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

ANDREW HAZELTON,[a,b] JUN A. ZHANG,[a,b] AND SUNDARARAMAN GOPALAKRISHNAN[b]

[a] *Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, Florida*
[b] *NOAA/AOML/HRD, Miami, Florida*

(Manuscript received 26 July 2021, in final form 3 February 2022)

ABSTRACT: Better representation of the planetary boundary layer (PBL) in numerical models is one of the keys to improving forecasts of TC structure and intensity, including rapid intensification. To meet this goal, our recent work has used observations to improve the eddy-diffusivity mass flux with prognostic turbulent kinetic energy (EDMF-TKE) PBL scheme in the Hurricane Analysis and Forecast System (HAFS). This study builds on that work by comparing a modified version of EDMF-TKE (MEDMF-TKE) with the hybrid EDMF scheme based on a *K*-profile method (HEDMF-KP) in the 2020 HAFS-globalnest model. Verification statistics based on 101 cases in the 2020 season demonstrate that MEDMF-TKE improves track forecasts, with a reduction in a large right bias seen in HEDMF-KP forecasts. The comparison of intensity performance is mixed, but the magnitude of low bias at early forecast hours is reduced with the use of the MEDMF-TKE scheme, which produces a wider range of TC intensities. Wind radii forecasts, particularly the radius of maximum wind speed (RMW), are also improved with the MEDMF-TKE scheme. Composites of TC inner-core structure in and above the PBL highlight and explain differences between the two sets of forecasts, with MEDMF-TKE having a stronger and shallower inflow layer, stronger eyewall vertical velocity, and more moisture in the eyewall region. A case study of Hurricane Laura shows that MEDMF-TKE better represented the subtropical ridge and thus the motion of the TC. Finally, analysis of Hurricane Delta through a tangential wind budget highlights how and why MEDMF-TKE leads to faster spinup of the vortex and a better prediction of rapid intensification.

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

1. Introduction

The primary goals of the ongoing Hurricane Forecast Improvement Project (HFIP; Gopalakrishnan et al. 2020) are to improve forecasts of TC intensity change and rapid intensification (RI), and also to improve forecasts of TC structure. Recent work (Gopalakrishnan et al. 2013, 2021; Tallapragada et al. 2014; Zhang et al. 2015) has demonstrated that improvements to hurricane structure representation in the forecast models, especially in the planetary boundary layer (PBL), lead to intensity forecast improvements. Those works have also illustrated that inner-core observations obtained from NOAA P3 flight missions may be key to the success of improving PBL schemes in hurricane forecast models, and demonstrate the need to use high-quality observations in future and ongoing model development.

In response to the proposed framework for addressing section 104 of the Weather Research Forecasting Innovation Act of 2017, NOAA is developing the Hurricane Analysis and Forecast System (HAFS). HAFS is a multiscale model and data assimilation package currently in development, capable of providing analyses and forecasts of the inner core structure of the TC out to 7 days, which is key to improving size and intensity predictions, as well as the large-scale environment that is known to influence the TC's motion (Dong et al. 2020; Hazelton et al. 2021a,b). The goal of HAFS is to provide reliable, robust, and skillful guidance on TC track and intensity (including RI), storm size, genesis, storm surge, rainfall, and tornadoes associated with TCs within the Unified Forecast System (UFS).

Recent studies have examined the impact of various modifications to PBL schemes to improve model forecasts of TC structure and intensity, and have shown how sensitive TC structure is to the choice of PBL parameterization, and configuration of the parameterization. Upgraded schemes have improved the utility of high-resolution models for both TC forecasting and research. Braun and Tao (2000) was an early example of such a study, using simulations from the fifth-generation Pennsylvania State University–National Center for Atmospheric Research (PSU–NCAR) Mesoscale Model (MM5) to examine the impact of different PBL schemes in simulations of Hurricane Bob (1991). They found that the intensity and precipitation structure of the TC were both very sensitive to the choice of PBL scheme. Smith and Thomsen (2010) noted the importance of differences in eddy diffusivity in modulating PBL structure in their idealized framework, which was supported by observational studies and real-case simulations that will be discussed later. Nolan et al. (2009a,b) tested and modified two PBL schemes in the Weather Research and Forecasting Model for simulations of Hurricane Isabel (2003), and found that ocean surface roughness, vertical resolution, and horizontal resolution are all important aspects of producing realistic PBL structure in TCs. Kepert (2010) showed that a slab PBL model produced an unrealistic TC inflow structure compared to a scheme that fully resolved

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

DOI: 10.1175/WAF-D-21-0124.1

the height in the PBL, highlighting the importance of fully resolving TC PBL processes in order to accurately depict the structure of the inner core.

One of the key ways that PBL schemes have been improved is through the use of state-of-the-art observational data. NOAA collects observations in the inner core of TCs that are critical to understanding the details of TC structural evolution (e.g., Zawislak et al. 2022) and can be applied to PBL schemes. One of the first studies to take advantage of this observational data was Gopalakrishnan et al. (2013). This study used flight-level data that estimated eddy diffusivity, and based on this data, introduced a parameter to scale the vertical diffusivity in the Hurricane Weather Research and Forecasting (HWRF) Model. Zhang et al. (2015) and Zhang et al. (2017) further showed how the observations-based modifications to reduce eddy diffusivity in the HWRF PBL scheme allowed for better prediction of TC structure and intensity (including rapid intensification). Building on this earlier work with HWRF, Gopalakrishnan et al. (2021) examined two different PBL schemes with modifications based on observations in HAFS simulations of Hurricane Michael (2018) and demonstrated how different boundary layer schemes applied to HAFS produced a wide variety of forecasts of structure and intensity change, based on uncertainty in key variables like eddy diffusivity. As new PBL schemes are developed for NOAA's UFS, it is important to test these schemes using a variety of TC cases (beyond just a single case study) to understand and improve their ability to represent TC structure. This will build on the foundation of work with the HWRF model to guide the implementation of optimized PBL schemes within the newer HAFS model for skillful prediction of TC structure and intensity change. The prior studies have clearly demonstrated that the choice of PBL schemes can play a big role in how the model simulates the TC inner core, and tweaks to diffusivity and other parameters within the schemes can lead to large changes in the prediction of TC evolution. It is important to evaluate how different schemes and modifications alter TC structure in this newly developed forecast system.

The current study builds on Gopalakrishnan et al. (2021) and other prior studies by examining the performance of a modified scheme over a wide variety of TC cases spanning a large range of structures and intensities. The main focus of this paper is comparing two different PBL schemes in otherwise identical versions of HAFS-globalnest. Many of the recent studies examining PBL impact on TC structure have focused on idealized cases (e.g., Gopalakrishnan et al. 2013; Zhang et al. 2020), or one or two real-world cases (e.g., Braun and Tao 2000; Nolan et al. 2009a,b). This study uniquely composites a large dataset of cases from the 2020 hurricane season to compare how the mean TC structure is impacted by the details of the PBL scheme using the HAFS-globalnest forecast model. This study will help motivate and guide future development of HAFS for both operational prediction and TC research, and motivate further observation-based improvement to physics parameterizations in HAFS. We attempt to isolate the impact on track, intensity, and structure biases from two different PBL schemes using 101 cases from the 2020 season.

2. Data and model setup

a. *Model setup and PBL schemes compared*

HAFS-globalnest (Hazelton et al. 2021a,b) was used in this study. The model layout includes a large static nest covering the North Atlantic, with two-way feedback in the global domain (e.g., Harris and Lin 2012); and 75 vertical levels with finer vertical resolution in the PBL compared to the 2019 version of HAFS-globalnest (Hazelton et al. 2021a). In addition to the PBL schemes described below, key parameterizations include the GFDL 6-class microphysics (Chen and Lin 2013; Zhou et al. 2019) and scale-aware SAS convective scheme (Han et al. 2017, for the global domain only, which tests showed improved track forecasts, not shown). Hazelton et al. (2021b) describes the model configuration in more detail.

The effects of two different PBL schemes on TC forecasts are examined in this study. One scheme is the hybrid eddy diffusivity mmass flux scheme (HEDMF-KP) scheme (Han et al. 2016). This parameterization uses the K-profile (KP) scheme in well-mixed boundary layers and also uses a simple first-order turbulence scheme based on the Richardson number for stable boundary layers. HEDMF-KP (also sometimes known as the hybrid EDMF scheme) was used in the operational Global Forecast System (GFS) until an upgrade in early 2021. Prior work has made modifications to this scheme to improve the structure of eddy diffusivity in high-wind conditions, leading to overall improvements in TC structure forecasts in HWRF (Wang et al. 2018). This has led to improvements in inflow angle, inflow depth, TC maximum intensity, as well as TC size. This modified version of HEDMF-KP was used in the operational HWRF model and was also used in initial experimental tests of HAFS during the 2019 Atlantic Hurricane Season (Dong et al. 2020, Hazelton et al. 2021a). More recently a turbulent kinetic energy (TKE)-based scheme (Han and Bretherton 2019) was developed for the GFS. This parameterization bases the eddy diffusivity in the PBL on the prognostic TKE and mixing length, with the prognostic TKE parameterized by an equation including shear and buoyancy productions, turbulence, dissipation, and mixing length. Based on the importance of mixing length in this scheme, a modification was made by lowering the cap on the saturation mixing length to 100 m to better match observations of TC mixing length in hurricane environments (Gopalakrishnan et al. 2021). This modification produced a sharper inflow angle, lower eddy diffusivity, and a stronger TC in the simulations of Hurricane Michael (2018) performed in that study. This modified EDMF-TKE scheme is referred to as MEDMF-TKE hereafter, which we compare with HEDMF-KP. Gopalakrishnan et al. (2021) compared MEDMF-TKE with a simpler version of HEDMF-KP (only using the alpha parameter to adjust eddy diffusivity), but only for the Hurricane Michael case. Here, we compare the most up-to-date versions of MEDMF-TKE and HEDMF-KP using a large set of cases from the 2020 Atlantic hurricane season.

Figure 1 shows the eddy diffusivity for momentum at 500-m height as a function of wind speed (Fig. 1a), and inflow angle as a function of normalized radius (Fig. 1b), for individual forecast runs from the two schemes compared in this study. Observational eddy diffusivity derived from flight-level data

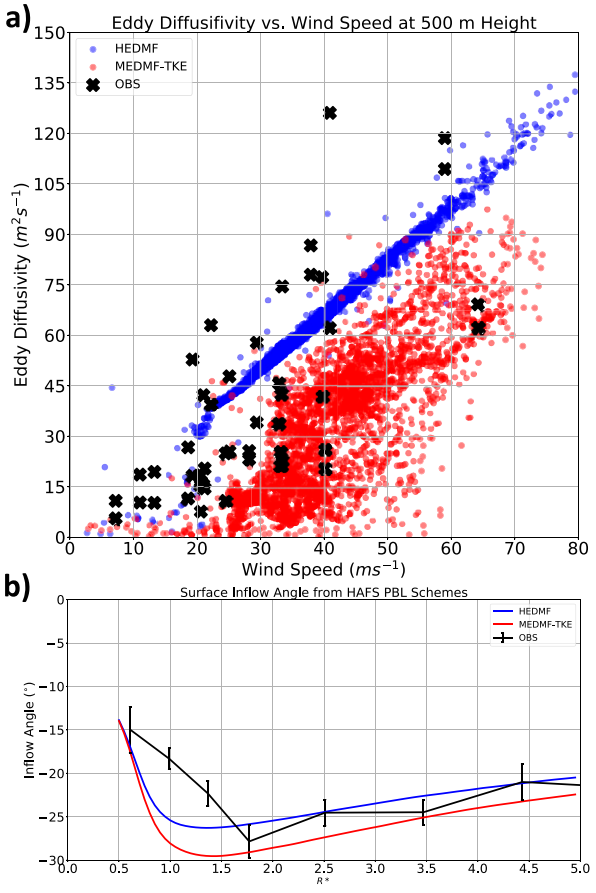

FIG. 1. (a) Eddy diffusivity ($m^2 s^{-1}$) as a function of wind speed ($m s^{-1}$) for the HEDMF-KP (blue) and MEDMF-TKE (red) schemes. Observational estimates of eddy diffusivity from Zhang et al. (2011) are shown in black. (b) Surface inflow angle as a function of radius normalized by the radius of maximum wind for the HEDMF-KP (blue) and MEDMF-TKE (red) schemes. Observational inflow angle (including confidence intervals) from Zhang and Uhlhorn (2012) is shown in black.

in several strong hurricanes (Zhang et al. 2011) and observational composites of inflow angle from GPS dropwindsonde data (Zhang and Uhlhorn 2012) are also overlaid for comparison. A similar framework was used to compare MEDMF-TKE with a simplified version of HEDMF-KP (using a simple parameter to adjust eddy diffusivity) by Gopalakrishnan et al. (2021) in a case study of Hurricane Michael (2018). The comparison in this study is slightly different from that in Fig. 1 of Gopalakrishnan et al. (2021) due to using a different TC and also, more importantly, because this version of HEDMF-KP already includes some modifications to eddy diffusivity at high wind speeds (Wang et al. 2018). Still, the eddy diffusivity for MEDMF-TKE in Fig. 1a is lower than that for HEDMF-KP. Figure 1b shows the inflow angle for the two schemes and the dropwindsonde composite (over a range of intensities) from Zhang and Uhlhorn (2012). The inflow angle has been shown to be a key parameter for understanding differences between PBL schemes (e.g., Wang et al. 2018; Gopalakrishnan et al. 2021). The inflow angle for the two schemes both generally fall within the bounds of the observational range for radii outside the RMW, although the inflow angle is larger for MEDMF-TKE than HEDMF-KP, consistent with the lower eddy diffusivity. The impact of these general characteristics of each scheme on TC evolution, including track, structure, and intensity, will be examined in the rest of this paper.

