NOAA

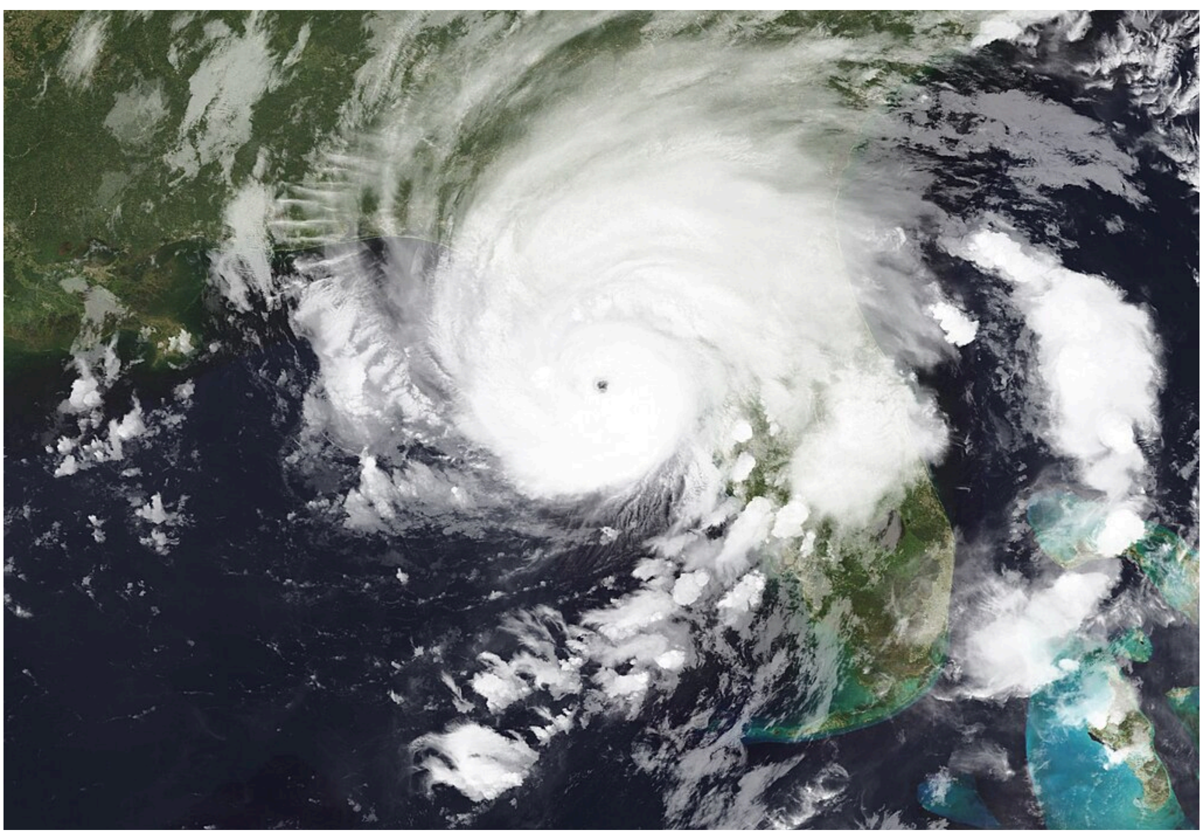
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
United States Department of Commerce

NOAA HURRICANE FORECAST IMPROVEMENT PROJECT

Hurricane Forecast Improvement Program (HFIP) Annual Report 2023

25 February 2025

NOAA technical memorandum HFIP-2023
https://doi.org/10.25923/qc0s-9150



Cover: GOES-16 geocolor image of Hurricane Idalia at 0715 UTC 30 August 2023. Image courtesy of NOAA/NESDIS/STAR.

Annual Report Writing Team

Aaron Poyer
William Komaromi
Jason Anderson
Sundararaman Gopalakrishnan

Ghassan Alaka
Jordan Dale
Andrew Hazelton
Wallace Hogsett
Frank Marks
Avichal Mehra
Kathryn Newman
Jamie Rhome
Vijay Tallapragada
Xuejin Zhang
Zhan Zhang

Table of Contents

Executive Summary

This technical report describes the activities and results of the Hurricane Forecast Improvement Program (HFIP) that occurred in the 2023 hurricane season. This year's report is intended to be more streamlined in general with less background/historical information and instead provide the key highlights of progress and achievement that the program has made in 2023 as well as the future direction and development.

In 2023, the HFIP program actively engaged in significant briefings and outreach initiatives with NOAA's management, executive leadership and Congressional committee, reinforcing the program's commitment to transparent communication and alignment with organizational priorities. The major focus of this report is to highlight the success of bringing the Hurricane Analysis and Forecast System (HAFS) v1 to operations on June 27, 2023. Additionally, this report will look at the success and challenges that HAFSv1 encountered with certain storms during the 2023 season as well as improvements that began development for v2 at the end of the year. We will also detail the thought process and groundwork for the new HFIP Strategic Plan, an overview of the annual meeting, results of various real time experiments, and highlight publications that came out in 2023 related to HFIP work and development.

The 2023 North Atlantic hurricane season was above average, as the fourth most active season on record, tied with 1933. There were 20 named storms, of which 7 developed into hurricanes, with 3 of those becoming major hurricanes. There were 3 landfalls in the U.S. from 2 tropical storms and 1 hurricane. The eastern North Pacific hurricane season also featured above average activity, with 17 named storms, 10 hurricanes, and 8 major hurricanes. Across the NHC area of responsibility, 14 tropical cyclones underwent rapid intensification (RI) , defined as an intensification of 30 kt or more in 24 hours, from 3 tropical cyclones (Franklin, Idalia, and Lee) in the Atlantic basin and 11 tropical cyclones in the Eastern North Pacific (Adrian, Beatriz, Calvin, Dora, Eugene, Fernanda, Hilary, Jova, Lidia, Norma, Otis).

The major highlights of 2023 were:

1. A joint development between NOAA operations and research, HAFSv1 became the first major coupled UFS based regional model used in operations when implemented on June 27, 2023. HAFS improved forecasts of track by >15% after day 2, & intensity by >12% after day 3 in NHC basins, compared to HWRF. As a cautionary note, HAFS underperformed relative to HWRF for days 1-2. Improving short-term intensity is part of ongoing research on multiscale interactions.

2. An experimental HAFS ensemble was run on the AWS Cloud in real-time for the first time during the 2023 hurricane season, and was found to be highly beneficial to probabilistic hurricane prediction by developers and forecasters alike.

3. HFIP Real-time Experiments (HREx) for 2023 was a collaborative effort across EMC, AOML, GFDL, and GSL, exploring a variety of configurations for HAFSv2 upgrades. An

updated HAFS configuration from storm-centric to multi-storm and basin-centric will be explored for the 2024 HREx.

4. While HAFSv1 has proven to be superior to HWRF and HMON by most metrics and at most lead times, the HAFS day 1-3 intensity forecast remains a concern to NHC. NHC would like to see this concern addressed prior to retiring the legacy models.
5. Moving forward, the focus is shifting towards optimizing a limited number of “top priorities” in a resource limited environment, based upon user and stakeholder feedback, agency priorities, and projected reduction (FY25+) in funding and continued HPC constraints.
6. Accelerated progress on the development of the HAFS Ensemble and the transition of the Data Assimilation (DA) system to JEDI are top priorities. The existing DA system (GSI) is no longer being developed.

1. History of HFIP

1.1. Introduction

This report describes the Hurricane Forecast Improvement Program (HFIP), its goals, proposed methods for achieving those goals, and the most recent results from the program, with an emphasis on advances in the skill of operational hurricane forecast guidance. Section 1 of this report describes the background, goals, and baselines for measuring success within the HFIP program. Section 2 focuses upon the initial operating capability (IOC) of the Hurricane Analysis and Forecasting System (HAFS), highlights high-resolution hurricane modeling successes from the 2023 hurricane season, and highlights experimental and developmental versions of the model, including a HAFS ensemble, that are in the testing and evaluation stages for possible future transitions. Section 3 highlights the engagement of HFIP with the community and summarizes the HFIP Annual Meeting 2023. Section 4 summarizes this report, and previews a new direction for the future of HFIP that will be elaborated upon in further detail in the upcoming HFIP Strategic Plan 2025-2035. For more background information, readers are referred to earlier reports available on the HFIP website.

1.2. The Hurricane Forecast Improvement Program (HFIP)

Originally established as the Hurricane Forecast Improvement Project, authorized in 2007 and beginning in 2009, HFIP was created within NOAA in response to the particularly damaging landfalling hurricanes (e.g., Charley, 2004; Wilma, Katrina, Rita, 2005) in the first half of that decade. HFIP’s original 5-year (for 2014) and 10-year goals (for 2019) were to:

- Reduce average track errors by 20% in 5 years, and by 50% in 10 years for days 1-5
- Reduce average intensity errors by 20% in 5 years, and 50% in 10 years for days 1-5

- Increase the probability of detection (POD)[1] for RI to 90% at Day 1, decreasing linearly to 60% at day 5, and decreasing the false alarm ratio (FAR) for rapid intensity change to 10% for day 1, increasing linearly to 30% at day 5. [The focus on RI change is the highest-priority forecast challenge identified by the National Hurricane Center (NHC)].
- Extend the lead-time for hurricane forecasts out to Day 7 (with accuracy equivalent to that of the Day 5 forecasts when those were introduced in 2003).

For more than a decade, HFIP has been providing the unified organizational infrastructure, funding, and compute resources for NOAA, university, and private partnerships to coordinate the hurricane research needed to achieve the above goals, improve storm surge forecasts, and accelerate the transition of model codes, techniques, and products from research to operations. HFIP focuses on multi-organizational activities to research, develop, demonstrate, and implement enhanced operational modeling capabilities, dramatically improving the numerical forecast guidance made available to the NHC, as well as enhancing the interpretation of that guidance. Through HFIP, NOAA continues to improve the accuracy of hurricane forecasts, with applied research using advanced computer models.

In 2017, Congress passed the Weather Research and Forecasting Innovation Act including Section 104, reauthorizing HFIP as the Hurricane Forecast Improvement *Program*. Under HFIP, this Congressional Act instructed NOAA to maintain a project to improve hurricane forecasting with the goal of developing and extending accurate hurricane forecasts and warnings in order to reduce loss of life, injury, and damage to the economy. HFIP has a particular focus on improving the prediction of rapid intensification and track of hurricanes, improving the forecast and communication of surges from hurricanes, and incorporating risk communication research to create more effective watch and warning products. In response to this charge, the HFIP Strategic Plan 2019-2024 was updated outlining the research and development needed to continue improving hurricane forecast guidance, enhance probabilistic hazard products, and design a more effective tropical cyclone (TC) product suite to better communicate risk to the public and emergency management community. Under the updated plan, HFIP will continue to address the original goals of reducing track and intensity forecast errors by 20% within 5 years and 50% within 10 years, and to extend forecasts out to 7 days, particularly with focus on rapid intensification guidance. In addition, the updated plan extends HFIP's purview to improving guidance on predicting storm structure and all hurricane hazards (surge, rain, associated severe weather, gusts as well as sustained winds) at actionable lead times for emergency managers (e.g., 72 hours). Improved hazard guidance will derive from dynamical model ensembles enabling probabilistic hazard products and improved track, intensity change and structure (radii to maximum and 35-knot winds) predictions before formation and throughout the storm's life cycle. Using social science research, HFIP will design a more effective tropical cyclone product suite to better communicate risk and transition all current tropical hazards products.

---

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

One of the key strategies defined in the revised hurricane forecast improvement strategic plan in response to the proposed framework for addressing the Weather Act of 2017, is to advance an operational HAFS. HAFS is a multi-scale model and data assimilation package capable of providing high-resolution analyses and forecasts of the inner core structure of the TC out to a lead time of 7 days, which is key to improving size and intensity predictions, as well as the large-scale environment that is known to steer TCs and provides favorable/unfavorable dynamic (e.g., vertical wind shear) and thermodynamic (e.g., mid-tropospheric moisture) conditions. HAFS will provide an operational analysis and forecast system out to 7 days for hurricane forecasters with reliable, robust and skillful guidance on TC track and intensity (including RI), storm size, genesis, storm surge, rainfall and tornadoes associated with TCs. It will provide an 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 UFS. HAFS is supported under several Hurricane Supplemental projects, (i) 1A-4a: Accelerate Development of Moving Nest for HAFS; (ii) 3A-1: Accelerate implementation of the updated HFIP Plan; (iii) 3A-2: Accelerate Re-engineering of HAFS; (iv) 2019 Disaster Supplemental Improving Forecasting of Hurricanes, Floods and Wildfires HU-2 project (v) 2022 Disaster Relief Supplemental Act HURR1 project.

HFIP is organized along two lines of activities: Stream-1 and Stream-2. While Stream-1 works within presumed operational computing resource limitations, Stream-2, also called as HFIP Real-time Forecasting Experiments (HREx), activities assume that resources will be provided to increase the available computer capability in operational settings, above the one that is already planned for the next five years. The purpose of Stream-2 is to demonstrate that the application of advanced and innovative science, technology, and increased computing will lead to the desired increase in accuracy, and other improvements in forecast performance. Because the level of computing necessary to perform such a demonstration is larger than can be accommodated by current operational computing resources, HFIP leverages the *Jet* supercomputer located at the David Skaggs Research Center (DSRC) in Boulder, Colorado.

2. HFIP in 2023

2.1. Background

This section summarizes the activities and results of the Hurricane Forecast Improvement Program (HFIP) that occurred in 2023. The major focus of this report is the deployment of the Hurricane Analysis and Forecast System (HAFS) within the Unified Forecast System (UFS) and its first operational implementation.

Much recent progress in tropical cyclone forecasting can be attributed to the success of HFIP over the last 15 years. In Section 2.2, we will provide more detailed background on the HFIP program and summarize the success of HFIP since its inception, highlighting the establishment of new goals as previous goals have been met. In Section 2.3, we will summarize the performance of the National Hurricane Center (NHC) and available real-time forecast guidance, with particular emphasis on the newly operationalized HAFS-A and HAFS-B mesoscale hurricane models that were developed primarily under the HFIP program, as well as comparisons against the legacy HWRF and HMON mesoscale models. In Section 2.4, we will discuss the development of the state-of-the-science next generation of models and ensembles for providing numerical guidance as part of the HREx experimental suite and beyond.

