Performance of an Improved TKE-Based Eddy-Diffusivity Mass-Flux (EDMF) PBL Scheme in 2021 Hurricane Forecasts from the Hurricane Analysis and Forecast System

XIAOMIN CHEN,[a,b] ANDREW HAZELTON,[a,c] FRANK D. MARKS,[a] GHASSAN J. ALAKA JR.,[a] AND CHUNXI ZHANG[d,e]

[a] *NOAA/OAR/Atlantic Oceanographic and Meteorological Laboratory, Miami, Florida*
[b] *Northern Gulf Institute, Mississippi State University, Stennis Space Center, Mississippi*
[c] *Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, Florida*
[d] *NOAA/NWS/NCEP Environmental Modeling Center, College Park, College Park, Maryland*
[e] *I.M. Systems Group, Inc., College Park, College Park, Maryland*

(Manuscript received 26 July 2022, in final form 24 November 2022)

ABSTRACT: Continuous development and evaluation of planetary boundary layer (PBL) parameterizations in hurricane conditions are crucial for improving tropical cyclone (TC) forecasts. A turbulence kinetic energy (TKE)-based eddy-diffusivity mass-flux (EDMF-TKE) PBL scheme, implemented in NOAA's Hurricane Analysis and Forecast System (HAFS), was recently improved in hurricane conditions using large-eddy simulations. This study evaluates the performance of HAFS TC forecasts with the original (experiment HAFA) and modified EDMF-TKE (experiment HAFY) based on a large sample of cases during the 2021 North Atlantic hurricane season. Results indicate that intensity and structure forecast skill was better overall in HAFY than in HAFA, including during rapid intensification. Composite analyses demonstrate that HAFY produces shallower and stronger boundary layer inflow, especially within 1–3 times the radius of maximum wind (RMW). Stronger inflow and more moisture in the boundary layer contribute to stronger moisture convergence near the RMW. These boundary layer characteristics are consistent with stronger, deeper, and more compact TC vortices in HAFY than in HAFA. Nevertheless, track skill in HAFY is slightly reduced, which is in part attributable to the cross-track error from a few early cycles of Hurricane Henri that exhibited ~400 n mi (1 n mi = 1.852 km) track error at longer lead times. Sensitivity experiments based on HAFY demonstrate that turning off cumulus schemes notably reduces the track errors of Henri while turning off the deep cumulus scheme reduces the intensity errors. This finding hints at the necessity of unifying the mass fluxes in PBL and cumulus schemes in future model physics development.

KEYWORDS: Boundary layer; Hurricanes/typhoons; Tropical cyclones; Forecast verification/skill; Numerical weather prediction/forecasting; Subgrid-scale processes

1. Introduction

Large sensitivities of modeled tropical cyclone (TC)'s intensity and structure to planetary boundary layer (PBL) parameterizations have been frequently documented in numerical modeling studies (e.g., Braun and Tao 2000; Smith and Thomsen 2010; Hill and Lackmann 2009; Nolan et al. 2009a,b; Bryan 2012; Gopalakrishnan et al. 2013; Bu et al. 2017; Zhang and Pu 2017; Chen et al. 2021b, 2022; Hazelton et al. 2022), which can be attributable to the fact that PBL schemes are rarely designed for hurricane conditions and parameterizations of vertical diffusion notably differ between different PBL schemes (Kepert 2012; Chen et al. 2021a; Chen 2022). Vertical diffusion of momentum in TC boundary layers affects inflow layer depth and inflow strength, which tie closely to radial advection of absolute angular momentum and, thereby, the inner-core contraction and TC intensification (Zhang et al. 2015; Chen and Bryan 2021). Vertical diffusion of water vapor in the boundary layer affects outer-core convective activity by modulating the transport of water vapor to the top of the boundary layer and further impacts the forecasts of the radius of gale-force wind (Bu et al. 2017). The intrinsic uncertainties of vertical diffusion within PBL schemes result in biased forecasts of TC structure and intensity, which further alters the way TCs interact with large-scale environmental flow and contributes to track errors. Thus, accurately representing boundary layer turbulent mixing in PBL schemes, especially in hurricane conditions, is highly desirable to improve TC forecasts.

Efforts have been made in recent years to improve PBL schemes in hurricane conditions by using in situ observations and large-eddy simulations (LES) (e.g., Chen and Bryan 2021; Chen et al. 2022; Chen 2022). As PBL schemes are essentially designed in a single-column fashion, assessing the vertical profile of turbulence variables is key to revealing the deficiencies of these schemes. However, observational profiles of turbulence metrics such as eddy viscosity are entirely missing where the surface wind speed is greater than 30 m s^{-1}. To fill in this gap in the observations and to evaluate PBL schemes in hurricane conditions, Chen et al. (2021a) developed a modeling framework that can be used for a small-domain [$O(5)$ km] LES and also single-column modeling (SCM) using different PBL schemes. This framework is built upon Bryan et al. (2017), which uses a few input parameters to represent the TC structure, and further includes a "large-scale" nudging technique that can anchor the thermodynamic profiles derived

Chen's current affiliation: Department of Atmospheric and Earth Science, University of Alabama in Huntsville, Huntsville, Alabama.

Corresponding author: Xiaomin Chen, xc0011@uah.edu

DOI: 10.1175/WAF-D-22-0140.1

from actual major hurricanes during the simulations. This special setting warrants the same thermodynamic conditions in the LES and SCM tests. Thus, the domain-averaged LES data can provide valuable profile information of turbulence properties under realistic hurricane conditions and also can be treated as the benchmark to evaluate different types of PBL schemes, ranging from *K*-profile parameterization schemes and a Louis-type scheme (Chen et al. 2021a) to high-order closure PBL schemes (Chen and Bryan 2021; Chen et al. 2022; Chen 2022) including Mellor–Yamada–Nakanishi–Niino (MYNN) and a turbulence kinetic energy (TKE)-based eddy-diffusivity mass-flux (EDMF-TKE) scheme.

The EDMF-TKE scheme is a recently updated PBL scheme in the Global Forecast System (GFS) version 16.0, a key part of NOAA's Unified Forecast System (UFS). As NOAA's next-generation hurricane application of the UFS, the Hurricane Analysis and Forecast System (HAFS) uses the same PBL scheme. The EDMF-TKE scheme includes special parameterizations for moist processes, including enhanced buoyancy in clouds and the interaction between TKE and cumulus convection (Han and Bretherton 2019). Additionally, TKE in EDMF-TKE is defined at the mass point or half layer such that it can be advected by gridscale flows, which can represent the inhomogeneous boundary layer conditions like in TCs (Chen and Bryan 2021). Nevertheless, the EDMF-TKE scheme was mainly designed for low-wind, unstable boundary layer conditions. Evaluation of EDMF-TKE in hurricane conditions based on LES results revealed a substantial overestimation of TKE below 500-m height (Chen 2022), and a modified EDMF-TKE scheme was proposed to address this issue (Chen et al. 2022). Three-dimensional idealized simulations and ensemble HAFS forecasts of Hurricane Michael (2018) consistently demonstrated that the modified EDMF-TKE produces stronger boundary layer inflow within the inner-core region that contributes to stronger TC intensity and a smaller radius of maximum wind (RMW) (Chen et al. 2022).

To extend the work of Chen et al. (2022) and to systematically examine the performance of the modified EDMF-TKE during the hurricane season, this study compares the performance of the original and modified EDMF-TKE schemes from a stand-alone version of HAFS during the 2021 North Atlantic hurricane season. With a large sample size, we compare the track, intensity, and structure forecast skill and examine the composite thermodynamic and kinematic structures of TC vortices. Results will guide the future development of PBL schemes and other physics components in HAFS and UFS.

2. Data, method, and model setup

a. *Model setup and a brief overview of the modified EDMF-TKE*

The 2021 baseline version of the stand-alone-regional HAFS (hereafter, HAFS-SAR; Dong et al. 2020) is used in this study. HAFS-SAR features a large static nest over the North Atlantic, Gulf of Mexico, Caribbean Sea, and eastern United States, with a horizontal grid spacing of ~3 km. Additionally, HAFS-SAR adopts a hybrid, sigma-pressure vertical coordinate, and the version of HAFS-SAR analyzed here uses 91 vertical levels spanning from the surface to 10 hPa. There are 31 levels below 700 hPa, and the layer thickness below the bottom model level is ~20 m. Key physics parameterizations used in HAFS-SAR include the GFDL 6-class microphysics scheme (Chen and Lin 2013), scale-aware SAS convective scheme (Han et al. 2017), and RRTMG shortwave and longwave radiation schemes (Iacono et al. 2008). HAFS-SAR is coupled to the Hybrid Coordinate Ocean Model (HYCOM; Bleck 2002). Two sets of experiments are compared in this study: one experiment (HAFA) uses the original EDMF-TKE (Han and Bretherton 2019) and the other (HAFY) uses the modified EDMF-TKE (Chen et al. 2022).

The modified EDMF-TKE scheme incorporates four major changes based on boundary layer theories and LES results (Chen et al. 2022): 1) determining values of two coefficients in the TKE dissipation term and eddy viscosity to match the surface layer and PBL parameterizations [see Eqs. (12)–(15) and the related discussions in Chen et al. 2022], 2) reducing the maximum allowable mixing length from 300 to 40 m (which also agrees with observational values, see Fig. 3 in Chen et al. 2022), 3) implementing a new definition of PBL height that performs better in high-wind conditions [see Eq. (16) in Chen et al. 2022], and 4) tapering and then turning off mass fluxes from the nonlocal portion of the PBL scheme in high-wind conditions ($\geq$20 m s^{-1}). These changes reduce the excessive vertical mixing in hurricane conditions that occurs when using the original EDMF-TKE scheme. Profiles of boundary layer tangential and radial winds as well as eddy viscosity are thereby substantially improved and match well with the LES results. Of note, the first two changes were found to have the largest impact on the improvements (see Fig. 6 and section 4a in Chen et al. 2022).

b. *Case analyzed, verification methods, and observational data*

HAFA was run in near–real time by NOAA's Environmental Modeling Center (EMC) during the 2021 North Atlantic hurricane season. Four cycles were run each day at 0000, 0600, 1200, and 1800 UTC. In comparison, HAFY was a hindcast over the period from 2 July to 7 November 2021, with two cycles each day at 0000 and 1200 UTC except for two high-impact landfalling hurricanes (i.e., Ida and Nicholas, https://www.ncei.noaa.gov/access/billions/) that had four cycles each day. Table 1 shows the matched cases from HAFA and HAFY that will be used in the following analysis. Following the standard National Hurricane Center (NHC) verification rules (Franklin 2008), only tropical and subtropical cyclones of at least depression intensity at initialization and verification time are included in the verification statistics. The total number of cycles is 179, of which 120 cycles are weak storms (i.e., tropical depressions or tropical storms) at the initial time. We started the verifications at 12 h since both HAFA and HAFY were cold-started using the GFS analysis. For 2022 and beyond, a more sophisticated initialization and data assimilation scheme has been developed.

Verification metrics used in this study include the mean absolute error (MAE), mean error (i.e., bias), and forecast skill

TABLE 1. List of matched cycle numbers and initialization periods between HAFA and HAFY.

