2019 Atlantic Hurricane Forecasts from the Global-Nested Hurricane Analysis and Forecast System: Composite Statistics and Key Events

ANDREW HAZELTON,[a,b] ZHAN ZHANG,[c,d] BIN LIU,[c,d] JILI DONG,[c,d] GHASSAN ALAKA,[b] WEIGUO WANG,[c,d] TIM MARCHOK,[e] AVICHAL MEHRA,[c] SUNDARARAMAN GOPALAKRISHNAN,[b] XUEJIN ZHANG,[b] MORRIS BENDER,[f] VIJAY TALLAPRAGADA,[c] AND FRANK MARKS[b]

[a] *Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, Florida*
[b] *NOAA/AOML/Hurricane Research Division, Miami, Florida*
[c] *NOAA/Environmental Modeling Center, College Park, Maryland*
[d] *IMSG, College Park, Maryland*
[e] *NOAA/GFDL, Princeton, New Jersey*
[f] *Princeton University, Princeton, New Jersey*

(Manuscript received 16 March 2020, in final form 19 January 2021)

ABSTRACT: NOAA's Hurricane Analysis and Forecast System (HAFS) is an evolving FV3-based hurricane modeling system that is expected to replace the operational hurricane models at the National Weather Service. Supported by the Hurricane Forecast Improvement Program (HFIP), global-nested and regional versions of HAFS were run in real time in 2019 to create the first baseline for the HAFS advancement. In this study, forecasts from the global-nested configuration of HAFS (HAFS-globalnest) are evaluated and compared with other operational and experimental models. The forecasts by HAFS-globalnest covered the period from July through October during the 2019 hurricane season. Tropical cyclone (TC) track, intensity, and structure forecast verifications are examined. HAFS-globalnest showed track skill superior to several operational hurricane models and comparable intensity and structure skill, although the skill in predicting rapid intensification was slightly inferior to the operational model skill. HAFS-globalnest correctly predicted that Hurricane Dorian would slow and turn north in the Bahamas and also correctly predicted structural features in other TCs such as a sting jet in Hurricane Humberto during extratropical transition. Humberto was also a case where HAFS-globalnest had better track forecasts than a regional version of HAFS (HAFS-SAR) due to a better representation of the large-scale flow. These examples and others are examined through comparisons with airborne tail Doppler radar from the NOAA WP-3D to provide a more detailed evaluation of TC structure prediction. The results from this real-time experiment motivate several future model improvements, and highlight the promise of HAFS-globalnest for improved TC prediction.

KEYWORDS: Hurricanes/typhoons; Tropical cyclones; Numerical weather prediction/forecasting

1. Introduction

The Hurricane Analysis and Forecast System (HAFS), a component of the Unified Forecast System (NOAA 2020) is a novel hurricane modeling system being developed collaboratively within the National Oceanic and Atmospheric Administration (NOAA) and with academic partners, under the guidance of the Hurricane Forecast Improvement Program (HFIP, http://www.hfip.org). It has a fully compressible nonhydrostatic dynamical core coupled with multiple physics suites. The current dynamic core and nesting capability with the finite volume solver (shortened to FV3 hereafter) on the cubed sphere is summarized in a series of publications (e.g., Lin and Rood 1996; Lin 1997; Lin and Rood 1997; Putman and Lin 2007; Harris and Lin 2013).

A few recent studies have begun to demonstrate the utility of nested FV3-based modeling systems for TC prediction. Hazelton et al. (2018a) used a 2-km stretched and nested configuration of the Geophysical Fluid Dynamics Laboratory (GFDL) fvGFS, FV3 dynamic core with the physics from the National Centers for Environmental Prediction (NCEP) operational global model for high-resolution hindcasts of several tropical cyclones (TCs) that were also sampled by NOAA WP-3D (P-3) flights, and compared the forecast model structure with the P-3 radar data. That study highlighted the capability for skillful prediction of TC structure with such a modeling system and demonstrated the utility of airborne observations for model evaluation. Building off of these findings, a configuration of fvGFS with a 3-km nested domain covering most of the North Atlantic was used to perform real-time forecasts during the 2017 Atlantic hurricane season. This included many high-impact TCs like Harvey, Irma, Maria, and Nate, all of which made landfall in the United States or its territories. Nested fvGFS had comparable track and intensity skill to operational models for this sample, and the 3-km nest also demonstrated improvement over a lower-resolution (13 km) globally uniform fvGFS. These results are summarized in Hazelton et al. (2018b). Further demonstrating the value of FV3-based TC forecasts, the globally uniform (no nested domain) configuration of fvGFS was tested with European Centre for Medium-Range Weather Forecasts (ECMWF) initial conditions, and showed improvement in TC track prediction over both the operational Global Forecast System

Corresponding author: Andrew Hazelton, andrew.hazelton@noaa.gov

DOI: 10.1175/WAF-D-20-0044.1

(GFS) and ECMWF models (Chen et al. 2019) for a set of TC tracks. Based on these promising early results of a global-nested hurricane prediction system using fvGFS, real-time forecasts were performed using a prototype version of HAFS during the 2019 Atlantic hurricane season.

One of the motivations for the model configuration described in the next section is the fact that the synoptic environment around a TC is critical to aspects of its evolution, and so predicting this environment (and the TC–environment interaction) correctly is critical. For example, Aberson (2010) showed how GFS forecasts were improved by synoptic surveillance flights using high-altitude aircraft and dropsondes in the Atlantic Basin. Aberson (2011) found similar improvements to TC track forecasts in both the Atlantic and west Pacific when assimilating data from dropsondes into the operational Global Forecast System (GFS). Other studies have demonstrated the importance of remote tropical cyclones to modifying the environment and affecting the track of other TCs. For example, Alaka et al. (2017) showed that the Basin-Scale Hurricane Weather Research and Forecasting (HWRF) Model, which has a much larger outer domain than the operational HWRF, improved the forecasts of TC track relative to operational HWRF when two or more other TCs were located far away in the same domain. Finally, tropical cyclones can also have significant impact on global circulation patterns, making proper representation of the TC–environment interaction of paramount importance. As an example, Archambault et al. (2013) showed how recurving typhoons in the Western North Pacific significantly amplified the extratropical flow in the area, and that this amplification propagated downstream. It is clear from these studies that having an accurate representation of the large-scale environment, and representing the details of TC interaction through a large static nest, are important for prediction.

The model configuration described next provides a framework for analysis and prediction that represents small-scale processes as well as the synoptic processes discussed above. The rest of the paper is organized as follows: section 2 describes the model setup, section 3 presents results including composite statistics as well as exploration of a few case studies, section 4 summarizes the findings, and section 5 discusses opportunities for future work.

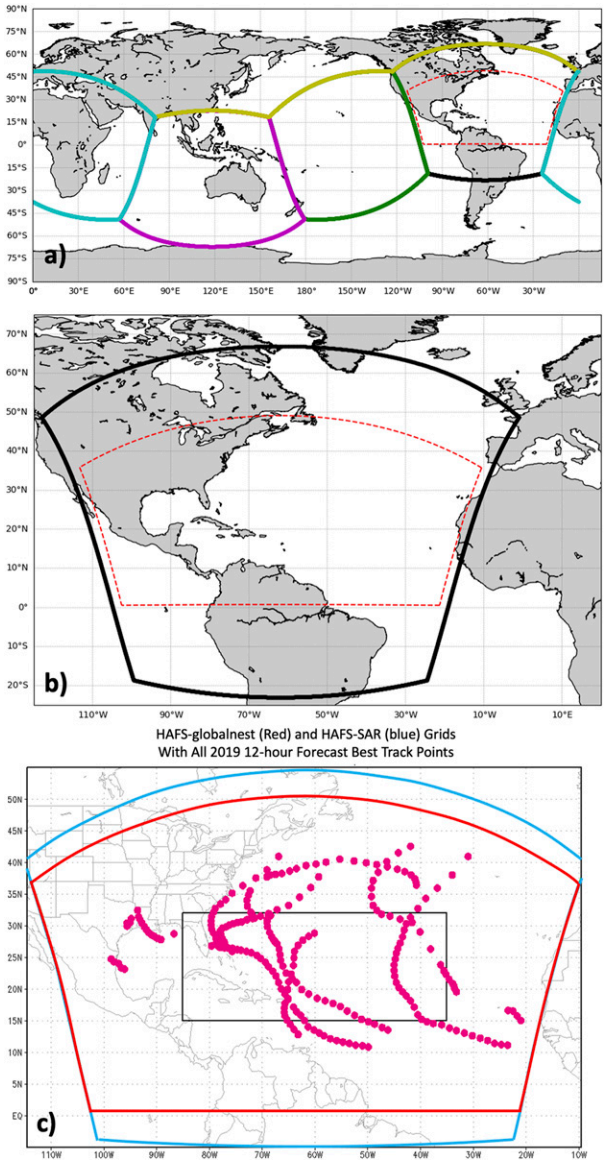

FIG. 1. (a) The global layout of the six tiles forming the cubed sphere domain of HAFS-globalnest. The 3-km static nest is shown in the red dashed lines. (b) A zoomed-in view of the Atlantic tile and 3-km nest. (c) The Atlantic domain of HAFS-globalnest (red), HAFS-SAR (blue), with the ''box'' used for later analysis of errors near and away from the domain edge shown in black. The pink dots show 12-h HAFS-globalnest forecasts.

2. Model configuration and data

a. *Model configuration*

The HAFS configuration described in this study is based on global-nested FV3 demonstrated in Harris and Lin (2013). This configuration enables two-way feedback between the global and static nested domains, with the global and nested domains running concurrently This allows for the analysis of the impact of two-way feedback on TC and environmental prediction. It is called HAFS-globalnest for the rest of the paper. It is different from the stand-alone-regional configuration of HAFS (called HAFS-SAR, Dong et al. 2020) which features a static regional nest that gets initial and boundary conditions (one-way feedback) from the operational GFS. HAFS-globalnest is a bit more computationally expensive than HAFS-SAR (121 core-hours per forecast hour versus 85 core-hours per forecast hour based on last year's test), as expected because HAFS-globalnest runs a global and nested forecast concurrently. For this season, HAFS-globalnest was initialized directly by the 13-km operational GFS analysis, i.e., a ''cold start.'' Development of a high-quality data assimilation system is an ongoing effort. HAFS-globalnest uses 64 vertical levels on both the global and nested domains.

TABLE 1. Configurations of HAFS-globalnest and the other operational and experimental models it is compared with.

Model	Domain	Boundary conditions	Finest grid spacing	PBL physics	Microphysics	Dynamic ocean (Y/N)
GFS	Global	—	13 km	EDMF	GFDL	N
HAFS-globalnest	Global with Atlantic static nest	—	3 km	Modified EDMF	GFDL	N
HAFS-SAR	Atlantic static nest	GFS	3 km	Modified EDMF	GFDL	N
HWRF	Storm-following nest	GFS	1.5 km	Modified EDMF	Ferrier–Aligo	Y
HMON	Storm-following nest	GFS	2 km	Modified EDMF	Ferrier–Aligo	Y

In this version of HAFS-globalnest, the global six tiles (at ~12–13-km horizontal grid spacing) making up the cubed-sphere grid are arranged such that one of six tiles is centered over the North Atlantic. Inside of this tile is a 3-km horizontal resolution nest, as shown in Fig. 1. This nest size covers the entire North Atlantic from western Africa to North America, although it is shifted slightly west compared to the nest used in Hazelton et al. (2018b) for the 2017 forecasts. The domain of HAFS-SAR is also shown (Fig. 1c), along with a "box" at the center of the HAFS-globalnest domain that is used for analysis of errors near and away from the domain edges in a later section.

b. Physics

HAFS-globalnest uses many of the same model physics schemes that are used in the operational GFS. The convective parameterization is the scale-aware scheme used in both the operational GFS and HWRF Models (Han et al. 2017). This scheme was turned on for the global domain, but was turned off for the nest to let the model explicitly resolve convection. This is a difference from the GFDL fvGFS configuration described in Hazelton et al. (2018b), and was based on earlier tests that showed turning the convective scheme off reduced track error, as well as a desire to have a baseline for later versions of HAFS that will have higher horizontal resolution where a cumulus parameterization is even less necessary. HAFS-globalnest uses the GFS hybrid eddy-diffusivity mass-flux (EDMF) planetary boundary layer (PBL) scheme (Han et al. 2016), with height-dependent modifications to the eddy diffusivity as described in Wang et al. (2018). The radiation scheme is the Rapid Radiative Transfer Model for global climate models (GCMs) (RRTMG; Iacono et al. 2008). For microphysics, HAFS-globalnest uses the GFDL 6-class single-moment microphysics scheme (e.g., Chen and Lin 2013). The surface drag is also similar to that used in the HWRF, which is based off of the GFDL surface layer scheme (Bender et al. 2007). There is no ocean coupling in this version of HAFS, although that capability is in development. For the cases shown here, HAFS-SAR also used the same physics configuration as the nested domain of HAFS-globalnest. Table 1 summarizes the model configuration (including physics) of HAFS-globalnest and the other operational and experimental models examined in this study.

c. Forecast period

HAFS-globalnest was run four times daily (0000, 0600, 1200, and 1800 UTC) from 5 July 2019 to 1 November 2019. The forecasts were run out to 168 h, to analyze the skill of the model both in the traditional 5-day forecast window as well as an extended 7-day window (in contrast, HAFS-SAR was only run out to 126 h). TCs were tracked using the Marchok tracker (Marchok 2002). Note that current operational regional hurricane models integrate only to 5 days, whereas operational global models typically integrate beyond 7 days. As seen below, not all TC forecasts actually extend out to 168 h, because there were many short-lived TCs during the 2019 season.

d. Observational data

TC track, maximum wind, and wind radii from model forecasts are verified against the working "best track" data from the National Hurricane Center (Landsea and Franklin 2013). The verifications shown below only include times when the TC was tracked in the model and was also classified as a TC (at least a tropical depression) in the best track data. To explore more complete structural evaluations and take advantage of the wealth of observational data collected in several of the 2019 TCs, the model data are also compared in several case studies with tail Doppler radar data from the NOAA P-3 flights. This radar data are collected in the TC inner core (e.g., within 150 km of the TC center), and the wind data are derived using the method of Gamache et al. (2004). The merged analysis, calculated by averaging the radar data from a flight over several flight legs (e.g., Rogers et al. 2013a, Reasor et al. 2013, DeHart et al. 2014), is used for the radar comparisons in this study. Similar data for past TCs was used in a quantitative and qualitative comparison of TC structure from high-resolution nested FV3 and P-3 radar data in Hazelton et al. (2018a).