2.2. Legacy of Successes within HFIP

As outlined in a recent HFIP Executive Oversight Board (HEOB) brief to the NOAA Assistant Administrators (AAs) and the Portfolio Directors, HFIP has been a quantifiable success. Since the inception of HFIP, model hurricane track errors have been reduced by 50%, intensity forecast errors have been reduced by 56%, and intensity errors during rapid intensification (RI) have been reduced by 47%. With the support of HFIP, the Hurricane Weather Research and Forecasting (HWRF) model became the best deterministic intensity guidance used worldwide in tropical cyclone prediction. In response to the Weather Act of 2017, a new set of HFIP goals were established in order to maintain ongoing research to improve hurricane forecasting. This new set of goals included: (1) further reduction of track and intensity forecast guidance errors by half (50%) from those set in 2007, including for rapid intensification; (2) improve forecasts and guidance for storm surge and other storm-induced hazards; and (3) incorporate risk communication research to create more effective watch and warning products.

HFIP has been a cross-cutting effort across NOAA. NWS/OSTI leads a collaborative effort to carry out the goals of HFIP, including, but not limited to, the invaluable collaboration between NWS/EMC for transitioning model innovations into operations, NWS/NHC for operational forecasts and products, and OAR/HRD for research and development. More recently, hurricane modeling has begun to look to the future, with a forthcoming transition to the Unified Forecast System (UFS) through development of the Hurricane Analysis and Forecast System (HAFS), which first became operational in 2023.

Model track forecast errors are closing the gap to meet the original 2007 HFIP 10-year and 2017 Weather Act 5-yr error reduction goals (Figure 1). Further development of the HAFS model is needed to close the gaps between observed track error and the original goals, as well as meet the Weather Act 10-year goal by 2027. The results have been even more impressive for intensity. Model track error has met the original 10-year goal, and even exceeded the Weather Act 5-year goal (Figure 2). The Weather Act 10-year goal for intensity is ambitious, and further development of HAFS will be needed to meet this goal by 2027.

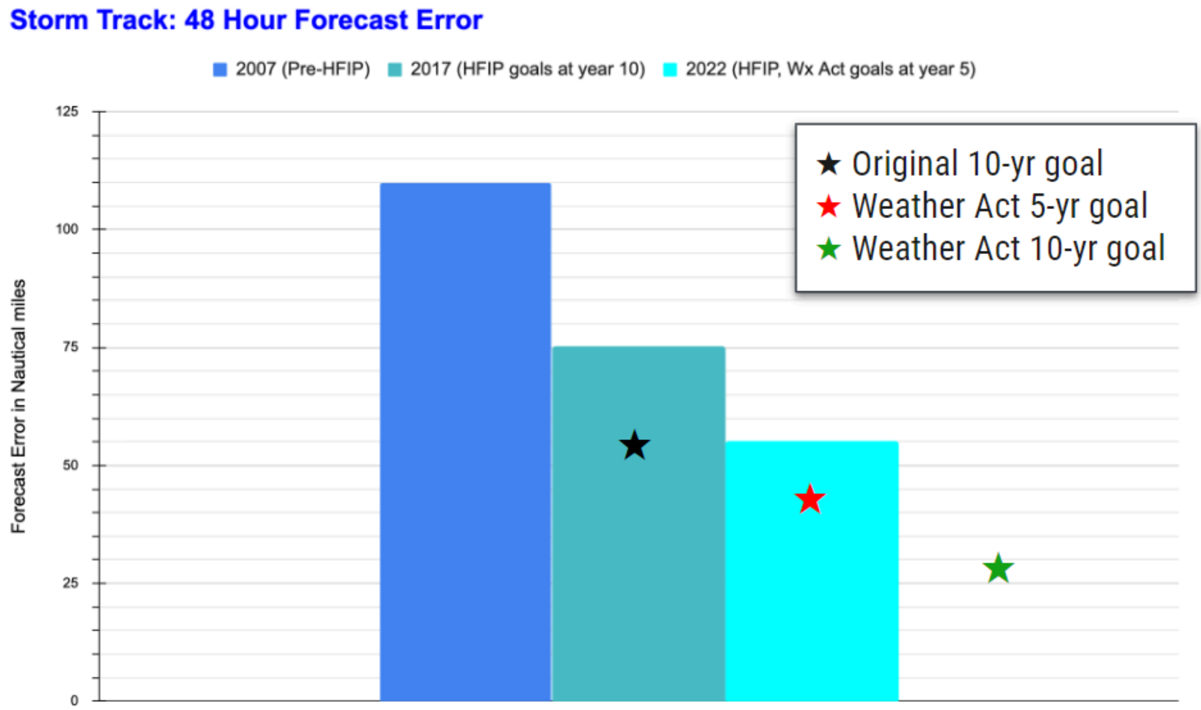


Figure 1: Observed track forecast error (nmi; bar graph) at 48-h lead time, pre-HFIP in 2007, when HFIP goals reached the year 10 mark in 2017, and the Weather Act goals reached year 5 in 2022, compared to the original 10-year goal, the Weather Act 5-year goal, and the Weather Act 10-year goal (black, red, and green stars, respectively).

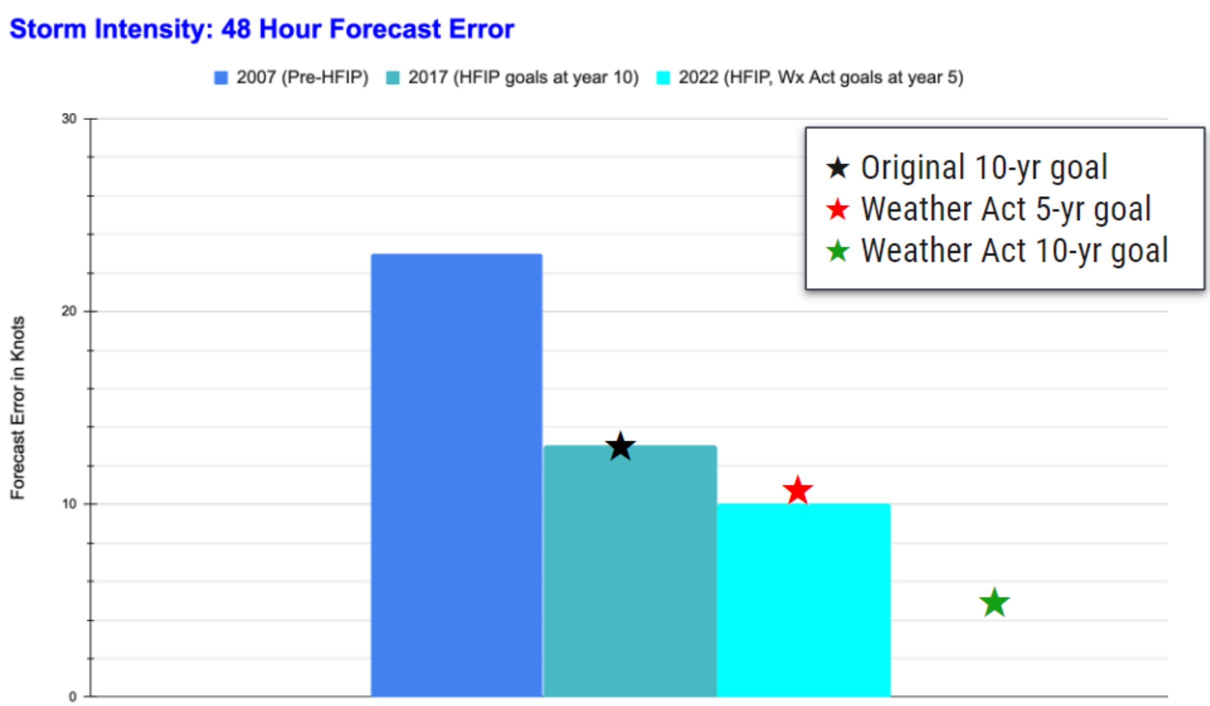


Figure 2: As in Figure 1, but for intensity error (kt).

The original HFIP 2007 goals pertained to the probability of detection of rapid intensification, as opposed to a specific improvement in error when RI occurs. Quantifiable 5-year and 10-year goals in terms of the reduction in intensity forecast error, conditional on RI being observed, were established in the Weather Act or 2017. Model-predicted intensity errors during periods of rapid intensification are currently approaching the Weather Act 5-year goal (Figure 3). As was the case for the HFIP track error objectives, further development of HAFS is needed to close the Weather Act 5-year goal gap and meet the Weather Act 10-year goal by 2027.

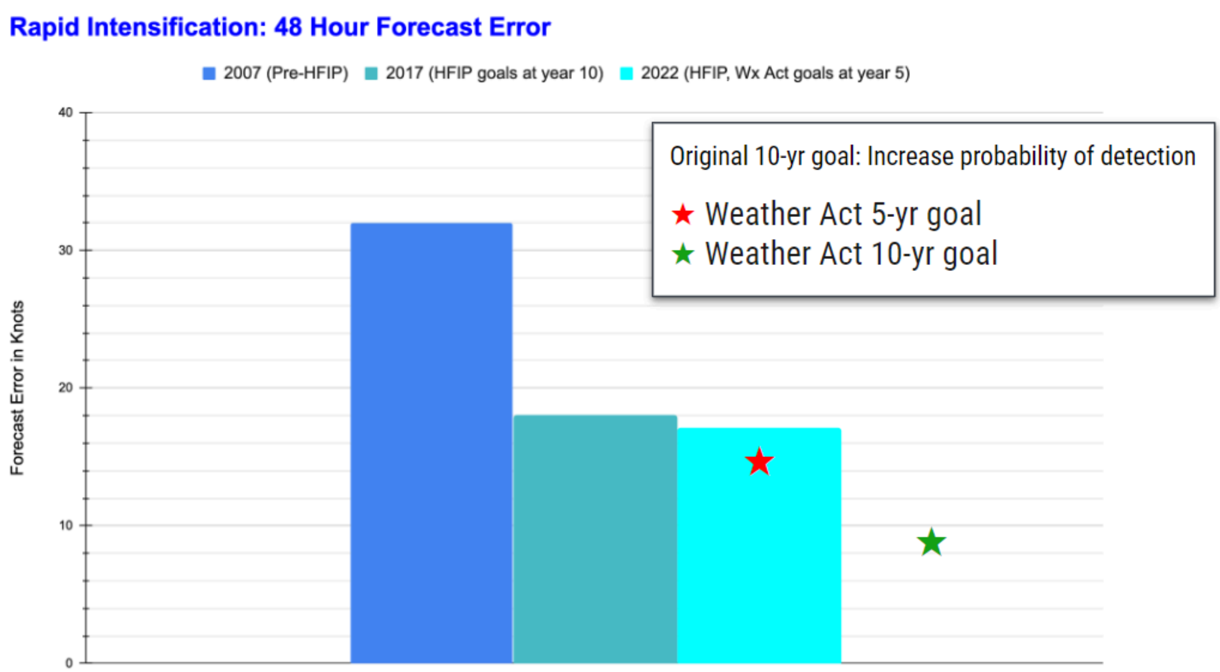


Figure 3: Observed intensity forecast error (kt; bar graph) at 48-h lead time, conditional upon rapid intensity being observed. Bars correspond to pre-HFIP in 2007, when HFIP goals reached the year 10 mark in 2017, and the Weather Act goals reached year 5 in 2022, compared to the Weather Act 5-year goal, and the Weather Act 10-year goal (red and green stars, respectively).

Aligned with the new 2017 Weather Act goals, HFIP is supporting a series of critical intermediary steps, including working towards having real-time (but not yet operational) predictive guidance from a HAFS ensemble by 2023, improved pre-formation disturbance guidance by 2026, and multiple moving nest capability in HAFS for all tropical ocean basins by 2027. Ongoing challenges, such as the recent Hurricane Ian disaster, highlight the need for continuing HFIP. In addition to focusing on the development of the next-generation HAFS probabilistic and ensemble systems, the future of HFIP also seeks to advance the social sciences component of risk communication in hurricane science. Critical advancements towards HFIP strategic goals related to risk communication are being made, including the operational implementation of Tropical Storm Force Winds - Time of Arrival product. HFIP will achieve

Social Behavioral and Economic Science (SBES) goals to further improve risk communication through the tropical product suite by integrating research outcomes into new and existing internal and public facing tropical products and services. Recent work on the development of HAFS ensemble looks to address continued challenges in communicating probabilistic information to forecasters, emergency managers, and the public.

2.3. Operational Highlights from the 2023 Hurricane Season

Assessments of operational real-time model performance were performed by NHC in the North Atlantic, Eastern North Pacific, and the Central North Pacific. In this report, we also include a brief summary from the Joint Typhoon Warning Center (JTWC), which includes all other global basins outside of NHC's area of responsibility (AOR). Additional model verification data is provided by EMC and HRD.

HAFSv1 became the first major coupled UFS based regional model used in operations when implemented on June 27, 2023. Two configurations of HAFS were implemented for the initial operating capability (IOC), referred to as HAFS-A and HAFS-B. A summary of the differences between HAFS-A and HAFS-B appears below in Table 1. Both versions of HAFS are running with a 6-km horizontal resolution outer regional grid, with a 2-km resolution storm-following mobile inner nest, and 81 vertical levels with a ~2 hPa model top. Both versions of HAFS run with a warm-cycled vortex initialization (VI) and a 4DEnVar data assimilation (DA). However, the threshold intensity in which the VI is enabled is slightly higher for HAFS-A (50 kt versus 40 kt for HAFS-B). HAFS-A runs in all basins globally, similar to HWRF, while HAFS-B only runs in the NHC and CPHC basins, similar to HMON, due to HPC resource limitations. Lastly, HAFS-A and HAFS-B run off of slightly different physics packages. Most notably, HAFS-A is running with GFDL microphysics, while HAFS-B uses Thompson microphysics in 2023.

