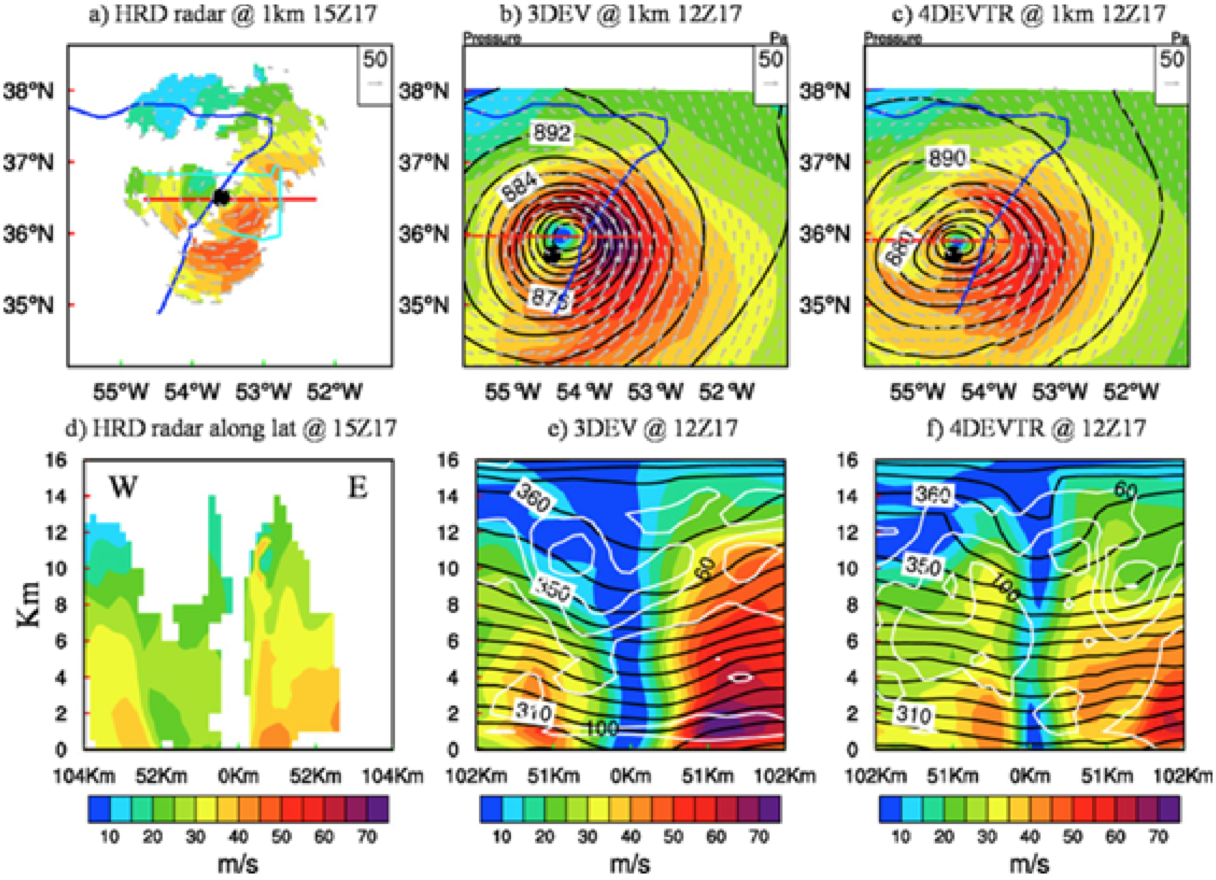


Figure 10. A comparison of observations and analyses of Hurricane Edouard in terms of (a-c) Horizontal plots at 1-km height and (d-f) vertical cross sections through the center on 17 September 2014. All horizontal plots show wind speed (filled) overlain with wind vectors, while the analyses also show pressure (black contour). All vertical plots show wind speed, and the analyses additionally show potential temperature (black contour) and relative humidity (white contour). The blue line in (a-c) denotes the flight track when TDR data is available between 1258-1417 UTC, and the turquoise line in (a) denotes the flight track when TDR data is available between 1617-1708 UTC. Due to the data distribution, the HRD radar composite is valid at 1500 UTC 17 September 2014, while the analyses are at 1200 UTC. Figure taken from Lu et al. (2017).

3.1.3.2.3 Improved use of currently assimilated data

Though a great deal more data are assimilated in HWRF than several years ago, particularly in the inner core, these data are very likely not being used in an optimal way. For example, there has been very limited recent tuning of assumed observation errors, or thinning, both of which are known to be very important for DA system functionality. Another major shortcoming is in the handling of satellite radiances in HWRF. The moving nest configuration in operational HWRF makes satellite radiance bias correction impossible, so the operational system relies on bias correction coefficients borrowed from the global Global Data Assimilation System (GDAS) system. Experiments with a version of the Hurricane Research Division (HRD) HWRF basin-scale system which assimilates data on a single, large domain show that while this does work for some satellite channels (i.e. the coefficients are similar; see Fig. 11), it is a poor assumption for others (i.e., the large coefficient in the "basin-scale HWRF" is not present in GDAS).

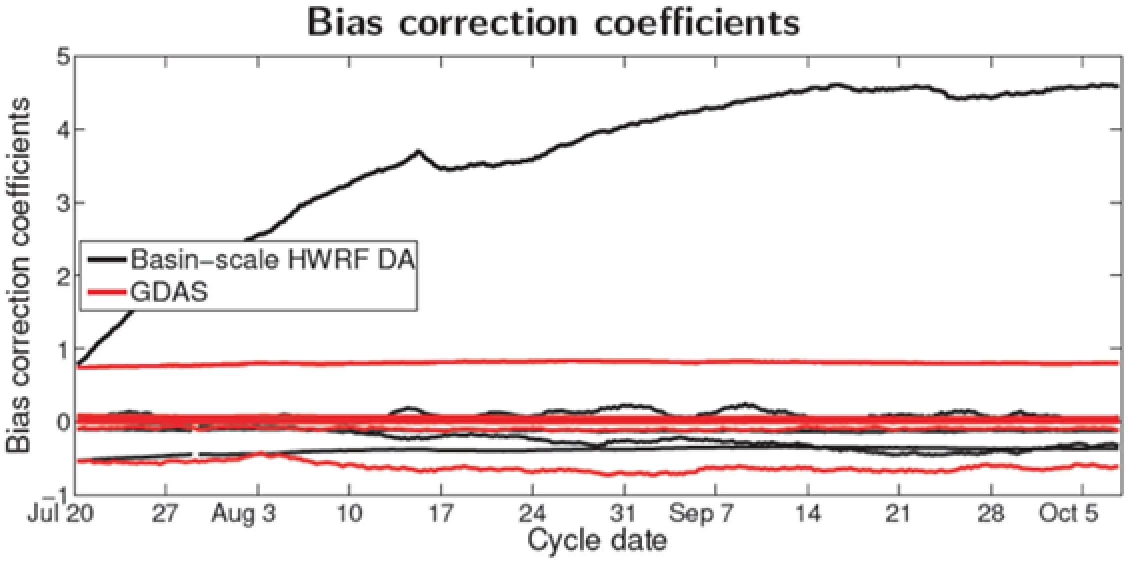


Figure 11. Radiance bias correction coefficients as calculated in GDAS (red) as compared with coefficients calculated in an experimental HWRF basin-scale system.

The appropriate solution to the bias correction problem in HWRF is to migrate to a basin-scale approach with a large, static outer domain on which a large sample of radiance data is assimilated. Testing with the aforementioned experimental basin-scale system has shown that HWRF physics are generally unbiased on the large-scale with the exception of near-surface moisture over the NATL basin. Given this finding, concerns regarding model drift with the basin-scale approach should be alleviated.

3.1.3.2.4 Using additional data

There is a large amount of data not used optimally by the operational HWRF. For example, a major source of unused data for near-land TCs comes from NOAA WSR-88D Doppler radars. Doppler velocity is particularly useful for improving TC forecasts (e.g., Zhang et al. 2009). Improving use of this data will most benefit forecasts of TCs that linger in coastal areas for long periods of time (e.g., Harvey and Irma in 2017). The data type requiring the most investment to take advantage of is all-sky radiances. As previously mentioned, the current operational HWRF system does not cycle condensate and cannot do so until upgrades are implemented. This upgrade, which represents a major investment, will have to be tested before any work on radiance assimilation can commence in HWRF. Furthermore, it is quite likely cloud-contaminated radiances will require computation of bias-correction coefficients native to the HWRF system. As outlined above, such an approach will necessitate a migration to a large, static outer domain on which the radiance data is assimilated.

3.1.3.3 Migration to FV3-HAFS

Though the work outlined above is necessarily focused on the HWRF implementation of HAFS, the ultimate goal is to move to an FV3 implementation using the knowledge gained from improvements to HWRF. Development and configuration selection (e.g., standalone regional or nested global) for the FV3-based HAFS will evolve over the next several years, after which the integration of a DA system can take place. A fully functional prototype FV3-based hurricane system should be available to or compared with the operational HWRF within three years. It is anticipated that several more years of DA development will be required for the FV3-based system at that point to catch up with ongoing HWRF developments, some of which have been discussed above.

3.1.4 Observations

Combined with numerical models and data assimilation, observations form the crucial third component to a comprehensive strategy for developing the HAFS system. Observations provide important information on the structure of the TC and its environment to forecasters. They are also invaluable for data assimilation/model initialization and model evaluation. They allow for studies on the impact of data on forecast skill and provide a means for optimizing data collection and sampling strategies to produce the best forecasts. Finally, observations can be used to develop and test theories of TC motion, structure, and intensity change through process studies that can result in better model diagnostics and guidance.

3.1.4.1 *Current Observational Capabilities*

Current observational capabilities, as discussed below, are presently used in the HWRF analysis. Next-generation observing systems and strategies will be incorporated into the HAFS system to produce improved analyses and forecasts.

3.1.4.1.1 Satellite Observations

Geostationary and polar-orbiting satellites provide crucial information on cloud structures, cloud-top temperatures, and moisture and precipitation distribution. Such information provides information used operationally by tropical meteorologists to estimate TC position and intensity over all of the oceanic basins through the “Dvorak technique”, and are especially important for locations where aircraft cannot reach or are not routinely flown. In addition, satellite-derived products, such as cloud-drift winds, provide information on the characteristics of the environment, which are vital in assessing environmental controls on TC track and intensity.

The recent successful launch of Geostationary Operational Environmental Satellite (GOES) 16 provides tremendous opportunities to observe TCs at high spatial and temporal resolution. The ability to observe features at 1-minute resolution allows forecasters to identify and track low-level circulations in formative systems and monitor rapidly evolving features like convective bursts that can play a significant role in RI. The lightning sensor on GOES-16/17 also allows forecasters to identify regions of lightning outbreaks, often associated with deep convection, in the inner core and outer bands of TCs. Derived multi-channel products from GOES-16 and other satellites are now allowing forecasters to examine the interaction of the TC and the environment during its entire life cycle from genesis to extratropical transition or decay.

Low-earth orbiting (LEO) microwave imagers on NOAA-20, Defense Meteorological Satellite Program (DMSP) F16-18, Global Precipitation Measurement (GPM), AQUA, and Global Change Observation Mission (GCOM)-W1 {e.g., Advanced Technology Microwave Sounder (ATMS), Advanced Microwave Sounding Unit (AMSU), Special Sensor Microwave Imager/Sounder (SSMIS), GPM Microwave Imager (GMI), Advanced Microwave Scanning Radiometer (AMSR-2)} provide forecasters with a radar-like view of the TC rainband structure and observations of structural changes such as eyewall replacement cycles. Microwave imagery can also provide critical information about the structure and location of the TC's low-level center, particularly at night and in areas not sampled by aircraft. Proper positioning of the TC low-level center is important for determining initial motion and initiating track model guidance.

LEO microwave sounders such as AMSU, ATMS, and SSMIS provide information on the vertical thermal structure of the TC warm core, which are used to derive TC intensity estimates along with estimates of 34-, 50-, and 64-knot wind radii.

Satellite-based scatterometers, such as Advanced Scatterometer (ASCAT), provide wide-swath satellite ocean vector wind observations that are heavily used by operational forecasters in subjective analyses of TC location, intensity, and structure. While current scatterometers cannot sample the inner-core winds of a hurricane due to spatial resolution limitations, scatterometers can provide a snapshot of the entire TC wind field and are used to inform intensity estimates for developing systems, especially for those systems outside the range of aircraft. Satellite altimeters provide valuable information on the marine wave field that is crucial for maritime interests, as well as allow estimates of the subsurface ocean temperature and salinity fields that can help to identify regions of deep mixed layers and warm core eddies important in sustaining TC intensification.

3.1.4.1.2 Airborne Observations

In the Atlantic basin Air Force Reserve C-130 and NOAA WP-3D aircraft are used to sample TCs whenever possible to provide critical observations of the location, strength, and structure of the storm circulation. Sampling of the environment is typically accomplished by the NOAA G-IV aircraft. These manned aircraft are equipped with a variety of instruments that sample the wind, temperature, moisture, pressure, precipitation, and ocean surface and subsurface temperature, current, and wave fields within and around TCs (e.g., with flight-level measurements, dropwindsondes, airborne Doppler radar, Stepped Frequency Microwave Radiometer, lower fuselage radar, and airborne expendable bathythermographs/current profilers).

Experimental airborne observing technologies, such as Light Detection and Ranging (LIDAR), have the ability to sample the wind field in the absence of precipitation scatterers Unmanned aerial systems, such as the Coyote and Global Hawk can sample temperature, moisture, and pressure fields in the planetary boundary layer of hurricanes, and over vast areas at very high altitudes for extended periods of time, areas that can't be reached by manned aircraft because of safety and/or aircraft performance limitations. These experimental observing technologies could potentially fill gaps in the current observing system, providing critical measurements needed to more fully capture the structures important to TC structure and intensity change.

3.1.4.1.3 Ground-based Observations

Ground-based observing platforms include the Weather Surveillance Radar (WSR)-88D network, rawinsondes, and manual and automated surface observing stations (ASOS). These platforms provide important information on the TC environment prior to landfall and later on the surface wind field, precipitation, microphysical structure, and three-dimensional wind fields of the TC inner core near and after landfall.

Ground-based water level sensors (e.g. tide stations) and real-time sensors deployed ahead of TCs {e.g. United States Geological Survey (USGS) Rapid Deployment Gages (RDGs) and storm tide sensors} are crucial to measuring water levels during storm surge events. These observing platforms, in addition to subjectively determined high-water marks, provide important sources of information for validating surge forecasts.

Mobile platforms are frequently deployed when there is a significant landfalling TC event. These platforms include mobile polarimetric radar and profilers. These platforms augment the existing ground-based observing network and can fill in key gaps during landfalling events.

3.1.4.2 *Proposed Observational Strategy for Implementation of HAFS*

HFIP will take advantage of advancements in these observing technologies to optimize sampling of the TC inner-core and environment and provide the needed support for forecaster analysis, model initialization and evaluation, current and future data impact studies (OSEs and OSSEs), and process studies.