TABLE 1. List of cases and date ranges analyzed.

Storm ID	Storm name	No. of cases	First/last initialization date
AL07	Gonzalo	3	1200 UTC 7 Jul/1200 UTC 23 Jul
AL08	Hanna	2	0000 UTC 23 Jul/1200 UTC 23 Jul
AL09	Isaias	3	0000 UTC 30 Jul/0000 UTC 31 Jul
AL11	Josephine	1	0000 UTC 12 Aug
AL13	Laura	10	0000 UTC 20 Aug/1200 UTC 24 Aug
AL14	Marco	9	0000 UTC 21 Aug/1200 UTC 24 Aug
AL17	Paulette	18	0000 UTC 7 Sep/0000 UTC 16 Sep
AL18	Rene	10	1200 UTC 9 Sep/1200 UTC 14 Sep
AL19	Sally	9	0000 UTC 12 Sep/0000 UTC 16 Sep
AL20	Teddy	7	0000 UTC 13 Sep/0000 UTC 16 Sep
AL21	Vicky	2	1200 UTC 15 Sep/0000 UTC 16 Sep
AL25	Gamma	3	0000 UTC 5 Oct/0000 UTC 6 Oct
AL26	Delta	6	0000 UTC 5 Oct/0000 UTC 8 Oct
AL27	Epsilon	3	0000 UTC 20 Oct/1200 UTC 25 Oct
AL28	Zeta	5	0000 UTC 25 Oct/0000 UTC 27 Oct
AL29	Eta	5	0000 UTC 1 Nov/0000 UTC 3 Nov
AL31	Theta	5	1200 UTC 13 Nov/1200 UTC 15 Nov

b. *Cases analyzed*

HAFS-globalnest was run in real-time during the 2020 Atlantic hurricane season using the MEDMF-TKE PBL scheme. For comparison purposes, a large subset of runs was performed with HEDMF-KP (with an otherwise identical configuration), totaling 64 runs of HAFS-globalnest for key 0000 and 1200 UTC cycles, with a total of ~120 individual TC forecasts for comparison with the real-time runs. Note that only 101 cases were included in the verification statistics, because the other 19 were for "invests," which were not classified as TCs yet. Table 1 lists the TCs analyzed and the date ranges for each. The 2020 Atlantic hurricane season featured a record number of named TCs in the Atlantic Basin, and there was a wide variety of TCs of different sizes, intensities, and structures covering a wide date range during the season.

3. Results

a. *Basic verification statistics*

1) TRACK

Figure 2 shows the track forecast performance from the HAFS-globalnest forecasts with MEDMF-TKE (real-time), and the hindcasts using HEDMF-KP. The results clearly indicate that MEDMF-TKE had better track forecast skill during the 2020 season than HEDMF-KP at every forecast hour after 36 h. Breaking down the performance into along-track and across-track errors reveals more details. The along-track skill

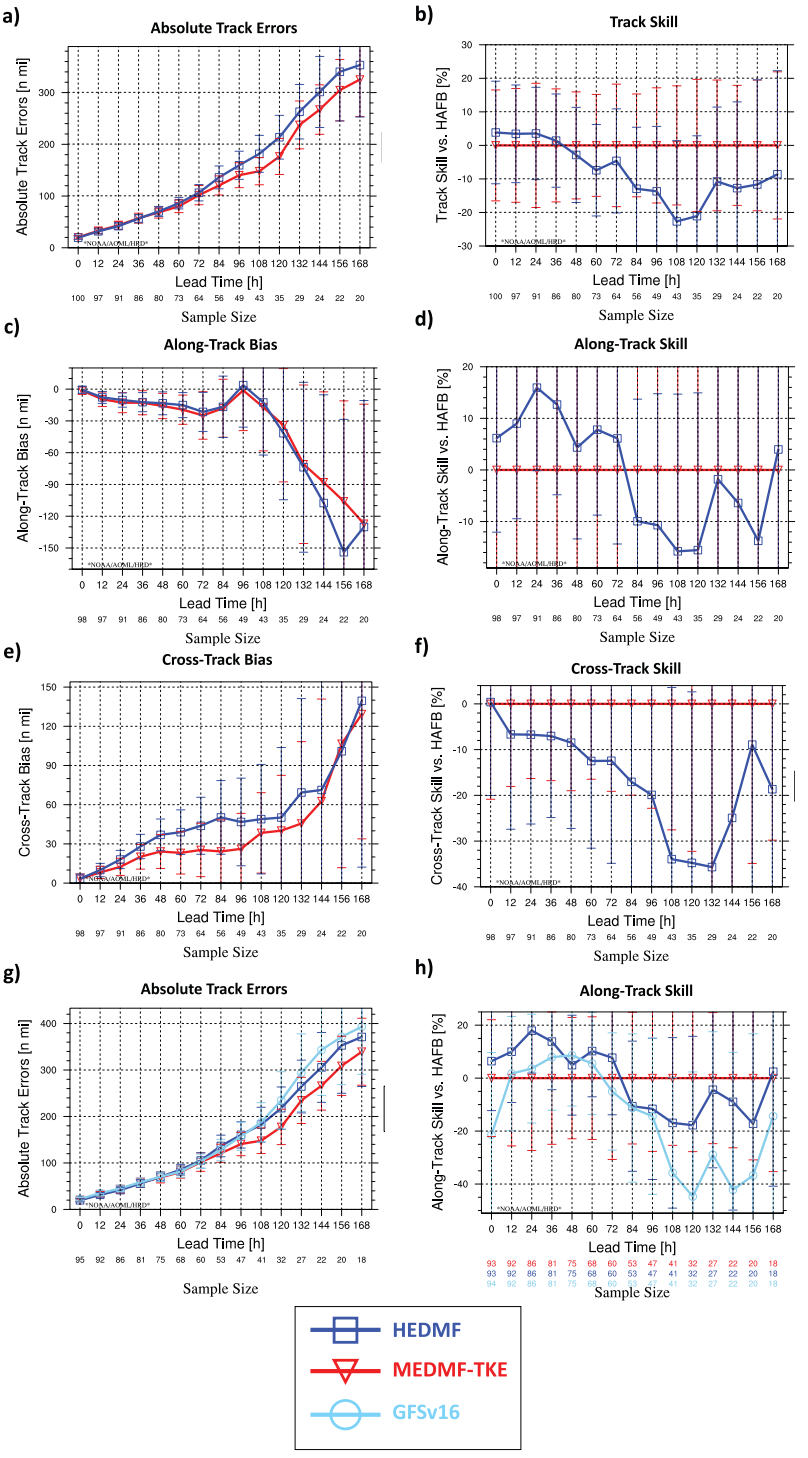

FIG. 2. (a) Track error (n mi; 1 n mi = 1.852 km) for HAFS-globalnest with MEDMF-TKE (red) and HEDMF-KP (blue). (b) Track skill of HAFS-globalnest with HEDMF-KP relative to MEDMF-TKE. (c) Along-track bias for HAFS-globalnest with MEDMF-TKE (red) and HEDMF-KP (blue). (d) Along-track track skill of HAFS-globalnest with HEDMF-KP relative to MEDMF-TKE. (e) As in (c), but for across-track bias. (f) As in (d), but for across-track skill. (g) As in (a), but also including GFSv16 results (light blue). (h) As in (f), but also including GFSv16 results (light blue). The black numbers show the sample size at each lead time. The error bars show the 95% confidence interval.

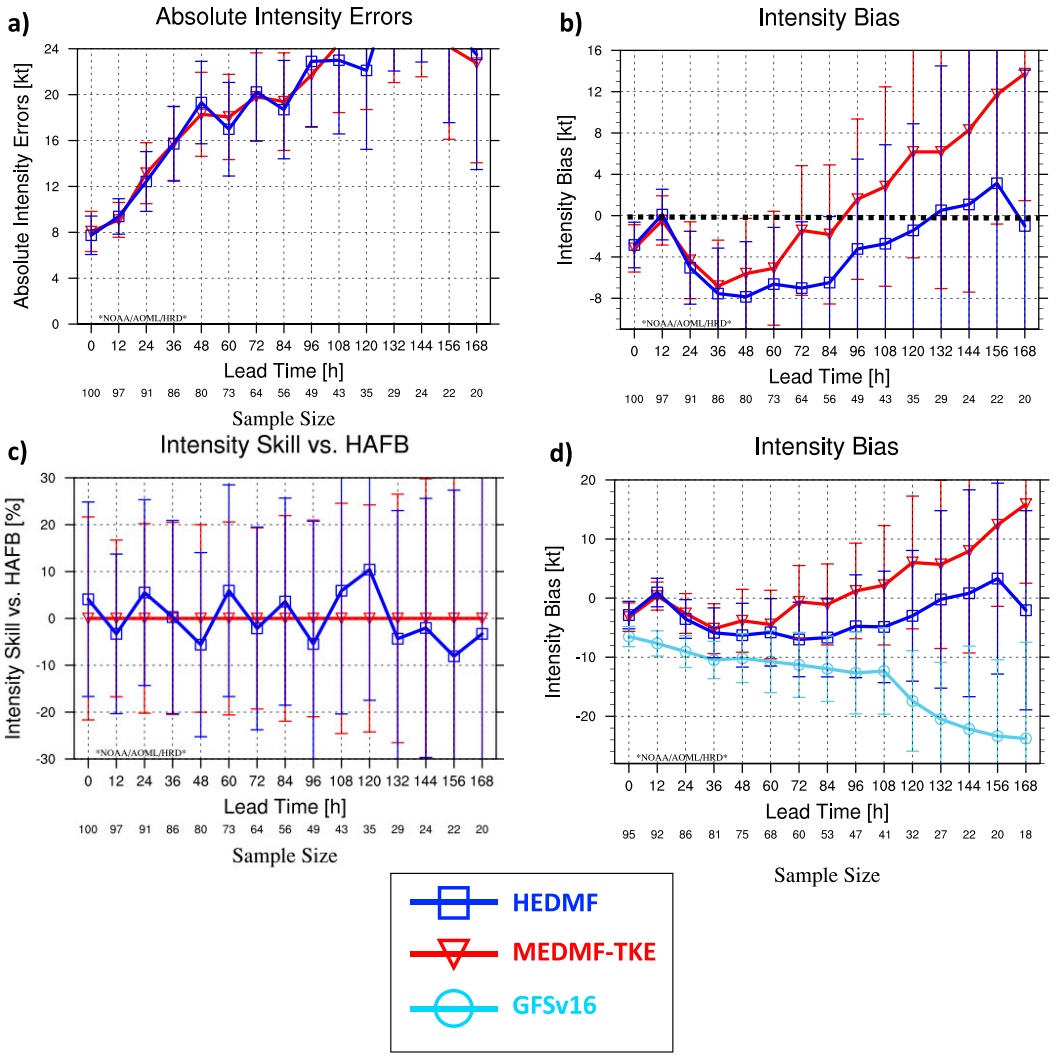

FIG. 3. (a) Mean intensity error (kt) for HAFS-globalnest with MEDMF-TKE (red) and HEDMF-KP (blue). (b) Mean intensity bias (kt) for HAFS-globalnest with MEDMF-TKE (red) and HEDMF-KP (blue). The thick dashed black line shows the 0 bias line. (c) Intensity skill of HAFS-globalnest with HEDMF-KP relative to MEDMF-TKE. The red numbers show the sample size at each lead time. (d) As in (b), but also including GFSv16 results (light blue). The black numbers show the sample size at each lead time. The error bars show the 95% confidence interval.

was relatively comparable, although HEDMF-KP produced a larger negative (slow) bias at longer lead times. The much larger difference came from the cross-track bias, although it should be noted that the sample size is small at longer lead times. The real-time HAFS-globalnest (with MEDMF-TKE) had a notable right bias at all forecast hours. The results here, however, indicate that HAFS-globalnest with HEDMF-KP had an even worse right bias, with cross-track skill 40% lower than MEDMF-TKE at 120 h. A case study of Hurricane Laura presented later will examine one of the most extreme examples of this right bias and some of the reasons for it, and the differences in the track bias between the two schemes. Overall, the new PBL scheme, MEDMF-TKE, led to improvements in the track forecasts in the Atlantic despite the biases present. The track forecasts for HAFS-globalnest with MEDMF-TKE also showed skill relative to the parallel GFSv16 forecasts (the version of the GFS that became operational in 2021) from 2020 (Figs. 2g,h), which included the (unmodified) version of EDMF-TKE. On the global domain of HAFS-globalnest, the track results were generally neutral in the east Pacific (not shown). MEDMF-TKE had slightly worse track forecasts at early lead times in the west Pacific (not shown), but slightly better at longer lead times. The sample sizes for the other basins were too small for robust analysis; however, the present paper mostly focused on Atlantic cases.

