Evaluating the Impact of Improvement in the Horizontal Diffusion Parameterization on Hurricane Prediction in the Operational Hurricane Weather Research and Forecast (HWRF) Model

Jun A. Zhang

NOAA/AOML/Hurricane Research Division, and University of Miami/Cooperative Institute for Marine and Atmospheric Studies, Miami, Florida

Frank D. Marks, Jason A. Sippel, and Robert F. Rogers

NOAA/AOML/Hurricane Research Division, Miami, Florida

Xuejin Zhang

NOAA/AOML/Hurricane Research Division, and University of Miami/Cooperative Institute for Marine and Atmospheric Studies, Miami, Florida

Sundararaman G. Gopalakrishnan

NOAA/AOML/Hurricane Research Division, Miami, Florida

Zhan Zhang and Vijay Tallapragada

NOAA/NCEP/Environmental Modeling Center, College Park, Maryland

ABSTRACT

Improving physical parameterizations in forecast models is essential for hurricane prediction. This study documents the upgrade of horizontal diffusion parameterization in the Hurricane Weather Research and Forecasting (HWRF) Model and evaluates the impact of this upgrade on hurricane forecasts. The horizontal mixing length L_h was modified based on aircraft observations and extensive idealized and real-case numerical experiments. Following an earlier work by the first two authors, who focused on understanding how the horizontal diffusion parameterization worked in HWRF and its dynamical influence on hurricane intensification using idealized simulations, a series of sensitivity experiments was conducted to simulate Hurricane Earl (2010) in which only L_h was varied. Results from the Earl forecasts confirmed the findings from previous theoretical and idealized numerical studies, in that both the simulated maximum intensity and intensity change rate are dependent on L_h. Comparisons between the modeled and observed structure of Hurricane Earl, such as storm size, boundary layer heights, warm-core height and temperature anomaly, and eyewall slope, suggested that the L_h used in the HWRF Model should be decreased. Lowering L_h in HWRF has a positive impact on hurricane prediction based on over 200 retrospective forecasts of 10 Atlantic storms. Biases in both storm intensity and storm size are significantly reduced with the modified L_h.

1. Introduction

As the horizontal resolution of operational hurricane models becomes smaller, improving physical parameterizations of subgrid-scale physical processes such as turbulent mixing becomes more important for hurricane prediction. Turbulent mixing in the vertical direction, especially in the atmospheric boundary layer, is well known to be critical in simulations and forecasts of hurricane intensity and structure (Braun and Tao 2000; Foster 2009; Smith and Thomsen 2010; Kepert 2012; Bao et al. 2012; Gopalakrishnan et al. 2013; Zhu et al. 2014; Zhang et al. 2015, 2017). As part of the

Hurricane Forecast Improvement Project (HFIP), recent advances in the parameterization of vertical turbulent mixing (i.e., vertical eddy diffusivity) in the operational Hurricane Weather Research and Forecasting (HWRF) Model, based on aircraft observations reported by Zhang et al. (2011a) and Zhang and Drennan (2012), have led to significant improvements in intensity and track forecasts (Tallapragada et al. 2014). The abovementioned physics improvement in HWRF followed the developmental framework proposed by Zhang et al. (2012), which consisted of four components: model diagnostics, physics development, physics implementation, and further evaluation. Building on the success of this approach, here we adapt the same framework to improve the parameterization of horizontal turbulent mixing (i.e., horizontal diffusion) in HWRF.

Previous studies have demonstrated that horizontal diffusion is an important aspect of the model physics for hurricane simulations. Bryan and Rotunno (2009, hereafter BR09) used an axisymmetric numerical cloud model [Cloud Model 1 (CM1)] to point out that numerical simulations of hurricane maximum potential intensity (MPI) are sensitive to horizontal diffusion in terms of horizontal mixing length L_h. Bryan et al. (2010) confirmed the findings of BR09 by using the three-dimensional CM1 simulations. Bryan (2012) showed that the impact of L_h on MPI can be as significant as the ratio of the surface exchange coefficient that was emphasized in the MPI theory of Emanuel (1995). Rotunno and Bryan (2012, hereafter BR12) further investigated the effects of both vertical and horizontal diffusion on hurricane MPI and structure, offering interpretations of why horizontal diffusion is an important element of hurricane dynamics related to intensity.

Motivated by BR09 and RB12, Zhang and Marks (2015, hereafter ZM15) studied the effect of L_h on hurricane intensity change and structure in idealized 3D simulations using the HWRF Model. They presented a series of 5-day simulations within the context of operational forecasts, instead of simulations with longer periods (i.e., ≥12 days) as used by BR09 and RB12. Nonetheless, the results of ZM15 supported the conclusions of BR09 and RB12 that larger values of L_h were associated with smaller maximum intensity. These studies (i.e., BR09, RB12, and ZM15) all agreed that the main impact of larger L_h on the intensification of a hurricane is to smooth the radial gradient of angular momentum and equivalent potential temperature in the eyewall region. In the simulations with smaller L_h, both the radial inflow and radial gradient of angular momentum are larger than in simulations with larger L_h,

which leads to a larger radial advection of the angular momentum and, in turn, more rapid hurricane intensity change, consistent with the spinup dynamics as discussed in Ooyama (1969), Smith et al. (2009), and Smith and Montgomery (2015). ZM15 also suggested that the extent of the sensitivity of maximum intensity and intensity change rate to L_h became smaller when the model horizontal resolution (i.e., grid spacing) was smaller than L_h.

Observational studies on the horizontal diffusion in hurricane conditions have been limited. Zhang and Montgomery (2012) presented the first observational analysis of horizontal momentum flux, horizontal eddy diffusivity, and L_h using flight-level data collected during four hurricanes (Allen, 1980; David, 1986; Hugo, 1989; and Frances, 2004) at an altitude of ~500 m. Their estimates of L_h were conducted within the framework of a simple K theory that assumed the stationarity and homogeneity of turbulent flow in the intense eyewall, while momentum was assumed to be transferred downgradient of the mean flow. On average, Zhang and Montgomery (2012) found that L_h is approximately 750 m in the hurricane eyewall. They also found that while there is a weak tendency for L_h to increase with the wind speed, this increase was not statistically significant. Note also that the optimal value of L_h recommended by Bryan et al. (2010) for hurricane model prediction is 1000 m, a result based on a large number of 3D CM1 simulations. This value of L_h is much closer to the observational estimate of Zhang and Montgomery (2012) than the value of L_h recommended by BR09 (1500 m) based on 2D simulations. The abovementioned theoretical, numerical, and observational studies provided guidance for improving L_h in the operational HWRF model.

As pointed out by ZM15, the value of L_h used in the 2015 version of the operational HWRF model (H215) was 1900 m, and L_h was >2000 m in earlier versions of HWRF because of the larger horizontal model resolution. Following the recommendation of ZM15 and real-case HWRF simulations presented in section 3, L_h was reduced in the 2016 version of the operational HWRF model (H216). The value of L_h used in H216 is 800 m, which is much closer to the observational estimate of Zhang and Montgomery (2012) and the value recommended by Bryan et al. (2010), than that used in H215. The present study documents the model upgrade process for the horizontal diffusion parameterization in HWRF and evaluates the impact of this improvement in L_h on HWRF's ability to predict track, intensity, and structure using extensive retrospective forecasts. For the first time, the effects of L_h on hurricane intensity and structure are evaluated in real-case hurricane forecasts and simulations.

