

2017 Atlantic Hurricane Forecasts from a High-Resolution Version of the GFDL fvGFS Model: Evaluation of Track, Intensity, and Structure

ANDREW T. HAZELTON^a AND MORRIS BENDER

Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, New Jersey

MATTHEW MORIN

University Corporation for Atmospheric Research, Boulder, Colorado

LUCAS HARRIS AND SHIAN-JIANN LIN

NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

(Manuscript received 5 April 2018, in final form 21 August 2018)

ABSTRACT

The 2017 Atlantic hurricane season had several high-impact tropical cyclones (TCs), including multiple cases of rapid intensification (RI). A high-resolution nested version of the GFDL finite-volume dynamical core (FV3) with GFS physics (fvGFS) model (HifvGFS) was used to conduct hindcasts of all Atlantic TCs between 7 August and 15 October. HifvGFS showed promising track forecast performance, with similar error patterns and skill compared to the operational GFS and HWRf models. Some of the larger track forecast errors were associated with the erratic tracks of TCs Jose and Lee. A case study of Hurricane Maria found that although the track forecasts were generally skillful, a right-of-track bias was noted in some cases associated with initialization and prediction of ridging north of the storm. The intensity forecasts showed large improvement over the GFS and global fvGFS models but were somewhat less skillful than HWRf. The largest negative intensity forecast errors were associated with the RI of TCs Irma, Lee, and Maria, while the largest positive errors were found with recurring cases that were generally weakening. The structure forecasts were also compared with observations, and HifvGFS was found to generally have wind radii larger than the observations. Detailed examination of the forecasts of Hurricanes Harvey and Maria showed that HifvGFS was able to predict the structural evolution leading to RI in some cases but was not as skillful with other RI cases. One case study of Maria suggested that the inclusion of ocean coupling could significantly reduce the positive bias seen during and after recurvature.

1. Introduction

Both track and intensity forecasts of tropical cyclones (TCs) have improved over the last 5–10 years, especially at longer lead times (e.g., Cangialosi and Franklin 2017). However, forecasts of rapid intensification (RI; 30 kt or more of intensification in 24 h, where 1 kt $\approx 0.51 \text{ m s}^{-1}$; e.g., Kaplan et al. 2010) continue to prove difficult for both

numerical weather models and operational forecasters, and improving model performance in these cases is a critical component of TC research. In addition, despite the improvements in track forecasts, there are still cases where models struggle with the evolution of steering and TC motion even in the short term, such as Hurricane Joaquin in 2015 (Berg 2016).

Several recent studies have examined the importance of high-resolution modeling to forecast RI, as well as other ways to improve RI forecasts. Davis et al. (2008) found that RI forecasts in the Advanced Research version of the Weather Research and Forecasting (WRF-ARW) Model were improved as the model resolution increased from 4 to 1 km. Bender et al. (2017) examined the impact of including objective wind radii estimates in the initialization of the GFDL hurricane model and

^a Current affiliation: Cooperative Institute for Marine and Atmospheric Studies, University of Miami, and NOAA/Atlantic Oceanographic and Meteorological Laboratory/Hurricane Research Division, Miami, Florida.

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

found that this technique led to forecast improvements of RI, especially at leads of 1–2 days. The current standard for high-resolution nested hurricane models is the Hurricane Weather Research and Forecasting (HWRF) Model. Multiple studies have examined this model and ways to improve its forecasts. [Gopalakrishnan et al. \(2013\)](#) examined the impact of changes to vertical diffusion on TC intensity and boundary layer structure in HWRF and found that reducing the diffusion in the model produced a shallower PBL that was more consistent with aircraft observations, in addition to better intensity forecasts. Similarly, [Zhang et al. \(2017\)](#) examined HWRF and found that improved PBL structure led to better RI forecasts. [Alaka et al. \(2017\)](#) discussed an experimental basin-scale version of HWRF and found that it had improved track forecast skill relative to the operational HWRF when multiple TCs were present, due to its ability to use multiple high-resolution moving nests.

In addition to track and intensity verifications, recent studies have begun to examine forecasts of storm structure. [Tallapragada et al. \(2014\)](#) showed that the 2012 upgraded HWRF had improved wind radii forecasts and a reduced high-intensity bias compared to older versions of the model. [Cangialosi and Landsea \(2016\)](#) performed a verification of 34-, 50-, and 64-kt wind radii forecasts from the National Hurricane Center (NHC) and dynamical models and found that the official NHC forecasts tended to be skillful compared to climatology, whereas the dynamical model forecasts were not. This study noted, however, that the magnitudes of the radii forecast errors were of similar magnitude to the observational uncertainty in these parameters.

In this study, we aim to further the goals of improved nested model performance in TC forecasting, particularly in the difficult outlier cases (such as RI). This is done through evaluation of the finite-volume dynamical core (FV3) with Global Forecast System (GFS) physics (fvGFS) model (HifvGFS), a relatively new high-resolution nested modeling system in development at the National Oceanic and Atmospheric Administration's (NOAA) Geophysical Fluid Dynamics Laboratory (GFDL). [Hazelton et al. \(2018\)](#) demonstrated the capability of this model for high-resolution TC intensity and structure forecasting in a small subset of cases from 2010 to 2016. Herein, the model is evaluated across a large set of cases from the 2017 Atlantic hurricane season. This study also analyzes model TC structure in 2order to illustrate the skill of the forecasts beyond the typical track and intensity metrics. The analyses performed here will help motivate further improvements to the model, with the goal of significantly increasing the skill of the model for both TC forecasting and research into physical processes.

2. Model setup and cases analyzed

a. Model configuration

This study examines the forecasts of a high-resolution nested version (nest configuration described below) of the fvGFS model. This model used the finite-volume cubed-sphere dynamical core ([Lin and Rood 1997](#); [Lin 1997, 2004](#)). The physical parameterizations are mostly similar to those used in the operational GFS run by the National Centers for Environmental Prediction (NCEP), except for a significant upgrade in the microphysics scheme to the GFDL 6-class single-moment scheme. The model is initialized using the GFS global analyses interpolated onto the fvGFS grid, as a “cold start” with no additional TC-specific data assimilation or vortex initialization. This is in contrast to the HWRF model, which uses both an advanced vortex initialization and vortex-scale data assimilation (e.g., [Biswas et al. 2017](#)).

The model uses the Rapid Radiative Transfer Model for GCMs (RRTMG; [Iacono et al. 2008](#)) radiation scheme, and the [Han and Pan \(2011\)](#) parameterization for the planetary boundary layer (PBL). Convection is parameterized using the recently updated scale-aware cumulus parameterization that is operational in both the GFS ([Han et al. 2017](#)) and HWRF models. The GFS microphysics (based on [Zhao and Carr 1997](#)) is replaced with a 6-class single-moment microphysics scheme ([Chen and Lin 2013](#)) developed at GFDL. This scheme is similar to the 6-class scheme of [Lin et al. \(1983\)](#). [Chen et al. \(2018, manuscript submitted to *Mon. Wea. Rev.*\)](#) and [Bender \(2017\)](#) found that this microphysics scheme improved TC forecasts of track, intensity, and genesis in a 13-km global version of the model. For most of the forecasts, no ocean coupling was included (like the GFS but in contrast to HWRF, which is coupled to an ocean model), although a test case with a developmental one-dimensional ocean model is examined later.

The fvGFS is a global model that can be regionally refined through grid stretching and nesting. [Hazelton et al. \(2018\)](#) demonstrated the capability of a high-resolution version of fvGFS in forecasts of TC intensity and structure. For this study, a nested version of fvGFS was used in which the global grid was C768 (~13-km resolution), and a factor-of-4 static nest ([Harris and Lin 2013](#)) was placed over the North Atlantic to reach a resolution of just over 3 km. This is higher resolution than the 13-km GFS, but lower resolution than the 2-km HWRF. Both the global and nested domains used 63 vertical levels. [Figure 1](#) shows the area covered by the Atlantic nest: from Africa to the western Gulf of Mexico at 3-km resolution. For the remainder of the study, the high-resolution nested fvGFS will be referred to as HifvGFS.

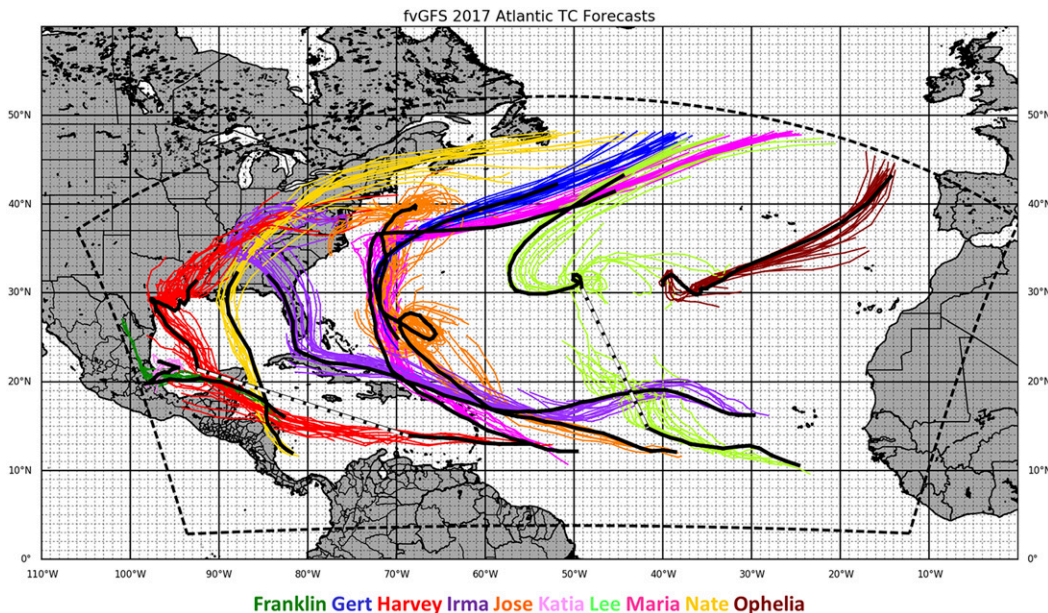


FIG. 1. Tracks of all forecasts analyzed in this study, color coded by storm (see legend). The best tracks are plotted in black. Where Harvey and Lee were temporarily not classified as tropical cyclones, the best tracks are indicated with dashed segments. The dashed black line around the edges of the figure outlines the HifvGFS nested 3-km domain.

b. TCs analyzed

This analysis covers the bulk of the 2017 hurricane season, including most cases from August through mid-October. The model was run four times daily from 7 August through 15 October (0000, 0600, 1200, and 1800 UTC initializations of GFS), for a total of 280 forecasts. All forecasts were run for 132 h. The TCs forecasted by the model were tracked for comparison with observations. Figure 1 shows all of the TCs in this study. The black lines are the “best track” (e.g., Miller et al. 1990, Sampson and Schrader 2000) observed tracks for each case, and the colors show the HifvGFS forecasts for each case (see legend). Table 1 lists all the TCs included in the analysis, as well as the number of forecasts made for each storm. The August–October peak period of the 2017 Atlantic season was very active with TCs in multiple parts of the basin, providing an excellent dataset for study of the HifvGFS forecast skill.

3. Results

a. Track forecast performance

Figure 2 shows several different evaluations of track forecast errors and skill, including statistical significance based on a Student’s *t* test (Neumann et al. 1977). Figure 2a shows that the differences among the three models are very small at most forecast hours, although HWRF and HifvGFS slightly outperformed the GFS at

day 5. For the three highest-profile cases of the season, Harvey, Irma, and Maria (Fig. 2b), HifvGFS had the best track forecast at long range. Although the differences compared to HWRF were small and not statistically significant, the improvements compared with the GFS were statistically significant at days 4 and 5. Figures 2c and 2d show the skill relative to the climatology and persistence statistical models (CLIPER; Aberson 1998), which paint a very similar picture. The skill of all track forecast models peaks around days 2–3. The GFS was slightly more skillful than the other two models for Harvey, Irma, and Maria at earlier forecast hours, but HifvGFS performed better at long range. Since these results show that the track forecasts from HifvGFS were comparable to those from these skillful

TABLE 1. The 2017 Atlantic TCs included in the analysis. The right column lists the number of forecasts for each TC.

