BAMS
Article

The G-IV Inner Circumnavigation

A Story of Successful Organic Interactions Between Research and Operations at NOAA

Jason A. Sippel, Sarah D. Ditchek, Kelly Ryan, and Christopher W. Landsea

KEYWORDS:
Tropical cyclones;
Aircraft observations;
Dropsondes;
Sensitivity studies;
Numerical weather prediction/forecasting;
Operational forecasting

ABSTRACT: This study describes the research-to-operations process leading to a recent change in tropical cyclone (TC) reconnaissance sampling patterns as well as observing-system experiments that evaluated the impact of that change on numerical weather prediction model forecasts of TCs. A valuable part of this effort was having close, multipronged connections between the TC research and operational TC prediction communities at the National Oceanic and Atmospheric Administration (NOAA). Related to this work, NOAA's Atlantic Oceanographic and Meteorological Laboratory (AOML) and National Hurricane Center (NHC) have a long history of close collaboration to improve TC reconnaissance. Similar connections between AOML and NOAA's Environmental Modeling Center (EMC) also laid a foundation for the observing-system experiments conducted here. More specifically, AOML and NHC collaborated in 2018 to change how NHC uses NOAA's Gulfstream-IV (G-IV) jet during TC synoptic surveillance missions. That change added a second circumnavigation at approximately 1.5° from TC centers, when possible. Preliminary experiments suggest that the change improved track forecasts, though the intensity results are more mixed. Despite the somewhat small sample size over a 3-yr period, the track improvement does agree with prior work. This effort has led to additional work to more fully examine G-IV sampling strategies.

SIGNIFICANCE STATEMENT: This study highlights a successful research-to-operations implementation and evaluation, facilitated by having close, multipronged connections between the tropical cyclone (TC) research and operational TC prediction communities at the National Oceanic and Atmospheric Administration (NOAA). In particular, in 2018 the National Hurricane Center (NHC) added a second circumnavigation at a radius of about 1.5° from TC centers to NOAA's Gulfstream-IV (G-IV) jet reconnaissance pattern. This change was motivated by growing evidence from scientists at the Atlantic Oceanographic and Meteorological Laboratory (AOML) that focusing G-IV surveillance closer to TCs would improve track forecasts. We present a preliminary assessment of the impact of that change and confirm in practice the findings of previous theoretical and tangentially related studies that led to the change—adding dropsondes near TCs tends to improve the track forecasts.

https://doi.org/10.1175/BAMS-D-23-0084.1

Corresponding author: Jason A. Sippel, jason.sippel@noaa.gov

In final form 17 November 2023

AFFILIATIONS: Sippel—NOAA/AOML/Hurricane Research Division, Miami, Florida; **Ditchek and Ryan**—NOAA/AOML/Hurricane Research Division, and Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, Florida; **Landsea**—NOAA/NWS National Hurricane Center, Miami, Florida

Research-to-operations (R2O) transitions can take on many different forms and have been crucial for improving weather forecasts at the National Oceanic and Atmospheric Administration (NOAA). Many transitions take place through formal entities, such as the Developmental Testbed Center, which was established as an R2O pathway for numerical weather prediction (Bernardet et al. 2008). Other transitions take less formal pathways, such as through annual NOAA testbed experiments (e.g., Gallo et al. 2017), and depend heavily on interpersonal interactions.

R2O transitions have also played a vital role in improving tropical cyclone (TC) forecasts at NOAA. Formal mechanisms such as NOAA's Hurricane Forecast Improvement Program (HFIP; Gall et al. 2013), Joint Technology Transfer Initiative (JTTI; Kondragunta et al. 2022), and Joint Hurricane Testbed (JHT; Rappaport et al. 2009, 2012) have advanced R2O efforts in modeling, observational analysis, postprocessing, and even observing-system strategies. Predating all of these mechanisms, internal R2O efforts at NOAA's Atlantic Oceanographic and Meteorological Laboratory (AOML) also contributed to advances, particularly in using airborne reconnaissance to improve TC forecasts.

R2O and TC reconnaissance

The Hurricane Research Division (HRD) of AOML led the earliest efforts in using reconnaissance data to improve TC forecasts. In particular, HRD organized a series of experiments with NOAA's WP-3D (hereafter P-3) aircraft to optimize a data collection strategy in the 1980s and early 1990s (Burpee et al. 1996; Franklin and DeMaria 1992). Burpee et al. (1996) found that assimilating dropsonde observations collected from those missions improved track forecasts by up to 30%. Given the large impact that dropsondes had on TC forecasts, NOAA procured a Gulfstream-IV (hereafter G-IV) jet aircraft in 1996 to conduct operational synoptic surveillance missions.

Some of the very first JHT projects expanded upon the success of Burpee et al. (1996) by improving how the G-IV was used. Aberson (2003) described part of this effort, which established that sampling only certain sensitive regions benefited TC track forecasts. One result endemic to all related studies was the high value of sampling around a TC (e.g., Aberson and Etherton 2006). This research ultimately led NHC to use a G-IV pattern that both sampled the environment and encircled TCs with a single ring of dropsondes (e.g., Fig. 1a). Aberson (2010) subsequently reviewed the impact of dropsondes during periods of G-IV sampling from 1997 to 2006. For forecasts initialized with G-IV dropsonde data, dropsondes had improved TC track predictions in the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) by about 10% through 2–3 days.

Since then, formal R2O mechanisms at NOAA have continued to improve both airborne data usage and observing-system strategies. Around 2010, HFIP funded a major initiative to expand the amount of reconnaissance data assimilated by operational models. This led to the first real-time assimilation of airborne Doppler data into NOAA's Hurricane Weather Research

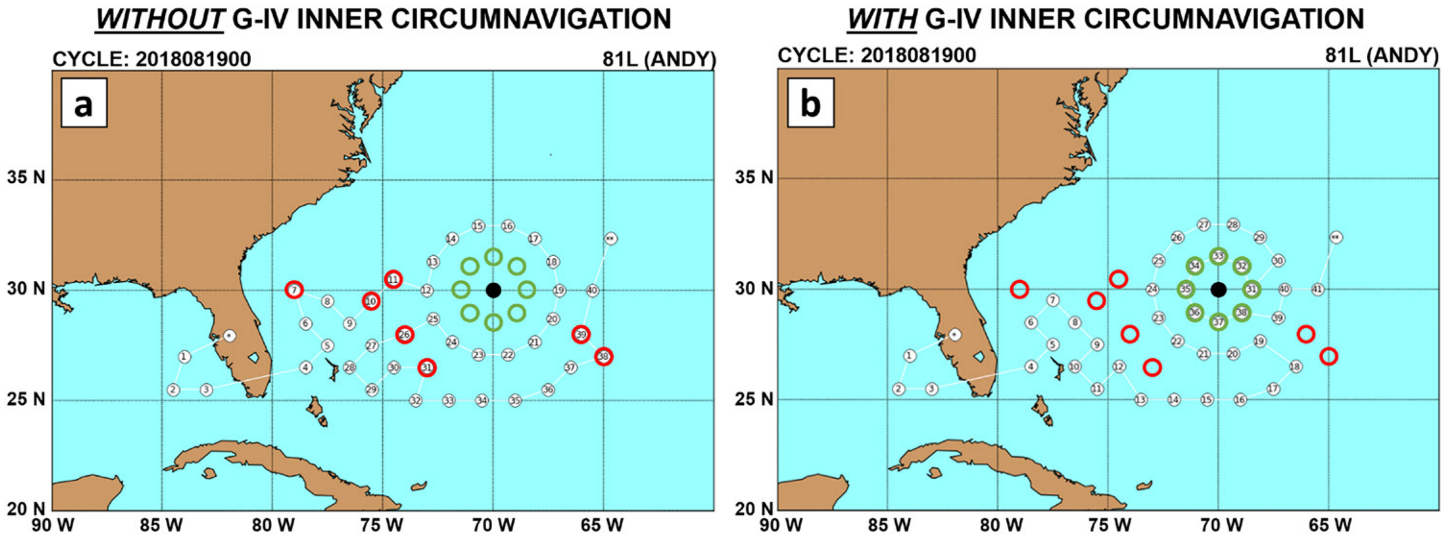


Fig. 1. Prototype G-IV flight-track pattern and dropsonde locations (a) without an inner circumnavigation and (b) with an inner circumnavigation. Note that both patterns include environmental targeting following Aberson (2003). Dropsondes that were removed (red) and added (green) given the flight-track change are also indicated.

and Forecasting (HWRF) model in 2013 as well as other airborne datasets in later years. JTTI subsequently supported work that expanded the use of inner-core dropsondes in HWRF (Winterbottom et al. 2018). As a result of these and other efforts, HWRF began assimilating all operationally transmitted reconnaissance data, which improved its intensity forecasts by 10%–15% on average through 2–3 days (Zawislak et al. 2022). Further, JHT has supported ongoing efforts to improve G-IV sampling strategies (Torn 2022), which benefits all NCEP models.