Storm ID	Storm name	No. of cycles	Initialization periods
AL05	Elsa	12	1200 UTC 2 Jul–1200 UTC 8 Jul
AL06	Fred	13	1200 UTC 11 Aug–0000 UTC 18 Aug
AL07	Grace	16	0000 UTC 14 Aug–1200 UTC 21 Aug
AL08	Henri	13	1200 UTC 16 Aug–1200 UTC 22 Aug
AL09	Ida	14	1200 UTC 26 Aug–0000 UTC 30 Aug
AL10	Kate	9	0600 UTC 28 Aug–1200 UTC 1 Sep
AL11	Julian	5	0000 UTC 29 Aug–0000 UTC 30 Aug
AL12	Larry	19	0000 UTC 1 Sep–0000 UTC 10 Sep
AL13	Mindy	3	0000 UTC 9 Sep–0000 UTC 10 Sep
AL14	Nicholas	10	1800 UTC 12 Sep–0000 UTC 15 Sep
AL16	Peter	8	1200 UTC 19 Sep–0000 UTC 23 Sep
AL17	Rose	8	1200 UTC 19 Sep–0000 UTC 23 Sep
AL18	Sam	23	0000 UTC 23 Sep–0000 UTC 4 Oct
AL19	Teresa	2	0000 UTC 25 Sep–1200 UTC 25 Sep
AL20	Victor	10	1200 UTC 29 Sep–0000 UTC 4 Oct
AL21	Wanda	14	1200 UTC 31 Oct–0000 UTC 7 Nov

improvement of MAE (Rappaport et al. 2009; Sippel et al. 2022). The skill improvement is measured by Brier skill score (BSS): BSS = 1 − (*F*/REF), where REF is the skill baseline (i.e., MAE of HAFA) and *F* is the MAE of HAFY. To identify notable outliers that affect the statistics, the spatial distribution of MAE of specific variables is examined. The structure metrics, including the radius of hurricane-force wind (R64), the radius of 50-kt (1 kt ≈ 0.51 m s^{-1}) wind (R50), and the radius of gale-force wind (R34), are calculated as the mean radius in all quadrants around the TC center.

To verify the inflow structure from HAFA and HAFY, we used a radial profile of observed inflow angle at z = 10 m based on a composite of 1600 Global Positioning System (GPS) dropsondes collected in 18 different hurricanes (Zhang and Uhlhorn 2012). Additionally, we used a recently developed airborne Doppler radar database (referred to as TC-RADAR; Fischer et al. 2022) to verify the TC structure of Hurricane Ida in section 4a. TC-RADAR includes a post-processed merged analysis of tail Doppler radar swaths, which provides greater azimuthal coverage of observations than a single swath and therefore is optimal for the evaluation of model vortex structure.

3. Results

In this section, we will first compare the performance of HAFA and HAFY in terms of track, intensity, and structure forecast metrics. Then, the impact of PBL schemes on vortex structure and thermodynamics as well as boundary layer structure will be discussed.

a. Verification statistics

Figure 1a compares the track performance of HAFA and HAFY. Track MAE is comparable between the two experiments before 48 h, and then differences in track MAE gradually increase with time. At t = 120 h, the track MAE of HAFY is ~25 n mi (1 n mi = 1.852 km) greater than that of HAFA. Figure 1b shows the track forecast skill of HAFY relative to HAFA. Consistent with track MAE, HAFY has degraded track skill compared to HAFA, with the extent of degradation increasing with time and reaching the maximum of 14%–15% over t = 108–120 h. Decomposing the track MAE into along-track and cross-track bias in Figs. 1c and 1d shows that the larger track MAE in HAFY is mostly attributable to a larger right-of-track bias, especially after t = 84 h.

Comparison of the intensity MAE in Fig. 1e indicates that HAFY has a slightly smaller intensity error than HAFA before t = 90 h and a slightly larger intensity error afterward. This result is well supported by the improved intensity skill of HAFY relative to HAFA before t = 90 h (see Fig. 1f), with a maximum improvement of ~17% at t = 36 h. After t = 90 h, HAFY has reduced intensity skill relative to HAFA, and the degradation increases with time. Figures 1g and 1h further show that HAFY reduces the weak bias in terms of both maximum sustained 10-m wind speed (VMAX) and minimum sea level pressure (PMIN) compared to HAFA.

The notable degradation in both intensity and track skill at longer lead times motivates us to investigate the potential causes. Figures 2a and 2b compare the spatial distribution of track MAE for all cases from HAFA and HAFY, indicated by their storm ID, at t = 96 h, when degradation of track skill in HAFY exceeds 9% relative to HAFA and for the first time the relative intensity skill of HAFY becomes negative (Figs. 1c,d). One important finding in Figs. 2a and 2b is that the larger track MAE at longer lead times from HAFY is in part driven by consistent track errors of >360 n mi from a few forecast cycles of Henri (08L) (see circled purple storm ID in Fig. 2b), as the track MAE of Henri in HAFA is generally <320 n mi (Fig. 2a). The histogram of track errors for all Henri cases at t = 96 h in Figs. 2c and 2d support this finding and further show that the maximum (mean) track error in HAFY is ~85 (58) n mi greater than that in HAFA. Details of the factors that contributed to the large track error in HAFY will be discussed in section 4.

To quantify the contribution of these anomalies to the overall performance of forecast skill, we removed four cycles of Henri circled in Fig. 2b or the rightmost four cycles in Fig. 2d

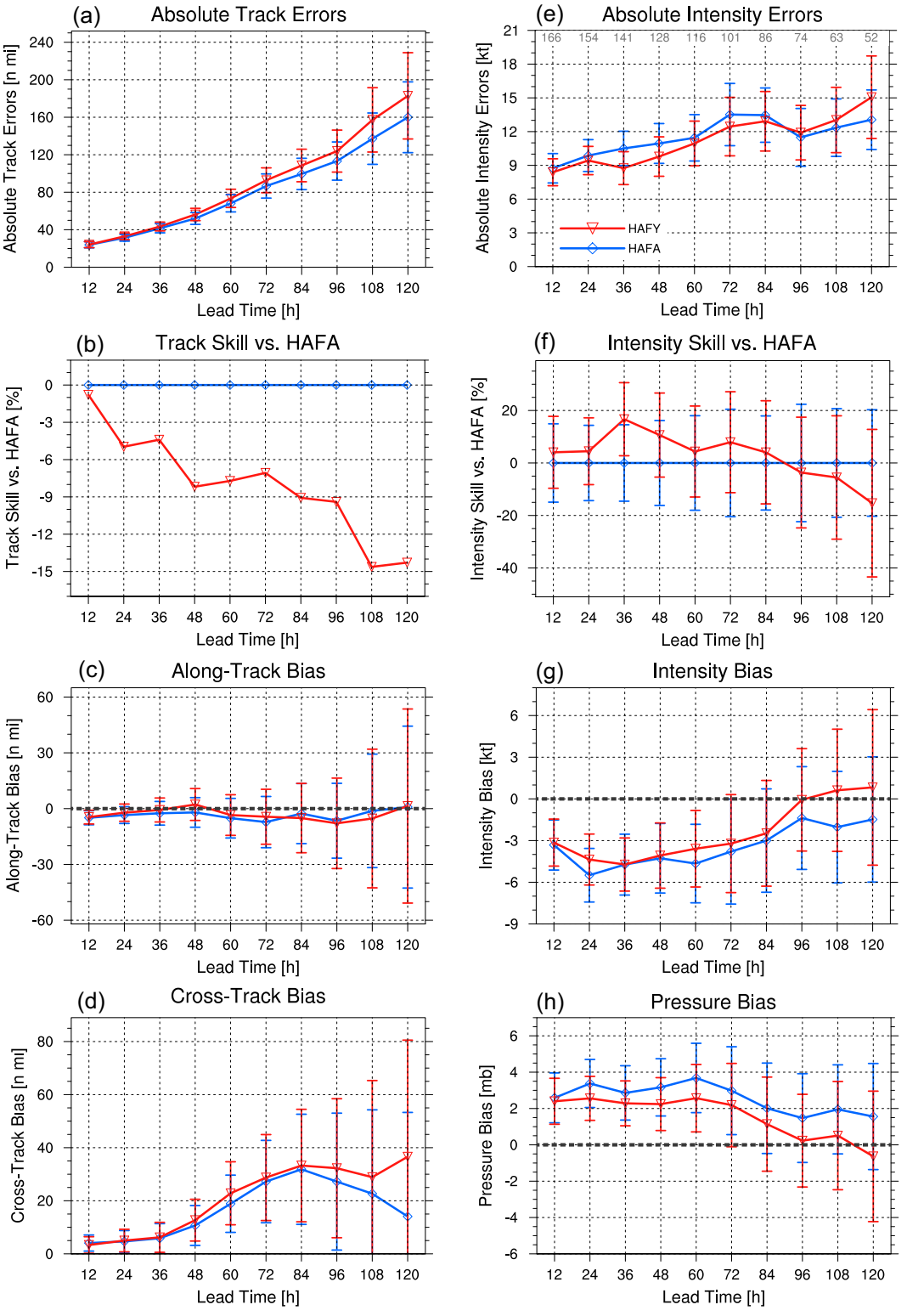

FIG. 1. Intensity and track errors for HAFA (blue) and HAFY (red) experiments, showing (a) track MAE (n mi), (b) relative track skill of HAFY, (c) along-track bias (n mi), (d) cross-track bias (n mi), (e) intensity MAE (kt), (f) relative intensity skill of HAFY, (g) intensity bias (kt), and (h) pressure bias (mb). Bars in MAE and bias panels indicate the 95% confidence intervals. Thick black dashed lines in the bias panels mark the zero value. The sample size at each lead time and the legend are shown in (b).