2. Horizontal diffusion parameterization in HWRF

The description of the horizontal diffusion parameterization in the HWRF Model given below parallels that given by ZM15. Details of other aspects of the HWRF Model physics can be found in Tallapragada et al. (2014) and Zhang et al. (2015). The horizontal diffusion scheme in HWRF is essentially the typical second-order nonlinear Smagorinsky-type parameterization as detailed by Janjić (1990). Horizontal eddy diffusivity K_h is used to parameterize the horizontal turbulent momentum flux[1] F_h in the form of

$$F_h = \rho K_h S_h, \tag{1}$$

where S_h is the horizontal strain rate of the mean flow (e.g., Stevens et al. 1999), which has the form of

$$S_h = \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right), \tag{2}$$

where x and y are the distances in the longitudinal and latitudinal directions, respectively, and u and v are the velocity components in the longitudinal and latitudinal directions, respectively. In addition, K_h is parameterized as a function of L_h and the horizontal deformation D_h in the form of

$$K_h = L_h^2 D_h, \tag{3}$$

where D_h has the form of

$$D_h^2 = \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right)^2. \tag{4}$$

In HWRF's dynamical core (the so-called NMM) K_h is set to be proportional to the sum of the deformation and another term related to turbulent kinetic energy (TKE). Because the planetary boundary layer (PBL) parameterization scheme in HWRF is based on a K-profile method, which is not a TKE-type scheme, the term related to TKE is set to a very small value that is close to zero and is thus not shown in Eq. (3). Furthermore, L_h is tuned through a parameter "coac" in HWRF that was named as the Smagorinsky constant in Janjić (1990). Note that L_h is a function of both coac and horizontal spacing dx in HWRF in the form of $L_h^2 =$ coac dx, following Janjić (1990) and Janjić et al. (2010). Values of coac are set to be different for the three

---

[1] Note that the horizontal eddy diffusivities for heat K_t and moisture K_q are set to be the same as that for momentum (i.e., $K_t = K_q = K_h$).

domains to reflect the dependence of K_h on model grid spacing.

Without TKE, the parameterization of K_h in HWRF was essentially the same as that used in the CM1 model as well as in NCAR's Weather Research and Forecasting (WRF) Model (Janjić 2003), except that the method of tuning the Smagorinsky constant is slightly different. A similar horizontal diffusion parameterization was also used in the Tropical Cyclone Model version 4 (TCM4; Wang 2007). We also note that the equations used by Zhang and Montgomery (2012) for observational estimates of K_h and L_h follow the same concept of momentum-flux parameterization as those used in the abovementioned numerical models, except that there is no scale dependence in the L_h estimation in that study, because a constant cutoff frequency was used in their momentum flux calculation. Nonetheless, the observational data reported by Zhang and Montgomery (2012) would provide useful guidance for the range of values of L_h to be tuned in HWRF, being aware of the relatively large uncertainty (~50%) of L_h in their estimates. Developing a new horizontal diffusion parameterization using observational data including a TKE term would be ideal for a hurricane model that uses a TKE-type boundary layer scheme, but this is beyond the scope of the present study.

3. Case study on the sensitivity of Hurricane Earl (2010) forecasts to L_h

To confirm the findings of ZM15, we conducted a series of sensitivity experiments to simulate Hurricane Earl (2010). We chose the Earl case because excellent aircraft observations were gathered during this storm (e.g., Montgomery et al. 2014; Rogers et al. 2015) that provided the ground truth for evaluating the modeled hurricane structure. The same model configuration as in the idealized simulations of ZM15 was used in the Earl forecasts. The brief description of the model and numerical experiments below parallels that in ZM15. The HWRF model used here is triply nested, with the parent domain covering $50° \times 50°$, while the two inner nests cover $\sim 15° \times 15°$ and $\sim 5° \times 5°$, respectively. The grid spacings of the three domains are 27, 9, and 3 km, respectively. Two-way interactions are configured for the inner two domains. The model has 61 hybrid vertical levels in total and more than 10 vertical levels below 850 hPa. Details of the model configuration and other aspects of the physics can be found in ZM15.

Five HWRF forecasts of Earl were run by varying L_h only while keeping other physics options the same. Hurricane Earl originated from a tropical wave off the west coast of Africa on 23 August 2010 and underwent

rapid intensification between 28 and 30 August. The forecasts were initialized at 1800 UTC 26 August 2010. This cycle was chosen because of the excellent observational data available for model evaluation after the 24-h forecast. In the control simulation, L_h was set to 1900 m, which is the same as in H215. In the four other simulations, we set L_h to 375, 750, 1500, and 3000 m, respectively (referred to as L375, L750, L1500, and L3000 hereafter). The observational data used for model evaluation were collected by multiple research aircraft during Earl (Montgomery et al. 2014; Rogers et al. 2015). Doppler radar data were used to evaluate the vortex-scale structure. High-altitude dropsonde data from NASA's DC-8 aircraft were used to verify the warm-core structure. Dropsonde data at low levels collected by both the NASA DC-8 and NOAA WP-3D aircraft were used to evaluate the boundary layer structure. Note that Chen and Gopalakrishnan (2015) and Smith et al. (2017) used the same observational dataset to verify their Earl forecast of the same cycle that was run using an earlier version of the HWRF model (H212).

The simulated track and intensity for these five Earl forecasts are shown in Fig. 1. There is not much difference in the track forecasts (Fig. 1a), while the intensity forecast in terms of both maximum 10-m wind speed (Fig. 1b) and minimum sea level pressure (Fig. 1b) shows strong sensitivity to L_h. Figure 1 indicates that smaller values of L_h are associated with stronger storms. Additionally, the maximum hurricane intensity of the 5-day forecast is found to increase with decreasing L_h (Fig. 2a). The mean 24-h intensity change of the 5-day forecast also increases with decreasing L_h (Fig. 2b). This result indicates that storms simulated with smaller L_h tended to intensify faster than those with larger L_h, supporting the results of ZM15 and BR09. It is evident from Figs. 1 and 2 that both the maximum intensity and the mean intensification rate in L750 better match the observed values from the National Hurricane Center's (NHC) best-track data than those in the other experiments. This value of L_h is close to the observational estimates of Zhang and Montgomery (2012) and the numerical recommendation of Bryan et al. (2010), but is much smaller than that used in H215 (i.e., 1900 m). In the control experiment, where L_h as in H215 was used, both the maximum intensity and intensity change rate are underestimated compared to the best track. Below, we evaluate the simulated structure from these five experiments using observations.[2]

---

[2] Note that the remainder of this section describes structural sensitivity that has a basic agreement with stronger storms. For example, stronger storms have smaller RMW values, less tilt, a stronger warm-core anomaly, and less eyewall slope.
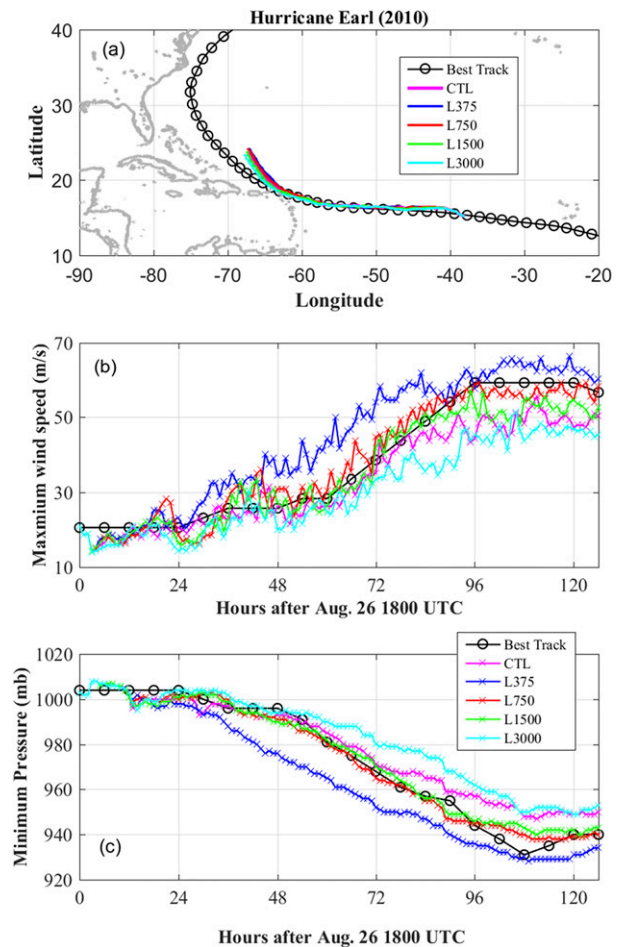

FIG. 1. Plots of (a) storm track, (b) maximum 10-m wind speed as a function of forecast time, and (c) minimum sea level pressure as a function of forecast time from five sensitivity experiments of Hurricane Earl (2010) varying L_h. Each line color represents each experiment, and the experiment name contains the value of L_h (m) used in the experiment. The control experiment (CTL) used L_h = 1900 m.