3.1.4.2.1 Forecaster Analysis and Situational Awareness

Forecasters provide the critical first step in the forecast process through subjective analysis of the TC location, intensity, and structure. This analysis is critical for a proper understanding of the TC itself and for initialization of TC model guidance, which use parameters analyzed by forecasters as input through data assimilation and other processes. Effective real-time TC analysis by a forecaster is also important for scrutinizing model analyses and short-term forecasts to determine if models have a proper representation of the TC and the surrounding environment. Observational data form the basis of the TC "best track" analysis process[18], which is the

---

[18] NHC and CPHC perform a post-storm analysis of all available observations that constitutes the official historical record for the TC.

foundation of verification data for evaluation of model guidance for TC track, intensity and structure.

HFIP will improve the use of these observations within the operational forecast environment by making more of these observations available to the forecaster in real-time. The increased temporal and spatial resolution of satellite and aircraft observations in the TC core will assist the forecaster to better recognize situations when RI may occur and monitor fluctuations in intensity due to eyewall replacement cycles or other structural changes. In particular, knowledge of the radius of maximum winds (RMW) is especially critical to forecasting storm surge magnitude and extent. RMW is presently detectable almost exclusively by aircraft. HFIP will focus on techniques to bring these high-resolution observations together within a common framework within the second generation Advanced Weather Interactive Processing System (AWIPS II) used by NHC, Central Pacific Hurricane Center (CPHC), and all the NWS Weather Forecast Offices (WFO).

A more complete analysis of the TC wind field, such as that available from better coverage of wide-swath, high-resolution satellite scatterometers, will improve the timing and placement of TC wind watches and warnings, while also providing better information to initialize storm surge guidance models, which are the foundation for the issuance of storm surge watches and warnings and serve as guidance to emergency managers for evacuation planning purposes.

3.1.4.2.2 Model Initialization

The ability to optimally specify the initial TC inner-core and environmental structure is crucial to realizing the best numerical model forecast guidance. Efforts to improve the model initial state used measurements of TC position and intensity to specify the initial TC structure using vortex initialization. More recently, sophisticated efforts have been developed to incorporate many observed fields into the model initial state using data assimilation efforts of varying degrees of complexity. HFIP will continue this strategy with a focus on improving the assimilation of satellite, aircraft, and upper ocean data to improve forecast guidance for RI and pre-formation disturbances.

Improved use of satellite observations to improve forecast guidance is a major focus under the HAFS development. Satellite data provides a wealth of information to facilitate significantly improved specification of the TC inner-core and surrounding environment. The assimilation of atmospheric motion vectors (AMVs), enhanced by the rapid-scan capabilities of GOES-16 and Himawari satellites, are showing positive impacts on the analysis and forecast of TC track and intensity. HFIP will continue the evaluation of these satellite observations in improving the forecast guidance for RI and pre-formation disturbances. Another HAFS priorities will be the evaluation of the assimilation of infrared and microwave radiances from NOAA-20 in order to improve the initial analysis.

TCs form over the ocean and derive energy through processes that act across the air-sea interface. Hence, the use of upper ocean observations is critical to improving the analysis of the ocean in the coupled HAFS, particularly for RI. Synthetic upper ocean profiles of temperature and salinity from satellite altimetry provides the bulk of observations used to initialize the upper ocean in current models. However, these synthetic profiles are based on climatological ocean observations and need to be improved for use in HAFS. New ocean observing technologies such as floats and gliders need to be utilized to provide detailed observations of the upper ocean thermal and salinity profiles that can be used to improve the analysis of the upper ocean in HAFS.

3.1.4.2.3 Model Evaluation

Observations provide a key dataset for evaluating the performance of numerical models. A well-constructed model evaluation study will compare model output and observations in a common framework; identify potential biases in the model based on these comparisons; propose, develop, and implement modifications to the numerical model based on this identification; and compare the modified numerical model with observations to test the impact of these modifications. Providing an improved analysis utilizing all of the available observations is a major goal under the HAFS strategy. This analysis will take advantage of new data assimilation techniques to provide the location, intensity, and structure of a TC to initialize HAFS and also should provide an improved TC best track to use in evaluating HAFS.

HFIP has supported several efforts that have demonstrated this approach and produced improvements in numerical models. For example, the vertical eddy diffusivity in the hurricane boundary layer, evaluated by comparing the prescribed HWRF mixing parameter with flight-level observations from hurricane eyewall penetrations, was reduced in the model to be more consistent with observations. The result of this modification was a reduction in depth of the PBL inflow layer and an increase in the strength of the inflow, resulting in an improvement in the ability of HWRF to predict RI events.

Similar evaluations will be made to advance HAFS using the various observational platforms described above. HFIP will focus on three key areas in order to improve forecast guidance for RI cases: (1) horizontal mixing as a function of model resolution; (2) microphysics and radiative interactions and their role in the vertical thermodynamic structure and heating profile; and (3) the air-sea interface, particularly the role upper ocean thermal structure plays on the exchange of energy to the atmosphere.

3.1.4.2.4 OSEs and OSSEs

The ability to assess the impact of observations on TC analyses and forecasts is a key step in evaluating the value of specific observing systems and strategies. Observing System Experiments (OSEs) combine observations, models, and data assimilation systems to evaluate the impact of existing observations on the subsequent analyses and forecasts of TC track, intensity, and/or structure. Observing System Simulation Experiments (OSSEs) use simulated observations from a nature run to test the impact of future observations on TC track, intensity, or structure analyses and forecasts. The OSE and OSSE approaches will be utilized to assess the impact of the existing or proposed observations, and to develop improved observing strategies, respectively, within the HAFS infrastructure. These assessments will inform decisions on the cost-effective utilization of these observing systems.

3.1.4.2.5 Process Studies

Considerable research has been conducted using observations to better understand the important physical processes associated with TC track, structure, and intensity. HAFS will be used to accelerate this research, combining observations with HAFS analyses to test theories underlying these processes. This work will continue under HFIP and will augment efforts in the development of HAFS.

3.1.5 Summary of Envisioned HAFS:

- The forecasting system will consist of a global model (FV3GFS) and regional models (e.g. HWRF, HMON, FV3CAM) capable of tracking multiple hurricanes at 1-2 km resolution using sophisticated telescopic, two-way interactive, moving nests within all global basins.

- HAFS will consist of a high-resolution analysis (at <3 km) for the Atlantic TC area (capable of being expanded to other global TC basins).

- HAFS Tropical Atlantic (TA) domain analyses will ingest all available airborne and satellite data, building upon advanced DA methods within JEDI/GSI framework for reconnaissance data for both the TA and vortex-following SC moving nests.

- HAFS will include advanced high-resolution ensemble forecasts for TA and SC domains, using stochastic physics, for representation of initial conditions and model uncertainties, thereby allowing for the provision of improved deterministic and probabilistic guidance for forecasters.

- HAFS TA domain analyses will facilitate initialization of the forecast system with very high-resolution SC nests for 7–10-day hurricane forecasts and will also provide the best state for validation of model output.

- HAFS will enable high-resolution 7-10 day forecast guidance for TC genesis and potential TCs in the TA domain initially, and expand it for all global TC domains

- HAFS will incorporate all science (modeling, physics, DA, coupling) advancements of HWRF, all infrastructure (NEMS, NUOPC, Earth System Modeling Framework (ESMF) advancements, and other R2O efforts related to TC prediction within the NOAA's UFS.

- HAFS will be extended to include ocean, waves, storm-surge, inundation and severe weather associated with landfalling hurricanes (with multi-way model coupling) to produce hurricane-related numerical guidance on these phenomena.

3.2 Improve probabilistic guidance

HFIP will use the FACETs framework as an overarching strategy to guide the modernization of the tropical cyclone product suite. HFIP will produce improved, actionable information for preparedness activities, including evacuations, that quantifies uncertainty for all TC hazards, including storm surge, and the development of advanced probabilistic forecast techniques will enhance that effort. The work will employ physical, social, and behavioral science research. Current deterministic and probabilistic guidance will be improved through physical science research, supporting improved track, intensity (including rapid intensification), and size predictions before formation, and throughout the storm's life cycle. Social and behavioral science research will be applied to construct a TC product suite that better communicates the forecast, risk and uncertainty for all TC hazards (surge, rain, associated severe weather, gusts as well as sustained winds). HFIP will improve actionable guidance by improving upon existing methods used to quantify uncertainty for storm surge, rainfall, severe weather, and sustained winds and gusts, as described below.

3.2.1 Storm Surge

Due to its potential for the largest loss of life in TCs, storm surge drives most hurricane evacuation planning and decision making. Currently, real-time storm surge products, including the Potential Storm Surge Flooding Map and the Storm Surge Watch/Warning for the United States (U.S.) East and Gulf Coasts, first become available 48 h prior to the onset of storm surge or tropical-storm-force winds (whichever is expected to occur first). However, given increasing coastal populations, and therefore, time-sensitive evacuation logistics and requirements specific to each coastal location, evacuation and preparedness decisions are often made well in advance of the availability of these real-time storm surge products. Current real-time storm surge guidance is provided by the Probabilistic Hurricane Storm Surge (P-Surge) model, which is a statistically based ensemble model that evaluates a large distribution of storm surge model runs generated by the NWS Sea, Lake, and Overland Surges from Hurricanes (SLOSH) model and uses the NHC official forecast and historical error distribution to capture the uncertainty in TC track, intensity, and structure. However, this statistical information does not represent the near real-time uncertainty in the TC structure. For example, Fig. 12a shows the official forecast track for Hurricane Irma (2017) one day prior to landfall on the southwest coast of Florida. However, the best-track (Fig. 12b) shows Irma made landfall farther east, reducing storm surge potential near Fort Myers and points north by more than 6 feet. To account for this variability, P-Surge begins with the current NHC official forecast and then incorporates historical errors in both

NHC's official track and intensity forecasts to create a distribution of hypothetical hurricane scenarios that each have a chance of occurring (see Fig. 12c).

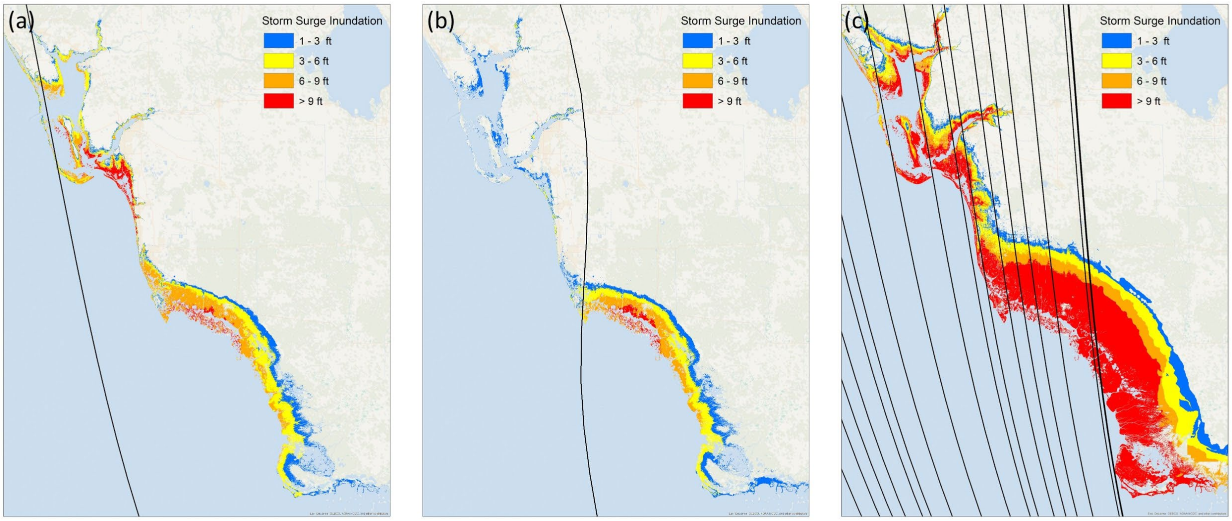


Figure 12. (a) the NHC official advisory one day before landfall, (b) the best-track. Shows the height of the water level above ground (units: ft), and (c) P-Surge track perturbations for Hurricane Irma (2017) one day prior to landfall, based on the NHC official advisory shown in (a).

Although this probabilistic modeling approach has proven integral to providing life-saving surge forecasts, several modifications need to be taken into account to ensure the best possible scientific approach is considered. In particular, the critical parameter of RMW is presently represented in a crude manner within P-Surge. HFIP will invest in improving initial conditions to P-Surge to more accurately represent the initial meteorological state. In particular, asymmetries in wind forcings within the model will be accounted for and a dynamical ensemble approach will be adopted. These improvements will lead to an increased lead time of reliable real-time storm surge guidance and forecasts from two to three days before the arrival of surge and wind hazards (Goal 4.1), as noted in Table A.2.4. In order to improve the initial conditions for P-Surge to accurately represent the meteorological state, real-time RMW information from HAFS will be incorporated into P-Surge and the parametric relationship in SLOSH will be modified to accommodate the real-time RMW information by utilizing V_{max} and RMW, leaving ΔP to be computed, thereby replacing the current methodology of using ΔP and V_{max} from the NHC official advisory to calculate the RMW. Additionally, HFIP will account for asymmetries in wind forcings within P-Surge (Fig.13) by assimilating asymmetries based on NHC official wind radii rather than a wind profile that assumes a symmetric response. This will be done by incorporating the NHC official 34-knot, 50-knot, and 64-knot wind radii. Furthermore, HFIP will incorporate a dynamical ensemble approach in P-Surge, whereby the uncertainty is quantified through perturbing the current (initial) state relative to nonlinear processes in both the atmosphere and ocean. Finally, as described in section A.2.1, HFIP's planned improvements to track, intensity and structure forecasts, along with the incorporation of forecast uncertainty estimates through ensembles and statistical post-processing, will reduce storm surge forecast uncertainty throughout the forecast period.

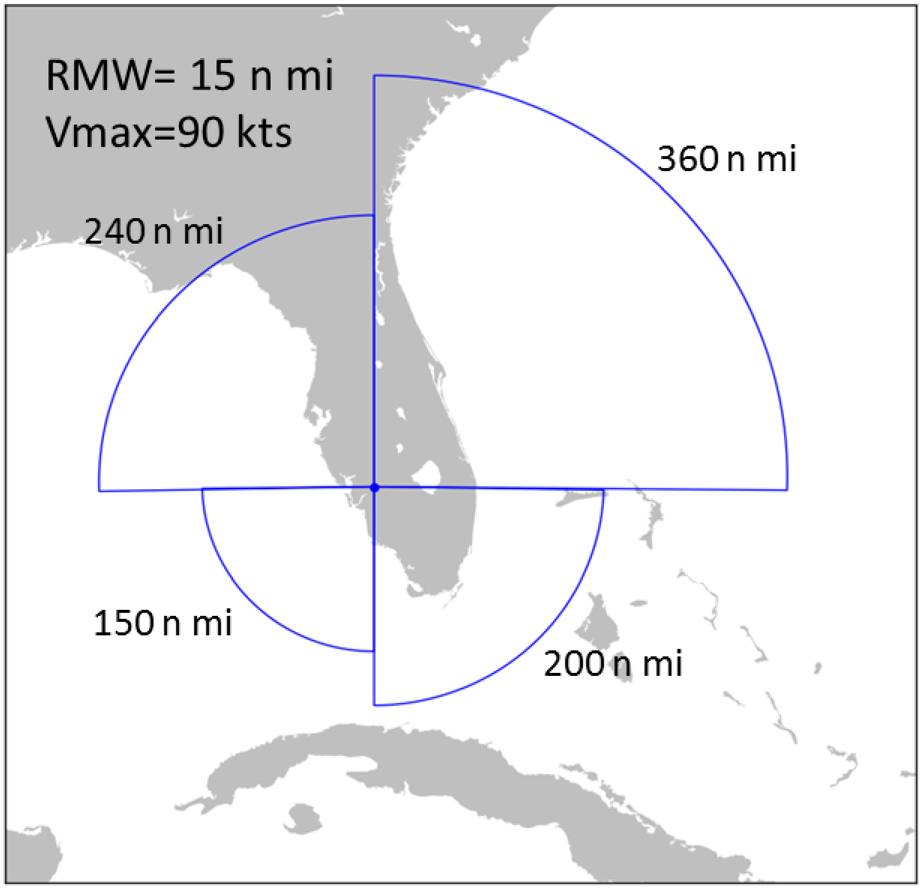


Figure 13. NHC 34-knot wind quadrants for Hurricane Irma on 00Z September 11. The quadrant analysis reveals the tropical storm force winds extends approximately 360 nm in the northeast quadrant, 200 nm in the southeast quadrant, 150 nm in the southwest sector and 240 nm in the northwest quadrant. As a result, the observed water levels reported in Charleston, SC were approximately 4 feet above mean higher high water (MHHW), a proxy for inundation. P-Surge guidance at any time in the forecast period never exceeded more than 2 feet because it could not account for the wind field asymmetry in that area.