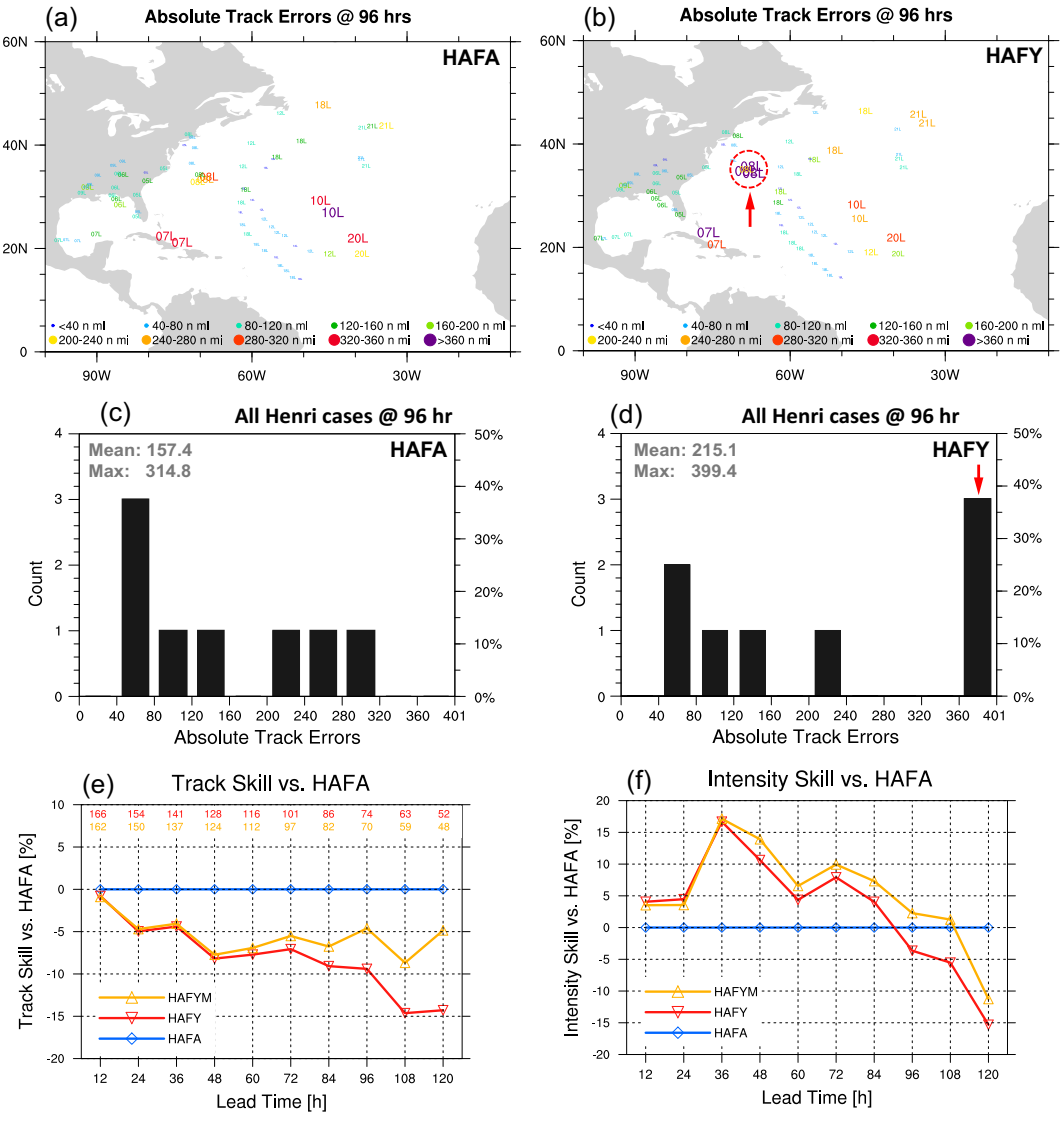

FIG. 2. (a) Spatial distribution of track MAE for all matched cycles at 96 h. Legend for the track MAE is shown at the bottom of (a) and (b); the larger the storm ID, the larger the track MAE. (c),(d) Histogram of the track MAE distribution for all Henri cycles at 96 h for HAFA and HAFY, respectively. Mean and maximum track MAE are provided in each panel. The red arrow or dashed circle in (b) and (d) denote large track MAE (>360 n mi) for Hurricane Henri (08L). (e) Track and (f) intensity skill of HAFY (red) and HAFYM (yellow, same as HAFY but excluding four cycles of Henri that have larger track MAE relative to HAFA (blue). The sample size of HAFY and HAFYM at each lead time is shown at the top of (e).

in the statistics and recomputed the track and intensity skill with a homogeneous sample size. To differentiate from HAFY, we term the HAFY experiment without the four cycles of Henri as HAFYM (yellow line in Figs. 2e,f). Interestingly, the trend of increasing degradation of track skill at longer lead times in HAFY disappears after removing the four cycles (Fig. 2e), and the mean degradation of track skill in HAFYM is 5.5% relative to HAFA. For reference, the mean degradation of track skill in HAFY is 8.1% relative to HAFA. The reduced degradation of track skill in HAFYM is associated with improved performance of intensity skill that extends to a longer lead time (i.e., t = 108 h, Fig. 2f). Compared to HAFA, the mean improvement in intensity skill before t = 108 h is 7.3% (4.8%) for HAFYM (HAFY). These results indicate that the mean error alone sometimes can be biased because of a few anomalous cases; thus, careful examination of the anomaly cases is needed to better understand the model performance and help to reduce future errors.

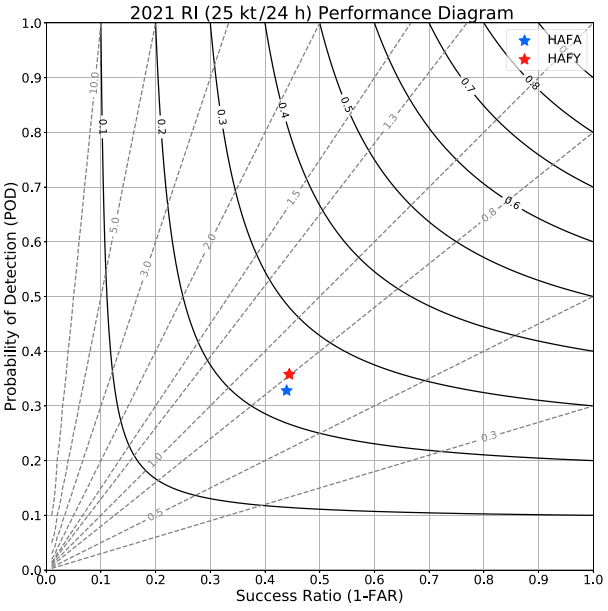

FIG. 3. (a) Performance diagram for RI forecasts for HAFA (blue) and HAFY (red) for the 25 kt per 24 h metric. The *x* axis denotes the success ratio (1 minus the false alarm ratio), and the *y* axis denotes the probability of RI detection. The dashed lines are the bias scores and the solid lines are the critical success index.

Since rapid intensification (RI) remains one of the most challenging TC intensity forecast issues (e.g., Fischer et al. 2019; Cangialosi et al. 2020), assessing the model performance on RI forecasts is important. Typically, the definition of RI is the increase of the VMAX exceeding 30 kt over 24 h, which approximately represents the 95th percentile of all 24-h over-water intensity changes for North Atlantic hurricanes (Kaplan and DeMaria 2003). By leveraging the uncertainties of the best track intensity noted by Torn and Snyder (2012), we use a RI definition of 25 kt per 24 h instead in this study. Of note, 25 kt per 24 h is one of the RI thresholds used in the Statistical Hurricane Intensity Prediction Scheme (SHIPS; DeMaria and Kaplan 1994; DeMaria et al. 2005). Figure 3 examines the RI forecast skill from the two experiments using the performance diagram introduced by Roebber (2009). Compared to HAFA, HAFY has a higher probability of detection (POD) and a slightly lower false alarm rate (FAR). Correspondingly, HAFY has a slightly smaller low bias and a higher critical success index than HAFA. Thus, it is safe to conclude that HAFY has slightly better RI forecast skill than HAFA. These findings together with Figs. 1e–h demonstrate the advantage of HAFY in TC intensity forecasts.

The performance of TC intensity forecasts is closely related to the structure forecasts. Figure 4 shows the MAE and bias of different structure metrics from HAFA and HAFY. HAFY generally has a smaller positive bias in RMW (Figs. 4a,b), R64 (Figs. 4c,d), and R34 (Figs. 4g,h) than HAFA. The behavior of the bias of R50 is different between the two experiments, with HAFA showing a marginally positive bias while HAFY showing a negative bias (Fig. 4f). Nevertheless, the magnitude of the R50 bias in HAFY is only ~2–3 n mi greater than that in HAFA and the R50 MAE in HAFY remains smaller than in HAFA (Fig. 4e). These findings demonstrate that HAFY has better skill in predicting TC structure than HAFA overall. The better TC structure in HAFY is an important contributor to its smaller weak bias (Figs. 1g,h) and better RI forecast skill (Fig. 3).

b. Comparison of vortex structure

To better understand the impact of PBL schemes on the vortex intensity and structure, this section examines a few more metrics for the two experiments. Figures 5a and 5b present the histogram of TC intensity and vortex depth[1] based on all the cases in each experiment. These cases are from the matched cycles documented in section 2b and have at least tropical depression intensity. There are 4203 cases for HAFA and 4000 cases for HAFY, and the differences in case number are attributable to the differences in landfall timing between the two experiments, as overland cases are excluded. Compared to HAFA, HAFY generally produces more major hurricanes (≥96 kt), especially within the intensity range of 105–130 kt (Fig. 5a). This finding suggests that the larger weak bias in HAFA (e.g., Fig. 1g) is associated with the poor forecasts of major hurricanes in HAFA. Probably due to a larger sample size of major hurricanes, HAFY produces more TCs with the vortex depth exceeding 13 km (Fig. 5b). In comparison, HAFA produces more TCs with a relatively shallower vortex depth (8–13 km). Figure 5c compares the histogram of the RMW at 2-km height from the two experiments. Both HAFY and HAFA have a peak of 35 km, and HAFY has a notable second peak of 55 km. Overall, HAFY has fewer cases with the 2-km RMW exceeding 90 km while generally having more cases with the RMW less than 90 km. This finding suggests that HAFY tends to produce a smaller RMW than HAFA, which is consistent with HAFY's smaller high bias in RMW (see Fig. 4b).

The composite boundary layer structure from HAFA and HAFY is examined in Fig. 6. In each panel, the radius is normalized by the RMW at 2-km height. Figures 6a–c compare the composite tangential wind (V_t) structure in the lowest 3 km between HAFA and HAFY. Consistent with previous results (Figs. 1g and 5a), the composite TC in HAFY generally has stronger V_t than in HAFA, with the greatest difference of ~3 m s^{-1} inside the RMW at 2-km height. In both experiments, maximum V_t occurs inside the RMW at a similar height. The HAFY TCs have shallower inflow layer depth, indicated by the contour of $V_r = -1$ m s^{-1}, with the greatest difference of 200–300 m within 1–3 × RMW (see red and blue dashed lines in Fig. 6f). Figures 6d–f also show that TCs in HAFY have stronger boundary layer inflow, especially radially inward of 3 × RMW. Another insightful measure of inflow strength is the inflow angle, defined as $\tan^{-1}(V_r/V_t)$. Figure 6g compares two experiment's inflow angles averaged

[1] Vortex depth is defined as the height where the tangential wind decays to 50% of its value at $z = 2$ km along the RMW (Hazelton et al. 2018).

