# A New Horizontal Mixing-Length Formulation for Numerical Simulations of Tropical Cyclones

WEIGUO WANG,[a] BIN LIU,[a] LIN ZHU,[a] ZHAN ZHANG,[b] AVICHAL MEHRA,[b] AND VIJAY TALLAPRAGADA[b]

[a] *I.M. System Group, Environment Modeling Center, NOAA/National Centers for Environmental Prediction, College Park, Maryland*

[b] *Environment Modeling Center, NOAA/National Centers for Environmental Prediction, College Park, Maryland*

(Manuscript received 2 August 2020, in final form 22 January 2021)

ABSTRACT: A new physically based horizontal mixing-length formulation is introduced and evaluated in the Hurricane Weather and Research Forecasting (HWRF) Model. Recent studies have shown that the structure and intensity of tropical cyclones (TCs) simulated by numerical models are sensitive to horizontal mixing length in the parameterization of horizontal diffusion. Currently, many numerical models including the operational HWRF Model formulate the horizontal mixing length as a fixed fraction of grid spacing or a constant value, which is not realistic. To improve the representation of the horizontal diffusion process, the new formulation relates the horizontal mixing length to local wind and its horizontal gradients. The resulting horizontal mixing length and diffusivity are much closer to those derived from field measurements. To understand the impact of different mixing-length formulations, we analyze the evolutions of an idealized TC simulated by the HWRF Model with the new formulation and with the current formulation (i.e., constant values) of horizontal mixing length. In two real-case tests, the HWRF Model with the new formulation produces the intensity and track forecasts of Hurricanes Harvey (2017) and Lane (2018) that are much closer to observations. Retrospective runs of hundreds of forecast cycles of multiple hurricanes show that the mean errors in intensity and track simulated by HWRF with the new formulation can be reduced approximately by 10%.

SIGNIFICANCE STATEMENT: To improve the representation of horizontal diffusion in numerical models, this study proposes a new formulation for horizontal mixing length, which calculates the mixing length as a function of local winds. In contrast, current operational models simply assume that the horizontal mixing length is a constant value or a fixed fraction of grid size, which is not realistic. The new formulation produces the horizontal mixing length and diffusivity much closer to those derived from observations than the formulation used in current models. Analyses of retrospective runs of hundreds of forecast cycles suggest that the errors in intensity and track simulated by HWRF with the new formulation can be reduced by 10%. Future work should focus on understanding how large-scale fields and tropical cyclone structure respond to horizontal diffusion parameterizations as well as their impacts on the forecasts of track and intensity.

KEYWORDS: Diffusion; Numerical weather prediction/forecasting; Parameterization; Subgrid-scale processes; Tropical cyclones

## 1. Introduction

Recent studies have indicated that the intensity and structure of numerically simulated tropical cyclones (TCs) are sensitive to horizontal mixing length $L_h$ in the parameterization of horizontal diffusion (e.g., Bryan 2012; Bryan and Rotunno 2009a,b; Bryan et al. 2010; Rotunno and Bryan 2012; Zhang and Marks 2015; Zhang et al. 2018). For example, Bryan and Rotunno (2009a) used an axisymmetric numerical model to investigate the impact of several uncertain parameters on the maximum azimuthal velocity in the simulated TC. They found that $L_h$ is a very important control factor for the simulated TC intensity. This is different from the conventional scale analysis in which the effect of horizontal diffusion is not thought important (Rotunno and Bryan 2012). To give a physical explanation, Rotunno and Bryan (2012) conducted a budget analysis and showed that horizontal diffusion may be a major contributor to the angular momentum budget near the eyewall area in the TC boundary layer. Through an experiment using the Hurricane Weather Research and Forecasting Model (HWRF) to simulate the evolutions of an idealized TC with different specified values of $L_h$, Zhang and Marks (2015) showed that both structure and intensity are sensitive to the choice of $L_h$ value. In another similar study but for real-case simulations, Zhang et al. (2018) showed that the HWRF Model with a horizontal mixing length of 750 m produced the best result in terms of TC structure and intensity when compared with observations.