TC	No. of cases
Franklin	14
Gert	20
Harvey	37
Irma	48
Jose	69
Katia	15
Lee	47
Maria	57
Nate	16
Ophelia	24

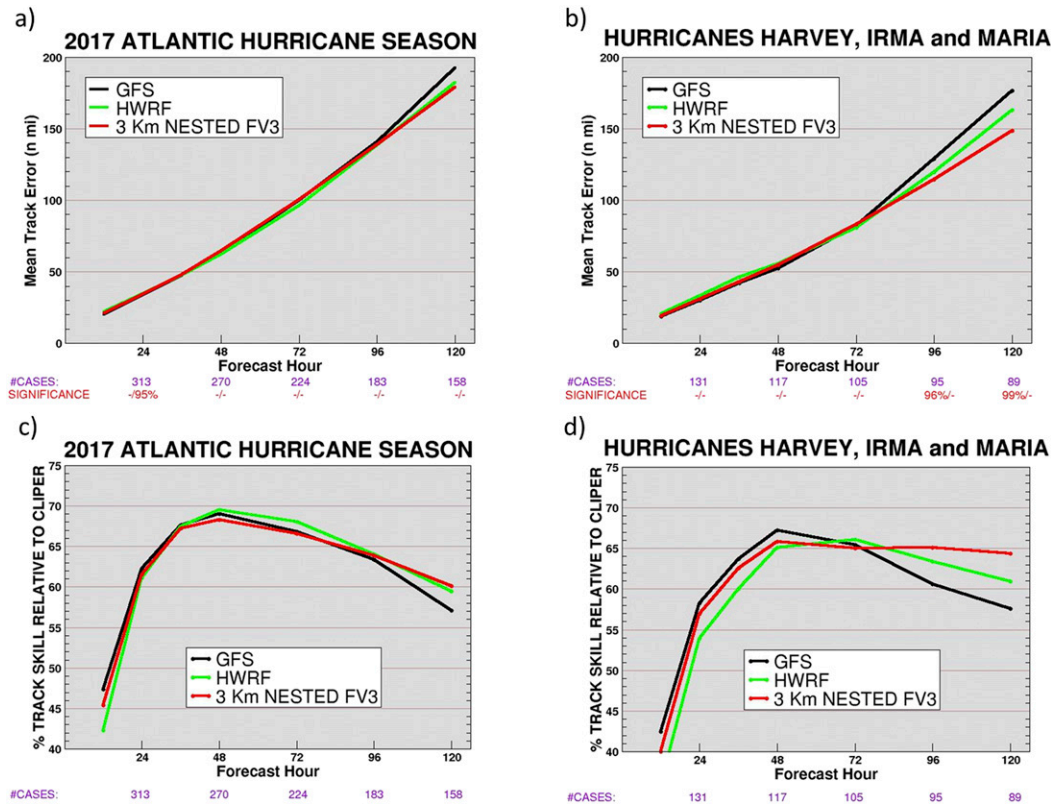


FIG. 2. (a) Mean 12–120-h track error (n mi) for the HifvGFS (red), operational GFS (black), and operational HWRF (green). The sample covers all TCs used in the study, and the number of cases at each forecast hour is shown along the bottom. The statistical significance (percentage level) of the differences between HifvGFS/GFS and HifvGFS/HWRF is also shown. (b) As in (a), but only for Hurricanes Harvey, Irma, and Maria. (c) As in (a), but for track skill (relative to CLIPER) instead of mean absolute error. (d) As in (c), but only for Hurricanes Harvey, Irma, and Maria.

operational models, they should not degrade further analysis of the TC intensity and structure evolution.

Figure 3 shows the along-track and across-track errors from HifvGFS, GFS, and HWRF. From the along-track error plot (Fig. 3a) we see that all models, including HifvGFS, have a slight slow bias that gets worse throughout the forecast. The across-track error plot (Fig. 3b) shows that HifvGFS has a fairly pronounced right-of-track bias that increases with forecast lead time. Interestingly, HifvGFS has the lowest along-track and highest across-track errors.

To look in more detail at where the largest track forecast errors were found, Fig. 4 shows maps of track forecast errors at 48, 72, and 120 h. The locations are also colored by different storms (see Fig. 1) to show which cases had the largest errors. Most of the biggest forecast errors resulted from cases relatively close to the boundary of the nest or with weak/complicated steering. At 48 h (Fig. 4a), the largest errors are mostly found for Hurricanes Gert and Maria during recurvature. There were also large errors in the forecasts for Hurricane Lee

during a period where it was making a small loop and some of the forecasts incorrectly sent the storm east (Fig. 1). For the 72-h forecasts (Fig. 4b), the Gert and Lee cases continued to show large errors. In addition, a few of the Harvey forecasts in the Gulf of Mexico and Irma cases in the eastern Atlantic showed larger errors. For the 120-h forecasts (Fig. 4c), there were several Harvey forecasts that had large errors from turning northeast too quickly. In addition, several Jose cases had large 5-day errors when the storm was making an anticyclonic loop, and several Lee forecasts had very large errors at 120 h.

To quantify the differences observed above, the mean track errors for different storm speeds and directions of motion were calculated at 48, 72, and 120 h. The results are shown in Table 2. “Fast moving” storms were defined as those with a best track forward motion greater than or equal to 10 m s^{-1} , while “slow moving” storms were defined as those with a forward motion less than 2.5 m s^{-1} . At 48 h, the storms with a southerly component of motion (mostly Jose and Lee) had significantly

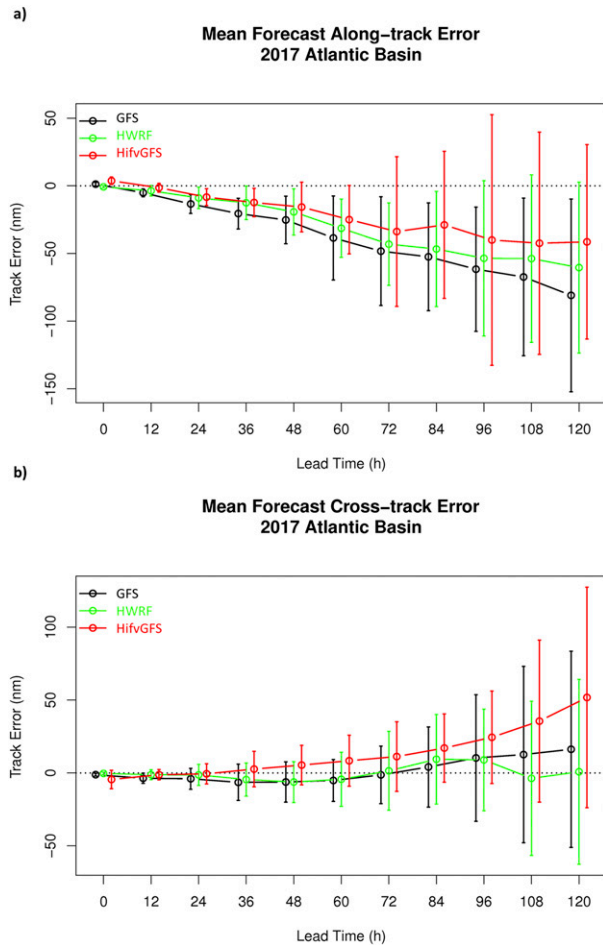


FIG. 3. (a) Along-track errors (n mi) from HifvGFS (red), HWRF (green), and GFS (black). Positive (negative) errors indicate a fast (slow) bias. The error bars show the 95% confidence interval. (b) Across-track errors (n mi) from HifvGFS (red), HWRF (green), and GFS (black). Positive (negative) errors indicate a right-of-track (left of track) bias. The error bars show the 95% confidence interval.

larger errors than those moving northeast or northwest. The slow-moving storms also had larger errors. At 72 h, the southwest/southeast-moving cases again had the largest errors (from a motion perspective), but northeast-moving cases also had larger errors than northwest, consistent with the observation of larger errors in some recurring cases. At 72 h there were no significant differences based on storm speed. At 120 h, although the northeast-moving recurring cases had the largest errors from a direction perspective, the relationship was not significant. However, the fast-moving cases did have a significantly higher error. These results imply that the largest short-term track errors were as a result of slow, often erratic motion (given the small sample size of the southeast/southwest-moving cases), while longer-term errors tended to be dominated by storm speed as TCs recurved.

b. Intensity forecast performance

Figure 5 shows the intensity forecast errors and skill for the HifvGFS forecasts from the 2017 Atlantic hurricane season. As was shown for track, the intensity errors are compared with GFS and HWRF results, for all cases (Figs. 5a,c) and also for only Harvey, Irma, and Maria (Figs. 5b, d). The intensity skill is computed relative to the Statistical Hurricane Intensity Forecast (SHIFOR) model (Knaff et al. 2003). In general, the forecasts from HifvGFS were somewhat less skillful than those from the operational HWRF, but much more skillful than the GFS forecasts. These differences are statistically significant ($p < 0.01$) at most forecast hours. The model did suffer from a spinup period of about 24 h where the skill was much lower than at later times (Fig. 5c). This is due to the initialization from the coarse GFS initial conditions. By 24 h, however, the model regains skill relative to climatology and peaks in skill around days 2–3.

To further examine the source of the intensity forecast errors, Fig. 6 shows the histograms of intensity errors for each of the three models at 48, 72 (near the peak of the intensity skill), and 120 h (a longer-range forecast). At 48 h (Fig. 6a), the HifvGFS errors are closer to a symmetric distribution about zero than HWRF, and there is a lower number of errors from -20 to -10 kt (negative bias). HifvGFS did have more forecasts with extreme negative errors (from ~ -50 to -80 kt) due to cases that failed to predict RI, although there were fewer cases than the coarser-resolution GFS. On the high-bias side, HifvGFS has a generally higher bias than GFS or HWRF partly due to the lack of ocean coupling, as it misses the generation of cold wakes by wind-induced mixing (GFS is also uncoupled but does not make storms as strong with its coarser resolution). The pattern at 72 h is similar (Fig. 6b), with HifvGFS having a larger positive bias in the 0–20-kt range and a larger number of forecasts with extreme negative errors but not nearly as many as the GFS. By 120 h (Fig. 6c), there is more overlap between the error distributions. However, HWRF still has a tendency for less negative bias, particularly in the range from -70 to -40 kt, while HifvGFS has a slightly higher frequency of positive bias cases. These figures suggest that although the model does capture some of the RI cases effectively, there are also a few significant misses, likely due in part to the lack of a vortex-scale initialization as well as a slightly lower resolution than the operational HWRF (3 vs 2 km). It should also be noted that the small sample of extreme events makes the significance of differences between the models somewhat questionable.

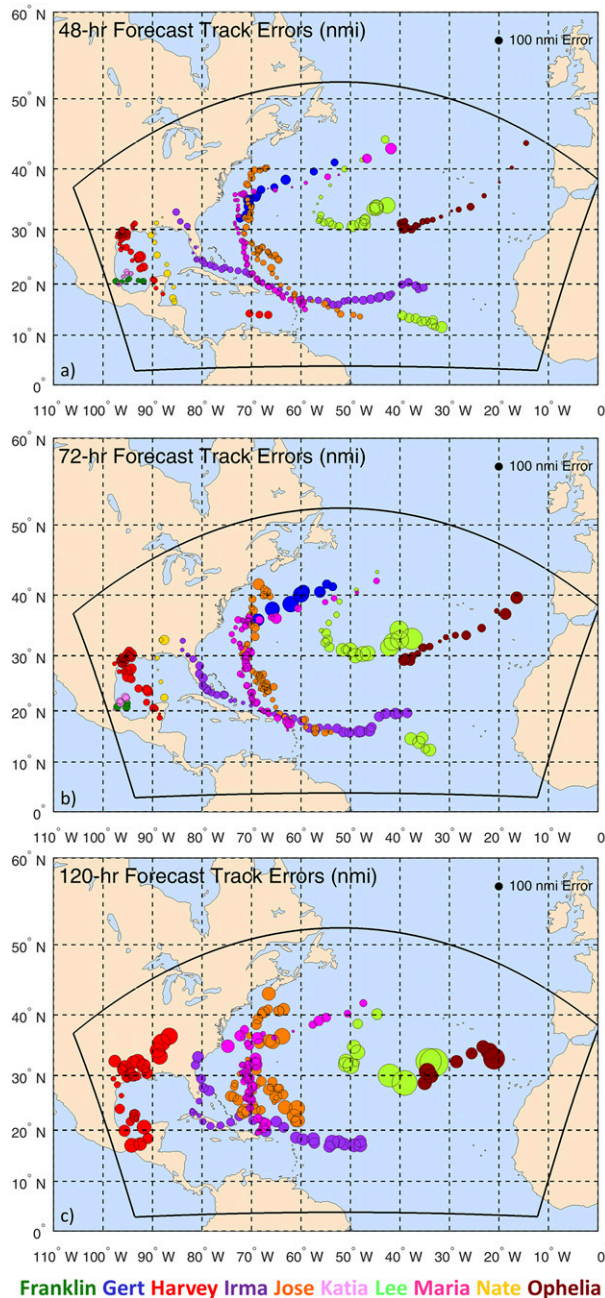


FIG. 4 The HifvGFS track errors (n mi) for (a) 48, (b), 72, and (c) 120 h. The latitude–longitude position from the forecast is plotted. The colors represent the different cases (see legend at bottom), and the diameter of the circle is proportional to the track error (see 100 n mi scale at top right). The nest domain is shown in black.

To further examine the sources of the intensity errors, the geographical distribution for the positive (Figs. 7b,d,f) and negative (Figs. 7a,c,e) errors is shown separately for each of the three forecast times discussed above (48, 72, and 120 h). At 48 h, the largest negative

TABLE 2. Mean track errors (n mi) of HifvGFS forecasts at forecast hours 48, 72, and 120, including errors separated by the direction of movement and speed. The number in parentheses is the sample size for each group. The relationships significant at the 95% level are shown in italics, and those significant at the 99% level are shown in boldface italics.