Over the last few years, HRD and NHC have closely collaborated to use observing-system changes to improve TC forecasts. Like the work of the 1980s and 1990s, this collaboration has occurred organically, outside the context of more formal R2O mechanisms. One example of such a change was end-point dropsondes added during United States Air Force (USAF) reserve 53rd Weather Reconnaissance Squadron C-130 transects (around 150–200 km from the TC center). A second change was releasing midpoint dropsondes at a radius of around 80 km from the TC center during P-3 missions. Ditchek et al. (2023a, hereafter D23A) more thoroughly documented both of these modifications. Another organic interaction between scientists at HRD and NHC led to some recent changes to G-IV sampling strategies. These G-IV changes provide excellent examples of successful R2O implementations and evaluations and provide a focus for this paper.

An overview of G-IV usage

NHC generally tasks NOAA's G-IV aircraft when a potential hurricane threatens the United States or its territories. These missions, which are designed by forecasters at NHC and sent to the NOAA/Office of Marine and Aviation Operations (OMAO) Aircraft Operations Center (AOC) typically a day in advance, have the following requirements specified in the National Hurricane Operations Plan (NHOP; NOAA 2022):

> 5.3.4. Synoptic Surveillance Data Requirements.
>
> When required, NHC will request sounding data on the periphery of systems approaching populated areas. CPHC may request sounding data on the periphery of those that may impact the Hawaiian Islands. For all synoptic-surveillance tasking requirements, NHC will be responsible for providing specific tracks including dropsonde locations pertaining to each designated synoptic time to CARCAH for coordination with the reconnaissance units.
>
> …

> 5.7.4. Synoptic Surveillance Missions.
>
> A synoptic surveillance mission is tasked to measure the large-scale wind and thermodynamic fields within approximately 800 nautical miles of tropical cyclones. Specific flight tracks will vary depending on storm location and synoptic situation, and multiple aircraft may be required to satisfy surveillance mission requirements.

Figure 1 provides a general example of the change in G-IV sampling strategy described in this paper. As described above, each G-IV mission typically has a portion of the flight that focuses on near-TC sampling as well as a portion that focuses on the environment. Before 2018 (e.g., Fig. 1a), this two-part flight track included dropsonde locations every 1°–2° designated by 1) a single circumnavigation at a radius of about 3° from the TC center and 2) environmental targets as described by Aberson (2003). Meanwhile, Fig. 1b demonstrates the changes implemented operationally in 2018. Specifically, NHC modified the pattern to include a second circumnavigation at a radius of about 1.5°, which comes at the expense of some environmental sampling. Hereafter, this second circumnavigation will be called the “inner” circumnavigation while the circumnavigation at 3° will be called the “outer” circumnavigation. Note that while the details of environmental targeting have changed (Torn 2022), this study focuses specifically on the G-IV inner circumnavigations.

The remainder of this work describes the R2O process that led to the addition of the G-IV inner circumnavigation as well as subsequent evaluations of the impact of the change on TC forecasts. Briefly, observation system simulation experiments (hereafter OSSEs; i.e., data-denial experiments with simulated data; Zeng et al. 2020) conducted at HRD from 2016 to 2018 suggested that adding a second G-IV circumnavigation could improve TC track forecasts. These results, together with prior relevant research, led NHC to modify G-IV flight strategies starting in 2018. We present a preliminary evaluation of the impact of those changes for the 2018–20 hurricane seasons in an experimental version of the HWRF model.

Support for the G-IV inner circumnavigation

Ryan et al. (2018, 2019, hereafter R18 and R19) introduced the OSSE framework relevant to this particular R2O effort. They ran a total of 10 OSSEs to evaluate the sensitivity of results to various plausible G-IV flight patterns. For each OSSE, simulated conventional and satellite observations from a well-documented hurricane nature run [i.e., a simulation considered to be the “ground truth”; here HNR1 from Nolan et al. (2013)] were assimilated into a system that was based on the 2013 version of HWRF. For the various plausible G-IV flight patterns, additional dropsonde observations were simulated. Note that HNR1 is a singular TC case, spanning the life cycle of a typical recurving Atlantic hurricane.

One common result among the OSSEs in R18 and R19 was that multiple concentric rings of dropsondes around the TC center resulted in the best forecasts. An example can be seen in Fig. 2, which shows track errors presented in Fig. 4c of R19 (i.e., CONTROL, SINGLE1, SINGLE2) in addition to results from another experiment described in R18 (i.e., DOUBLE). CONTROL did not assimilate any observations related to the G-IV, SINGLE1 and SINGLE2 each had one G-IV circumnavigation at different radii, and DOUBLE included two G-IV circumnavigations. The Fig. 2 result and others (not shown) suggested that a change in G-IV sampling to add a second circumnavigation might improve TC track forecasts.

Previous research available at the time (e.g., Fig. 3) suggested that sampling within or close to the vortex also improved TC track forecasts. For example, Harnisch and Weissman (2010) found that dropsondes in remote regions (i.e., 700–1,200 km from the TC center) had less of an impact on track forecasts than did dropsondes in the vicinity of TCs (Fig. 3a).

Further, observing-system experiments (i.e., data-denial experiments with real data) conducted at NCEP in 2017 also suggested a benefit of near-TC sampling from high-altitude dropsondes (Sippel et al. 2017; Wick et al. 2020). In particular, assimilating Global Hawk dropsonde data did not benefit GFS track forecasts for 2012–14 TCs, but it substantially benefited forecasts of 2016 TCs. One possible explanation for the greater benefit in 2016 was that NOAA used the Global Hawk in TC-focused survey patterns that year (Wick et al. 2020; Fig. 3b, bottom), whereas the flights from 2012 to 2014 focused generally on large-scale surveillance (Braun et al. 2016; Fig. 3b, top).