It is noticed that all the aforementioned studies simply assumed that $L_h$ is a fixed fraction of grid size or a constant value over a simulation domain, given that the nature of $L_h$ is unclear and it is difficult to determine from limited observations. In the literature, the most reasonable value of $L_h$ is usually estimated by a trial-and-error approach in TC simulations (Bryan and Rotunno 2009a; Zhang et al. 2018). However, the assumption of a constant $L_h$ is not realistic. Bryan and Rotunno (2009b) calculated $L_h$ in TCs based on simulations using a turbulence-resolving numerical model, suggesting that $L_h$ varies strongly with distance from the TC center. In another report, Zhang and Montgomery (2012) analyzed flight-level data collected by aircraft that penetrated the eyewalls of three intense hurricanes between the sea surface and 1 km. Their analyses suggested that there is large variability in $L_h$ over different areas and $L_h$ is far from constant. Given these investigations, to represent the horizontal diffusion process in a more realistic

*Corresponding author*: Weiguo Wang, weiguo.wang@noaa.gov

DOI: 10.1175/WAF-D-20-0134.1

way, we propose a flow-dependent horizontal mixing-length formulation for the parameterization of horizontal diffusion in this study. Section 2 introduces the new horizontal mixing-length formulation, along with the parameterization scheme of horizontal diffusion used in the HWRF Model. Section 3 describes the design of experiments using the HWRF Model to simulate idealized and real TCs. Section 4 discusses the impacts of the new formulation on the simulations of the idealized TC and two real TC cases (Hurricanes Harvey and Lane). In addition to the case studies, this section analyzes the results from retrospective runs using the latest operational HWRF system with the new formulation, where hundreds of forecast cycles of multiple hurricanes in the Northern Atlantic (NATL) and Eastern Pacific (EPAC) basins are simulated. Finally, a summary is given in section 5.

## 2. Horizontal mixing-length formulation

A second-order nonlinear Smagorinsky-type approach is used in the HWRF Model to parameterize the subgrid-scale horizontal diffusion (Janjić 1990). The horizontal turbulent flux $F_h$ is expressed as

$$F_h = -\rho K_h \left( \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right), \tag{1}$$

where $\rho$ is air density; $u$ and $v$ are horizontal velocity components in the $x$ and $y$ directions, respectively; and $K_h$ is horizontal diffusivity defined as

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

where $D_h$ is proportional to horizontal deformation. The term $D_h$ is calculated as

$$D_h = \left[ \left( \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right)^2 + \left( \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right)^2 \right]^{1/2}. \tag{3}$$

The horizontal mixing length $L_h$ is assumed to be comparable to grid size and specified as a fixed fraction of grid spacing:

$$L_h = c\,\Delta, \tag{4}$$

where $c$ is a constant and $\Delta$ is the model horizontal grid spacing. As a result, $L_h$ is approximately constant over a given domain. As mentioned above, such an invariant or flow-independent length scale is not consistent with the findings from observational and high-resolution modeling studies (Bryan and Rotunno 2009b; Zhang and Montgomery 2012). Physically, $L_h$ is dependent on local flow, especially for rotational motions. Intuitively, the mixing length in a rotational flow could vary with the distance from the center of rotation. To this end, we hypothesize that $L_h$ is a function of the shear and stretching of local horizontal wind:

$$L_h = 0.5(L_{h1} + L_{h2}), \tag{5}$$

where $L_{h1}$ and $L_{h2}$ are the length scales for shear and stretching, respectively:

$$L_{h1} = sW \left[ \left( \frac{\partial v}{\partial x} \right)^2 + \left( \frac{\partial u}{\partial y} \right)^2 \right]^{-1/2}, \tag{6}$$

and

$$L_{h2} = sW \left[ \left( \frac{\partial u}{\partial x} \right)^2 + \left( \frac{\partial v}{\partial y} \right)^2 \right]^{-1/2}, \tag{7}$$