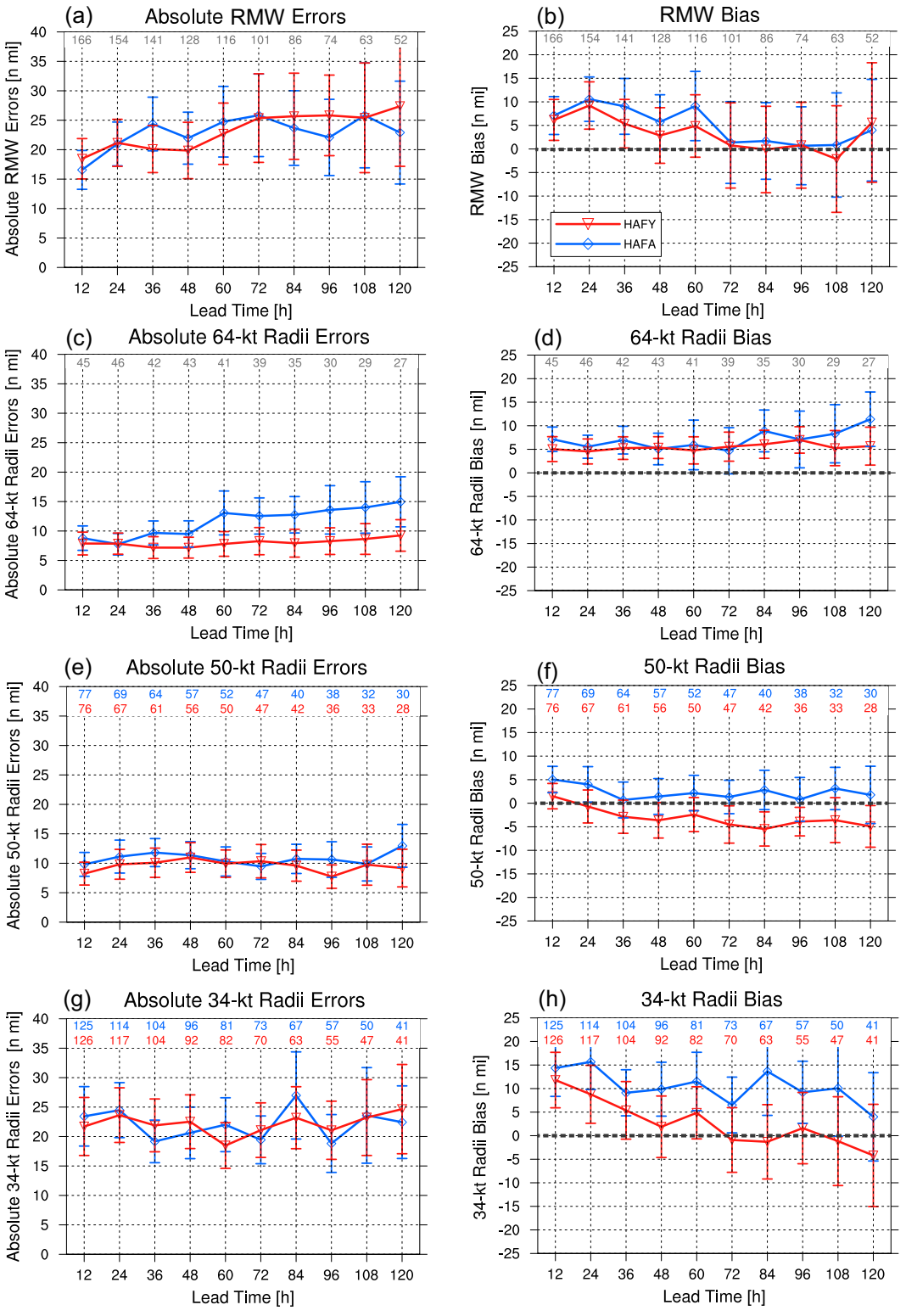

FIG. 4. Forecast MAE and bias of (a),(b) RMW; (c),(d) 64-kt wind radii; (e),(f) 50-kt wind radii; and (g),(h) 34-kt wind radii for HAFA (blue) and HAFY (red) experiments. The bars show the 95% confidence intervals. The gray numbers at the top of (a)–(d) show the uniform sample size between the two experiments at each lead time. The sample size differs in (e)–(h) and the blue and red numbers denote the sample size for HAFA and HAFY, respectively. Thick black dashed lines in the bias panels mark the zero value.

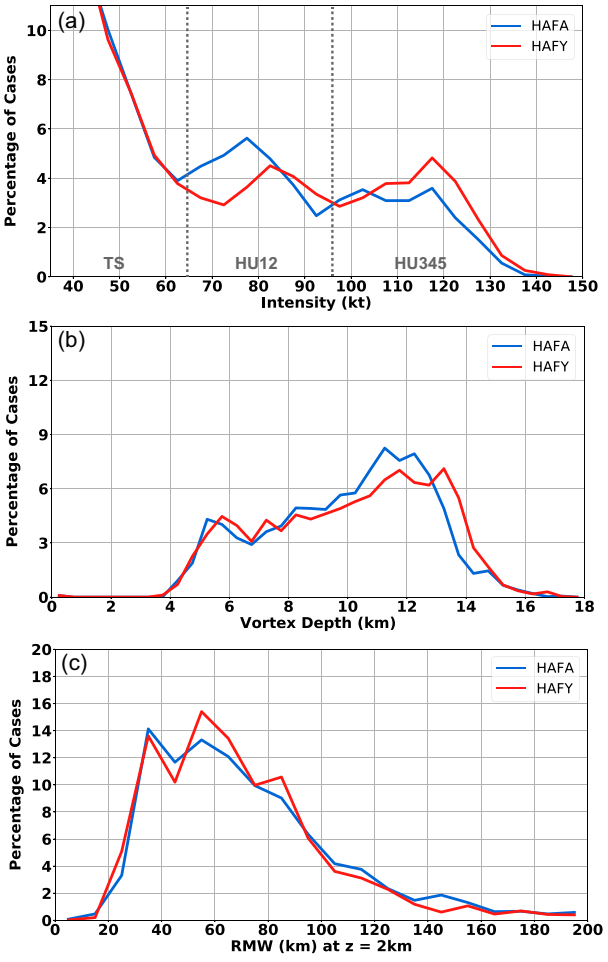

FIG. 5. (a) Percentage of cases stratified by (a) TC intensity, (b) vortex depth, and (c) RMW at $z = 2$ km for HAFA (blue) and HAFY (red). Percentage of cases is per bin widths of 5 kt in (a), 0.5 km in (b), and 5 km in (c). Vertical dashed lines in (a) delimit tropical storms (TS), category-1–2 hurricanes (HU12), and major hurricanes (HU345).

within the 10–500-m layer. Consistent with Fig. 6f, the boundary layer inflow is stronger in HAFY, especially within 1–3 × RMW. Differences in boundary layer inflow between HAFA and HAFY are in part attributed to the impact of reduced eddy viscosity with the modified EDMF-TKE on lowering inflow layer depth and also accelerating boundary layer inflow in the core region, which agrees well with the results from the idealized three-dimensional tests and ensemble HAFS forecasts of Hurricane Michael (2018) documented in Chen et al. (2022). Figures 6h and 6i further compare the radial profile of moist static energy (MSE) and specific humidity averaged within the boundary layer (<1 km). MSE is a useful variable to indicate the thermodynamic state within the core region (Chen et al. 2019; Fischer et al. 2023). Results demonstrate that TCs in HAFY also have higher MSE inside of 3 × RMW (Fig. 6h), which is largely attributable to higher specific humidity (Fig. 6i). The stronger boundary layer inflow with higher specific humidity in HAFY will lead to stronger moisture convergence in the eyewall region, which benefits the sustainment of eyewall convection and thereby supports stronger TC intensity, as seen in Fig. 5a.

To further understand the model performance compared to observations, Fig. 7 examines the 10-m inflow angle from both experiments and observations. The observed inflow angle is based on a dropsonde composite of category-1–5 hurricanes (Zhang and Uhlhorn 2012). To match the observations, only cases with hurricane intensity from both experiments are included, and they are normalized by the 10-m RMW (i.e., R^*). To remove the impact of landfalling processes on the surface inflow angle (Hlywiak and Nolan 2022), we exclude the cases whose distance to land is less than 5 times the RMW at $z = 2$ km. The number of cases included as well as the mean and median TC intensity from HAFA and HAFY are shown in Table 2. Both the median and mean TC intensity of the subset from HAFY are 2–4 m s^{-1} stronger than those from HAFA, and the differences are statistically significant at the 95% level based on a two-tailed Student's t test (not shown). The median TC intensity of the observations is 56.7 m s^{-1}, which is ~7 (11) m s^{-1} stronger than that in HAFY (HAFA). The case number at each radius from each experiment is over 1600, which is approximately a factor of 5–10 greater than the sample size of observations (see Fig. 2 in Zhang and Uhlhorn 2012).[2] Regardless of the differences in sample size, the inflow angle from both HAFA and HAFY are within 5° of the observations. Additionally, the 10-m inflow angle in HAFY remains larger than that in HAFA, and the radial location of the peak inflow angle in HAFY is closer to the observational value.

In summary, HAFY produces stronger, deeper, and more compact TC vortices than HAFA. This is closely related to the stronger, shallower, and more humid boundary layer inflow in the core region as a response to the reduced vertical turbulent mixing with the modified EDMF-TKE scheme.

4. Case studies for two hurricanes

The statistics and composite structure of TCs in section 3 provide an overview of the performance of the original and modified EDMF-TKE PBL schemes during the 2021 hurricane season. To illustrate how the impact of PBL schemes manifests in specific storms, we analyze Hurricanes Ida and Henri in this section. The former underwent a notable prelandfall RI and provides an excellent opportunity to examine the impact of PBL schemes on RI forecasts, while the latter significantly contributes to the larger track errors of HAFY (compared to HAFA) as discussed in section 3a.

a. *Hurricane Ida (2021)*

To closely examine the impact of the modified EDMF-TKE scheme on RI forecasts, we compare two sets of ensemble forecasts of Hurricane Ida (2021) from HAFA and

[2] The large sample size of hurricanes from HAFA and HAFY leads to a very small value of 95% confidence intervals (~0.1°) around the composite inflow angle, which are thereby not shown in Fig. 7.