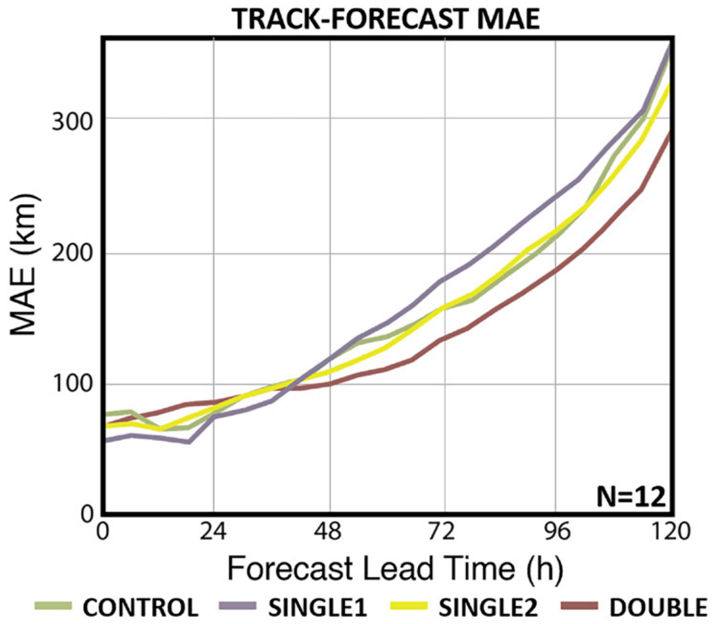
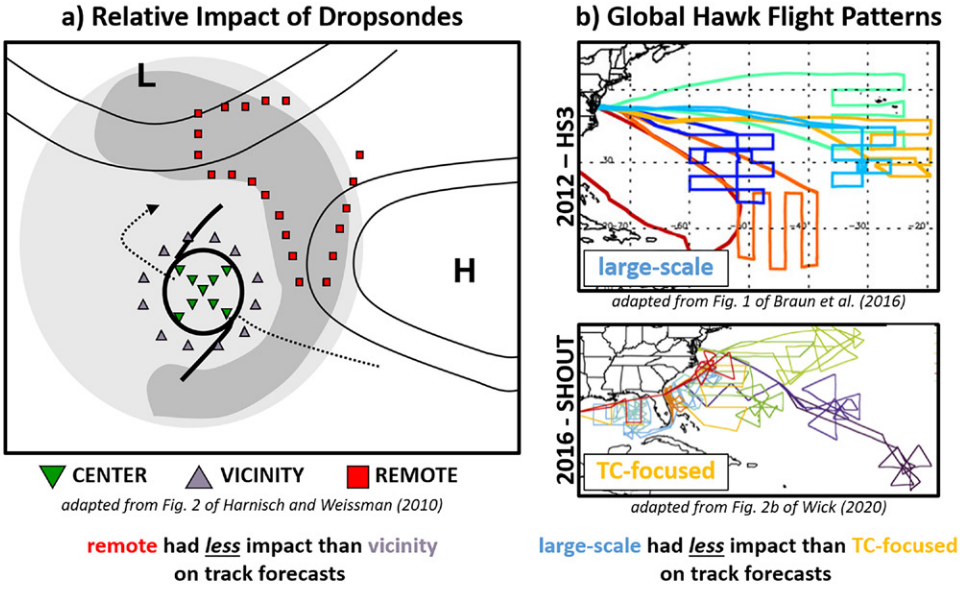


Fig. 2. The track-forecast mean absolute error (MAE) for CONTROL (green), SINGLE1 (purple), SINGLE2 (yellow), and DOUBLE (red). The sample size at 0 h is 12 forecasts. Note that this graphic is a modified version of Fig. 4c from the OSSEs included in R19 along with DOUBLE from R18.

Operational implementation and use

Conversations regarding adding an inner G-IV circumnavigation began during the 72nd Interdepartmental Hurricane Conference (IHC) in March 2018. Along with the potential benefits to TC track forecasts demonstrated in R18, R19, and Fig. 3, conversations


a) Relative Impact of Dropsondes

b) Global Hawk Flight Patterns

adapted from Fig. 2 of Harnisch and Weissman (2010)

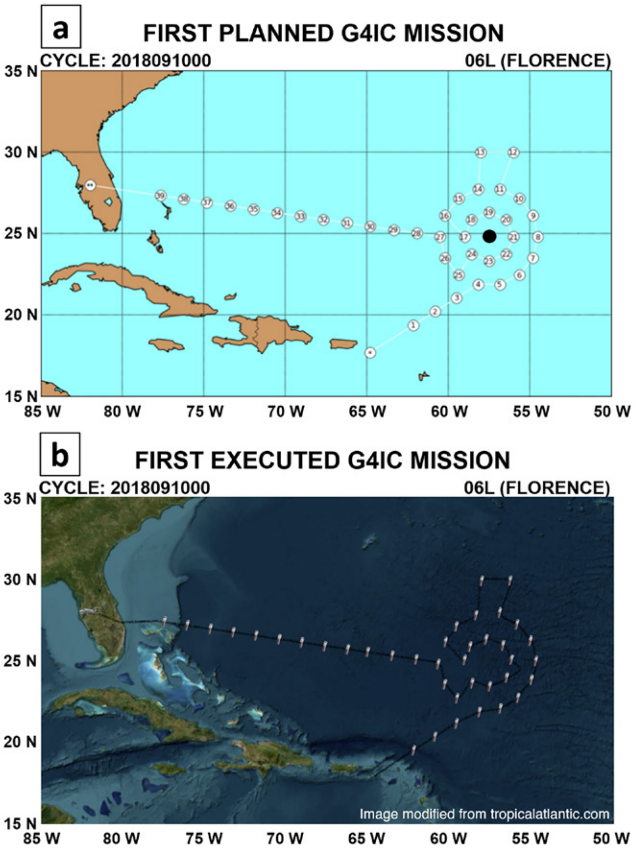
adapted from Fig. 1 of Braun et al. (2016)

adapted from Fig. 2b of Wick (2020)

remote had *less* impact than vicinity on track forecasts

large-scale had *less* impact than TC-focused on track forecasts

Fig. 3. Results from prior research that corroborated the Fig. 2 OSSE results. (a) The relative impact of dropsondes in the center, vicinity, and remote regions on TC track forecasts, adapted from Fig. 2 of Harnisch and Weissman (2010), and (b) Global Hawk flight patterns (top) during Hurricane and Severe Storm Sentinel (HS3) in 2012 and (bottom) during Sensing Hazards with Operational Unmanned Technology (SHOUT) in 2016, adapted from Fig. 1 of Braun et al. (2016) and Fig. 2b of Wick et al. (2020), respectively.



c

LIST OF G-IV MISSIONS

Year	TC	# missions	# w/ G4IC
2018	Florence (06L)	8	5
	Michael (14L)	2	2
2019	Dorian (05L)	10	6
	Jerry (10L)	2	2
2020	Isaias (09L)	3	0
	Laura (13L)	6	2
	Marco (14L)	2	1
	Delta (26L)	4	4
	Zeta (28L)	4	2
	Eta (29L)	3	2
2021	Elsa (05L)	1	0
	Henri (08L)	4	4
	Ida (09L)	3	3
	Sam (18L)	2	2
2022	Fiona (07L)	3	2
	Ian (09L)	7	5
	Nicole (17L)	3	2
2023	Franklin (08L)	1	0
	Idalia (10L)	3	0
	Lee (13L)	10	8
TOTAL:		81	52

Fig. 4. A summary of previous double-circumnavigation flights including the first (a) planned and (b) executed G-IV inner-circumnavigation (G4IC) mission into Hurricane Florence as well as (c) a list of all TCs with G-IV missions to date since the flight in (a). Numbered circles in (a) indicate planned dropsonde locations, and the black circle indicates the projected center position of Hurricane Florence. Note that all planned dropsondes were released during this mission.

focused on the benefits of additional sampling of the TC vortex. For example, Sippel (2018) showed at the same IHC that assimilating data from C-130 end-point dropsondes had improved intensity forecasts in HWRF. There was also interest in collecting airborne Doppler radar data from the G-IV during an inner circumnavigation, as radar data are not typically available from the 3° circumnavigation due to the lack of precipitation (and thus radar scatterers) at such a large radius. Radar data collected during an inner circumnavigation could help with NHC's situational awareness and also be assimilated to improve HWRF forecasts.

Discussions between three of the authors (Sippel, Ryan, and Landsea) regarding adding an inner G-IV circumnavigation continued after the IHC. These discussions focused on organizing logistics for and obtaining permissions to include the addition. The main logistical hurdle was revising the flight-track software so that the inner circumnavigation could be readily included for each mission. These revisions were provided by an NHC contractor (Andrew Penny). Other logistics involved determining safety considerations for both the G-IV crew as well as other aircraft flying in the TC environment. Many of these considerations are detailed below. The other aspect needed before adding an inner G-IV circumnavigation was buy-in from all relevant and interested parties [i.e., NHC, HRD, AOC, NOAA's Environmental Modeling Center (EMC), and the Air Force 53rd Weather Reconnaissance Squadron]. To that end, the same three authors provided briefings to officials at each office with background information on and evidence for including an inner G-IV circumnavigation.