where $W$ is wind speed, and $s$ is a scaling or correction factor due to the use of finite differences approximating the derivatives in the HWRF Model. The factor $s$ approaches to unity for higher spatial resolutions. Ideally, $L_h$ of a purely rotational motion with a constant angle velocity is proportional to the distance to the center of the rotation. The factor $s$ is estimated as

$$s = \left( \frac{\Delta_0}{\Delta} \right)^{\alpha}, \tag{8}$$

where $\Delta_0$ is the grid spacing of the order of 100 m with which a simulation can resolve the eddies that contain most of the kinetic energy and carry most of the flux in turbulent flows, and $\alpha$ (or $s$) is an empirical constant that is chosen so that the resulting length scale is comparable to that derived from observations. In our tests at $\Delta \sim 2$ km, $\alpha$ is chosen to be 1.45 with $\Delta_0 = 100$ m. For the subgrid-scale parameterization, $L_h$ is limited by the model grid spacing.

## 3. Experimental setup

The HWRF Model 2017 version is used to test the impact of the new horizontal mixing-length formulation. The HWRF Model is configured to simulate the evolution of an idealized TC on an $f$ plane without ocean coupling and data assimilation (Wang et al. 2018). The configuration includes three domains, with one parent grid and two telescopic and movable two-way nested grids. The parent domain covers approximately 80° × 80° with 18-km grid spacing, and the two nested grids cover 12° × 12° and 7° × 7° with 6- and 2-km horizontal grid spacing, respectively. The model uses 61 levels in the vertical (approximately 18 levels below 1000 m) with a top of 1000 Pa. Time integration took place at the 27-s interval for the parent domain, and 9- and 3-s intervals for nesting domains. Physics schemes include the GFDL surface-layer scheme (Bender et al. 2007) for calculating surface fluxes and the energy budget, Ferrier microphysics scheme (Ferrier et al. 2002) for cloud processes, Rapid Radiative Transfer Model (Iacono et al. 2008) for longwave and shortwave radiation, revised simplified Arakawa–Schubert convective scheme (Han and Pan 2011) for deep and shallow cumulus convections on the outer two domains, and an improved eddy-diffusivity mass-flux boundary layer scheme (Wang et al. 2018) for the vertical diffusion process.

The approach to generate the idealized TC at the beginning of the integration in this study is similar to that used in Liu et al. (2011). The idealized TC is produced by implanting a bogus vortex with a maximum wind speed of 30 m s$^{-1}$ and a radius of 50 km into a typical tropical atmosphere environment. Within

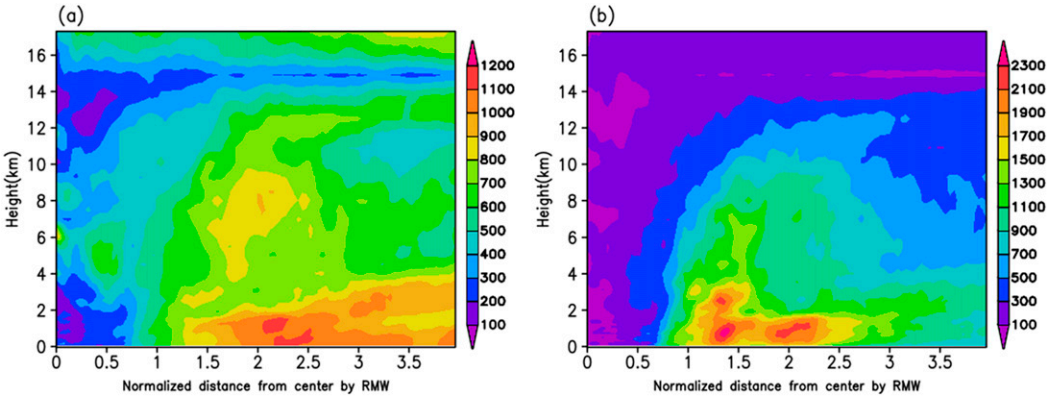


FIG. 1. (a) Azimuthally averaged vertical cross section of horizontal length scale (m) at the 72nd hour in the LVAR simulation. (b) Horizontal diffusivity ($m^2 s^{-1}$). Horizontal distance from the center is normalized by RMW near the surface.