HAFSv1 .0	Domain	Resolution	DA/VI	Ocean/Wave Coupling	Physics	Basins
HFSA	**Storm-centric with one moving nest, parent: ~78x75 deg, nest: ~12x12 deg**	**Regional (ESG), ~6/2 km, ~L81, ~2 hPa model top**	**Vmax > 50 kt warm-cycled VI and 4DEnVar DA**	**Two-way HYCOM, one-way WW3 coupling for NHC/CPHC basins**	**Physics suite-1**	**All global Basins NHC/CPHC/ JTWC Max 7 Storms to replace HWRF**

HFSB	Storm-centric with one moving nest, parent: ~75x75 deg, nest: ~12x12 deg	Regional (ESG), ~6/2 km, ~L81, ~2 hPa model top	Vmax > 40 kt warm-cycled VI and 4DEnVar DA	Two-way HYCOM No Waves	Physics suite-2	NHC/CPHC Max 5 Storms to replace HMON

Table 1: Comparison between the HAFS-A and HAFS-B configurations for the 2023 initial operating capability (IOC), implemented June 27, 2023.

Overall, in the 2023 season for all NHC basins, HAFS-A and HAFS-B performed superior to the legacy HWRF system at all lead times beyond 12 h for track (Figure 4a). The global models, such as NCEP's GFS and the ECMWF model are not particularly skillful for intensity due to horizontal resolution limitations relative to the mesoscale models. However, they remain quite competitive for track skill. When including the global models for comparison, HAFS-A and HAFS-B still had the best track forecast skill amongst available guidance except for the ECMWF for days 1-2, and superior to the ECMWF from days 3-5 in 2023 (not shown). For intensity, HAFS-A and HAFS-B outperformed HWRF from 60-120h, but is comparable to or in some cases lags behind HWRF (Figure 4b) as well as HMON (not shown) from 0-48 h. Also, on a case-by-case basis, the nature of the intensity bias for both versions of HAFS is also typically the same (either both strong bias or both weak bias). This comparative skill lag for intensity at early lead times and lack of differentiation in model physics, DA, and internal components between the two HAFS variants are areas of active research and development supported by HFIP.

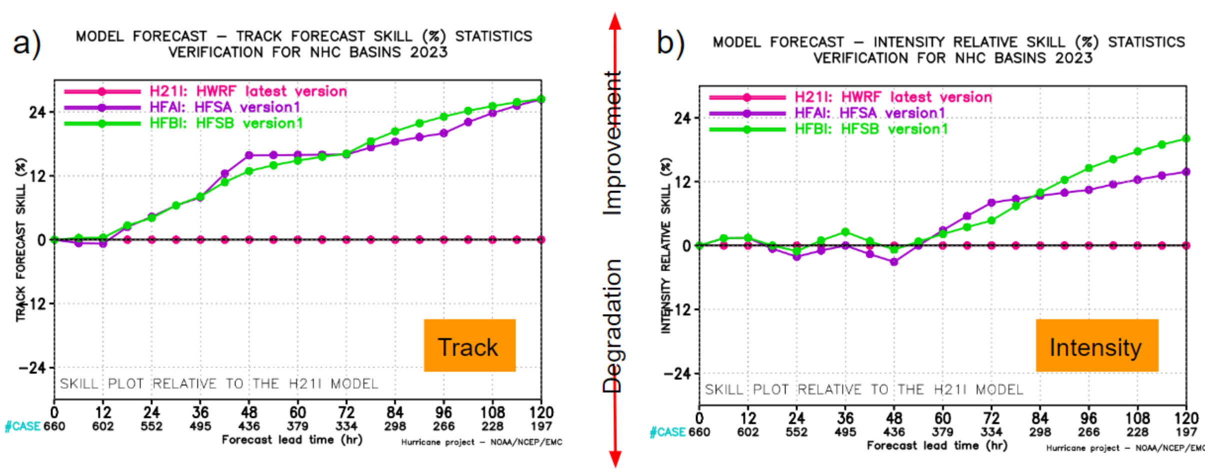


Figure 4: Model forecast skill (%) for (a) track, and (b) intensity in 6-hour increments from 0-120 hours for HWRF (magenta), HAFS-A (violet), and HAFS-B (green), relative to HWRF as the baseline for all storms in NHC basins in 2023.

Next, we stratify the results from 2023 by basin. For track, HAFS-A and HAFS-B were the most skillful models in both basins at most lead times (Figure 5a-b). GFS was comparable or superior to both HAFS variants from 12-18 h, while HMON was the top model for track from 24-36 h only. For intensity, HWRF and HMON outperformed HAFS-A and HAFS-B from 12-72 h, while HAFS-B was the most skillful from 96-120 h (Figure 5c). However, HAFS-A and HAFS-B were the top performing models for intensity in the East Pacific at all lead times, except for 12-24 h where HAFS-A is tied with HMON for most skillful (Figure 5d).

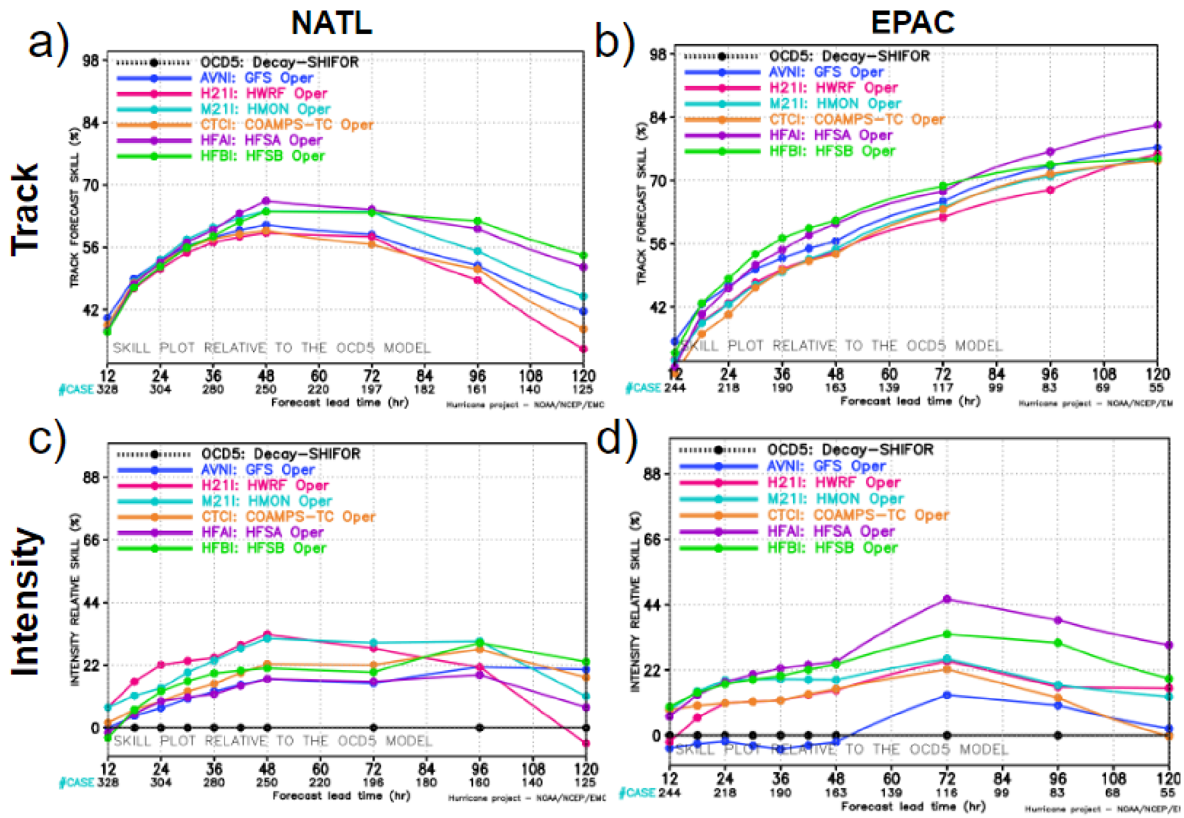


Figure 5: Forecast skill relative to a simple climatological model (OCD5 model) (a, b) for track, and (c, d) for intensity, stratified by basin, with results depicting results for (a, c) the North Atlantic, and (b, d) the East Pacific basins in 2023. Models included are OCD5 (black), GFS (blue) , HWRF (magenta), HMON (cyan), COAMPS-TC (orange), HAFS-A (violet), and HAFS-B (green).

While HFIP verification traditionally emphasizes the NHC basins due to the implications to impacts to the continental United States and Hawaii, the JTWC basins also include important U.S. territories and bases in Guam, Okinawa, as well as U.S. allies. As such, we also include verification statistics from the Western North Pacific (Figure 6). Note that amongst NWS mesoscale hurricane models, only HAFS-A and HWRF are run in the West Pacific. HAFS-B and HMON are only run in the NHC basins, due to HPC resource constraints. In the West Pacific in 2023, GFS was the clear leader for track, with the most skillful forecast at all lead times. However, HAFS-A also performed well, and was the second most skillful model for track from 48-120 h. For intensity, the Navy's COAMPS-TC model was the most skillful at most lead times, from 12-84 h. However, HAFS-A also performed adequately and was the second most skillful model for intensity in the West Pacific from 18-72 h in 2023.

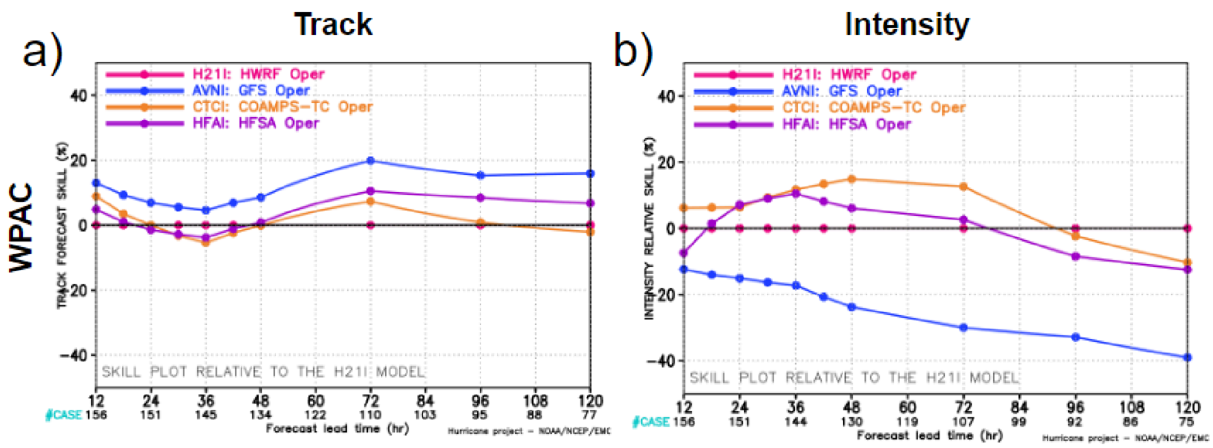


Figure 6: (a) Track, and (b) intensity forecast skill for the Western North Pacific in 2023. Models included for comparison are the HWRF (magenta), GFS (blue), COAMPS-TC (orange), and HAFS-A (violet).

There were several TCs in particular that HAFS struggled with in 2023. Hurricane Lee was arguably the most challenging TC for HAFS to predict in the Atlantic basin this season. In terms of positives, both HAFS variants produced stellar track forecasts, and correctly predicted that RI would occur. However, HAFS struggled significantly with the timing of the onset of RI, off by 12-18 hours in cases (Figure 7a,b,c). HAFS also over-intensified Lee for many consecutive forecast cycles, and was too slow during the rapid weakening phase. It is hypothesized that poor vortex initialization and DA issues contributed to the challenges in predicting Lee. It is also possible that other physics configurations would perform better for Lee. These hypotheses are being tested in retrospective runs for Lee in 2024.

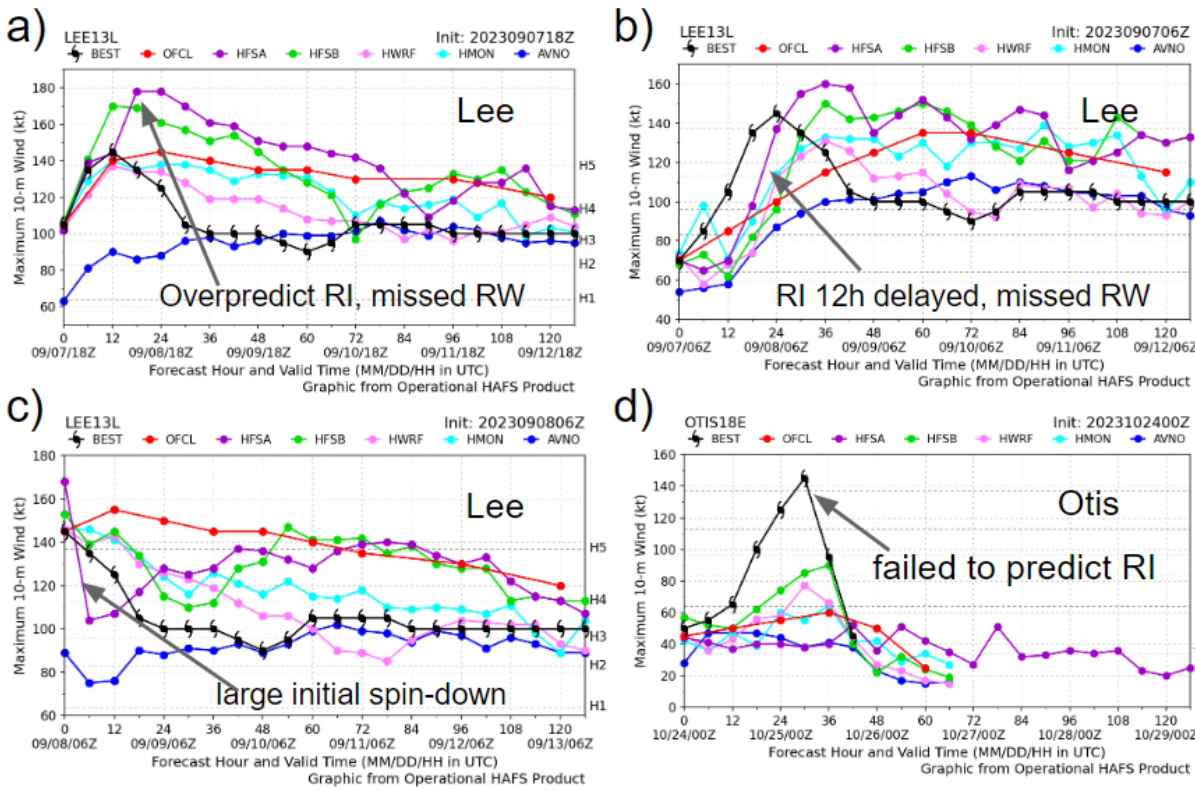


Figure 7: Forecast intensity (kt) for (a, b, c) Hurricane Lee from forecasts initialized (a) 18 UTC 07 Sep 2023, (b) 06 UTC 07 Sep 2023, (c) 06 UTC 08 Sep 2023, and (d) Hurricane Otis initialized 00 UTC 24 Oct 2023. Included are the verifying "best track" intensity (black), the NHC official forecast at the time (red), HAFS-A (violet), HAFS-B (green), HWRF (magenta), HMON (cyan), and GFS (blue).