Approval for systematically flying the G-IV double circumnavigation was obtained in the 2018 hurricane season. There are several items to note with this change, which fortunately did not require any modification to the NHOP:

- Adding the inner circumnavigation of eight additional dropsondes necessitated that some dropsondes in the periphery be removed. (Typically G-IV missions are all flown near maximum endurance, so an increase in sampling in one part of the mission requires a decrease elsewhere.)
- As the number of additional dropsondes in the added circumnavigation is nearly exactly matched by the number of dropsondes removed elsewhere in the flight track, there has been no additional cost for dropsondes, fuel, or personnel time.
- The 1.5° inner circumnavigation is closer to the TC center than the outer circumnavigation, but it is also generally outside of the eyewall and not over the TC center.
- The inner circumnavigation has been flown clockwise to avoid flying into deep convection associated with a rainband (a clockwise pattern could potentially lead the G-IV into the moat between a rainband and the eyewall).
- Flying closer to the TC center requires a higher altitude to avoid hazards (e.g., deep convection). Considering that the G-IV reaches its maximum cruising altitude toward the end of a flight, the inner circumnavigation has been flown as late in the mission as possible in order to maximize altitude.
- The addition of the 1.5° inner circumnavigation has led to more frequent revisions to the originally provided flight track due in part to errors in the forecasted TC center location.
- As with other G-IV missions, small modifications of the flight track can occur for safety (e.g., to avoid going through a towering cumulus).
- The G-IV flight crew has to coordinate more closely with low-level (P-3 and C-130) aircraft crews when dropping dropsondes in the 1.5° inner circumnavigation, as it is often spatially collocated with where those low-level aircraft typically fly.

The G-IV double circumnavigation was first used midway through the 2018 hurricane season with a flight into Hurricane Florence (Figs. 4a,b), and the approach has been routinely used since then. Among the 81 tasked G-IV missions to date (i.e., November 2023) since the initial Florence double-circumnavigation flight, 52 (64%) had inner circumnavigations. Ones that did not typically were due to flight duration limitations as well as the TCs being too close to land. Figure 4c summarizes the various tasked G-IV missions since late 2018 and indicates which TCs had the double circumnavigation flown.

Figure 5 provides an example of how meteorological data transmitted from a recent G-IV double circumnavigation complements the suite of existing reconnaissance data. The flight, which targeted Hurricane Ian (2022; while it was still a tropical storm), occurred simultaneously with missions from a P-3 and a C-130. The P-3 and C-130 repeatedly sampled low levels near the TC center, while the high-altitude G-IV sampled both the outer vortex and TC environment (for clarity, the G-IV track is shaded). Figure 5a shows the G-IV mission as planned, while Figs. 5b and 5c show the reconnaissance data[1] assimilated into the relevant cycles of the operational HWRF model. Data from the outer circumnavigation were assimilated in the first cycle (0600 UTC; Fig. 5b), while data from part of the outer circumnavigation and the entire inner circumnavigation were assimilated in the second cycle (1200 UTC; Fig. 5c). Note that the inner-circumnavigation dropsondes symmetrically sampled the vortex during the

---

[1] This includes high-density observations (HDOBs) from flight level (i.e., temperature, specific humidity, and wind) and from the stepped frequency microwave radiometer (SFMR), radial velocity from the tail Doppler radar (TDR), and pressure, temperature, relative humidity, and horizontal winds from global positioning system (GPS) dropsondes.

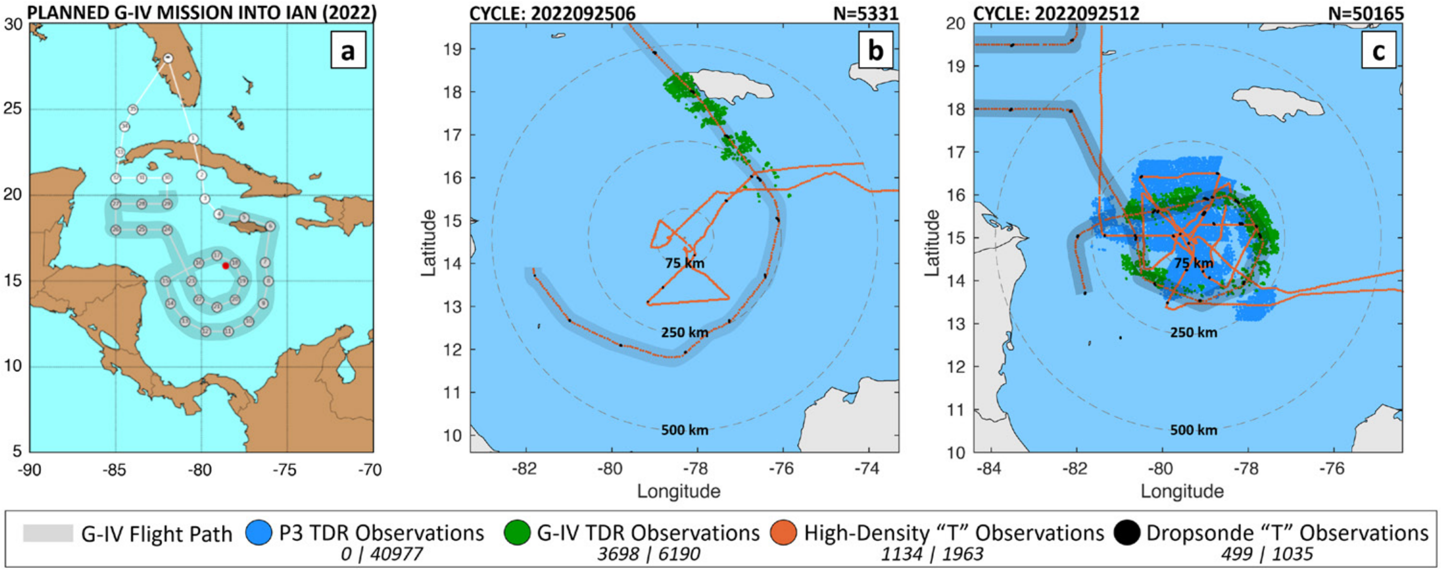


Fig. 5. **A representative example of (a) a G-IV mission with an inner circumnavigation and (b),(c) the corresponding assimilated reconnaissance observations in HWRF. The shaded portion of the planned flight track in (a) corresponds with the actual G-IV flight track shaded in (b) and (c). Note that (b) and (c) include reconnaissance from aircraft other than the G-IV. The number of assimilated reconnaissance observations overall is given in the top right of (b) and (c), and the number of select individual reconnaissance observation type is given in the legend.**

1200 UTC cycle with full-tropospheric profiles of wind, temperature, and humidity that otherwise would not have been available.

A preliminary evaluation of NWP impact

Before discussing this evaluation, we also wish to highlight the value of close collaboration between HRD and EMC, as it directly laid the groundwork for the assessment undertaken below. Much of the past interaction between HRD and EMC was funded by HFIP and focused on evaluation and development of the HWRF model. Some of those improvements related to DA are described above, and Alaka et al. (2023, manuscript submitted to *Bull. Amer. Meteor. Soc.*) provides a full discussion of the R2O efforts that improved HWRF. In addition to model improvements, HRD and EMC have occasionally worked closely together to evaluate observing systems relevant to TC prediction. Such efforts include some of the G-IV studies cited earlier as well as a more recent evaluation of the impact of Global Hawk dropsondes released during the Sensing Hazards with Operational Unmanned Technology (SHOUT) experiment (Wick et al. 2020) discussed above.