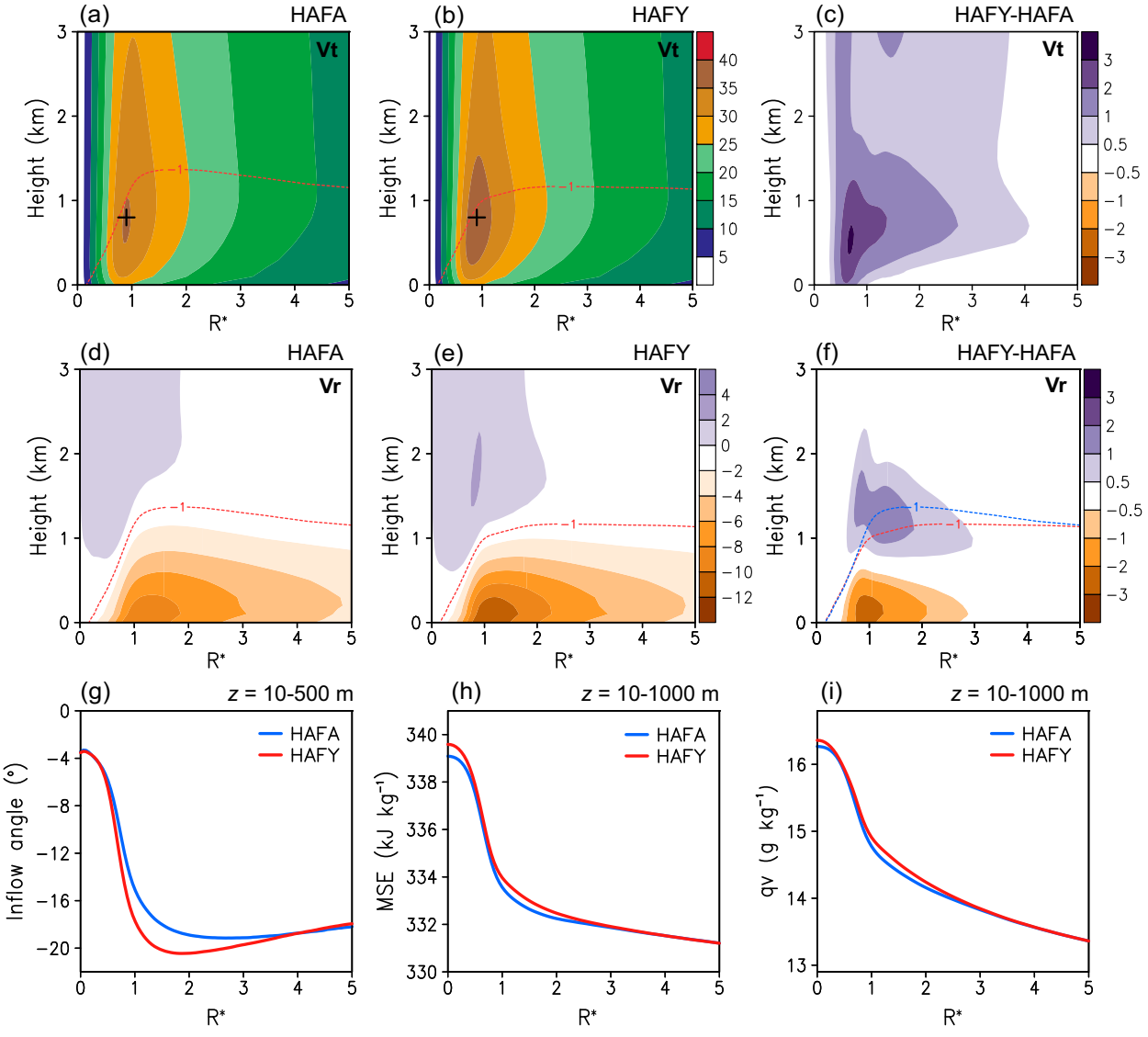

FIG. 6. (a),(b) Composite radius–height distribution of azimuthal-mean tangential wind (V_t; m s^{-1}) for HAFA and HAFY, respectively. (c) Differences in the distribution of V_t between the two experiments (i.e., HAFY − HAFA). (d)–(f) As in (a)–(c), but for radial wind (V_r). The red dashed line in (a), (b), (d), and (e) shows $V_r = -1$ m s^{-1}; red and blue dashed lines in (f) show $V_r = -1$ m s^{-1} from HAFY and HAFA, respectively. The black plus in (a) and (b) denotes the location of the maximum V_t. Composite radial profile of (g) inflow angle (°) averaged within the 10–500-m layer, and (h) MSE (kJ kg^{-1}) and (i) specific humidity (g kg^{-1}) averaged within 10–1000 m. Red and blue lines denote HAFY and HAFA experiments, respectively. The radius in each panel is normalized by the RMW at 2-km height.

HAFY. Each ensemble includes four members initialized at −12, −6, +6, and +12 h relative to 1200 UTC 28 August.[3] Of note, the RI period of Ida lasted from 0600 UTC 28 August to 1200 UTC 29 August according to the NHC best track data (gray line in Fig. 8b). Figure 8a shows that the track forecasts between the ensemble members of the two experiments are similar, with HAFA members initialized at −12 and −6 h relative to 1200 UTC 28 August having greater left-of-track errors as Ida moves inland. The averaged maximum TC intensity from HAFY ensemble members (red lines in Fig. 8b) is ~64 m s^{-1}, which is 4 m s^{-1} stronger than that from HAFA members (blue lines in Fig. 8b) and closer to the value of 67 m s^{-1} from the best track data. Additionally, the contraction of RMW is more notable in HAFY members, which accounts for a smaller RMW before TC landfall ($t \approx 41$ h) that agrees better with the best track data (Fig. 8c).

Additionally, we verify the TC structure before landfall by selecting a representative ensemble member initialized at 0600 UTC 28 August (i.e., the time of RI onset according to the NHC best track) from each experiment (thick lines in

[3] The cycle of 1200 UTC 28 August is not included due to a forecast failure caused by unknown HPC issues in HAFY.

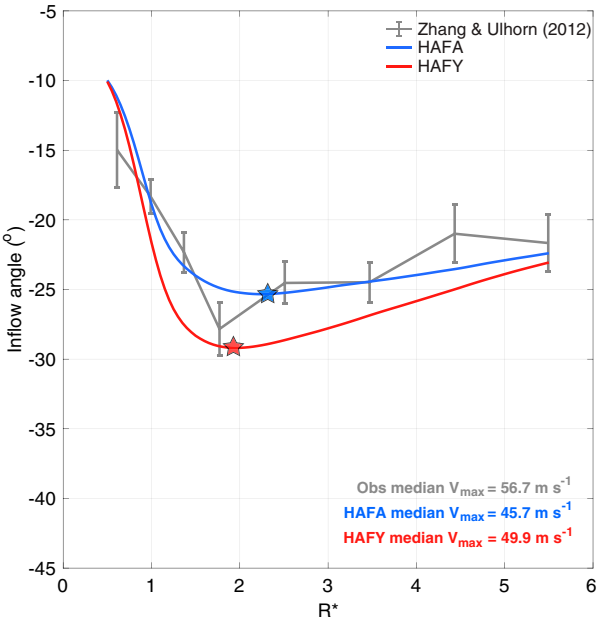

FIG. 7. The composite 10-m radial profile of inflow angle as a function of normalized R^* for hurricane cases from HAFA (blue) and HAFY (red). Stars mark the location of the peak inflow angle. The gray line denotes the 10-m radial profile of inflow angle based on a dropsonde composite of category-1–5 hurricanes (Zhang and Uhlhorn 2012). The gray bar denotes the 95% confidence intervals. The R^* from models is normalized by 10-m RMW to match the observations.

Figs. 8b,c) and comparing the composite TC structure over t = 33–39 h against the tail Doppler radar (TDR) merged analysis of the 20210829I1 mission in Ida (2021) (Fig. 9). The period of the TDR merged analysis is from 0835 to 1455 UTC 29 August (i.e., $t \approx$ 33–39 h). Comparison of the storm-centered, Earth-relative horizontal wind speed at z = 2 km in Figs. 9a–c shows that both HAFA and HAFY encouragingly reproduce the observed wavenumber-1 structure in the horizontal wind speed (with stronger wind speeds to the right of storm motion) and capture the observed radius of hurricane-force wind (see yellow shading) in different quadrants. The structure difference between HAFA and HAFY occurs mostly in the inner-core region: while both experiments produce greater RMW than the TDR analysis, the high bias in RMW is smaller in HAFY than in HAFA (see RMW values in Figs. 9a–c); additionally, the smaller eye in HAFY is comparable to the TDR analysis. Comparison of the azimuthal-mean tangential wind in Figs. 9d–f further shows that HAFY produces stronger peak tangential wind (>60 m s^{-1}) in the lower boundary layer of the eyewall and smaller RMW below 12-km height than HAFA, both of which match better to the TDR analysis. HAFY also produces a deeper TC vortex than HAFA, with strong tangential winds of >32 m s^{-1} in HAFY extending to a deeper layer, which is closer to the TDR analysis too. The stronger, deeper, and more compact TCs in HAFY members are closely related to boundary layer processes, which will be discussed shortly. Overall, the structure-based verification results demonstrate that HAFY produces more realistic TC intensity and structure before Ida's landfall.

TABLE 2. Mean and median TC intensity (m s^{-1}) as well as case number for the composite of hurricanes from HAFA and HAFY.

	Mean value	Median	Case No.
HAFA	47.5	45.7	1717
HAFY	50.2	49.9	1667

Figure 10 compares the radial inflow structure averaged over t = 18–24 h from the same ensemble member. The analysis period immediately follows the bifurcation point for the intensity evolution (see the coral arrow in Fig. 8b). The boundary layer inflow in HAFY is much stronger than that in HAFA, with the maximum difference of the inflow strength being >6 m s^{-1} inside of 4 × RMW (where RMW is indicated by the black line in Fig. 10). This finding is consistent with the composite structure of inflow strength or inflow angle discussed in section 3b and supports the results from Chen et al. (2022). The inflow layer depth is comparable between the two members over the analysis period, which is seemingly different from the composite results in Fig. 6f. The "unexpected" deeper inflow near the core in HAFY is likely a response to the stronger eyewall diabatic heating over the composite period, which can be inferred from a broader annulus of eyewall updraft in HAFY than in HAFA (cf. red contours in Figs. 10a,b). Outside the eyewall region, the inflow layer depth in both members is within 1–1.5 km. In HAFY, the stronger boundary layer inflow and the resulting enhanced boundary layer moisture convergence beneath the eyewall can support stronger eyewall convective updrafts. The stronger boundary layer inflow together with more eyewall diabatic heating benefits the radial advection of large absolute angular momentum (AAM) toward the TC center, which is favorable to TC intensification and RMW contraction (e.g., Smith and Montgomery 2015; Chen and Bryan 2021) and accounts for the stronger TC intensity and smaller RMW in HAFY than in HAFA. Meanwhile, stronger eyewall updrafts in HAFY lead to stronger upward advection of large AAM in the eyewall region (Chen et al. 2017), which accounts for the deeper vortex structure in HAFY (Fig. 9f).

b. *Hurricane Henri (2021)*

We discussed earlier in section 3a that HAFY has noticeably larger track errors than HAFA for Henri, with the track error of a few cycles in HAFY being greater than 360 n mi at 96 h. Examination of the track of all matched cycles of Henri from HAFA and HAFY in Figs. 11a and 11b together with the histogram of track errors in Figs. 2c and 2d indicates that 4 early cycles[4] (blue and green lines in Figs. 11a,b) are the main contributors to the larger track errors in HAFY. Figures 12a–c examine the track MAE and bias for Henri from HAFA and HAFY. Results show that track differences between HAFA and HAFY rapidly increase after t = 72 h (Fig. 12a), which is

[4] These four early cycles of Henri were initialized over the period from 0000 UTC 16 August to 1200 UTC 17 August 2021.