3. Results

a. Verification statistics

Figure 2 shows the track forecast errors for HAFS-globalnest as well as HAFS-SAR and other GFS-based guidance [operational GFS, HWRF, and the Hurricanes in a Multiscale Ocean-coupled Nonhydrostatic (HMON) model]. While the global domain of HAFS-globalnest is very similar to the operational GFS in terms of model configuration, the nested domain is configured very differently from HWRF and HMON. The static HAFS nest is much larger than the moving nests of HWRF and HMON, but the horizontal resolution is coarser (~3 km in HAFS, 2 km

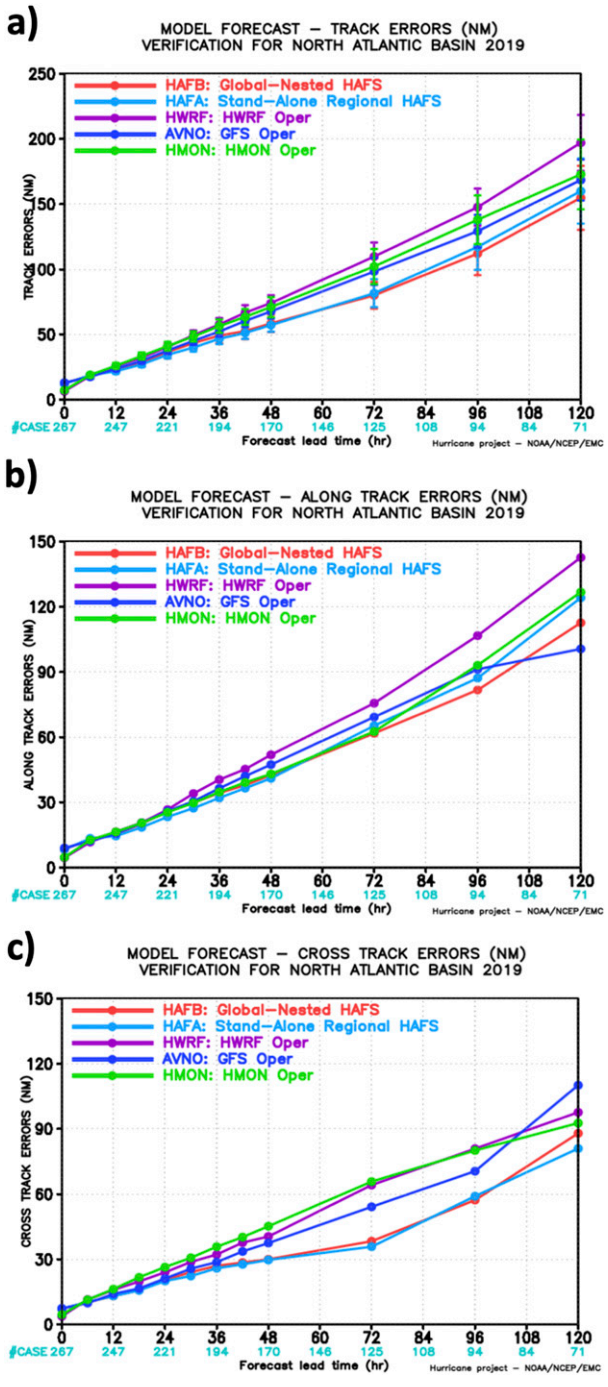

FIG. 2. (a) Mean track errors (n mi) out to 120 h for HAFS-globalnest (red), HAFS-SAR (light blue), HWRF (purple), GFS (blue), and HMON (green). The error bars show the standard error. (b) As in (a), but only for along-track errors. (c) As in (a), but only for cross-track errors.

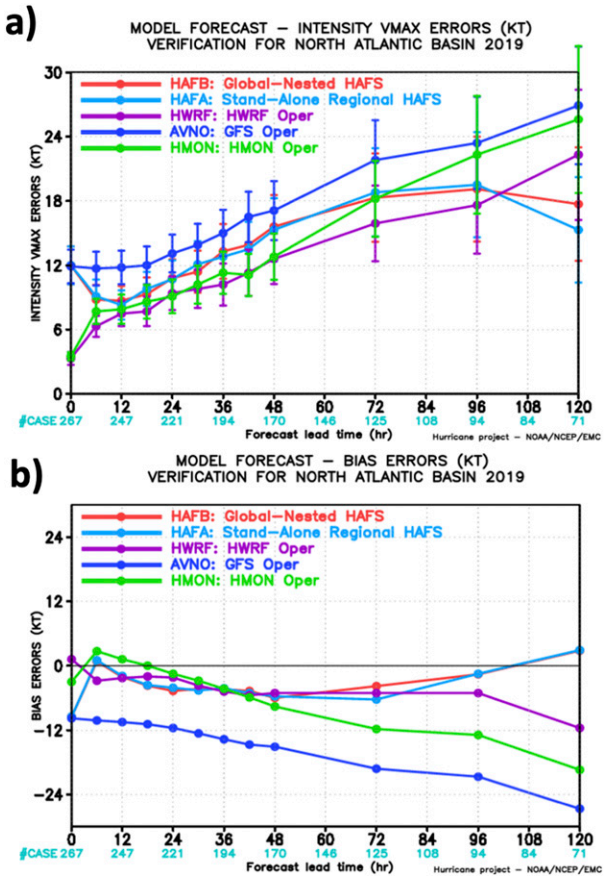

FIG. 3. (a) Mean intensity errors (kt) out to 120 h for HAFS-globalnest (red), HAFS-SAR (light blue), HWRF (purple), GFS (blue), and HMON (green). The error bars show the standard error. (b) As in (a), but for intensity bias.

in HMON, 1.5 km in HWRF) and there are also some differences in vertical resolution and physics that should be kept in mind when comparing these models. The mean errors (with 95% confidence interval) are shown in Fig. 2a, with the along-track and cross-track errors shown in Figs. 2b and 2c. Both configurations of HAFS consistently had the lowest errors of the GFS-based track models out through day 5, with HAFS-globalnest showing slightly lower errors than HAFS-SAR at day 5. Breaking down the errors, the biggest differences between HAFS-globalnest/HAFS-SAR and the other models seemed to be in the cross-track errors, which were noticeably smaller at days 2–5. Some of the individual case studies examined later will look at some of these tracks in more detail.

Looking at the intensity errors and bias (Fig. 3), HAFS-globalnest and HAFS-SAR have very similar bias, and both have large initial errors (a weak negative bias). This is not surprising, as both HAFS-globalnest and HAFS-SAR use the same coarse initialization from the GFS analysis (improving vortex-scale initialization is a subject of ongoing work). However, the errors show that the model ''spins up'' within about 12 h and has errors that are generally comparable to HMON and HWRF by days 4–5, although HWRF is slightly better at days 1–2. The almost immediate improvement in intensity forecasts compared to the global model highlights the ability of high-resolution FV3 within

the 3-km nest to spin up a TC despite a coarse initialization, and demonstrates the value of the high-resolution nest and TC-specific physics. HAFS-globalnest improves on both the track and intensity prediction of a global-only configuration of FV3 (the operational GFS). In terms of intensity bias, HAFS-globalnest has a general weak bias at early leads, similar to HWRF and HMON. While those models' weak bias increases with lead time, the HAFS-globalnest weak bias decreases by later forecast hours. Some of this difference is likely due to the lack of ocean coupling in this version of HAFS, which may be offsetting some of the low bias caused by missed RI events (discussed later). Overall, these results are encouraging for the ability of the model to predict the basic metrics generally used to operationally assess TC forecast skill.

To further examine the model forecast skill of TC structure compared to operational models, the wind radii [34, 50, and 64 kt (1 kt $\approx$ 0.51 m s^{-1})] are examined next, for HAFS-globalnest, HAFS-SAR, and the operational guidance (Fig. 4). Cangialosi and Landsea (2016) documented some of the forecast performance for these metrics from global models, and it is useful to examine the forecast structure of TCs despite observational uncertainties. For 50-kt (R50) and 64-kt (R64) radii, the models are all generally similar, and HAFS-globalnest generally has little bias in R64 at longer lead times. The biggest difference, however, is that HAFS-globalnest has too large of a 34-kt wind radius (R34) at all lead times after spinup (12 h), in contrast to the slight low bias in all of the other models. This means that HAFS-globalnest produces a larger outer TC wind field than the other models, especially for weaker storms. Despite being initialized from the GFS initial conditions and using the same dynamic core, HAFS-globalnest diverges almost immediately from the GFS R34 forecast. The advection scheme may play a role, as the current version of HAFS uses a more diffusive advection scheme than GFS. The surface physics could also be worth evaluating, although HAFS-globalnest used a very similar configuration to the operational HWRF, with drag coefficient decreasing and leveling off at high wind speeds, yet these two models were different in their R34 evolution. Thus, the role of PBL physics, microphysics, and model dynamics in this structural bias is worthy of further exploration.

The 5-day verification statistics were presented earlier in comparison with other experimental and operational models. However, since the HAFS-globalnest forecasts were run out to 7 days, it is worthwhile to compare the forecasts with a global-only forecast, specifically the operational GFS, as well as two other models [the Met Office (UKMO) and European Centre for Medium-Range Weather Forecasting (ECWMF) models] out to day 6 (Fig. 5). It should be noted that the UKMO and ECMWF forecasts were only run twice daily, so the homogeneous sample for this comparison is smaller than the comparison between GFS and HAFS-globalnest only. The improvement in track forecasts from HAFS-globalnest persists to about day 5.5, with the GFS performing better at day 7. UKMO and ECMWF perform somewhat better than both GFS and

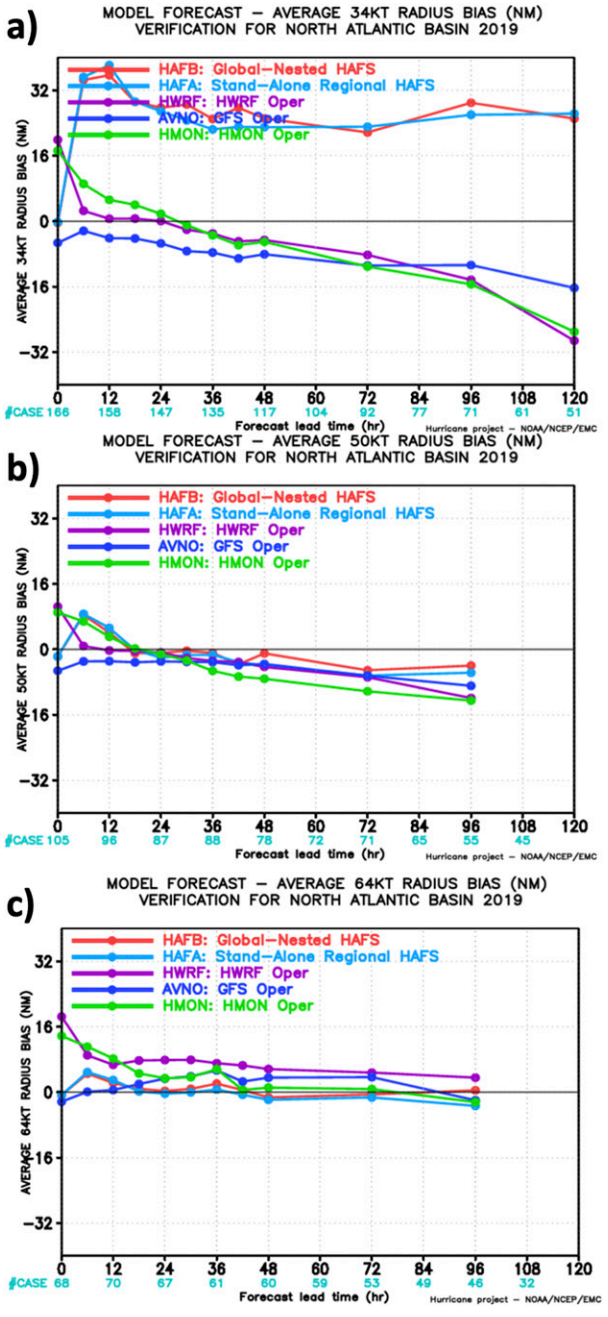

FIG. 4. (a) Mean radius of 34-kt winds bias out to 120 h for HAFS-globalnest (red), HAFS-SAR (light blue), HWRF (purple), GFS (blue), and HMON (green). (b) As in (a), but for 50-kt winds. (c) As in (a), but for 64-kt winds.