Specifically related to this work, HRD and EMC worked together to ensure an appropriate response to the Bipartisan Budget Act of 2018, which included supplemental appropriations for TC prediction (hereafter the FY18 Hurricane Supplemental). Given the considerable number of observing-system changes with dropsondes that were taking place at that time, HRD and EMC both agreed that a large systematic evaluation of their use was warranted. As a result, the FY18 Hurricane Supplemental funded both the procurement of additional dropsondes and the subsequent evaluation of how best to use them. To that end, D23A as well as Ditchek and Sippel (2023) performed general assessments of dropsonde usage, while the current study focuses specifically on the aforementioned G-IV changes.

What follows is a preliminary assessment of the impact of data from the G-IV inner circumnavigation on TC forecasts. This component of the FY18 Hurricane Supplemental effort used a "basin-scale" version of the 2020 operational HWRF (Zhang et al. 2016; Alaka et al. 2017, 2019, 2020, 2022) to compare an experiment that allowed all dropsonde data (hereafter ALL) to another that denied inner-circumnavigation dropsonde data (hereafter NOG4IC).

Table 1. Years run, forecast periods (FP), start and end dates (in YYYYMMDDHH format), number of cycles, TCs with G-IV inner circumnavigations (G4ICs), and number of forecasts (FCSTS) with G4ICs. A summary row is included at the bottom of the table. Note that two of the 106 forecasts were unverifiable (the last cycle in Michael as it was designated as "extratropical" and the last cycle in Marco as it was designated as a "low").

Year	FP	Start date	End date	Cycles	TCs with G4ICs	FCSTS with G4ICs
2018	1	2018091000	2018091600	25	06L (Florence)	5
	2	2018100900	2018101200	13	14L (Michael)	3
2019	3	2019082700	2019090512	39	05L (Dorian)	6
2020	4	2020082306	2020082800	20	13L (Laura)	2
		2020082306	2020082506	9	14L (Marco)	1
Summary						
3 years	4 FPs		106 cycles		5 TCs with a G4IC	17 FCSTS with a G4IC

Note that the ALL experiment here is identical to ALL in D23A. Experiments covered those TCs from 2018 to 2020 that had inner circumnavigations during the active North Atlantic basin (NATL; including the North Atlantic Ocean, the Gulf of Mexico, and the Caribbean Sea) periods defined in D23A. Experiments encompassed four forecast periods as outlined in Table 1. Each period began with the first inner circumnavigation in the TC of interest and ended eight cycles (i.e., 2 days) after the last inner circumnavigation. Note that the only difference between ALL and NOG4IC comes from the G-IV inner-circumnavigation dropsonde data assimilated.

Figure 6 explores changes in the spatial distribution of dropsonde data between the two experiments. Dropsonde temperature observations[2] assimilated in each TC's D02 are shown in two different frameworks. Overall, removing dropsondes from the G-IV inner circumnavigations reduced the number of individually assimilated dropsonde temperature observations by 11%. That reduction can be seen in Fig. 6 between 100 and 250 km, straddling the 165-km radius (i.e., around 1.5° from the TC center). For a detailed description of the distributions in Figs. 6a and 6b as they relate to reconnaissance sampling in general, please see section 2d in D23A.

Verification. As in D23A, the performance of each experiment was evaluated by verifying forecasts against the NHC "best track" (Landsea and Franklin 2013) available from NHC. Standard NHC forecast verification procedures (Cangialosi 2022) were applied to raw output from the vortex tracker (Marchok 2002, 2021), known as the "late model" in operational parlance. Note that we present verification of only track and two measures of TC intensity [maximum sustained 10-m wind speed (VMAX) and minimum sea level pressure (PMIN)], since eliminating inner-circumnavigation dropsonde observations did not meaningfully alter any of the operationally predicted significant wind radii on average (not shown). Further, the consistency metric introduced in Ditchek et al. (2023b, hereafter D23B) is used to identify lead times of consistent improvement or degradation in mean absolute error (MAE) skill, median absolute error (MDAE) skill, and frequency of superior performance (FSP; Velden and Goldenberg 1987; Goldenberg et al. 2015). Finally, results presented here are for homogeneous samples.

For each variable, this study stratifies forecast errors in two ways. First, we evaluate the net impact of the G-IV inner circumnavigation from the five TCs where at least one inner circumnavigation was flown (i.e., Florence, Michael, Dorian, Laura, and Marco). Again, this sample contains forecast errors from the time of the first inner circumnavigation until two days after the last one. Then, we isolate the immediate impact of the inner circumnavigation by examining forecast errors for just those forecasts initialized with dropsonde data from that pattern.

---

[2] Only temperature is shown for simplicity, though temperature, humidity, and winds were all assimilated.

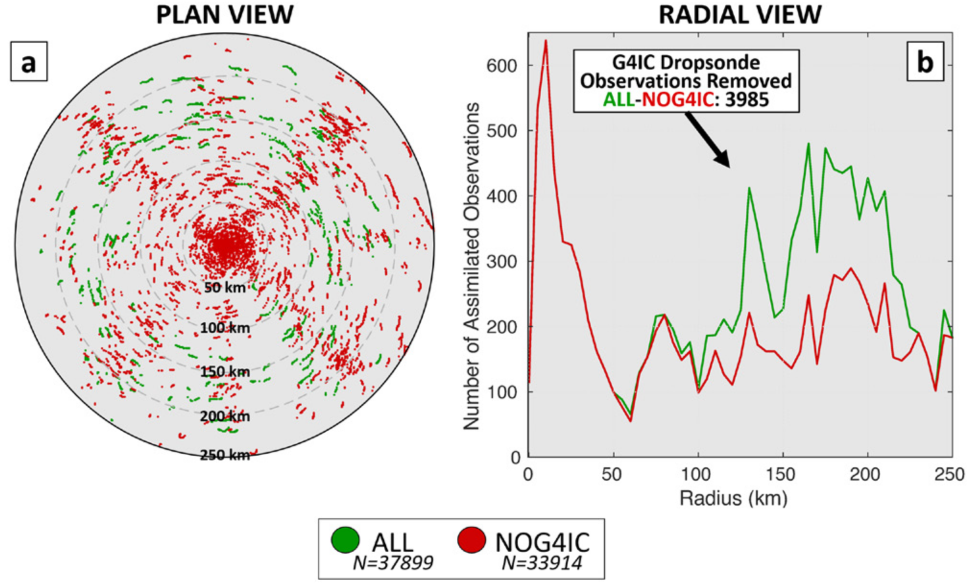


Fig. 6. **The number of individually assimilated dropsonde temperature observations in each TC's D02 from 0 to 250 km in (a) TC-relative plan view and (b) in TC-relative radial view for both ALL (green) and NOG4IC (red). Note that NOG4IC is plotted on top of ALL. Thus, any green markers seen indicate those dropsonde observations that were part of the G-IV inner circumnavigation (G4IC) that were present in ALL but denied in NOG4IC. Note that ALL had 37,899 individually assimilated dropsonde temperature observations while NOG4IC had 3,985 fewer. This difference is due to the individually assimilated G-IV inner circumnavigation dropsonde temperature observations that were removed in NOG4IC.**

Composite results. Perhaps the most important result from Fig. 7 is that the G-IV inner-circumnavigation dropsonde data consistently improved track forecasts. More specifically, Fig. 7a shows that track forecasts in TCs with inner circumnavigations improved with at least marginal consistency for all but three lead times between 24 and 96 h. This improvement averaged almost 4%, reached up to 6%, and ranged between 2 and 9 km. When considering only those forecasts that had dropsonde data from inner circumnavigations (Fig. 7d), the window of track-forecast improvement shifted to slightly longer lead times compared to those found in Fig. 7a. More specifically, the track forecast improvement occurred at all but two lead times between 48 and 108 h. This improvement averaged almost 3%, reached up to 6%, and ranged between 3 and 12 km (Fig. 7d). The similar improvement in both samples suggests that the inner-circumnavigation dropsonde data immediately improved analyses when it was assimilated (i.e., Fig. 7d) and improved the first guess for DA during subsequent cycles.