Hurricane Otis in the East Pacific was also a particularly challenging case for HAFS, as well as all other guidance. Otis underwent a period of extremely rapid intensification, from 55 kt to 145 kt in 24 h, which was vastly underpredicted by both HAFS variants. It is suspected that a lack of inner-core observations during and prior to the period of RI contributed to a struggle with the VI and DA for Otis. This TC also featured an extremely small core, which makes it even more difficult to capture in models. It is possible that even with HAFS 2-km horizontal resolution, even finer resolution inner nest is necessary to capture inner-core intensification processes for TCs such as Otis.

Alternately to many of the aforementioned challenging cases, while NHC struggled greatly with the prediction of Hurricane Philippe, the two versions of HAFS outperformed all other models for Philippe's track (Figure 8). While Philippe still posed a challenge for HAFS-A and HAFS-B for track, as well as intensity, Philippe's forecast demonstrates a clear win for HAFS during a particularly high uncertainty event.

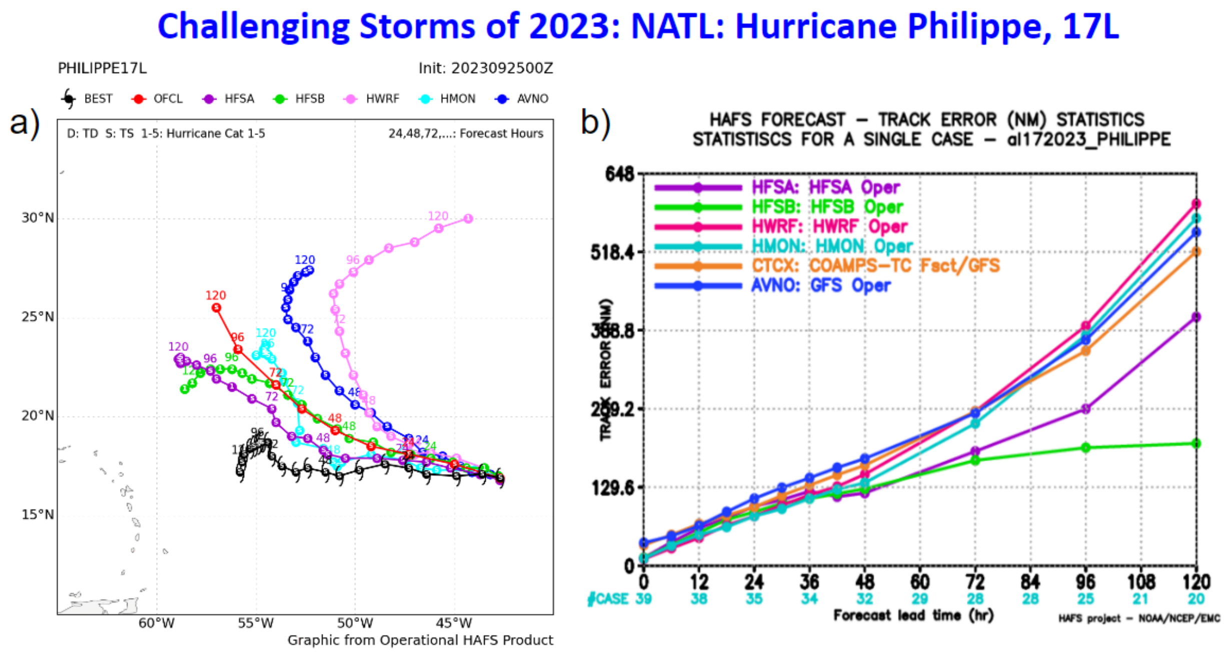


Figure 8: (a) Track forecasts for Hurricane Philippe initialized 00 UTC 25 Sep 2023. Tracks shown include the verifying "best track" (black), the NHC official forecast at the time (red), HAFS-A (violet), HAFS-B (green), HWRF (magenta), HMON (cyan), and GFS (blue). (b) Track forecast error (nautical miles; nm) averaged over 39 forecasts for Philippe, including HAFS-A (violet), HAFS-B (green), HWRF (magenta), HMON (cyan), COAMPS-TC (orange), and GFS (blue).

2.4 Highlights from the HFIP Real-time Experiments (HREx) and the HAFS Ensemble

a. HREx 2023

Each hurricane season, HFIP supports a handful of real-time experiments that are run on the Jet supercomputer with the goal of demonstrating high readiness level (RL) models and modeling subcomponents as part of the HFIP Real-time Experiments (HREx). The goal is to transition the most promising of the HREx performers during the off-season for operational implementation during the following hurricane season. In 2023, there were five HREx projects:

- HAFS v1.1A led by EMC
- HAFS v1.1B led by CIMAS and HRD
- MOM6 HAFS led by AOML and CIRA
- GFS Physics led by GSL
- T-SHiELD and SHiELD led by GFDL

We provide some basic details on each of the model configurations tested under HREx, and a simplified summary of the results. For further details, we refer the reader to the HREx slides on the HFIP Annual Meeting 2023 website.

HAFS v1.1A

The HAFS v1.1A experiment led by EMC is a modified version of HAFS-A v1.0, featuring a resolution increase from 6-2 km to 5.4-1.8 km (on grids 1 and 2, respectively), a switch from GFDL microphysics to Thompson microphysics, and MOM6 ocean coupling instead of HYCOM in coordination with the MOM6-HAFS HREx experiment (more details below). This experimental version also featured a number of DA advancements, including cloud hydrometeor relocation and cycling, a modified composite vortex for the vortex initialization (VI) scheme, a reduced max wind threshold from 50 kt to 40 kt for which warm cycling is enabled, and the addition of assimilating high-resolution GOES-R CONUS and mesoscale floater AMVs.

Overall, results from this experiment indicate improved skill for intensity on the order of ~5% from 48-120 h. The effect of these upgrades was essentially a wash in terms of track skill, with the experimental run having superior skill at 120 h, the control run outperforming the experimental run 72 h, and essentially identical results at all other lead times. However, the improvements to the model microphysics and in particular the hydrometeor cycling had a significant impact on improving the storm structure in HAFS. As can be seen in an example retrospective run from Hurricane Ian, the HAFS v1.1A experimental run (HFXA) is a much closer match to the NOAA P-3 tail Doppler radar analysis from HRD at the same time than the HAFS-A v1.0 control (CTRL) forecast (Figure 9). As such, due to clear improvements in TC intensity and structure, EMC plans to move forward with this experimental version of HAFS for implementation in HAFS v2 in 2024.

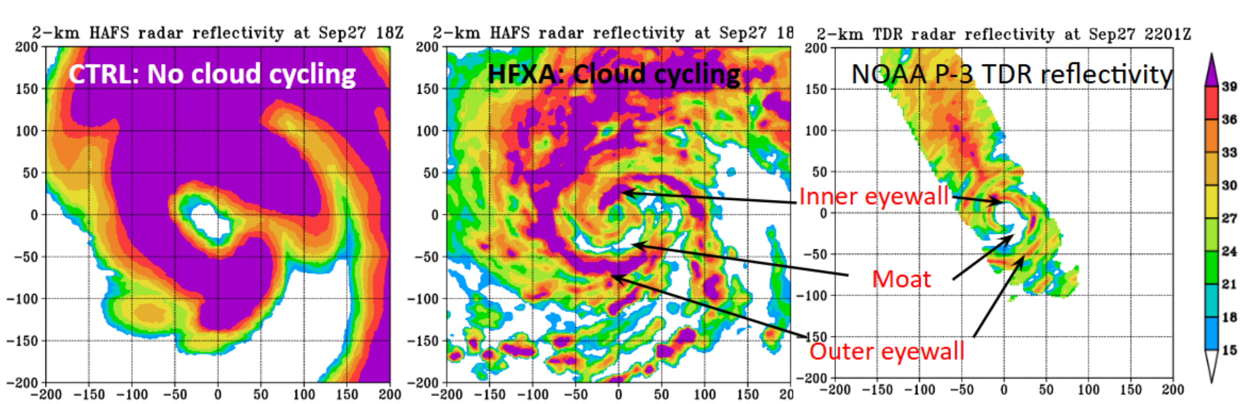


Figure 9: Comparison of simulated reflectivity at 2-km altitude for Hurricane Ian on 27 Sep 2022 from (a) HAFS-A v1.0 without cloud hydrometeor cycling, (b) HAFS v1.1A with cloud hydrometeor cycling, and (c) comparison against observed values from NOAA P-3 tail Doppler radar analysis.

HAFS v1.1B

The HAFS v1.1B experiment (HFXB) led by CIMAS and HRD is based upon a modified version of HAFS-B v1.0 (HFSB). This experiment featured a significant change in the Scale-Aware SAS convection scheme, namely using “prognostic sigma” prognostic closure (Bengtsson et al. 2022), along with other smaller modifications to the TC PBL and the HYCOM mixing. Comparing the results for intensity between HFXB and HFSB, the effect on intensity skill is neutral in the Atlantic, and slightly positive in the East Pacific from 78-120 h. There is also a slight improvement in storm structure, which shows up in simulated reflectivity but is also reflected in wind-radii verification. For 34-kt wind radii (R34), HFXB shows a slight improvement over HFSB. However, both versions have a low bias for R34 of 15-25 km at most lead times.

Overall, the greatest impact from the HFXB experiment was to track, with HFXB showing a 6-10% improvement over HFSB in the Atlantic from 24-120 h (Figure 10a), as well as an improvement of 5-10% in the East Pacific, to as much as a 30% improvement at 120 h (Figure 10b). Improvements of these magnitudes to track skill are fairly substantial, and are often not achieved from a single year of development and testing. As such, these improvements were slated for implementation into HAFS-B in 2024.

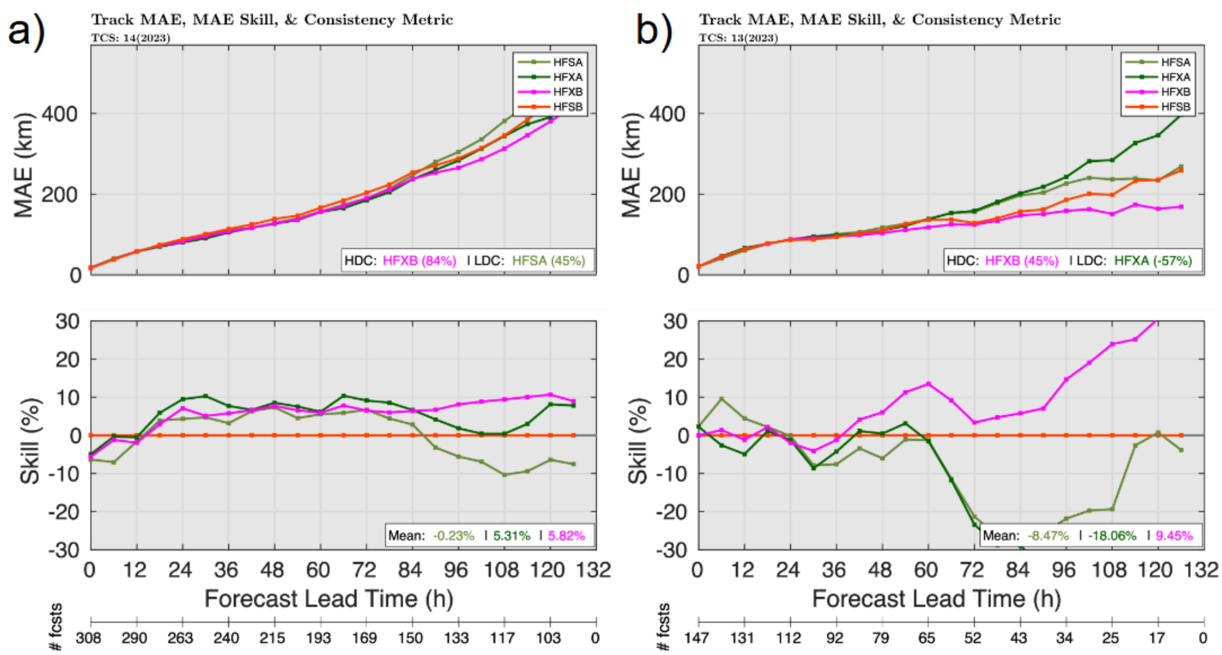


Figure 10: Track mean absolute error (MAE; top) and skill percentage change (%; bottom) as a function of forecast lead time (h) for (a) 308 forecast cases in the Atlantic, and (b) 147 forecast cases in the East Pacific in 2023. Models included in the comparison are the operational HAFSv1.0 -A (HFSA) and -B (HFSB), as well as experimental versions HAFSv1.1 -A (HFXA) and -B (HFXB).

MOM6 HAFS

The MOM6 HAFS experiment, led by a combination of AOML, EMC and CIRA, was run in coordination with EMC and the HAFS v1.1A experiment. While EMC mostly focused on the development and verification of the atmospheric component of this experiment, the AOML and CIRA group primarily focused on the development of the ocean model, and verification of the oceanic 3-D fields and atmosphere via coupled processes. In one such example from Hurricane Lee at 06 UTC 07 Sep 2023, a comparison is made for sea surface temperature (SST) and salinity between operational HAFS v1.0, three different iterations of HAFS v1.1

including two versions that are coupled to MOM6, HWRF coupled to the POM ocean model, and verification from a nearby saildrone (Figure 11). Overall, HAFS v1.1 coupled to MOM6 with the EPBL option is the best performing model for SST, and the second best for salinity.