HAFS-globalnest at day 6. The sample size is much lower at days 6–7, which makes interpretation of the results difficult, but the results suggest that further work is needed to evaluate and improve the HAFS-globalnest representation of synoptic features driving TC track. For example, it is important for future work to assess whether the large-scale flow forecasts themselves are skillful compared to the operational

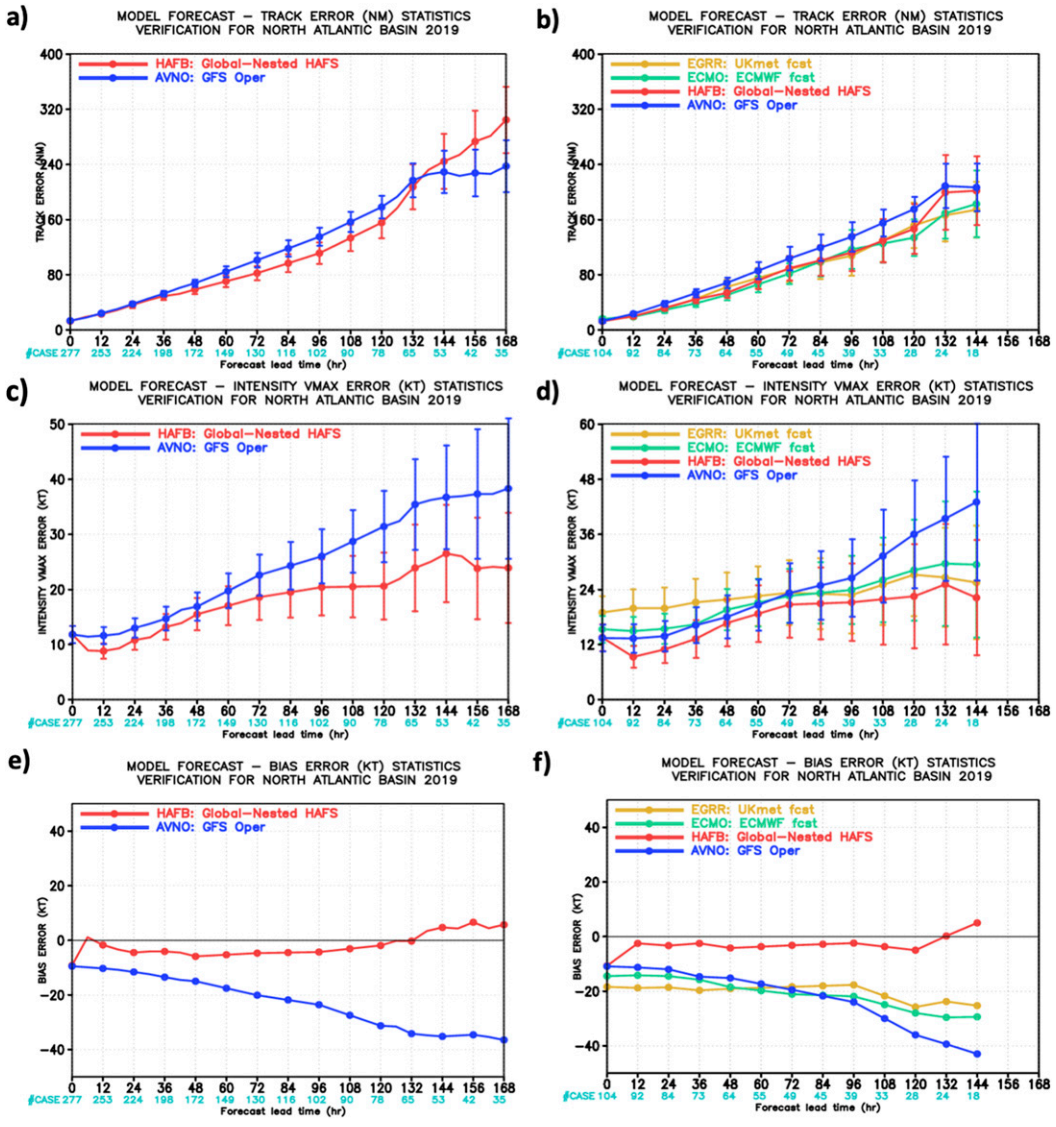

FIG. 5. (a) The 0–7-day track forecast errors (n mi) from HAFS-globalnest (red) and operational GFS (blue). The error bars show the standard error. (b) As in (a), but only to 6 days and also including operational ECMWF and UKMO. (c) The 0–7-day intensity forecast errors (kt) from HAFS-globalnest (red) and operational GFS (blue). The error bars show the standard error. (d) As in (c), but only to 6 days and also including operational ECMWF and UKMO. (e) The 0–7-day intensity bias (kt) from HAFS-globalnest (red) and operational GFS (blue). (f) As in (e), but only to 6 days and also including operational ECMWF and UKMO.

GFS. The intensity forecasts from HAFS-globalnest, however, demonstrate the value of the high-resolution nest. While GFS had increasing errors all the way out to day 7 due to an increasing negative bias (again, subject to the uncertainties associated with a small sample), the intensity errors for HAFS-globalnest seemed to saturate around day 4, with a relatively small bias throughout. ECMWF and UKMO had somewhat lower overall intensity error than GFS, but still had a pronounced negative bias increasing with lead time. The skill of HAFS-globalnest relative to GFS kept increasing with increasing lead, showing how important high-resolution TC structure (and the possible improvement of physics parameterizations at higher resolution) is to intensity forecasts.

One of the biggest goals of HFIP (Gall et al. 2013) is to improve forecasts of rapid intensification (RI), which is defined typically as a 30-kt increase in maximum sustained wind speed in 24 h (Kaplan and DeMaria 2003). Performance diagrams (Roebber 2009) illustrate the skill for RI forecasts from HAFS-globalnest and the operational HWRF and HMON (Fig. 6) for this 30 kt $(24\,\text{h})^{-1}$ threshold as well as a slightly stronger threshold [35 kt $(24\,\text{h})^{-1}$] and slightly lower threshold [25 kt $(24\,\text{h})^{-1}$], both of which are described in Kaplan et al. (2010). The operational GFS was not included

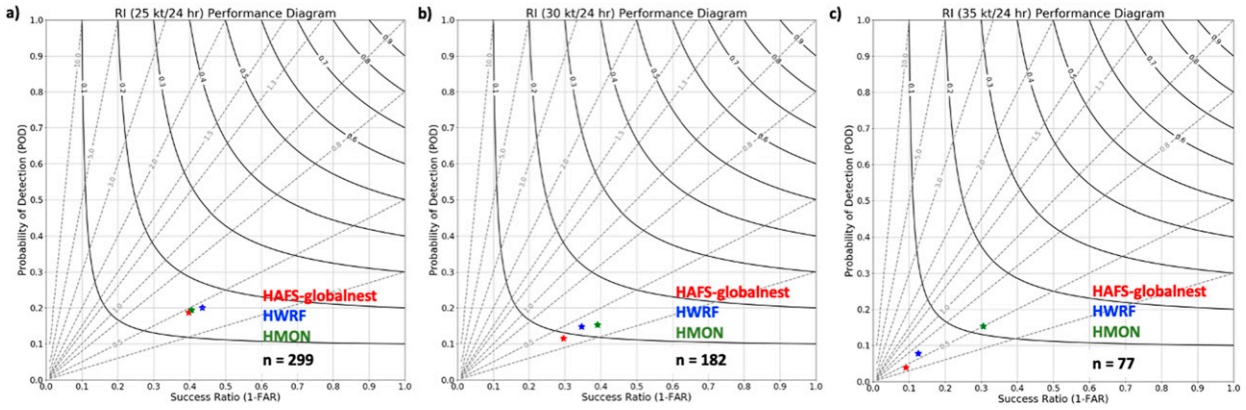

FIG. 6. (a) Performance diagram showing success ratio (1 − false alarm ratio) vs probability of detection for 25 kt $(24\,\text{h})^{-1}$ rapid intensification for HAFS-globalnest (red), HWRF (blue), and HMON (green). Black solid lines are contours of critical success index, and gray dashed lines are contours of bias. The total number of 24-h periods during which RI was observed is listed in bold. (b) As in (a), but for 30 kt $(24\,\text{h})^{-1}$. (c) As in (a), but for 35 kt $(24\,\text{h})^{-1}$.

in order to have a more realistic comparison of high-resolution nested model RI forecast performance. These performance diagrams are based on several metrics described in Roebber (2009) from a 2 × 2 contingency table including false alarm ratio (FAR), probability of detection (POD), and critical success index (CSI). These are summarized in Table 2. To improve sample size for evaluation, all forecast periods were evaluated together. In the performance diagrams, RI forecast skill improves as you go up (improved probability of detection) and to the right (reduced false alarm ratio). For all three definitions, HAFS-globalnest has worse overall performance with RI prediction than HWRF and HMON. The difference in performance is small for the 25 kt $(24\,\text{h})^{-1}$ definition, but increases with increasing RI magnitude. This comparison is perhaps slightly inconsistent, given the 3-km resolution compared to 1.5 km for the current HWRF and 2 km for HMON (and some differences in vertical resolution as well). The degree to which resolution and model physics improve RI prediction in HAFS-globalnest is a subject of ongoing research. It should also be noted that for all three models shown, the success ratio (1 − FAR) was higher than the POD, indicating that missed events were a bigger issue than false alarms for this year's TCs. In the case of Hurricane Dorian, the track forecasts likely played a role in the missed RI forecasts (e.g., Emanuel and Zhang 2016), as will be discussed later. The effects of vertical wind shear also tend to make RI prediction difficult, because of the importance of correctly capturing inner-core details (e.g., Zhang and Rogers 2019). This affected the forecasts for other storms in 2019 such as Hurricane Jerry and Hurricane Humberto.

A key question motivated by the 2019 real-time experiments was the difference in performance between HAFS-globalnest and HAFS-SAR. As mentioned previously, the main difference between the two configurations is that HAFS-globalnest has a global 13-km FV3 running concurrently in two-way feedback with the 3-km nested domain, while HAFS-SAR solely has the 3-km nest running, with one-way feedback providing boundary conditions from the operational GFS (also at 13 km). Given the increased computational requirements needed to concurrently run a global model with the nest, it is important to assess the difference in skill. Overall, as discussed above, both models had similar performance for track, intensity, and structure prediction. However, as discussed later, there were some noteworthy cases where differences were seen. To explore how the different boundary conditions may have impacted the track forecasts, track errors were calculated inside and outside a region (or "box") at the center of the domain, away from the boundary where feedback between the nest and boundary conditions would occur. This box covers approximately 15°–32°N and 35°–85°W. For the "inside the box" cases, which form the majority of the dataset, both configurations of HAFS are very similar in terms of track error (Fig. 7), and perform better than the operational GFS. However, for the smaller set of cases

TABLE 2. Description of the metrics used to calculate performance diagrams for rapid intensification (based on Roebber 2009).

Metric	Definition	Abbreviation
Hit	RI was forecast, and occurred	h
False alarm	RI was forecast, but did not occur	f
Miss	RI was not forecast, but occurred	m
Correct null	RI was not forecast, and did not occur	n
Probability of detection	$\frac{h}{h+m}$	POD
False alarm ratio	$\frac{f}{f+h}$	FAR
Critical success index	$\frac{h}{h+f+m}$	CSI
Success ratio	1 − FAR	SR
Bias	$\frac{\text{POD}}{\text{SR}}$	

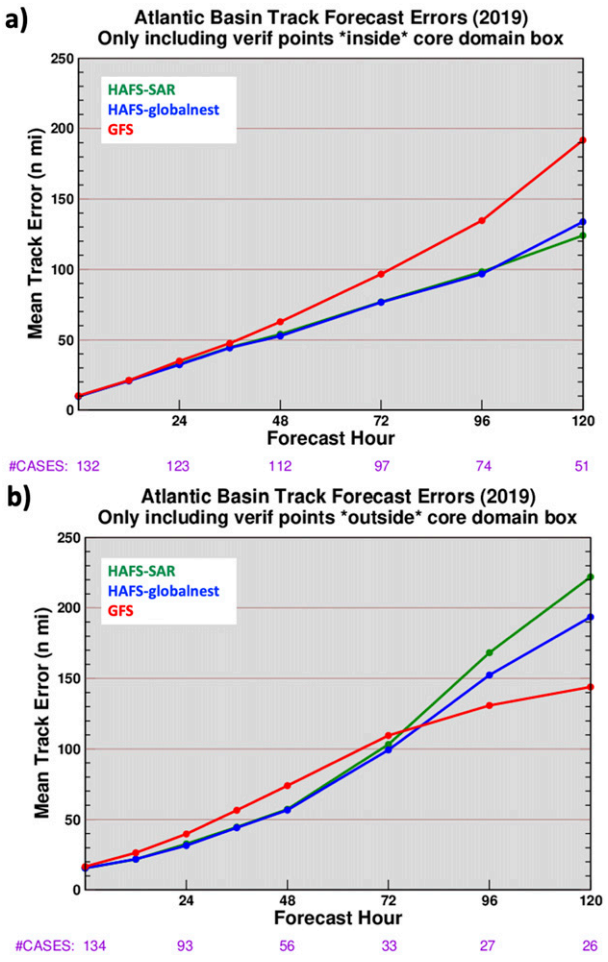

FIG. 7. (a) Track forecast errors (n mi) for HAFS-globalnest (red), HAFS-SAR (green), and operational GFS (blue) for the cases within the box at the center of the nested domain. (b) As in (a), but for the cases outside the box at the center of the nested domain.