The observed radius–height structure of the azimuthal wind speed is compared to that averaged between 72 and 78 h in the five Earl forecasts (Fig. 3). During this period (1800 UTC 29 August–0000 UTC 30 August), a total of four aircraft [i.e., NOAA WP-3D and Gulfstream-IV (G-IV), U.S. Air Force C-130, and NASA DC-8] were flown into Earl simultaneously, collecting Doppler radar and dropsonde data with extensive coverage. Here, the quality-controlled Doppler radar data from the NOAA WP-3D aircraft are used to validate the vortex-scale structure in the Earl forecasts (Fig. 3a). The depth and strength of the hurricane vortex increase with decreasing L_h. For instance, the contour of 64 kt ($\sim$33 m s^{-1}) extends to 12-km altitude and 130-km radius in L375 (Fig. 3c), while the same contour in L3000 only extends to 2-km altitude and 100-km radius

(Fig. 3f). The maximum azimuthally averaged wind speeds for these two forecasts are 54 and $33 \, m \, s^{-1}$, respectively, differing by as much as $21 \, m \, s^{-1}$. The experiment with $L_h = 750 \, m$ (Fig. 3d) has a contour of 64 kt ($\sim 33 \, m \, s^{-1}$) up to 10-km altitude and 150-km radius, which is close to that in the observed vortex. The Earl forecasts in L750 and L1500 match the observed wind structure in terms of both the maximum wind speed and the depth of the vortex, which is much better than that in other experiments. This result suggests the control experiment used too large a value of L_h to obtain the correct vortex structure.

The storm size in terms of the radius of the maximum wind speed (RMW) in L375 and L750 is found to be closer to the radar-observed value than the other experiments (Fig. 4). It is also evident from Fig. 4 that the RMW averaged during the same period as in Fig. 3 increases with L_h. This result again supports ZM15. This increase in RMW is likely due to the weaker inflow in the boundary layer in simulations with larger L_h, as shown in Fig. 5, following the dynamical argument of Kilroy et al. (2016). Less horizontal diffusion yields stronger radial pressure gradients and stronger inflow in the boundary layer. With stronger inflow, air parcels travel closer to the storm center before being lifted in the eyewall region, favoring the contraction of the storm.

Vertical profiles of tangential V_t and radial V_r velocity averaged at the RMW and $2 \times$ RMW are compared to averaged dropsonde[3] data and are shown in Fig. 5 for the same period as in Fig. 3. It is evident from Fig. 5 that the simulated vertical V_t and V_r profiles from L750 and L1500 are mostly within the error bar (i.e., 95% confidence interval calculated based on raw sample statistics) of the observations, while those in the other forecasts, including the control experiment, are outside of the error bar. These two forecasts also better captured the observed peak values of V_t and V_r. In the control experiment, the vertical profile of V_t lies at the edge of the error bar in the eyewall region (i.e., RMW), but the values of V_t in the boundary layer at the outer radii ($2 \times$ RMW) are significantly lower than those in the dropsonde wind profiles, again implying that L_h used in H215 is too large.

The height of the maximum V_t and the inflow layer depth[4] in L750 and L1500 are closer to the dropsonde observations than those in the other experiments (Fig. 5). This result is corroborated by Fig. 6, where the boundary

---

[3] Dropsondes located within 5 km from the RMW are used for comparison. Nine and 12 dropsondes are averaged at the RMW and $2 \times$ RMW, respectively.

[4] The inflow-layer depth is taken as the height of 10% of the maximum inflow strength (Zhang et al. 2011b).
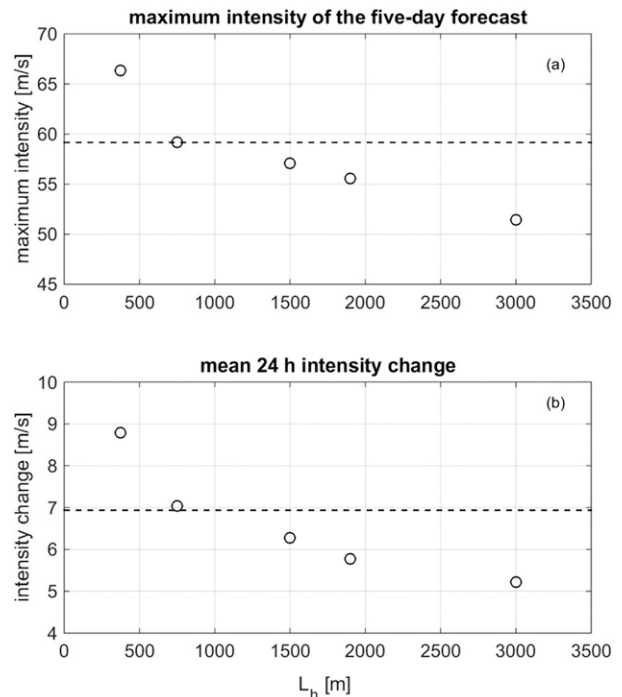

FIG. 2. Plots of (a) the maximum intensity in terms of maximum surface wind speed and (b) the mean 24-h intensity change as a function of L_h for the 5-day forecasts of Hurricane Earl (2010) initialized at 1800 UTC 26 Aug 2010. The black dashed line represents observed values. Note that $L_h = 1900 \, m$ was used in the control run.

layer heights are plotted as a function of L_h at the RMW. Both the height of the maximum V_t and the inflow layer depth are found to slowly increase with increasing L_h in the eyewall region for $L_h < 2000 \, m$, then increase much faster with L_h for larger L_h. Note that we also studied the thermodynamic mixed-layer depth that is taken as the height where the vertical gradient of the virtual potential temperature is higher than $3 \, K \, km^{-1}$ (Zeng et al. 2004; Zhang et al. 2011b) and found that the thermodynamic mixed-layer depth is generally independent of L_h (result not shown), in agreement with ZM15.

ZM15 also evaluated the vertical structure of the hurricane vortex in terms of eyewall slope, which is computed by linearly fitting the RMW at each level between 2 and 8 km, following Stern and Nolan (2009). According to balanced dynamics, the eyewall slopes outward because of the warm core that is maximized in the middle and upper troposphere (Shapiro and Willoughby 1982). When the eyewall slope is larger, the wind speed tends to decay faster with height. Thus, the eyewall slope can be a useful structural metric for evaluating the model depiction of the TC warm core and the subsequent response of the wind field. Figure 7 shows that the slope of the RMW gets larger with larger L_h, which again
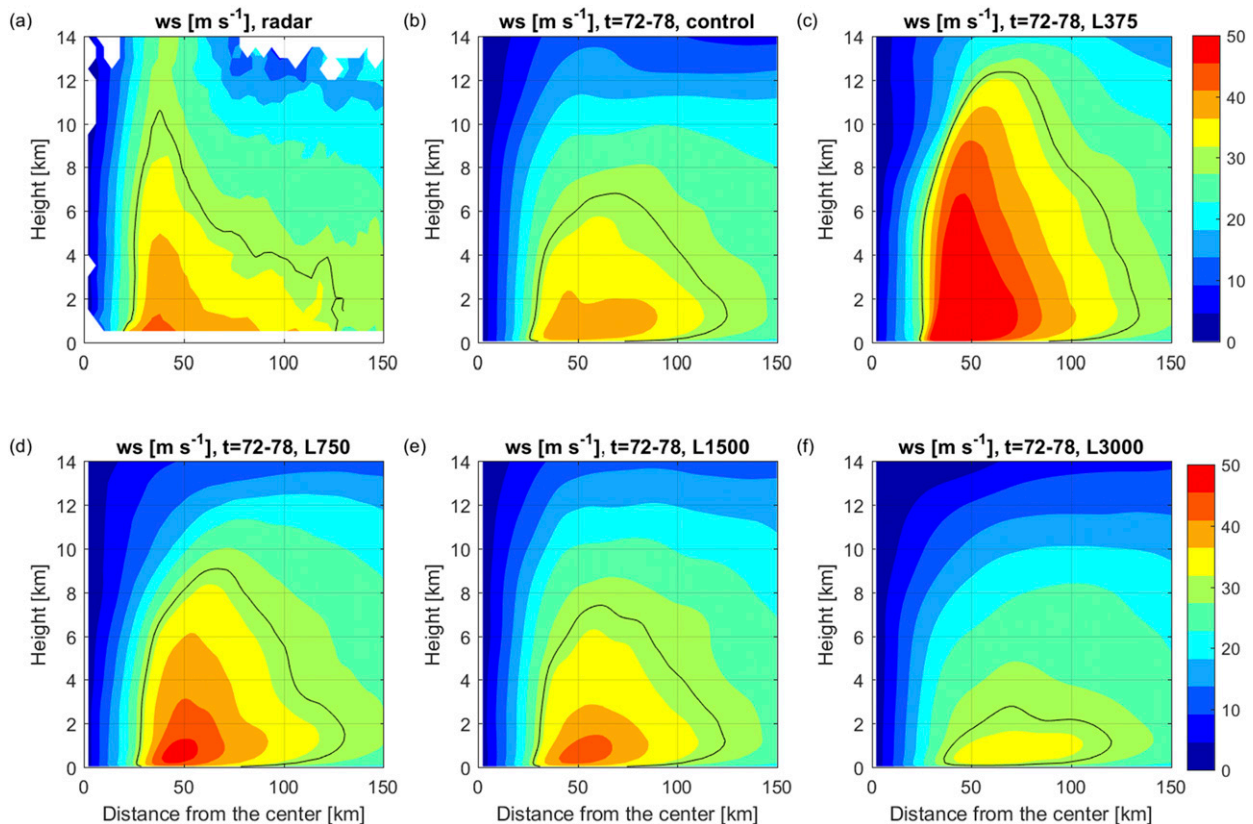
FIG. 3. Radius–height plots of the azimuthally averaged total wind speed over 72–78 h of the 5-day forecasts of Hurricane Earl (2010) initialized at 1800 UTC 26 Aug 2010, from (a) Doppler radar observations, (b) the control experiment (L_h = 1900 m), and the (c) L375, (d) L750, (e) L1500, and (f) L1500 experiments. The contour interval is 5 m s^{-1}, and the black contour has the value of 64 kt (33 m s^{-1}).