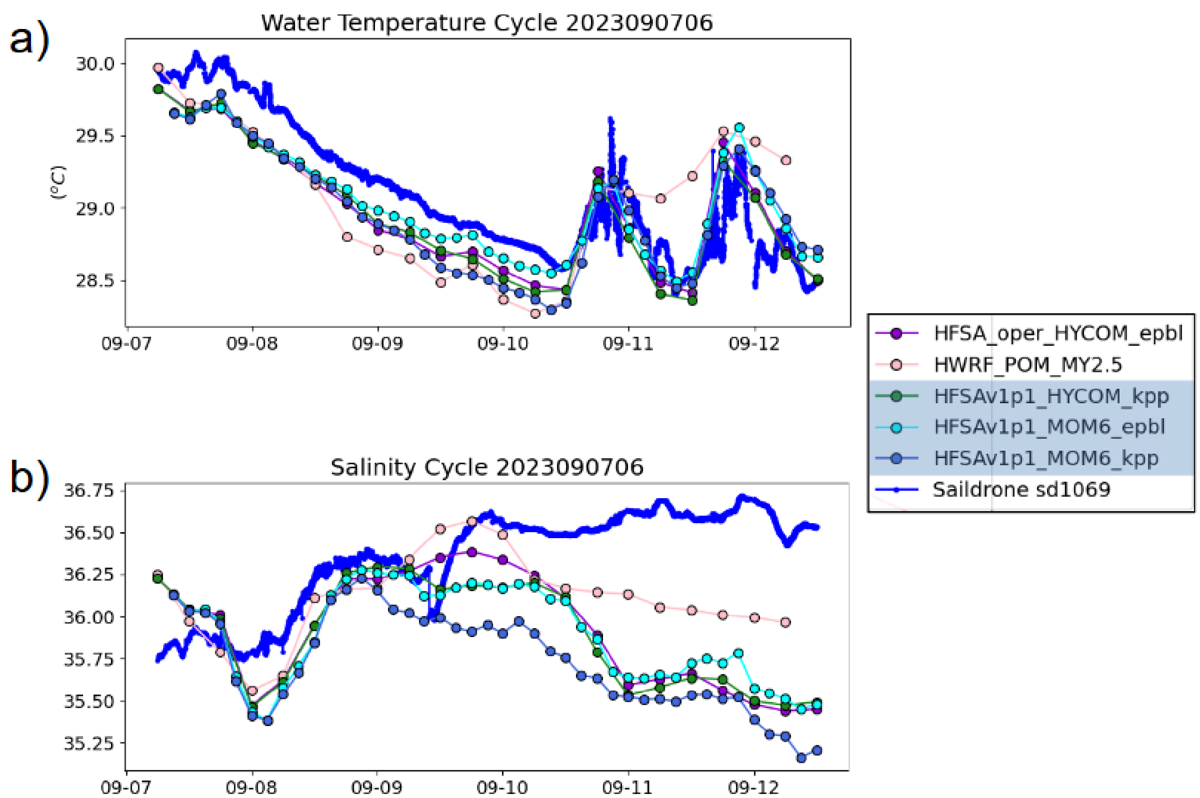


Figure 11: Model forecast versus observations for (a) SST (°C), and (b) salinity (psu). Included model forecasts are HAFS-A v1.0 with HYCOM and EPBL (magenta), HWRF with POM and MY2.5 (peach), HAFS-A v1.1 with HYCOM and KPP (green), HAFS-A v1.1 with MOM6 and EPBL (cyan), HAFS-A v1.1 with MOM6 and KPP (thin blue line with filled circles), and Saildrone observations from sd1069 (thick blue line).

GFS Physics

The GFS Physics experiment led by GSL differs from the aforementioned experiments in that it pertains to testing of model physics in the global model, as opposed to the regional hurricane model. However, given the fact that the GFS is still a top source of guidance information for NHC, particularly for track and genesis, but also increasingly for intensity trends (especially when bias corrected), a focus on improving the representation of TCs in GFS is also well within the real of HFIP goals. Additionally, it should be noted that all HAFS forecasts rely upon GFS initial conditions and boundary conditions. As such, improvements to GFS will also feed back into HAFS, potentially improving the HAFS forecast as well.

In this experiment, GSL tested a variety of different physics configurations in GFS, including comparing the existing TKE-EDMF PBL in GFSv17 to the MYNN-EDMF PBL, comparing both Community Convective Cloud (C3) and MYNN-EDMF shallow convection to the existing SAS shallow convection, and comparing the C3 deep convection to the existing SAS deep convection. Additionally, tests were conducted comparing the existing non-aerosol-aware Thompson microphysics scheme with the aerosol-aware version.

Using default physics parameters and options, none of the new schemes outperformed the baseline GFSv17 in the Atlantic. However, experiments with the C3 scheme outperformed the control GFS in the East Pacific. Some biases were identified in the precipitation patterns and cloud cover output using the new physics, which will require some adjustment of physics parameters to address. Overall these results show promise, and warrant further investigation.

T-SHiELD and SHiELD

In this HREx experiment, GFDL ran a real-time experiment using their 13-km resolution global SHiELD model, and their 3-km resolution regional nested T-SHiELD model. Note that T-SHiELD uses vortex initialization and explicitly resolved convection, akin to the HAFS inner nest, while SHiELD does not use VI and uses SA-SAS convective parameterization, more analogous to GFS.

The SHiELD model performed quite well in the Atlantic in 2023, outperforming the GFS from 24-120 h and comparable with the top-performing ECMWF at 120 h. The model also performed quite well for track in the East Pacific and West Pacific basins. T-SHiELD performed even better than SHiELD in terms of track skill, with comparable scores to HAFS-A and -B at most lead times.

T-SHiELD did not perform quite as well for intensity in 2023, with larger errors than either HAFS variant or the legacy HWRF and HMON models. The model featured a pronounced weak bias from 0-60 h, which resulted in larger errors. Interestingly, GFDL found the initial weak bias to be *more* pronounced in cases in which VI was used, a technique which is typically used to reduce spin-up issues. This suggests further tuning of the VI will be necessary in future versions of the model. Lastly, it is noted that both SHiELD and T-SHiELD exhibited a pronounced negative bias in terms of R34. TCs in the model are too small, and actually become smaller with time. Overall, this work suggests some promising results, particularly in terms of track prediction. However, additional tuning will be needed to mitigate the low intensity and small TC biases in future experiments.

b. HAFS Ensemble

In 2023, EMC debuted a proof-of-concept HAFS ensemble that ran in real-time. Prior to elaborating on the details, we will first address the *need* for the NWS to run an experimental ensemble for hurricane prediction. The HAFS Ensemble is motivated by NHC and JTWC operational requirements, along with the NWS Strategic Plan under “Ken’s 10”. As stated by NHC, “we currently lack a dynamical regional hurricane model ensemble system that can represent intensity and structure” (Mike Brennan, NHC, 9th NOAA Ensemble Users Workshop), and “a mesoscale NWP TC Ensemble – a missing link for hazard magnitude, timing and uncertainty communication” (Wallace Hogsett, NHC, 9th NOAA Ensemble Users Workshop). Similar concerns were also voiced by JTWC: “there is a lack of skillful high-resolution ensembles to predict TC intensity and structure change, to complement COAMPS-TC ensemble” (Matthew Kucas, JTWC, 9th NOAA Ensemble Users Workshop). Additionally, under the NWS Strategic Plan, a HAFS Ensemble directly addresses the need for probabilistic IDSS for hurricane forecasting.

Due to a lack of sufficient on-prem HPC resources, the HAFS Ensemble was run on the AWS cloud in 2023. The first task of running the HAFS Ensemble on the cloud was porting HAFS related codes, scripts, and workflow. Leveraging past and future UFS-based applications on AWS enhances avenues of collaboration and community contributions to future versions of HAFS. Year 1 of this experiment has demonstrated a need for further testing and evaluation to optimize strategy, including but not limited to exploring R&D issues related to domain configurations, moving nests, physics options, stochastic physics, ensemble size, and cadence of running the forecast. Probabilistic numerical guidance is an outlined goal for HFIP, along with the need to further explore products that convey forecast uncertainty. HFIP will also support the exploration of multi-model or multi-core ensembles in the future. This activity leverages existing Disaster Relief Supplemental projects for use of dynamic inputs via hurricane ensembles for storm surge and flooding downstream applications within NHC and NOS.

In 2023, the HAFS Ensemble ran 4 cycles per day (00, 06, 12, and 18 UTC) with 21 members based upon HAFS-A v1.0, with one control member and 20 members with perturbed initial conditions and stochastic physics. The ensemble was run using a static 6-km horizontal resolution grid that encompassed the tropical North Atlantic and part of the Eastern North Pacific, with 66 vertical levels and a model top of 2 hPa. IC/BC perturbations were pulled from the operational GEFS, and stochastic physics perturbations included SPPT, SKEB, and SHUM. For simplicity in this “proof of concept” version, the ensemble ran without vortex initialization or data assimilation. The computer resource requirement in the AWS cloud was 6 nodes or 576 cores per member per cycle.

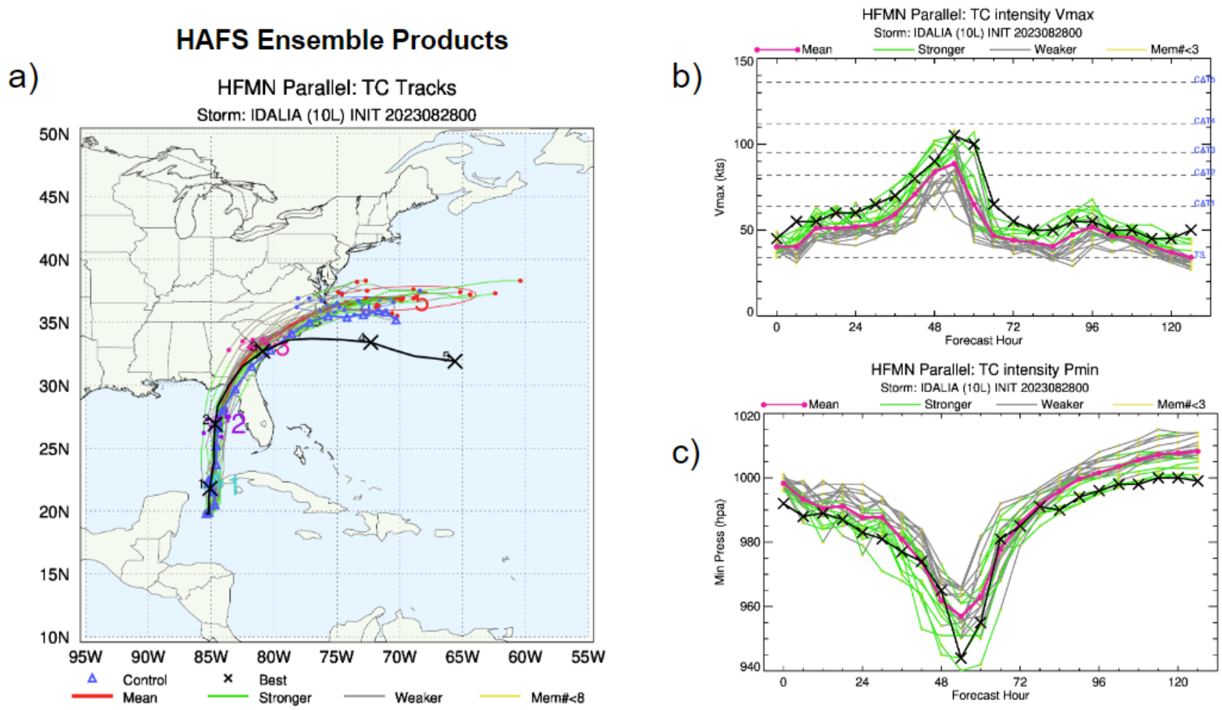


Figure 12: Experimental 21-member HAFS Ensemble forecast for Hurricane Idalia (10L) from 00 UTC 28 Aug 2023 depicting probabilistic forecast information for (a) track, (b) intensity (maximum 10-m wind; kt), and (c) MSLP (hPa). In (a), colors depict members stronger than the median intensity (green), members weaker than the median intensity (grey), the ensemble mean (red), the control member (blue), the verifying best track (black); ellipses depict ⅔ of the distribution of members fitted to a Gaussian distribution along the ellipse major and minor axes; numbers label lead time (day). In (b) and (c), colors depict members stronger than the median intensity (green), members weaker than the median intensity (grey), the ensemble mean (magenta), and the verifying best track intensity (black).

An example forecast from the 00 UTC 28 Aug 2023 cycle for Hurricane Idalia is shown (Figure 12). In this forecast, the HAFS Ensemble performed extremely well for track for days 1-3, although it was north / left biased at days 4-5. Note that the ellipses depicting ⅔ of the track variance grow with time, depicting greater forecast uncertainty at longer lead times. Some preliminary verification of spread-skill scores has also shown that, in general, larger ellipses correspond to greater uncertainty or greater mean error, while smaller ellipses correspond to lower mean error and lower uncertainty. As such, the ensemble is reasonably well calibrated. The ensemble also performs reasonably well for intensity in this forecast, despite the challenges associated with predicting RI. Approximately 70% of the ensemble correctly predicts that RI will occur in this case, with accurate timing for onset of RI. However, most ensemble members were too weak in terms of the maximum intensity. More generally, the HAFS Ensemble sometimes misses RI, particularly for smaller hurricanes, where 6-km horizontal resolution may be insufficient to resolve the developing eyewall.