"outside the box," the results are different. Through 72 h, both configurations of HAFS are very similar and better than GFS. However, at longer leads, GFS is actually better, and the two configurations of HAFS diverge, with HAFS-globalnest

TABLE 3. Fraction of HAFS-globalnest cases that predicted TC genesis (defined as 12 consecutive forecast hours at an intensity of 35+ kt) for 10 nondeveloping invests.

Storm ID	Genesis predicted
AL932019a	0/11
AL952019a	0/6
AL962019a	8/11
AL902019b	0/1
AL922019b	9/13
AL942019b	0/15
AL962019b	0/2
AL912019c	4/4
AL922019c	2/2
AL952019c	3/3
Total	26/67

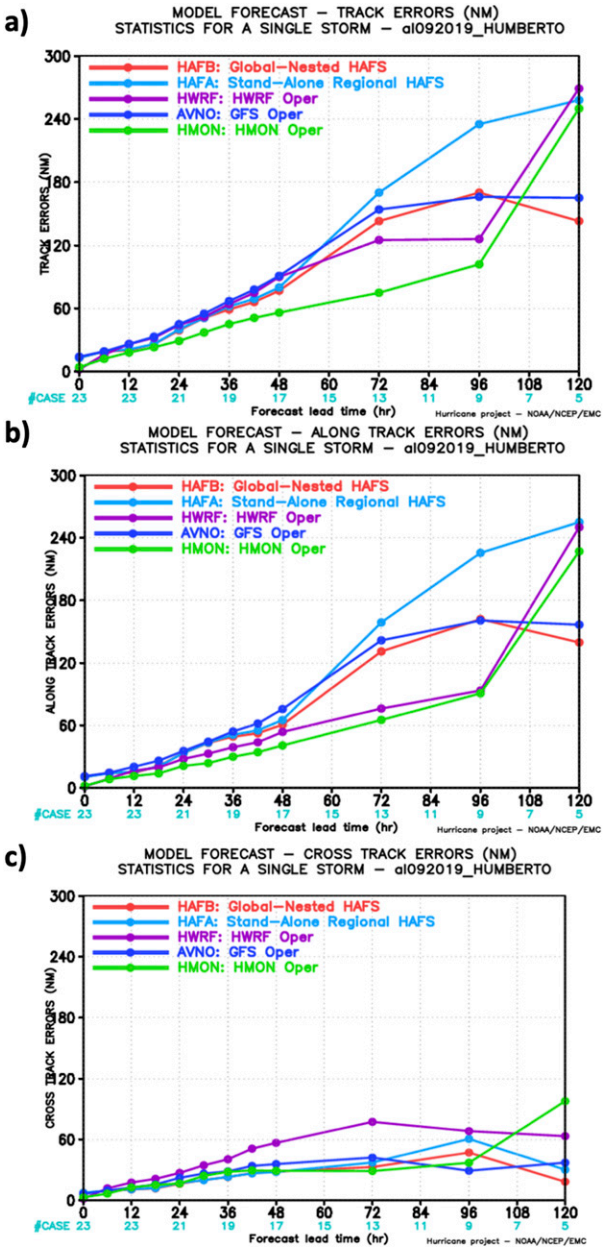

FIG. 8. (a) Mean track errors (n mi) for Humberto out to 120 h for HAFS-globalnest (red), HAFS-SAR (light blue), HWRF (purple), GFS (blue), and HMON (green). The error bars show the standard error. (b) As in (a), but only for along-track errors. (c) As in (a), but only for cross-track errors.

having lower track errors. An example of a case where the track errors for HAFS-globalnest were lower than for HAFS-SAR in this region near the nest boundaries is examined in the case study of Hurricane Humberto in section 3b(1), and illustrates a case where the midlatitude westerly flow near the northern edge of the nest appeared to be better predicted in HAFS-globalnest.

One other key question in these composite results was whether or not the large high-resolution nest spuriously

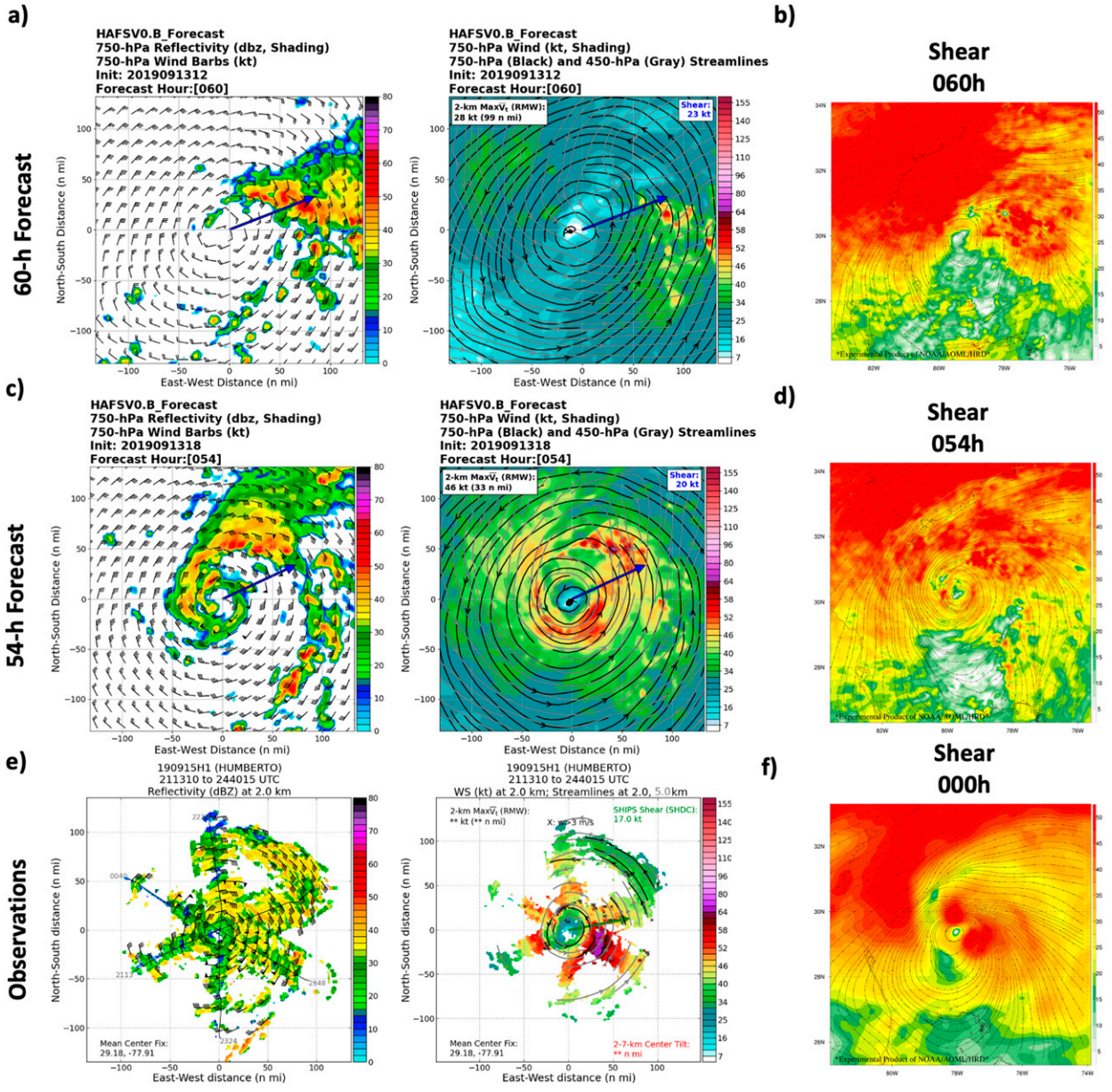

FIG. 9. (a) (left)The 750-hPa reflectivity (shaded) and wind (kt) from the HAFS-globalnest forecast initialized at 1200 UTC 13 Sep 2019 for Humberto, 60 h. (right) The 750-hPa wind (shaded; kt), with 750-hPa streamlines (black) and 450-hPa streamlines (gray) overlaid. The shear vector is shown in blue in both images. (b) The 850–200-hPa vertical shear (kt) from the 60-h HAFS forecast initialized at 1200 UTC 13 Sep 2019. (c) As in (a), but for the forecast initialized at 1800 UTC 13 Sep 2019, 54 h. (d) As in (b), but for the forecast initialized at 1800 UTC 13 Sep 2019, 54 h. (e) (left) NOAA P-3 observed 2-km reflectivity (shaded) and wind (kt). (right) NOAA P-3 observed 2-km wind (shaded; kt), with 2-km streamlines (black) and 5-km streamlines (gray) overlaid. The shear vector is shown in blue (green for the observed) on the right. (f) Analyzed 850–200-hPa shear (kt) at 0000 UTC 16 Sep 2019.

developed TCs that did not develop in the real world. To help answer this question, the HAFS-globalnest forecasts from 10 different ''invests'' (areas of interest, numbered 90–99, marked by the National Hurricane Center, which may or may not later be classified as TCs) that did not become TCs were examined, to see how many HAFS forecasts showed these developing into TCs. For the purposes of this study, the invest was defined as a TC in the model if the intensity from the Marchok tracker was 35 kt or more for 12 h (four consecutive forecast points, output every 3 h). The results are shown in Table 3. For five of the systems, none of the HAFS-globalnest forecasts showed development into a TC (correctly). For the other five systems, HAFS-globalnest did incorrectly show TC genesis for a majority of the forecasts. Overall, however, only 26 of 67 forecasts for these nondeveloping invests showed TC formation. The continued evaluation of TC

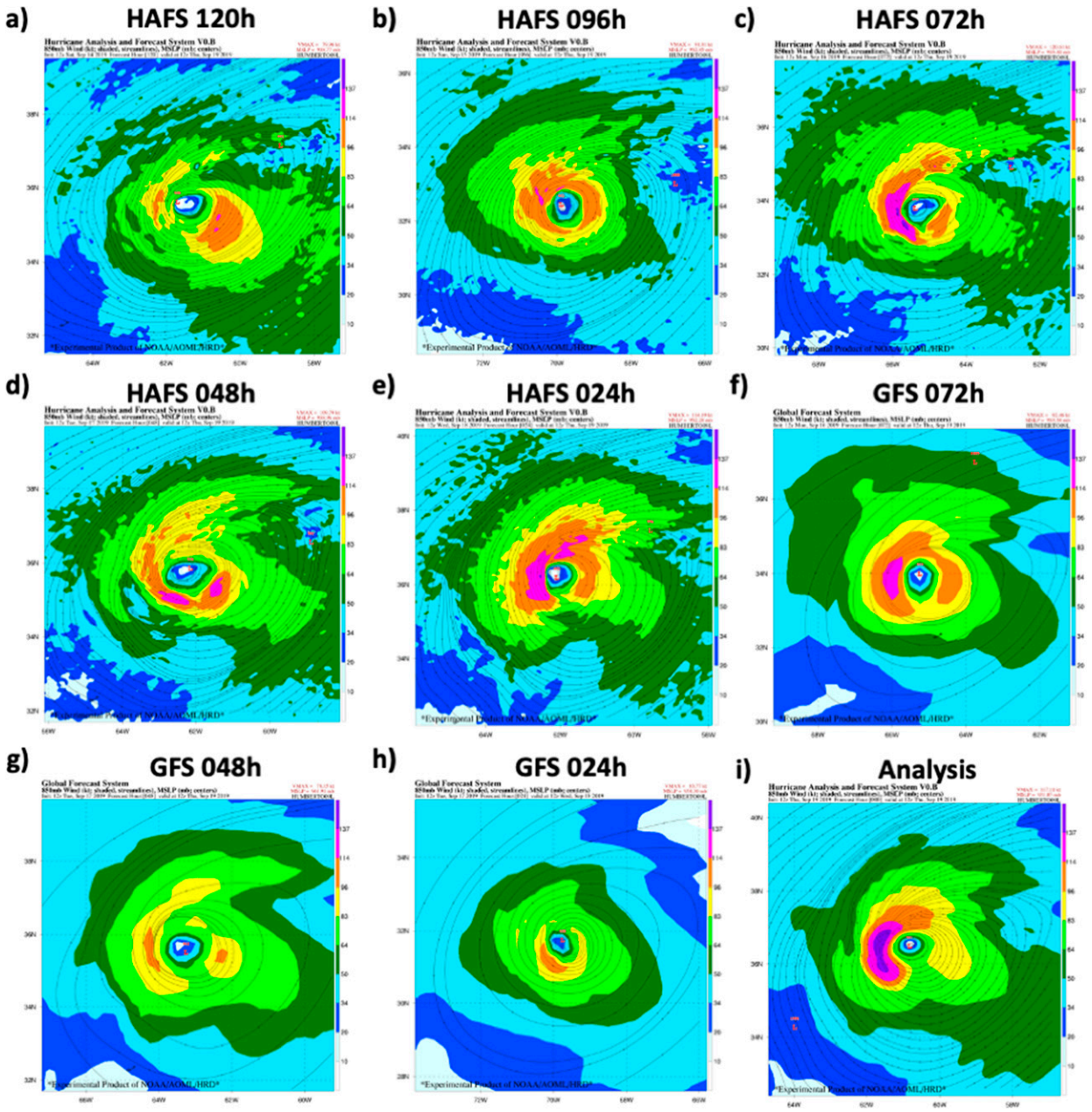

FIG. 10. (a) The 850-hPa wind for Humberto valid at 1200 UTC 12 Sep 2019, from a 120-h forecast from HAFS-globalnest. (b) As in (a), but from a 96-h forecast. (c) As in (a), but from a 72-h forecast. (d) As in (a), but from a 48-h forecast. (e) As in (a), but from a 72-h forecast. (f) As in (c), but for the operational GFS. (g) As in (d), but for the operational GFS. (h) As in (e), but for the operational GFS. (i) Analyzed 850-hPa wind.