2) INTENSITY

Figure 3 shows basic intensity (maximum instantaneous model wind) forecast statistics from the real-time HAFS-

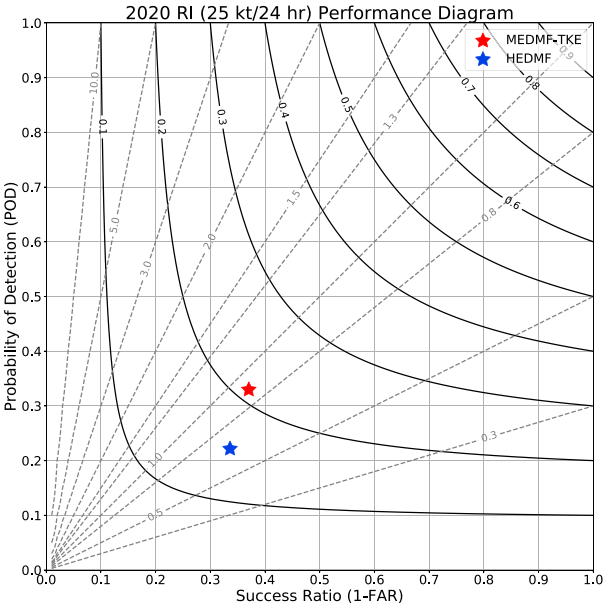

FIG. 4. Performance diagrams for rapid intensification forecasts for HAFS-globalnest with MEDMF-TKE (red) and HEDMF-KP (blue) for the 25 kt (24 h)$^{-1}$ metric. The x axis shows the success ratio (1 minus the false alarm ratio), and the y axis shows the probability of RI detection. The dashed lines are the bias scores, and the solid lines are the critical success index (CSI). See Roebber (2009) for more details.

globalnest forecasts using MEDMF-TKE and the experiment using HEDMF-KP (and also GFSv16). The overall intensity error is very similar between the two sets of HAFS forecasts, and both versions, not surprisingly, performed better than the global model. The skill of HEDMF-KP relative to MEDMF-TKE oscillates around 0 for all lead times. However, the bias characteristics are very different. MEDMF-TKE has a much smaller bias than HEDMF-KP (which had a substantial negative bias) around days 3–4. However, at longer lead times a large positive bias develops in MEDMF-TKE. It should be noted, however, that the sample size is fairly small at longer leads, which results in large error bars and increased uncertainty in the robustness of the results for days 5 and after. The relationship between the intensity differences and structure differences caused by the different PBL schemes is explored in coming sections, and an example of a case with large intensity differences is discussed.

One of the key parts of TC intensity change prediction is to correctly predict its rapid intensification. Improving forecasts of RI is one of the major goals of HFIP. To assess the skill of RI forecasts using the two different PBL schemes, a performance diagram (Roebber 2009) for the two schemes is shown in Fig. 4. For the sake of generating a large enough sample size for a meaningful result, intensity changes from different forecast hour periods (e.g., 12–36, 24–48, and 96–120, etc.) are included together in this diagram. Also, due to similar sample size constraints, a slightly lowered threshold is used for rapid intensification, 25 kt (24 h)$^{-1}$ (1 kt $\approx$ 0.51 m s^{-1}) rather than the traditional 30 kt (24 h)$^{-1}$, is one of the RI thresholds used in the Statistical Hurricane Intensity Prediction Scheme (SHIPS; Kaplan et al. 2010). The model with MEDMF-TKE has better skill in RI forecasts using this metric, with a much higher probability of detection (POD) and also a slightly lower false alarm rate (FAR), indicating that despite the similar performance in overall intensity skill, MEDMF-TKE was able to improve forecasts during critical RI periods.

3) WIND RADII

When evaluating the impacts of a PBL scheme on TC forecasts, it is important to analyze metrics of TC structure alongside standard track and intensity. One such metric is the wind radii. Wind radii forecasts have been examined in recent models and official forecasts by Cangialosi and Landsea (2016), and are important for the communication of TC impacts as well as for prediction of overall TC structure. Figure 5 shows the wind radii verification metrics for these two experiments, including four different radii (based on the 10-m winds): 34-kt, 50-kt, 64-kt, and radius of maximum winds (RMW). The verifications are performed against the "best track" radii. For the first three metrics, the errors are comparable between the MEDMF-TKE and HEDMF-KP forecasts. However, the RMW bias (with RMW determined from the GFDL vortex tracker, Marchok 2002, 2021) is somewhat reduced in the forecasts with MEDMF-TKE compared to HEDMF-KP, especially at longer lead times. Since the contraction of the RMW in TCs tends to be associated with intensification (Shapiro and Willoughby 1982), there is likely a connection between the differences in intensity bias and RMW bias between the two schemes. Both schemes have a negative (too small) bias in 64-kt wind radius at longer lead times, which could be a reflection of the model failing to fully capture eyewall replacement cycles or extratropical transition at longer leads. This effect appears to be especially prevalent in MEDMF-TKE, as the RMW bias also becomes negative at longer leads. However, the sample size is low so that the uncertainty is high at these later forecast hours (as evidenced by the large error bars). More work will be needed to fully explore these possibilities (e.g., forecasts of eyewall replacement cycle) in the future, and examine how differences in model resolution and changes in the PBL interact together when impacting the structure forecasts. More details of the wind structure differences, including those above the surface, will be explored in the next section.

b. Distributions of intensity and structure biases

To further examine the characteristics of HAFS-globalnest forecasts with each PBL scheme, Fig. 6 shows the distribution of intensity bias, RMW bias, and 34-kt wind radii bias for the forecasts with MEDMF-TKE and HEDMF-KP. The forecasts at 72 h are chosen because they are a few days into the forecast but at a period when the sample size is still relatively large. The distribution of intensity bias is consistent with the mean error plots: MEDMF-TKE is shifted to the right, with fewer cases that have a low intensity bias than HEDMF-KP. For RMW, the peak of the distribution is positive for HEDMF-KP and negative for MEDMF-TKE, indicating that MEDMF-TKE tends to make smaller inner cores than

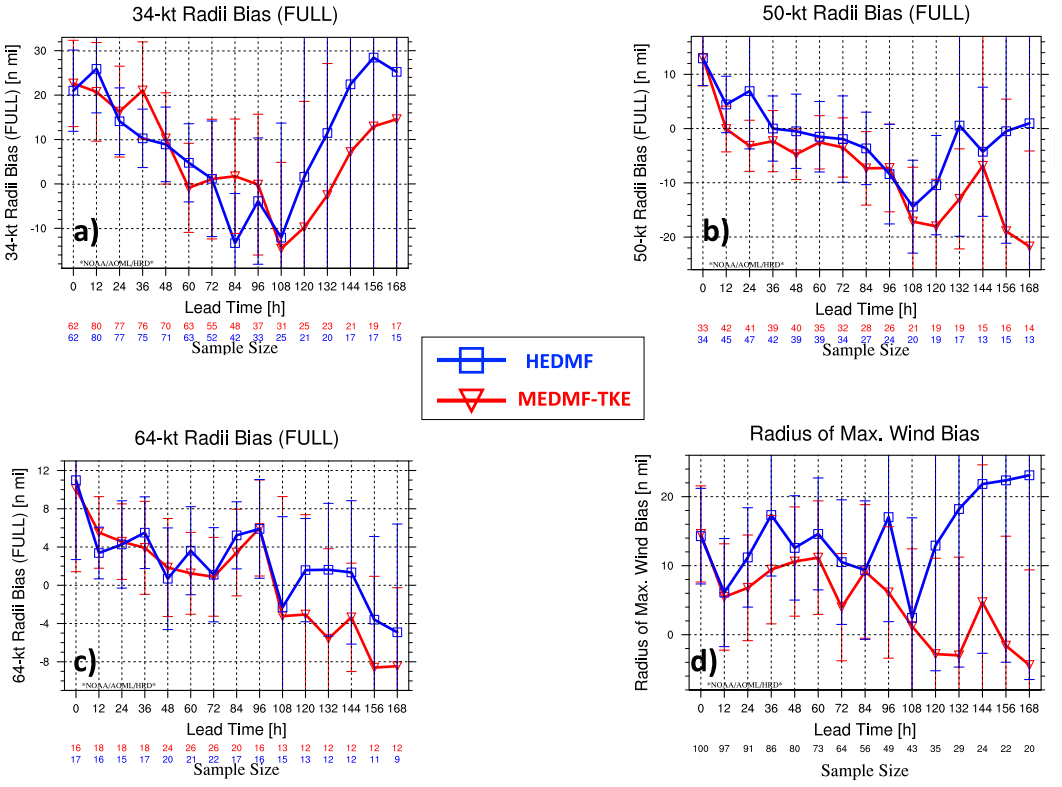

FIG. 5. (a) Mean 34-kt wind radii bias for HAFS-globalnest with MEDMF-TKE (red) and HEDMF-KP (blue). (b) As in (a), but for 50-kt wind radii bias. (c) As in (a), but for 64-kt wind radii bias. (d) As in (a), but for radius of maximum winds (RMW) bias. The numbers along the bottom show the sample size at each lead time. The error bars show the 95% confidence interval.

HEDMF-KP (consistent with the better POD for RI). There does not appear to be a large difference in the overall distribution for R34, consistent with the similar mean bias at 72 h seen in Fig. 5c. The next section will explore these structure and intensity differences in more detail, and discuss some of the likely causes.

c. *Structure composites*

One useful method for capturing the impact of PBL schemes across a wide range of cases is to examine composites of TC structure. The impact of modifications to the PBL scheme on the azimuthal mean TC structure has been demonstrated in prior studies (e.g., Gopalakrishnan et al. 2013, 2021; Zhang et al. 2015). In this section, several variables are composited in an azimuthal-mean framework to examine the impact of the different PBL schemes on the TC primary and secondary circulations, as well as precipitation and thermodynamic structure. As in other composite studies using model and radar data (e.g., Rogers et al. 2013; Zhang et al. 2015), the data are normalized relative to the RMW at $z = 2$ km (in normalized bins of 0.05RMW) to prevent artifacts from averaging storms of different sizes. In addition, since quite a few 2020 cases formed or moved near land, we removed data points where the TC was within 3RMW of a major landmass, to avoid complications and biases from land. A final filter was applied to remove all TCs of less than tropical storm intensity (35 kt, 18 m s^{-1}) to avoid including cases where the TC circulation was broad and/or diffuse, and thus the large RMW was not representative and suitable for normalization. The median intensities of the two sets of composites were 63 kt (32 m s^{-1}) for HEDMF-KP and 64 kt (33 m s^{-1}) for MEDMF-TKE, indicating that the storm strength samples are fairly comparable for purposes of structural analysis.

The first composite examined was the TC tangential wind, representative of the primary circulation (Fig. 7). For both schemes, the composite shows a deep, broad cyclonic circulation, with a peak in the eyewall at around z = 500–1000 m, generally consistent with the dropsonde composite of Franklin et al. (2003). However, the composite for MEDMF-TKE shows a stronger maximum, especially in the lowest 1 km of the troposphere. A difference plot between the two schemes confirms this point, with a composite difference of 3–4 m s^{-1} in the PBL near the RMW. This difference extends upward throughout the troposphere, with the MEDMF-TKE composite 1–2 m s^{-1} stronger up to ~ z = 16 km. It is clear that the MEDMF-TKE scheme tends to produce deeper TCs than the HEDMF-KP, which is likely linked to the RI forecast differences (e.g., Zhang et al. 2017).