HFIP will also expand real-time storm surge forecasting capabilities, including the Storm Surge Watch/Warning, to ensure consistent operational storm surge model guidance, products, and services are available for all areas of NWS responsibility, including off the mainland(e.g., Puerto Rico, the U.S. Virgin Islands, southern California, Hawaii, Guam, and American Samoa) (Goal 4.2), as described in Table A.2.4. This gap in operational guidance for Puerto Rico, for example, was particularly apparent during Hurricane Maria (2017). To extend the present operational surge forecasting capability from mild-sloped coastal areas such as the U.S. East and Gulf of Mexico coasts to steep-sloped areas such as Caribbean and Pacific islands will require a coupled parametric-wave model be included to account for the contribution of waves. Unlike the U.S. Gulf of Mexico and mainland Atlantic coastlines, where the bathymetric profile is shallow (limiting wave height), waves can be a significant contributor to the total water level rise in regions where the bathymetric profile is quite steep.

3.2.2 Rainfall

Over the past 10 years, precipitation forecasts have improved for TCs. The improvement is due, in part, to a significant reduction in track forecast error over that period. Yet, significant errors in the forecast of precipitation that can produce life-threatening impacts still exist, as was seen with Hurricanes Harvey (2017) (Figs. 14 and 15) and Matthew (2016) (Fig. 16). HFIP anticipates better capturing these events with improved track and intensity forecasts through upgrades to regional TC models. More specifically, improvements to the precipitation forecasts are expected through the use of HWRF/HAFS ensembles at varying temporal and spatial scales. The improved probabilistic model guidance is expected to directly contribute toward improved skill of probabilistic Quantitative Precipitation Forecast (QPF) products presently issued by Weather Prediction Center (WPC) during a TC event, from tactical (0-6 hours) to preparedness (1-5 day) timescales (Fig. 16), and hence better excessive rainfall forecasts for TCs. Further investigation into properly representing terrain and its influence on precipitation patterns are imperative for improving extreme rainfall associated with TCs, as well.

HFIP will improve excessive rainfall forecasts for TCs by improving the accuracy and lead time of the Excessive Rainfall Outlook product through day 3 (Goal 4.5); determining and improving the skill (by 10%) of QPF for landfalling TCs over the contiguous U.S. (CONUS), including Puerto Rico and the U.S. Virgin Islands (Goal 4.6); and creating and disseminating a Probabilistic Tropical QPF for CONUS, including Puerto Rico and the U.S. Virgin Islands, based on HAFS/HWRF output (Goal 4.7), as described in Table A.2.4. It is anticipated that these improvements will lead to enhanced flood-related decision support services.

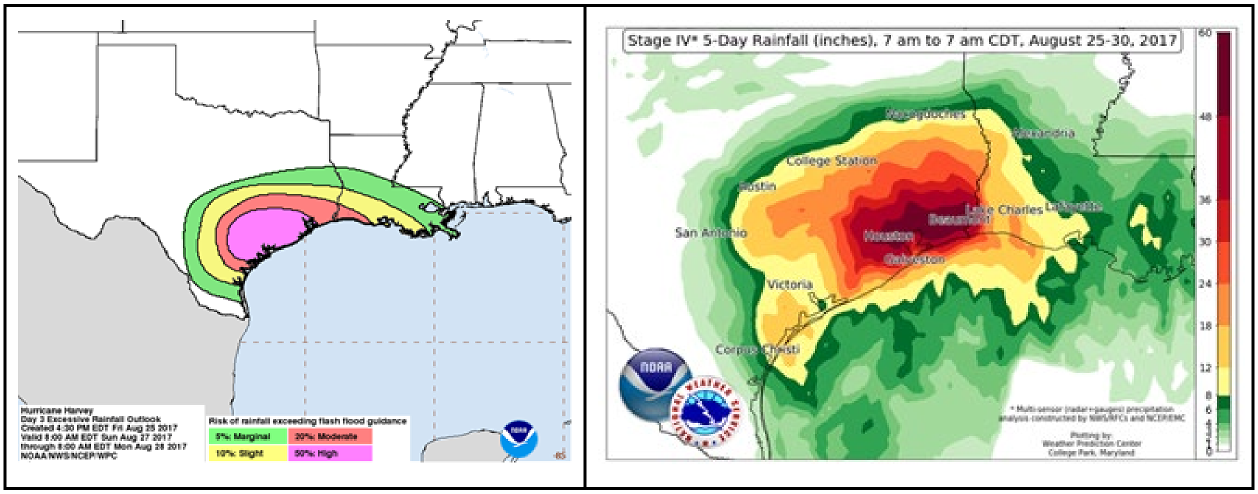


Figure 14. WPC Day 3 Excessive Rainfall Outlook for Hurricane Harvey on the left. On the right, 5-day Total Rainfall for Hurricane Harvey (August 25-30, 2018).

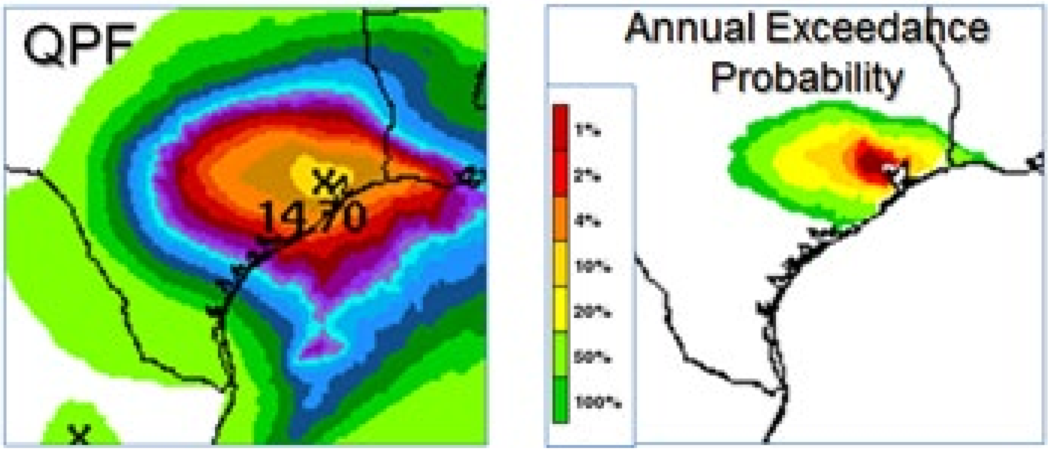


Figure 15. On the left, 24hr Quantitative Precipitation Forecast (QPF) issued on August 26th. On the right, an image of the rainfall forecast on the right reaching a 1% Annual Exceedance Probability (> 100 yr rainfall event).

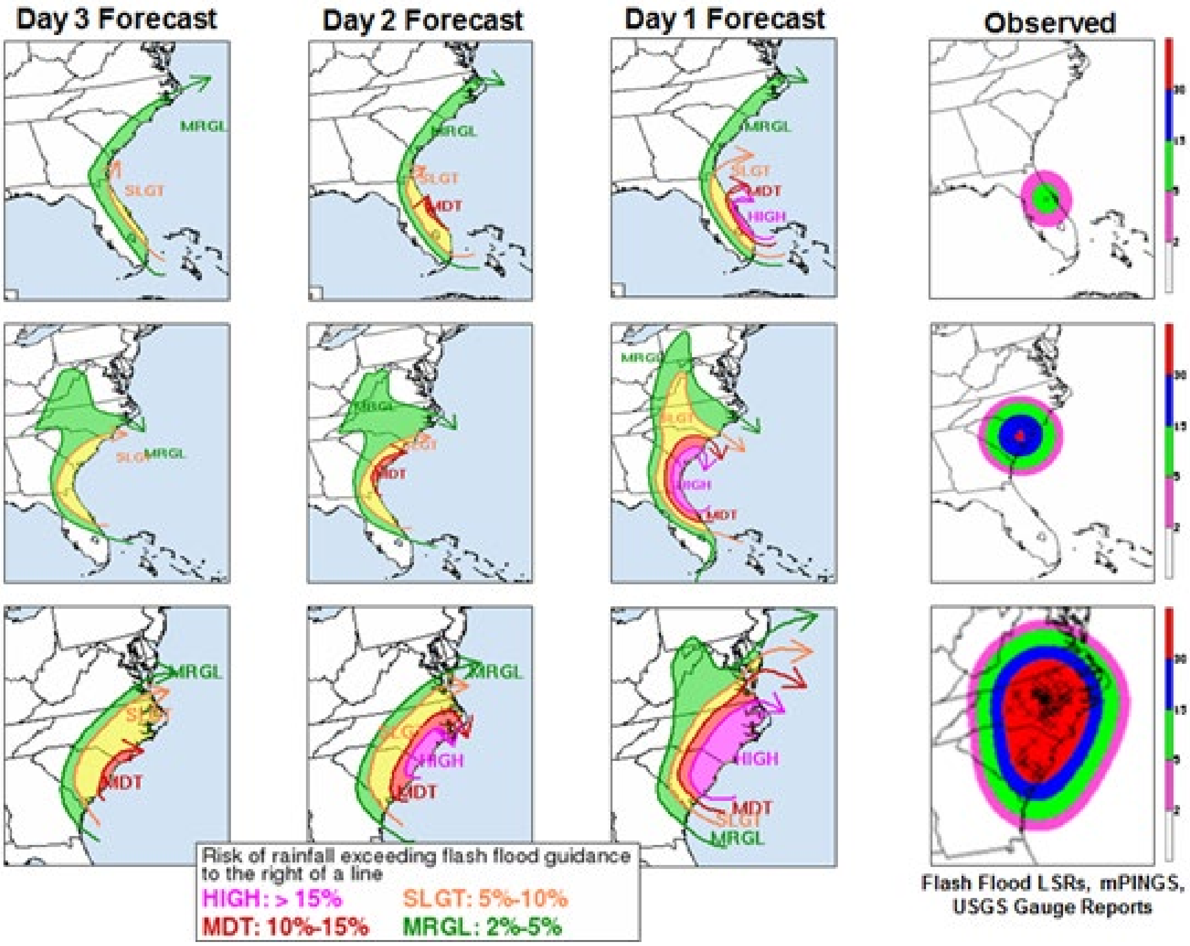


Figure 16. Excessive Rainfall Outlooks for Hurricane Matthew with plotted observations on the right panels from Flash Flood Local Storm Reports (LSRs), Meteorological Phenomena Identification Near the Ground (mPING), and USGS Gauge Reports.

3.2.3 Severe Weather

In addition to the high winds of the TC's primary circulation, storm surge and rainfall, tropical cyclones can produce tornadoes away from the TC's inner core. Storm Prediction Center (SPC) provides probabilistic outlooks for severe weather, including those from tornadoes. These are created subjectively by forecasters based on model forecasts and other objective guidance. The day 1 severe weather outlook provides separate probabilities for large hail, high winds and tornadoes, while the longer-range outlooks provide a single probability for all severe weather hazards.

An HFIP goal is to improve the 24-h probabilistic tornado outlook product by 10% (Goal 4.8 in Table A.2.4.4). To date, there has been limited focused research on the prediction of TC tornado likelihood, and hence SPC probabilistic outlooks for landfalling TCs have historically been rather smooth with limited resolution or sharpness. In order to achieve Goal 4.8 and to make additional forecast improvements beyond this goal, focused operationally relevant research will be needed, including:

- Establishing the baseline reliability of SPC tornado probabilities associated with landfalling TCs;
- Determining the climatology of tornadic versus non-tornadic TC supercells, and if observable environmental differences exist and can be reliably forecast;
- Assessing the operational applicability of CAMs in TC environments to resolve small-scale supercells and provide proxy tornado indicators; and
- Exploring the role of specialized hurricane CAMs in the TC tornado forecast challenge.

3.2.4 Sustained Winds and Gusts

Tropical storm and hurricane watches and warnings are issued on the basis of the wind threat, having traditionally been determined from the official track, size and intensity forecasts, coupled with the forecaster's subjective determination of uncertainty, as well as non-meteorological factors. While existing TC wind speed probabilities provides objective guidance for the placement of wind watches and warnings, the probabilities rely heavily on climatological forecast error distributions and are only weakly situationally dependent. Incorporating dynamical model ensemble information could improve the utility of the wind probabilities for watch/warning placement. Doing so is an HFIP objective.

Probabilistic products will be evaluated for a set of points located along and just inland of the U.S. Gulf and Atlantic coasts, with the points deemed to have been affected by the winds of the various thresholds determined from NHC best-track positions and wind radii. Wind speed probabilities at both two days and five days will be evaluated (Goal 4.3).

TC wind warnings, intensities, and wind probability products all refer to sustained (1-min mean) surface (10-m) winds occurring over an unobstructed exposure. Wind gusts are not considered in any of these products, and furthermore are analyzed and forecast only in the most rudimentary manner. In addition, the ability of intensity forecast models to provide wind gust guidance is not well known. A second wind-related goal (Goal 4.4) is to perform an evaluation of dynamical model wind-gust forecasts as a first step towards development of new wind-gust hazard products.

3.3 Enhance communication of risk and uncertainty

A review of societal impacts was conducted and synthesized in a 2005 report that outlined research required to better convey risk and uncertainty associated with TCs. Such recommendations included supporting research on the following areas: warning process, user impacts, decision-making, risk quantification and perception, behavior response, evacuation processes, and economic impacts. Following this report, a number of other research efforts, including a report entitled *Assessing Current Storm Surge Information from the Public Perspective*, spurred NOAA toward a concerted effort to better integrate the social and behavioral sciences into its products, information, and services.

HFIP recognized and acted upon the need to better convey uncertainty and risk information to reduce loss to life and property. For example, it supported extensive social and behavioral research on the storm surge threat.. That research ultimately resulted in the Potential Storm Surge Flooding Map (Fig. 17), which depicts the potential storm surge flooding that a TC could produce, and the Storm Surge Watch/Warning Graphic, which highlights areas that have a significant risk of life-threatening storm surge. The Potential Storm Surge Flooding Map became operational during the 2016 Hurricane Season, though the social and behavioral science research began in 2010. The successful development and operational implementation of this new product was largely due to a five-year iterative process between social and behavioral scientists and the operational forecast community. The Storm Surge Watch/Warning Graphic became operational in 2017.