genesis skill will be an important aspect of analyzing HAFS-globalnest in further development.

b. Case studies

1) HURRICANE HUMBERTO

Hurricane Humberto was one of three major hurricanes to form in the North Atlantic Basin during the 2019 season. It formed east of Florida, and drifted north parallel to the coast before accelerating northeast, impacting Bermuda as it did so. HAFS-globalnest initially had a left-of-observed track bias, showing a weak vortex heading into the Florida Peninsula. As a result, there was also a very pronounced weak bias in the early forecasts. After these first few incorrect forecasts, however, HAFS-globalnest correctly showed the slow north motion, followed by acceleration to the northeast. As a result, the intensity forecasts also improved, with many cycles correctly showing intensification into a major hurricane. The later track forecasts were generally correct (HAFS-globalnest had lower day-5 track errors (Fig. 8) for Humberto than GFS, HWRF,

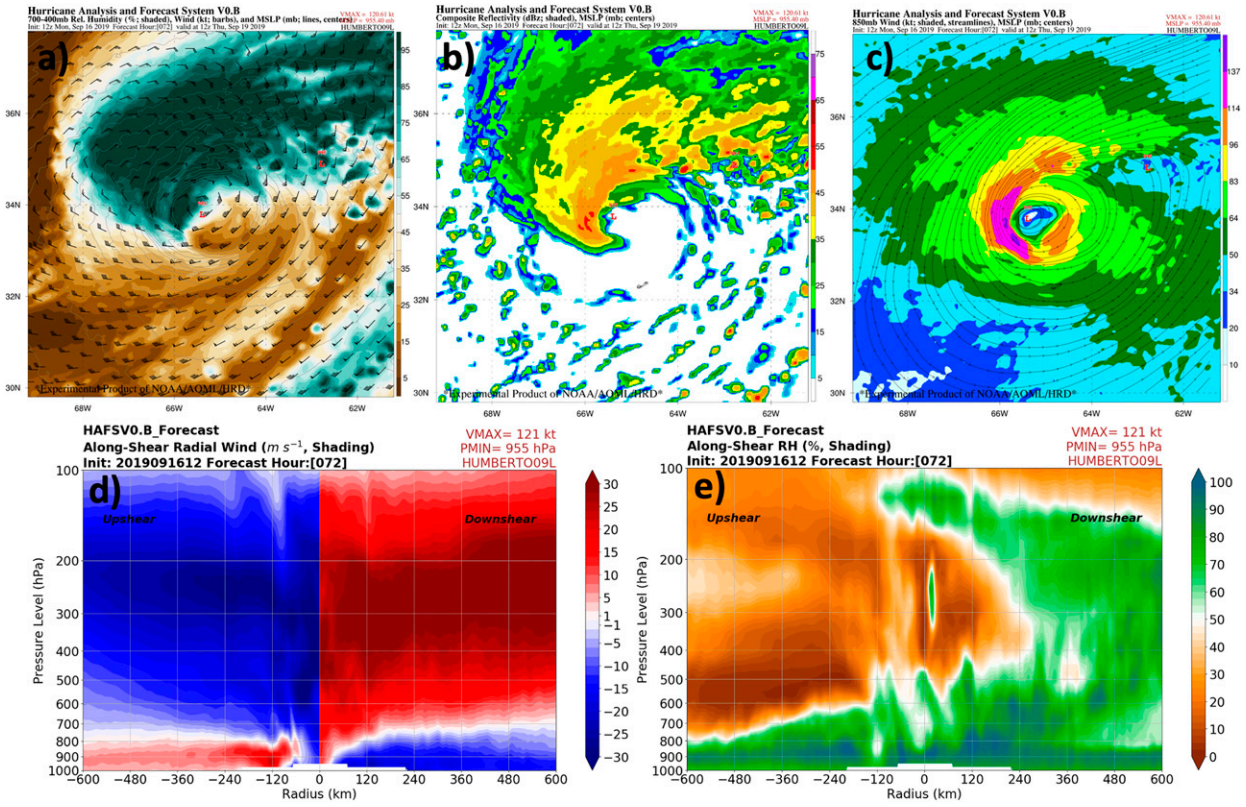

FIG. 11. (a) The 700–400-hPa average relative humidity (%) from HAFS-globalnest initialized 1200 UTC 16 Sep 2019 for Humberto, valid at 1200 UTC 19 Sep 2019. (b) As in (a), but for composite reflectivity (dBZ). (c) As in (a), but for 850-hPa wind (kt). (d) Along-shear cross section of radial wind ($m\ s^{-1}$) for the same valid time. (e) Along-shear cross section of relative humidity (%) for the same valid time.

and HAFS-SAR), although a few cycles showed too much interaction with a trough, leading to an incorrect forecast of a turn back toward the northwest. This will be examined in more detail.

In addition to the early uncertainties due to track, there were some uncertainties in the forecast structure and intensity of Humberto due to the relative roles of shear and the vortex structure. Figure 9 shows two forecasts, 1200 UTC 13 September

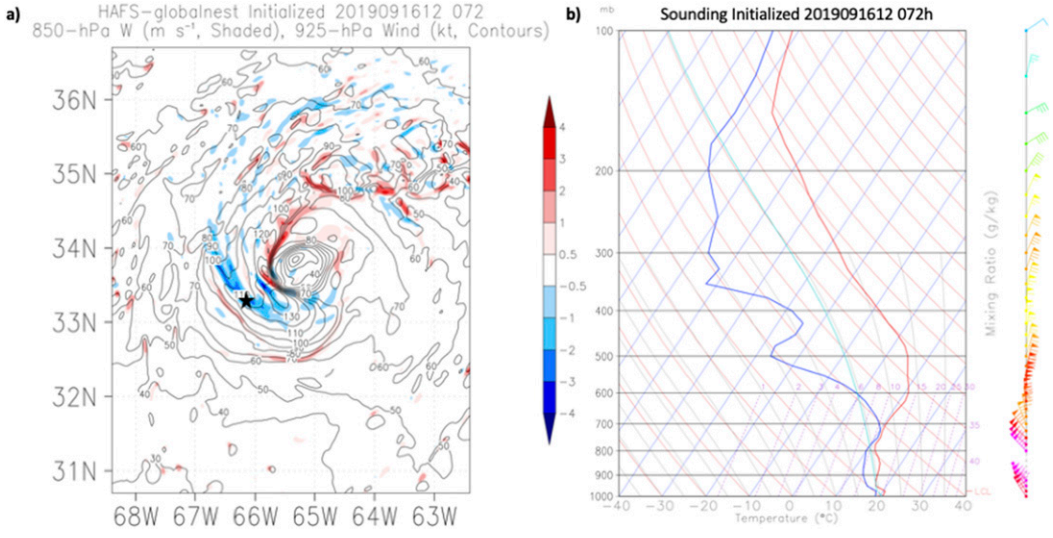

FIG. 12. (a) HAFS-globalnest initialized 1200 UTC 16 Sep 2019 72-h forecast 850-hPa vertical velocity ($m\ s^{-1}$; shaded) and 925-hPa wind (kt; contours). The star represents the location of the sounding. (b) Sounding from the HAFS-globalnest forecast. The blue line shows the dewpoint and the red line shows the temperature. The wind barbs are on the right side.

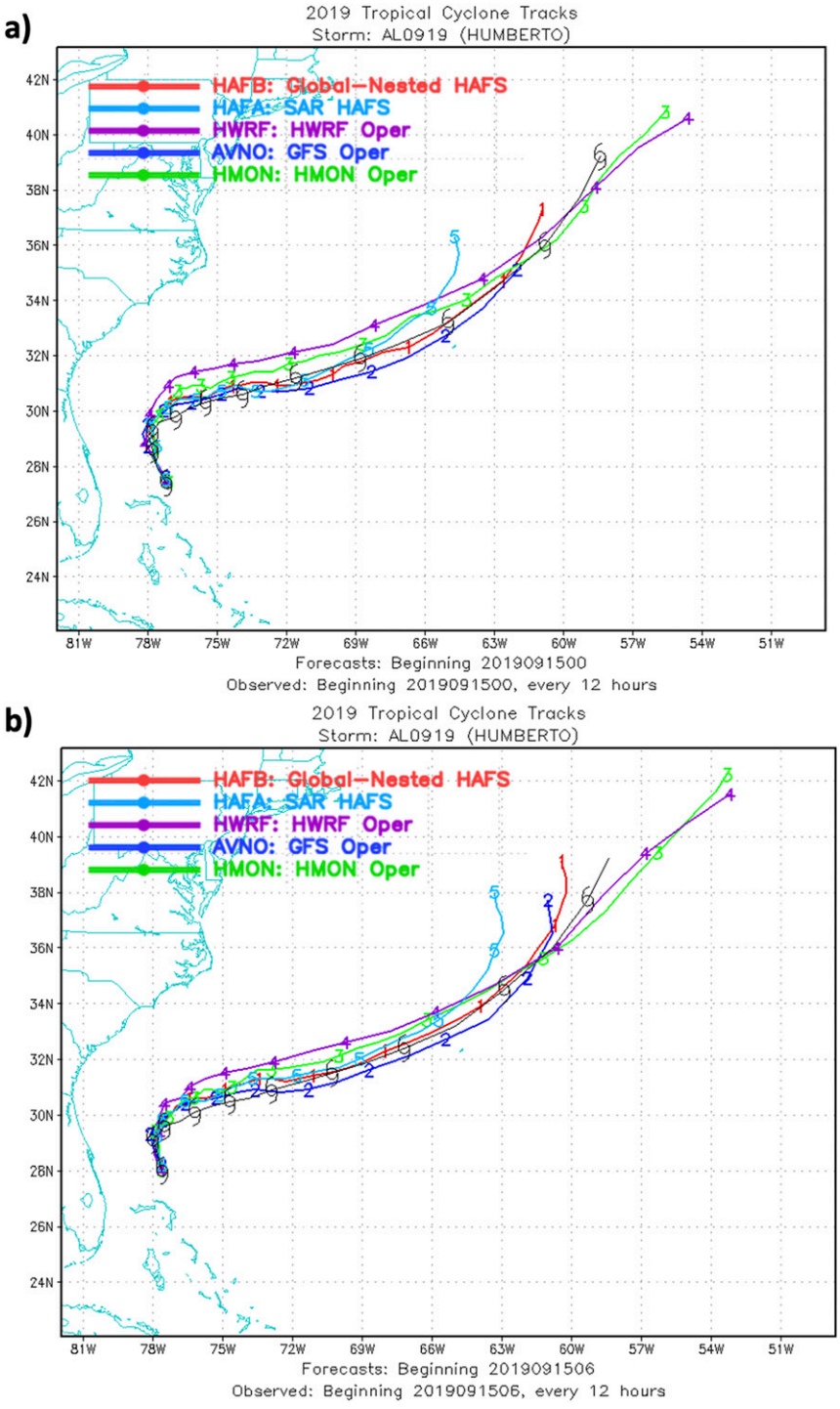

FIG. 13. (a) The 5-day track forecasts for Humberto initialized at 0000 UTC 15 Sep 2019 from HAFS-globalnest (red), HAFS-SAR (light blue), operational HWRF (purple), operational HMON (green), and operational GFS (dark blue). (b) As in (a), but initialized at 0600 UTC 15 Sep 2019.