	48 h	72 h	120 h
All cases	67.7 (277)	104.6 (236)	186.9 (168)
NE moving	60.3 (75)	101.5 (73)	203.0 (55)
NW moving	61.3 (145)	82.5 (117)	177.9 (88)
SE/SW moving	94.9 (56)	165.6 (46)	183.2 (25)
Relationships	<i>SE/SW > NE</i> <i>SE/SW > NW</i>	<i>NE > NW</i> <i>SE/SW > NE</i> <i>SE/SW > NW</i>	
Slow moving	86.8 (40)	116.3 (33)	156.8 (25)
Moderate moving	64.2 (211)	100.1 (178)	188.1 (126)
Fast moving	66.4 (26)	121.3 (25)	222.9 (17)
Relationships	<i>Slow > medium</i>		<i>Fast > slow</i>

errors (Fig. 7a) are associated with Maria, Lee, and Harvey. There are also some larger errors along the track of Irma. Most of these negative errors were from times when the TCs were intensifying quickly and/or reaching peak intensity. However, there are some small errors in these deep-tropical cases as well. The positive errors at 48 h (Fig. 7b) are generally smaller in magnitude than the negative examples, and most are associated with recurring TCs. The largest positive errors are associated with Jose, which looped (see Fig. 1) over its own cold wake. The ocean cooling is not represented in HifvGFS, leading to a high bias in these cases. At 72 h (Figs. 7c,d), the picture is similar. There is a mix of small and large negative errors along the track of Irma, with a cluster of larger errors near the Antilles with Maria. There was also a negative bias along the track of Gert as it turned NE. Once again, the positive biases were generally smaller and associated with recurving cases due to the lack of ocean coupling, particularly for Jose and Maria (which moved over the cold wake generated by Jose). At 120 h (Figs. 7e,f), the large negative errors are dominated by Harvey, Irma, and Lee. The errors for Jose, Maria, and Ophelia were generally smaller. The positive errors at longer range are dominated by Jose and Maria in the subtropics and midlatitudes, with the largest errors during the loop Jose made. For example, if Jose were removed from the verifications, the mean error from HifvGFS (shown in Fig. 5a) would be 2 kt less than that from the GFS, as the Jose cases in HifvGFS had an average positive bias of 21.1 kt at hour 120. These results are consistent with the lack of ocean coupling, although other processes could also be at work.

To demonstrate the value of the 3-km nest in improving the intensity prediction, Fig. 8 compares the intensity forecast error and bias of the HifvGFS with a global fvGFS (referred to as gfvGFS) run at 13-km

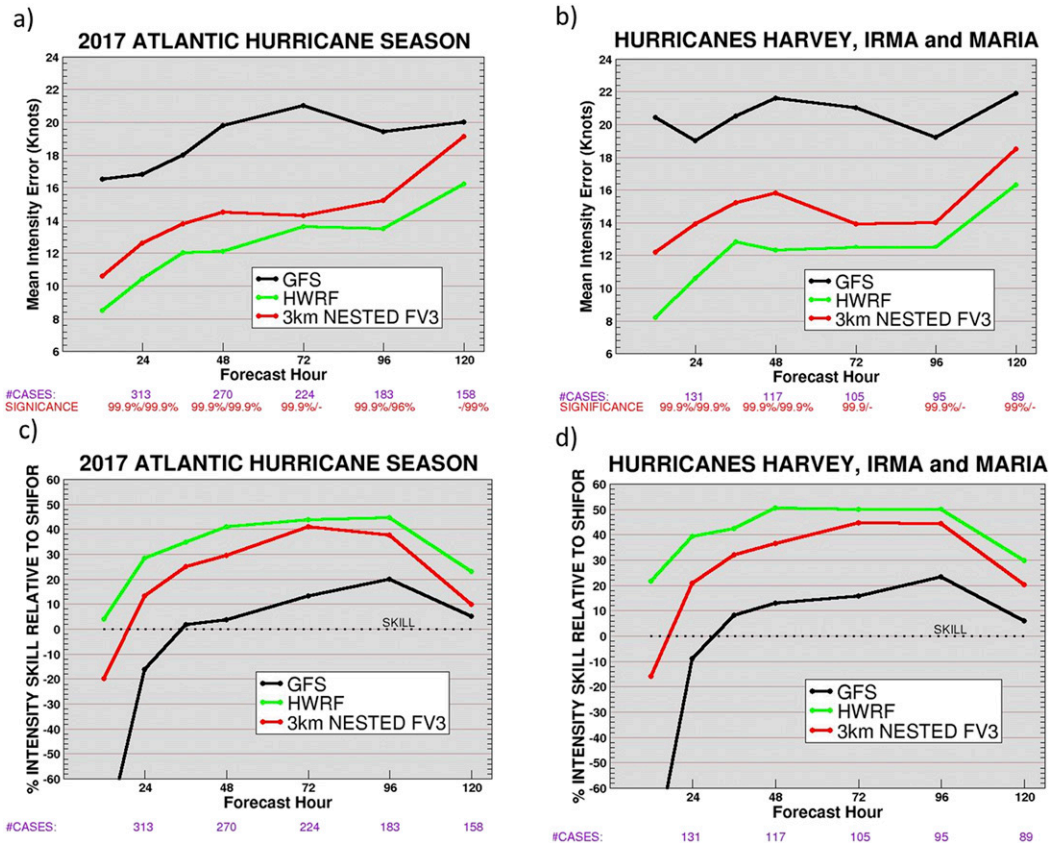


FIG. 5. (a) Mean intensity error (kt) at each lead time from 12 to 120 h for the HifvGFS (red), operational GFS (black), and operational HWRF (green). The sample covers all TCs used in the study, and the number of cases at each forecast hour is shown along the bottom. The statistical significance (percentage level) of the differences between HifvGFS/GFS and HifvGFS/HWRF is also shown. (b) As in (a), but only for Hurricanes Harvey, Irma, and Maria. (c) As in (a), but for intensity skill (relative to SHIFOR) instead of mean absolute error. (d) As in (c), but only for Hurricanes Harvey, Irma, and Maria.

resolution for the same set of TCs. This global model was only run twice daily (at 0000 and 1200 UTC), so the sample size is smaller for this comparison. Figure 8a shows that, after the initial 12–24 h of spinup, HifvGFS has lower errors than gfvGFS at all forecast hours except 120 (due to the large high biases in Jose), with the largest improvements at forecast hour 72. This is also apparent when comparing the skill relative to SHIFOR, which shows an increase of ~15% at 72 h with the nested model. The improvement is made more clear when considering the intensity bias (Fig. 8c), which is 5–6 kt smaller at all forecast hours for the nested model, indicating that the biggest increase in skill comes from reduction of the large negative bias seen in the global model. By 120 h, HifvGFS actually had a slight positive bias when including the entire sample size (0600 and 1800 UTC cases). This is potentially due to the lack of ocean coupling and the large errors in Jose (an issue that is being addressed in ongoing model upgrades). Figure 9 shows the intensity error histograms for the

HifvGFS and gfvGFS homogenous comparison. The high-resolution model has more positive errors than the global model, but fewer large negative errors, indicating better performance with rapidly intensifying TCs. The large number of HifvGFS forecasts with positive bias that resulted from a lack of ocean coupling in storms such as Jose and Maria helps to explain why the global model had slightly higher skill at day 5. These results suggest the need for the inclusion of ocean coupling in high-resolution models, particularly for storms where the intensity is strongly impacted by a large ocean response (e.g., slow-moving storms, storms moving over a shallow mixed layer, or looping storms). The impacts of ocean coupling are examined in a case study of Hurricane Maria in a later section. Nevertheless, the overall results from the nested forecasts show that, despite little impact on track skill (not shown), increased horizontal resolution can significantly improve intensity forecasts.

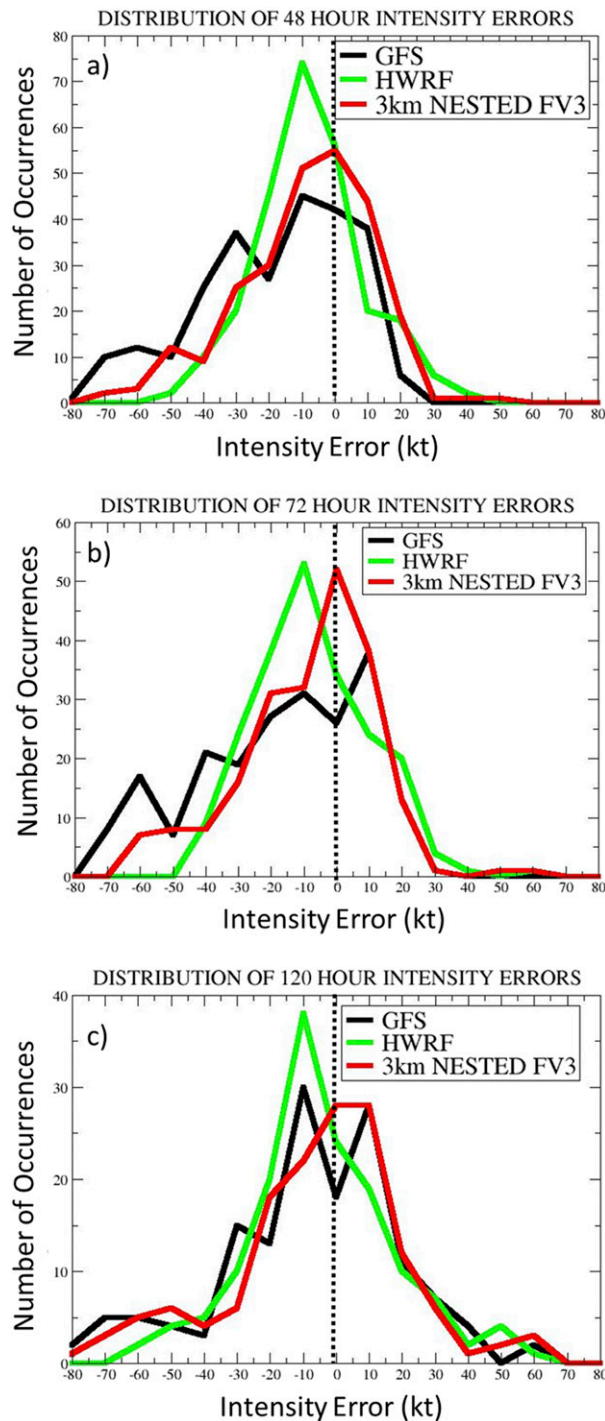


FIG. 6. Histograms of intensity forecast errors (kt) for HifvGFS (red), HWRf (green), and GFS (black) at (a) 48, (b) 72, and (c) 120 h.

c. Structure forecast performance

Hazelton et al. (2018) demonstrated the utility of the structure validation of high-resolution fvGFS TC forecasts using airborne radar data. This type of analysis

allows the model to be analyzed beyond the “single score” metrics of track and intensity error. Such a comprehensive analysis is not possible for every case in this dataset, but some validation of TC forecast structure is critical for understanding model strengths and weaknesses. Thus, the HifvGFS forecasted radii of 34-kt (R34), 50-kt (R50), and 64-kt winds (R64) are evaluated and compared with the best track wind radii data. Cangialosi and Landsea (2016) pointed out that wind radii measurements are not as accurate in cases in systems without aircraft reconnaissance. However, in order to have a sample size large enough for robust results (since these cases are only from one season), we did not remove cases without aircraft observations.

Figure 10 shows the wind radii and wind radii errors for the three different wind thresholds (34, 50, and 64 kt) for HifvGFS as well as the gfvGFS, HWRf, and GFS models. HifvGFS had the largest wind radii errors for all wind thresholds and at all forecast lead times. The larger HifvGFS errors appeared to have originated from a large bias that developed during the initial 12–24-h spinup of the model vortex. Figure 11 further examines the structure forecasts from HifvGFS, to give a better idea about the distribution of wind radii compared with that seen in observations, by comparing the histograms of 34-kt wind radii from HifvGFS and the best track for forecast hours 48, 72, and 120. The observed radii distribution is somewhat bimodal, with a primary peak around 70–100 n mi (1 n mi = 1.852 km) and a secondary peak around 175 n mi. At 48 h, HifvGFS shows some of the double peak, but the distribution is generally shifted to larger radii (i.e., the model wind radii tend to be too large). The pattern is similar at 72 h. By 120 h, the model has a single peak at a much larger radius than the observed peak. Overall, these results indicate that HifvGFS has a tendency to make the wind radii too large compared to the observed best track data with this tendency getting progressively worse throughout the forecast. Much of this error likely results from the spinup process due to the “cold start” from GFS initial conditions, particularly since the large error bias is much worse in the 3-km fvGFS compared to the 13-km fvGFS, as adjustment is made to the much higher resolution. However, other factors (such as model resolution and diffusion) also likely play a role that is currently being investigated in upgraded versions of the model physics. This will be discussed further in a later section and will be summarized in detail in future work.

d. Case studies

Next, some individual TC cases are examined, to look at the HifvGFS forecasts in more detail to examine the strengths and weaknesses of the model forecasts.