the vortex, geopotential heights are modified according to the gradient wind balance. The ambient horizontal wind profile is based on Chan and Williams (1987), with the maximum tangential wind appearing at the 850-hPa level. The ambient temperature and humidity profiles were derived from the monthly averaged vertical profiles for September at the location 20°N, 145°E from the NCEP–DOE AMIP-II Reanalysis product (Kanamitsu et al. 2002). Environmental geopotential heights are then determined through geostrophic balance. Sea surface temperature is set to 28°C.

Three simulations of the idealized TC are performed using the configured HWRF Model with different formulations of $L_h$ in the innermost domain. The first simulation (L750) uses a constant horizontal mixing length of 750 m, which is close to that used in the current operational HWRF Model. For comparison, a constant value of 2 km is used in the second simulation (L2KM). The third simulation, denoted as LVAR, uses the proposed new flow-dependent $L_h$ formulation. The model is integrated for 5 days in each simulation. The same experimental design is also used in the case studies of Hurricanes Harvey (2017) and Lane (2018).

In the retrospective simulations of hundreds of forecast cycles of multiple hurricanes, the latest version (2021) of the operational HWRF system is used to test the impact of the new formulation, where the horizontal grid spacing values are 1.5, 4.5, and 13.5 km over the three domains, respectively.

## 4. Results

This section discusses the results from the case studies of the idealize TC and Hurricanes Harvey (2017) and Lane (2018), followed by the analysis of the retrospective simulations.

### *a. Idealized case*

As an example, Fig. 1 shows the radius–height distributions of azimuthally averaged $L_h$ and $K_h$ at the 72nd hour in the LVAR simulation. In contrast to the constant $L_h$ values used in L750 and L2KM, $L_h$ in LVAR tends to generally increase as the distance $R$ from the vortex center increases, despite local fluctuations. In the eyewall area, the value of $L_h$ ranges from 500 to 800 m. This is consistent with that derived from field measurements (Zhang and Montgomery 2012). The horizontal mixing length $L_h$ is smaller in the area with stronger horizontal gradients of winds. Large $L_h$ values appear in the area far from the vortex center where the horizontal gradients of winds are weak. The diffusivity, $K_h$, increases with $R$ and reaches its maximum value in the area outside the eyewall [from roughly 1.2 to 2 times the radius of the maximum wind (RMW)]. However, the maximum $K_h$ in the simulations (L750 and L2KM) using constant $L_h$ values appears to be in the eyewall area where wind is strongest (figure not shown); this is not supported by observations (Zhang and Montgomery 2012). In the vertical, all simulations show that $K_h$ decreases with height due to more turbulent flow in the lower levels.

There are very limited data available to verify the parameterized $L_h$ and $K_h$. Zhang and Montgomery (2012) estimated $L_h$ and $K_h$ using the flight-level data collected by research aircraft that penetrated the eyewalls of a few strong hurricanes between the sea surface and 1 km. Following their analysis, we examine the variations of $L_h$ and $K_h$ with wind speed[1] at approximately 500 m above the surface over the vortex area within the radius of twice the RMW in several HWRF simulations using the new formulation. To compare the results from different simulations, the wind speed is normalized by the maximum wind speed at the same level. The results shown in Fig. 2 are derived from a range of simulations under different scenarios. They include the analyses of model outputs over the innermost domain with the 2-km horizontal resolution (i.e., 2-km domain hereafter) of the idealized simulation (LVAR), the 1.5-km domain of one cycle simulation of Hurricane Karina (2020), an EPAC weak storm, using the latest HWRF Model

[1] Here $L_h$ and $K_h$ are displayed against wind speed just for comparisons with those data shown in Zhang and Montgomery (2012). They might be dependent more reasonably on the distance from the vortex center or wind gradients based on our proposed formulation.