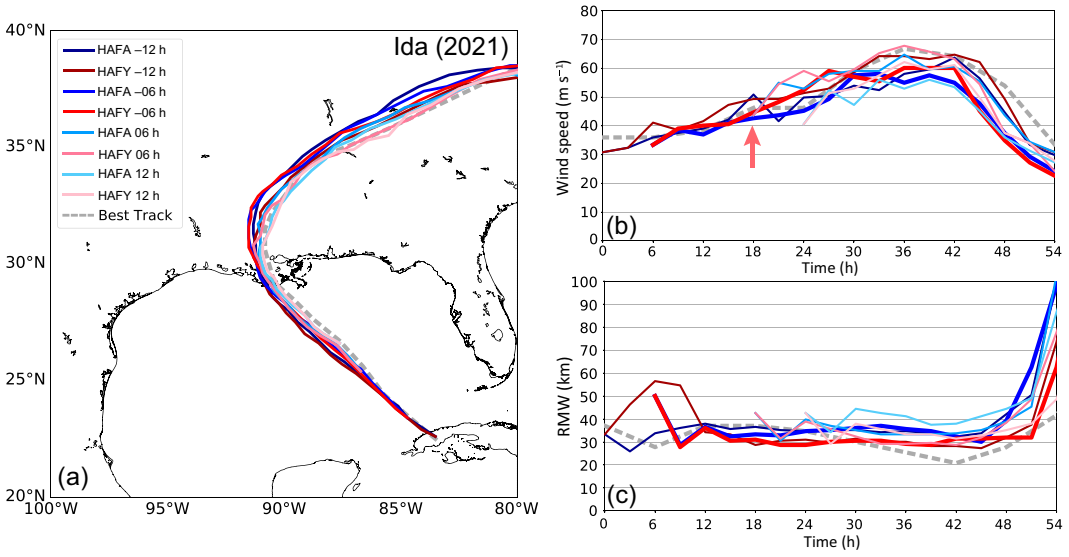

FIG. 8. Comparison of 4-member ensemble HAFA and HAFY forecasts of Hurricane Ida (2021) initialized at −12, −6, +6, and +12 h relative to 1200 UTC 28 Aug, respectively, showing the evolution of (a) TC track, (b) 10-m maximum wind (m s^{-1}), and (c) RMW (km). The dashed gray line denotes the best track; blue and red lines denote HAFA and HAFY experiments, respectively. Thick red and blue lines in (b) and (c) denote the forecasts initialized at 0600 UTC 28 Aug, and the coral arrow in (b) denotes the bifurcation point of the two forecasts.

mostly attributable to greater right-of-track bias in HAFY (Fig. 12c), as track deflection in those early cycles of Henri in HAFY occurs much earlier and the simulated storm subsequently moves much faster toward the northeast (Figs. 11a,b). While there are potentially multiple factors causing the larger track error of Henri in HAFY, one factor we focused on is the cumulus scheme, which was motivated by two considerations. First, PBL schemes affect boundary layer instability and indirectly impact the mass-flux calculation within the cumulus scheme, as the updraft in the SAS cumulus scheme starts from the level of maximum MSE in the boundary layer. In addition, an impact of the cumulus scheme on the TC track was noted in previous studies (e.g., Ma and Tan 2009; Sun et al. 2014; Bassill 2014).

To investigate the impact of the cumulus scheme on the track forecasts of Henri, we ran additional two sets of experiments for Henri based on HAFY, one turning off both SAS shallow and deep cumulus schemes (experiment HFY2) and the other only turning off the SAS shallow cumulus scheme (experiment HFY3). Figures 11c and 11d show the tracks of all cycles of Henri from HFY2 and HFY3, respectively. Comparison of Figs. 11c, 11d, and 11b indicates that the right-of-track bias in the early cycles of Henri is substantially reduced in both HFY2 and HFY3. This finding is supported by the cross-track bias in Fig. 12c, as at $t = 120$ h the cross-track bias for HFY3 is reduced to half compared to HAFY and the cross-track bias for HFY2 becomes minimal. A similar reduction of the along-track bias at longer lead times is found in HFY2 and HFY3 (Fig. 12b), as the storm moves slower without a shallow cumulus scheme or both cumulus schemes. Since both cross-track and along-track biases are reduced, it is not surprising to find that the track MAE is reduced in these experiments (Fig. 12a). The track MAE for HFY3 is very similar to HAFA, while the track MAE for HFY2 is the smallest among the four experiments, which is <100 n mi before $t = 96$ h. Importantly, the intensity forecasts are impacted by cumulus schemes too (Figs. 12d–f). While the VMAX MAE is similar between HFY3 and HAFY, the VMAX MAE for HFY2 is reduced by 10 kt at longer lead times compared to HAFY (Fig. 12d). In comparison, the weak bias in terms of VMAX (PMIN) in both HFY2 and HFY3 is reduced by ~10 kt (~8 mb; 1 mb = 1 hPa) over days 4–5 compared to HAFY (Figs. 12e,f). The consistently reduced weak bias in HFY2 is closely related to its improved track skill, as TCs in HFY2 have notably reduced right-of-track bias and stay over the region with relatively higher ocean heat content (OHC) (not shown). Comparison of the results from HFY2 and HFY3 indicates that the deep cumulus scheme affects both intensity and track errors.

In summary, the above analysis demonstrates the significant impact of cumulus schemes on Henri's intensity and track forecasts. Further examination of the cumulus schemes including the parameterizations of different physical processes and scale-aware effects is needed, which is beyond the scope of this study and will be left for an upcoming study.

5. Conclusions and discussion

A new turbulence kinetic energy (TKE)-based moist eddy-diffusivity mass-flux (EDMF-TKE) planetary boundary layer scheme has been implemented in NOAA's next-generation hurricane forecast model, the Hurricane Analysis and Forecast System (HAFS). Using a recently developed modeling framework based on large-eddy simulation (LES) (Chen et al. 2021a),

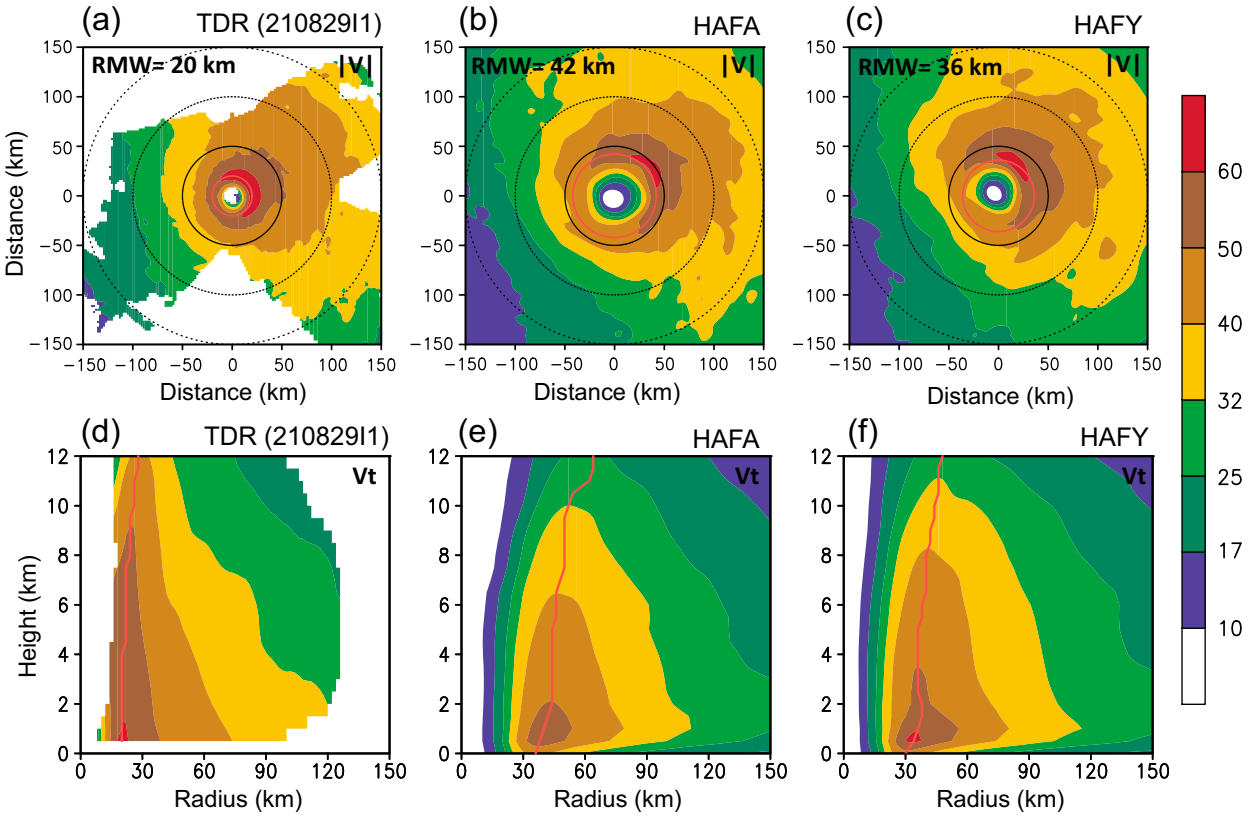

FIG. 9. (a)–(c) Plan view of storm-centered, Earth-relative horizontal wind speed (shading; m s^{-1}) at $z = 2$ km from the merged TDR analysis of the 20210829I1 mission in Hurricane Ida (2021), HAFA, and HAFY, respectively. (d)–(f) As in (a)–(c), but for azimuthal-mean tangential wind (shading; m s^{-1}). The HAFA and HAFY results are averaged over $t = 33$–39 h for the ensemble member initialized at 0600 UTC 28 Aug. Black circles in (a)–(c) denote 50-, 100-, and 150-km radii. The red circle and red line denote the RMW.

the EDMF-TKE scheme has been optimized for hurricane conditions (Chen et al. 2022). This follow-up study evaluates the performance of the modified EDMF-TKE PBL scheme during 2021 North Atlantic hurricane forecasts against the original EDMF-TKE from the regional stand-alone HAFS. Based on a large sample of cases, this study examines the impact of PBL schemes on TC forecast metrics, inner-core structure, as well as boundary layer structure and thermodynamics. The key findings based on statistics, composite analyses, and case studies are summarized below:

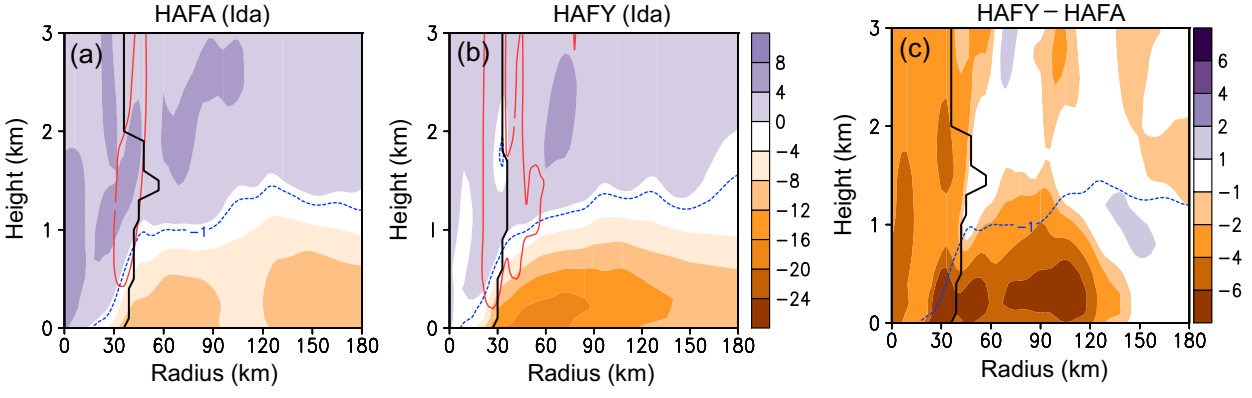

FIG. 10. The radial–height plot of azimuthally averaged radial wind (shading; m s^{-1}) averaged over $t = 18$–24 h for the (a) HAFA and (b) HAFY forecasts of Hurricane Ida (2021) initialized at 0600 UTC 28 Aug. (c) The difference in radial winds (i.e., HAFY − HAFA). In (a) and (b), the dashed blue line denotes $V_r = -1$ m s^{-1}. The red contour in (a) and (b) denotes $w = 0.3$ m s^{-1} and the black line denotes the mean RMW. The RMW and the contour of $V_r = -1$ m s^{-1} in (c) are from HAFA.