Building off the 2008 HFIP Strategic Plan, which highlights the importance of an increased understanding of how TCs impact society, NOAA aims to further incorporate social and behavioral sciences to ensure its products and services best inform decision makers and the public.

3.3.1 Vision

HFIP plans to build upon the success in the Potential Storm Surge Flooding Map and Storm Surge Watch/Warning Graphic and more fully incorporate the social and behavioral sciences into the development and/or assessment of the TC products, information, and services for all hazards. These HFIP efforts will also look at successes gained by the Forecasting a Continuum of Environment Threats (FACETs) initiative. FACETs is an NWS-Oceanic and Atmospheric

Research (OAR) effort to transition NOAA hazard information services from a deterministic framework to a probabilistic framework, with social and behavioral sciences fully integrated to ensure the development of easy to understand products that effectively communicate risk and impacts. The heart of FACETs is an iterative, collaborative physical, social and behavioral research effort to reach the end goal of a full suite of products that are optimized to result in the maximum benefit to society for all hazards.

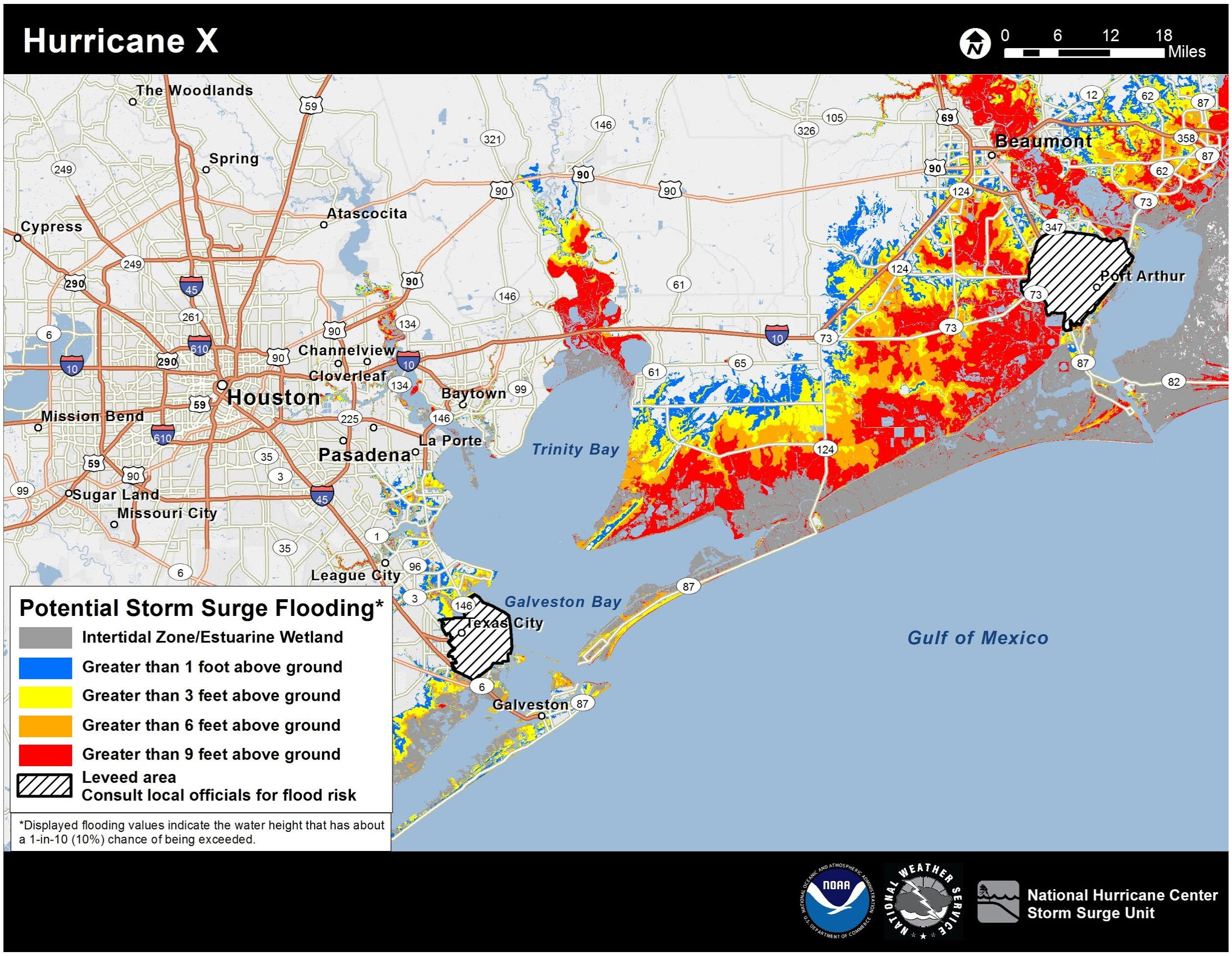


Figure 17. Static example of the Potential Storm Surge Flooding Map

HFIP will support the use of social and behavioral science methodologies for TC hazards (e.g., storm surge, rain, associated tornadoes, gusts as well as sustained winds) by assessing partner and stakeholder information needs, and characteristics of that needed information including physical science (risk, confidence, uncertainty, etc.), technological (formats, interactivity) and messaging (graphics, interactive, apps, etc.) parameters to effectively modernize the TC product suite (including those from NHC, CPHC, WFOs, River Forecast Centers (RFCs), and NCEP). Over the next ten years, through the use of social and behavioral science to guide product modernization, NOAA will not only collect a baseline understanding of the TC product suite, but will also assess broader information needs and technologies to help guide the necessary upgrades, enhancements, and new products the TC suite needs to modernize.

3.3.2 Approach to Incorporate Risk Communication For More Effective Products

Section 104 of the Weather Research and Forecasting Innovation Act of 2017 instructs NOAA to maintain a project to incorporate risk communication research in the design and communication of its products. By 2021 NOAA will complete a baseline understanding of partner and stakeholder needs relating to the TC product suite. By 2023, through social and behavioral science research, NOAA intends to improve communicating the forecasted risks by transitioning 2-3 TC hazard guidance products per year and, by 2028, modernize all products in the TC product suite.

The envisioned pathway in using social and behavioral science to modernize all products in the TC product suite includes the following:

- Understanding intended messages of the TC product and communication suite to help frame measurable social and behavioral science research objectives;
- Understanding partner and stakeholder needs through the use of literature reviews and use studies, to include (but not limited to) interviews, focus groups, surveys, and web data:
    - Understanding how NWS partners and stakeholders *currently* interpret and use the current TC product suite, including probabilistic information, for decision-making;
    - Identifying NWS partner and stakeholder *future* TC information requirements at various timescales (i.e., weeks, days, minutes prior to an event) and characteristics of that information, including physical science (risk, confidence, uncertainty, etc.), technological (formats, interactivity), and messaging (graphics, interactive, apps, etc.) parameters, to effectively modernize the TC product suite;
    - Evaluating which TC products and messages weather information providers use, modify, and show to public audiences in order to identify what changes to the TC product suite are needed;
- Use NOAA's Hazardous Weather Testbed (HWT) and Operations Proving Ground to evaluate current and future products and operational paradigms in a naturalistic environment that includes NWS forecasters and core partners;
- Iterating between social and behavioral sciences and the operational community to develop and/or enhance new and/or current TC products;
- During TC product development, incorporating feedback and conducting product testing with NWS partners and stakeholders on proposed changes to the TC product suite utilizing social and behavioral science;
- Including America's Weather Industry (AWI) in the shared responsibility of effective conveyance of risk and uncertainty; and
- Improving outreach and training on new TC products, services, and messaging.

Table A.2.4 in Appendix, contains the specific goals and metrics for product assessment and improvement through the use of social and behavioral science, to enhance the communication of risk and uncertainty. Priorities and objectives in achieving these goals are listed in Table 2, below.

3.4 Support Dedicated High Performance Computing Allocation

3.4.1 History of dedicated computing allocation in support of HFIP

Since 2009, NOAA dedicated \$32M (~\$4M per year) to HFIP exclusively for increasing allocation and the maintenance of research and development (R&D) high performance computing for hurricanes. From 2009 through 2016, annual expansions, by way of incremental hardware procurements, resulted in the Jet High Performance Computing (HPC) machine ultimately totaling 45,000 processors and 4.4 Petabytes of storage.

HFIP was organized around two "streams": Stream-1 (operational hurricane model development) and Stream-2 (development, testing and evaluation of experimental models and variants of the operational models, evaluation of new techniques and strategies for hurricane model forecast improvements prior to testing for possible operational implementation). Since a larger community, including key academic partners, are usually involved in hurricane model testing and evaluation, HFIP's strategy was to develop a dedicated HPC capacity at NOAA's Earth System Research Laboratory (ESRL) in Boulder, Colorado, which led to the establishment of HFIP's Jet HPC system dedicated solely to hurricane research, including a 3-month demonstration phase during hurricane season. Part of the success of HFIP may be attributed to the provision that dedicated large computer resources to it. Almost all high-resolution HWRF deterministic and ensemble model advancements, and testing and evaluations are done on the Jet HPC system. Jet resources are also currently used for the development, testing and evaluation of the FV3-based NGGPS for hurricanes. The availability of dedicated computing resources for hurricanes eliminates competition with other high-priority computing needs across NOAA's broad programs.

3.4.2 Importance of "Stream 2" research computing allocation in R2O

A major component of NOAA's HFIP is the support of the Stream-2 real-time experiments. Since the level of computing necessary to perform such a demonstration is larger than can be accommodated by current operational computing resources, HFIP uses the Jet computing facility. A major component of Stream 2 (also known as the Demonstration Project) is an Experimental Forecast System (EFS) that HFIP runs each hurricane season. It is used 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 after each season to identify techniques that appear particularly promising to operational forecasters and/or modelers. These potential advances can be blended into operational implementation plans (i.e., Stream 1 activities) or further developed outside of operations within Stream 2. Stream 2 activities have been responsible for all major R2O transitions related to the HWRF. HFIP intends to continue Stream 2 activities, as long as an adequate dedicated computing allocation can be secured.

3.4.3 Moving forward, securing dedicated computing allocation

Continued use of dedicated HFIP computing by re-capitalizing the previously dedicated, but aging, NOAA R&D supercomputer (Jet) ($6M per year) is recommended. Ongoing support will continue to be required ($2M per year) for annual operation and maintenance (O&M). The dedicated NOAA HPC capacity will be used to further improve hurricane predictions, placing priority on development, testing and evaluation of the HAFS system and high-resolution ensembles.

NAO 216-110 defines HPC as the unified system for solving NOAA's largest computational problems, composed of supercomputer systems and associated communications, analysis, visualization and storage systems, and application and systems software with all components well-integrated and linked over a high-speed network. To conduct the required R&D over the next 5 years approximately 80M dedicated computational core hours per month and 13,600 TB of archival space are required starting in 2019. Increases to approximately 184M dedicated computational core hours per month and roughly 32,600 TB archival space will be needed by 2023 (see Table 1).

Table 1: High Performance Computing Requirements (FY18 - FY23)

Compute	**(core hr/month)**	FY 2018	FY 2019	FY 2020	FY 2021	FY 2022	FY 2023	Notes
Hurricane	Prediction (R&D)	63.06M	79.92M	98.36M	150.4M	166M	184.16M	RDHPC
Hurricane	Operations (NCEP)	1.54M	1.85M	2.21M	2.66M	3.20M	3.84M	WCOSS
Storm surge	NHC/SLOSH/ SWAN	4.8M	6.6M	8.4M	10.2M	12.0M	13.8M	RDHPC
	MDL	0.36M	1.58M	2.02M	3.32M	6.85M	7.09M	RDHPC
	NOS		0.45M	0.45M	0.55M	0.55M	0.71M	RDHPC
Disk	**(TB)**							
Hurricane	Prediction	9,820	13,625	17,365	20,155	27,600	32,600	RDHPC
Hurricane	Operations (NCEP)	800	960	1152	1383	1660	1990	WCOSS

Storm surge	NHC/SLOSH/ SWAN	80	110	140	170	200	230	RDHPC
	MDL	32	44	56	68	80	92	RDHPC
	NOS	6	88	91	101	104	140	RDHPC

Tape	**(TB)**							
Hurricane	Prediction	32800	46100	61100	72217	99250	115257	RDHPC
Storm surge	NOS	80	82	82	92	92	128	RDHPC

Table 1. Depiction of dedicated computational (core hours/month) and archival space (TB) required to execute the research, development, and operational requirements of the HFIP.

Specific priorities and objectives for securing a dedicated HPC allocation may be found in Table 2, below.

3.5 R2O Enhancements

The transfer of research to operations requires robust interaction between the research and operational community as well as a strong interface with the user community. It also requires a healthy infrastructure for the transition. That includes resources and processes for evaluation and demonstration, operational implementation and operations and maintenance. The strategy revolves around the notion that to accomplish R2O, we must support "Operation to Research (O2R)" needs to ensure that the transitions are successful. This strategy is imperative to realize the benefits of targeted research in operations. HFIP will continue to accelerate the transition of research and new observing systems and platforms to operations by adhering to the NOAA Administrative Order 216-105B[19] "Policy on Research and Development Transitions" in assessing the maturity of R&D projects from R2O via Readiness Levels (Fig. 18), developing and maintaining transition plans, and utilizing NOAA's Testbeds and Proving Grounds.

HFIP will establish R2O activities within the Joint Hurricane Testbed (JHT), broadening the JHT charter by increasing support to include HAFS and social and behavioral science R&D. HFIP proposes to develop interactions between the JHT, Hydrometeorological Testbed (HMT), HWT, DTC, Joint Center for Satellite Data Assimilation (JCSDA), and Quantitative Observing System

---

[19] http://www.corporateservices.noaa.gov/ames/administrative_orders/chapter_216/Handbook_NAO216-105B_03-21-17.pdf

Analysis Project (QOSAP) to address HAFS R&D. As discussed in section 3.3, HFIP will also be closely linked with R2O activities taking place as part of the FACETs initiative. HFIP will collaborate with the FACETs Working Groups in NWS and OAR to ensure that physical and social science research conducted under HFIP is consistent with FACETs-related efforts, and is integrated into efforts taking place in other NWS service program areas.

R2O will be enhanced by way of the following:

1. Following the guidelines of NOAA administrative orders (NAO) 216-105B, develop transition plans identifying and documenting requirements, develop technology and evaluating prototype options, conduct a demonstration and evaluation, develop and operationally implement the system, and secure sustained operations and maintenance.