On the other hand, the impact of inner-circumnavigation dropsondes on intensity forecasts was more nuanced. First, Figs. 7b and 7c shows that on average, intensity forecasts did not improve in TCs with inner-circumnavigation dropsondes. Turning to only the cycles that assimilated inner-circumnavigation dropsonde data, results were a bit different. In particular, the added data improved MAE skill for both VMAX (Fig. 7e) and PMIN (Fig. 7f) by up to 20%. Despite the elevated MAE skill in ALL, little consistency occurred due to similar MDAE and FSP between the experiments (not shown). This suggests that the improvement in ALL relative to NOG4IC in Figs. 7e and 7f resulted from a few particularly bad intensity forecasts in NOG4IC. Further analysis revealed a few forecasts from Hurricane Dorian (2019) drove the improvement found (not shown).

Discussion

This study both describes the R2O process that led to some recent changes in how NOAA uses the G-IV jet for TC surveillance missions and then presents a preliminary evaluation of the

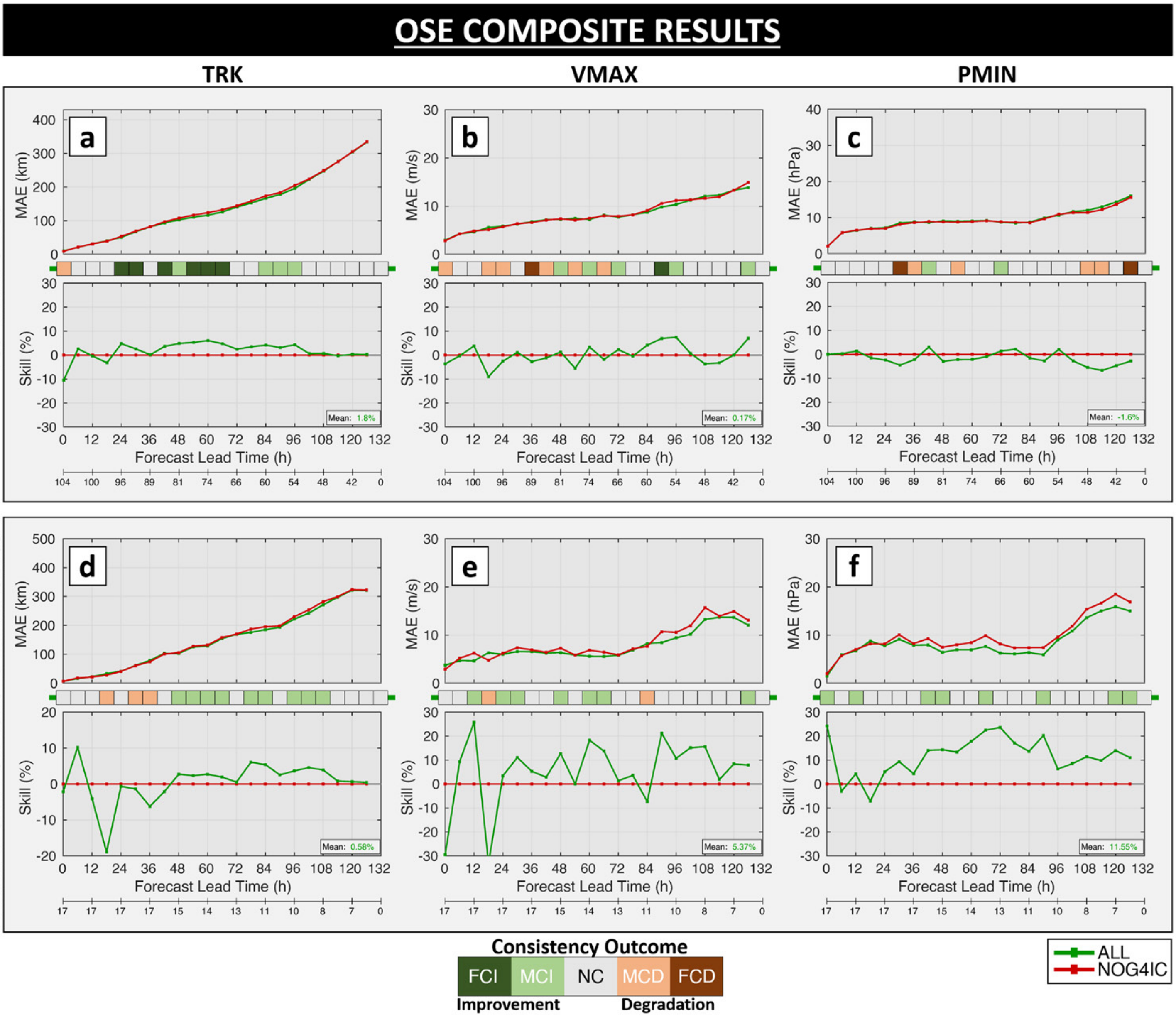


Fig. 7. The MAE and MAE skill for NATL TCs for the ALL (green) and NOG4IC (red) experiments for (a)–(c) those TCs that had at least one G-IV inner circumnavigation (G4IC) and (d)–(f) for just those individual forecasts initialized with G-IV inner-circumnavigation dropsonde data (G4ICs) for (a),(d) track (TRK), (b),(e) VMAX, (c),(f) PMIN. Shaded boxes between the MAE and MAE skill panels indicate forecast lead times where results were fully consistent, marginally consistent, or not consistent, based on the consistency-metric criteria found in D23B. The sample size is given below the *x* axis in each panel.

impact of that change. Past R2O efforts for improving TC prediction have come through both formally funded projects as well as through organic interactions between NHC forecasters and other scientists, particularly at NOAA's AOML. This particular effort fits into the later category and began with a series of exchanges at the 2018 Interdepartmental Hurricane Conference (IHC), which themselves stemmed from ongoing research at AOML. More specifically, emerging evidence at the time suggested that focusing G-IV surveillance closer to TCs would improve track forecasts. In response, NHC changed their G-IV tasking plans to include two complete circumnavigations around a TC, when possible.

Subsequent observing-system experiments provided a preliminary assessment of the impact of inner circumnavigation dropsonde data with an experimental version of NOAA's HWRF model. While preliminary, findings represent the first published assessment of the impact of inner-circumnavigation dropsondes. Results from a series of relevant TCs from the 2018–20 hurricane seasons suggest that the additional data have improved track forecasts, though the impact on intensity forecasts was less consistent. Yet, intensity errors in a few particularly bad forecasts were substantially reduced during periods of inner-circumnavigation sampling.

The present study most closely relates to D23A, as both projects were part of the same FY18 Hurricane Supplemental effort described earlier. D23A evaluated the impacts of all dropsondes over a very large sample of 2017–20 TCs and found improvements in forecasts of TC track, intensity, and outer wind radii. Notably, the impact on intensity in D23A was inconsistent, whereas the impact on track and outer wind radii was more consistent. Thus, the finding here that inner-circumnavigation dropsondes predominantly improved track forecasts aligns with D23A, but the lack of impact on outer radii does not. A possible reason for the differences between results here and D23A is the limited sample size available for this work since the inner G-IV circumnavigation occurs relatively infrequently compared to the entirety of dropsonde sampling. Note that other studies (e.g., Zawislak et al. 2022; Sippel 2018) have found that adding reconnaissance data typically benefits intensity forecasts, but those results are hard to compare due to differences present (e.g., model version, data types, and samples).

Broadly speaking, the infrequency of G-IV inner circumnavigations and the concomitant small sample size achieved here is the main limitation of this study. One of the key findings of D23A was that a large sample was needed due to strong variability in impacts between TCs and even over entire seasons. Achieving such a large sample with the G-IV inner circumnavigation pattern will take many years at the current rate the pattern is used. Thus, this study can only provide a preliminary assessment of the change.