The radial flow (Fig. 8) comparison also shows some key structural differences between the two forecasts. MEDMF-

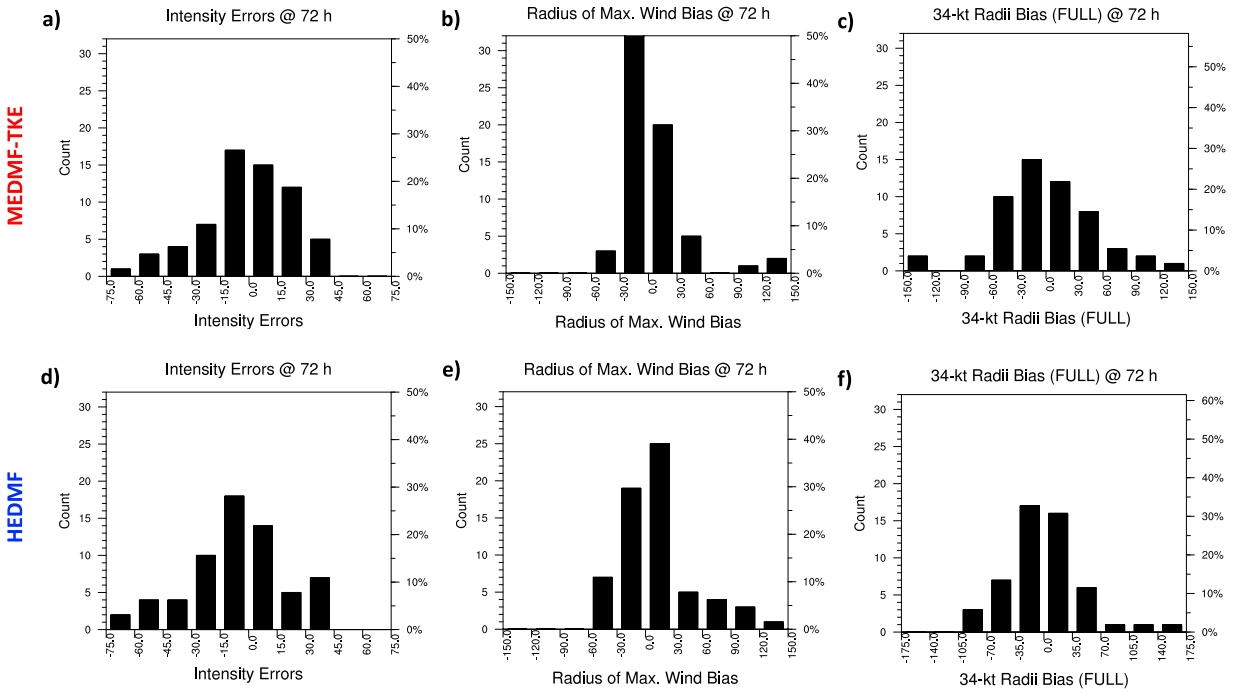

FIG. 6. (a) Distribution of intensity bias at 72 h for HAFS-globalnest with MEDMF-TKE. (b) Distribution of RMW bias at 72 h for HAFS-globalnest with MEDMF-TKE. (c) Distribution of R34 bias at 72 h for HAFS-globalnest with MEDMF-TKE. (d) As in (a), but with HEDMF-KP. (e) As in (b), but with HEDMF-KP. (f) As in (d), but with HEDMF-KP.

TKE produces a stronger and slightly shallower inflow layer than the HEDMF-KP, which should be associated with stronger supergradient flow at the top of the PBL near the RMW (e.g., Kepert and Wang 2001). The stronger near-surface inflow in MEDMF-TKE indicates that there is more inward transport of angular momentum from outer radii, which is likely one of the main reasons for a greater rate of intensification in MEDMF-TKE (e.g., Gopalakrishnan et al. 2013; Smith and Montgomery 2015; Zhang and Rogers 2019). There is also stronger outflow aloft in MEDMF-

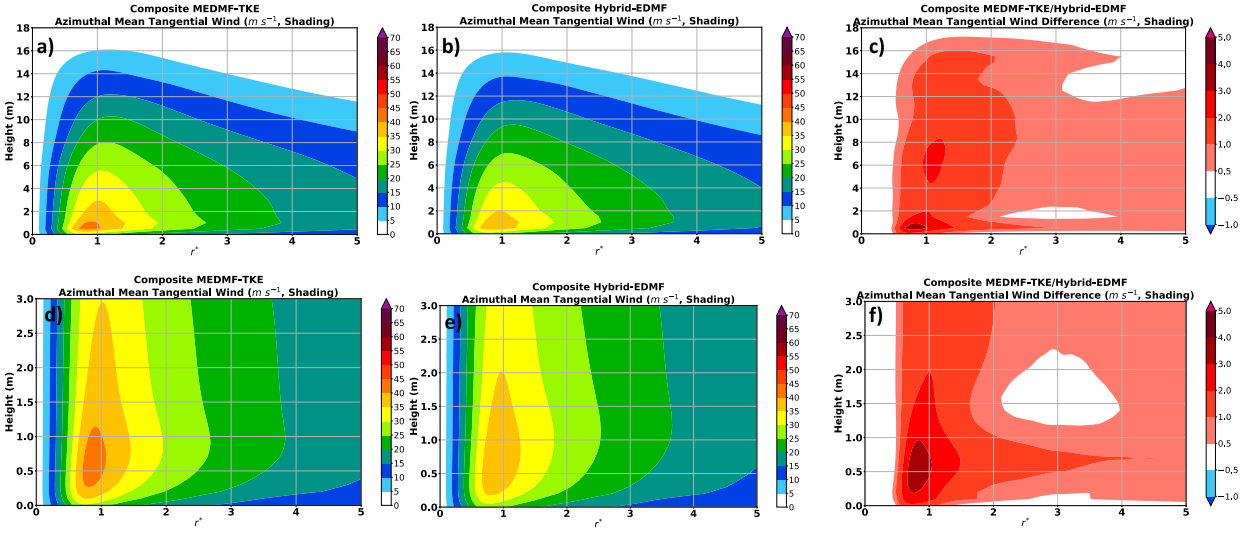

FIG. 7. (a) Composite tangential wind (m s^{-1}) for the forecasts using the MEDMF-TKE scheme from the surface to $z = 18$ km. (b) As in (a), but for the forecasts using the HEDMF-KP scheme. (c) Difference in tangential wind (m s^{-1}) between the forecasts using MEDMF-TKE and the forecasts using the HEDMF-KP from the surface to $z = 18$ km. (d) As in (a), but only up to $z = 3$ km. (e) As in (b), but only up to $z = 3$ km. (f) As in (c), but only up to $z = 3$ km.

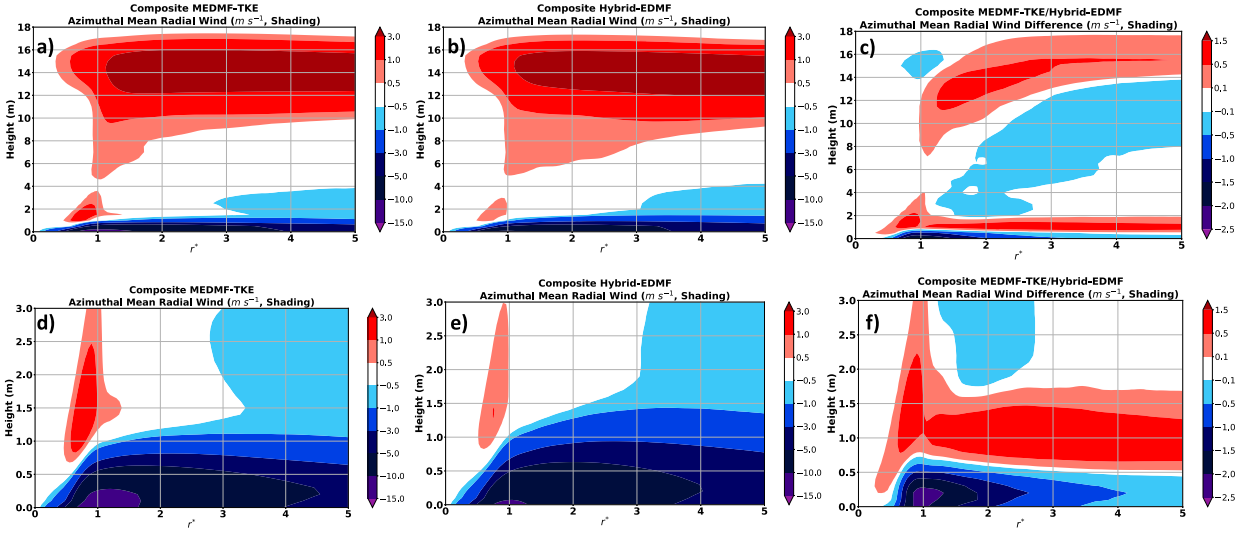

FIG. 8. (a) Composite radial wind (m s^{-1}) for the forecasts using MEDMF-TKE scheme from the surface to $z = 18$ km. (b) As in (a), but for the forecasts using the HEDMF-KP scheme. (c) Difference in radial wind (m s^{-1}) between the forecasts using MEDMF-TKE and the forecasts using the HEDMF-KP from the surface to $z = 18$ km. (d) As in (a), but only up to $z = 3$ km. (e) As in (b), but only up to $z = 3$ km. (f) As in (c), but only up to $z = 3$ km.

TKE, indicating that the entire secondary circulation is stronger.

Evidence of a more robust and concentrated secondary circulation in MEDMF-TKE than in HEDMF-KP was also seen in composites of vertical velocity (Figs. 9a–c). In both schemes, there is a strong updraft peaking in the 5–10-km layer and sloping outward with height (as seen in, for example, Stern et al. 2014). However, the composite peak updraft is clearly stronger in MEDMF-TKE. Figures 9d–f explores forcing for the stronger updrafts seen in MEDMF-TKE, showing the composite radial divergence in the two schemes. The overall pattern is relatively similar in both forecasts, with the peak in radial convergence inside the RMW in the PBL (favoring intensification, Rogers et al. 2013; Ahern

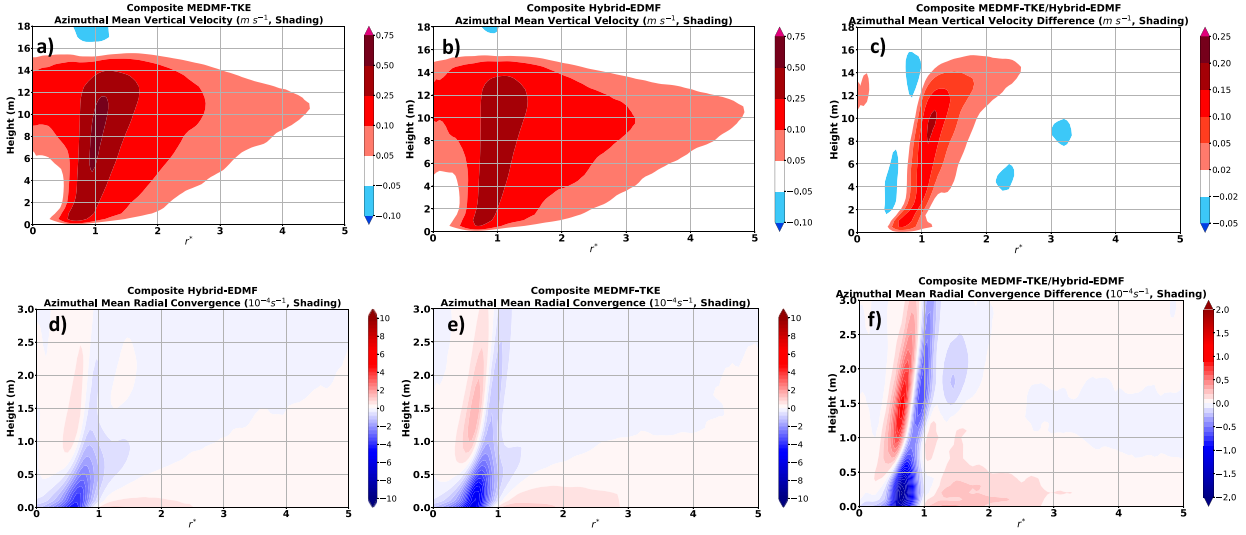

FIG. 9. (a) Composite vertical velocity (m s^{-1}) for the forecasts using MEDMF-TKE scheme from the surface to $z = 18$ km. (b) As in (a), but for the forecasts using the HEDMF-KP scheme. (c) Difference in vertical velocity (m s^{-1}) between the forecasts using MEDMF-TKE and the forecasts using the HEDMF-KP from the surface to $z = 18$ km. (d) Composite radial divergence (10^{-4} s^{-1}) for the forecasts using MEDMF-TKE scheme from the surface to $z = 3$ km. (e) As in (d), but for the forecasts using the HEDMF-KP scheme. (f) Difference in radial divergence (10^{-4} s^{-1}) between the forecasts using MEDMF-TKE and the forecasts using the HEDMF-KP from the surface to $z = 3$ km. In all panels, radial convergence is negative and radial divergence is positive.