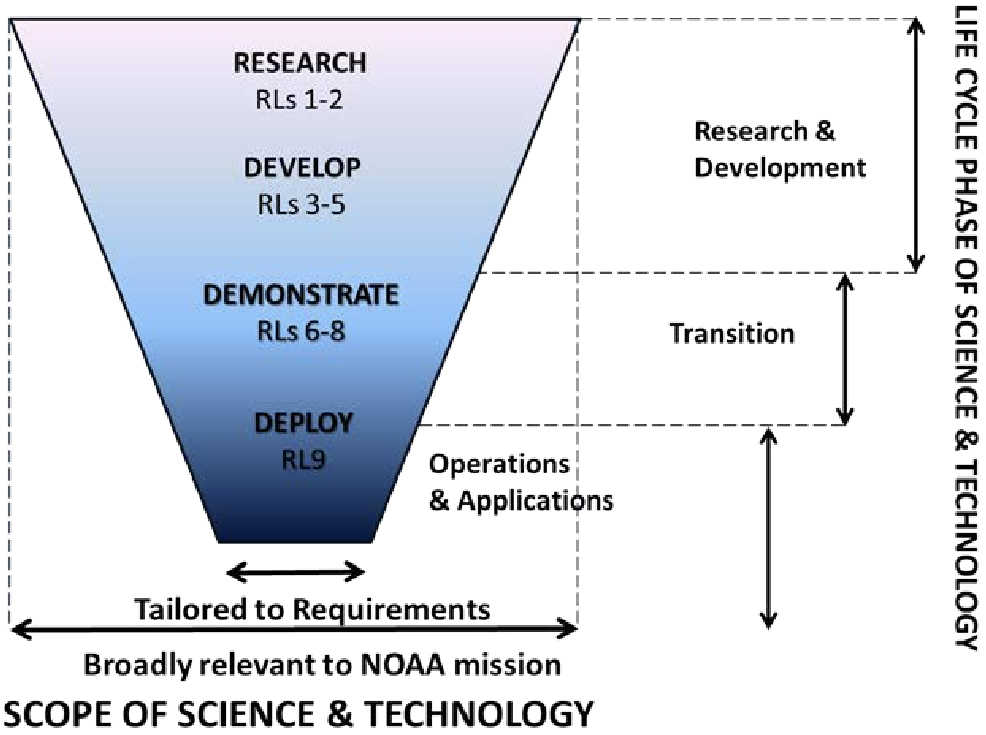


Figure 18. The NOAA transition funnel

2. Broaden the JHT charter and increase support to put more emphasis on HAFS research & development issues including R&D for specific HAFS improvements, to evaluate models, to develop new tools for forecasters, and to test the utility of observing systems for operational platforms.
3. Broaden the JHT charter and increase support to put more emphasis on social and behavioral science R&D issues including R&D for specific improvements to TC hazard guidance products, to develop new products and services for forecasters,

and to test the utility of new products, information and services for operational platforms.

4. Take advantage of and expand upon existing testbeds and proving grounds in the following manner:
   - Develop interactions between JHT,HWT, HMT, DTC, JCSDA, and QOSAP to address HAFS issues.
   - Expand DTC activities to support making HAFS available to the research community by:
     - Providing the current version of operational model system in the repository, developing documentation and training material for HAFS model system
     - Supporting and conducting workshops on operational HAFS model system
     - Supporting research grants to test and improve HAFS model system capability, particularly to investigate model predictability, trade-offs between decreased grid-spacing and ensembles, impacts of improved physics packages/parameterization, impacts of different model cores, ensembling approaches, and improved numerical techniques
     - Facilitating the transition of research developments into operational model system
   - Expand the JCSDA and QOSAP activities to support HAFS improvements through R&D (including OSSEs) to address data assimilation issues related to increasing diversity in remote sensing (e.g., satellites, radar) and in-situ data collection platforms (e.g., aircraft, buoys, etc.)
   - Utilize HWT, HMT and the Operations Proving Ground (OPG) to evaluate current and future products and operational paradigms in a naturalistic environment that includes NWS forecasters and core partners.

Specific priorities and objectives for the enhancement of R2O activities may be found in Table 2, below.

3.6 Broaden expertise and expand interaction with external community

NOAA recognizes the broad scope of the scientific challenges associated with understanding and predicting hurricanes. Addressing these challenges and improving the forecasts of TC track and intensity will therefore involve significant interaction with the external community. In order to broaden and access the necessary expertise, HFIP envisions enhancing its TC-related research approach and expertise supported through its annual hurricane research program activities through the following actions:

- Maintain a Scientific Review Committee composed of representatives from NOAA and the external R&D community to provide insight on the scientific relevancy of HFIP activities with respect to the current state of the science. Note: Federal agencies may not request or accept consensus opinions, advice, or recommendations from the Science Advisory Committee. Instead, Science Advisory Committee members will be invited to provide their individual insight on the scientific direction and merits of HFIP activities;

- Grants/Contracts to support research and technology development and training activities for external community at NOAA operational facilities;
- Outreach and education for operational modeling development activities (e.g., workshops, conferences, publications);
- Continue routine workshops and meetings for internal and external reviews on HFIP priorities and progress;
- Establish a process to increase operational numerical modeling and TC expertise through workshops, symposia, conferences, visiting scientist;
- Establish a Technical Advisory Committee composed of representatives from NOAA and the external R&D community to advise the HFIP on standardized metrics and test cases for testing and acceptance of HAFS updates;
- Exploit the resources of existing mechanisms (e.g., via NOAA Science Advisory Board, Earth Prediction Innovation Center (EPIC), Unified Forecast System (UFS) Strategic Implementation Plan (SIP), HFIP Socio-Economic Working Group, and NOAA's internal Councils) to evaluate the progress and direction of the HFIP internally and with the external community. This evaluation is intended to include annual HAFS R&D workshops focused on reviewing our progress, determining applications and techniques for R2O and O2R, and evaluating our goals and metrics;
- Establish a mechanism to promote interaction of outside research community on HAFS improvements (e.g. annual HAFS research review at the Office of the Federal Coordinator for Meteorology (OFCM) Interdepartmental Hurricane Conference);
- Support and advise NOAA leadership on interactions with the broader community with respect to TC forecast improvement issues; and
- Engage, align and take advantage of ongoing efforts of other R&D programs.

Specific priorities and objectives for broadening expertise and expanding interaction with the external community may be found in Table 2, below.

4. Strategy Priorities and Objectives

Table 2 includes the priorities and objectives for the new HFIP Strategic Plan. Each set of priorities are categorized by implementation strategy. The objectives described are interdependent. If the short-term investments are not made, it will take longer to reach the outcomes in the next tier.

The investments and timelines for each activity listed in Table 2 are being developed through

NOAA and are contingent upon resource availability. A fully supported plan is designed to achieve the following key strategies for NOAA:

- Advance operational hurricane analysis and forecast system (HAFS)
- Improve probabilistic guidance
- Enhance communication of risk and uncertainty
- Increase HPC Allocation
- R2O Enhancements
- Broaden expertise and expand interaction with external community

Table 2: Priorities and Objectives for Strategy Implementation

Strategy: Advance an operational hurricane analysis and forecast system (HAFS)	
Priorities	**Objectives**
▪ R&D for HAFS to advance deterministic and ensemble prediction capabilities to seven days ▪ R&D for fusion of modeling, data assimilation and observations to produce an analysis of record ▪ R&D for statistical post-processing to extract guidance and uncertainty information	▪ Annual upgrades to HAFS implemented operationally prior to transition to UFS ▪ Targeted data assimilation improvements for HAFS prior to transition to UFS ▪ Produce forecast track and intensity guidance to seven days based on HAFS ▪ Complete demonstration of impact of forecast performance using the HAFS based on UFS ▪ HAFS based on UFS implemented operationally
Strategy: Improve probabilistic guidance	
Priorities	**Objectives**
▪ R&D for calibrating HAFS guidance ▪ Incorporate dynamically-based uncertainty into hazard models and products ▪ Increase lead time of real-time storm surge guidance and forecasts from two to three days before the arrival of surge and wind hazards	▪ Test and evaluate high resolution ensemble model for use in generating probabilistic guidance on track, intensity, and structure for use in improved storm surge and other hazard guidance ▪ Update SLOSH's parametric wind model, incorporate wind structure information into P-Surge ensembles, and move from a statistical to a dynamical ensemble

▪ Expand real-time storm surge forecasting capabilities to areas outside of the contiguous United States, including the Storm Surge Watch / Warning ▪ Create a Probabilistic Tropical Quantitative Precipitation Forecast (QPF) product using HAFS output ▪ Create probabilistic tornado guidance for TC events using HAFS output	▪ Couple operational storm surge model with wave model to account for steep-sloped areas, while developing SLOSH Maximum of the Maximum (MOMs) / Maximum Envelope of Water (MEOWs) for OCONUS ▪ Utilize HAFS probabilistic output to update existing Probabilistic QPF for TC events ▪ Utilize HAFS probabilistic output to update existing 1-day tornado probabilities for TC events ▪ Perform a systematic evaluation of dynamical model forecasts for wind gusts associated with TCs
Strategy: Enhance communication of risk and uncertainty	
Priorities ▪ Evaluate TC products for the effective communication of risk ▪ Determine operationally viable ideas collected from NWS partners and stakeholders ▪ Iterate between social and behavioral scientists and operational community to develop and/or enhance new and/or current TC products	**Objectives** ▪ Conduct baseline assessment of NWS TC product suite for effective communication of risk and prioritize those products informed by social and behavioral science ▪ Modernize TC product suite based on the baseline assessment and prioritization ▪ Collect longitudinal data from users to ensure increased efficacy as product modernization occurs ▪ Gather NWS partner and stakeholder feedback on product changes
Strategy: Increase HPC Allocation	
Priorities ▪ NOAA R&D and operational computing to support HAFS development ▪ Sustain modeling and software engineering expertise ▪ Match technological innovations (i.e. keep pace with new technologies)	**Objectives** ▪ Staffing and computing infrastructure established to test and evaluate HAFS improvements and generating probabilistic guidance ▪ Evaluate and implement a cloud-computing approach to test and evaluate HAFS improvements and generating probabilistic guidance

	▪ Increase HPC capacity each year
Strategy: R2O Enhancements	
Priorities ▪ Accelerate R2O using NOAA Testbeds by following NOAA’s best practices for promoting readiness levels (RLs) ▪ Develop a process to prioritize research targeted for operational improvements	**Objectives** ▪ Staff and infrastructure established for enhanced R2O ▪ Involvement with external community for modeling R&D through JHT, HMT, HWT, DTC, and JCSDA
Strategy: Broaden expertise and expand interaction with external community	
Priorities ▪ Collaborate with key agencies and professional organizations to ensure effective training and outreach ▪ Collaborate/coordinate with social and behavioral sciences ▪ Work with commercial and other weather-oriented organizations to disseminate and/or develop improved NWS TC products conveying risk or uncertainty	**Objectives** ▪ Maintain advisory committees and community workshops ▪ Maintain a visiting scientist program at research and operational centers ▪ Conduct workshop with social, behavioral, and physical scientists to assess operational viability of partner and stakeholder needs ▪ Re-invigorate the grants program

5. Requirements for Success

The success of the next phase of HFIP in reaching the four goals outlined in Section 2 requires sufficient funding to support the activities outlined here. NOAA made significant progress toward achieving HFIP goals in the first 5-6 years of the program. Starting in FY 2015, however, NOAA dedicated fewer resources to HFIP due to competing budget priorities across the agency. This slowed the rate of progress towards HFIP goals by restricting the capacity to test and evaluate new research and delaying transition of potential new analysis and forecast applications into operations. The lower funding levels also hindered engagement with the academic community that dramatically slowed model improvements. With the passage of the Weather Act by Congress in 2017, NOAA is now dedicated to reinvigorating HFIP to move towards meeting the requirements of the Act. Resource requirements are still being considered within the agency and will be reflected in NOAA's future year budget requests. The FY18 Appropriations and "Hurricane Supplemental" bill funding are being allocated to support HFIP. Some key requirements and dependencies associated with HFIP priorities/strategies are detailed below:

1. **High-performance computing:** Support for both research and operational HPC support is critical to accelerating HFIP R&D and achieving the goals laid out in the Weather Act. NOAA's R&D HPC capacity is currently much lower than that available for operations. An increase in R&D capacity as well as sustained support for operations and maintenance of current systems will be essential to meeting the goals and timelines of HFIP going forward.
2. **HAFS development:** The prototype HAFS will take advantage of ongoing HWRF developments to migrate towards an advanced analysis and forecast system with cutting-edge research on modeling, physics, data assimilation, and coupling to earth system components for high-resolution TC predictions. This will be done within the outlined NGGPS/SIP objectives of the FV3 Dynamical Core-based UFS. Current model guidance from the HWRF only extends to five days, so in order to produce forecast guidance to seven days HFIP must continue to closely collaborate with the NGGPS Program.
3. **Improved observations:** To succeed, HFIP will need to incorporate improved observations from both satellites and aircraft. This will depend on the observational platforms themselves (e.g., NOAA satellites (i.e., GOES-R Series and JPSS satellites); and take advantage of domestic and international satellite observations (i.e., NASA Earth Observatory, Japanese, European, French satellites) as well as data assimilations advancements. These activities will require sustained, dedicated resources in order to meet HFIP timelines.
4. **Development of probabilistic guidance:** This is a new strategy driven by the need for enhanced ensemble forecasts to provide probabilistic guidance on storm track, intensity, and storm size to quantify uncertainty for all TC hazards. Additional support will be necessary to accelerate those activities in order to meet HFIP timelines. HPC capacity will be essential to generating large enough ensemble forecasts representative of the true uncertainty in the guidance. The current HWRF ensemble contains 20 members at the highest spatial resolution of three kilometers. R&D is needed to evaluate whether this number of ensembles is sufficient; however, current research suggests twice as many members are required to provide sufficient probabilistic guidance. Additional HPC support beyond current levels is critical to this R&D.
5. **Improved guidance on pre-formation and extended forecast skill from 5 to 7 days:** The former is a new goal driven by NHC's operational requirement to provide guidance, products and services in a 48-hour actionable time-frame for disturbances that may result in hazardous conditions to land areas before becoming a tropical storm. The latter is

driven by an ongoing HFIP science and R2O challenge. Additional resources will be essential to advancing the goals related to these activities. Any delays in implementing the HAFS and/or decrease in guidance skill will negatively impact this activity.

6. **Improved quantitative precipitation for landfalling TCs:** There are a number of NOAA activities related to our ability to accelerate progress in this area (Fig. 19). For example, some of HFIP's goals and objectives overlap with the NOAA Water Initiative,[20] the Consumer Option for an Alternative System to Allocate Losses (COASTAL) Act[21]. How well all of these activities can work together toward these common goals, taking advantage of allocated resources while not competing for support, will be a major challenge to the success of this effort.

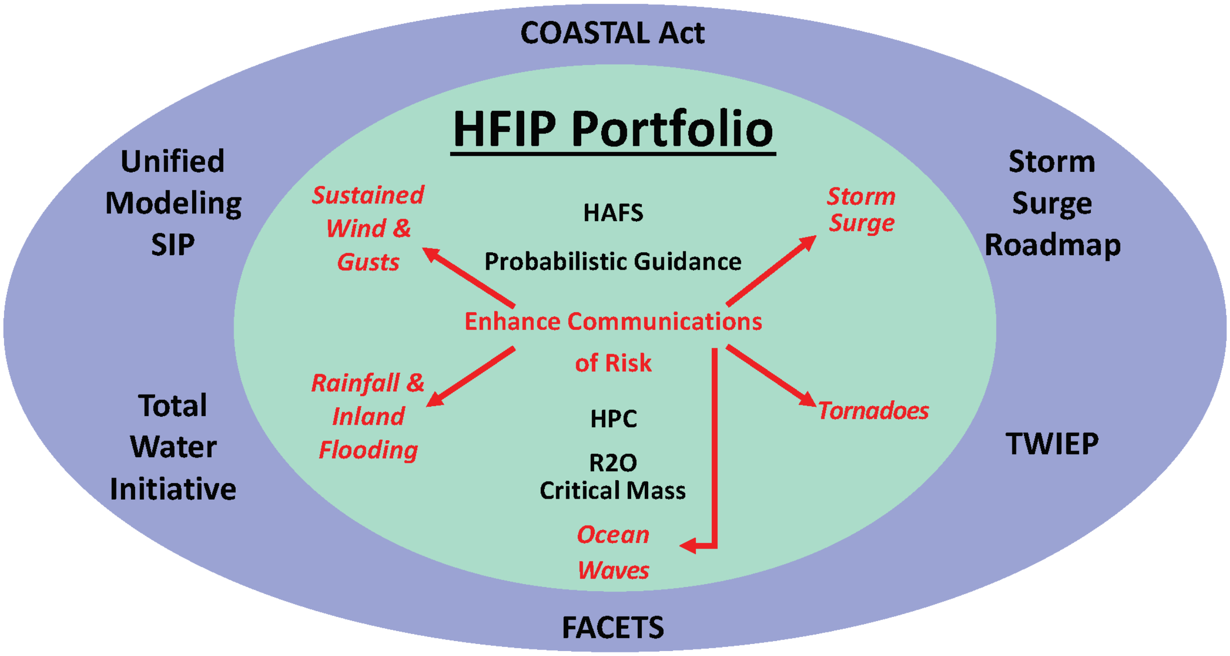


Figure 19. The HFIP portfolio. HFIP activities are within the green ellipse and the other NOAA initiatives that have overlapping interests are within the blue ellipse.