supports the finding of ZM15. Figures 4 and 7 together indicate that the eyewall slope increases with the RMW as the RMW also increases with L_h, which agrees with the observations documented by Stern and Nolan (2009) and Stern et al. (2014). It is also evident from Fig. 7 that the eyewall slope simulated in L750 is closest to that estimated based on the Doppler radar data.

With observations from the high-altitude NASA DC-8 aircraft, additional structures such as the warm-core anomaly can be evaluated. Following the same method as Stern and Nolan (2012), we calculated the warm-core anomaly as the difference between the mean potential temperature within the 15-km radius of the storm center and that in an annulus 200–300 km away from the storm center. Figure 8a compares the peak warm-core anomaly as a function of L_h with the observed warm-core anomaly obtained from high-altitude dropsonde data from the same period (1800 UTC 29 August–0000 UTC 30 August). A similar calculation of the warm-core anomaly using dropsonde data was documented by Durden (2013). It is evident from Fig. 8 that the peak warm-core

anomaly, which is also correlated with storm intensity (cf. Fig. 1), decreases with L_h. The warm-core anomaly simulated in L750 and L1500 is found to be
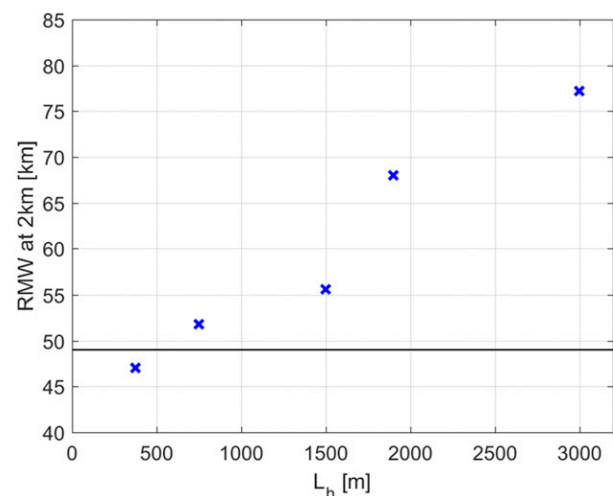

FIG. 4. Plots of the 2-km RMW as a function of L_h averaged between 72 and 78 h (1800 UTC 29 Aug–0000 UTC 30 Aug). The blue crosses are from the HWRF forecasts. The solid line is an estimate of the parameters based on Doppler radar observations.
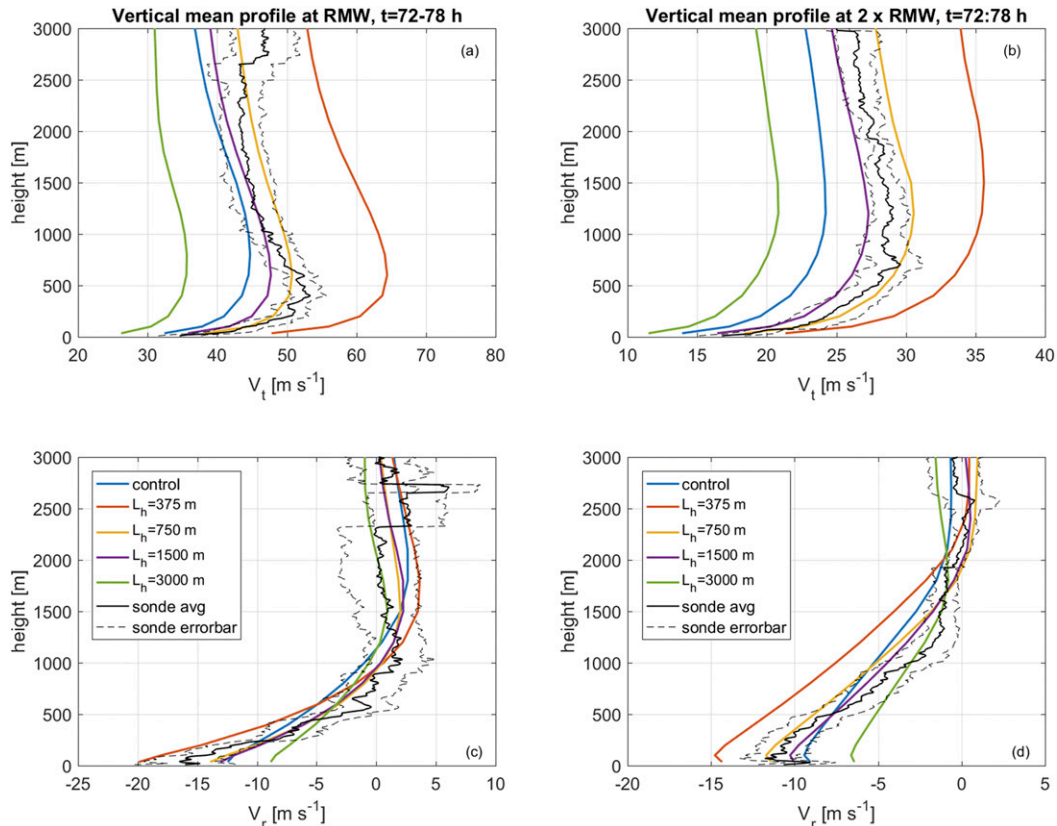
FIG. 5. Plots of the vertical mean profiles of (a),(b) V_t and (c),(d) V_r at the RMW from the Hurricane Earl (2010) forecasts averaged between 72 and 78 h (1800 UTC 29 Aug–0000 UTC 30 Aug). Comparisons at (left) the RMW and (right) $2 \times$ RMW. The black solid and dashed lines, respectively, in each plot show the mean and 95% confidence interval of the wind profiles from dropsonde observations. Note that $L_h = 1900$ m was used in the control run.

closer to the observed value than that in the other experiments.

The warm-core height, defined as the height of the peak warm-core anomaly, is shown in Fig. 8b as a function of L_h. Here, the warm-core height slowly increases with L_h and is only weakly correlated with the storm intensity. This is consistent with Durden (2013) based on a large number of high-altitude dropsonde data in the eye and ambient regions of multiple hurricanes. For the warm-core height forecast, once again L750 and L1500 performed better than the other experiments.

As previous studies have suggested, vortex tilt is an important aspect of sheared storms (Jones 1995; Reasor et al. 2000), and we evaluated how vortex tilt may be linked to L_h. Here, the tilt is computed as the displacement of the circulation center between 1- and 8-km altitude, where the circulation center is defined as the location of minimum horizontal wind speed. Figure 9 shows that the vortex tilt increases with L_h.