2019 and 1800 UTC 13 September 2019, and the associated precipitation and vortex structure (both of these cycles correctly kept Humberto offshore). The P-3 observations are also shown, and the earlier forecast had a much more asymmetric precipitation pattern, and weaker vortex. Just 6 h later, the forecast precipitation was still somewhat asymmetric compared to observations, but a much more symmetric and compact inner core was forecast to form and closer to reality. Looking at the shear fields in the

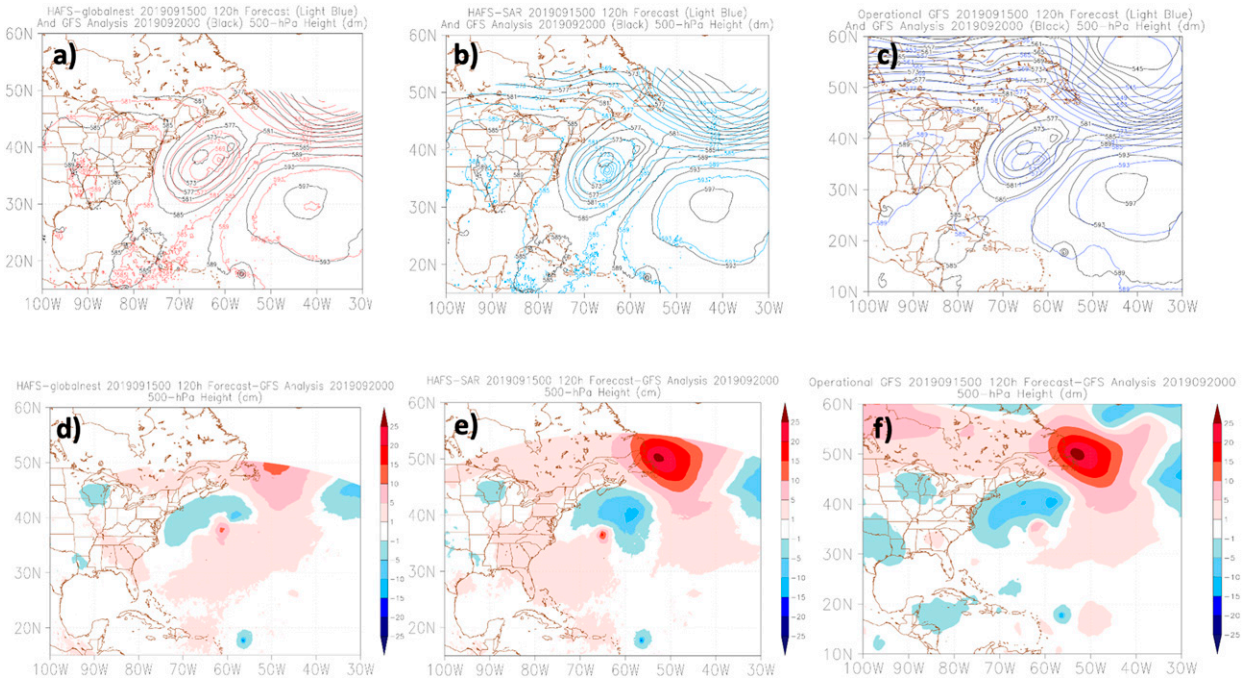

FIG. 14. (a) The 120-h 500-hPa height (dam; red) from HAFS-globalnest, initialized at 0000 UTC 15 Sep 2019, and the GFS analysis of 500-hPa height (dam; black) at 0000 UTC 20 Sep 2019. (b) As in (a), but for HAFS-SAR (light blue). (c) As in (a), but for operational GFS (dark blue). (d) 500-hPa height difference (dam; shaded) between the GFS analysis at 0000 UTC 20 Sep 2019 and the 120-h forecast from HAFS-globalnest. (e) As in (d), but for HAFS-SAR. (f) As in (d), but for the operational GFS.

forecast and analysis (Figs. 9b,d,f), the large-scale shear was similar in both forecasts (and the analysis), and the main differences were the convective-scale symmetry differences seen in the radar data. Similar precipitation processes were shown in Hazelton et al. (2020) to be important in the intensification of Hurricane Michael (2018) in shear. The later forecast peaked with a maximum 10-m wind of 116 kt at 126 h, while the earlier forecast only had a peak 10-m wind of 80 kt at the same synoptic time.

The interesting aspect of the evolution of Humberto, later during its life cycle, was a wind maximum that developed on the west side of the TC as it accelerated to the northeast. This asymmetry was somewhat peculiar because it was not on the southeast side of the TC where it would be expected due to motion asymmetry. In general, this asymmetry was well forecast by HAFS-globalnest (Figs. 10a–e), especially at 072 h and shorter lead times, but the operational GFS generally missed the magnitude of the asymmetric wind maximum (Figs. 10f–h). The structure of the wind maximum, as well as the position of Humberto in the right-entrance region of a jet streak, suggest the presence intrusion described as a ''sting jet'' (e.g., Browning and Field 2004; Stewart 2020). Taking a closer look at some of the relevant variables, including relative humidity and reflectivity (Fig. 11) show that the wind maximum was indeed collocated with a ''comma head'' of reflectivity in a region of high RH gradient where evaporative cooling would be maximized. Figures 11d and 11e, which are cross sections along the large-scale shear vector, illustrate the low-level inflow and drying originating on the upshear side and reaching lower levels just upshear (southwest) of the TC center. This region near the ''comma head'' was also associated with significant midlevel drying and low-level subsidence (Fig. 12) bringing strong winds toward the surface. The evaporative sinking and banded structure are consistent with the sting jet structure described in Browning and Field (2004). This wind maximum allowed the storm to maintain major hurricane intensity despite strong shear and the beginning of extratropical transition (ET), and the TC produced wind gusts over 100 kt (51 m s^{-1}) on Bermuda around this time (Stewart 2020).

One other key aspect of Hurricane Humberto is that it was one of the cases where there were some major track forecast differences between HAFS-globalnest and HAFS-SAR. HAFS-globalnest produced markedly better track forecasts (Fig. 8) for this case. Looking at some of the individual track forecasts (Fig. 13), there were a few cycles where both configurations of HAFS were showing Humberto curving back toward the northwest toward the United States. This was especially pronounced in HAFS-SAR, and was not seen in any of the other operational or experimental guidance, although a few operational ensemble members (not shown) did hint at this possibility. This is one case where the R34 high bias in HAFS [which was over 30 n mi (1 n mi = 1.852 km) larger for this case than the bias in HWRF/HMON] may have contributed to excessive interaction with synoptic features to the northwest of the TC. The observed storm continued northeast without any such bend back to the northwest, leading to large longer-term track errors. These errors occurred in the outside the box

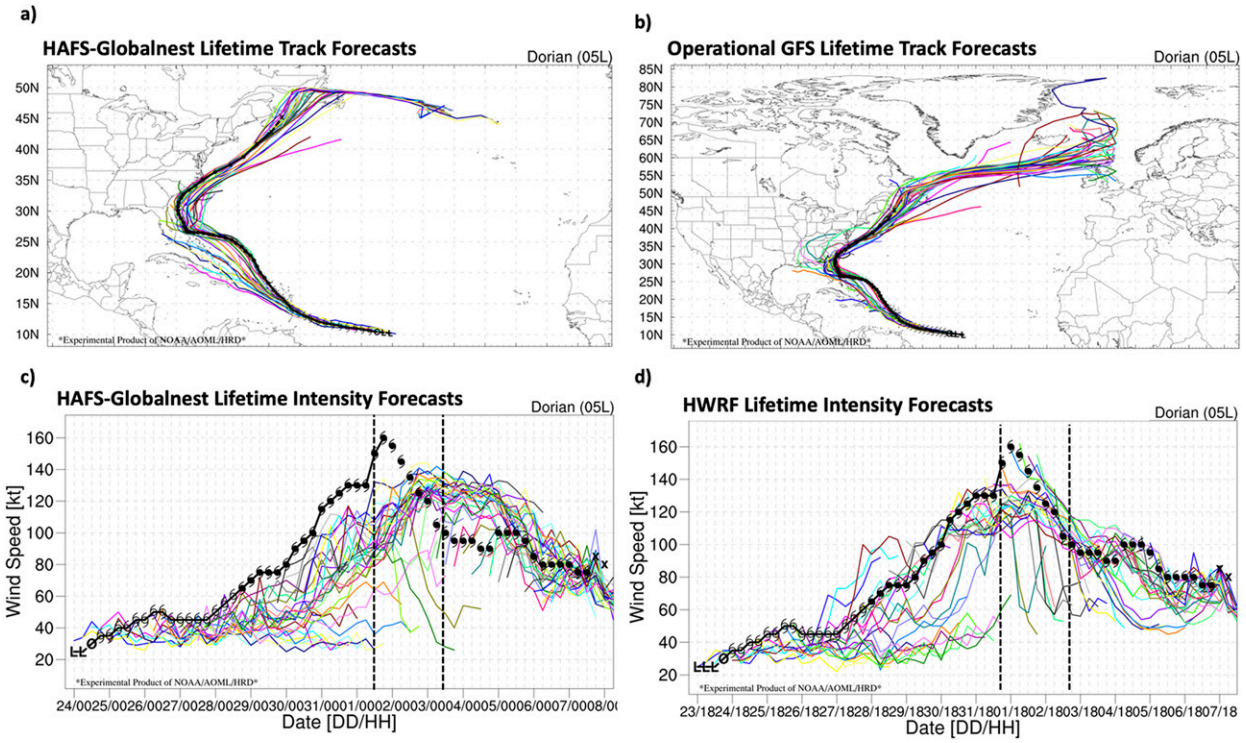

FIG. 15. (a) All HAFS-globalnest track forecasts for Dorian. The black line represents best track. (b) As in (a), but for operational GFS. (c) All HAFS-globalnest intensity forecast for Dorian. The black line represents best track. The dashed black lines show the period where Dorian was stalling in the Bahamas. (d) As in (c), but for operational HWRF.

region fairly close to the northern boundary of the nest. The synoptic evolution along this region was more realistic in this case for HAFS-globalnest with two-way feedback than HAFS-SAR with one-way feedback (Fig. 14). In HAFS-globalnest, the cutoff low interacting with Humberto stayed separate to the west as was observed (although they did get closer than in reality). In HAFS-SAR, the cutoff low appeared to phase with Humberto, swinging it unrealistically to the northwest. HAFS-globalnest also had the smallest error (compared with the analysis) in the ridge to the northeast of Humberto, while the operational GFS and HAFS-SAR (which received boundary conditions from the GFS) had too much ridging in this area, promoting an unrealistically amplified flow. This improved interaction in HAFS-globalnest led to the more realistic track forecast for this case.

2) HURRICANE DORIAN

Hurricane Dorian was one of the most impactful tropical cyclones of the 2019 Atlantic season, devastating the Bahamas as a category-5 hurricane and also making landfall in Cape Hatteras, North Carolina, as a category-1 hurricane (Fig. 15). The track forecasts showed the adjustment from initial forecasts bringing the storm through the Greater Antilles to a more rightward track into the Western Atlantic. Later on, most HAFS-globalnest forecasts were consistently showing the storm staying east of Florida (Fig. 15a) and turning north in the Bahamas, which turned out to be the correct forecast. This was in sharp contrast to the operational GFS (Fig. 15b), which more consistently predicted a Florida landfall. The intensity forecasts generally showed steady intensification (Fig. 15c), although most missed the initial deepening rate, and none captured the 160-kt peak. After the peak, there was a slight high bias as Dorian stalled in the Bahamas (indicated by the dashed black lines in Figs. 15c and 14d).

To highlight the difference in tracks between HAFS-globalnest (and HAFS-SAR) and the other GFS-based guidance for Dorian, two forecasts are shown (Fig. 16): one initialized at 0000 UTC 29 August 2019 and one initialized at 0000 UTC 30 August 2019. In both of these forecasts, GFS, HWRF, and HMON all showed Dorian making landfall along the East Coast of Florida, while HAFS-globalnest (and HAFS-SAR) was able to show the slower motion and turn to the north in the Bahamas. This difference was consistent over multiple forecasts in this period.

In the real-time operations, one of the biggest issues for Dorian was track forecasts in the eastern Caribbean. Early forecasts showed the storm would stay weak and pass near or over Puerto Rico and/or Hispaniola. In reality, the center redeveloped farther northeast (National Hurricane Center 2019a), and the observed track was outside the track forecast spread of operational and experimental guidance (deterministic and ensemble). There were some hints in these real-time HAFS-globalnest forecasts of the importance of the development of a small-scale core to the track and intensity forecasts during this period, as illustrated in the precipitation and wind structure in a Dorian forecast initialized at 0000 UTC

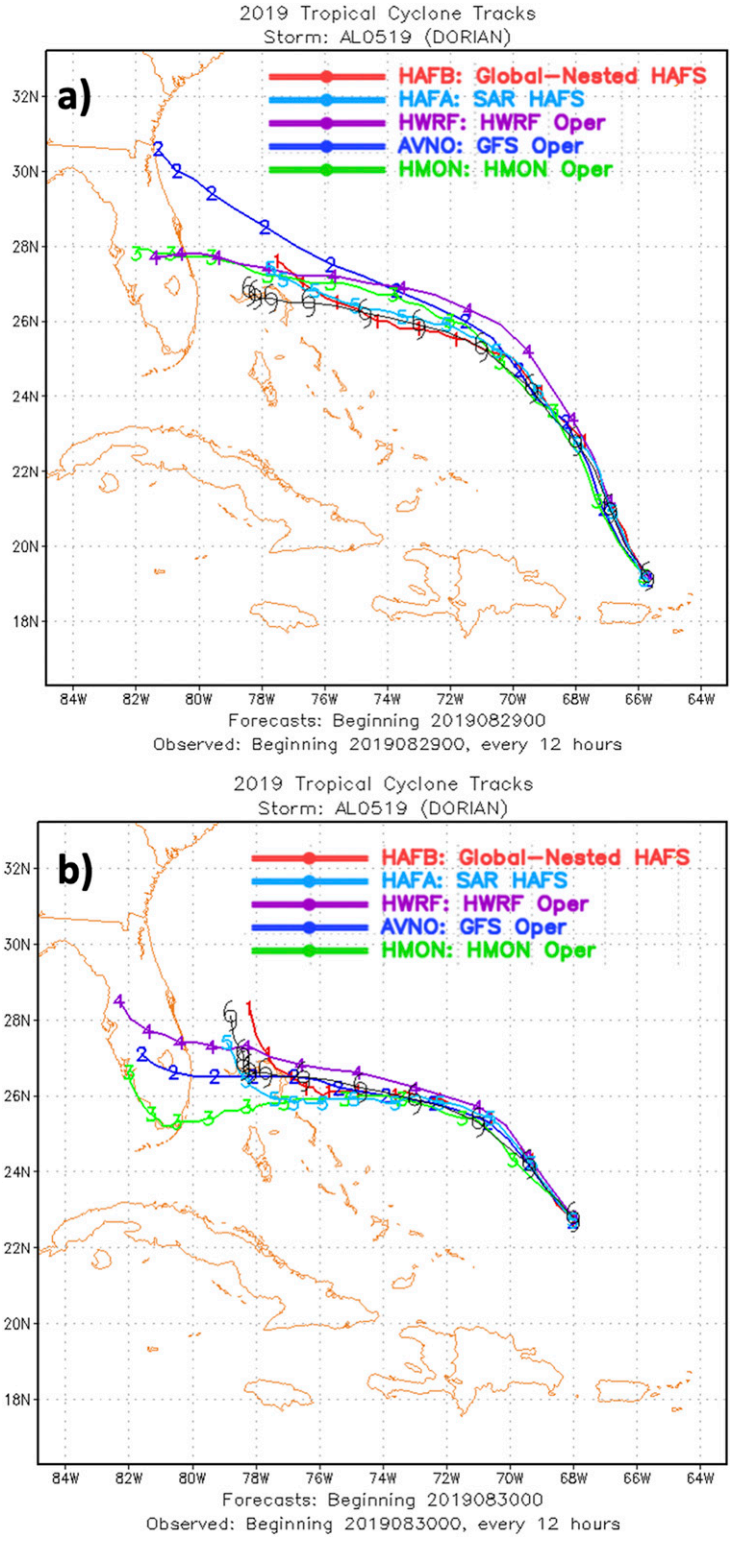

FIG. 16. (a) The 5-day track forecasts for Dorian initialized at 0000 UTC 29 Aug 2019 from HAFS-globalnest (red), HAFS-SAR (light blue), operational HWRF (purple), operational HMON (green), and operational GFS (dark blue). (b) As in (a), but initialized at 0000 UTC 30 Aug 2019.