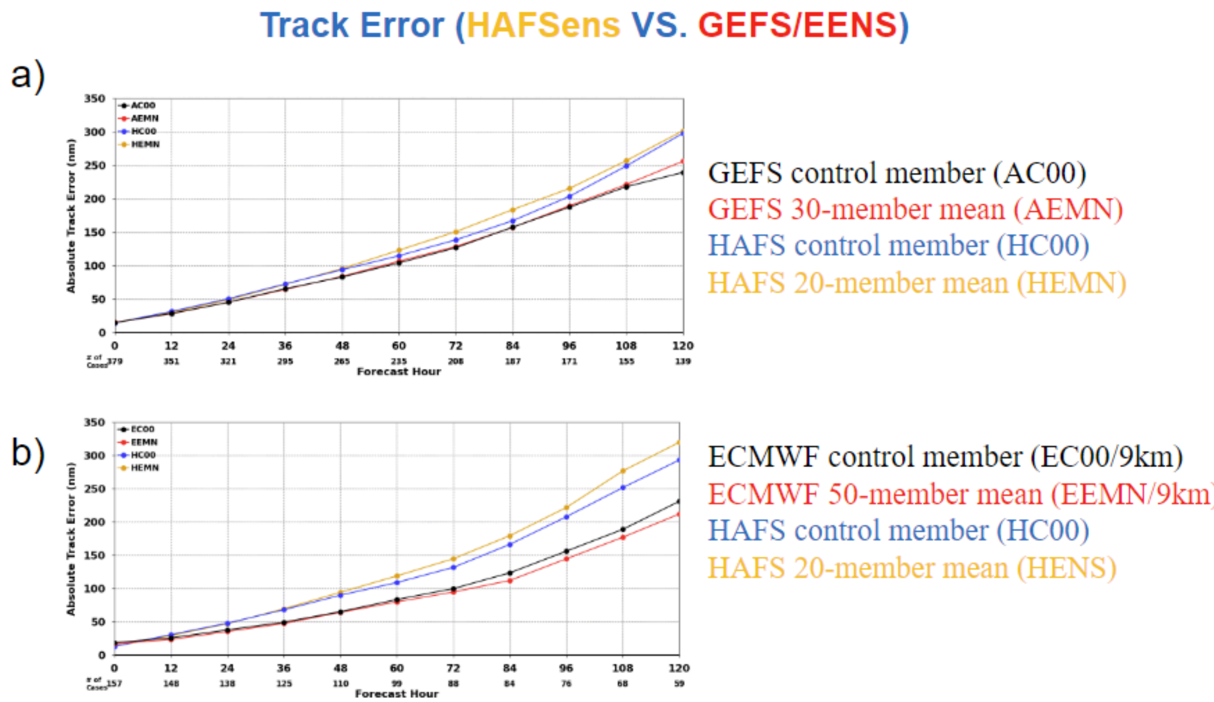


Figure 13: Track forecast error (nm) as a function of lead time (h) for the HAFS Ensemble control member (blue), the HAFS Ensemble mean (yellow), compared against (a) the GEFS control (black) and ensemble mean (red), and against (b) the ECMWF control (black) and ensemble mean (red).

Across a larger sample that includes the entire 2023 season, 379 cases total, error and biases for the HAFS Ensemble control member and the ensemble mean are compared with GEFS and the ECMWF ensembles and control members. For track, the HAFS Ensemble mean and control track error are comparable to GEFS (Figure 13a) through 60 h, and associated with slightly greater error from 72-120 h. The HAFS Ensemble is associated with greater track error than the ECMWF Ensemble at all lead times (Figure 13b); however, NHC found the ECMWF Ensemble to have the most accurate track forecast amongst all guidance in 2023, so this is a high bar to cross. Interestingly, the HAFS Ensemble control member is associated with slightly lower mean track error than the ensemble mean, which is atypical for an ensemble. This suggests that perhaps the GEFS ensemble perturbations are not the optimal perturbations for track spread in the HAFS Ensemble.

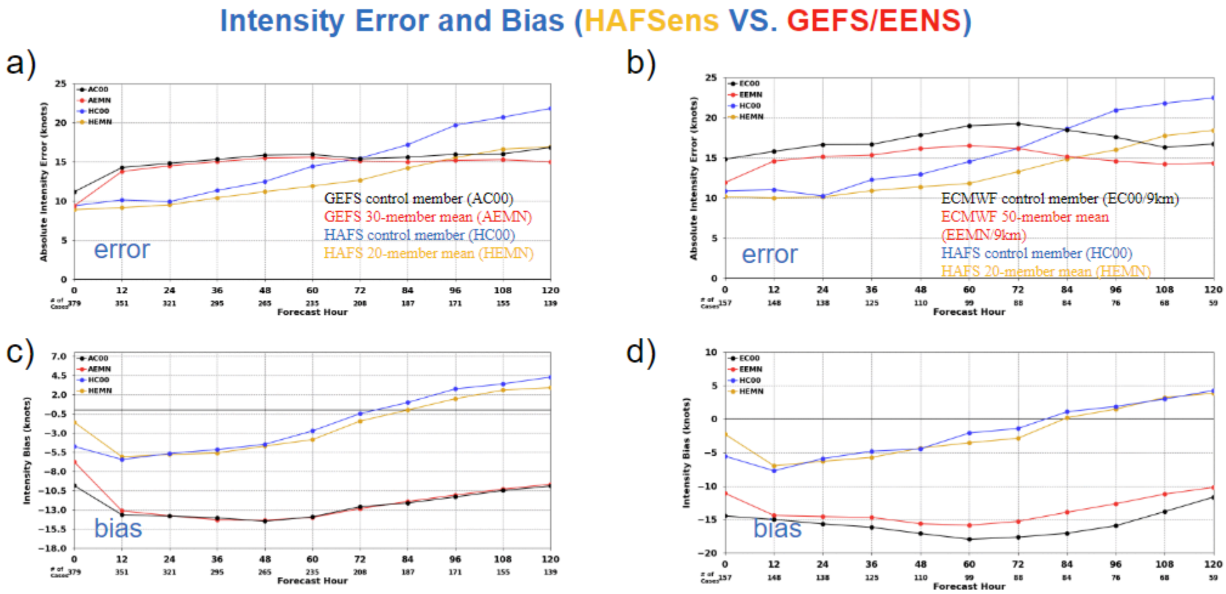


Figure 14: Intensity error (kt; a,b) and bias (kt; c,d) of the HAFS Ensemble control member (blue) and mean (yellow) compared to the GEFS control (black) and ensemble mean (red) in (a,c), and against the ECMWF control (black) and ensemble mean (red) in (b,d).

Lastly, we examine the intensity error and bias of the HAFS Ensemble compared to the GEFS and ECMWF ensemble means and control members (Figure 14). Unlike for track, the advantages of running a high-resolution mesoscale ensemble become much more apparent for intensity. The HAFS Ensemble, particularly when using the mean intensity, is superior to the GEFS or ECMWF ensemble means or control members from 0-72 h, and is comparable to both global ensemble suites from 84-120 h. The HAFS Ensemble is also much closer to zero mean bias, whereas the GEFS and ECMWF ensembles are associated with significant weak biases for intensity. We expect the advantage of the HAFS Ensemble to grow even further in future versions that utilize higher resolution and/or nesting as is done in deterministic operational HAFS.

2.5 Experimental Product Visualization and IDSS

Various other products were tested and debuted in 2023 that are consistent with HFIP and the 2017 Weather Act goals of improved tools for forecasters and emergency managers to make crucial decisions related to a variety of hazards that come with a landfalling TC. Some highlights from recent advancements are highlighted below.

The AOML Hurricane Model Viewer now provides a full array of forecast graphics on an easy-to-use interface for a total of eleven different operational and experimental models.

NHC’s storm surge watch and warning system is testing a new polygon-based approach, which will be live on AWIPS for further testing in 2024. The new system allows storm surge watches and warnings to be applied specifically to regions along the coastline that are at threat without unnecessary over-warning of regions with sufficiently high elevation. An example showing the intricacies of the polygon based warning system in the Fort Myers region of Florida is shown in Figure 15.

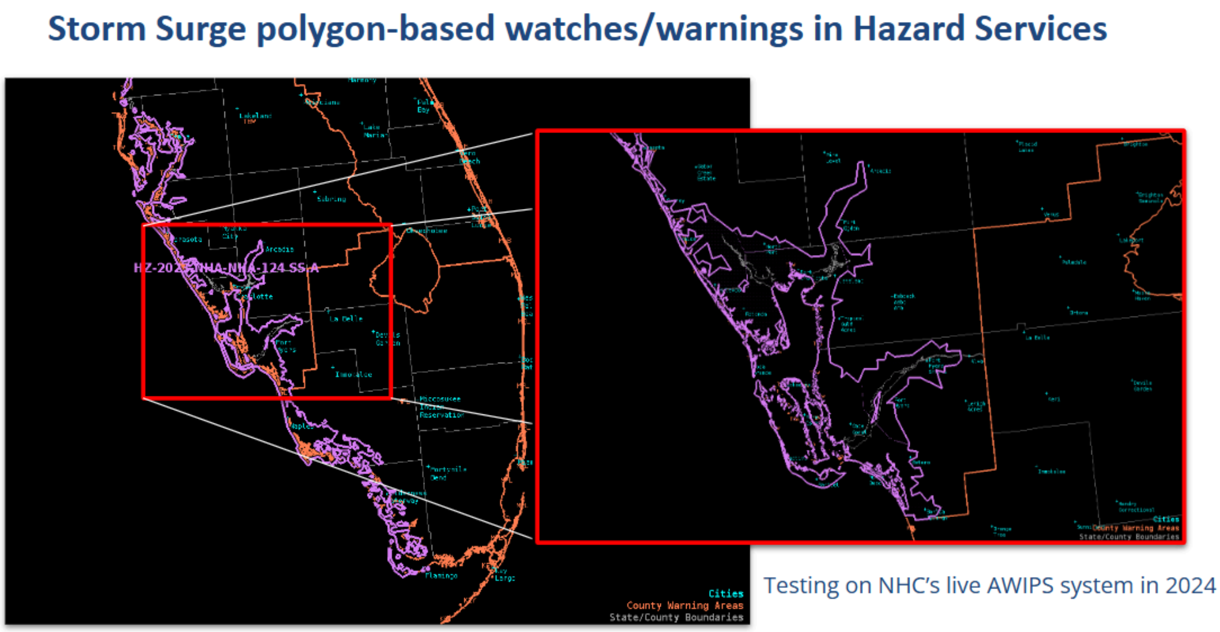


Figure 15: An example of the new polygon-based storm surge watch and warning system over the Fort Myers region of Florida.

NHC has also adjusted their Wind Speed Probability model (WSP) to account for land reduction (due to friction), while retaining the higher wind values over inland bodies of water, to give more realistic values that are more consistent with observations. An example comparing the legacy WSP to WSP 2.0 for Hurricane Ian is shown in Figure 16. The new and improved version, WSP 2.0, is being tested in the Hurricane and Ocean Testbed (HOT) and is targeted for operational implementation in the next few years.

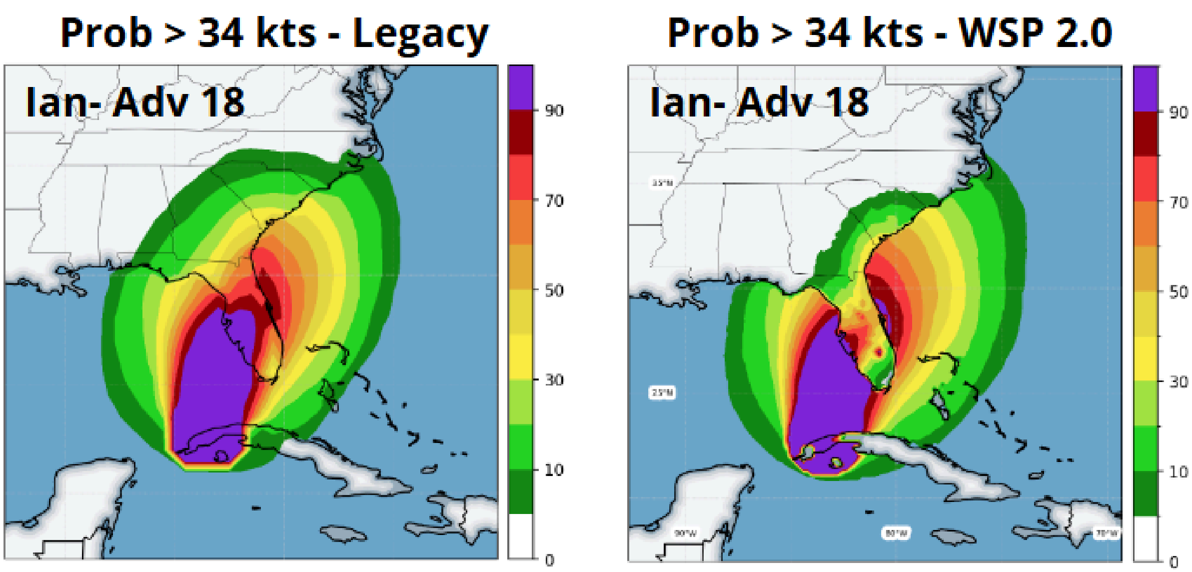


Figure 16: NHC's Wind Speed Probability model (WSP) example for Hurricane Ian advisory 18, showing the legacy system on the left and the new experimental WSP 2.0 on the right.

NHC also began testing the integration of hurricane ensemble guidance in DESI, a cloud-based ensemble interrogation platform developed by NOAA/GSL. This platform facilitates forecaster analysis of ensemble systems, and the example below (Figure 17) depicts the arrival time of winds during Hurricane Idalia, as predicted by the HAFS ensemble in the cloud. This prototype demonstration showed promise for future operational use of this visualization and analysis technology with hurricane ensemble systems.

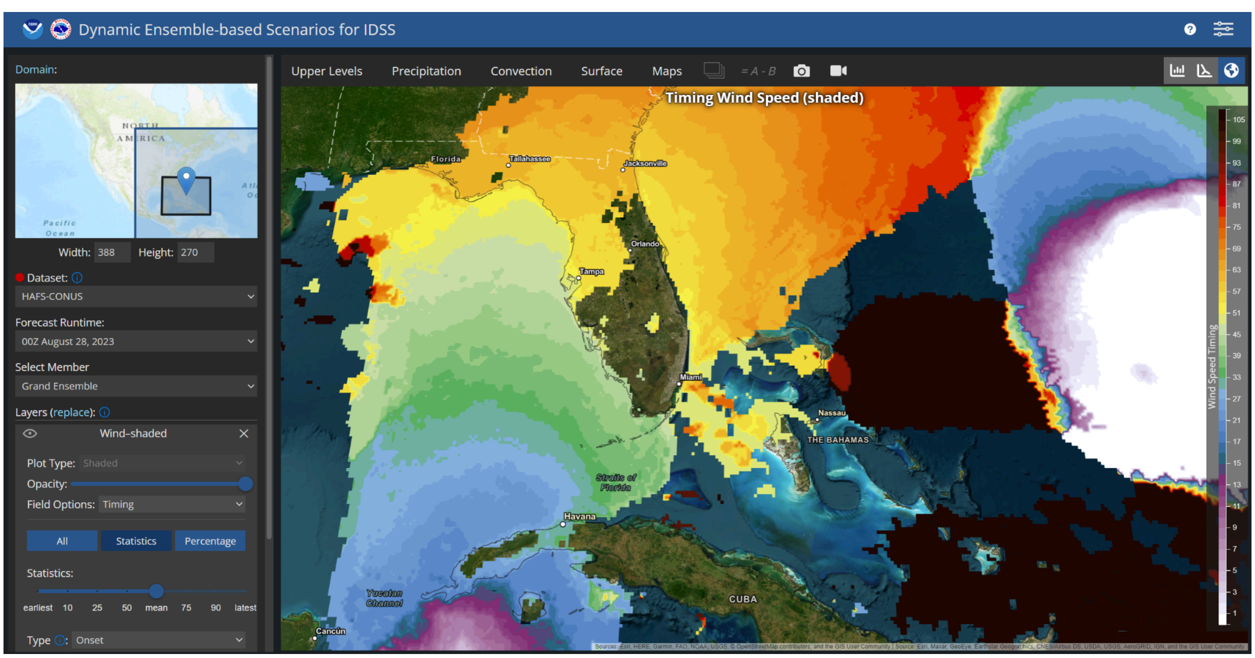


Figure 17: HAFS ensemble data integrated with the DESI analysis platform, depicting the "time of arrival" of tropical storm winds during Hurricane Idalia.