The observed tilt during the period of interest (1800 UTC 29 August–0000 UTC 30 August) was obtained from Doppler radar data. Once again, L750 and L1500 captured the tilt magnitude better than the other experiments. Previous theoretical and observational studies (Reasor et al. 2004; Reasor and Eastin 2012) have suggested that a stronger, deeper vortex tends to be more resilient against the environmental wind shear and from being tilted. Our results are consistent with this, as the hurricane vortex is much stronger and deeper in the Earl forecasts with smaller L_h than those in the Earl forecasts with larger L_h, as discussed earlier (cf. Fig. 3). This result thus suggests that as the vortex-scale structure is sensitive to L_h, the vortex response to environmental shear may also be sensitive to L_h. It would be of interest to understand the underlying dynamics for the tilt reduction when L_h is small, which may be related to vortex Rossby waves and deep convection. However, this topic is beyond the scope of the present study.
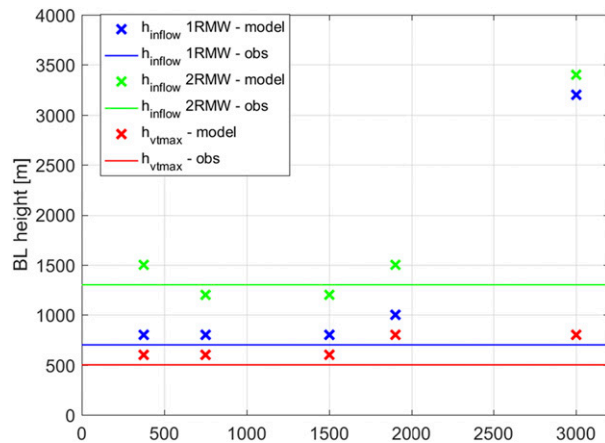
FIG. 6. Plots of the boundary layer height as a function of L_h averaged between 72 and 78 h (1800 UTC 29 Aug–0000 UTC 30 Aug). The blue color represents the inflow layer depth at the RMW, the green color represents the inflow layer depth at $2 \times$ RMW, and the red color represents the height of the maximum tangential wind speed. The solid lines are estimates of the parameters based on dropsonde observations. Note that $L_h = 1900$ m was used in the control run.

4. Verification of retrospective HWRF forecasts of multiple storms

The sensitivity experiments for Hurricane Earl shown above and the idealized simulations of ZM15 suggest that the L_h used in the operational H215 was too large and should be reduced to values between 750 and 1500 m. During the model upgrade in 2016, L_h was lowered from 1900 to 800 m by tuning the coac parameter in HWRF. The value of L_h (800 m) used is close to the median value of the observationally estimated L_h reported by Zhang and Montgomery (2012) and is also consistent with the recommendations of Bryan et al. (2010) based on extensive CM1 simulations and comparisons of the modeled MPI and storm structure to climatology and observational composites.

To further evaluate the impact of the modification of L_h in H216, we ran two sets of retrospective forecasts: one using the operational H216 HWRF in which L_h was set to 800 m (referred to as HOAC hereafter) and the other using H216 but with L_h set to 1900 m (i.e., the value as in H215, referred to as COAC hereafter) for 10 Atlantic tropical cyclones. These included all eight storms in the 2014 season (Arthur, Bertha, Cristobal, Dolly, Edouard, Fay, Gonzalo, and Hanna) and two storms in the 2016 season (Earl and Matthew). Note that the data assimilation system was turned on in these retrospective forecasts in a similar manner as previous studies on model physics upgrades, such as in the European Centre for Medium-Range Weather Forecasts
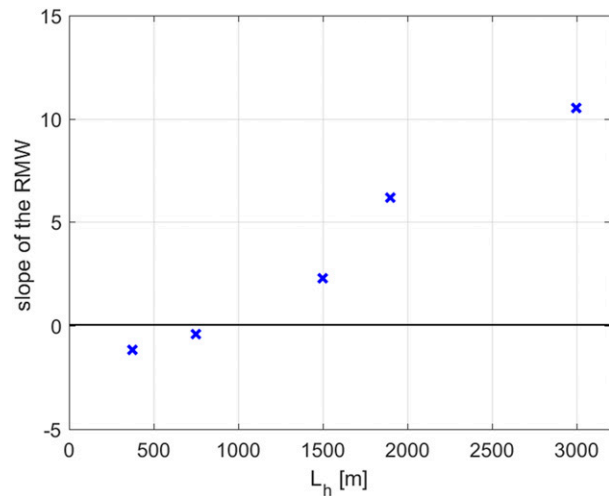

FIG. 7. As in Fig. 4, but for the slope of the RMW as a function of L_h.

model (Rodwell and Palmer 2007; Rodwell and Jung 2008; Klocke and Rodwell 2014), and the WRF Model (Cavallo et al. 2016). Data assimilated into HWRF included observations from radiosondes, aircraft, surface ships and buoys, surface land stations, pilot balloons, wind profilers, scatterometers, aircraft-released dropsondes, satellites recording radiances and winds, and the tail Doppler radar on the NOAA WP-3D (Tallapragada et al. 2014).

To verify the retrospective forecasts, the post-processed best-track data from NHC were used. These data and the HWRF-generated track outputs are shown in Fig. 10a, which shows the mean track forecast errors for HOAC and COAC for the retrospective forecasts. There is a small improvement in the track forecast for HOAC compared to COAC at a lead time of 36–48 h, although this improvement is not statistically significant at the 95% confidence interval. This small improvement in the track forecast appears in both the along-track (Fig. 10b) and cross-track (Fig. 10c) error verifications during the 36–48-h forecast period. During other periods of the forecast, the performance levels of the along-track and cross-track forecasts are reversed. This result basically suggests that lowering L_h in HWRF did not degrade the hurricane track forecast. Of note, we found that the improvement in the track error by reducing L_h in HWRF is much larger for strong storms than weak storms (not shown).

The mean intensity (VMAX) errors for HOAC and COAC for the retrospective forecasts are shown in Fig. 11, where the error bar represents the 95% confidence interval. It is evident from Fig. 11 that substantial improvements in the intensity error were made when we lowered L_h in the HWRF Model. The overall
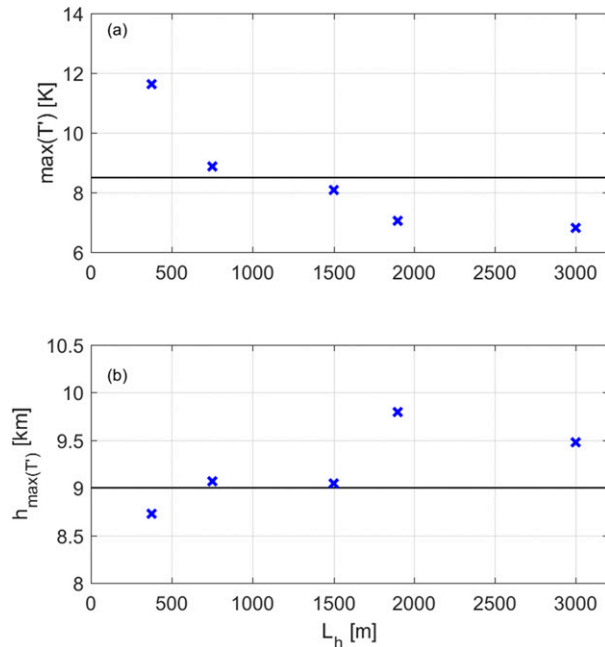
FIG. 8. As in Fig. 4, but for the (a) maximum warm-core anomaly and (b) height of the maximum warm-core anomaly as a function of L_h The solid line in each panel is an estimate of the parameters based on dropsonde observations.

FIG. 9. As in Fig. 4, but for the vortex tilt as a function of L_h.

improvement is 10%–15% for HOAC compared to COAC for almost all of the verification times. Although statistical analysis (*t* test) showed that the improvement is not statistically significant at the 95% confidence level, this result suggests that lowering L_h in HWRF has a positive impact on the intensity forecasts. Note that this improvement is statistically significant at the 80% confidence interval for lead times of 6–30, 40–60, and 84–120 h.