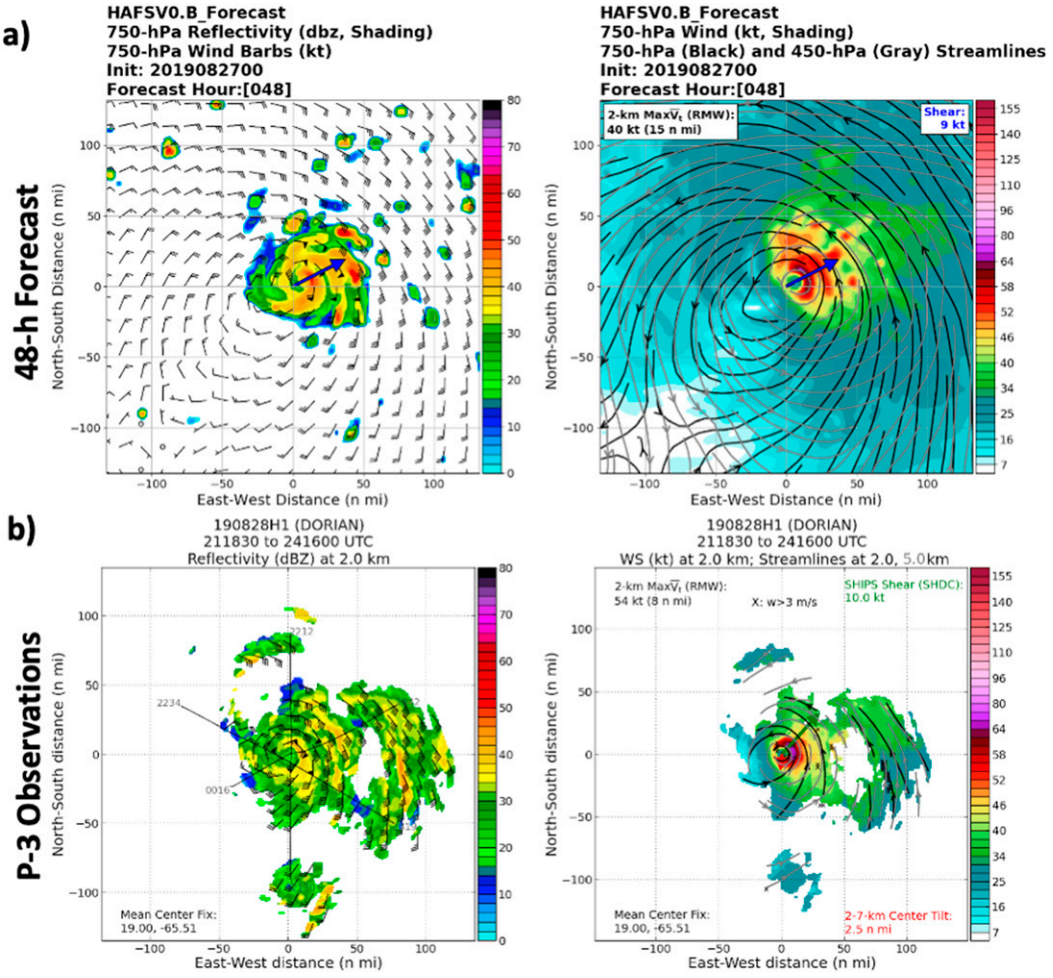

FIG. 17. (a) (left) The 750-hPa reflectivity (shaded) and wind (kt) from the HAFS-globalnest forecast initialized at 1800 UTC 26 Aug 2019 for Dorian. (right) The 750-hPa wind (shaded; kt), with 750-hPa streamlines (black) and 450-hPa streamlines (gray) overlaid. The shear vector is shown in blue in both images. (b) (left) NOAA P-3 observed 2-km reflectivity (shaded) and wind (kt). (right) NOAA P-3 observed 2-km wind (shaded; kt), with 2-km streamlines (black) and 5-km streamlines (gray) overlaid. The shear vector is shown in blue (green for the observed) on the right.

27 August 2019 (Fig. 17). The equivalent observed data are also shown, from the merged analysis from the NOAA P-3 flight into Dorian around this time. In this forecast, the core precipitation, although confined to a small area, is more symmetric. In addition, a small, strong core of winds is evident. This forecast matched the observed structure of Dorian, which also showed a small but robust core. The degree to which the prediction of this early inner-core evolution determined the long-term fate of Dorian is an important question, beyond the scope of this paper, which is being examined in a follow-up study using an ensemble of HAFS forecasts initialized at the time mentioned above.

3) HURRICANE LORENZO

Hurricane Lorenzo was the third major hurricane of the 2019 Atlantic hurricane season. It became a category-5 hurricane farther to the north and east than any other storm in the Atlantic on record (National Hurricane Center 2019b). The track forecasts were generally consistent with the observed track, showing the turn to the north over the eastern Atlantic (Fig. 18). There were a few cases with a left bias initially, but the rest of the forecasts showed a correct timing and speed of the recurve. For intensity, HAFS-globalnest correctly predicted the initial intensification into a major hurricane (although the peak was slightly underestimated). The later intensity forecasts had some issues, however. HAFS-globalnest generally showed a long period of near-steady intensity or slow intensification before eventual decay. However, what really occurred was a secondary peak at category-5 intensity, followed by a precipitous decline in maximum wind speed. Comparison with observations (Fig. 19) shows that the wind field of Lorenzo was correctly predicted to be broad by HAFS-globalnest, but the peak wind was too strong. HAFS-globalnest had a strong inner wind maximum around $r = 30$ n mi, while the broader wind maximum in observations was around $r = 50$ n mi. The degree to which this was a function of HAFS-globalnest missing inner-core

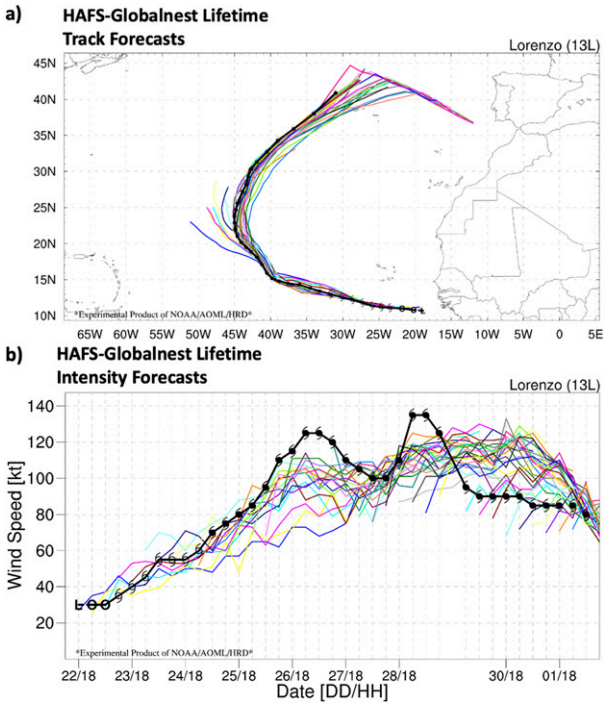

FIG. 18. (a) All HAFS-globalnest track forecasts for Lorenzo. The black line represents best track. (b) All HAFS-globalnest intensity forecast for Lorenzo.

processes (like an eyewall replacement cycle) versus other processes (like ocean cooling) is something that would be interesting to explore in further work, ideally with a fully coupled version of HAFS-globalnest.

4. Conclusions

Real-time forecast results from HAFS-globalnest from the 2019 Atlantic hurricane season show promising forecast skill in track and intensity. The track forecasts from HAFS-globalnest were better than most operational guidance, including GFS and HWRF, out to day 5. At longer lead times (day 7 especially), however, the operational GFS performed slightly better. The two configurations of HAFS (HAFS-globalnest and HAFS-SAR) were similar overall, but for some of the cases near the nest edges, HAFS-globalnest, with two-way feedback, showed a better track forecast. For intensity, HAFS-globalnest was slightly worse than HWRF and HMON at early leads, but performed similarly at longer leads. The intensity error was smaller than the operational GFS at all lead times, and the significant negative intensity bias of the operational GFS was reduced in HAFS-globalnest. The value of the high-resolution nest for intensity prediction is obvious as others also suggested. For TC structure, there was generally little bias for 64- and 50-kt winds, but the radius of 34-kt winds had a significant high bias. Further inspection and investigation of this issue, and possible solutions (including dynamics and surface/PBL physics), is a subject of ongoing efforts. Rapid intensification results show that HAFS-globalnest, like other hurricane models, struggles to predict RI.

A closer look at a few of the more high-profile and/or challenging Atlantic cases from 2019 illustrates both the strengths of HAFS-globalnest as well as opportunities for further improvement. Hurricane Humberto's track was generally well-predicted by HAFS-globalnest after the model was able to pick up on the storm staying east of Florida. Earlier forecasts over land suffered from a low bias in intensity. The model correctly forecast some unique structural features of Humberto, including a "sting jet" leading to strong winds intrusion as the storm underwent ET near Bermuda. Also, Humberto was one of the cases near the nest edge of HAFS-SAR where HAFS-globalnest had better track forecasts than HAFS-SAR, showing less of a turn back toward the northwest. Further detailed examination of this case could be useful for understanding the differences in different nesting approaches. Hurricane Lorenzo demonstrated the capability of HAFS-globalnest to correctly predict intensification in some scenarios, although it did not capture some temporal evolution, such as the secondary peak in intensity. The track forecasts for Lorenzo were also able to capture the correct location and speed of the turn to the north over the Central Atlantic. As a whole, the case studies show that HAFS-globalnest shows promising forecast skill of TC track and the large-scale environment, and is also capable of predicting small-scale structure and intensity changes, although more work is needed to improve the consistency of these forecasts.

5. Future work

HAFS development is an ongoing effort under collaboration among AOML, NCEP/EMC, GFDL, and other NOAA organizations and universities. Several upgrades are planned to HAFS-globalnest in the near future, many of which are already ongoing. The bias in R34 (with the storms being too large in HAFS-globalnest) motivates more careful examination of several aspects of the model, including boundary layer and surface physics, to see their role in this persistent bias. The size bias will also be examined above the surface through different metrics, to determine how robust this bias is. Also, as was noted in the case study of Hurricane Dorian, a skillful data assimilation system to allow for accurate initialization of the TC inner core region (without any artificial procedures like vortex bogusing) is under development, which will lead to improvements in both structure and intensity prediction. For both Dorian and Lorenzo, HAFS-globalnest was able to capture some intensification, but missed the category-5 intensity peak. Further work is needed to improve resolution and model physics to improve storm structure, resolve extreme winds, and improve predictability of rapid intensification. Also, as different grid and nesting strategies are explored in HAFS-globalnest, the importance of two-way feedback for correctly predicting storm-to-storm and storm-environment interactions, both of which can affect track forecasts, will be an important topic for future study. For storms like Dorian, this will allow for the exploration of questions such as what is the degree to which there was a feedback between the track, structure, and intensity (i.e., which caused which?), and how can this be probed further using high-resolution HAFS experiments?