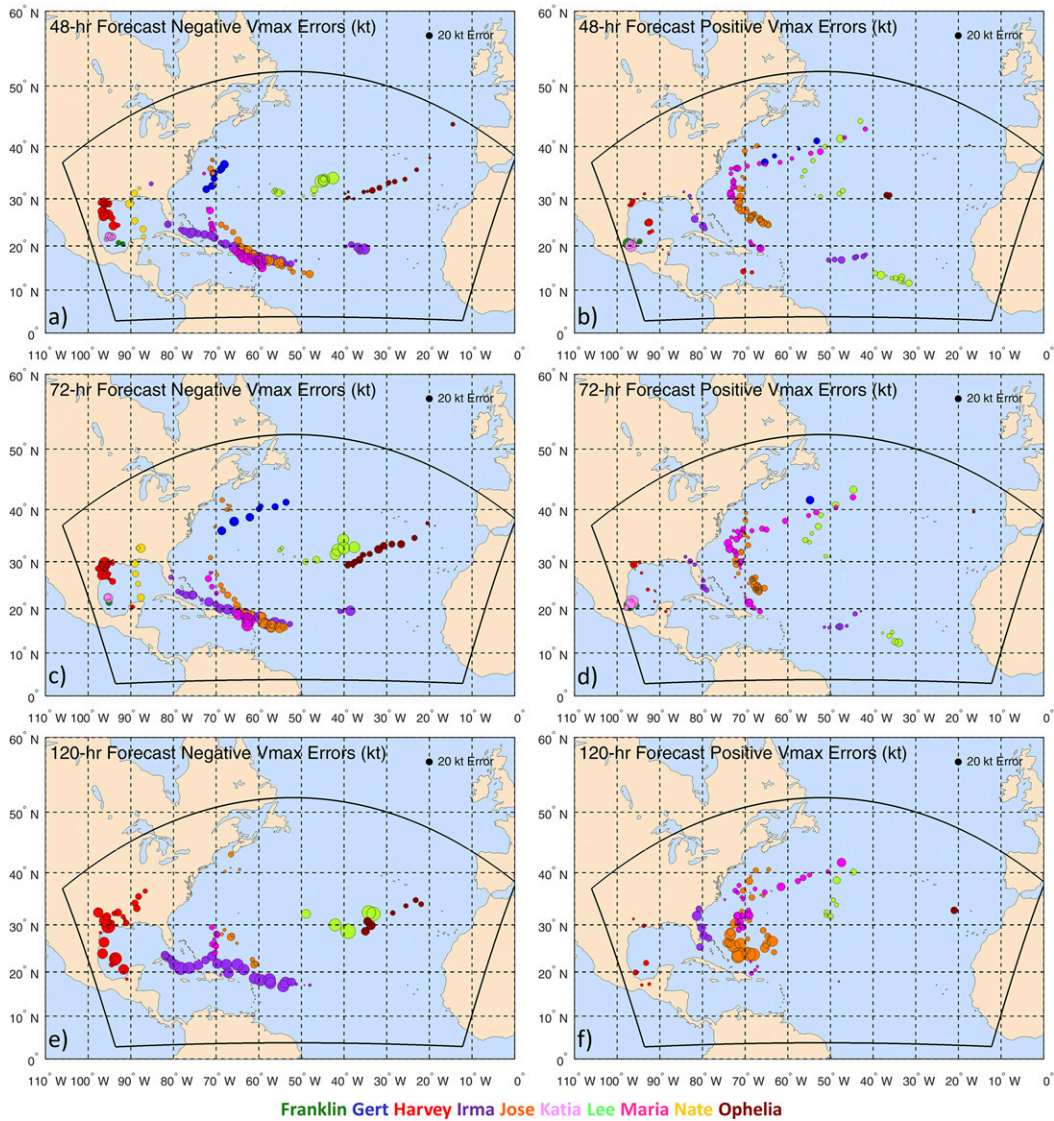


FIG. 7. Cases where HifvGFS had a negative intensity bias at (a) 48, (c) 72, and (e) 120 h. The latitude–longitude position from the forecast is plotted. The colors represent the different cases (see legend at bottom), and the diameter of the circle is proportional to the intensity error (see 20-kt scale at top right). The nest domain is shown in black. (b),(d),(f) As in (a), (c), and (e), but for cases where with a positive intensity bias.

1) HURRICANE HARVEY (GULF OF MEXICO)

The first case study examines the HifvGFS forecasts of Hurricane Harvey, specifically focusing on the forecasts after regeneration in the Gulf of Mexico, between 1200 UTC 23 August 23 and 0000 UTC 31 August. Figure 12 shows the track and intensity forecasts from each of these forecasts, as well as the observed tracks covering the entire 120-h period of these forecasts. The track forecasts were quite accurate until landfall (Fig. 12a), consistently showing the storm coming ashore near or just north of Corpus Christi,

Texas. Many of the forecasts also showed the drift along the Texas coast, although none showed the full reemergence over the Gulf and some of the later forecasts accelerated NE too quickly (these cases had some of the larger 120-h forecast errors seen in Fig. 4). Some of the HWRf forecasts (Fig. 12b) had the same issue, in addition to a few forecasts that incorrectly headed west into southern Texas. The early intensity forecasts (Fig. 12c) did not show enough deepening, but several of the forecasts initialized on 24 August (36–60 h in Fig. 12c) correctly predicted RI into a major hurricane,

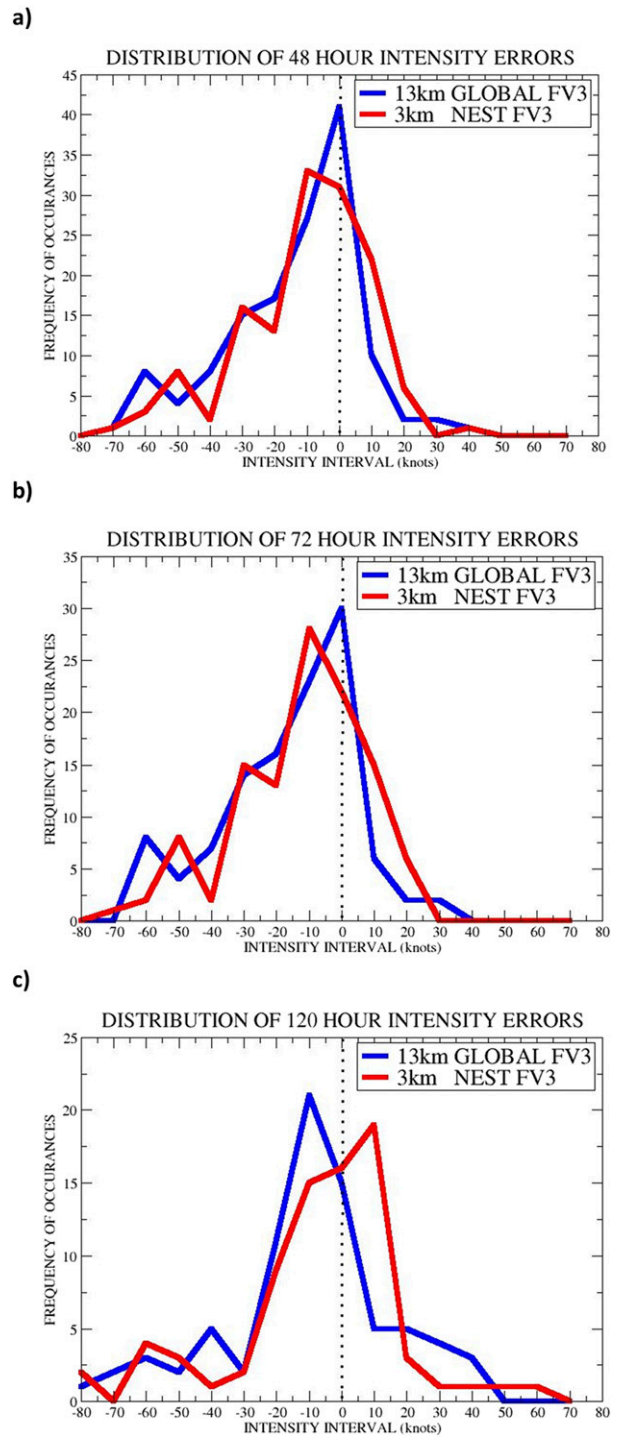
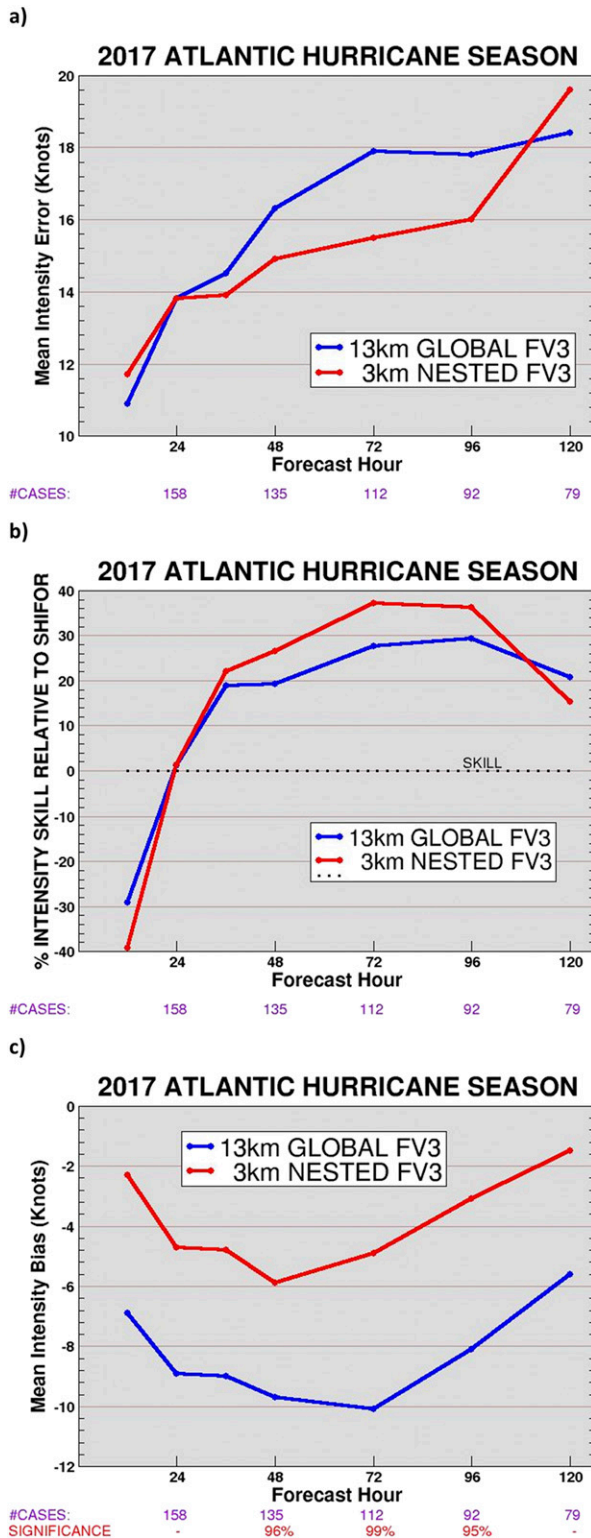


FIG. 9. Histograms of intensity forecast errors (kt) for HifvGFS (red) and gfvGFS (blue) at (a) 48, (b) 72, and (c) 120 h.

FIG. 8. (a) Mean intensity error (kt) at each forecast hour from 12 to 120 for the HifvGFS (red) and gfvGFS (blue). The number of cases at each forecast hour is shown along the bottom. The statistical significance (percentage level) of the differences between HifvGFS and gfvGFS is also shown. (b) As in (a), but for intensity skill (relative to SHIFOR) instead of mean absolute error. (c) As in (a), but for intensity bias instead of mean absolute error.

and actually outperformed some of the corresponding HWRF forecasts (Fig. 12d). However, overall HWRF did better with the intensity forecasts initialized close to major hurricane intensity, likely due to its advanced

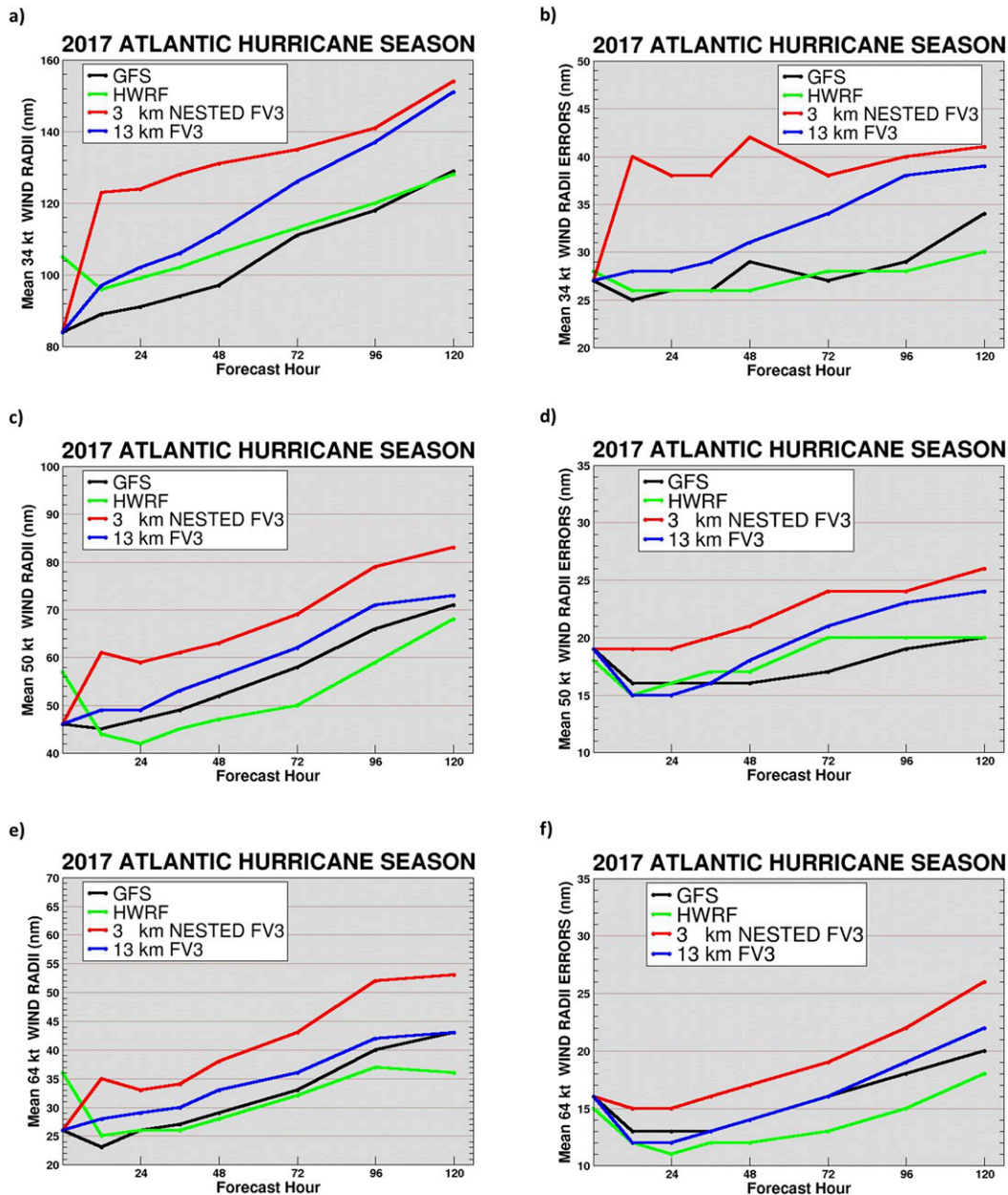


FIG. 10. Mean (a) 34-, (c) 50-, and (e) 64-kt wind radii (n mi) for HifvGFS, gfvGFS, HWRF, and GFS. (b),(d),(f) As in (a), (c), and (e), but for mean error of wind radii for HifvGFS, gfvGFS, HWRF, and GFS.

vortex-scale initialization (e.g., Gopalakrishnan et al. 2012) and higher horizontal resolution.