An alternative way to measure the forecast impact is to examine the frequency of superior performance (FSP), which is a measure (in percent) of how frequently one model yields a better forecast than another. The FSP for the intensity forecasts from HOAC and COAC are compared to the 2014 operational version of the HWRF model (H214) in Fig. 12. The improvement in the intensity forecasts of HOAC compared to COAC is consistent throughout the 120-h forecast period except at 72-h lead time. Figure 12 also indicates that both HOAC and COAC performed much better in terms of intensity forecast compared to H214. The 10%–15% improvement in the intensity forecasts of HOAC compared to COAC is also clearly shown in Fig. 12.

In addition to the intensity error improvement due to the modification of L_h, the bias in intensity is largely improved in HOAC compared to COAC (Fig. 13). The low bias in the intensity forecast of ~8 kt (after a 12-h
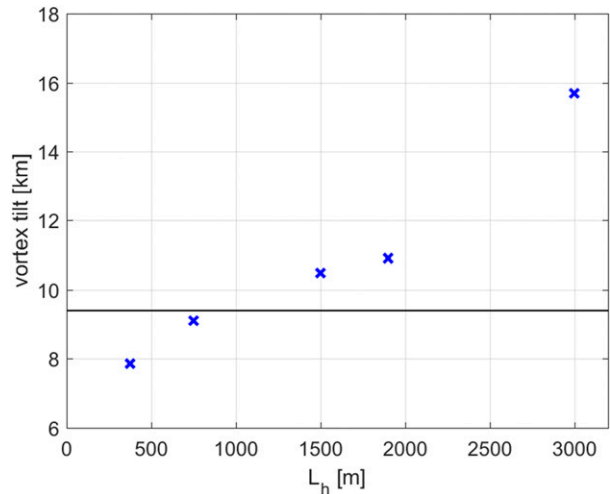
lead time; 1 kt = 0.51 m s^{-1}) was reduced to ~3 kt from COAC to HOAC because of the modification of L_h only. This improvement in the hurricane intensity bias forecast is statistically significant at the 95% confidence interval, which is quite encouraging. Our previous work on improving the vertical eddy diffusivity based on aircraft observations (Zhang et al. 2015) also led to improvement in the intensity bias forecast, but the change was not as large as the improvement caused by improving L_h shown here.

Using NHC data, we also validated the storm size forecasts for COAC versus HOAC. A statistically significant improvement was found in the RMW forecast in HOAC compared to COAC (Fig. 14). The high bias in the RMW forecast of HWRF was reduced by 6 n mi (~10 km) on average (Fig. 14). There is only a small improvement in the forecast of other outer storm size metrics, such as 64- and 34-kt radii, while the improvement is not statistically significant (result not shown). Of note, there is still a relatively large positive bias in the RMW in H216, even after the improvement in L_h (Fig. 14), which is consistent with the continued weak bias. This result suggests that other physics may be required to further reduce the storm size bias in HWRF.

5. Conclusions and discussion

This study documents the process of upgrading the horizontal diffusion parameterization in HWRF. Idealized numerical experiments were first conducted to understand the effects of horizontal diffusion on hurricane intensification and the associated dynamics in HWRF in an earlier study (ZM15). We then conducted a real-case
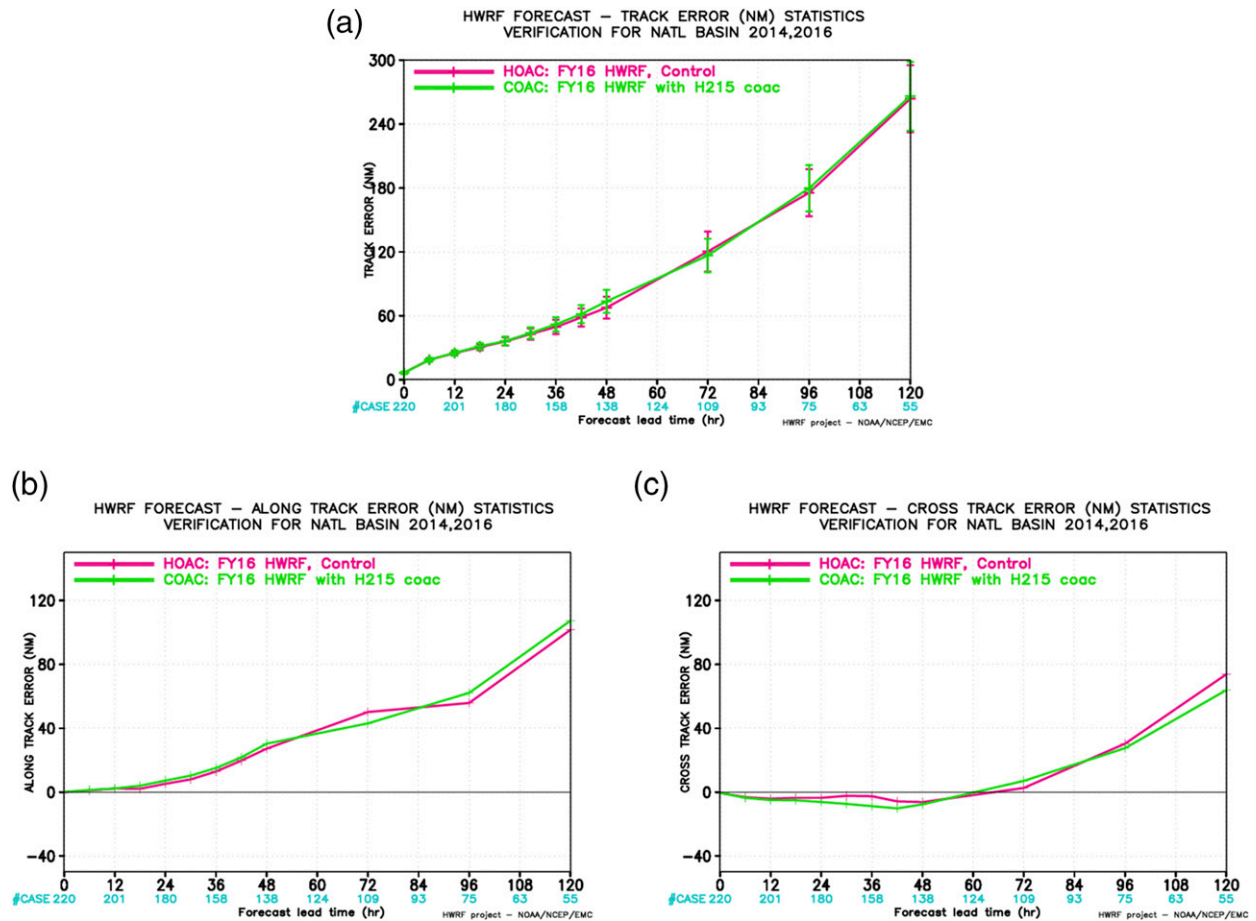
FIG. 10. Plots of the (a) absolute track error, (b) along-track error, and (c) cross-track error in HWRF retrospective forecasts with two different setups for the parameterization of horizontal diffusion in terms of L_h. In the control experiment (referred to as HOAC), L_h = 800 m as in H216, while the other experiment (COAC) used L_h = 1900 m and the same coac value as in H215 but with the H216 model. The error bar represents the 95% confidence interval. FY16 means year 2016 in the legend.

study with HWRF using the same setup as in the idealized experiments to verify the findings from the idealized study. In the sensitivity experiments of real-case HWRF forecasts (Hurricane Earl), we compared the modeled hurricane structure to extensive aircraft observations using structural metrics developed in our previous model evaluation studies (i.e., Zhang et al. 2015). Results from the Earl forecasts confirmed that storm intensity decreases with L_h in agreement with the results reported by BR09 and ZM15. In addition, storm structural characteristics including storm size in terms of RMW, kinematic boundary layer heights, warm-core height and temperature anomaly, and eyewall slope are found to be sensitive to L_h. Comparisons between the modeled and observed structures suggest that the experiments with L_h = 750–1500 m performed better than the control experiment. These values of L_h are much smaller than that used in H215, but they are closer to the

observational estimates of Zhang and Montgomery (2012) and the numerical estimates of Bryan et al. (2010). Based on the results of our sensitivity
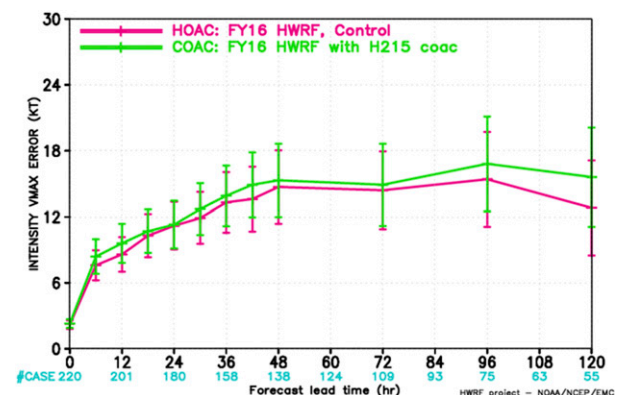