7. **Improved tornado guidance for landfalling TCs:** There are many NOAA activities related to our ability to accelerate progress in this area (Fig. 19). For example, some of HFIP's goals and objectives overlap with the Verification of the Origins of Rotation in Tornadoes EXperiment-Southeast (VORTEX-SE),[22] and the Tornado Warning Improvement and Extension Program (TWIEP) plan to address Section 103 of the Weather Act[23] on improved tornado guidance. How well all of these activities can work

---

[20] http://www.noaa.gov/explainers/noaa-water-initiative-vision-and-five-year-plan

[21] https://www.weather.gov/sti/coastalact

[22] https://www.nssl.noaa.gov/projects/vortexse/

[23] https://www.congress.gov/115/plaws/publ25/PLAW-115publ25.pdf

together toward these common goals, taking advantage of allocated resources while not competing for support, will be a major challenge to the success of this effort.

8. **Improved NWS TC products:** At this point, there is no dedicated funding allocated to the incorporation of social and behavioral science research to create a more effective TC hazard product suite. Without additional funding, NOAA will not be able to address this HFIP focus area. If provided with additional funding, the success of the ongoing HFIP-supported storm surge product development will serve as a model to follow in order to modernize the product suite for all TC hazards.
9. **Grant support for community engagement:** External partnerships augment HFIP's capacity to conduct the R&D necessary to address HFIP's physical, social, and behavioral science challenges. Investments and collaboration with partners (federal, academic, external, international), testbeds, and the community (e.g. federal-funding opportunities) are essential to the success of the project.

6. Summary

The HFIP has been highly successful over the past 10 years, reducing errors in track and intensity forecasts. However, much remains to be done. The next phase of HFIP will continue its mission to reduce impacts of TCs, especially loss of life and damage to property, through the implementation of key strategies designed to improve forecasts and warnings. NWS and OAR will continue to address existing science and R2O challenges by improving regional and global models, transferring promising innovations from research to operations, and collaborating with academia, the commercial weather industry, media, and the emergency response community to achieve the objectives outlined for the next phase of HFIP.

Appendix A

Specific Goals and Metrics

The *Weather Research and Forecasting Innovation Act of 2017* requires HFIP to improve three areas: forecasts of RI, storm surge products, and communication of the risks associated with TC hazards. In section A.1 of this Appendix, the current suite of TC forecast, warning, and hazard products is summarized to provide background for the discussion that follows in section A.2 about the specific HFIP goals, and the measures HFIP will use to assess progress toward those goals. Note that these goals concern not only the accuracy and reliability of TC forecasts and warnings, but also the use of those products to enhance mitigation and preparedness decisions by emergency management officials at all levels of government, the media, and by individuals.

A.1 Summary of Current Products

The NHC and CPHC produce numerous TC forecast products, including track and intensity forecasts out to five days, wind field structure forecasts (the maximum radial extent of winds of specific thresholds) for 34-kt and 50-kt winds out to 72 h and for 64-kt winds (beginning in 2018) out to two days. NHC and CPHC also issue two-day and five-day probabilistic TC genesis forecasts.

NHC and CPHC produce a standard set of text and graphical products every six hours.. Two of the better-known text products are the Tropical Cyclone Public Advisory (TCP) and the Tropical Cyclone Discussion (TCD). The TCP summarizes current storm attributes, lists coastal watches and warnings for wind and storm surge, describes the 5-day forecast in general terms, and also broadly describes the areas at risk for wind, storm surge, rainfall, and tornado hazards. The TCD provides the reasoning behind the forecast, as well as a subjective assessment of the forecast's level of uncertainty. When TCs pose a significant threat to land, a section of the TCD contains "Key Messages", a bulleted list of essential takeaways (many of them hazard related) aimed both at members of the media and the general public.

The TC "track forecast cone" graphic depicts the official track forecast and an uncertainty cone, as well as the coastal wind watches and warnings and categorical intensity forecast information. The size of the cone that surrounds the official track forecast is determined from the 67th percentile of the previous 5 years' official track forecast errors; thus, the cone size is constant over the course of a given hurricane season. The cone graphic also displays the extent of hurricane- and tropical-storm-force winds at the advisory time, but does not otherwise provide any specific information on TC hazards.

Probabilities that individual locations will experience sustained 34-, 50-, or 64-kt winds during the 5-day forecast period are provided with every advisory package; these probabilities are determined from Monte Carlo simulations based on the official TC forecast, past official track and intensity forecast errors, and climatological wind field information. The wind speed probabilities are available in text, graphical, and gridded formats; in addition, they provide the underlying data for graphics that display for any given location the most-likely and earliest-reasonable times of arrival of sustained 34-kt winds, a key threshold for evacuation and preparation timing.

Probabilistic storm surge guidance is also generated for the U.S. Atlantic and Gulf coasts during the watch/warning time period (beginning typically 48-60 h prior to landfall); this guidance is also based on the official TC forecast along with historical track and intensity forecast errors, although the methodology to create the statistical sample of possible outcomes is not as sophisticated as the Monte Carlo technique used to compute the wind speed probabilities. Through a process known as P-Surge, the statistical sample of TC outcomes is then coupled with the SLOSH model to generate the probabilistic surge guidance. P-Surge is the primary objective input for the storm surge watches and warnings issued for the Atlantic and Gulf coasts and for the Potential Storm Surge Flooding Map, a graphic that displays for any individual location a reasonable worst-case scenario for storm surge flooding (specifically, it displays the 10% exceedance inundation height).

It’s worth emphasizing that the current probabilistic products for storm surge are based entirely on historical error characteristics, rather than situationally specific uncertainty, and that the wind speed probabilities only vary situationally in a very crude manner (dynamical track model spread is used to draw the statistical sample from three fixed terciles of error characteristics).

National-level rainfall products for TCs are provided by the WPC. These include deterministic and probabilistic QPFs, and an excessive rainfall outlook for the CONUS that indicates the likelihood of rainfall exceeding flash flood guidance. Watches and warnings for floods or flash floods are issued by NWS WFOs, with river flood guidance provided by RFCs.

National-level tornado information is provided by the SPC, which issues probabilistic severe weather outlooks, mesoscale convective discussions, and tornado watches for the CONUS. Tornado warnings are issued by WFOs.

WFOs issue a variety of TC-related products and hazard-specific information that build on the official NHC or CPHC forecast information and the probabilistic products described above, but which are tailored for smaller geographical areas. The primary WFO text product is the Tropical Cyclone Local Watch/Warning Product (WFO TCV), a segmented Valid Time Event Code (VTEC) product with each segment corresponding to an individual NWS forecast zone. Each segment lists the TC wind and storm surge watches/warnings in effect, forecasted conditions for wind, storm surge, flooding rain, and tornadoes, and their potential impacts. The product is generated from local gridded forecast information and national guidance.

A second WFO text product is the Hurricane Local Statement (HLS), preparedness product that conveys a succinct message on land-based local impacts from a TC. When paired with the WFO TCV the two products provide a complete, localized TC hazard and impact forecast.

Hurricane Threats and Impacts (HTI) Graphics are issued by WFOs when TC wind or surge watches or warnings are in effect within a WFO’s area of responsibility. Through the use of probability data, color-coded HTI graphics depict the potential conditions to protect against with accompanying descriptions of potential impacts needed for effective preparations. The HTI graphics are based on the latest forecast for specific locations while also including a reasonable safety margin to account for forecast uncertainty.

Selected NWS probabilistic products are summarized in Table A.1. The three focus areas of the next phase of HFIP will provide new capabilities to improve these products through more accurate forecasts, more sophisticated measures of uncertainty, and a comprehensive evaluation of the form and content of these products.

Table A.1. Summary of selected NWS probabilistic products.

Product	**Inputs**	**Source of Uncertainty Information**
Wind Speed Probabilities	NHC/CPHC official forecast track and intensity; climatology/persistence model for wind structure.	Historical official track and intensity errors; climatological variability for wind structure; dynamical model spread terciles.
Storm Surge Probabilities	NHC official forecast of track and intensity	Historical official track and intensity errors; Climatological variability for radius of maximum wind.
Tornado Probabilities	SPC Forecaster	Subjective determination by forecaster
Excessive Rainfall Probabilities	WPC QPF, Flash Flood Guidance, WPC forecaster	Subjective determination by forecaster
Probabilistic QPF	WPC QPF “Most Likely Outcome”, PQPF ensemble, PDF generation	Multi-model ensemble information

A.2 Detailed Goals and Metrics

Section 2 defined the four broad goals (restated below with somewhat greater precision) for the next phase of HFIP. The first three are directed toward improving the forecast guidance of TC track, intensity, wind structure, and genesis, while the fourth goal is directed toward improvements in hazard prediction and risk communication. The four goals and the metrics that will be used to track progress towards their achievement are discussed in sections A.2.1-A.2.4 and presented in Tables A.2.1-A.2.4.

A.2.1 Improve forecast guidance for TC track, intensity, and structure

Guidance for track, intensity and structure forecasts is obtained from dynamical models (both global and regional), statistical models, and statistical post-processing of model output. The NHC/CPHC official track forecast is expressed as the latitude and longitude of the storm center from the initial (analysis) time out to five days. Intensity is forecast in terms of the TC's maximum 1-min sustained surface (10-m) wind, and is also given out to five days. Official track and intensity forecasts, as well as those from the guidance models are verified against the NHC/CPHC "best track", a post-storm analysis of all available observations that constitutes the official historical record for the TC.

Official forecasts for the maximum radial extent of 34-, 50- and 64-kt winds are given for each of the four quadrants surrounding the storm. Although NHC provides a post-storm analysis of wind radii as part of the best track, they do not formally verify the wind structure forecasts, since the data available for constructing the best track are sparse and even best-track wind radii have large uncertainties. Best-track radii do have improved reliability when reconnaissance aircraft data are available, however. In a relative sense, the 34-kt best-track radii have less uncertainty than those for 50 kt or 64 kt, in part because satellite scatterometers can be used to estimate the extent of 34-kt winds. As a result, HFIP has chosen to focus on an evaluation of the 34-kt wind structure forecasts, recognizing the inherent uncertainty in the available ground truth.

Another wind structure parameter important for storm surge and wind speed probability (WSP) models is the RMW. NHC/CPHC provide an initial estimate of the RMW but do not forecast this parameter. Lacking an RMW forecast, storm surge and WSP models use statistical methods to estimate the RMW and its forecast uncertainty, although these methods do not always properly represent the evolution of storm structure or its uncertainty. Because of its relevance to improved hazard prediction, the RMW has been chosen as an additional structure parameter in HFIP forecast improvement assessments.

Progress toward meeting the HFIP track, intensity, and structure goals will be assessed using Atlantic basin forecasts, primarily because the more frequent availability of aircraft reconnaissance data significantly enhances the quality of Atlantic best-track analyses relative to those from other basins. Although assessed using Atlantic basin forecasts, new and improved modeling techniques are also expected to benefit forecasts made for other global basins. The period 2015-2017 was chosen as the baseline period to mitigate against variations in forecast difficulty from season to season; this three-year period was also an active one in the Atlantic that featured a wide variety of storm tracks, an ample supply of intense hurricanes, and several notable U.S. landfalls.

HFIP projects are primarily directed toward improving the guidance models used by operational forecasters; it is therefore reasonable to define HFIP metrics in terms of the performance of the guidance (rather than, for example, the performance of the NHC Atlantic basin official TC forecast). Furthermore, as noted above, not every parameter being evaluated by HFIP is currently forecast by NHC. History tells us, however, that as the guidance models improve, NHC/CPHC official forecast improvements naturally follow. Consistent with this notion, HFIP track, intensity, and structure metrics will be based only on models available to forecasters in time to meet operational deadlines – this means that only interpolated, or "early" versions of the models will be used to construct HFIP baselines and metrics.

Table A.2.1 summarizes the five HFIP Goal 1 metrics that will be used to measure improvements in TC track, intensity and structure forecasts. These five metrics are discussed individually below.

A blend or consensus of several top-performing track models usually has lower forecast errors than any individual model, and the performance of the NHC official forecast tends to closely track the performance of the multi-model consensus. NHC's simplest operational multi-model track consensus is known as TVCN, and the TVCN performance over the 2015-2017 period has been chosen as the HFIP TC track baseline. In 2017, TVCN comprised an equally weighted average of the early versions of three global models (ECMWF, GFS, and UKMET) and two regional models (HWRF and COAMPS-TC). Note that the composition of TVCN can change from year to year as the relative performance and availability of models evolve, and in the future TVCN could also include single-model ensemble forecasts and statistical post-processed forecasts developed under HFIP. Progress will be measured as improvements in future versions of TVCN relative to the baseline; the target for Metric 1.1 is to reduce TVCN mean errors by 50% relative to the baseline.

Similarly, the multi-model intensity consensus IVCN, evaluated over the 2015-2017 period, has been chosen as an HFIP intensity baseline. In 2017 IVCN comprised an equally weighted average of the HWRF and COAMPS-TC regional models, and the LGEM and D-SHIPS statistical models. The members of IVCN will also likely evolve in the future. Progress will be measured in terms of improvements in future versions of IVCN relative to the baseline, and the target for Metric 1.2 is to reduce IVCN mean errors by 50% relative to the baseline.

Rapid intensification (RI) was a leading focus of the original HFIP effort and still remains a significant forecast challenge for both the model guidance and forecasters. Generally defined as the 95th percentile of the intensity change distribution at a given forecast lead time, here it specifically refers to a 30-kt or more increase in intensity over a 24-h period. Because of the importance of RI, one or more distinct intensity guidance metrics will be developed. A number of options are currently being considered; these include:

- Mean Atlantic basin IVCN errors for the subset of forecasts in which a 30-kt increase in intensity occurred during any 24-h period prior to the verification time. It's worth noting that while attaining a 50% improvement in IVCN errors over all forecasts might be an unrealistic goal (in part because such a reduction would impinge upon the inherent uncertainty in estimating intensity), attaining a 50% improvement for the RI cases is both detectible in principle and potentially of much greater value to users.
- Mean forecast error, of IVCN or a designated HFIP model, evaluated only at verification times when RI is ongoing.
- Mean forecast error, of IVCN or a designated HFIP model, evaluated only at verification times when RI is either ongoing or was forecast.
- Probability of detection / false alarm rate (POD/FAR) for 0-24 h, 24-48 h, 48-72 h, etc.for either IVCN or a designated HFIP model.