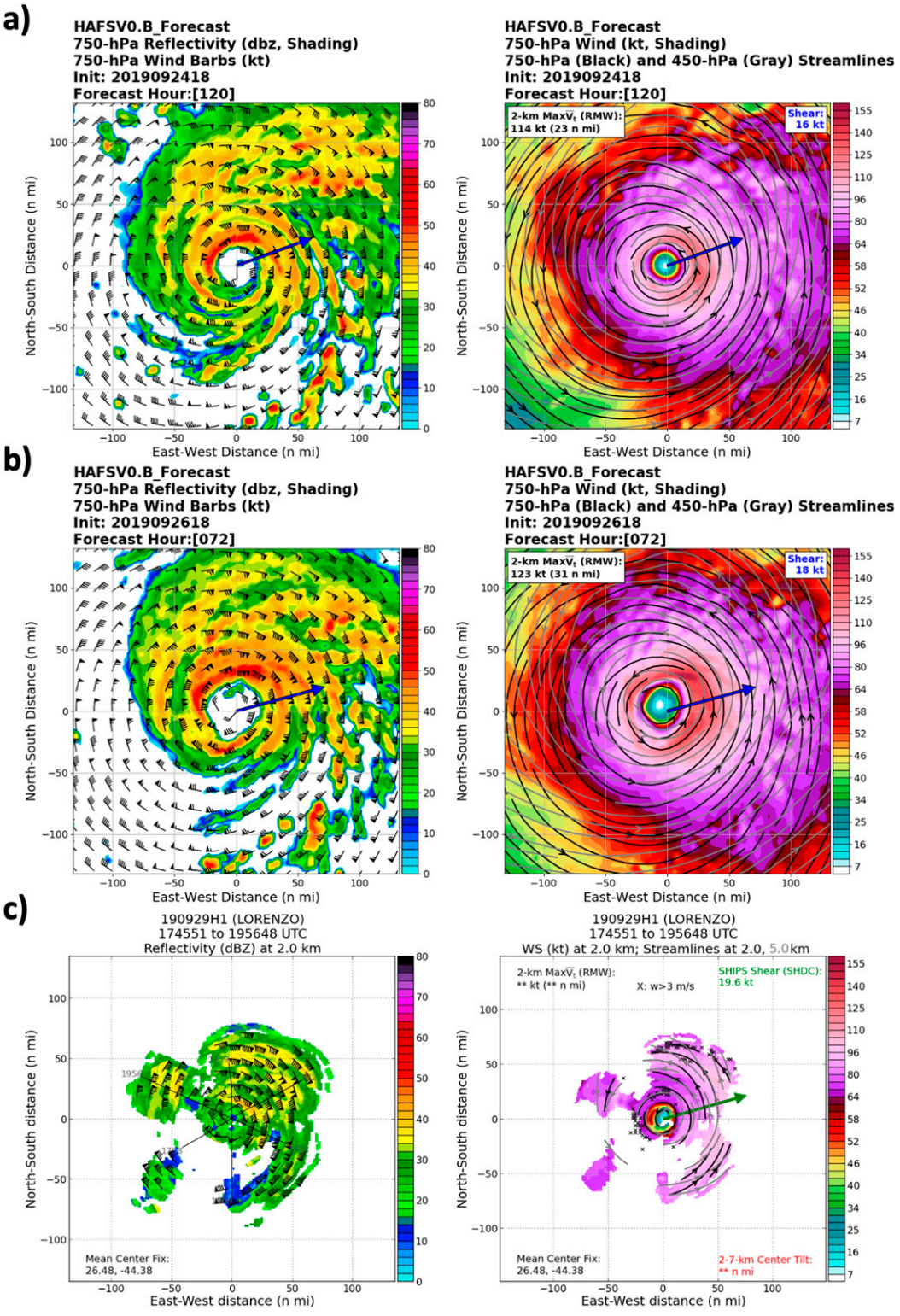

FIG. 19. (a) (left) The 750-hPa reflectivity (shaded) and wind (kt) from the HAFS-globalnest forecast initialized at 1800 UTC 24 Sep 2019, 120 h. (right) 750-hPa wind (shaded; kt), with 750-hPa streamlines (black) and 450-hPa streamlines (gray) overlaid. The shear vector is shown in blue in both images. (b) As in (a), but for the forecast initialized at 1800 UTC 26 Sep 2019, 72 h. (c) (left) NOAA P-3 observed 2-km reflectivity (shaded) and wind (kt). (right) NOAA P-3 observed 2-km wind (shaded; kt), with 2-km streamlines (black) and 5-km streamlines (gray) overlaid. The shear vector is shown in green on the right.

In addition, the two-way feedback will be evaluated through its impact on large-scale verification metrics such as the 500-hPa anomaly correlation. This is a subject of ongoing investigation. Finally, since HAFS-globalnest can be run regardless of whether there are active TCs or areas of interest, a detailed exploration of the skill of the model in TC genesis forecasts will be another area of study.

Acknowledgments. The authors thank the NOAA Research and Development High Performance Computing Systems (RDHPCS) staff for support with the real-time reservation used to run these forecasts on the NOAA Jet computer system. In addition, the authors thank the crews of the NOAA P3 aircraft for collecting the data used for the observational comparisons in this study. The Hurricane Intensity Forecast Experiment (IFEX; Rogers et al. 2013b) is supported by HFIP and conducted by NOAA/AOML/HRD scientists. Levi Cowan, Xiaomin Chen, Mike Fiorino, and three anonymous reviewers provided helpful comments that improved the manuscript. The lead author was supported by NOAA Hurricane Supplemental Grant NA19OAR0220187.

Data availability statement. The complete HAFS-globalnest data from 2019 are available on the NOAA RDHPCS computing system archive. The track files, as well as individual cases, are also available upon request.

REFERENCES

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

——, 2011: The impact of dropwindsonde data from the THORPEX Pacific Area Regional Campaign and the NOAA hurricane field program on tropical cyclone forecasts in the Global Forecast System. *Mon. Wea. Rev.*, **139**, 2689–2703, https://doi.org/10.1175/2011MWR3634.1.

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.

Archambault, H. M., L. F. Bosart, D. Keyser, and J. M. Cordeira, 2013: A climatological analysis of the extratropical flow response to recurving western North Pacific tropical cyclones. *Mon. Wea. Rev.*, **141**, 2325–2346, https://doi.org/10.1175/MWR-D-12-00257.1.

Bender, M. A., I. Ginis, R. Tuleya, B. Thomas, and T. Marchok, 2007: The operational GFDL coupled hurricane–ocean prediction system and a summary of its performance. *Mon. Wea. Rev.*, **135**, 3965–3989, https://doi.org/10.1175/2007MWR2032.1.

Browning, K. A., and M. Field, 2004: Evidence from Meteosat imagery of the interaction of sting jets with the boundary layer. *Meteor. Appl.*, **11**, 277–289, https://doi.org/10.1017/S1350482704001379.

Cangialosi, J. P., and C. W. Landsea, 2016: An examination of model and official National Hurricane Center tropical cyclone size forecasts. *Wea. Forecasting*, **31**, 1293–1300, https://doi.org/10.1175/WAF-D-15-0158.1.

Chen, J.-H., and S.-J. Lin, 2013: Seasonal predictions of tropical cyclones using a 25-km-resolution general circulation model. *J. Climate*, **26**, 380–398, https://doi.org/10.1175/JCLI-D-12-00061.1.

——, and Coauthors, 2019: Advancements in hurricane prediction with NOAA's next-generation forecast system. *Geophys. Res. Lett.*, **46**, 4495–4501, https://doi.org/10.1029/2019GL082410.

DeHart, J. C., R. A. Houze, and R. F. Rogers, 2014: Quadrant distribution of tropical cyclone inner-core kinematics in relation to environmental shear. *J. Atmos. Sci.*, **71**, 2713–2732, https://doi.org/10.1175/JAS-D-13-0298.1.

Dong, J., and Coauthors, 2020: The evaluation of real-time Hurricane Analysis and Forecast System (HAFS) Stand-Alone Regional (SAR) model performance for the 2019 Atlantic hurricane season. *Atmosphere*, **11**, 617, https://doi.org/10.3390/atmos11060617.

Emanuel, K., and F. Zhang, 2016: On the predictability and error sources of tropical cyclone intensity forecasts. *J. Atmos. Sci.*, **73**, 3739–3747, https://doi.org/10.1175/JAS-D-16-0100.1.

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.

Gamache, J. F., J. S. Griffin Jr., P. P. Dodge, and N. F. Griffin, 2004: Automatic Doppler analysis of three-dimensional wind fields in hurricane eyewalls. *26th Conf. on Hurricanes and Tropical Meteorology*, Miami, FL, Amer. Meteor. Soc., 5D.4, http://ams.confex.com/ams/pdfpapers/75806.pdf.

Han, J., M. Witek, J. Teixeira, R. Sun, H.-L. Pan, J. K. Fletcher, and C. S. Bretherton, 2016: Implementation in the NCEP GFS of a hybrid eddy-diffusivity mass-flux (EDMF) boundary layer parameterization with dissipative heating and modified stable boundary layer mixing. *Wea. Forecasting*, **31**, 341–352, https://doi.org/10.1175/WAF-D-15-0053.1.

——, W. Wang, Y. C. Kwon, S. Hong, V. Tallapragada, and F. Yang, 2017: Updates in the NCEP GFS cumulus convection schemes with scale and aerosol awareness. *Wea. Forecasting*, **32**, 2005–2017, https://doi.org/10.1175/WAF-D-17-0046.1.

Harris, L. M., and S.-J. Lin, 2013: A two-way nested global-regional dynamical core on the cubed-sphere grid. *Mon. Wea. Rev.*, **141**, 283–306, https://doi.org/10.1175/MWR-D-11-00201.1.

Hazelton, A. T., L. Harris, and S.-J. Lin, 2018a: Evaluation of tropical cyclone forecasts in a high-resolution version of the multiscale GFDL fvGFS model. *Wea. Forecasting*, **33**, 419–442, https://doi.org/10.1175/WAF-D-17-0140.1.

——, M. Bender, M. Morin, L. Harris, and S.-J. Lin, 2018b: 2017 Atlantic hurricane forecasts from a high-resolution version of the GFDL fvGFS model: Evaluation of track, intensity, and structure. *Wea. Forecasting*, **33**, 1317–1337, https://doi.org/10.1175/WAF-D-18-0056.1.

——, X. Zhang, W. Ramstrom, S. Gopalakrishan, F. D. Marks, and J. A. Zhang, 2020: High-resolution ensemble HFV3 forecasts of Hurricane Michael (2018): Rapid intensification in shear. *Mon. Wea. Rev.*, **148**, 2009–2032, https://doi.org/10.1175/MWR-D-19-0275.1.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, **113**, D13103, https://doi.org/10.1029/2008JD009944.

Kaplan, J., and M. DeMaria, 2003: Large-scale characteristics of rapidly intensifying tropical cyclones in the North Atlantic basin. *Wea. Forecasting*, **18**, 1093–1108, https://doi.org/10.1175/1520-0434(2003)018<1093:LCORIT>2.0.CO;2.

——, ——, and J. A. Knaff, 2010: A revised tropical cyclone rapid intensification index for the Atlantic and eastern North Pacific basins. *Wea. Forecasting*, **25**, 220–241, https://doi.org/10.1175/2009WAF2222280.1.

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.

Lin, S.-J., 1997: A finite-volume integration method for computing pressure gradient force in general vertical coordinates. *Quart. J. Roy. Meteor. Soc.*, **123**, 1749–1762, https://doi.org/10.1002/qj.49712354214.

——, and R. B. Rood, 1996: Multidimensional flux-form semi-Lagrangian transport schemes. *Mon. Wea. Rev.*, **124**, 2046–2070, https://doi.org/10.1175/1520-0493(1996)124<2046:MFFSLT>2.0.CO;2.

——, and ——, 1997: An explicit flux-form semi-Lagrangian shallow-water model on the sphere. *Quart. J. Roy. Meteor. Soc.*, **123**, 2477–2498, https://doi.org/10.1002/qj.49712354416.

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, https://ams.confex.com/ams/25HURR/techprogram/paper_37628.htm.

National Hurricane Center, 2019a: Tropical Storm Dorian Discussion Number 14. Accessed 24 February 2020, https://www.nhc.noaa.gov/archive/2019/al05/al052019.discus.014.shtml.

——, 2019b: Hurricane Humberto Discussion Number 27. Accessed 25 February 2020, https://www.nhc.noaa.gov/archive/2019/al13/al132019.discus.027.shtml.

NOAA, 2020: Unified Forecast System. Accessed 2 June 2020, https://vlab.ncep.noaa.gov/web/environmental-modeling-center/unified-forecast-system.

Putman, W. M., and S.-J. Lin, 2007: Finite-volume transport on various cubed-sphere grids. *J. Comput. Phys.*, **227**, 55–78, https://doi.org/10.1016/j.jcp.2007.07.022.

Reasor, P. D., R. Rogers, and S. Lorsolo, 2013: Environmental flow impacts on tropical cyclone structure diagnosed from airborne Doppler radar composites. *Mon. Wea. Rev.*, **141**, 2949–2969, https://doi.org/10.1175/MWR-D-12-00334.1.

Roebber, P. J., 2009: Visualizing multiple measures of forecast quality. *Wea. Forecasting*, **24**, 601–608, https://doi.org/10.1175/2008WAF2222159.1.

Rogers, R., and Coauthors, 2013a: NOAA's Hurricane Intensity Forecasting Experiment: A progress report. *Bull. Amer. Meteor. Soc.*, **94**, 859–882, https://doi.org/10.1175/BAMS-D-12-00089.1.

——, P. Reasor, and S. Lorsolo, 2013b: Airborne Doppler observations of the inner-core structural differences between intensifying and steady-state tropical cyclones. *Mon. Wea. Rev.*, **141**, 2970–2991, https://doi.org/10.1175/MWR-D-12-00357.1.

Stewart, S., 2020: Tropical cyclone report: Hurricane Humberto (13–19 September 2019). NHC Tech. Rep. AL092019, 44 pp., https://www.nhc.noaa.gov/data/tcr/AL092019_Humberto.pdf.

Wang, W., J. A. Sippel, S. Abarca, L. Zhu, B. Liu, Z. Zhang, A. Mehra, and V. Tallapragada, 2018: Improving NCEP HWRF simulations of surface wind and inflow angle in the eyewall area. *Wea. Forecasting*, **33**, 887–898, https://doi.org/10.1175/WAF-D-17-0115.1.

Zhang, J. A., and R. F. Rogers, 2019: Effects of parameterized boundary layer structure on hurricane rapid intensification in shear. *Mon. Wea. Rev.*, **147**, 853–871, https://doi.org/10.1175/MWR-D-18-0010.1.