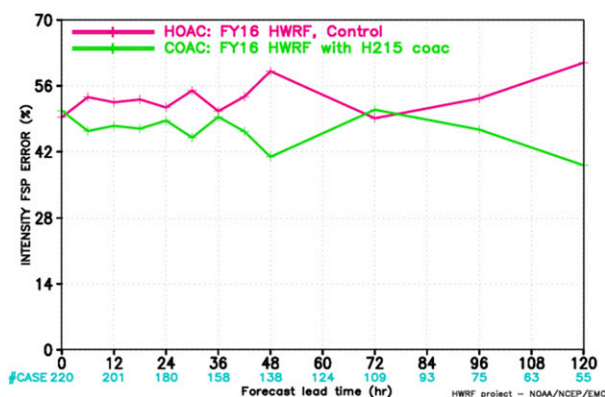FIG. 11. As in Fig. 10, but for the absolute error for intensity.

FIG. 12. As in Fig .10, but for the FSP for intensity error relative to the 2014 version of HWRF.
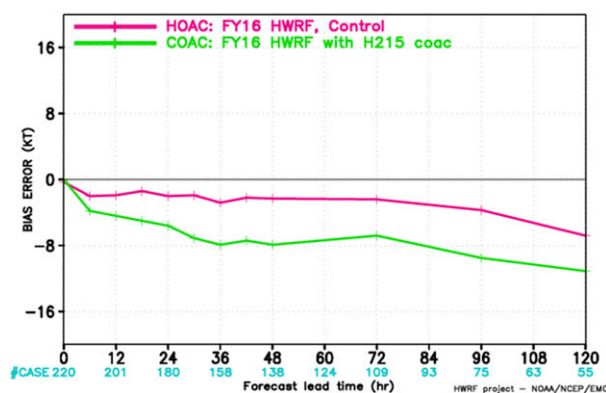

FIG. 13. As in Fig. 10, but for the biases in intensity.

experiments of both idealized and real-case simulations, as well as the previous studies on horizontal diffusion mentioned above, a decision was made to reduce L_h to 800 m in H216.

We then further evaluated the impact of the reduced L_h on track, intensity, and structure forecasts using extensive retrospective forecasts from over 200 cases for 10 Atlantic tropical cyclones. To identify the impact of the improvement in L_h and optimize the model initial conditions, the HWRF data assimilation system was turned on during the retrospective experiments following previous studies of model physics improvement of other forecast models (e.g., Rodwell and Jung 2008; Klocke and Rodwell 2014; Cavallo et al. 2016). Our results show that lowering L_h in HWRF significantly reduced the bias in intensity forecasts. The improvement in the mean intensity error is 10%–15% for almost all lead time for the 5-day forecasts by reducing L_h, and this improvement is statistically significant at 80% for lead times of 6–30, 40–60, and 84–120 h. The track forecast is only slightly improved and the improvement is not statistically significant. Small improvements in both along- and cross-track errors are found at lead times of 24–36-h forecasts. The storm size forecast is found to be significantly improved in H216 using the modified L_h, with the bias in RMW reduced by 6 n mi (~10 km) on average. We also note that the RMW in HWRF is still much larger than the observations based on the best track, indicating further improvement in HWRF's model physics beyond horizontal diffusion is required.

Our results suggest that the impact of lowering L_h in HWRF is independent of the version of the HWRF used here, as sensitivity experiments using both H215 and H216 showed similar positive impacts on intensity and structure forecasts, when L_h was

lowered to values close to previous observational and numerical estimates. Besides the improvement in the vertical eddy diffusivity in HWRF using aircraft data as documented by Zhang et al. (2015), the present study further emphasizes the important role of aircraft observations in model diagnostics and improvement. We demonstrate the utility of the developmental framework as proposed by Zhang et al. (2012) to improve TC model physics. This framework illustrates a process for identifying model deficiencies through a comparison with observations, new model physics development and implementation, and additional validation of the new physics. Here, we show the usefulness of this framework to improve the horizontal diffusion parameterization in the operational HWRF. It should be invaluable for evaluating and advancing other aspects of the model physics beyond horizontal diffusion in the future.

In the end, it is worthwhile to note that the current parameterization of horizontal diffusion in HWRF is far from perfect, given the limited knowledge of L_h and the uncertainty in the tuning method used here. As mentioned earlier, the observation-based estimate of L_h given by Zhang and Montgomery (2012) did not take into account the model resolution associated with subgrid-scale turbulent parameterization in the numerical models. Note also that the value of L_h recommended by Bryan et al. (2010) was based on a different numerical model from HWRF, which utilized different model resolution and physics configurations. This study is only a first attempt to improve the parameterization of subgrid turbulent mixing in NWP models like HWRF, while further proving the importance of horizontal-diffusion parameterization in NWP models for real-case hurricane forecasts. Future work is aimed at developing a new parameterization of L_h that includes
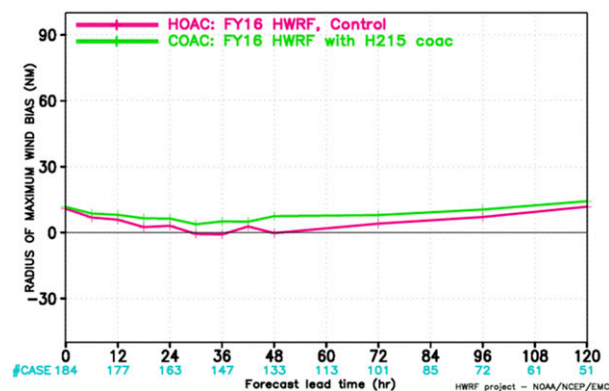
FIG. 14. As in Fig .10, but for the bias in the RMW.

the scale-aware feature. Including a TKE component into the horizontal-diffusion parameterization is also recommended, which will benefit a hurricane model that utilizes a TKE-type PBL scheme. Given the limited number of turbulence observations in hurricanes, a field program especially designed for low-level hurricane measurements, likely with unmanned aircraft, is desirable for further calibration of the turbulent parameterizations in hurricane models.

REFERENCES

Bao, J.-W., S. G. Gopalakrishnan, S. A. Michelson, F. D. Marks, and M. T. Montgomery, 2012: Impact of physics representation in the HWRFX on simulated hurricane structure and pressure–wind relationship. *Mon. Wea. Rev.*, **140**, 3278–3299, https://doi.org/10.1175/MWR-D-11-00332.1.

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., 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.

——, and R. Rotunno, 2009: The maximum intensity of tropical cyclones in axisymmetric numerical model simulations. *Mon. Wea. Rev.*, **137**, 1770–1789, https://doi.org/10.1175/2008MWR2709.1.

——, ——, and Y. Chen, 2010: The effects of turbulence on hurricane intensity. *29th Conf. on Hurricanes and Tropical Meteorology*, Tucson, AZ, Amer. Meteor. Soc., 8C.7, https://ams.confex.com/ams/pdfpapers/167282.pdf.

Cavallo, S. M., J. Berner, and C. Snyder, 2016: Diagnosing model error from time-averaged tendencies in the Weather Research and Forecasting Model. *Mon. Wea. Rev.*, **144**, 759–779, https://doi.org/10.1175/MWR-D-15-0120.1.

Chen, H., and S. G. Gopalakrishnan, 2015: A study on the asymmetric rapid intensification of Hurricane Earl (2010) using the HWRF system. *J. Atmos. Sci.*, **72**, 531–550, https://doi.org/10.1175/JAS-D-14-0097.1.