One of the key aspects of Hurricane Harvey’s evolution over the Gulf of Mexico was a change in structure and subsequent RI. This will be examined in the forecast initialized at 0000 UTC 24 August. At the beginning of this forecast, Harvey was a tropical depression with a very tilted vertical structure in both the observations and the model (Fig. 13a). The midlevel center was significantly displaced to the north of the low-level center. Such a tilt is

usually due to shear and is not conducive for TC development (e.g., DeMaria 1996). As the forecast continued, however, the vortex became more vertically aligned (Figs. 13b,c) and was almost perfectly aligned by hour 12 (Fig. 13d). The modeled TC rapidly intensified during this 12-h period by ~25 kt and continued to deepen (although at a slightly slower rate) after the vortex became aligned (Fig. 12c).

The reason for the vertical alignment of the vortex and subsequent commencement of the storm intensification

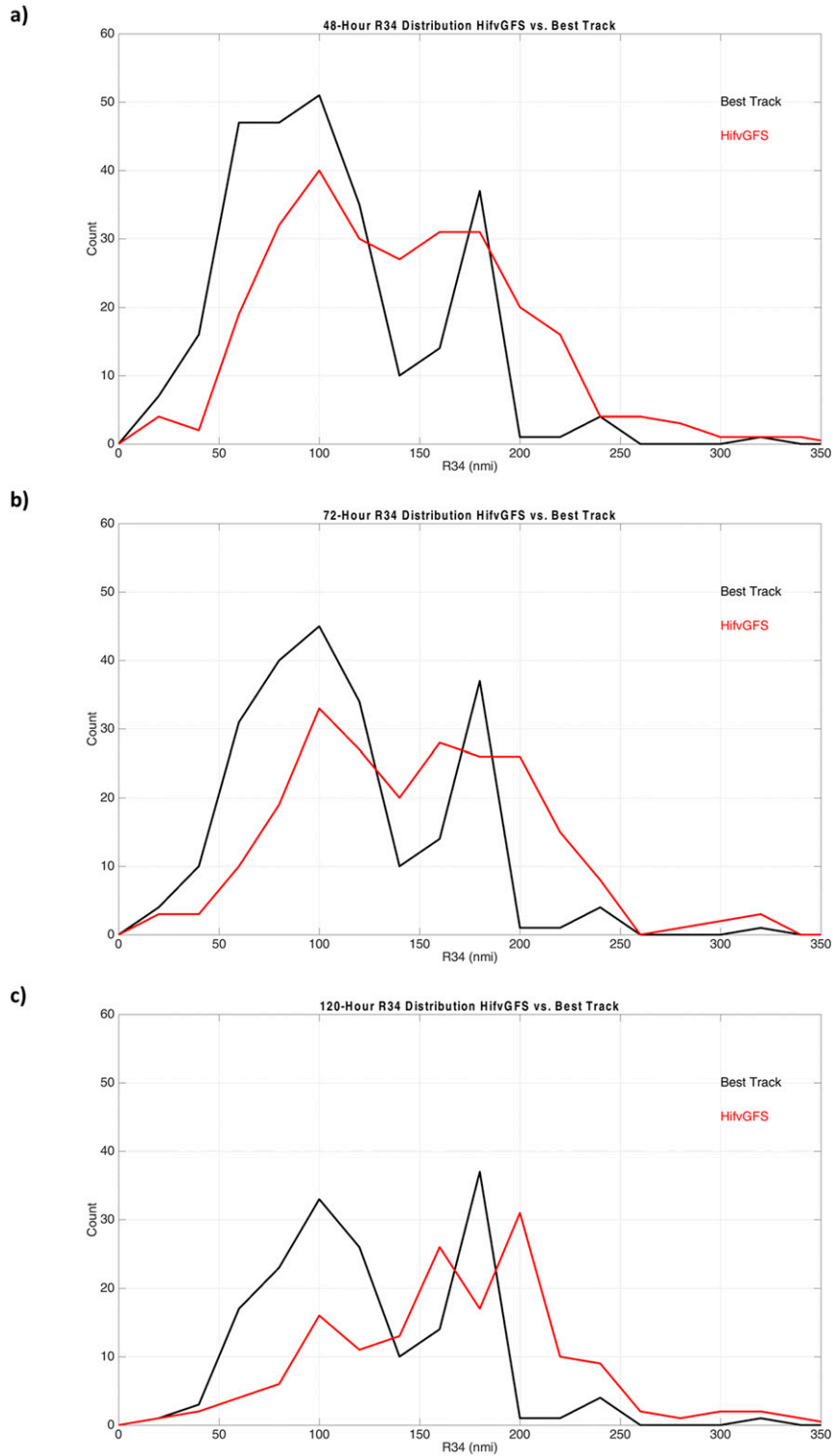


FIG. 11. Distribution of 34-kt wind radii (n mi) for HifvGFS at hour (a) 48, (b) 72, and (c) 120 and the corresponding best track distribution from the same forecast times.

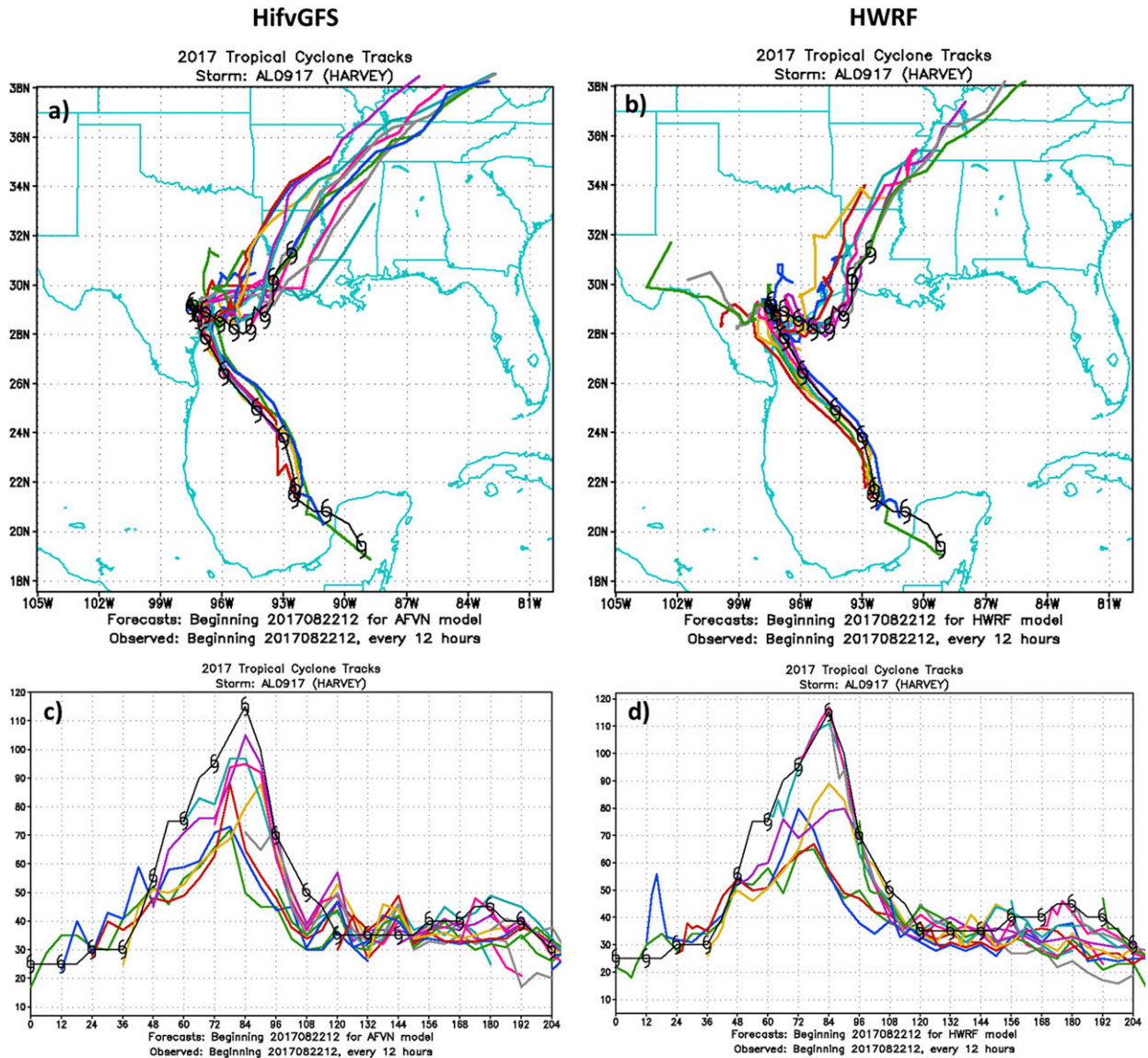


FIG. 12. (a) Composite of HifvGFS 120-h forecast tracks of Hurricane Harvey initialized between 1200 UTC 22 Aug and 0000 UTC 31 Aug. The black line with hurricane symbols indicates the best track positions covering this period. (b) As in (a), but for the HWRF Model. (c) Nested fvGFS 120-h intensity forecasts of 10-m maximum winds (kt) for Hurricane Harvey initialized between 1200 UTC 22 Aug and 0000 UTC 31 Aug. The black line with hurricane symbols indicates the best-track intensities (kt) covering this period. (d) As in (c), but for the HWRF Model.

was investigated next. Since the vertical wind shear did decrease slightly during this period and was less than 10 kt for the entire time per the SHIPS (e.g., DeMaria and Kaplan 1994) analysis, it appears that vortex-scale processes were responsible for aligning the vortex and triggering RI. Figure 13 also shows the 500-hPa vertical velocity (at contour levels of 3, 5, and 7 m s^{-1}) during the period of vortex alignment, at 3, 6, 9, and 12 h. Throughout this period, a region of convective bursts (CBs; defined here as 500-hPa vertical velocity of at least 3 m s^{-1}) was found on the north side of the system,

near the initial midlevel vortex, with vertical velocities as high as 10 m s^{-1} . These CBs helped to simultaneously align the vortex and trigger RI. The effect of the CBs on TC structure can also be seen through simulated satellite imagery (Fig. 14), which is compared with observational imagery. Initially, there was a large concentration of convection associated with the CBs along the northern part of the TC, and the convection became more axisymmetric in both the model and the observations as the TC organized. CBs have been shown to be associated with intensification and RI in other studies, both numerical

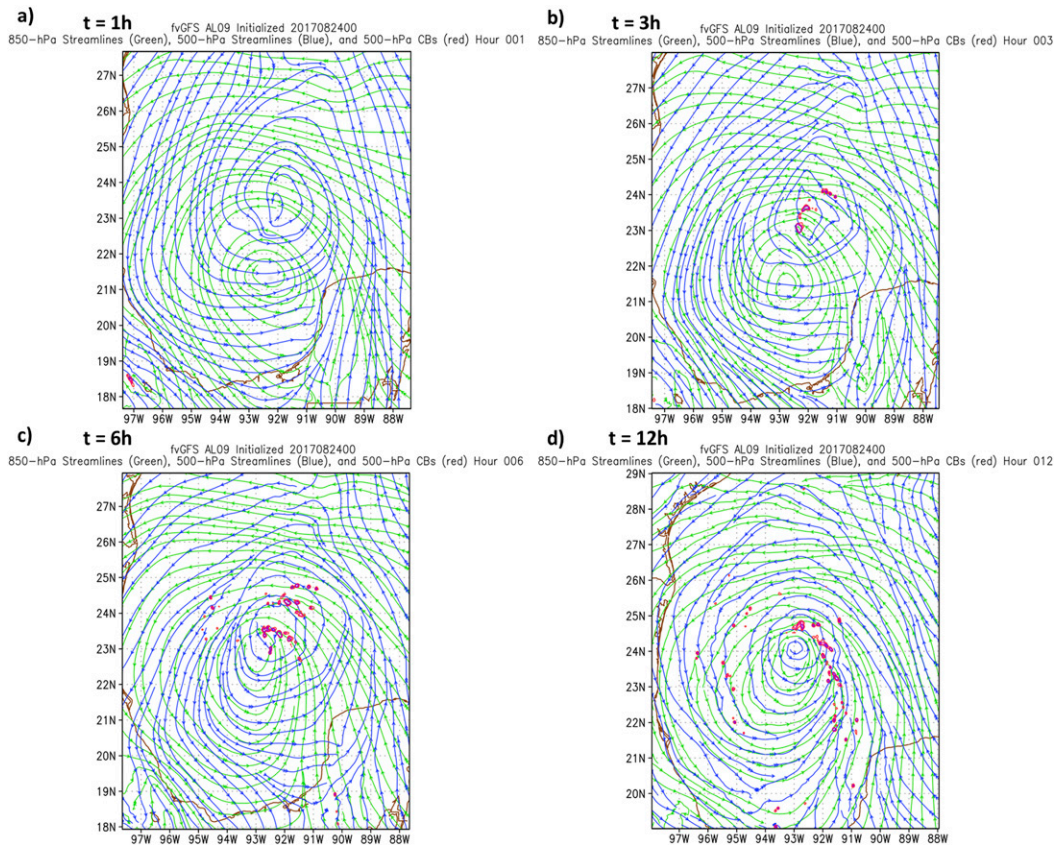


FIG. 13. The 850-hPa winds (green streamlines) and 500-hPa winds (blue streamlines) at (a) 1, (b) 6, (c) 9, and (d) 12 h of the HifvGFS Harvey forecast initialized at 0000 UTC 24 Aug 2017. Vertical velocity at 500 hPa is also contoured, with red contours denoting 3 m s^{-1} , magenta contours denoting 5 m s^{-1} , and purple contours denoting 7 m s^{-1} .

and observational (e.g., Guimond et al. 2010; Hazelton et al. 2017). A similar alignment process was also documented in studies of Hurricane Earl, both numerical (Chen and Gopalakrishnan 2015) and observational (Rogers et al. 2015).