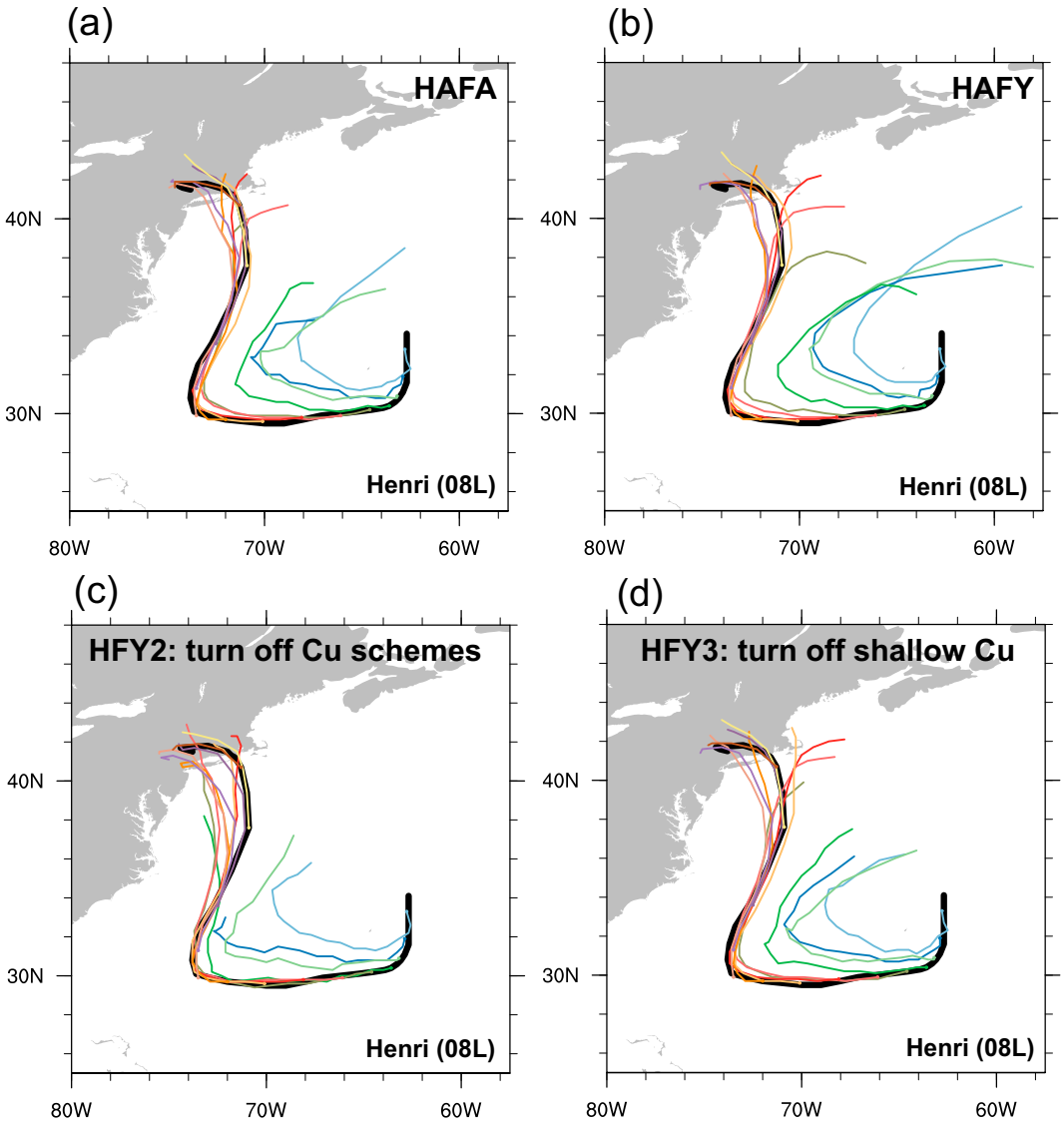

FIG. 11. (a)–(d) Track plots of all matched cycles of Henri from HAFA, HAFY, and HAFY-based sensitivity experiments HAY2 and HAY3, respectively. Different colors represent different forecast cycles. The early four cycles are highlighted by blue and green colors.

1) Compared to the original EDMF-TKE, the modified EDMF-TKE improves the forecast skill of intensity and structure overall; specifically, the modified EDMF-TKE has better rapid intensification (RI) forecast skill due to a higher probability of detection and a slightly lower false alarm rate.
2) The modified EDMF-TKE slightly reduces the track skill. The slightly larger track error with the modified EDMF-TKE mainly comes from the larger cross-track bias, which is in part attributable to a few early cases of Hurricane Henri.
3) The modified EDMF-TKE forecasts produce stronger, deeper, and more compact TC vortices.
4) Consistent with the reduced eddy viscosity (see details in Chen et al. 2022), the modified EDMF-TKE forecasts produce accelerated boundary layer inflow outside the radius of the maximum wind, as seen in the composite analysis and the RI case of Hurricane Ida. The enhanced boundary layer inflow and the associated enhanced moisture convergence beneath the eyewall can support stronger eyewall convection.
5) In both sets of forecasts, the magnitude of inflow angle for hurricanes is comparable to the dropsonde composite to within 5°, and the inflow angle in the modified EDMF-TKE forecasts peaks at a smaller radius that is closer to observations.

These findings demonstrate encouraging improvements of model forecast skill in TC intensity and structure due to the improvement of the EDMF-TKE scheme. There is still room for the improvement of track forecasts using this modified

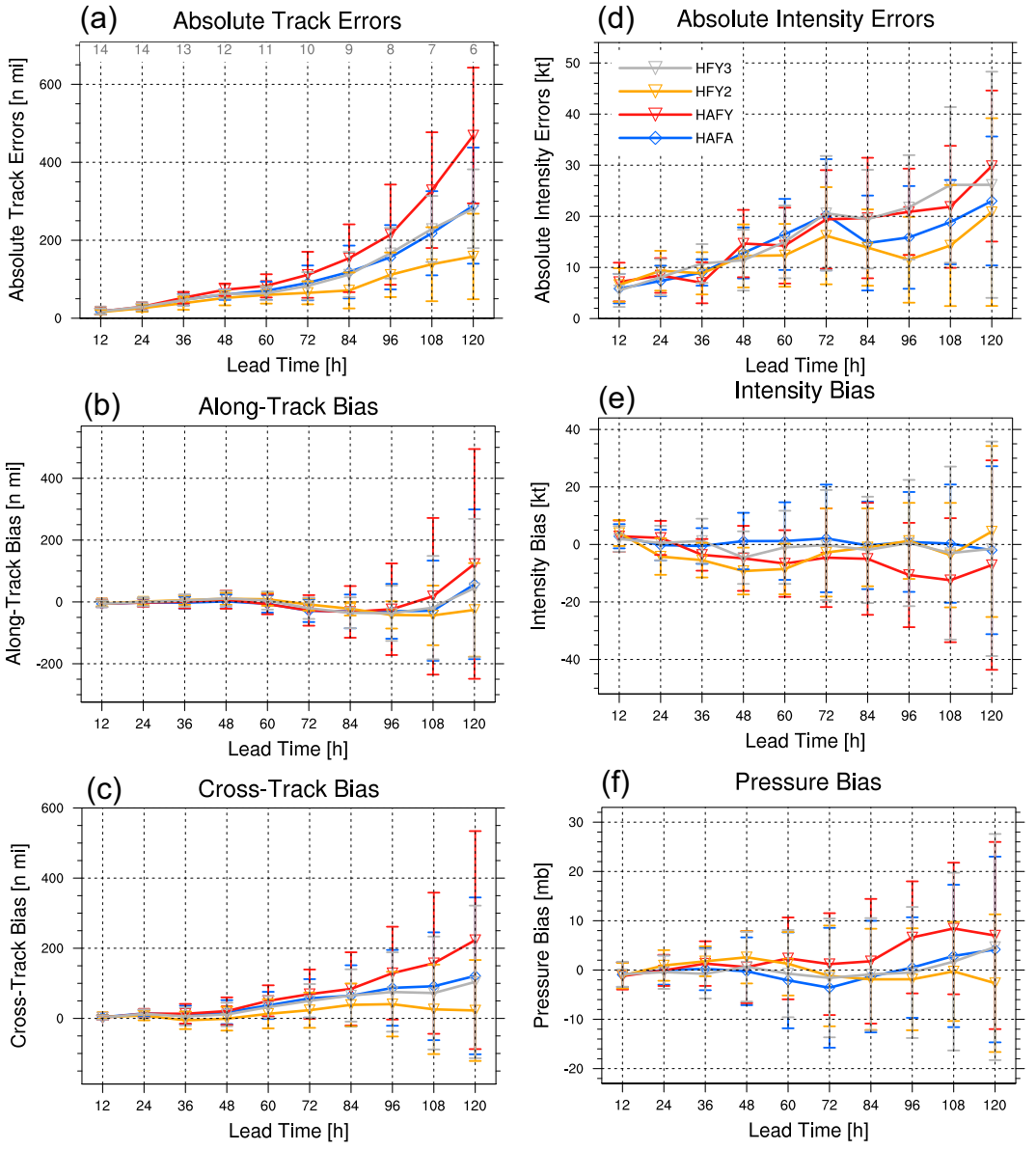

FIG. 12. As in Fig. 1, but for all matched cycles of Henri only. The sample size at each lead time is shown at the top of (a).

EDMF-TKE scheme. One identified key factor affecting track forecasts is the cumulus scheme. Preliminary results in this study demonstrate that the track error of Henri is linked closely to both the shallow and deep cumulus schemes, as turning off cumulus schemes leads to notably improved track performance. Additionally, the intensity error for Henri is also closely related to the deep cumulus scheme. Further investigation into the physics parameterizations within cumulus schemes that affected the forecasts of Henri is ongoing.

The identified impact of cumulus schemes on the model track performance is aligned with the findings in previous studies (e.g., Ma and Tan 2009; Sun et al. 2014; Bassill 2014), although those studies mainly focused on the horizontal grid spacings of $O(10)$ km that are well beyond the convective gray zone (Hong and Dudhia 2012). The 3-km (HAFS v0.2 in 2021) and 2-km horizontal grid spacing (HAFS v0.3 in 2022) are well within the convective gray zone, where explicit model dynamics can only partly resolve convective features that are parameterized at coarser scales. Given this, special attention is needed to investigate the effect of cumulus schemes in the gray zone. Additionally, mass-flux profiles in the SAS cumulus schemes depend on the highest-MSE parcels within the boundary layer, while in TCs these parcels are typically near the surface; thus, cumulus mass flux becomes analogous to the surface-driven mass flux in the EDMF-TKE PBL scheme under TC conditions. Efforts to unify the mass-flux parameterizations between PBL and cumulus schemes are encouraged for future model physics development for hurricane applications.

Last, the verification of inflow angle also highlights the need to collect more boundary layer data in the future.

Acknowledgments. We would like to acknowledge Dr. Jun Zhang for sharing the inflow angle data and Dr. Bin Liu (EMC) for the help on the model setup. This project benefits from the discussion with Drs. Jason Sippel and Michael Fischer. The authors are grateful for the constructive suggestions and comments from Drs. Sarah Ditchek, Trey Alvey, and two anonymous reviewers that help improve the clarity and presentation of the manuscript. Xiaomin Chen is supported by the Award NA21OAR4320190 to the Northern Gulf Institute at Mississippi State University from NOAA's Office of Oceanic and Atmospheric Research, U.S. Department of Commerce. Andrew Hazelton is supported by NOAA Grant NA19OAR0220187.