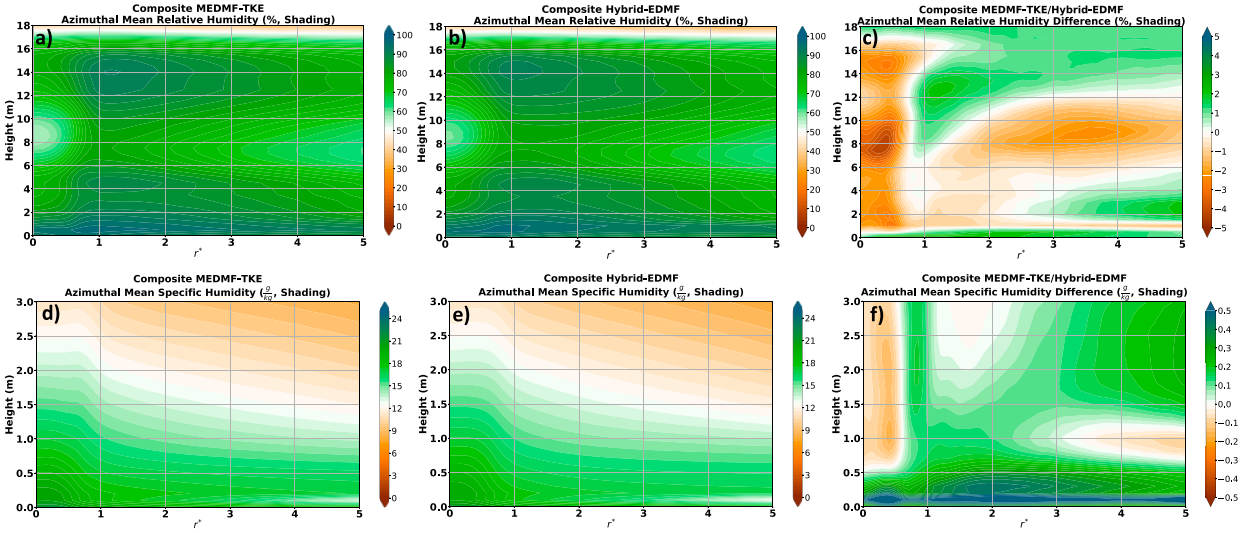

FIG. 10. (a) Composite temperature (K) for the forecasts using MEDMF-TKE from the surface to $z = 18$ km. (b) As in (a), but for the forecasts using HEDMF-KP. (c) Difference in temperature (K) between the forecasts using MEDMF-TKE and the forecasts using the HEDMF-KP from the surface to $z = 18$ km. (d) Composite relative humidity (%) for the forecasts using MEDMF-TKE scheme from the surface to $z = 18$ km. (e) As in (a), but for the forecasts using HEDMF-KP. (f) Difference in relative humidity (%) between the forecasts using MEDMF-TKE and the forecasts using the HEDMF-KP from the surface to $z = 18$ km. (g) Composite specific humidity (g kg^{-1}) for the forecasts using MEDMF-TKE from the surface to $z = 3$ km. (h) As in (g), but for the forecasts using HEDMF-KP. (i) Difference in specific humidity (g kg^{-1}) between the forecasts using MEDMF-TKE and the forecasts using the HEDMF-KP from the surface to $z = 3$ km.

et al. 2019; Zhang et al. 2017) and divergence at the top of the eyewall updraft. The difference plots show that these features are more pronounced in MEDMF-TKE: The radial convergence inside the RMW in the PBL is stronger and also narrower. Note that there is slightly more divergence around r^* (where $r^* = R/\mathrm{RMW}_{2\mathrm{km}}$) = 1–2 in MEDMF-TKE above $z = 1.5$ km. This, along with stronger divergence aloft (not shown), favors stronger and more focused updrafts inside the RMW in MEDMF-TKE, as was seen in the composites of vertical velocity.

Finally, composites of specific humidity (in the PBL) and relative humidity (full troposphere) are shown in Fig. 10. Looking at the full-troposphere humidity, the general structure of the two composites is similar, but the differences show that the boundary layer inflow and eyewall updraft are more moist in MEDMF-TKE, and the eye is drier due to the increased subsidence and warming. According to the specific humidity comparison in the PBL, MEDMF-TKE has a more moist inflow layer, likely due to the stronger and shallower inflow layer resulting from lower diffusivity. These thermodynamic composites illustrate that in addition to the increased inward transport of angular momentum, MEDMF-TKE favors more rapid TC intensification than HEDMF-KP through increasing moisture availability to eyewall updrafts, which along with increased subsidence in the eye results in a warmer TC core (not shown).

Together, the composite results paint a picture of TC structure that is more consistently conducive to intensification and contraction of the inner core in MEDMF-TKE. These results are consistent with the comparison of RI forecast verifications between the forecasts with these two PBL schemes.

d. *Additional structure metrics in all individual cases*

In addition to composites created by averaging the data from a large number of separate forecasts at different times, it is worthwhile to examine the aggregate distribution of TC structure calculated over a large set of individual cases. Figure 11 shows the distribution of a variety of metrics used to assess the TC structure for the entire set of cases from HEDMF-KP and MEDMF-TKE (the same subset used to create the composites above). The first eight metrics are direct measures of the TC structure, while the final two metrics provide a measure of how the PBL schemes are producing different PBL structures within the TC.

The metrics examined are as follows:

1) Maximum wind speed (Fig. 11a)
2) 2-km RMW (Fig. 11b)
3) Vortex depth (Fig. 11c), defined as the height at which the tangential wind along the RMW decreases to 50% of its value at $z = 2$km (Hazelton et al. 2018)
4) "Alpha parameter" (Fig. 11d) describing the rate of radial decay of the tangential wind between the RMW and 3RMW (Mallen et al. 2005)
5) RMW slope (Fig. 11e) $\Delta R/\Delta Z$ (Stern and Nolan 2009; Stern et al. 2014)
6) Local Rossby number (Fig. 11f, Chen et al. 2018)

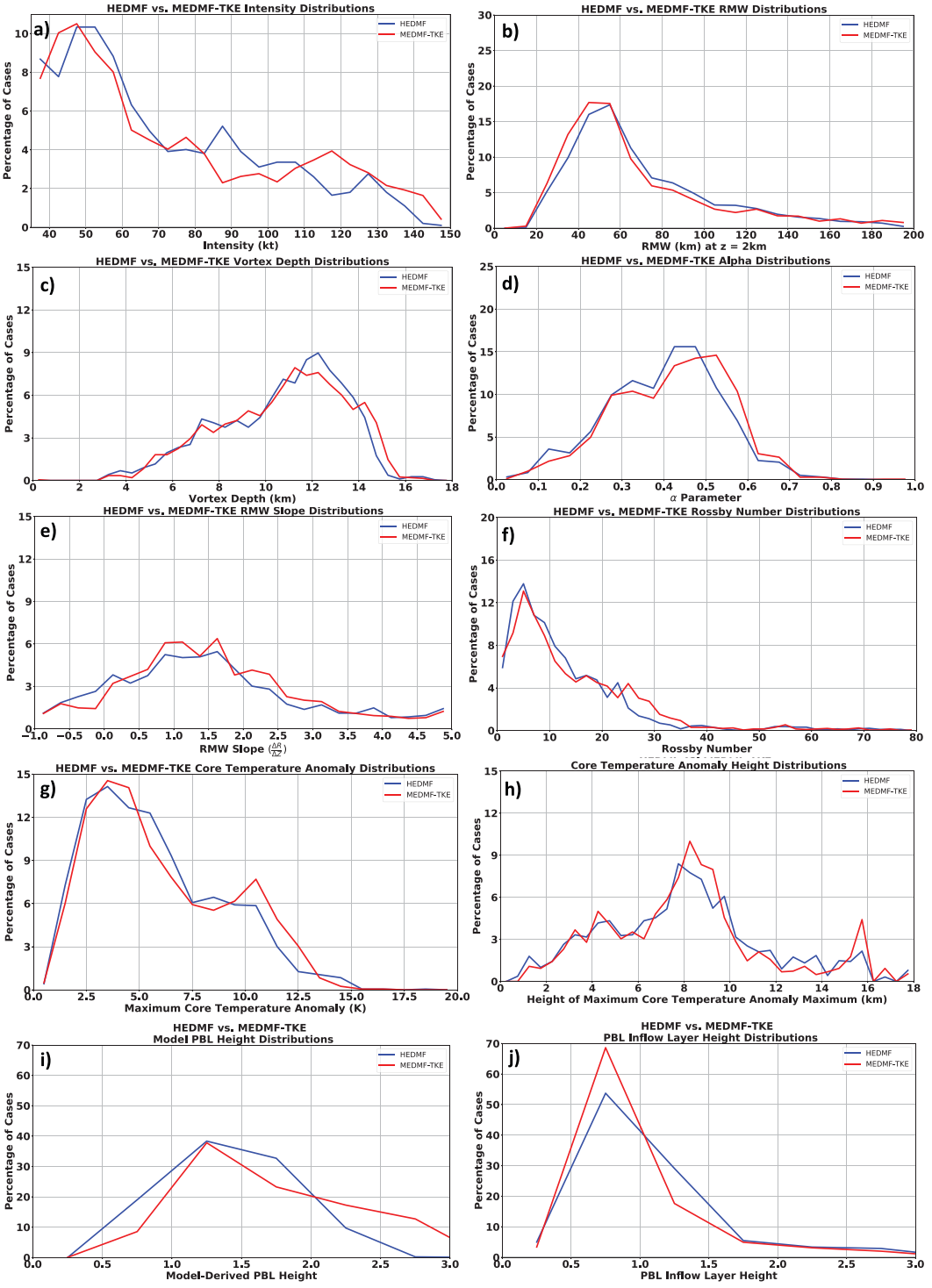

FIG. 11. (a) Histogram (normalized by total number of cases) of intensity for HAFS-globalnest with HEDMF-KP (blue) and MEDMF-TKE (red). (b) As in (a), but for 2-km RMW. (c) As in (a), but for vortex depth. (d) As in (a), but for the alpha decay parameter. (e) As in (a), but for RMW slope. (f) As in (a), but for Rossby number. (g) As in (a), but for core temperature anomaly. (h) As in (a), but for core temperature anomaly height. (i) As in (a), but for model-derived PBL height. (i) As in (a), but for PBL height based on 10% of the peak inflow magnitude.

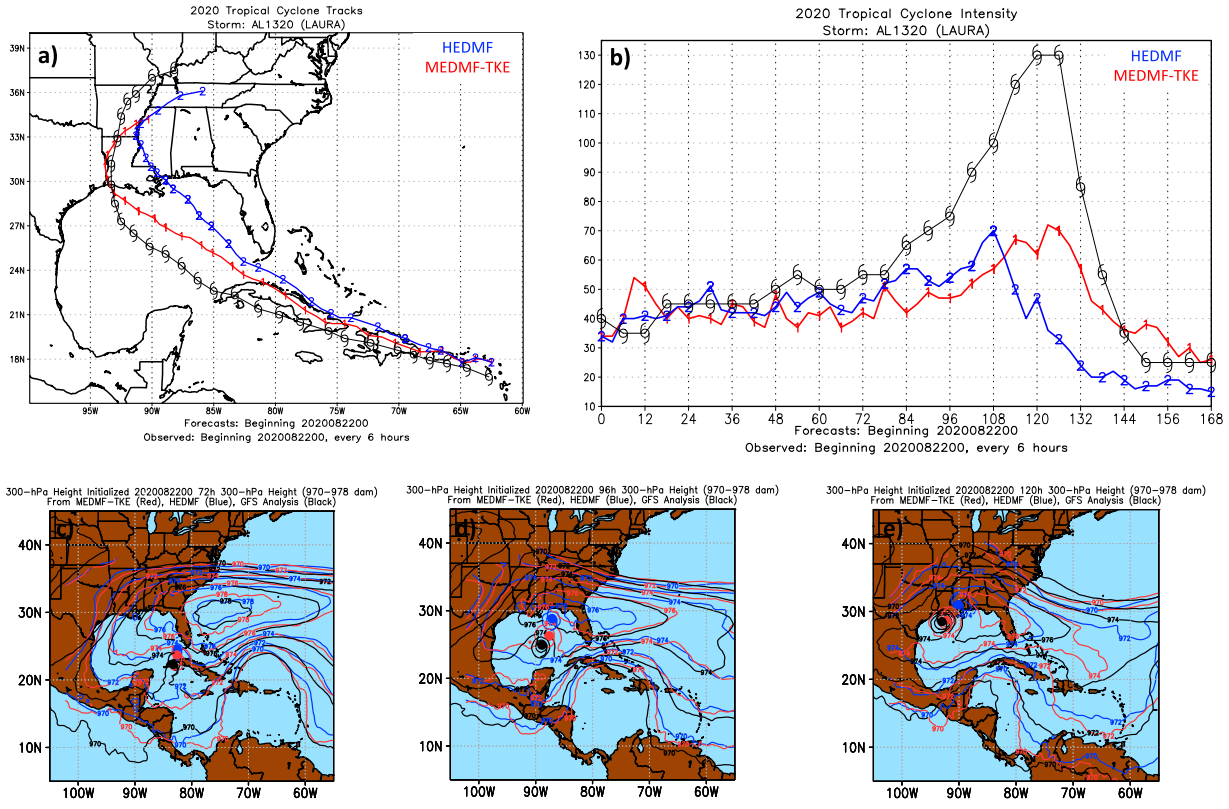

FIG. 12. (a) Tracks of Hurricane Laura for the HAFS-globalnest forecasts with MEDMF-TKE (red) and HEDMF-KP (blue) initialized at 0000 UTC 22 Aug 2020, and run for 168 h. The best track is shown in black. (b) As in (a), but for intensity (kt). (c) 500-hPa height (dam) for the HAFS-globalnest forecasts with MEDMF-TKE (red) and HEDMF-KP (blue) initialized at 0000 UTC 22 Aug 2020 and valid at 72 h. The GFS analysis is shown in black. The TC symbols show the forecast (red, blue) and observed (black) positions of Laura at this time. (d) As in (c), but valid at 96 h. (e) As in (c), but valid at 120 h.