Additional research funded by NOAA’s Hurricane and Ocean Testbed (HOT; formerly JHT) is more fully investigating the impact of G-IV sampling strategies on TC forecasts during the 2020–23 seasons (Ditchek 2023). The current study has only assessed the impact of *dropsondes* from the G-IV inner circumnavigation on TC forecasts, but as seen in Fig. 5, other reconnaissance data (e.g., HDOBs and TDR) are collected during inner G-IV circumnavigations. Thus, efforts are currently underway to quantify the impact of *all* reconnaissance data from the inner circumnavigation on TC forecasts. That impact will then be directly compared to the impact of environmental targeting (i.e., Torn 2022) and the impact of the G-IV overall. In doing so, guidance will be provided on potential optimizations for the current G-IV flight-track strategy.

A valuable part of this effort was having close, multipronged connections between TC research and operational TC prediction. As described earlier, NOAA’s HRD and NHC have a long history of close collaboration. A major contributor to this success has been that a number of current and past employees at NHC, including the fourth author, began their careers at HRD. Further contributing to this success is that the first author jointly leads the NHC-based HOT along with the Science and Operations Officer at NHC. Thus, while the conversations regarding the suggested pattern change started during the 2018 IHC, the close personal connections between HRD and NHC kept the conversation going. Similar connections between HRD and EMC also laid a foundation for this effort. For example, the first author participated in the SHOUT experiment first as an employee of EMC and then HRD. As described above, the evaluation of the impact of SHOUT dropsondes motivated conversation regarding how operational reconnaissance sampling might be changed. Finally, these same connections helped ensure that the FY18 Hurricane Supplemental funded both the procurement of additional dropsondes and the subsequent evaluation (i.e., D23A; Ditchek and Sippel 2023) of how best to use them. Thus, this study is a culmination of two parallel research and operational engagements, one between NHC and HRD to change the G-IV flight track and another one between HRD and EMC to fund the subsequent evaluation as part of the FY18 Hurricane Supplemental.

In summary, the exploratory work that led to organic interactions at the IHC in March 2018 represented only the initial stages of a renewed effort to optimize high-altitude sampling of TCs. The operational change that this paper focuses on was part of a less formal process, whereas the subsequent evaluation presented above as well as the in-progress HOT project are important formal iterations within the overall effort. This pattern of optimizing and improving observing systems to improve TC forecasts will likely be important for the foreseeable future as NOAA expands investments into TC reconnaissance. In particular, NOAA will replace the now-aging G-IV jet with two Gulfstream G550s in the near future. This advancement poses new opportunities to explore innovative observing strategies as models and DA systems evolve.

Acknowledgments. Both the OSSE and OSE components of this research were carried out in part under the auspices of the Cooperative Institute for Marine and Atmospheric Studies (CIMAS), a Cooperative Institute of the University of Miami and the National Oceanic and Atmospheric Administration, Cooperative Agreement NA20OAR4320472. The OSSE-component funding came from the Quantitative Observing System Assessment Program (QOSAP) and supported the third author (Kelly Ryan). The OSE-component funding came from the FY18 Hurricane Supplemental (NOAA Award NA19OAR0220188) and supported the second author (Sarah Ditchek). Further, thank you to Sim Aberson, Eric Blake, Michael Brennan, and Frank Marks, and three *BAMS* reviewers (Sharan Majumdar, Jonathan Zawislak, and one anonymous) for their constructive comments on the manuscript. The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect those of OAR or the Department of Commerce.

Data availability statement. OSE experiments were performed on the NOAA RDHPCS supercomputers Hera and Orion, with output archived on NCEI's High Performance Storage System (HPSS) for a 5-yr term. This output is not publicly available, however, those interested in the output can contact the second author (Sarah Ditchek). The final B-decks (i.e., best tracks) used for verification are available from NHC and can be found at https://www.nhc.noaa.gov/data/hurdat. Dropsonde data can be found on HRD's Hurricane Field Program's public-facing website at https://www.aoml.noaa.gov/data-products/hurricanedata. Finally, the GRaphics for OS(S)Es and Other modeling applications on TCs (GROOT) verification package developed by the second author (Sarah Ditchek) and funded by the Quantitative Observing System Assessment Program (QOSAP) and the FY18 Hurricane Supplemental (NOAA Award NA19OAR0220188) was used to generate some of the graphics in this publication. It can be found at https://github.com/sditchek/GROOT.

References

Aberson, S. D., 2003: Targeted observations to improve operational tropical cyclone track forecast guidance. *Mon. Wea. Rev.*, **131**, 1613–1628, https://doi.org/10.1175//2550.1.

——, 2010: 10 years of hurricane synoptic surveillance (1997–2006). *Mon. Wea. Rev.*, **138**, 1536–1549, https://doi.org/10.1175/2009MWR3090.1.

——, and B. J. Etherton, 2006: Targeting and DA studies during Hurricane Humberto (2001). *J. Atmos. Sci.*, **63**, 175–186, https://doi.org/10.1175/JAS3594.1.

Alaka, G. J., X. Zhang, S. G. Gopalakrishnan, S. B. Goldenberg, and F. D. Marks, 2017: Performance of basin-scale HWRF tropical cyclone track forecasts. *Wea. Forecasting*, **32**, 1253–1271, https://doi.org/10.1175/WAF-D-16-0150.1.

——, ——, ——, Z. Zhang, F. D. Marks, and R. Atlas, 2019: Track uncertainty in high-resolution HWRF ensemble forecasts of Hurricane Joaquin. *Wea. Forecasting*, **34**, 1889–1908, https://doi.org/10.1175/WAF-D-19-0028.1.

——, D. Sheinin, B. Thomas, L. Gramer, Z. Zhang, B. Liu, H.-S. Kim, and A. Mehra, 2020: A hydrodynamical atmosphere/ocean coupled modeling system for multiple tropical cyclones. *Atmosphere*, **11**, 869, https://doi.org/10.3390/atmos11080869.

——, X. Zhang, and S. G. Gopalakrishnan, 2022: High-definition hurricanes: Improving forecasts with storm-following nests. *Bull. Amer. Meteor. Soc.*, **103**, E680–E703, https://doi.org/10.1175/BAMS-D-20-0134.1.

Bernardet, L., and Coauthors, 2008: The Developmental Testbed Center and its Winter Forecasting Experiment. *Bull. Amer. Meteor. Soc.*, **89**, 611–628, https://doi.org/10.1175/BAMS-89-5-611.

Braun, S. A., P. A. Newman, and G. M. Heymsfield, 2016: NASA's Hurricane and Severe Storm Sentinel (HS3) investigation. *Bull. Amer. Meteor. Soc.*, **97**, 2085–2102, https://doi.org/10.1175/BAMS-D-15-00186.1.

Burpee, R. W., J. L. Franklin, S. J. Lord, R. E. Tuleya, and S. D. Aberson, 1996: The impact of Omega dropwindsondes on operational hurricane track forecast models. *Bull. Amer. Meteor. Soc.*, **77**, 925–934, https://doi.org/10.1175/1520-0477(1996)077<0925:TIOODO>2.0.CO;2.

Cangialosi, J. P., 2022: National hurricane center forecast verification report: 2021 hurricane season. NOAA/NWS, 76 pp., https://www.nhc.noaa.gov/verification/pdfs/Verification_2021.pdf.

Ditchek, S. D., 2023: The overall and individual impact of two recent G-IV flight-track changes on TC forecasts using HAFS. *HFIP Annual Meeting*, Miami, FL, NOAA, 6B.2D, https://hfip.org/events/2023/annual-review-meetings/20231114.

——, and J. A. Sippel, 2023: A comparison of the impacts of inner-core, in-vortex, and environmental dropsondes on tropical cyclone forecasts during the 2017–20 hurricane seasons. *Wea. Forecasting*, **38**, 2169–2187, https://doi.org/10.1175/WAF-D-23-0055.1.