Another key aspect of Hurricane Harvey was the extreme precipitation observed over SE Texas and Louisiana. The map of observed precipitation ending at 1200 UTC 28 August is compared to a corresponding 4-day forecast of accumulated precipitation from the HifvGFS forecast initialized 4 days prior (Fig. 15). Although the model did not capture the maximum of 30+ in. during this 4-day period, it did correctly show much of the observed precipitation pattern, including the maximum near Houston, Texas, with a large area over 20 in., a secondary maximum near the landfall point NE of Corpus Christi, and the large swath of 6–12 in. extending into SW Louisiana. This highlights the ability of HifvGFS to forecast many aspects of a high-impact hurricane including not only track and intensity, but also the kinematic structure (e.g., vortex evolution) and thermodynamic/moisture structure (as seen in the precipitation).

2) HURRICANE MARIA

The second case study focuses on the HifvGFS forecasts of Hurricane Maria. Figure 16 shows the HifvGFS (and HWRF) track and intensity forecasts covering the entire life cycle of Hurricane Maria. Most of the HifvGFS tracks were relatively accurate, although there was a consistent right-of-track bias, especially in the early forecasts. This was also seen in the HWRF forecasts (Fig. 16b). However, the forecasts later in the life cycle of Maria correctly showed the eastward turn and quick acceleration to the north and east.

The intensity forecasts were somewhat more problematic. While most forecasts did show the storm becoming a major hurricane, and a few indicated RI, none correctly captured the timing or magnitude of the RI. HWRF did slightly better on the RI, but most HWRF forecasts were still 30–40 kt too low. The 2-km resolution in HWRF likely allowed for better resolution of the core structure than the 3-km HifvGFS (e.g., Fierro et al. 2009), but the tiny core of Maria was difficult for both models to fully resolve. Later in Maria's

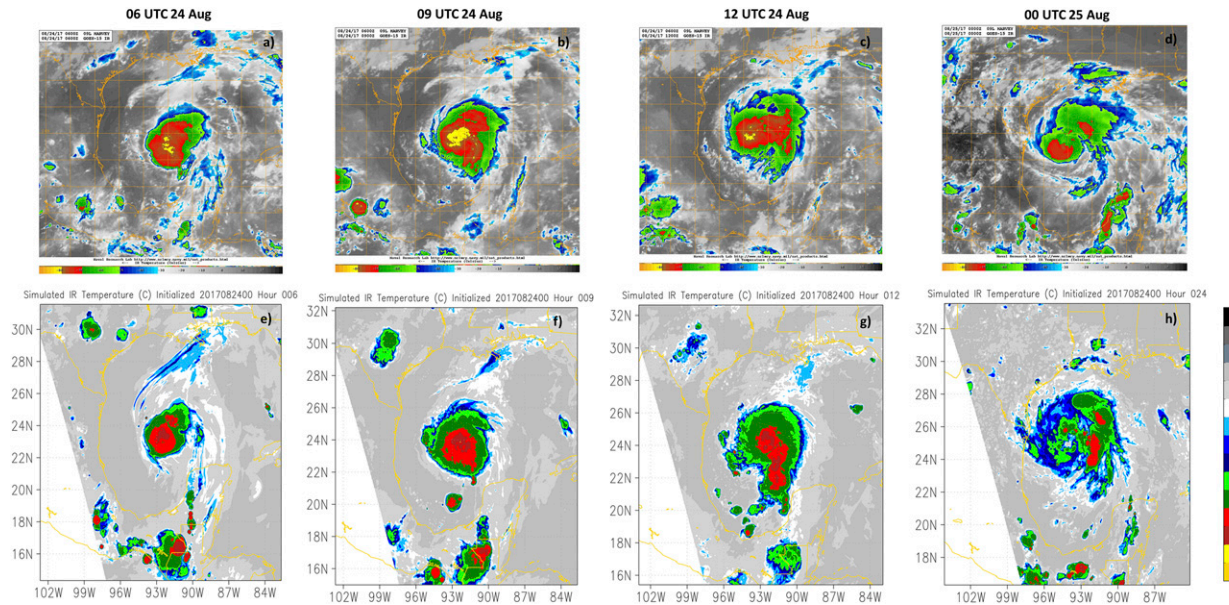


FIG. 14. Observed infrared imagery of Hurricane Harvey at (a) 0600, (b) 0900, and (c) 1200 UTC 24 Aug and (d) 0000 UTC 25 Aug 2017. [Image from the Naval Research Laboratory Monterey Marine Meteorology Division (<https://www.nrlmry.navy.mil/TC.html>).] HifvGFS forecast of simulated IR imagery initialized at 0000 UTC 24 Aug 2017 and valid at (e) 0600, (f) 0900, and (g) 1200 UTC 24 Aug and (h) 0000 UTC 25 Aug 2017.

life cycle, most of the HifvGFS forecasts were too strong, although they did correctly predict decay.

One of the curiosities of Maria’s forecast was the aforementioned right bias in track, even in short-term forecasts. This was also noted in the HWRF forecasts (see Fig. 15b) and GFS forecasts (not shown). A similar bias was seen in forecasts of Hurricane Irma while the storm was in the western Atlantic (e.g., Fig. 1). The 500-hPa heights from the HifvGFS runs initialized at 0000 UTC 18 September at 6, 12, 24, and 48 h are presented in Fig. 17 and compared with the GFS operational analysis. At 6 h, as Maria was approaching the Leeward

Islands, the ridge to the north of the storm was already slightly larger and more pronounced in the observations than in the forecast. This is clearly evident by 12 h, with the area covered by the 590-dam line north of the TC nearly twice as large in the analysis compared to the forecast. The ridging differences persist at the 24- and 48-h forecast lead times, with the 590-dam line covering a much larger area (and having more of a zonal orientation) in observations compared with the model forecasts. The amplitude of the ridge is underestimated, leading to the storm moving more poleward than in reality. It is possible that some of this is due to the model

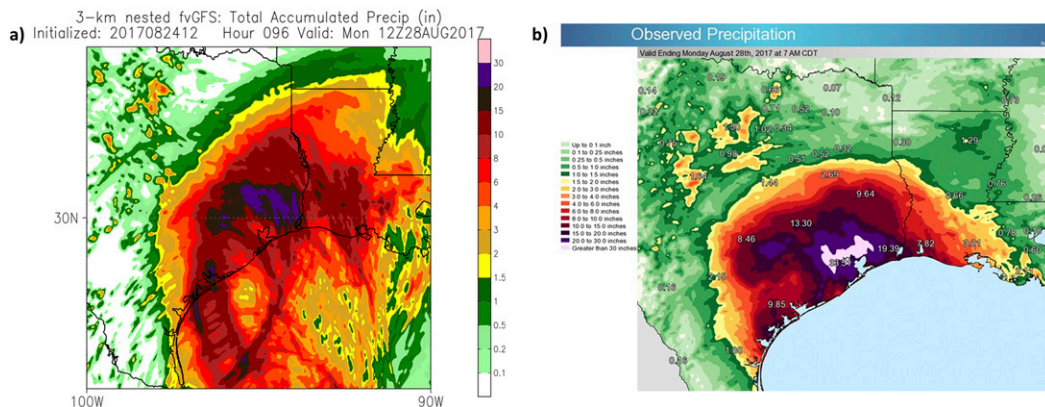


FIG. 15. (a) Harvey 96-h HifvGFS forecasted accumulated precipitation initialized at 1200 UTC 24 Aug 2017. (b) Observed Harvey precipitation ending at 1200 UTC 28 Aug 2017 (<https://www.weather.gov/hgx/hurricaneharvey>).

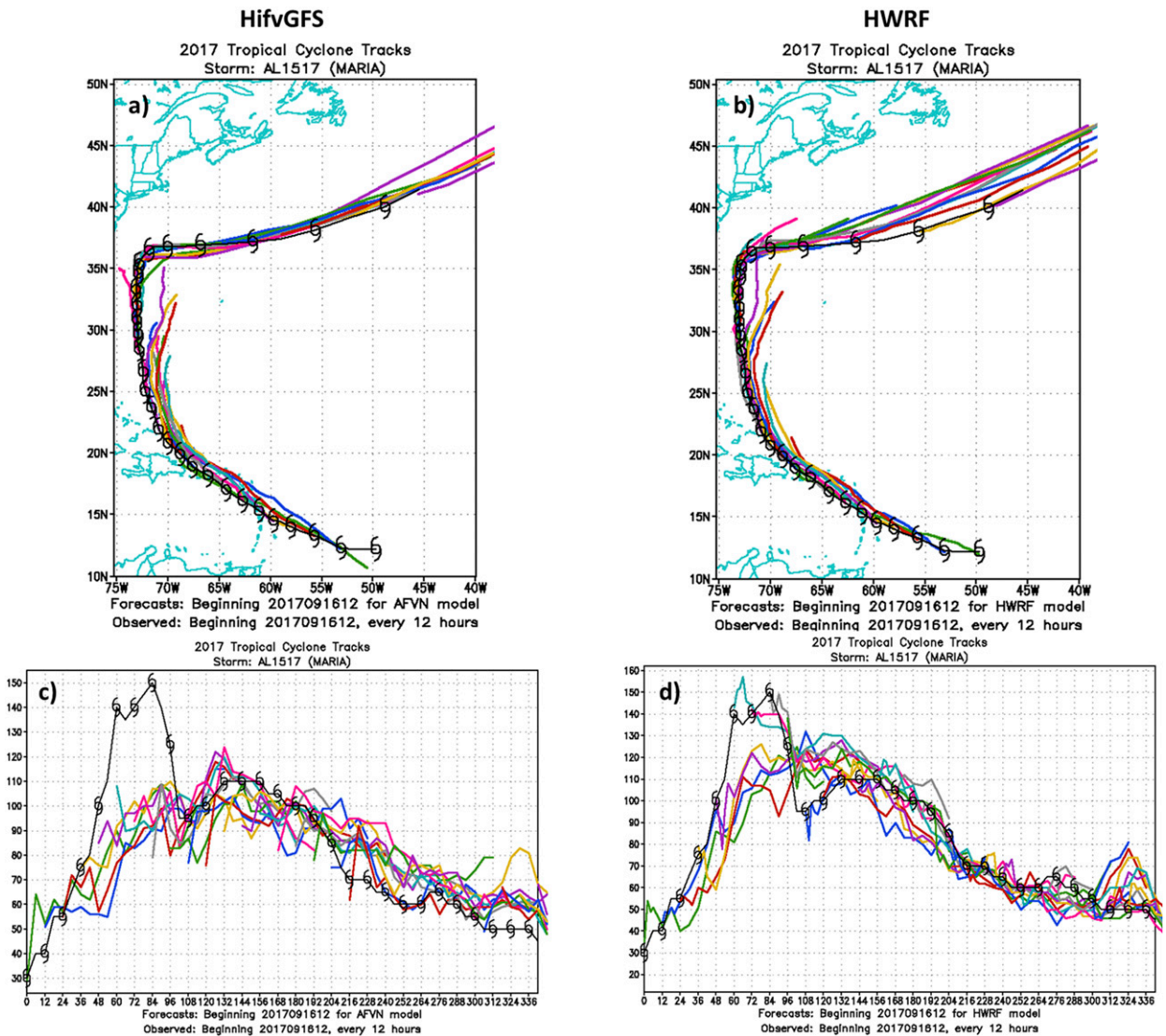


FIG. 16. (a) Composite of HifvGFS 120-h forecast tracks of Hurricane Maria initialized between 1200 UTC 16 Sep and 0000 UTC 1 Oct. The black line with hurricane symbols indicates the best track positions covering this period. (b) As in (a), but for the HWRf Model. (c) The HifvGFS 120-h intensity forecasts of 10-m maximum winds (kt) for Hurricane Maria initialized between 1200 UTC 16 Sep and 0000 UTC 1 Oct. The black line with hurricane symbols indicates the best track intensities (kt) covering this period. (d) As in (c), but for the HWRf Model.

physics, but tests with no cumulus parameterization (not shown) had a similar track. The fact that this bias was already pronounced by 6 h and appeared in other models using the GFS initialization (the operational GFS and HWRf models) implies an issue with the initialization of the ridge. The HifvGFS model currently cannot be run with any other initial conditions, and so this hypothesis cannot be directly tested in this study. However, these results do suggest that improved satellite data assimilation over the tropical and subtropical Atlantic is necessary for improved TC track forecasts.