7) Warm-core magnitude (Fig. 11g), defined as the maximum temperature difference between the inner 15-km and the outer 200–300-km ring. (Zhang et al. 2015)
8) Warm core height (Fig. 11h), defined as the height coincident with the warm-core magnitude
9) PBL height derived from the model PBL scheme (Han et al. 2016), evaluated at the 2-km RMW
10) PBL height, defined as the height in the inflow layer where inflow becomes weaker than 10% of the peak inflow (e.g., Zhang et al. 2011, 2015), and evaluated at the RMW of 2 km altitude

Although the median intensity was similar as noted above, the distribution of TC intensity slightly skewed more toward stronger TCs in MEDMF-TKE than in HEDMF-KP, consistent with the composite forcing noted above. The difference in the distribution of RMW is consistent with the intensity differences, with MEDMF-TKE having a smaller mean RMW in general, and a more narrow range of forecast RMW. In terms of vortex depth, MEDMF-TKE also tended to produce TCs that were vertically deeper. HEDMF-KP had a sharp peak at around 12.5 km, while the peak in the distribution for MEDMF-TKE was broader from 12 to 14 km. Similarly, for the alpha decay parameter, HEDMF-KP had a peak at around 0.4, while it was ~0.5 for MEDMF-TKE, indicating that the vortices had a sharper (or narrower) radial wind profile. The sharper and deeper vortices seen in MEDMF-TKE are consistent with the composite structure described above.

RMW slope comparison did not show a large signal, which could be due to the fact that there is typically a weak correlation between RMW slope and TC intensity (Stern and Nolan 2009). The local Rossby number, a measure of both vortex size and intensity, was larger for MEDMF-TKE than HEDMF-KP, consistent with the tendency for stronger and smaller vortices. As with several of the metrics, the peak of Rossby number was broader for MEDMF-TKE than HEDMF-KP, indicating a larger range of vortex structures. The magnitude of the warm core anomaly was typically larger for MEDMF-TKE than HEDMF-KP, with the relative magnitudes of peaks at around 5 and 10 K flipped between the two schemes. However, the warm core height distributions show an interesting pattern: the peak frequency at 8-km height was larger for MEDMF-TKE than HEDMF-KP, which was skewed toward slightly higher values. This is consistent with Zhang et al. (2015), who found a difference in warm-core magnitude in HWRF forecasts with varying vertical diffusion in the PBL scheme, but not a major difference in warm-core height. It is worth noting that

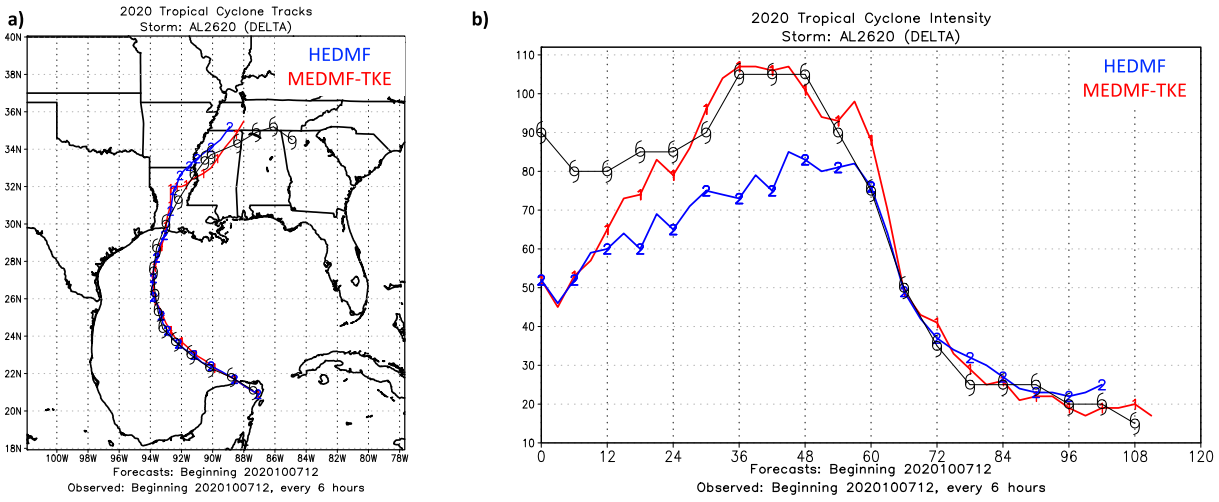

FIG. 13. (a) Track forecasts for Hurricane Delta initialized at 1200 UTC 7 Oct 2020 from HAFS-globalnest with MEDMF-TKE (red) and HEDMF-KP (blue). The best track observations are shown in black. (b) As in (a), but for intensity forecasts (kt).

MEDMF-TKE had a sharper secondary peak at around z = 16 km, however. This peak appears in the composite difference plot of MEDMF-TKE versus HEDMF-KP (Fig. 10c). This is consistent with the full-physics simulations of Stern and Nolan (2012), which found a double maximum warm-core structure. These structure metrics paint a picture consistent with the composites: MEDMF-TKE leads to a more robust secondary circulation with stronger low-level inflow that leads to smaller TCs, sharper wind maxima, larger warm-core anomaly, and stronger TCs.

The PBL-specific metrics highlight some of the ways in which the two schemes examined produce different PBL structures. The PBL height based on the inflow layer depth (Fig. 11j) is smaller in the runs with MEDMF-TKE than in the runs with HEDMF-KP, consistent with the smaller eddy diffusivity (e.g., Zhang et al. 2015) and also similar to what was seen in the inflow composites (Fig. 8). However, the model-derived PBL height (Fig. 11i) was different. In fact, MEDMF-TKE had a wider distribution with more cases with a larger PBL height as calculated in the scheme itself. As discussed in Han et al. (2016; Han and Bretherton 2019), for both of these schemes the PBL height is derived based on the Critical Richardson number and the virtual temperature at the lowest model level and the surface. The derived height may be larger in MEDMF-TKE due to the more moist (and warmer, not shown) low-level inflow and larger vertical advection of TKE. This is consistent with the stronger vertical velocity and greater eyewall reflectivity above the PBL seen in the composites. It is somewhat counterintuitive that the two metrics of PBL height differ in this way, but given the fact that the two schemes calculate height the same way, it is another illustration of how the MEDMF-TKE scheme tends to produce shallower and stronger inflow structure in TC environments, which also leads to greater influx of heat and moisture as well as larger TKE extending to a deeper layer.

e. TC case studies

The composites and structure comparison statistics of all individual cases discussed above are useful to examine how MEDMF-TKE and HEDMF-KP performed throughout the 2020 hurricane season, and to understand the overall characteristics of the different schemes. To further illustrate some of these findings and how they manifest in individual TCs, two case studies are examined next. These provide more insight into the behavior of the two schemes and their impacts, including possible reasons for relatively large differences in some cases.

1) HURRICANE LAURA

The first case study examined is Hurricane Laura. Laura was one of the strongest Atlantic TCs of 2020, making landfall near Lake Charles, Louisiana, as a category-4 hurricane in late August. The forecasts for Laura from HAFS-globalnest with MEDMF-TKE and HEDMF-KP highlighted some of the differences in track between the two configurations and the observed track (Fig. 12a). Both versions of the model had a right bias in early forecasts. However, the forecasts with MEDMF-TKE kept the TC closer to the observed track. Both forecast intensities were too weak (Fig. 12b) in this case, at least partly due to excessive land interaction caused by the track differences, but the forecast with MEDMF-TKE made landfall in approximately the correct place with a peak intensity at the correct time. Some of the track differences were attributable to differences in the subtropical ridge strength between the two schemes (Figs. 12c–e). At 72 h (initialized at 0000 UTC 22 August), the 976-dam 300-hPa height contour showing the western edge of the subtropical ridge was too far north compared to the GFS analysis in both versions of the model. Some of the differences in ridge positioning may have been due to the differences in the track itself, but the center of the ridge (denoted by the 978-dam line) was notably

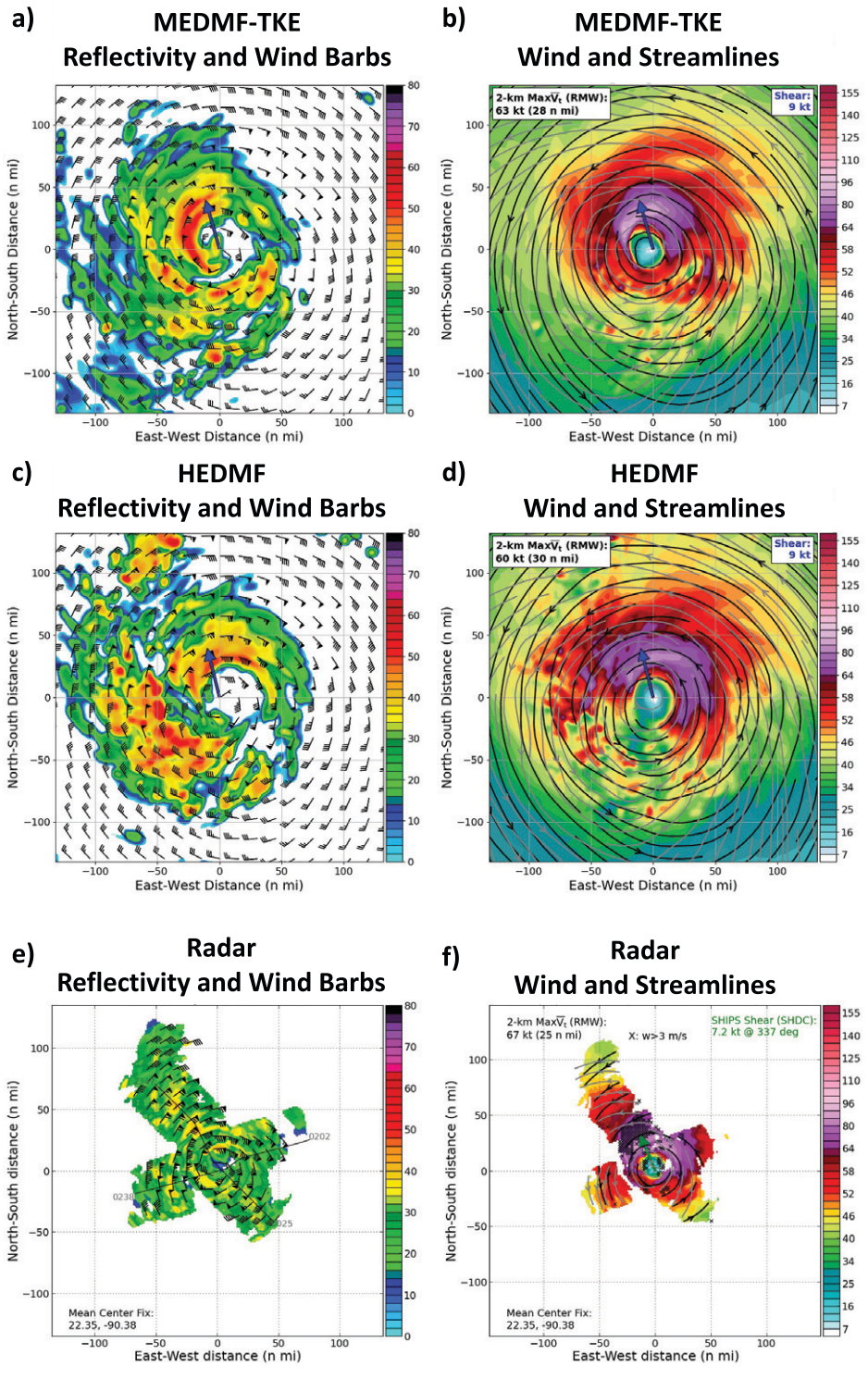

FIG. 14. (a) 2-km reflectivity (dBZ) and wind barbs from the HAFS-globalnest forecast for Hurricane Delta with MEDMF-TKE initialized at 1200 UTC 7 Oct, valid at 12 h. The arrow shows the 850–200-hPa shear vector. (b) From the same forecast as in (a), but showing 2-km wind speed (shaded), and 2-km (black) and 5-km (gray) streamlines. (c) As in (a), but for the forecast with HEDMF-KP. (d) As in (b), but for the forecast with HEDMF-KP. (e) As in (a), but for P-3 observations of Delta between 2337 UTC 7 Oct and 0238 UTC 8 Oct. (f) As in (b), but for P-3 observations of Delta between 2337 UTC 7 Oct and 0238 UTC 8 Oct.