All-hazard risk communication has also been a thrust of HFIP that has been receiving increased attention in recent years. NHC is coordinating with SPC, WPC, and local WFOs to implement new risk maps that depict threats from a variety of hazards across localities, including wind, storm surge, freshwater flooding from rainfall, and tornadoes. An example of these new products from Hurricane Ian is shown in Figure 18. Preliminary results have been mostly positive in terms of the public's perception of the new products, but there are still issues in terms of misunderstanding that need to be addressed.

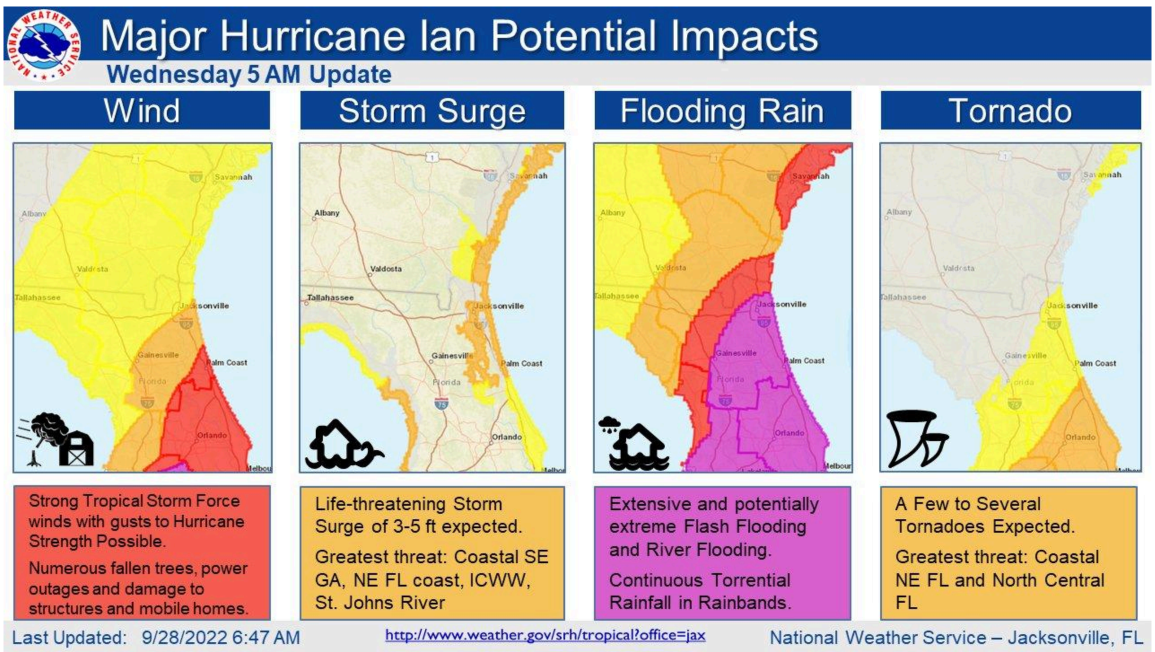


Figure 18: A new NWS Hurricane Potential Impacts summary bulletin, which depicts severity of impacts across different areas in a WFO area of responsibility. In this example, from left to right, impacts from wind, storm surge, flooding rain, and tornadoes are included.

Additionally, it should be emphasized that the NWS has identified changing trends in terms of how Americans receive their weather information. Fewer and fewer people rely on radio, print, or even television anymore, while a majority (~53%) get their weather information from a digital device. The issue of communication becomes a challenge when the source of information is from social media, where many people in the general public cannot differentiate between official information and unofficial or misleading information.

3. Community and Annual Engagement

3.1 Summary of the HFIP Annual Meeting 2023

The Hurricane Forecast Improvement Program (HFIP) annual meeting took place in Miami, FL, November 14th through the 17th. The meeting hosted 144 participants from the HFIP community spanning the public, private, and academic sectors, as well as international partners.

The primary objective of the annual meeting is to discuss the current operational Hurricane Analysis and Forecast System (HAFSv1) assessments, current and future hurricane modeling and product development activities, and update key HFIP strategic priorities.

The meeting discussed early results & lessons learned from real-time developmental experiment results for future upgrades to HAFS. In 2023 hurricane season, HAFSv1 has demonstrated superior track and intensity forecast skills over HWRF, with track forecast skill improved by more than 15% from days 2-5, and intensity forecast skill by ~10-20% from days 3-5 for storms within the National Hurricane Center (NHC) area of responsibility. While having two operational variants of HAFS, HAFS-A and HAFS-B, provided some insight as to forecast uncertainty, the results from 2023 demonstrate that a full HAFS ensemble will be necessary to adequately quantify uncertainty. This information will be used to inform and outline the work required to progress HAFS toward multiple high-resolution moving storm-centric nests across all global basins, anticipated in 2027. For the first time in the history of the HFIP annual meeting, presentations from an emergency manager and a cruise line meteorologist provided additional context as to how significantly local government and private industry have benefitted from HFIP, while also providing additional context as to how the National Weather Service (NWS) can better tailor forecast products and services to a broader set of users.

The collaboration at the annual meeting, along with the presented data will foster developing efficient pathways forward to progressing a world-leading, reliable, and skillful model guidance on TC track and intensity (including rapid intensification), storm size, genesis, storm surge, rainfall, and tornadoes associated with TCs and Socio-Economic impacts. Additional objectives include further development of ideas to be included in the new 2025 HFIP Strategic Plan with revised 5-year and 10-year goals. These goals will focus on advancing forecast and communication of all hazards from TCs, incorporate risk communication research to create more effective watches & warnings, and produce improved probabilistic risk products with a focus on vulnerable communities and industries through the use of social, behavioral, and economic sciences. Lastly, a variety of approaches were discussed at the annual meeting as to how to further enhance the role of HAFS as the UFS Hurricane Application while fostering even deeper integration of ideas and potential from other aspects of the larger UFS community both inside and outside of tropical meteorology. Overall, the annual meeting proved to be a great success through contributions from many across and beyond the HFIP community, summarizing the highly successful history of the program, and setting the stage for a realigned focus for HFIP over the next 10 years.

3.2 Major Actions and Takeaways

- **Actions**
    - Strategic Plan 2025: We need to create a new Strategic Plan for 2025, with updated 5-year and 10-year goals. Some things to think about when formulating these updated goals based upon discussions at HFIP Annual Meeting:
        - Advance the operational hurricane analysis and forecast system (HAFS),

including a HAFS-JEDI transition plan

- Improve probabilistic forecast guidance, by quantifying intensity, track, structure (RMW, R34) uncertainty (how do we measure R34 in model vs obs?), and greater emphasis on leveraging tools such as reanalysis and machine learning
- Enhance communication of risk and uncertainty. Application of triangulation results to product suite (AFS Tropical Roadmap)
- Support Dedicated High Performance Computing Allocation - Dedicated RDHPC vs Cloud
- R2O Enhancements (HOT, DTC, EPIC, UFS R2O)
- Broaden expertise and expand interaction with external community via EPIC and UFS

- Tiger Teams: We will form two tiger teams, one for data assimilation (DA) and the other for uncertainty quantification (UQ)
  - The DA team will focus on the need to replace GSI with JEDI. Which model infrastructure, physics, etc updates will need to be delayed in order to test, evaluate, and implement JEDI? If other model development continues at same speed, it will continue to delay JEDI implementation. DA tiger team will develop an implementation plan and oversee progress and eventual transition. They will target the HAFSv3 FY25 transition, coincident with GFSv18.
  - *Ultimately the HFIP-specific DA tiger team was superceded by a NWS-wide DA tiger team, which has drafted a new strategic plan for DA across NWS*
  - The UQ tiger team will provide a proof of concept: Provide storm specific ensemble model statistics to produce PDFs of model parameters (track, wind, RMW, R34, etc.) to develop, test & evaluate uncertainty information guidance for WTCM, WSP, P-surge, & P-Rain operational forecast guidance products (HAFS ensemble, COAMPS-TC ensemble, GEFS, etc.) using ML approaches e.g., DESI, TCANE. First develop a proof of concept, develop new probabilistic tools and products, and then calibrate and verify these products. Explore cost versus benefit of single-model versus multi-model ensemble. Leverage ML approaches, such as DESI and TCANE.
  - *Some discussion as to how to get the most value out of the UQ tiger team has been discussed, including what are its goals and priorities. However, this tiger team has not officially been designated yet.*
- HAFS Training Tutorial: HFIP will work with the community, EPIC, and DTC to coordinate a HAFS tutorial to increase the user base, as was done with HWRF in the past.

- **Takeaways**
  - Need to establish new 2025 strategic plan with updated 5 and 10 year goals. First strategic plan was in 2009, last updated in 2019. Need clear goals, as was done in the past.
  - Transition from GSI to JEDI will be a major step that we need to take.
  - Improved probabilistic guidance, quantification of uncertainty. Can we better leverage reanalysis, or machine learning for this?
  - Enhance communication, particularly with respect to risk uncertainty
  - Need dedicated HPC support, and strategy for continued adoption of cloud.
  - R2O enhancements and collaboration with HOT, DTC, EPIC, and UFS R2O. Broaden expertise and expand community, particularly through EPIC and UFS.
  - Need to increase spread in the models, particularly between HAFS-A and -B. Physics perturbations and different package options helps, but also need spread in the initial conditions.
  - Need more focus on storm structure, particularly in terms of size of the wind field. Not only does a larger wind field impact a larger region, but it also is the main driver of storm surge.
  - Why are there differences in model forecasts in large bust cases? For example, why was HAFS-B consistently better than HAFS-A for Philippe (2023)?
  - What can the global system do to make our lives easier on the TC side? At which point is the GFS cycling good enough that HAFS does not have to run its own cycling? 6-km resolution DA using JEDI is consistent with GFSv18 development ~5 years out.
  - NHC would like to see us optimize and verify the wind speed probability thresholds. Need to make sure we get data into AWIPS.
  - Need to think about physics issues, particularly in the “gray zone” for parameterized deep convection.
  - NHC: top priority should be to make sure spread-error score is appropriately tuned for track and intensity at all lead times.
  - Need to develop improved precipitation forecasts for TCs. Inland freshwater flooding is a major source of damage and loss of life.
  - Need to verify probabilistic wind swath and probabilistic precipitation forecasts. Does a 40% chance of wind above X threshold or precipitation above Y threshold verify 40% of the time?

4. Summary and Concluding Remarks

The 2023 HFIP Annual Meeting represented somewhat of an inflection point for the program. The theme of the meeting was a review of the progress that has been made since the beginning of HFIP, including a history of how the program has adjusted to changing forecaster needs, fluctuations in funding and personnel, and guidance from leadership. It was an eventful year for HFIP, as 2023 was the year of the operationalization of HAFS and the streamlining of EMC's hurricane suite under the UFS, and also the debut of an experimental 21-members HAFS Ensemble run on the AWS cloud. While the ultimate goal is to retire the legacy HWRF and HMON systems in order to free up personnel and CPU to dedicate entirely to more advanced, future versions of HAFS and the experimental HAFS Ensemble, the decision will not be made lightly without the support of NHC. Discussions emerging from the 2023 annual meeting have also highlighted the need for better ways to quantify the impact of hurricane research and improved hurricane forecasts. How many lives do we save each hurricane season with more accurate forecasts and improved forecast communication? How many dollars are saved by reducing the area unnecessarily under a hurricane warning, thereby reducing the size of the evacuation? These are just examples of some of the metrics by which we seek to measure the success of HFIP, moving forward.

Exhaustive discussions of lessons learned have set the stage for 2024, which will be forward-thinking and focus on the development of a new Strategic Plan with revised 5-year and 10-year goals. The new strategic plan will outline a new vision for HFIP, aligning our goals with the NWS 2023-2033 Strategic Plan, particularly in terms building expertise and tools to increase our capacity to understand, interpret, and communicate risk-based/probabilistic information to drive probabilistic Impact-Based Decision Support Services (prob-IDSS) and accelerating the transition from product and service development to deployment with rapid prototyping, operations proving grounds, and testbeds. As data assimilation (DA) currently sits near the forefront of Numerical Weather Prediction (NWP) development at the NWS, and since the translation from GSI to JEDI represents a major paradigm shift that will have major implications for hurricane modeling, care has also been taken to align the vision of DA for HFIP with EMC's Data Assimilation Strategy (Kleist et al. 2024), and a soon-to-be-released NWS Data Assimilation Strategy. The new HFIP Strategic Plan 2025-2035 will be close to final form in time for the 2024 HFIP Annual Meeting, and will be available from the NOAA Central Library by early 2025.

5. List of HFIP Publications

Journals and Periodicals

Munsi, A., A. P. Kesarkar, J. N. Bhate, and V. Tallapragada, 2024: Helicity: A Possible Indicator of Negative Feedback Initiation of Tropical Cyclone–Ocean Interaction. *Earth and Space Science*, **11**, e2023EA003211. https://doi.org/10.1029/2023EA003211

Shin, J., Z. Zhang, B. Liu, Y. Weng, Q. Liu, A. Mehra, and V. Tallapragada, 2024: The implementation of cloud and vertical velocity relocation/cycling system in the vortex initialization of the HAFS. *Atmosphere,* **15,** 1006, doi:10.3390/atmos15081006.