The Naval Research Laboratory (NRL) developed a consensus model for wind radii prediction (RVCN), which includes input from dynamical and statistical models. RVCN has been designated as the measure for 34-kt wind forecast performance. RVCN was available beginning in 2016, and so the baseline sample will be all the 2016-2017 Atlantic storm. Progress will be measured as improvements in future versions of RVCN relative to the baseline (Metric 1.4).

As noted above, NHC does not forecast RMW. To support the WSP model, however, a simple climatological RMW model known as C-RMW had been previously developed. C-RMW was run operationally during 2015-17, and a subset of those forecasts will establish the baseline level of performance. Because NHC does not do a post-storm best-track evaluation of RMW, only operational RMW estimates are available to evaluate the performance of C-RMW. As a result, the subset of cases forming the baseline will be restricted to those cases when aircraft reconnaissance data were available; this eliminates from the baseline sample cases for which ground-truth uncertainty is highest. Progress will be measured by comparing RMW forecasts from HWRF or other advanced prediction systems against the C-RMW baseline (Metric 1.5).

Table A.2.1. Metrics for HFIP Goal 1

Goal 1	**Reduce track, intensity, and structure forecast guidance errors by 50% relative to a 2017 baseline. Reduce intensity forecast guidance errors by 50% for RI events.**
Metric 1.1	Mean absolute error (MAE) of TVCN track consensus
Baseline	TVCN MAE, Atlantic basin 2015-17, at 12, 24, 36, 48, 72, 96, and 120 hr: 23.1, 35.6, 47.6, 63.1, 101.5, 146.4, 195.9 n mi
Target	11.6, 17.8, 23.8, 31.6, 50.8, 73.2, 98.0 n mi
Metric 1.2	MAE of IVCN intensity consensus
Baseline	IVCN MAE, Atlantic basin 2015-17, at 12, 24, 36, 48, 72, 96, and 120 hr: 6.1, 8.5, 10.2, 11.4, 12.6, 14.4, 17.1 kt
Target	3.1, 4.3, 5.1, 5.7, 6.3, 7.2, 8.6 kt
Metric 1.3	Mean forecast error, of IVCN or a designated HFIP model, evaluated only at verification times when RI is either ongoing or was forecast
Baseline	IVCN MAE, Atlantic and eastern Pacific combined, 2015-17, at 12, 24, 36, 48, 72, 96, and 120 h, evaluated at only those times when RI occured: 17, 26.1, 28.6, 31.4, 36.9, 31.3, 32.1 kt
Target	8.5, 13.1, 14.3, 15.7, 18.5, 15.6, 16.1 kt
Metric 1.4	MAE of 34-kt RVCN radii consensus
Baseline	34-kt RVCN MAE, of Atlantic basin 2016-17, at 12, 24, 36, 48, 72, 96, and 120 hr: 21, 22, 22, 22, 25, 28, 36 n mi
Target	10.5, 11, 11, 11, 12.5, 14, 18 n mi
Metric 1.5	MAE of dynamical model (HWRF or follow-on models) RMW forecasts, evaluated when an aircraft reconnaissance fix was made within 6 h of the initial and verifying times.

Baseline	C-RMW MAE, Atlantic basin reconnaissance-restricted sample for 2015-17, at 12, 24, 36, 48, 72, 96, and 120 h: 18.4, 18.7, 19.0, 19.0, 20.0, 20.5, 20.9 n mi
Target	9.2, 9.4, 9.5, 9.5, 10.0, 10.3, 10.5 n mi

A.2.2 Improve extended range track and intensity forecasts

Currently, official forecasts of TC track and intensity are provided out to five days. Some users, however, engage in planning activities at longer lead times. In addition, NWS WFOs are required to issue seven-day forecasts for their areas of responsibility, but with little guidance provided by NHC it is difficult for the NWS to have consistent messaging on potential TC impacts at days six and seven. NHC has conducted internal testing on seven-day forecasts for several years, but the reliability of those forecasts has thus far been insufficient to consider making them publicly available. To improve service in this area, the guidance for TC track and intensity on days six and seven must become more reliable.

Table A.2.2 summarizes the HFIP Goal 2 metrics, discussed individually below, that will be used to measure improvements in extended-range track and intensity forecasts.

For track, the HFIP goal is to improve the Atlantic basin day seven TVCN mean error to current day 5 levels (Metric 2.1). If that can be attained, it is likely that the forecasts would be good enough for public dissemination without negative impacts. That said, there is some concern that the inevitable handful of extremely large seven-day errors could undermine public confidence in the entirety of the track forecasts, so it will be important to also be cognizant of the distribution of seven-day forecast errors. Note that the target in this case is not defined in terms of a percentage improvement relative to current TVCN levels. Even so, to assess progress it is convenient to define a baseline level of current skill and for this purpose the baseline sample will again be the three-year period 2015-17.

For intensity, there is currently only very limited quantitative intensity guidance available beyond five days, since the current statistical and regional dynamical hurricane models only run out that far. Thus, an essential first step will be to extend the multi-model intensity consensus IVCN out to seven days. That accomplished, the HFIP goal for extended-range intensity forecasts will be to reduce the day-seven Atlantic basin IVCN mean error to current day-five levels, as measured by the five-day mean IVCN error for the period 2015-17 (Metric 2.2). Since there currently is no IVCN available for day seven, a baseline has been defined as the 2015-17 mean GFS (AVNI) intensity error (although we note that as with track, this baseline plays no role in determining the target level of accuracy).

Table A.2.2. Metrics for HFIP Goal 2

Goal 2	**Produce 7-day track and intensity forecast guidance as accurate as a 2017 5-day baseline.**
Metric 2.1	MAE of 7-day TVCN track consensus
Baseline	TVCN MAE, Atlantic basin 2015-17, at 178 h: 222.9 n mi
Target	195.9 n mi
Metric 2.2	MAE of 7-day IVCN intensity consensus
Baseline	AVNI MAE, Atlantic basin 2015-17, at 178 h: 22.2 kt
Target	17.3 kt

A.2.3 Improve guidance on pre-genesis disturbances

As noted in section A.1, NHC and CPHC issue probabilistic forecasts of TC genesis covering the two-day and five-day forecast periods. In 2017, NHC began issuing its full suite of TC forecast products for “Potential Tropical Cyclones” (PTCs), defined as disturbances that are not yet a tropical cyclone, but which pose the threat of bringing tropical storm or hurricane conditions to land areas within 48 hours. The effective issuance of PTC advisory packages by NHC requires that high-quality track, intensity, and wind radii guidance be available to the forecaster for weather systems that have not yet become TCs.

Table A.2.3 summarizes the HFIP Goal 3 metrics, discussed individually below, that concern TC formation forecast improvements.

Currently, model track forecast errors for PTCs and other pre-genesis disturbances are larger, on average, than those for TCs. Such systems are, virtually by definition, less well-organized than TCs, often lacking a well-defined center of circulation, and are generally shallower and weaker compared to TCs. These factors make it more difficult for models to properly analyze their location and structure, and to forecast their interactions with the surrounding steering flow. In addition, we would expect a greater interdependence between track and intensity forecast error with pre-genesis disturbances, making track progress dependent on historically hard-to-come-by improvements in intensity forecasts. For these reasons, HFIP is setting more modest goals for pre-genesis forecast improvements relative to the TC goals. For track (Metric 3.1), the HFIP goal is to improve the Atlantic basin TVCN guidance for pre-genesis systems by 20% relative to those values over the period 2015-17.

Intensity prediction for pre-genesis systems is also problematic. The current TC model guidance suite has been developed explicitly for systems that are already TCs. For example, the developmental data sets for the SHIPS and LGEM statistical-dynamical models are restricted to TC cases; applying these models to systems excluded from the developmental sample results in a high intensity forecast bias for PTCs. The HWRF model was also developed with TCs in mind and imposes a vortex structure that often will poorly represent the true structure of a disturbance. As a result, the utility of the current intensity models for pre-genesis systems is limited. The HFIP intensity goal (Metric 3.2) is therefore similarly modest, aiming to improve the Atlantic basin IVCN guidance for pre-genesis systems by 20% relative to those values over the period 2015-17.

Operational probabilistic genesis forecasts have until recently been based largely on subjective (forecaster) interpretations of global model fields, satellite imagery, and other observations. In model analysis and forecast fields, TCs can be objectively identified through the setting and evaluation of certain thresholds for parameters such as low-level circulation and mid-level temperature anomalies. A “TC tracker”, such as the one developed by GFDL, coupled with statistical post-processing of model forecast fields, has been used in recent years to develop objective guidance on the timing and likelihood of TC genesis, and this objective guidance has been playing an increased role in the genesis forecasts issued by NHC. HFIP has established a goal to improve the guidance on the timing of genesis by 20%, with the baseline level of skill defined by the genesis timing errors for Atlantic basin pre-genesis systems, as identified by the GFDL tracker applied to GFS model fields, for the period 2015-17 (Metric 3.3).

Verification of forecasts for pre-genesis systems poses its own set of challenges. Guidance models such as TVCN or IVCN are only run on Atlantic disturbances that are designated as “invests” by NHC in the Automated Tropical Cyclone Forecast System (ATCF). Furthermore, NHC does not perform a post-storm best-track analysis on invests. Only if an invest becomes a designated PTC by NHC and advisories are issued will NHC construct a best track, and even then the best track may not cover the entire period for which guidance was run.

For purposes of evaluating HFIP Goal 3, the sample of pre-genesis disturbances will be restricted to officially designated invests and PTCs. Verifying positions and intensities will be taken from a post-storm best track, when available, and otherwise from the final operational working best track. Some work will be required to develop the capability to conduct these verifications, and until that occurs the baseline error statistics will be unavailable.

Table A.2.3. Metrics for HFIP Goal 3

Goal 3	**Improve forecast guidance on pre-genesis disturbances, for track, intensity, and the timing of genesis, by 20% relative to a 2017 baseline.**
Metric 3.1	MAE of TVCN track consensus for invests and PTCs at analysis time
Baseline	TVCN MAE, Atlantic basin 2015-17 for invests and PTCs at analysis time, at 12, 24, 36, 48, 72, 96, and 120 hr: [x] n mi
Target	0.8*baseline
Metric 3.2	MAE of IVCN intensity consensus for invests and PTCs at analysis time
Baseline	IVCN MAE, Atlantic basin 2015-17 for invests and PTCs at analysis time, at 12, 24, 36, 48, 72, 96, and 120 hr: [y] kt
Target	0.8* baseline kt
Metric 3.3	MAE of the predicted time of tropical (or subtropical) cyclone genesis
Baseline	MAE of (sub-)tropical cyclone genesis time from GFDL tracker applied to the GFS for 2015-17 Atlantic basin invests and PTCs at analysis time: [z] h
Target	0.8* baseline h

A.2.4 Improve hazard guidance and risk communication

The four primary hazards associated with TCs are storm surge, wind, rainfall and severe weather. The sections below discuss each hazard in turn and then efforts to improve the NWS's tropical cyclone product suite. Table A.2.4 summarizes the goals and metrics discussed below.

A.2.4.1 Storm surge

Due to its potential for large loss of life, the storm surge hazard drives most TC evacuation planning and decision making. Although tremendous progress has been made over the past decade to develop new products for decision makers, important gaps remain. Real-time storm surge products for the U.S. Atlantic and Gulf coasts, for example, are currently initiated roughly 48 h before the expected onset of life-threatening storm surge or tropical-storm-force winds. However, evacuation and preparedness decisions often must be made at longer lead times, before the real-time surge products are available. A second important limitation is the inability of the P-Surge framework to model the wide array of possible tropical cyclone structures that play a fundamental role in determining the surge threat posed by individual storms. In particular, the critical parameters of 34-kt wind radii and RMW are currently handled only very crudely by P-Surge.

Fortunately, storm surge prediction (especially within the P-Surge framework) benefits directly from improvements to TC track, intensity, and structure forecasts – those being addressed by HFIP Goal 1. These improvements, coupled with the availability of forecast uncertainty estimates through ensembles and statistical post-processing, should allow the storm surge product suite to be extended from two to three days. Because cross-track forecast error is such a large contributor to P-surge output, lessening this uncertainty has been chosen as the first storm surge goal, listed as Goal 4.1 in Table A.2.4. A second storm surge goal (Goal 4.2) will be to extend the current surge products, watches, and warnings beyond the Gulf and east coasts to all U.S. coastal regions at risk for TC storm surge.

A.2.4.2 Sustained winds and gusts

Tropical storm and hurricane watches and warnings are issued on the basis of the wind threat, having traditionally been determined from the official track, size and intensity forecast, coupled with the forecaster's subjective determination of uncertainty as well as non-meteorological factors. While TC wind probabilities have the potential to provide objective guidance for the placement of wind watches and warnings, they rely heavily on climatological forecast error distributions and are only weakly situationally dependent. Incorporating dynamical model ensemble information could improve the utility of the wind probabilities for watch/warning placement.

Probabilistic products can be verified using several metrics, including bias, Brier skill score, and threat score. Because of the intended use of the wind speed probabilities as guidance for watch/warning placement, the threat score will be used, as the threat score applied here would measure the overlap between areas with high probabilities of 34-, 50-, or 64-kt winds, and those areas that actually received those winds. The evaluation will be conducted on a regular grid covering the domain 15-50 °N, 60-100 °W , with the points deemed to have been affected by the winds of the various thresholds determined from NHC best-track positions and wind radii. Wind speed probabilities at both two days and five days will be evaluated (Goal 4.3).

TC wind warnings, intensities, and wind probability products all refer to sustained (1-min mean) surface (10-m) winds occurring over an unobstructed exposure. Wind gusts are not considered in any of these products, and furthermore are analyzed and forecast only in the most rudimentary manner. In addition, the ability of intensity forecast models to provide wind gust guidance is not well known. A second wind-related goal (Goal 4.4) is to perform an evaluation of dynamical model wind-gust forecasts as a first step towards development of new wind-gust hazard products.

A.2.4.3 Rainfall

One of the largest sources of error associated with TC rainfall forecasts, especially beyond two days, has been the TC track forecast error. Even though track errors have improved significantly in recent years, there are still events such as Hurricanes Harvey (2017) and Matthew (2016), in which the cumulative precipitation locations and amounts had significant errors because of track forecast errors. We expect that upgrades to the regional TC models will produce improved rainfall forecasts through improved track and intensity performance. In addition, HFIP plans to support running ensembles (HAFS and HWRF) at different time and spatial scales; these ensembles are expected to more reliably capture the range of possible solutions.

Thus, we expect improvements over the current probabilistic QPF products produced at WPC for TC events, both at the tactical (0-6 h) and preparedness (1-5 day) time scales, leading directly to improved flood and flash flood forecasts and warnings. Table A.2.4 shows three goals (Goals 4.5-4.7) for improving rainfall products that will enhance flood-related decision support services and help mitigate flood impacts.

A.2.4.4 Severe local weather

In addition to the high winds, storm surge and rainfall, TCss can produce other locally severe weather, sometimes very far from the storm center. Tornadoes are the primary severe weather hazard from TCs. As described in section 3.2.3, SPC provides probabilistic outlooks for severe weather, including those from tornadoes. These are created subjectively by forecasters based on model forecasts and other objective guidance. The day-one severe weather outlook provides separate probabilities for large hail, high winds and tornadoes, while the longer-range outlooks provide a single probability for all severe weather hazards.