——, ——, G. J. Alaka Jr., S. B. Goldenberg, and L. Cucurull, 2023a: A systematic assessment of the overall dropsonde impact during the 2017–2020 hurricane seasons using the basin-scale HWRF. *Wea. Forecasting*, **38**, 789–816, https://doi.org/10.1175/WAF-D-22-0102.1.

——, ——, P. J. Marinescu, and G. J. Alaka Jr., 2023b: Improving best-track verification of tropical cyclones: A new metric to identify forecast consistency. *Wea. Forecasting*, **38**, 817–831, https://doi.org/10.1175/WAF-D-22-0168.1.

Franklin, J. L., and M. DeMaria, 1992: The impact of omega dropwindsonde observations on barotropic hurricane track forecasts. *Mon. Wea. Rev.*, **120**, 381–391, https://doi.org/10.1175/1520-0493(1992)120<0381:TIOODO>2.0.CO;2.

Gall, R., J. Franklin, F. Marks, E. N. Rappaport, and F. Toepfer, 2013: The hurricane forecast improvement project. *Bull. Amer. Meteor. Soc.*, **94**, 329–343, https://doi.org/10.1175/BAMS-D-12-00071.1.

Gallo, B. T., and Coauthors, 2017: Breaking new ground in severe weather prediction: The 2015 NOAA/Hazardous Weather Testbed Spring Forecasting Experiment. *Wea. Forecasting*, **32**, 1541–1568, https://doi.org/10.1175/WAF-D-16-0178.1.

Goldenberg, S. B., S. G. Gopalakrishnan, V. Tallapragada, T. Quirino, F. Marks, S. Trahan, X. Zhang, and R. Atlas, 2015: The 2012 triply-nested, high-resolution operational version of the Hurricane Weather Research and Forecasting System (HWRF): Track and intensity forecast verifications. *Wea. Forecasting*, **30**, 710–729, https://doi.org/10.1175/WAF-D-14-00098.1.

Harnisch, F., and M. Weissmann, 2010: Sensitivity of typhoon forecasts to different subsets of targeted dropsonde observations. *Mon. Wea. Rev.*, **138**, 2664–2680, https://doi.org/10.1175/2010MWR3309.1.

Kondragunta, C., and Coauthors, 2022: Joint technology transfer initiative: A research-to-operations transition program in NOAA. *12th Conf. on Transition of Research to Operations*, Houston, TX, Amer. Meteor. Soc., 5B.1, https://ams.confex.com/ams/102ANNUAL/meetingapp.cgi/Paper/393218.

Landsea, C. W., and J. L. Franklin, 2013: Atlantic hurricane database uncertainty and presentation of a new database format. *Mon. Wea. Rev.*, **141**, 3576–3592, https://doi.org/10.1175/MWR-D-12-00254.1.

Marchok, T. P., 2002: How the NCEP tropical cyclone tracker works. *25th Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., P1.13, http://ams.confex.com/ams/pdfpapers/37628.pdf.

——, 2021: Important factors in the tracking of tropical cyclones in operational models. *J. Appl. Meteor. Climatol.*, **60**, 1265–1284, https://doi.org/10.1175/JAMC-D-20-0175.1.

NOAA, 2022: National hurricane operations plan. Office of the Federal Coordinator for Meteorological Services and Supporting Research (OFCM) Doc. FCM-P12-2022, 30 pp., https://repository.library.noaa.gov/view/noaa/47707.

Nolan, D. S., R. Atlas, K. T. Bhatia, and L. R. Bucci, 2013: Development and validation of a hurricane nature run using the joint OSSE nature run and the WRF Model. *J. Adv. Model. Earth Syst.*, **5**, 382–405, https://doi.org/10.1002/jame.20031.

Rappaport, E. N., and Coauthors, 2009: Advances and challenges at the National Hurricane Center. *Wea. Forecasting*, **24**, 395–419, https://doi.org/10.1175/2008WAF2222128.1.

——, J. Jiing, C. W. Landsea, S. T. Murillo, and J. L. Franklin, 2012: The Joint Hurricane Testbed: Its first decade of tropical cyclone research-to-operations activities reviewed. *Bull. Amer. Meteor. Soc.*, **93**, 371–380, https://doi.org/10.1175/BAMS-D-11-00037.1.

Ryan, K., L. Bucci, J. Delgado, R. Atlas, and S. Murillo, 2018: Sensitivity of G-IV dropsonde configuration on tropical cyclone prediction using a regional OSSE framework. *22nd Conf. on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface*, Austin, TX, Amer. Meteor. Soc., 612, https://ams.confex.com/ams/98Annual/webprogram/Paper336603.html.

——, ——, ——, ——, and ——, 2019: Impact of Gulfstream-IV dropsondes on tropical cyclone prediction in a regional OSSE system. *Mon. Wea. Rev.*, **147**, 2961–2977, https://doi.org/10.1175/MWR-D-18-0157.1.

Sippel, J., 2018: The value of extra dropsondes deployed by AFRES. *72nd Interdepartmental Hurricane Conf*., Miami, FL, Interagency Council on Advancing Meteorological Services, 11 pp., https://www.icams-portal.gov/meetings/TCORF/ihc18/session_8/8-8-sippel.pdf.

——, K. Howard, and V. Tallapragada, 2017: The impact of Global Hawk dropsondes on NCEP hurricane forecasts. *18th Cyclone Workshop*, Sainte Adele, QC, Canada, State University of New York at Albany, 1 p., https://www.atmos.albany.edu/facstaff/kristen/CW18/Monday/Sippel.pdf.

Torn, R. D., 2022: Ensemble-based sensitivity guidance for reducing TC track forecast uncertainty. *35th Conf. on Hurricanes and Tropical Meteorology*, New Orleans, LA, Amer. Meteor. Soc., 6B.4, https://ams.confex.com/ams/35Hurricanes/meetingapp.cgi/Paper/401886.

Velden, C., and S. Goldenberg, 1987: The inclusion of high density satellite wind information in a barotropic hurricane-track forecast model. Preprints, *17th Conf. on Hurricanes and Tropical Meteorology*, Miami, FL, Amer. Meteor. Soc., 90–93.

Wick, G. A., and Coauthors, 2020: NOAA's Sensing Hazards with Operational Unmanned Technology (SHOUT) experiment observations and forecast impacts. *Bull. Amer. Meteor. Soc.*, **101**, E968–E987, https://doi.org/10.1175/BAMS-D-18-0257.1.

Winterbottom, H., J. A. Sippel, A. Mehra, B. Zhang, and Z. Zhang, 2018: Recent and planned upgrades to operational HWRF data assimilation. *33rd Conf. on Hurricanes and Tropical Meteorology*, Ponte Vedra, FL, Amer. Meteor. Soc., 3C.5, https://ams.confex.com/ams/33HURRICANE/webprogram/Paper338953.html.

Zawislak, J., and Coauthors, 2022: Accomplishments of NOAA's airborne hurricane field program and a broader future approach to forecast improvement. *Bull. Amer. Meteor. Soc.*, **103**, E311–E338, https://doi.org/10.1175/BAMS-D-20-0174.1.

Zeng, X., and Coauthors, 2020: Use of observing system simulation experiments in the United States. *Bull. Amer. Meteor. Soc.*, **101**, E1427–E1438, https://doi.org/10.1175/BAMS-D-19-0155.1.

Zhang, X., S. G. Gopalakrishnan, S. Trahan, T. S. Quirino, Q. Liu, Z. Zhang, G. Alaka, and V. Tallapragada, 2016: Representing multiple scales in the hurricane weather research and forecasting modeling system: Design of multiple sets of movable multilevel nesting and the basin-scale HWRF forecast application. *Wea. Forecasting*, **31**, 2019–2034, https://doi.org/10.1175/WAF-D-16-0087.1.