Data availability statement. Datasets for the stand-alone-regional HAFS cycles in the 2021 hurricane season are available on the NOAA RDHPCS computer system, or by request to Andrew Hazelton (Andrew.Hazelton@noaa.gov).

REFERENCES

Bassill, N. P., 2014: Accuracy of early GFS and ECMWF Sandy (2012) track forecasts: Evidence for a dependence on cumulus parameterization. *Geophys. Res. Lett.*, **41**, 3274–3281, https://doi.org/10.1002/2014GL059839.

Bleck, R., 2002: An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates. *Ocean Modell.*, **4**, 55–88, https://doi.org/10.1016/S1463-5003(01)00012-9.

Braun, S. A., and W.-K. Tao, 2000: Sensitivity of high-resolution simulations of Hurricane Bob (1991) to planetary boundary layer parameterizations. *Mon. Wea. Rev.*, **128**, 3941–3961, https://doi.org/10.1175/1520-0493(2000)129<3941:SOHRSO>2.0.CO;2.

Bryan, G. H., 2012: Effects of surface exchange coefficients and turbulence length scales on the intensity and structure of numerically simulated hurricanes. *Mon. Wea. Rev.*, **140**, 1125–1143, https://doi.org/10.1175/MWR-D-11-00231.1.

——, R. P. Worsnop, J. K. Lundquist, and J. A. Zhang, 2017: A simple method for simulating wind profiles in the boundary layer of tropical cyclones. *Bound.-Layer Meteor.*, **162**, 475–502, https://doi.org/10.1007/s10546-016-0207-0.

Bu, Y. P., R. G. Fovell, and K. L. Corbosiero, 2017: The influences of boundary layer mixing and cloud-radiative forcing on tropical cyclone size. *J. Atmos. Sci.*, **74**, 1273–1292, https://doi.org/10.1175/JAS-D-16-0231.1.

Cangialosi, J. P., E. Blake, M. DeMaria, A. Penny, A. Latto, E. Rappaport, and V. Tallapragada, 2020: Recent progress in tropical cyclone intensity forecasting at the National Hurricane Center. *Wea. Forecasting*, **35**, 1913–1922, https://doi.org/10.1175/WAF-D-20-0059.1.

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.

Chen, X., 2022: How do planetary boundary layer schemes perform in hurricane conditions: A comparison with large-eddy simulations. *J. Adv. Model. Earth Syst.*, **14**, e2022MS003088, https://doi.org/10.1029/2022MS003088.

——, and G. H. Bryan, 2021: Role of advection of parameterized turbulence kinetic energy in idealized tropical cyclone simulations. *J. Atmos. Sci.*, **78**, 3593–3611, https://doi.org/10.1175/JAS-D-21-0088.1.

——, Y. Wang, K. Zhao, and D. Wu, 2017: A numerical study on rapid intensification of Typhoon Vicente (2012) in the South China Sea. Part I: Verification of simulation, storm-scale evolution, and environmental contribution. *Mon. Wea. Rev.*, **145**, 877–898, https://doi.org/10.1175/MWR-D-16-0147.1.

——, J. A. Zhang, and F. D. Marks, 2019: A thermodynamic pathway leading to rapid intensification of tropical cyclones in shear. *Geophys. Res. Lett.*, **46**, 9241–9251, https://doi.org/10.1029/2019GL083667.

——, G. H. Bryan, J. A. Zhang, J. J. Cione, and F. D. Marks, 2021a: A framework for simulating the tropical cyclone boundary layer using large-eddy simulation and its use in evaluating PBL parameterizations. *J. Atmos. Sci.*, **78**, 3559–3574, https://doi.org/10.1175/JAS-D-20-0227.1.

——, M. Xue, B. Zhou, J. Fang, J. A. Zhang, and F. D. Marks, 2021b: Effect of scale-aware planetary boundary layer schemes on tropical cyclone intensification and structural changes in the gray zone. *Mon. Wea. Rev.*, **149**, 2079–2095, https://doi.org/10.1175/MWR-D-20-0297.1.

——, G. H. Bryan, A. Hazelton, F. D. Marks, and P. Fitzpatrick, 2022: Evaluation and improvement of a TKE-based eddy-diffusivity mass-flux (EDMF) planetary boundary layer scheme in hurricane conditions. *Wea. Forecasting*, **37**, 935–951, https://doi.org/10.1175/WAF-D-21-0168.1.

DeMaria, M., 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.

——, M. Mainelli, L. K. Shay, J. A. Knaff, and J. Kaplan, 2005: Further improvements to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Wea. Forecasting*, **20**, 531–543, https://doi.org/10.1175/WAF862.1.

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.

Fischer, M. S., B. H. Tang, and K. L. Corbosiero, 2019: A climatological analysis of tropical cyclone rapid intensification in environments of upper-tropospheric troughs. *Mon. Wea. Rev.*, **147**, 3693–3719, https://doi.org/10.1175/MWR-D-19-0013.1.

——, P. D. Reasor, R. F. Rogers, and J. F. Gamache, 2022: An analysis of tropical cyclone vortex and convective characteristics in relation to storm intensity using a novel airborne Doppler radar database. *Mon. Wea. Rev.*, **150**, 2255–2278, https://doi.org/10.1175/MWR-D-21-0223.1.

——, ——, B. H. Tang, K. L. Corbosiero, R. D. Torn, and X. Chen, 2023: A tale of two vortex evolutions: Using a high-resolution ensemble to assess the impacts of ventilation on a tropical cyclone rapid intensification event. *Mon. Wea. Rev.*, **151**, 297–320, https://doi.org/10.1175/MWR-D-22-0037.1.

Franklin, J. L., 2008: National Hurricane Center forecast verification. *28th Conf. on Hurricanes and Tropical Meteorology*, Orlando, FL, Amer. Meteor. Soc., 13A.1, https://ams.confex.com/ams/28Hurricanes/webprogram/Paper137870.html.

Gopalakrishnan, S. G., F. Marks, 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.

Han, J., and C. S. Bretherton, 2019: TKE-based moist eddy-diffusivity mass-flux (EDMF) parameterization for vertical turbulent mixing. *Wea. Forecasting*, **34**, 869–886, https://doi.org/10.1175/WAF-D-18-0146.1.

——, W. Wang, Y. C. Kwon, S.-Y. 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.

Hazelton, A. T., 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.

——, J. A. Zhang, and S. Gopalakrishnan, 2022: Comparison of the performance of the observation-based hybrid EDMF and EDMF-TKE PBL schemes in 2020 tropical cyclone forecasts from the global-nested Hurricane Analysis and Forecast System. *Wea. Forecasting*, **37**, 457–476, https://doi.org/10.1175/WAF-D-21-0124.1.

Hill, K. A., and G. M. Lackmann, 2009: Analysis of idealized tropical cyclone simulations using the Weather Research and Forecasting Model: Sensitivity to turbulence parameterization and grid spacing. *Mon. Wea. Rev.*, **137**, 745–765, https://doi.org/10.1175/2008MWR2220.1.

Hlywiak, J., and D. S. Nolan, 2022: The evolution of asymmetries in the tropical cyclone boundary layer wind field during landfall. *Mon. Wea. Rev.*, **150**, 529–549, https://doi.org/10.1175/MWR-D-21-0191.1.

Hong, S.-Y., and J. Dudhia, 2012: Next-generation numerical weather prediction: Bridging parameterization, explicit clouds, and large eddies. *Bull. Amer. Meteor. Soc.*, **93**, ES6–ES9, https://doi.org/10.1175/2011BAMS3224.1.

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.

Kepert, J. D., 2012: Choosing a boundary layer parameterization for tropical cyclone modeling. *Mon. Wea. Rev.*, **140**, 1427–1445, https://doi.org/10.1175/MWR-D-11-00217.1.

Ma, L.-M., and Z.-M. Tan, 2009: Improving the behavior of the cumulus parameterization for tropical cyclone prediction: Convection trigger. *Atmos. Res.*, **92**, 190–211, https://doi.org/10.1016/j.atmosres.2008.09.022.

Nolan, D. S., J. A. Zhang, and D. P. Stern, 2009a: Evaluation of planetary boundary layer parameterizations in tropical cyclones by comparison of in situ observations and high-resolution simulations of Hurricane Isabel (2003). Part I: Initialization, maximum winds, and the outer-core boundary layer. *Mon. Wea. Rev.*, **137**, 3651–3674, https://doi.org/10.1175/2009MWR2785.1.

——, D. P. Stern, and J. A. Zhang, 2009b: Evaluation of planetary boundary layer parameterizations in tropical cyclones by comparison of in situ observations and high-resolution simulations of Hurricane Isabel (2003). Part II: Inner-core boundary layer and eyewall structure. *Mon. Wea. Rev.*, **137**, 3675–3698, https://doi.org/10.1175/2009MWR2786.1.

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

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

Sippel, J. A., X. Wu, S. D. Ditchek, V. Tallapragada, and D. Kleist, 2022: Impacts of assimilating additional reconnaissance data on operational GFS tropical cyclone forecasts. *Wea. Forecasting*, **37**, 1615–1639, https://doi.org/10.1175/WAF-D-22-0058.1.

Smith, R. K., and G. L. Thomsen, 2010: Dependence of tropical-cyclone intensification on the boundary-layer representation in a numerical model. *Quart. J. Roy. Meteor. Soc.*, **136**, 1671–1685, https://doi.org/10.1002/qj.687.

——, and M. T. Montgomery, 2015: Toward clarity on understanding tropical cyclone intensification. *J. Atmos. Sci.*, **72**, 3020–3031, https://doi.org/10.1175/JAS-D-15-0017.1.

Sun, Y., Z. Zhong, W. Lu, and Y. Hu, 2014: Why are tropical cyclone tracks over the western North Pacific sensitive to the cumulus parameterization scheme in regional climate modeling? A case study for Megi (2010). *Mon. Wea. Rev.*, **142**, 1240–1249, https://doi.org/10.1175/MWR-D-13-00232.1.

Torn, R. D., and C. Snyder, 2012: Uncertainty of tropical cyclone best-track information. *Wea. Forecasting*, **27**, 715–729, https://doi.org/10.1175/WAF-D-11-00085.1.

Zhang, F., and Z. Pu, 2017: Effects of vertical eddy diffusivity parameterization on the evolution of landfalling hurricanes. *J. Atmos. Sci.*, **74**, 1879–1905, https://doi.org/10.1175/JAS-D-16-0214.1.

Zhang, J. A., and E. W. Uhlhorn, 2012: Hurricane sea surface inflow angle and an observation-based parametric model. *Mon. Wea. Rev.*, **140**, 3587–3605, https://doi.org/10.1175/MWR-D-11-00339.1.

——, D. S. Nolan, R. F. Rogers, and V. Tallapragada, 2015: Evaluating the impact of improvements in the boundary layer parameterization on hurricane intensity and structure forecasts in HWRF. *Mon. Wea. Rev.*, **143**, 3136–3155, https://doi.org/10.1175/MWR-D-14-00339.1.