Kim, H.-S., B. Liu, B. Thomas, D. Rosen, W. Wang, A. Hazelton, Z. Zhang, X. Zhang, and A. Mehra, 2024: Ocean component of the first operational version of Hurricane Analysis and Forecast System: Evaluation of HYbrid Coordinate Ocean Model and hurricane feedback forecasts. *Front. Earth Sci.,* **12,** 1399409, doi:10.3389/feart.2024.1399409.

Gramer, L. J., J. Steffen, M. Aristizabal, and H.-S. Kim, 2024: The Impact of Coupling a Dynamic Ocean in the Hurricane Analysis and Forecast System. *Front. Earth Sci.,* **12,** 1418016, doi:10.3389/feart.2024.1418016.

Wang, W., J. Han, J. Shin, X. Chen, A. Hazelton, L. Zhu, H.-S. Kim, X. Li, B. Liu, Q. Liu, J. Steffen, R. Sun, W. Zheng, Z. Zhang, and F. Yang, 2024: Physics schemes in the first version of NCEP operational hurricane analysis and forecast system (HAFS). *Front. Earth Sci.,* **12,** 1379069, doi:10.3389/feart.2024.1379069.

Aristizábal Vargas, M. F., H.-S. Kim, M. L. Hénaff, T. Miles, S. Glenn, and G. Goni, 2024: Evaluation of the ocean component on different coupled hurricane forecasting models using upper-ocean metrics relevant to air-sea heat fluxes during Hurricane Dorian (2019). *Front. Earth Sci.,* **12,** 1342390, doi:10.3389/feart.2024.1342390.

Hazelton, A., G. J. Alaka, L. Gramer, W. Ramstrom, S. Ditchek, X. Chen, B. Liu, Z. Zhang, L. Zhu, W. Wang, B. Thomas, J. Shin, C.-K. Wang, H.-S. Kim, X. Zhang, A. Mehra, F. Marks, and S. Gopalakrishnan, 2023: 2022 real-time Hurricane forecasts from an experimental version of the Hurricane analysis and forecast system (HAFSV0.3S). *Front. Earth Sci.,* **11,** 1264969, doi:10.3389/feart.2023.1264969.

Alaka Jr., G,. J. A. Sippel, Z. Zhang, H.-S. Kim, F. Marks, V. Tallapragada, A. Mehra, X. Zhang, A. Poyer, and S. G. Gopalakrishnan, 2024: Lifetime performance of the operational Hurricane Weather Research and Forecasting model (HWRF) for North Atlantic tropical cyclones. *Bull. Amer. Meteor. Soc.,* **105,** doi:10.1175/BAMS-D-23-0139.1.

Wang, W., L. Zhu, B. Liu, Z. Zhang, A. Mehra, V. Tallapragada, 2023: A Forecast Cycle–Based Evaluation for Tropical Cyclone Rapid Intensification Forecasts by the Operational HWRF Model. *Wea. Forecasting*, **38,** 125-138, doi:10.1175/WAF-D-22-0007.1.

Wang, W., Z. Zhang, J. P. Cangialosi, M. Brennan, L. Cowan, P. Clegg, H. Takuya, I. Masaaki, A. K. Das, M. Mohapatra, M. Sharma, J. A. Knaff, J. Kaplan, T. Birchard, J. D. Doyle, J. Heming, J. Moskaitis, W. A. Komaromi, S. Ma, C. Sampson, L. Wu, and E. Blake, 2023: A review of recent advances (2018-2021) on tropical cyclone intensity change from operational perspectives, Part 2: Forecasts by Operational Centers. *Tropical Cyclone Research and Review*, **12(1)**, 50-63. doi:10.1016/j.tcrr.2023.05.003.

Zhang, Z., W. Wang, J. D. Doyle, J. Moskaitis, W. A. Komaromi, J. Heming, L. Magnusson, J. P. Cangialosi, L. Cowan, M. Brennan, S. Ma, A. K. Das, H. Takuya, P. Clegg, T. Birchard, J. A. Knaff, J. Kaplan, M. Mohapatra, M. Sharma, I. Masaaki, L. Wu, and E. Blake, 2023: A review of recent advances (2018-2021) on tropical cyclone intensity change from operational perspectives, Part 1: Dynamical model guidance, *Tropical Cyclone Research and Review*, **12(1)**, 30-49. doi:10.1016/j.tcrr.2023.05.004.

Wang, W., J. Han, F. Yang, J. Steffen, B. Liu, Z. Zhang, A. Mehra, and V. Tallapragada, 2023: Improving the Intensity Forecast of Tropical Cyclones in the Hurricane Analysis and Forecast System. Wea. Forecasting, 38, 2057–2075, doi:10.1175/WAF-D-23-0041.1.

Publications in Print

Xie, L. and B. Liu, 2024: Marine meteorology. *Encyclopedia of Atmospheric Sciences, 3nd Edition,* Elsevier Ltd., accepted.

Peng, J., Z. Zhang, W. Wang, R. Panda, B. Liu, Y. Weng, A. Mehra, V. Tallapragada, X. Zhang, S. Gopalakrishnan, W. Komaromi, and A. Poyer, 2024: HAFS Ensemble Forecast in AWS. *Front. Earth Sci.,* accepted.

Publications in Review

Chen, J.-H., T. Marchok, M. Bender, K. Gao, S. Gopalakrishnan, L. Harris, A. Hazelton, B. Liu, A. Mehra, M. Morin, F. Yang, X. Zhang, Z. Zhang, and L. Zhou, 2024: Closing the gap - advances in US models' hurricane predictions. *Bull. Amer. Meteor. Soc.,* in review.

Technical Reports:

Wang, W., Y. Weng, C. Wang, J. Peng, B. Liu, and Z. Zhang, 2024, An idealized hurricane analysis and forecast system (I-HAFS), WMO blue Book, Section 9.

Worthen, D., J. Wang, R. Montuoro, D. Heinzeller, B. Li, G. Theurich, U. Turuncoglu, D. Rosen, D. Jovic, B. Curtis, R. Mahajan, H. Lei, A. Richert, A. Chawla, J. Wang, J. Meixner, A. Abdolali, M. Masarik, L. Pan, M. Barlage, B. Liu, M. Vertenstein, T. Craig, R. Benson, T. Robinson, T. Clune, W. Jiang, N. Barton, G. Vandenberghe, M. Potts, J. Kim, N. Perlin, C. Book, L. Bernardet, F. Yang, S. Sun, H.-S. Kim, B. Baker, J. Huang, C.-H. Jeon, and I. Stajner, 2024: Coupling infrastructure capability in UFS weather model. *Office note (National Centers for Environmental Prediction (U.S.)) 519,* 115pp, doi:10.25923/dvv2-3g03.

6. References

Adkins, J., 2022. GeoXO benefit analysis. https://doi.org/10.25923/7tqj-r641.

Considine, T. J., C. Jablonowski, B. Posner, and C. H. Bishop, 2004: The Value of Hurricane Forecasts to Oil and Gas Producers in the Gulf of Mexico. J. Appl. Meteor. Climatol., 43, 1270–1281, https://doi.org/10.1175/1520-0450(2004)043%3C1270:TVOHFT%3E2.0.CO;2.

Kleist, D., J.R. Carley, A. Collard, E. Liu, S. Liu, C.R. Martin, C. Thomas, R. Treadon, G. Vernieres, 2024: Data Assimilation Strategy and Development Plan for NCEP's Environmental Modeling Center. Series: Office note (National Centers for Environmental Prediction (U.S.)) ; 517. https://doi.org/10.25923/kqra-6z66.

Martinez, A.B., 2020: Forecast Accuracy Matters for Hurricane Damage. Econometrics, 8, https://doi.org/10.3390/econometrics8020018.

Molina, R., D. Letson, B. McNoldy, P. Mozumder, and M. Varkony, 2021: Striving for Improvement: The Perceived Value of Improving Hurricane Forecast Accuracy. Bull. Amer. Meteor. Soc., 102, E1408–E1423, https://doi.org/10.1175/BAMS-D-20-0179.1.

Molina, R. and I. Rudik, 2024: The Social Value of Hurricane Forecasts. National Bureau of Economic Research, 23548, https://doi.org/10.3386/w32548.

National Weather Service 2023-2033 Strategic Plan. Available online at https://www.weather.gov/media/wrn/NWS-2023-Strategic-Plan.pdf.

NOAA Office of Coastal Management, 2024: Hurricane Costs. Available online at https://coast.noaa.gov/states/fast-facts/economics-and-demographics.html.

Peng, J., Z. Zhang, W. Wang, R. Panda, B. Liu, Y. Weng, A. Mehra, V. Tallapragada, X. Zhang, S. Gopalakrishnan, W. Komaromi, J. Anderson, and A. Poyer, 2024: HAFS Ensemble Forecast in AWS Cloud. Frontiers in Earth Science, 12, https://doi.org/10.3389/feart.2024.1396612.

Poyer, A., W. Komaromi, S. Gopalakrishnan, and coauthors, 2023: 2021-2022 HFIP R&D Activities Summary: Recent Results and Operational Implementation. HFIP Technical Report: HFIP2023-1, https://doi.org/10.25923/exgj-1n68.

United Nations, 2023: Global Survey Report on Persons with Disabilities and Disasters. Online, https://www.undrr.org/media/90432.

Weinkle, J., C. Landsea, D. Collins, R. Musulin, R. Crompton, P. Klotzbach, and R. Pielke Jr., 2018: Normalized hurricane damage in the continental United States 1900–2017. Nat. Sustain., 1, 808–813, https://doi.org/10.1038/s41893-018-0165-2.

Zachry, B., W. Booth, J. Rhome, and T. Sharon, 2015: A National View of Storm Surge Risk and Inundation. Wea. Climate Soc., 7, 109–117, https://doi.org/10.1175/WCAS-D-14-00049.1.

Appendix A: Table of Acronyms

4DEnVar	4-Dimensional Ensemble Variance-based data assimilation
AFS	Analyze, Forecast and Support office
AI	Artificial Intelligence
AMV	Atmospheric Motion Vector
AOML	Atlantic Oceanographic and Meteorological Laboratory

AOR	Area of Responsibility
AWIPS	Advanced Weather Interactive Processing System
AWS	Amazon Web Services
C3	Community Convective Cloud
CIMAS	Cooperative Institute For Marine And Atmospheric Studies
CIRA	Cooperative Institute for Research in the Atmosphere
CLP5	5-day Climatology and Persistence Track Forecast
COAMPS-TC	Coupled Ocean-Atmosphere Mesoscale Prediction System for Tropical Cyclones
CONUS	Continental United States
CPHC	Central Pacific Hurricane Center
DA	Data Assimilation
DESI	Dynamic Ensemble-based Scenarios for IDSS
DSHF	Decay SHIFOR Model Intensity Forecast
DSRC	David Skaggs Research Center
DTC	Developmental Testbed Center
ECMWF	European Centre for Medium-Range Weather Forecasts
EDMF	Eddy-Diffusivity/Mass-Flux
EMC	Environmental Modeling Center
EPAC	East Pacific
EPIC	Earth Prediction Innovation Center
ERC	Eyewall Replacement Cycle
FAR	False Alarm Ratio
FY	Fiscal Year
GEFS	Global Ensemble Forecast System
GFDL	Geophysical Fluid Dynamics Laboratory
GFS	Global Forecast System
GOES	Geostationary Operational Environmental Satellites
GSI	Gridpoint Statistical Interpolation
GSL	Global Systems Laboratory
HAFS	Hurricane Analysis and Forecast System
HEOB	HFIP Executive Oversight Board
HFIP	Hurricane Forecast Improvement Program
HMON	Hurricanes in a Multi-scale Ocean-coupled Non-hydrostatic model
HOT	Hurricane Ocean Testbed
HPC	High Performance Computing

HRD	Hurricane Research Division
HREx	HFIP Real-time Experiments
HWRF	Hurricane Weather Research and Forecast model
HYCOM	Hybrid Coordinate Ocean Model
IDSS	Impact-based Decision Support Services
IOC	Initial Operating Capability
JEDI	Joint Effort for Data assimilation Integration
JTWC	Joint Typhoon Warning Center
ML	Machine Learning
MOM6	Modular Ocean Model v6
MYNN	Mellor–Yamada–Nakanishi–Niino
NATL	North Atlantic
NCEP	National Centers for Environmental Prediction
NESDIS	National Environmental Satellite, Data, and Information Service
NHC	National Hurricane Center
NOAA	National Oceanic and Atmospheric Administration
NOS	National Ocean Service
NWS	National Weather Service
OAR	Oceanic and Atmospheric Research
OCD5	Operational CLP5 and DSHF Blended Intensity Forecast
OSTI	Office of Science and Technology Integration
PBL	Planetary Boundary Layer
PDF	Probability Density Function
POD	Probability of Detection
R&D	Research and Development
R2O	Research-to-Operations
R34	Radius of 34-kt wind
RDHPC	Research and Development High Performance Computing
RI	Rapid Intensification
RL	Readiness Level
RMW	Radius of Maximum Wind
SA-SAS	Scale-Aware Simplified Arakawa-Schubert
SAS	Simplified Arakawa-Schubert
SBES	Social Behavioral and Economic Science
SHiELD	System for High-resolution prediction on Earth-to-Local Domains
SHIFOR	Statistical Hurricane Intensity Forecast

SHUM	Stochastic Humidity perturbations
SKEB	Stochastic Kinetic Energy Backscatter
SPC	Storm Prediction Center
SPPT	Stochastically Perturbed Parametrization Tendencies
STAR	Center for Satellite Applications and Research
T-SHiELD	Tropical System for High-resolution prediction on Earth-to-Local Domains
TC	Tropical Cyclone
TCANE	Tropical Cyclone Artificial Neural Network Error
TKE	Turbulent Kinetic Energy
UFS	Unified Forecast System
UQ	Uncertainty Quantification
UTC	Coordinated Universal Time
VI	Vortex Initialization
WFO	Weather Forecast Office
WPAC	West Pacific
WPC	Weather Prediction Center
WSP	Wind Speed Probability
WTCM	Windspeed Tropical Cyclone Model
WW3	Wavewatch III