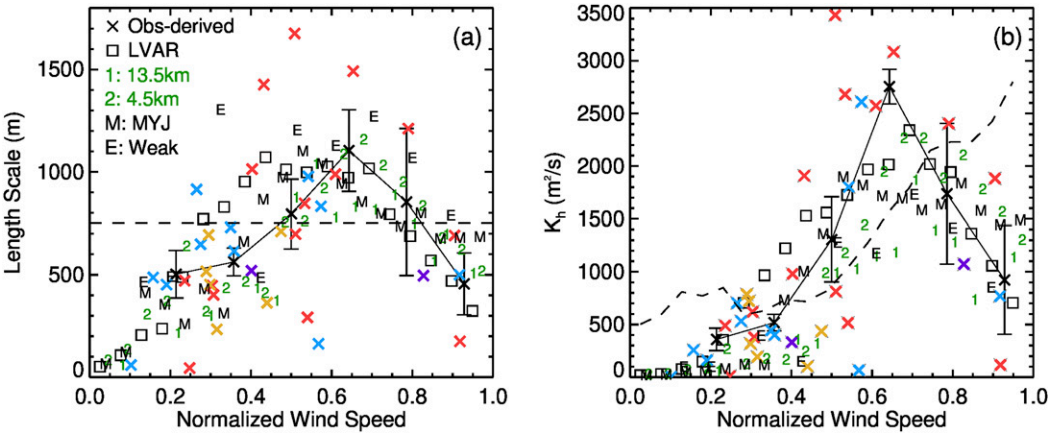


FIG. 2. (a) Horizontal length scale $L_h$ in HWRF simulations, varying with the normalized mean wind speed at approximately 500 m above the surface. Cross symbols in different colors denote data collected from flights during Hurricanes David (1979) (red), Allen (1980) (blue), Hugo (1989) (purple), and Frances (2004) (orange). The data were collected during some periods within ~2 h for each hurricane. Black crosses and vertical bars denote bin-averaged values and standard deviation of observations. See details in Zhang and Montgomery (2012). Squares denote the idealized simulation (LAR) results over the innermost domain with 2-km grid spacing. Green symbols "1" and "2" stand for results from a simulation over domains with 13.5-km and 4.5-km horizontal grid spacing values, respectively; "M" stands for results over the 1.5-km domain from a simulation using MYJ boundary layer scheme; and "E" stands for results over the 1.5-km domain from the simulation for a weak hurricane. (b) Horizontal diffusivity $K_h$.

(symbol E), and the 4.5- and 13.5-km domains of one cycle simulation of Hurricane Laura (2020) in the NATL basin (green symbols). Also included is the output over the 1.5-km domain of one cycle simulation of Hurricane Dorian (2019) using a different boundary layer scheme (MYJ) (Janjić 1990), given the possible impacts of boundary layer schemes (Tang et al. 2018; Wang et al. 2016). For reference, $L_h$ and $K_h$ of L750 are also shown (broken lines). Crosses with different colors denote the results derived from the data collected by aircraft flying into different hurricanes. First, observation-derived results indicate that $L_h$ is not constant over a range of wind speeds and it tends to increase with mean wind speed up to 60% of the maximum wind speed (approximately 40 m s$^{-1}$), then decreases with wind speed (Fig. 2a). Such a variation in $L_h$ is well simulated by the new formulation used in the HWRF experimental runs with different grid resolutions and PBL schemes, while the L750 overestimates $L_h$ in the regions of low and strong wind speeds and underestimates in between. Second, the observation-derived $K_h$ has a peak value in the regions of 60%–70% of the maximum wind speed (40–50 m s$^{-1}$) (Fig. 2b), rather than in the eyewall area containing the strongest winds. This feature is also well captured in all HWRF experimental runs. In a different manner, $K_h$ generally increases with wind speed when a constant $L_h$ is used. These comparisons indicate that the new formulation produces more realistic spatial distributions of $L_h$ and $K_h$, though the magnitudes of $L_h$ and $K_h$ of both LVAR and L750 are comparable to those derived from observations.