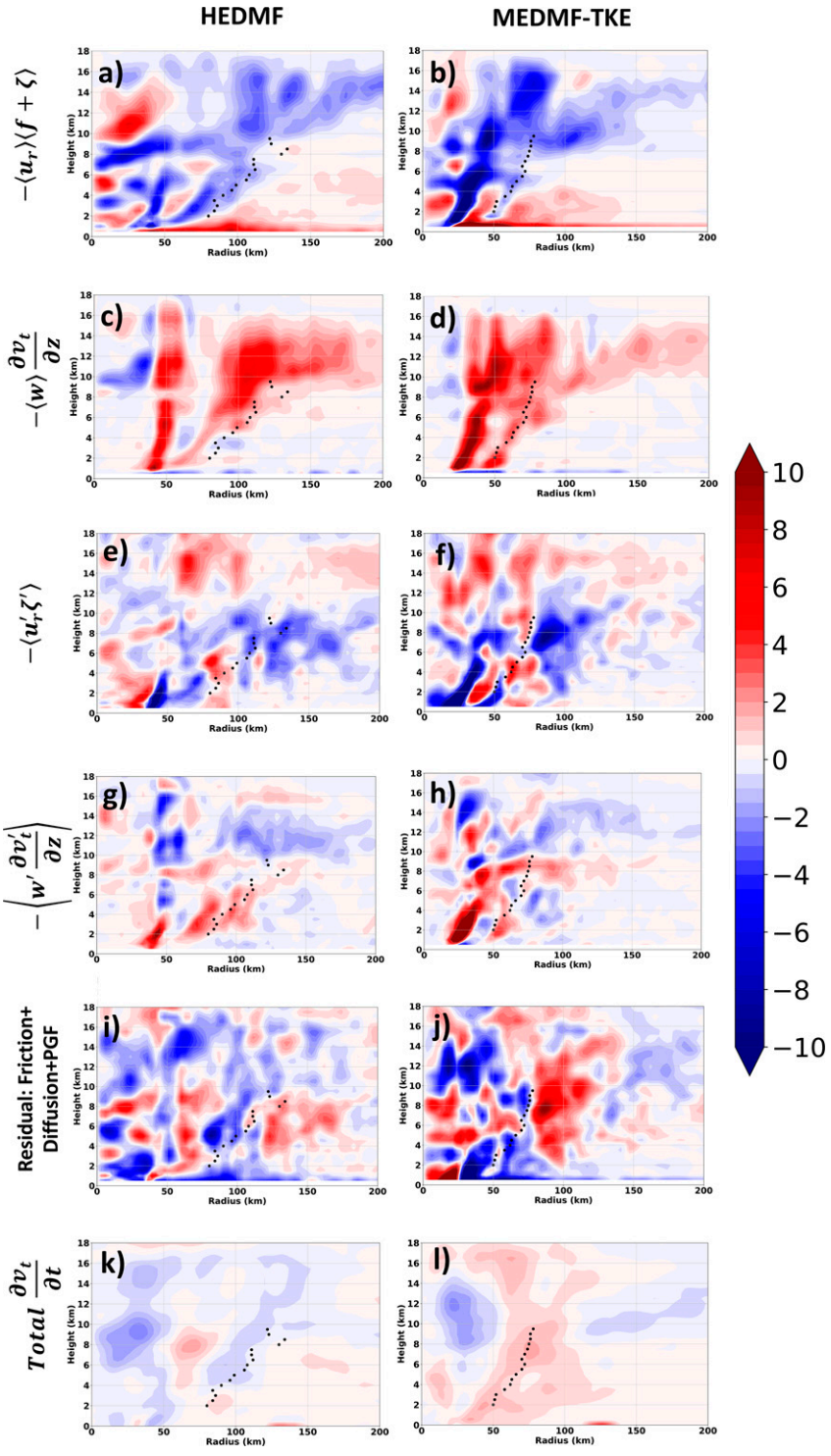

FIG. 15. Tangential wind tendency equation terms (m s^{-1} h^{-1}) for the forecast with (right) MEDMF-TKE and (left) HEDMF-KP for Hurricane Delta initialized at 1200 UTC 7 Oct, valid at 15 h. The terms are the following: (a),(b) mean radial influx of mean absolute vorticity; (c),(d) mean vertical advection of mean tangential momentum; (e),(f) eddy flux of eddy momentum; (g),(h) vertical eddy flux of tangential momentum; (i),(j) the residual term including the pressure gradient force and diffusivity effects; and (k),(l) the total tangential wind tendency. The azimuthal mean RMW is shown in the black dots.

stronger in the run with MEDMF-TKE, which kept Laura farther southwest than in the HEDMF-KP run. At 96 h and then 120 h (Figs. 12d,e), as the tracks began to diverge significantly, the ridge position in the run with MEDMF-TKE was notably stronger closer to that of the GFS analysis. It is difficult to totally disentangle the cause and effect of differences in the position of the western edge of the ridge and the location of the TC itself, but the large differences in the ridge over the Bahamas and Western Atlantic (away from the TC) show that the difference in PBL schemes was enough to affect the placement and strengths of ridges, and not just the TC structure. Future work will explore the physical basis for this ridge difference in detail.

2) Hurricane Delta

Hurricane Delta was an October hurricane that formed in the Caribbean and underwent an initial period of RI followed by rapid weakening, and made landfall as a category-2 hurricane along the Yucatan Peninsula. It intensified into a major hurricane again over the Gulf of Mexico and then weakened slightly due to vertical shear and made landfall as a category 2 near Lake Charles, Louisiana. The analysis here focuses on the intensification period over the Gulf of Mexico, where there were large differences in the intensity between the forecasts with the two PBL schemes. Figure 13 shows the track and intensity forecasts for Delta from 1200 UTC 7 October from HAFS-globalnest with MEDMF-TKE and HEDMF-KP. These forecasts were initialized right after Delta had made landfall in the Yucatan Peninsula. The track forecasts were very accurate in both versions of the model. The forecasts were very different in structure and intensity, however. The forecast with MEDMF-TKE correctly showed Delta quickly intensifying into a major hurricane with peak winds of 105 kt within 36 h before weakening just prior to landfall at 60 h), while the storm in the HEDMF-KP forecast intensified much more slowly and only showed a peak intensity of approximately 80 kt. Note that the large apparent difference between the storm intensity of HAFS and best track at initial time is because the TC was initialized over land, which affected the derivation of model-derived 10-m wind that is used for these intensity plots. Examining the impact of the HAFS land surface model on TC wind over land is a planned subject of future investigation.

There were several NOAA P-3 flights into Hurricane Delta, which provided the opportunity to evaluate the forecast structure through comparison with radar data collected by the aircraft. A comparison for this model cycle is shown in Fig. 14. The large-scale shear is nearly the same in both forecasts and observations, with less than 10 kt of shear from the south-southeast. The storm structure of the two forecasts is also fairly similar, but there are some subtly key differences that likely impacted the intensification of Delta. The wind field was more compact and focused near the inner core in MEDMF-TKE than HEDMF-KP (Figs. 14b,d). The heavy convection and precipitation was focused near Delta's eyewall in MEDMF-TKE and observations, while in HEDMF-KP the precipitation in the rainbands outside of the eyewall was heavier (Figs. 14a,c). As discussed in several studies (e.g., Rogers et al. 2013; Hazelton et al. 2017; Zhang and Rogers 2019), more symmetric convection concentrated inside the RMW tends to favor intensification more than outer core convection. This is also consistent with the composite results above, with reflectivity typically higher in the eyewall in MEDMF-TKE. The precipitation was also more symmetric on the upshear (southeast) side in MEDMF-TKE, which was more consistent with observations (Fig. 14e). Symmetric precipitation like this has been shown to be an important precursor for RI in some cases (e.g., Fischer et al. 2018; Alvey et al. 2020).

The comparison with airborne radar demonstrated the importance of the smaller core in MEDMF-TKE in the Hurricane Delta case. To illustrate exactly how the two schemes differed in the spinup of Hurricane Delta during this forecast cycle, we examine a budget of the tangential wind tendency, given by

$$\frac{\partial v_t}{\partial t} = -\langle u_r \rangle \langle f + \zeta \rangle - \langle w \rangle \frac{\partial \langle v_t \rangle}{\partial z} - \langle u_r' \zeta' \rangle - \left\langle w' \frac{\partial v_t'}{\partial z} \right\rangle + \left\langle \frac{1}{\rho r} \frac{\partial p'}{\partial \lambda} \right\rangle + \langle D_h \rangle + \langle D_v \rangle. \quad (3.1)$$

The terms on the right-hand side represent, respectively, the mean spinup by mean radial influx of mean absolute vorticity, the mean vertical advection of mean tangential momentum, the spinup due to eddy flux of vorticity by eddy radial flow, the eddy vertical advection of the eddy angular momentum, the perturbation pressure gradient force, and the diffusion terms (both horizontal and vertical). The angle brackets indicate terms that are azimuthally averaged. A similar form of this equation has been used to understand TC spinup in both azimuthal mean (Smith et al. 2017) and asymmetric (Leighton et al. 2018; Green et al. 2021) frameworks. Figure 15 shows the first four individual terms from Eq. (3.1), a residual term encompassing the last three terms, and the total tangential wind tendency for both schemes used to forecast Hurricane Delta at 15 h. The tangential wind tendency is much greater in the run with MEDMF-TKE than the run with HEDMF-KP and focused inside the RMW, indicating more efficient spinup and contraction of the vortex. In the PBL, the term that contributes most strongly is the mean radial flux of vorticity. This term is positive in the PBL in both schemes but much stronger in the run with MEDMF-TKE due to stronger inflow (consistent with the composite results). The mean radial eddy flux term and the residual terms show a more random pattern and largely tend to offset. Above the PBL, the mean and eddy vertical advection are both positive, showing the importance of these terms for spinning up the TC in the midlevels. As the comparison with radar showed, these terms are likely larger in MEDMF-TKE due to the concentration of upward vertical motion near the TC core. In MEDMF-TKE, the stronger PBL inflow brought more angular momentum inward and led to more upward motion near the core, helping to spin up the low-level vortex and then advect that momentum vertically. Some of the same processes occurred in HEDMF-KP, but the different PBL

structure and weaker influx of momentum meant that they were slower to spin up the vortex. Turbulence above the PBL in the eyewall can be a key source of differences in intensity and structure, as shown by Zhu et al. (2019, 2021), and this appears to be reflected in the differences in the spinup aloft between their schemes.

4. Discussion, conclusions, and future work

This study demonstrates the performance of two PBL schemes in the NOAA's Hurricane Analysis and Forecast System and helps to motivate further improvements. Performances of the *K*-profile Hybrid-EDMF (HEDMF-KP) scheme, used in the operational HWRF and initial tests of HAFS, and the EDMF scheme with the prognostic TKE (EDMF-TKE), used in the operational GFS as of early 2021 and tested (in a modified form, MEDMF-TKE, with reduced mixing length) in the 2020 real-time HAFS-globalnest experiment (Hazelton et al. 2021b) are evaluated. A large sample of cases from the 2020 season is examined, to see how the biases in TC forecast metrics, as well as TC structure, vary between the two sets of forecasts. This will inform future development of PBL schemes in HAFS.

Track statistics demonstrated that MEDMF-TKE performed better than HEDMF-KP overall, with a reduced right-of-track bias in several cases, including Hurricane Laura in August 2020. In the Laura case, this was due to a better representation of the subtropical ridge, specifically a slightly stronger ridge to the north of Laura, in the MEDMF-TKE forecasts. The overall comparison of intensity statistics was mixed, but MEDMF-TKE produced better forecasts of TC rapid intensification. The lack of ocean coupling in the current configuration of HAFS-globalnest may be responsible for some of the positive intensity bias at longer lead times, as some preliminary tests with a coupled version of stand-alone-regional HAFS (HAFS-SAR, not shown) showed very promising intensity improvement, with reduced bias at longer lead times (forecast days 4 and onward).