To examine the possibility that the high bias in intensity seen in many of the later Maria forecasts (as well as other cases, especially during recurvature at higher latitudes) was impacted by the lack of ocean coupling in these runs, one of the Maria forecasts was rerun with a simple one-dimensional ocean model included. This simple ocean model allows for wind-induced mixing and also cooling due to rainfall, and is similar to that of [Pollard et al. \(1973\)](#). The test case was initialized at 1200 UTC 23 September with monthly climatological mixed layer depths.

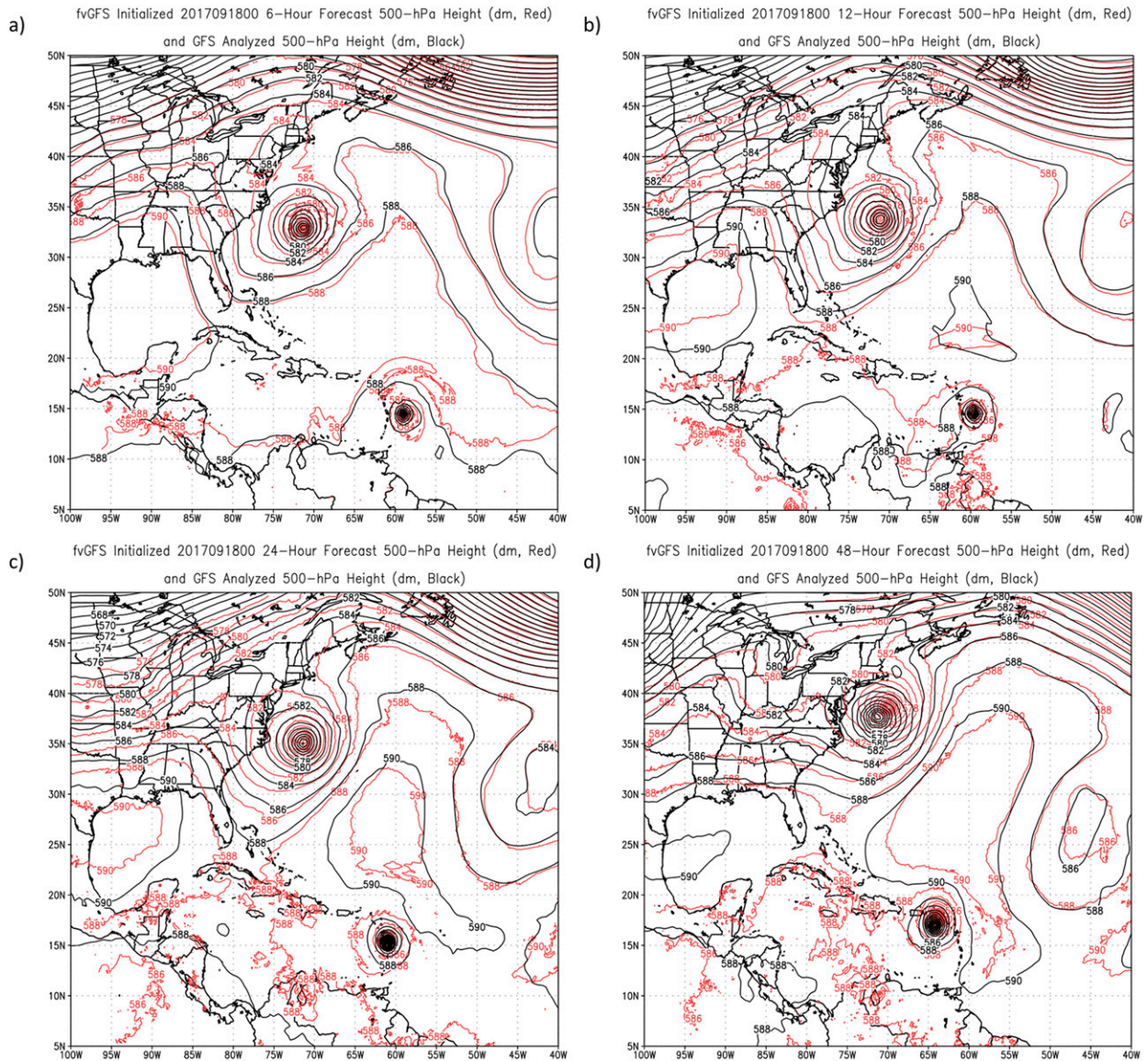


FIG. 17. (a) The HifvGFS 6-h forecast of 500-hPa geopotential height (dam; red) initialized at 0000 UTC 18 Sep 2017 (valid at 0600 UTC 18 Sep), with GFS-analyzed 500-hPa height (dam; black) valid at 0600 UTC 18 Sep. (b) As in (a), but valid at 1200 UTC 18 Sep (12-h forecast). (c) As in (a), but valid at 0000 UTC 19 Sep (24-h forecast). (d) As in (a), but valid at 0000 UTC 20 Sep (48-h forecast).

The coupled fvGFS forecast shows a much smaller high bias in intensity than the original uncoupled forecast, with improvements of ~ 10 kt in wind speed (Fig. 18a) and ~ 10 hPa in pressure (Fig. 18b). In addition, the forecasted SST cooling was consistent with the observations (cf. Figs. 18c and 18d), which showed an area of 1° – 2° C cooling along the track, with a maximum around 3° C along a swath east of Florida. The forecast did not show as wide a cool wake as in the observations, which could be due to the one-dimensional model not capturing three-dimensional ocean processes like advection, Ekman upwelling, or the full extent of the prior

cooling due to Hurricane Jose (especially since the one-dimensional model uses a climatological mixed layer depth). In addition, some areas of cooling and warming outside the TC track were missed. However, the general accuracy of the observed TC-induced cooling and improvement in the intensity forecast is promising. This one-dimensional ocean mixed layer model is still being tuned and refined (and thus was not used on the full suite of cases), but shows the potential for significant improvement of HifvGFS intensity forecasts through this and other physics upgrades that are currently being tested.

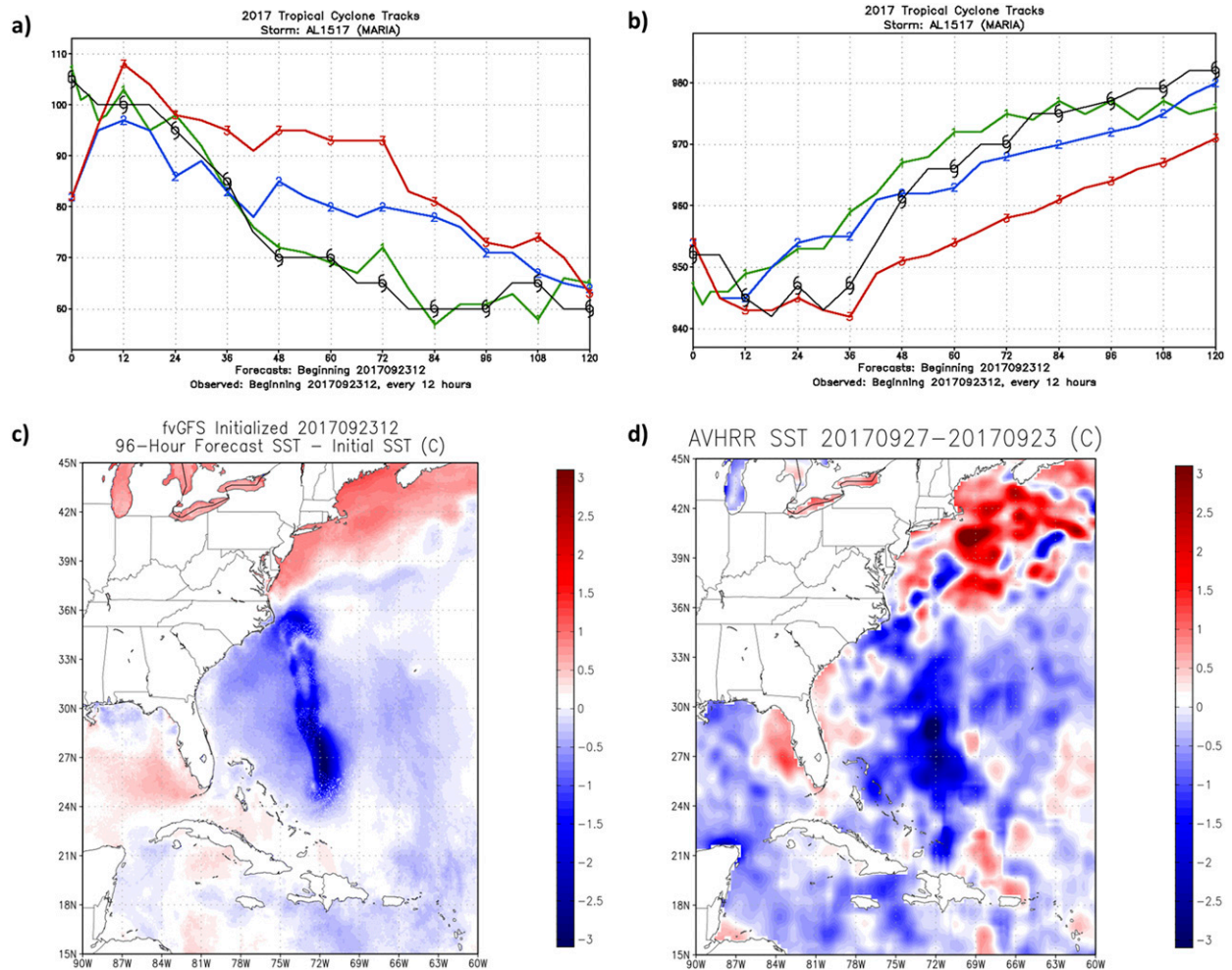


FIG. 18. (a) Intensity forecast of 10-m maximum winds (kt) for forecasts initialized at 1200 UTC 23 Sep 2017 from the original fvGFS run (red), the HifvGFS run with one-dimensional ocean coupling (blue), and HWRf (green). The best track is shown in black with hurricane symbols. (b) As in (a), but for minimum central pressure (hPa). (c) The 4-day change in SST ($^{\circ}\text{C}$) for the HifvGFS run initialized at 1200 UTC 23 Sep 2017 and run with one-dimensional ocean coupling. (d) The 4-day observed change in SST ($^{\circ}\text{C}$) between 23 and 27 Sep 2017 using $1/4^{\circ}$ NOAA daily Optimum Interpolation Sea Surface Temperature (OISST; Banzon et al. 2016) data from the Advanced Very High Resolution Radiometer (AVHRR).

4. Conclusions and future work

The analysis presented in this study demonstrates the forecast capability of the HifvGFS model for improved forecasts of TC track, intensity, and structure. Results are shown for a sample of 347 forecasts (including 313 individual TC cases at 24 h) in the Atlantic, initialized between early August and mid-October 2017, which included the entire life cycles of Hurricanes Harvey, Irma, and Maria. The model's track forecast skill is comparable to or better than other GFS-based models (the operational GFS and HWRf models) at all forecast hours. The cases that did have larger track errors were the erratic tracks of Hurricanes Jose and Lee in the subtropics. However, the track errors were smaller in

the deep tropics, and the overall track skill was similar to other hurricane models.

The HifvGFS showed large intensity improvement over the 13-km gfVgFS as well as the global GFS but was less skillful than HWRf, mainly due to a larger negative bias. The error maps provided more information about the source of these biases. Most of the large negative errors came from Maria and Lee. The model also struggled with the underprediction of the intensity of Hurricane Irma. This is likely due to the fact that Irma remained as a very strong category 4/5 hurricane for about five consecutive days (Cangialosi et al. 2018), and unlike the HWRf Model (e.g., Gopalakrishnan et al. 2012), no TC-specific initialization is currently available in fvGFS, and the model is initialized from 13-km GFS

analyses. This was particularly problematic because Irma was an extremely strong category 4/5 hurricane during this period. Most of the positive intensity forecast bias cases were associated with Jose during its cyclonic loop and with Maria during its recurvature over the North Atlantic. However, the intensity forecasts for Harvey showed promise as HifvGFS correctly predicted the processes that lead to RI.

To analyze the forecasts beyond just the track and intensity metrics, the wind radii forecasts were also compared with observations. It is worth noting that there is some observational uncertainty (e.g., Cangialosi and Landsea 2016), although the findings are generally consistent with the direct comparison with radar data in a smaller set of HifvGFS cases in an earlier paper by Hazelton et al. (2018). It was found that the model wind radii were too large compared to the observations, resulting in larger errors with HifvGFS compared to the operational HWRF or the GFS. This size bias begins early in the forecast period (around hours 12–24), indicating that there may be some spinup issues leading to unrealistic structures. This highlights the need to add storm-scale initialization to avoid the “cold start” from the GFS initial conditions. The impact of this spinup bias could be tested in future work by initializing HifvGFS with the HWRF initial conditions. After these early forecast times, the increase in error with time is generally consistent with the other models. However, recent tests (not shown) show that changes to the tracer advection and boundary layer schemes have a large impact on wind radii in HifvGFS, and such changes will continue to be explored. Specifically, these changes produce smaller errors in the 50- and 64-kt wind radii, although the 34-kt wind radii are not impacted significantly (perhaps implying that the 34-kt wind radii are more sensitive to the initial conditions and spinup). Such changes will continue to be explored during the 2018 hurricane season and in further tests as the model is developed.