Durden, S. L., 2013: Observed tropical cyclone eye thermal anomaly profiles extending above 300 hPa. *Mon. Wea. Rev.*, **141**, 4256–4268, https://doi.org/10.1175/MWR-D-13-00021.1.

Emanuel, K. A., 1995: Sensitivity of tropical cyclones to surface exchange coefficients and a revised steady-state model incorporating eye dynamics. *J. Atmos. Sci.*, **52**, 3969–3976, https://doi.org/10.1175/1520-0469(1995)052<3969:SOTCTS>2.0.CO;2.

Foster, R. C., 2009: Boundary-layer similarity under an axisymmetric, gradient wind vortex. *Bound.-Layer Meteor.*, **131**, 321–344, https://doi.org/10.1007/s10546-009-9379-1.

Gopalakrishnan, S. G., F. D. Marks Jr., J. A. Zhang, X. Zhang, J.-W. Bao, and V. Tallapragada, 2013: A study of the impacts of vertical diffusion on the structure and intensity of the tropical cyclones using the high-resolution HWRF system. *J. Atmos. Sci.*, **70**, 524–541, https://doi.org/10.1175/JAS-D-11-0340.1.

Janjić, Z. I., 1990: The step-mountain coordinate: Physical package. *Mon. Wea. Rev.*, **118**, 1429–1443, https://doi.org/10.1175/1520-0493(1990)118<1429:TSMCPP>2.0.CO;2.

——, 2003: A nonhydrostatic model based on a new approach. *Meteor. Atmos. Phys.*, **82**, 271–285, https://doi.org/10.1007/s00703-001-0587-6.

——, R. Gall, and M. E. Pyle, 2010: Scientific documentation for the NMM solver. NCAR Tech. Note NCAR/TN-477+STR, 54 pp., https://dx.doi.org/10.5065/D6MW2F3Z.

Jones, S. C., 1995: The evolution of vortices in vertical shear: Initially barotropic vortices. *Quart. J. Roy. Meteor. Soc.*, **121**, 821–851, https://doi.org/10.1002/qj.49712152406.

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.

Kilroy, G., R. K. Smith, and M. T. Montgomery, 2016: Why do model tropical cyclones grow progressively in size and decay in intensity after reaching maturity? *J. Atmos. Sci.*, **73**, 487–503, https://doi.org/10.1175/JAS-D-15-0157.1.

Klocke, D., and M. J. Rodwell, 2014: A comparison of two numerical weather prediction methods for diagnosing fast-physics errors in climate models. *Quart. J. Roy. Meteor. Soc.*, **140**, 517–524, https://doi.org/10.1002/qj.2172.

Montgomery, M. T., J. A. Zhang, and R. K. Smith, 2014: An analysis of the observed low-level structure of rapidly intensifying and mature Hurricane Earl, 2010. *Quart. J. Roy. Meteor. Soc.*, **140**, 2132–2146, https://doi.org/10.1002/qj.2283.

Ooyama, K., 1969: Numerical simulation of the life cycle of tropical cyclones. *J. Atmos. Sci.*, **26**, 3–40, https://doi.org/10.1175/1520-0469(1969)026<0003:NSOTLC>2.0.CO;2.

Reasor, P. D., and M. D. Eastin, 2012: Rapidly intensifying Hurricane Guillermo (1997). Part II: Resilience in shear. *Mon. Wea. Rev.*, **140**, 425–444, https://doi.org/10.1175/MWR-D-11-00080.1.

——, M. T. Montgomery, F. D. Marks Jr., and J. F. Gamache, 2000: Low-wavenumber structure and evolution of the hurricane inner core observed by airborne dual-Doppler radar. *Mon. Wea. Rev.*, **128**, 1653–1680, https://doi.org/10.1175/1520-0493(2000)128<1653:LWSAEO>2.0.CO;2.

——, ——, and L. D. Grasso, 2004: A new look at the problem of tropical cyclones in vertical shear flow: Vortex resiliency. *J. Atmos. Sci.*, **61**, 3–22, https://doi.org/10.1175/1520-0469(2004)061<0003:ANLATP>2.0.CO;2.

Rodwell, M. J., and T. N. Palmer, 2007: Using numerical weather prediction to assess climate models. *Quart. J. Roy. Meteor. Soc.*, **133**, 129–146, https://doi.org/10.1002/qj.23.

——, and T. Jung, 2008: Understanding the local and global impacts of model physics changes: An aerosol example. *Quart. J. Roy. Meteor. Soc.*, **134**, 1479–1497, https://doi.org/10.1002/qj.298.

Rogers, R. F., P. D. Reasor, and J. A. Zhang, 2015: Multiscale structure and evolution of Hurricane Earl (2010) during rapid intensification. *Mon. Wea. Rev.*, **143**, 536–562, https://doi.org/10.1175/MWR-D-14-00175.1.

Rotunno, R., and G. H. Bryan, 2012: Effects of parameterized diffusion on simulated hurricanes. *J. Atmos. Sci.*, **69**, 2284–2299, https://doi.org/10.1175/JAS-D-11-0204.1.

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.

——, ——, and S. V. Nguyen, 2009: Tropical cyclone spin-up revisited. *Quart. J. Roy. Meteor. Soc.*, **135**, 1321–1335, https://doi.org/10.1002/qj.428.

——, J. A. Zhang, and M. T. Montgomery, 2017: The dynamics of intensification in an HWRF simulation of Hurricane Earl (2010). *Quart. J. Roy. Meteor. Soc.*, **143**, 297–308, 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.

Stevens, B., C. Moeng, and P. P. Sullivan, 1999: Large-eddy simulations of radiatively driven convection: Sensitivities to the representation of small scales. *J. Atmos. Sci.*, **56**, 3963–3984, https://doi.org/10.1175/1520-0469(1999)056<3963:LESORD>2.0.CO;2.

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

Wang, Y., 2007: A multiply nested, movable mesh, fully compressible, nonhydrostatic tropical cyclone model—TCM4: Model description and development of asymmetries without explicit asymmetric forcing. *Meteor. Atmos. Phys.*, **97**, 93–116, https://doi.org/10.1007/s00703-006-0246-z.

Zeng, X., M. A. Bruke, M. Zhou, C. Fairall, N. A. Bond, and D. H. Lenschow, 2004: Marine atmospheric boundary layer height over the eastern Pacific: Data analysis and model evaluation. *J. Climate*, **17**, 4159–4170, https://doi.org/10.1175/JCLI3190.1.

Zhang, J. A., and W. M. Drennan, 2012: An observational study of vertical eddy diffusivity in the hurricane boundary layer. *J. Atmos. Sci.*, **69**, 3223–3236, https://doi.org/10.1175/JAS-D-11-0348.1.

——, and M. T. Montgomery, 2012: Observational estimates of the horizontal eddy diffusivity and mixing length in the low-level region of intense hurricanes. *J. Atmos. Sci.*, **69**, 1306–1316, https://doi.org/10.1175/JAS-D-11-0180.1.

——, and F. D. Marks, 2015: Effects of horizontal diffusion on tropical cyclone intensity change and structure in idealized three-dimensional numerical simulations. *Mon. Wea. Rev.*, **143**, 3981–3995, https://doi.org/10.1175/MWR-D-14-00341.1.

——, ——, M. T. Montgomery, and S. Lorsolo, 2011a: 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.

——, R. F. Rogers, D. S. Nolan, and F. D. Marks, 2011b: On the characteristic height scales of the hurricane boundary layer. *Mon. Wea. Rev.*, **139**, 2523–2535, https://doi.org/10.1175/MWR-D-10-05017.1.

——, S. G. Gopalakrishnan, F. D. Marks, R. F. Rogers, and V. Tallapragada, 2012: A developmental framework for improving hurricane model physical parameterization using aircraft observations. *Trop. Cyclone Res. Rev.*, **1**, 419–429, https://doi.org/10.6057/2012TCRR04.01.

——, 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.

Zhu, P., K. Menelaou, and Z.-D. Zhu, 2014: Impact of subgrid-scale vertical turbulent mixing on eyewall asymmetric structures and mesovortices of hurricanes. *Quart. J. Roy. Meteor. Soc.*, **140**, 416–438, https://doi.org/10.1002/qj.2147.