There were notable differences in TC structure as well, shown in both the composites and the calculated structure metrics in all individual cases. MEDMF-TKE tended to produce storms that had smaller RMW, deeper and narrower vortices, and a more robust warm-core response, especially at upper levels. The composites of radial wind and radial convergence showed stronger and narrower inflow in MEDMF-TKE than HEDMF-KP, and enhanced convergence radially inward of the RMW. This was reflected in the vertical structure as well, with stronger and more concentrated eyewall updrafts and a more moist eyewall in MEDMF-TKE than in HEDMF-KP. These structure differences are consistent with the intensity bias differences, with MEDMF-TKE producing stronger TCs on average. The enhanced inflow advecting more absolute angular momentum toward the TC center promotes greater spinup of the azimuthal mean vortex (Smith and Montgomery 2015), and the stronger convergence inside the RMW favors stronger updrafts and therefore a stronger TC (Zhang et al. 2017; Zhang and Rogers 2019). Similar results have been seen in observational composites using both dropwindsonde and radar data (e.g., Ahern et al. 2019; Rogers et al. 2013). This results in a TC with stronger maximum tangential wind near the RMW and a drier but warmer eye in MEDMF-TKE.

The analysis of the tangential wind tendency budget illustrates how several of the processes shown in the composite structure contribute to the spinup of the TC vortex, a process that was typically stronger in MEDMF-TKE:

1) Azimuthal mean radial inflow transports absolute angular momentum inward and helps spin up the mean TC circulation. Stronger and shallower inflow tends to enhance this process, as seen in MEDMF-TKE.
2) Stronger azimuthal mean inflow leads to stronger radial convergence and an enhancement of the mean eyewall updraft, which helps advect more momentum vertically and spin up the TC through the depth of the troposphere. Enhanced latent heat release and forced subsidence in the eye lead to a more robust warm core, lowering the central pressure and favoring further enhancement of the low-level inflow.
3) The stronger low-level inflow and convergence also favor small-scale updrafts that are focused inside the RMW, and these features contribute to eddy spinup of the tangential wind field, and also tend to warm the TC core and help lower the pressure when they are focused inside the RMW.

Ongoing work is focused on continued improvement to the EDMF-TKE scheme. The mixing length tuning method described in Gopalakrishnan et al. (2021) will be evaluated through comparison with additional observations (e.g., Zawislak et al. 2022) and large eddy simulations (e.g., Bryan et al. 2017). Methods for including parameters in the PBL schemes to account for the effects of wind shear on PBL structure (Rodier et al. 2017) will also be explored. Also, Fig. 1 showed that the inflow angle (and to some extent the eddy diffusivity) for MEDMF-TKE was actually farther away from observations than HEDMF, yet it produced better forecasts of rapid intensification and structure. This indicates that there needs to be further examination of the details of how the TC structure is reacting to the PBL physics (and other physics), including comparisons with more extensive observational datasets that can provide better information about the vertical profile of eddy diffusivity as well as variability across different kinds of TCs. We also plan to develop a larger and more detailed dropsonde composite of inflow angle to examine the dependency of inflow angle on intensity and location in the TC, which will allow for better comparison with model data. Finally, a recent study using HAFS (Zhu et al. 2021) has demonstrated that, in the EDMF-KP scheme, the formulation of the Brunt–Väisälä frequency can impact turbulence in the eyewall above the PBL, and have a significant effect on TC structure and intensity. It will be worthwhile to examine how stability and turbulence above the PBL are handled in the MEDMF-TKE scheme and whether some of the same modifications applied to EDMF-KP will be useful in the TKE-based scheme as well. Investigation of possible interaction between

turbulent mixing schemes within and above the PBL and its effects on TC intensity change is also recommended. Finally, we plan to examine how the choice and configuration of PBL scheme, and the interaction between the PBL scheme and other physics (such as the convective scheme) impacts large-scale synoptic flow (and biases in the flow) in HAFS, to improve medium-to-long-range TC track forecasts.

Acknowledgments. The authors thank the crews of the NOAA P-3 aircraft for their tireless efforts to collect TC observations, including those used to evaluate the forecasts in this study. The authors thank Trey Alvey, Kyle Ahern, and two anonymous reviewers whose comments helped improve early versions of the manuscript. The lead author was supported by the NOAA Grant NA19OAR0220187. Jun Zhang was supported by NOAA Grant NA19OAR0220186 and ONR Grant N00014-20-1-2071.

Data availability statement. Track files, grb2 data files, and other data from these model runs are available on the NOAA RDHPCS computer system, or by request.

REFERENCES

Ahern, K., M. A. Bourassa, R. E. Hart, J. A. Zhang, and R. F. Rogers, 2019: Observed kinematic and thermodynamic structure in the hurricane boundary layer during intensity change. *Mon. Wea. Rev.*, **147**, 2765–2785, https://doi.org/10.1175/MWR-D-18-0380.1.

Alvey, G. R., E. Zipser, and J. Zawislak, 2020: How does Hurricane Edouard (2014) evolve toward symmetry before rapid intensification? A high-resolution ensemble study. *J. Atmos. Sci.*, **77**, 1329–1351, https://doi.org/10.1175/JAS-D-18-0355.1.

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., 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.

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

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

Chen, X., M. Xue, and J. Fang, 2018: Rapid intensification of Typhoon Mujigae (2015) under different sea surface temperatures: Structural changes leading to rapid intensification. *J. Atmos. Sci.*, **75**, 4313–4335, https://doi.org/10.1175/JAS-D-18-0017.1.

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, K. L. Corbosiero, and C. M. Rozoff, 2018: Normalized convective characteristics of tropical cyclone rapid intensification events in the North Atlantic and eastern North Pacific. *Mon. Wea. Rev.*, **146**, 1133–1155, https://doi.org/10.1175/MWR-D-17-0239.1.

Franklin, J. L., M. L. Black, and K. Valde, 2003: GPS dropwindsonde wind profiles in hurricanes and their operational implications. *Wea. Forecasting*, **18**, 32–44, https://doi.org/10.1175/1520-0434(2003)018<0032:GDWPIH>2.0.CO;2.

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.

——, and Coauthors, 2020: 2019 HFIP R&D activities summary: Recent results and operational implementation, HFIP Tech. Rep. HFIP2020-1, 42 pp., https://repository.library.noaa.gov/view/noaa/26468.

——, A. T. Hazelton, and J. A. Zhang, 2021: Improving hurricane boundary layer parameterization scheme based on observations. *Earth Space Sci.*, **8**, e2020EA001422, https://doi.org/10.1029/2020EA001422.

Green, A., S. G. Gopalakrishnan, G. J. Alaka, and S. Chiao, 2021: Understanding the role of mean and eddy momentum transport in the rapid intensification of Hurricane Irma (2017) and Hurricane Michael (2018). *Atmosphere*, **12**, 492, https://doi.org/10.3390/atmos12040492.

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.

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

——, W. Wang, Y. C. Kwon, S.-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.

Harris, L. M., and S.-J. Lin, 2012: 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. F. Rogers, 2017: Analyzing simulated convective bursts in two Atlantic hurricanes. Part II: Intensity change due to bursts. *Mon. Wea. Rev.*, **145**, 3095–3117, https://doi.org/10.1175/MWR-D-16-0268.1.

——, 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.

——, and Coauthors, 2021a: 2019 Atlantic hurricane forecasts from the global-nested Hurricane Analysis and Forecast System (HAFS): Composite statistics and key events. *Wea. Forecasting*, **36**, 519–538, https://doi.org/10.1175/WAF-D-20-0044.1.

——, and Coauthors, 2021b: 2020 real-time forecasts from the global-nested Hurricane Analysis and Forecast System (HAFSV0.1B). *Fourth Special Symp. on Tropical Meteorology and Tropical Cyclone Research and Forecasting: Multiscale Processes and Model Development Part II*,

virtual, Amer. Meteor. Soc., 12.5, https://ams.confex.com/ams/101ANNUAL/meetingapp.cgi/Paper/382062.

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

Kepert, J., 2001: The dynamics of boundary layer jets within the tropical cyclone core. Part I: Linear theory. *J. Atmos. Sci.*, **58**, 2469–2484, https://doi.org/10.1175/1520-0469(2001)058<2469:TDOBLJ>2.0.CO;2.

——, 2010: Slab-and height-resolving models of the tropical cyclone boundary layer. Part I: Comparing the simulations. *Quart. J. Roy. Meteor. Soc.*, **136**, 1686–1699, https://doi.org/10.1002/qj.667.

——, and Y. Wang, 2001: The dynamics of boundary layer jets within the tropical cyclone core. Part II: Nonlinear enhancement. *J. Atmos. Sci.*, **58**, 2485–2501, https://doi.org/10.1175/1520-0469(2001)058<2485:TDOBLJ>2.0.CO;2.

Leighton, H., S. Gopalakrishnan, J. A. Zhang, R. F. Rogers, Z. Zhang, and V. Tallapragada, 2018: Azimuthal distribution of deep convection, environmental factors, and tropical cyclone rapid intensification: A perspective from HWRF ensemble forecasts of Hurricane Edouard (2014). *J. Atmos. Sci.*, **75**, 275–295, https://doi.org/10.1175/JAS-D-17-0171.1.

Mallen, K. J., M. T. Montgomery, and B. Wang, 2005: Reexamining the near-core radial structure of the tropical cyclone primary circulation: Implications for vortex resiliency. *J. Atmos. Sci.*, **62**, 408–425, https://doi.org/10.1175/JAS-3377.1.

Marchok, T. P., 2002: How the NCEP tropical cyclone tracker works. *25th Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., P1.13, https://ams.confex.com/ams/25HURR/techprogram/paper_37628.htm.

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

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.

Rodier, Q., V. Masson, F. Couvreux, and A. Paci, 2017: Evaluation of a buoyancy and shear based mixing length for a turbulence scheme. *Front. Earth Sci.*, **5**, 65, https://doi.org/10.3389/feart.2017.00065.

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

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

Shapiro, L. J., and H. E. Willoughby, 1982: The response of balanced hurricanes to local sources of heat and momentum. *J. Atmos. Sci.*, **39**, 378–394, https://doi.org/10.1175/1520-0469(1982)039<0378:TROBHT>2.0.CO;2.

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.

——, J. A. Zhang, and M. T. Montgomery, 2017: The dynamics of intensification in a hurricane weather research and forecasting simulation of Hurricane Earl (2010). *Quart. J. Roy. Meteor. Soc.*, **143**, 293–308, https://doi.org/10.1002/qj.2922.

Stern, D. P., and D. S. Nolan, 2009: Reexamining the vertical structure of tangential winds in tropical cyclones: Observations and theory. *J. Atmos. Sci.*, **66**, 3579–3600, https://doi.org/10.1175/2009JAS2916.1.

——, and ——, 2012: On the height of the warm core in tropical cyclones. *J. Atmos. Sci.*, **69**, 1657–1680, https://doi.org/10.1175/JAS-D-11-010.1.

——, J. R. Brisbois, and D. S. Nolan, 2014: An expanded dataset of hurricane eyewall sizes and slopes. *J. Atmos. Sci.*, **71**, 2747–2762, https://doi.org/10.1175/JAS-D-13-0302.1.

Tallapragada, V., C. Kieu, Y. Kwon, S. Trahan, Q. Liu, Z. Zhang, and I. Kwon, 2014: Evaluation of storm structure from the operational HWRF Model during 2012 implementation. *Mon. Wea. Rev.*, **142**, 4308–4325, https://doi.org/10.1175/MWR-D-13-00010.1.

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

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

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.

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

——, F. D. Marks, M. T. Montgomery, and S. Lorsolo, 2011: An estimation of turbulent characteristics in the low-level region of intense Hurricanes Allen (1980) and Hugo (1989). *Mon. Wea. Rev.*, **139**, 1447–1462, https://doi.org/10.1175/2010MWR3435.1.

——, 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.

——, 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.

——, E. A. Kalina, M. K. Biswas, R. F. Rogers, P. Zhu, and F. D. Marks, 2020: A review and evaluation of planetary boundary layer parameterizations in Hurricane Weather Research and Forecasting model using idealized simulations

and observations. *Atmosphere*, **11**, 1091, https://doi.org/10.3390/atmos11101091.

Zhou, L., S.-J. Lin, J.-H. Chen, L. M. Harris, X. Chen, and S. L. Rees, 2019: Toward convective-scale prediction within the next generation global prediction system. *Bull. Amer. Meteor. Soc.*, **100**, 1225–1243, https://doi.org/10.1175/BAMS-D-17-0246.1.

Zhu, P., B. Tyner, J. A. Zhang, E. Aligo, S. Gopalakrishnan, F. D. Marks, A. Mehra, and V. Tallapragada, 2019: Role of eyewall and rainband eddy forcing in tropical cyclone intensification. *Atmos. Chem. Phys.*, **19**, 14 289–14 310, https://doi.org/10.5194/acp-19-14289-2019.

——, A. Hazelton, Z. Zhang, F. D. Marks, and V. Tallapragada, 2021: The role of eyewall turbulent transport in the pathway to intensification of tropical cyclones. *J. Geophys. Res. Atmos.*, **126**, e2021JD034983, https://doi.org/10.1029/2021JD034983.