The case studies presented highlighted both the model’s strengths and areas where it can be improved. The HifvGFS forecasts showed the evolution of Harvey’s structure from a weak TS with a large vertical tilt to a stacked, vertically coherent vortex primed for RI. This alignment was associated with strong convection in the core of Harvey, similar to the RI initiation of Hurricane Earl in 2010 (Rogers et al. 2015; Chen and Gopalakrishnan 2015). A majority of the HifvGFS forecasts of Hurricane Harvey correctly captured the RI. The model also correctly predicted the extreme rainfall near Houston and much of SE Texas and SW Louisiana. For Hurricane Maria, the overall track evolution was generally consistent with the observations, but a weak bias in the ridge strength north of the TC led to a consistent right bias in

the early forecasts. This appears to be due to the GFS initial conditions, as similar behavior was seen in the tracks of both the HWRF and GFS models as well (not shown). Although some of the Maria intensity forecasts did show intensification into a major hurricane, the model did not capture the peak intensity. This was a case where the resolution and initialization likely were key issues, as Maria was an extremely intense storm with a tiny inner core. Future work to improve TC initialization in HifvGFS will examine this possibility. As noted above, one of the other issues with the Maria forecasts was a tendency for a high bias in intensity later in the period. A test case using a one-dimensional mixed layer ocean model showed promise in alleviating this issue, with the ocean cooling producing a ~10-kt decrease in the positive intensity bias.

Future work will build on the positive results seen in this study and improve the model forecasts of track, structure, and intensity. Given the success of the Maria case study using a one-dimensional ocean model, this ocean coupling will be tuned and improved to help eliminate the high-intensity bias seen in some cases, especially during recurvature. This will lead toward the eventual goal of coupling to a fully three-dimensional ocean, such as that used by operational HWRF (e.g., Mellor 2004), that can properly account for all TC interaction processes, including upwelling due to slow-moving TCs (like Jose) or TCs moving over a cold wake as simulated by Bender and Ginis (2000). The improvements seen in HWRF (e.g., Biswas et al. 2017) as the model’s horizontal resolution was increased from 3 to 2 km and the vertical resolution was improved with the addition of more levels (from 61 to 75) motivates similar changes in future upgrades of HifvGFS. Versions of the model with higher horizontal and vertical resolution will be tested over a large sample set to ensure that they lead to improved forecasts, hopefully eliminating some of the large negative errors and improving RI forecasts. The structure (e.g., wind radii) forecasts from future versions of the model will also be further validated as observational data from the 2017 hurricane season becomes available. Different physics configurations (e.g., microphysics, boundary layer schemes) will be tested, and evaluation of other important structural metrics such as rainfall forecast skill will be performed. Finally, work will be undertaken on vortex-scale data assimilation and improved TC initialization, in order to allow the model to predict the short-term evolution of extremely strong TCs like Irma and Maria.

In conclusion, this paper has demonstrated that HifvGFS is a promising model capable of the skillful prediction of TC track, intensity, and structure, and continued improvements as noted above will enable it to become a top-flight high-resolution hurricane prediction system.

Acknowledgments. ATH was supported by NOAA Grant NA14OAR4830101. Baoqiang Xiang was very helpful in the design and implementation of the one-dimensional ocean mixed layer model for the Hurricane Maria test. Kieran Bhatia, Tim Marchok, Frank Marks, and two anonymous reviewers provided helpful comments that improved an earlier version of the manuscript.

REFERENCES

- Aberson, S. D., 1998: Five-day tropical cyclone track forecasts in the North Atlantic basin. *Wea. Forecasting*, **13**, 1005–1015, [https://doi.org/10.1175/1520-0434\(1998\)013<1005:FDTCTF>2.0.CO;2](https://doi.org/10.1175/1520-0434(1998)013<1005:FDTCTF>2.0.CO;2).
- 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>.
- Banzon, V., T. M. Smith, T. M. Chin, C. Liu, and W. Hankins, 2016: A long-term record of blended satellite and in situ sea-surface temperature for climate monitoring, modeling and environmental studies. *Earth Syst. Sci. Data*, **8**, 165–176, <https://doi.org/10.5194/essd-8-165-2016>.
- Bender, M. A., 2017: Evaluation of tropical cyclone forecasts with the GFDL FV3 dynamical core. *71st Interdepartmental Hurricane Conf.*, Miami, FL, Federal Coordinator for Meteorological Services and Supporting Research, http://www.ofcm.gov/meetings/TCORF/ihc17/Session_08/8-3-Bender_web.pdf.
- , and I. Ginis, 2000: Real-case simulation of hurricane–ocean interaction using a high-resolution coupled model: Effects on hurricane intensity. *Mon. Wea. Rev.*, **128**, 917–946, [https://doi.org/10.1175/1520-0493\(2000\)128<0917:RCOHO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<0917:RCOHO>2.0.CO;2).
- , T. P. Marchok, C. R. Sampson, J. A. Knaff, and M. J. Morin, 2017: Impact of storm size on prediction of storm track and intensity using the 2016 operational GFDL hurricane model. *Wea. Forecasting*, **32**, 1491–1508, <https://doi.org/10.1175/WAF-D-16-0220.1>.
- Berg, R., 2016: Hurricane Joaquin (AL112015) 28 September–7 October 2015. National Hurricane Center Tropical Cyclone Rep., 36 pp., http://www.nhc.noaa.gov/data/tcr/AL112015_Joaquin.pdf.
- Biswas, M., and Coauthors, 2017: Hurricane Weather Research and Forecasting (HWRF) Model: 2017 scientific documentation. NCAR Tech. Note NCAR/TN-544+STR, 99 pp., https://dtcenter.org/HurrWRF/users/docs/scientific_documents/HWRFv3.9a_ScientificDoc.pdf.
- 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>.
- , and J. L. Franklin, 2017: 2016 hurricane season. National Hurricane Center Forecast Verification Rep., 72 pp., http://www.nhc.noaa.gov/verification/pdfs/Verification_2016.pdf.
- , A. S. Latta, and R. Berg, 2018: Hurricane Irma (AL112017) 30 August–12 September 2017. National Hurricane Center Tropical Cyclone Rep., 111 pp., https://www.nhc.noaa.gov/data/tcr/AL112017_Irma.pdf.
- Chen, H., and S. G. Gopalakrishnan, 2015: A study on the asymmetric rapid intensification of Hurricane Earl (2010) using the HWRF system. *J. Atmos. Sci.*, **72**, 531–550, <https://doi.org/10.1175/JAS-D-14-0097.1>.
- 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>.
- Davis, C., and Coauthors, 2008: Prediction of landfalling hurricanes with the Advanced Hurricane WRF Model. *Mon. Wea. Rev.*, **136**, 1990–2005, <https://doi.org/10.1175/2007MWR2085.1>.
- DeMaria, M., 1996: The effect of vertical shear on tropical cyclone intensity change. *J. Atmos. Sci.*, **53**, 2076–2088, [https://doi.org/10.1175/1520-0469\(1996\)053<2076:TEOVSO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1996)053<2076:TEOVSO>2.0.CO;2).
- , and J. Kaplan, 1994: A Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic basin. *Wea. Forecasting*, **9**, 209–220, [https://doi.org/10.1175/1520-0434\(1994\)009<0209:ASHIPS>2.0.CO;2](https://doi.org/10.1175/1520-0434(1994)009<0209:ASHIPS>2.0.CO;2).
- Fierro, A. O., R. F. Rogers, F. D. Marks, and D. S. Nolan, 2009: The impact of horizontal grid spacing on the microphysical and kinematic structures of strong tropical cyclones simulated with the WRF-ARW Model. *Mon. Wea. Rev.*, **137**, 3717–3743, <https://doi.org/10.1175/2009MWR2946.1>.
- Gopalakrishnan, S. G., S. Goldenberg, R. Quirino, X. Zhang, F. Marks Jr., K.-S. Yeh, R. Atlas, and V. Tallapragada, 2012: Toward improving high-resolution numerical hurricane forecasting: Influence of model horizontal grid resolution, initialization, and physics. *Wea. Forecasting*, **27**, 647–666, <https://doi.org/10.1175/WAF-D-11-00055.1>.
- , F. Marks Jr., J. A. Zhang, X. Zhang, J.-W. Bao, and V. Tallapragada, 2013: A study of the impacts of vertical diffusion on the structure and intensity of the tropical cyclones using the high-resolution HWRF system. *J. Atmos. Sci.*, **70**, 524–541, <https://doi.org/10.1175/JAS-D-11-0340.1>.
- Guimond, S. R., G. M. Heymsfield, and F. J. Turk, 2010: Multiscale observations of Hurricane Dennis (2005): The effects of hot towers on rapid intensification. *J. Atmos. Sci.*, **67**, 633–654, <https://doi.org/10.1175/2009JAS3119.1>.
- Han, J., and H.-L. Pan, 2011: Revision of convection and vertical diffusion schemes in the NCEP Global Forecast System. *Wea. Forecasting*, **26**, 520–533, <https://doi.org/10.1175/WAF-D-10-05038.1>.
- , 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., R. E. Hart, and R. Rogers, 2017: Analyzing simulated convective bursts in two Atlantic hurricanes. Part II: Intensity change due to convective bursts. *Mon. Wea. Rev.*, **145**, 3073–3094, <https://doi.org/10.1175/MWR-D-16-0267.1>.
- , L. Harris, and S.-J. Lin, 2018: Evaluation of tropical cyclone structure 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>.
- 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., M. DeMaria, and J. 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/10.1175/2009WAF222280.1>.
- Knaff, J. A., M. DeMaria, B. Sampson, and J. M. Gross, 2003: Statistical, 5-day tropical cyclone intensity forecasts derived from climatology and persistence. *Wea. Forecasting*, **18**,

- 80–92, [https://doi.org/10.1175/1520-0434\(2003\)018<0080:SDTCIF>2.0.CO;2](https://doi.org/10.1175/1520-0434(2003)018<0080:SDTCIF>2.0.CO;2).
- 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>.
- , 2004: A “vertically Lagrangian” finite-volume dynamical core for global models. *Mon. Wea. Rev.*, **132**, 2293–2307, [https://doi.org/10.1175/1520-0493\(2004\)132<2293:AVLFDC>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<2293:AVLFDC>2.0.CO;2).
- , and R. B. Rood, 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>.
- Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22**, 1065–1092, [https://doi.org/10.1175/1520-0450\(1983\)022<1065:BPOTSF>2.0.CO;2](https://doi.org/10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2).
- Mellor, G. L., 2004: Users guide for a three-dimensional, primitive equation, numerical ocean model (June 2004 version). Princeton University, 56 pp., http://jes.apl.washington.edu/modsims_two/usersguide0604.pdf.
- Miller, R. J., A. J. Schrader, C. R. Sampson, and T. L. Tsui, 1990: The Automated Tropical Cyclone Forecasting system (ATCF). *Wea. Forecasting*, **5**, 653–660, [https://doi.org/10.1175/1520-0434\(1990\)005<0653:TATCFS>2.0.CO;2](https://doi.org/10.1175/1520-0434(1990)005<0653:TATCFS>2.0.CO;2).
- Neumann, C. J., M. B. Lawrence, and E. L. Caso, 1977: Monte Carlo significance testing as applied to statistical tropical cyclone models. *J. Appl. Meteor.*, **16**, 1165–1174, [https://doi.org/10.1175/1520-0450\(1977\)016<1165:MCSTAA>2.0.CO;2](https://doi.org/10.1175/1520-0450(1977)016<1165:MCSTAA>2.0.CO;2).
- Pollard, R. T., P. B. Rhines, and R. O. R. Y. Thompson, 1973: The deepening of the wind-mixed layer. *Geophys. Fluid Dyn.*, **3**, 381–404, <https://doi.org/10.1080/03091927208236105>.
- Rogers, R. F., P. D. Reasor, and J. A. Zhang, 2015: Multiscale structure and evolution of Hurricane Earl (2010) during rapid intensification. *Mon. Wea. Rev.*, **143**, 536–562, <https://doi.org/10.1175/MWR-D-14-00175.1>.
- Sampson, C. R., and A. J. Schrader, 2000: The Automated Tropical Cyclone Forecasting system (version 3.2). *Bull. Amer. Meteor. Soc.*, **81**, 1231–1240, [https://doi.org/10.1175/1520-0477\(2000\)081<1231:TATCFS>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<1231:TATCFS>2.3.CO;2).
- Tallapragada, V., C. Kieu, Y. Kwon, S. Trahan, Q. Liu, Z. Zhang, and I. Kwon, 2014: Evaluation of storm structure from the operational HWRF during 2012 implementation. *Mon. Wea. Rev.*, **142**, 4308–4325, <https://doi.org/10.1175/MWR-D-13-00010.1>.
- Zhang, J. A., R. F. Rogers, and V. Tallapragada, 2017: Impact of parameterized boundary layer structure on tropical cyclone rapid intensification forecasts in HWRF. *Mon. Wea. Rev.*, **145**, 1413–1426, <https://doi.org/10.1175/MWR-D-16-0129.1>.
- Zhao, Q., and F. H. Carr, 1997: A prognostic cloud scheme for operational NWP models. *Mon. Wea. Rev.*, **125**, 1931–1953, [https://doi.org/10.1175/1520-0493\(1997\)125<1931:APCSFO>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<1931:APCSFO>2.0.CO;2).