Figure 3 shows the time series of the maximum values of 10-m wind speed (intensity) and minimum values of surface pressure for all simulations. After a model spinup period, the intensity of each simulation generally increases with time, and reach a quasi-steady status after the 72nd hour. The intensity of L2KM is 10–15 m s$^{-1}$ weaker than L750 and LVAR. The intensity of L750 is only slightly stronger than that of LVAR. L2KM and L750 generate the highest and lowest minimum values of surface pressure, respectively. The minimum surface pressure of LVAR is higher than that of L750, though the intensities of both simulations are very close. This suggests that the minimum surface pressure of simulated TCs could be more sensitive to the $L_h$ formulation than the maximum wind.

Figure 4 shows a Hovmöller diagram of the azimuthally averaged tangential wind component at approximately 50 m above the surface for each simulation. Solid and broken lines denote the contours of hurricane-force (33 m s$^{-1}$) and destructive (26 m s$^{-1}$) winds, respectively. L750 produces a slower broadening of the vortex (in terms of hurricane-force wind and 20 m s$^{-1}$ contours) and a stronger inner core than L2KM. This is because L750 uses a smaller $L_h$ and produces weaker horizontal diffusion than L2KM (Zhang and Marks 2015). The vortex simulated in LVAR possesses the combined features of L750 and L2KM. The broadening of the vortex simulated in LVAR is close to that in L2KM, but LVAR produces a stronger inner core like L750. Inside the eyewall, $L_h$ and $K_h$ (near the surface) in LVAR are smaller than those in L2KM and L750. Outside the eyewall, $L_h$ in LVAR is close to that in L2KM and $K_h$ in L750 is generally smaller than those in L2KM and LVAR. This might partially explain the differences in vortex expansion in Fig. 4.

Next, we analyze the impact of the use of different $L_h$ formulations on the structure of the simulated TC. Figure 5 shows the radius–height cross sections of azimuthally averaged wind components (tangential wind, radial wind, and vertical velocity) averaged between the 90th and 96th hours of each

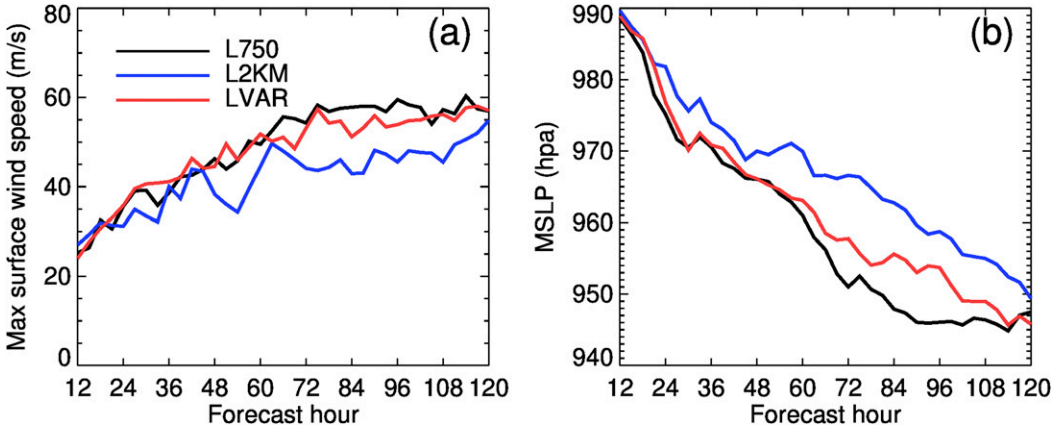


FIG. 3. (a) Time series of the maximum wind speeds at 10 m simulated from HWRF simulations using different $L_h$ formulations (L750, L2KM, and LVAR). (b) The minimum sea level pressure values.

simulation. All three simulations can reproduce typical features of an idealized TC, i.e., a boundary layer jet, radial inflow and outflow, and an updraft area near the eyewall. Differences, however, are evident in the vortex intensity and size. Among the three simulations, the simulated vortex of L2KM expands least in the vertical and most broadly in the horizontal (as shown by the dotted contour line of hurricane-force wind, 33 m s$^{-1}$), with the weakest jet, radial inflow and outflow wind and updraft. The magnitudes of the vortex intensity simulated by LVAR and L750 are close, but LVAR simulates a more compact vortex, possessing a stronger and narrower updraft. In addition, the vortex of LVAR is approximately 1 km lower in the vertical, and 20 km wider (outside of RMW) than that of L750. This can be attributed to the larger $L_h$ and stronger horizontal diffusion outside of RMW at low levels in LVAR, making vortex larger in the horizontal.