The goal is to improve the 24- hour probabilistic tornado outlook product by 10% as shown by Goal 4.8 in Table A.2.4.

A.2.4.5 Product suite modernization

A new HFIP focus area is to create more effective products through incorporation of risk communication research. To provide improved communication of TC hazard risk, NOAA must first undertake an assessment of its current TC product and services suite; this will allow NOAA to gain a baseline understanding of the ways in which its products are used by partners and stakeholders (e.g., emergency managers, broadcast media, the general public, etc.) in their decision-making. The effectiveness of such an assessment is enhanced if it integrates users at the outset, while continued partner/stakeholder involvement in product improvement and development results in co-produced outcomes that will better meet users' needs.

The following are some of the considerations and challenges associated with this substantial undertaking:

- The needs of a diverse internal and external user base will need to be balanced. These considerations include determining if any products can or should have limited availability, determining the proper balance between automated guidance and human interaction, and identifying which products/services should be streamlined, expanded, or discontinued.
- Although many users understand the need and value of probabilistic guidance, NOAA's commitment to an increased focus on probabilistic information will occasionally be at odds with some users' unfamiliarity with probabilities and their natural desire for deterministic forecasts. This assessment and development effort needs to ensure a

comprehensive effort to transition NOAA products and services toward the provision and effective utilization of probabilistic hazard information.

- Along those lines, research is needed to help clarify how probabilistic forecasts improve decision-making by core partners, and how such forecasts improve core partners' ability to distinguish between low- and high-impact events.
- Research will also be needed to determine the best mix of visualizations, stories, colors, etc.that best improves the communication of risk and uncertainty, resulting in improved decision support services, risk assessments, and preparedness.
- Product improvements will need to be iteratively tested and evaluated with users to ensure their effectiveness.

With these considerations in mind, Table A.2.4 identifies three goals and a series of actions that HFIP has identified for this process. The first (Goal 4.9) broadly speaks to activities needed to assess the current TC product suite, the second (Goal 4.10) addresses requirements for improved products, while the third (Goal 4.11) deals with the implementation process. Note that these goals are not inherently quantitative in nature (e.g., "Assess how the TC product suite is used"), and therefore the listed "metrics" represent the necessary actions or accomplishments that must be achieved in order to reach the stated goal.

Table A.2.4. Metrics and sub-goals for HFIP Goal 4

Goal 4	**Improve hazard guidance and risk communication for all of the TC hazards (wind, surge, rainfall, and tornadoes) at actionable lead times through the application of social and behavioral sciences, resulting in a modernized suite of TC products, information, and services.**
Goal 4.1	**Increase lead time of real-time storm surge products and services from two days to three days with no loss of skill.**
Metric 4.1	Mean three-day cross-track error used by P-surge, unadjusted for initial intensity (five-yr mean NHC-official cross-track error for all Atlantic basin forecasts that initiate and verify between 10-45N and 60-100W).
Baseline	Mean three-day NHC-official cross-track error for all Atlantic basin forecasts that initiate and verify between 10-45N and 60-100W for the period 2013-17: 57 n mi
Target	Mean three-day cross-track errors less than or equal to the mean two-day NHC-official cross-track error for all Atlantic basin forecasts that initiate and verify between 10-45N and 60-100W for the period 2013-17: 35 n mi
Goal 4.2	**Expand real-time storm surge forecast and warning capabilities to cover all areas served by the NWS that are vulnerable to TC storm surge, including Puerto Rico, the U.S. Virgin Islands, southern California, Hawaii, Guam, and American Samoa.**
Metric 4.2	Number of surge products and services that have been expanded to entire NWS area of responsibility vulnerable to TC storm surge
Baseline	Operational surge products and services available in 2017 for the U.S. Gulf and East coasts (i.e., P-surge, Potential Storm Surge Flooding Graphic, storm surge watches and warnings)
Target	Baseline products and services available to entire NWS area of responsibility vulnerable to TC storm surge
Goal 4.3	**Improve the accuracy of the 34-, 50-, and 64-kt wind speed probabilities at two and five days by 50%.**
Metric 4.3	The threat score of the 34-, 50-, and 64-kt wind speed probabilities, evaluated on a regular grid covering 15-50 °N and 60-100 °W, at days two and five.

Baseline	The threat score from the 2017 version of the wind speed probability model, run on the Atlantic 2015-2017 sample over the domain 15-50 °N, 60-100 °W, for the 34-, 50- and 64-kt thresholds: Day 2 [x, y, z] Day 5 [xx, yy, zz]
Target	Day 2 1.5*[x, y, z] Day 5 1.5*[xx, yy, zz]

Goal 4.4	**Perform a systematic evaluation of dynamical model forecasts for wind gusts associated with TCs.**
Metric 4.4	Publication or technical report.
Baseline	N/A
Target	Publication or technical report.
Goal 4.5	**Improve the accuracy and lead time of the WPC Excessive Rainfall Outlook for TCs.**
Metric 4.5	Brier Score of Day-3 Excessive Rainfall Outlook for landfalling Atlantic basin TCs
Baseline	Current Brier Score of Day-Three Excessive Rainfall Outlook, 2015-7 Atlantic basin CONUS-landfalling TCs. [x]
Target	Current Brier Score of Day-Two Excessive Rainfall Outlook: [y]
Goal 4.6	**Improve skill of Quantitative Precipitation Forecasts (QPF) for landfalling TCs.**
Metric 4.6	QPF Brier Score for TCs affecting CONUS, Puerto Rico, and U.S. Virgin Islands
Baseline	QPF Brier Score for TCs affecting CONUS, Puerto Rico, and U.S. Virgin Islands during 2015-17: [x]
Target	10% improvement over baseline
Goal 4.7	**Create a probabilistic tropical QPF product based upon HAFS/HWRF ensemble output.**
Metric 4.7	Dissemination of probabilistic tropical QPF for CONUS, Puerto Rico, and the U. S. Virgin Islands TC threats.
Baseline	N/A

Target	Dissemination of probabilistic tropical QPF for CONUS, Puerto Rico, and the U. S. Virgin Islands TC threats.
Goal 4.8	**Improve the SPC one-day probabilistic forecast for tornadoes by 10%.**
Metric 4.8	The resolution of the SPC one-day tornado probabilities associated with landfalling tropical cyclones, as measured by the Brier Score.
Baseline	Brier score of the SPC one-day tornado probabilities associated with landfalling tropical cyclones from 2015-201.
Target	Baseline*1.1
Goal 4.9	**Assess how the current TC product suite is used across America's Weather Enterprise, by NWS partners, and by end users.**
Metric/ Milestone 4.9.1	Documentation of the current TC product suite and completion of a baseline assessment of TC product use across the NWS.
Metric/ Milestone 4.9.2	Completion of a baseline assessment of TC product use across America's Weather Industry, with an emphasis on identification of high- versus low-use products, including an assessment of how the product suite is modified and dissemination to public audiences.
Metric/ Milestone 4.9.3	Completion of a baseline assessment of NWS core partners' and end-users' use and understanding of the current TC product suite, with particular emphasis on their numeracy skills and understanding of probabilities.
Baseline	N/A
Target	Completion of assessments described in the above metrics/milestones.
Goal 4.10	**Identify requirements for a modernized TC product suite.**

Metric/ Milestone 4.10.1	Completion of a baseline assessment of NWS partners' and user TC information needs. Such an assessment should consider potential time scales from minutes to weeks in advance of an event, as well as characteristics of the needed information such as risk, confidence, uncertainty, formats, interactivity, methods of delivery, etc.
Metric/ Milestone 4.10.2	Identification of intended communication objectives for social and behavioral science researchers.
Metric/ Milestone 4.10.3	Completion of a baseline social and behavioral science analysis of efficacy of current TC products to support key decision-making by NWS partners and users to: 1) meet intended communication objectives and ensure partners' and users' information needs are met.
Metric/ Milestone 4.10.4	Synthesis of baseline assessments and review by the HFIP Socio-Economic Working Group to determine operational viability of identified needs. Development of product prioritization, identifying which products/services should be streamlined, expanded, or discontinued.
Baseline	N/A
Target	Completion of the actions described in the above metrics/milestones.
Goal 4.11	**Develop and disseminate a modernized TC product suite that is informed by probabilistic information to better convey risk and uncertainty, and through which enables enhanced risk assessment and timely preparedness actions on the part of users, partners, and stakeholders to reduce loss of life and property, and which includes other weather-forecast-related organizations in our shared responsibility in the effective conveyance of risk and uncertainty.**
Metric/ Milestone 4.11.1	Creation of working groups, guided by the HFIP Socio-Economic Working Group, representing interdisciplinary expertise (e.g., social, behavioral, and physical science researchers, operational forecasters) necessary to manage, develop, implement, and disseminate the proposed product changes.
Metric/ Milestone 4.11.2	Gathering of NWS partner and user feedback on proposed product changes through the use of partner/user engagement, as well as social and behavioral science methodologies.

Metric/ Milestone 4.11.3	Development of an NWS partnership with other weather-forecast-related organizations that disseminates the modernized NWS TC product suite, and that empowers the other weather information providers to develop their own TC products to better convey risk and uncertainty.
Metric/ Milestone 4.11.4	Development of an NWS partnership with key agencies and professional organizations (e.g., FEMA, NEMA, IAEM, AMS, NWA, etc.) that ensures effective training and outreach is developed and available to key partners and users of the modernized TC product suite.
Baseline	N/A
Target	Completion of the actions described in the above metrics/milestones.

Appendix B.

List of Acronyms

3DEnVAR	Three-dimensional hybrid ensemble-variational data assimilation
4DEnVAR	Four-dimensional hybrid ensemble-variational data assimilation
AOML	Atlantic Oceanographic and Meteorological Laboratory
AMSR	Advanced Microwave Scanning Radiometer
AMSU	Advanced Microwave Sounding Unit
AMV	Atmospheric Motion Vectors
ASCAT	Advanced Scatterometer
ASOS	Automated Surface Observing Stations
ATMS	Advanced Technology Microwave Sounder
AWI	America's Weather Industry
CAM	Convection Allowing Model
COAMPS	Coupled Ocean/Atmosphere Mesoscale Prediction System
COASTAL	Consumer Option for an Alternative System to Allocate Losses Act
CONUS	Contiguous United States
CPHC	Central Pacific Hurricane Center
CYGNSS	Cyclone Global Navigation Satellite System
DA	Data Assimilation
DMSP	Defense Meteorological Satellite Program
DOD	Department of Defense
DTC	Developmental Testbed Center
EMC	Environmental Modeling Center
EFS	Experimental Forecast System
ESMF	Earth System Modeling Framework

ESRL	Earth System Research Laboratory
FACETs	Forecasting a Continuum of Environmental Threats
FEMA	Federal Emergency Management Agency
FV3	Finite-Volume Cube-Sphere Dynamical Core
GCOM	Global Change Observation Mission
GDAS	Global Data Assimilation System
GFDL	Geophysical Fluid Dynamics Laboratory
GFS	Global Forecast System
GMI	GPM Microwave Imager
GOES	Geostationary Operational Environmental Satellite
GPM	Global Precipitation Measurement
GSI	Gridpoint Statistical Interpolation
HAFS	Hurricane Analysis Forecast System
HEDAS	Hurricane Ensemble Data Assimilation System
HFIP	Hurricane Forecast Improvement Program
HLS	Hurricane Local Statement
HMON	Hurricanes in a Multi-scale Ocean-coupled Non-hydrostatic Model
HMT	Hydrometeorological Testbed
HPC	High Performance Computing
HRD	Hurricane Research Division
HTI	Hurricane Threats and Impacts
HWRF	Hurricane Weather Research and Forecast Model
HWT	Hazardous Weather Testbed
JCSDA	Joint Center for Satellite Data Assimilation
JEDI	Joint Effort for Data Assimilation Integration
JHT	Joint Hurricane Testbed
LEO	Low-Earth Orbiting

LIDAR	Light Detection and Ranging
LSR	Local Storm Reports
MAE	Mean Absolute Error
MEOW	Maximum Envelope of Water
MOM	Maximum of the Maximum
MPING	Meteorological Phenomena Identification Near the Ground
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NEMS	NOAA Environmental Modeling System
NGGPS	Next Generation Global Prediction System
NHC	National Hurricane Center
NOAA	National Oceanic and Atmospheric Administration
NRL	Naval Research Laboratory
NUOPC	National Unified Operational Prediction Capability
NWS	National Weather Service
O2R	Operations to Research
O&M	Operation & Maintenance
OAR	Oceanic and Atmospheric Research
OCONUS	Outside Contiguous United States
OFCM	Office of the Federal Coordinator for Meteorology
OPG	Operations Proving Ground
OSEs	Observing System Experiments
OSSEs	Observing System Simulation Experiments
PTC	Potential Tropical Cyclones
QOSAP	Quantitative Observing System Analysis Project
QPF	Quantitative Precipitation Forecast
R&D	Research and Development

R2O	Research to Operations
RDGs	Rapid Deployment Gages
RFC	River Forecast Centers
RI	Rapid Intensification
RL	Readiness Levels
RMW	Radius of Maximum Wind
SC	Storm Centric
SFMR	Stepped Frequency Microwave Radiometer
SIP	Strategic Implementation Plan
SLOSH	Sea, Lake and Overland Surges from Hurricanes Model
SPC	Storm Prediction Center
SSMIS	Special Sensor Microwave Imager/Sounder
TA	Tropical Atlantic
TB	Terabyte
TC	Tropical Cyclone
TCD	Tropical Cyclone Discussion
TCP	Tropical Cyclone Public Advisory
TDR	Tail Doppler Radar
TROPICS	Time-Resolved Observations of Precipitation structure and storm Intensity with a Constellation of Smallsats
TWIEP	Tornado Warning Improvement and Extension Program
U.S.	United States
USGS	United States Geological Survey
UFS	Unified Forecast System
VORTEX-SE	Verification of the Origins of Rotation in Tornadoes Experiment-Southeast
VM	Vortex Modification
VR	Vortex Relocation

VTEC	Valid Time Event Code
WFOs	Weather Forecast Offices
WPC	Weather Prediction Center
WRN	Weather Ready Nation
WSP	Wind Speed Probability
WSR	Weather Surveillance Radar

Appendix C

References

Hsiao, L., D. Chen, Y. Kuo, Y. Guo, T. Yeh, J. Hong, C. Fong, and C. Lee, 2012: Application of WRF 3DVAR to Operational Typhoon Prediction in Taiwan: Impact of Outer Loop and Partial Cycling Approaches. *Wea. Forecasting*, **27**, 1249–1263, https://doi.org/10.1175/WAF-D-11-00131.1

Lu, X., X. Wang, M. Tong, and V. Tallapragada, 2017: GSI-Based, Continuously Cycled, Dual-Resolution Hybrid Ensemble–Variational Data Assimilation System for HWRF: System Description and Experiments with Edouard (2014). *Mon. Wea. Rev.*, **145**, 4877–4898, https://doi.org/10.1175/MWR-D-17-0068.1

Zhang, F., Y. Weng, J.A. Sippel, Z. Meng, and C.H. Bishop, 2009: Cloud-Resolving Hurricane Initialization and Prediction through Assimilation of Doppler Radar Observations with an Ensemble Kalman Filter. *Mon. Wea. Rev.*, **137**, 2105–2125, https://doi.org/10.1175/2009MWR2645.1