It has been reported in the literature that RMW of a TC generally slopes outwards, with its magnitude being greatest in the upper troposphere. Stern and Nolan (2009) analyzed a dataset of three-dimensional Doppler wind analyses comprising seven hurricanes on 17 different days and found that the slope of RMW was a linear function of RMW at 2 km. Based on their results, we diagnosed the vertical profiles of RMW for the three simulations, which are shown by black dashed lines in Fig. 5. The RMWs of the simulated vortexes at different

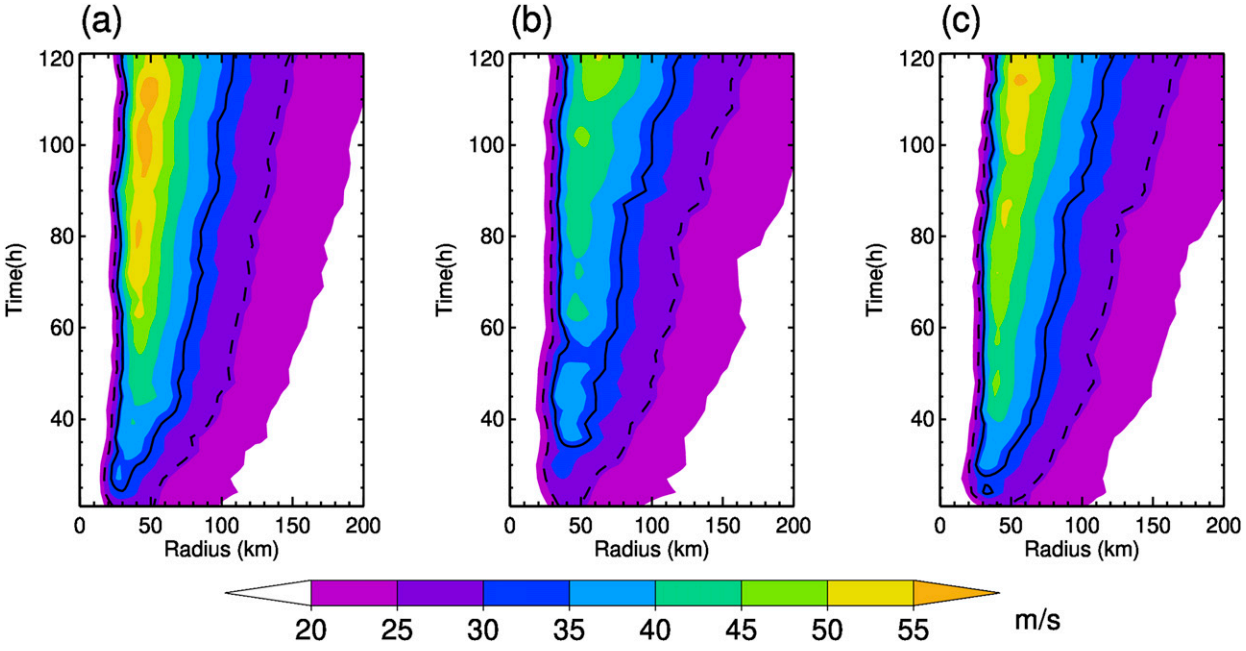


FIG. 4. Hovmöller diagram of azimuthally averaged tangential wind speed at approximately 50 m above the surface from the (a) L750, (b) L2KM, and (c) LVAR simulations. Solid line denotes the contour of hurricane-force wind (33 m s$^{-1}$) and broken line denotes the contour of destructive wind (26 m s$^{-1}$).

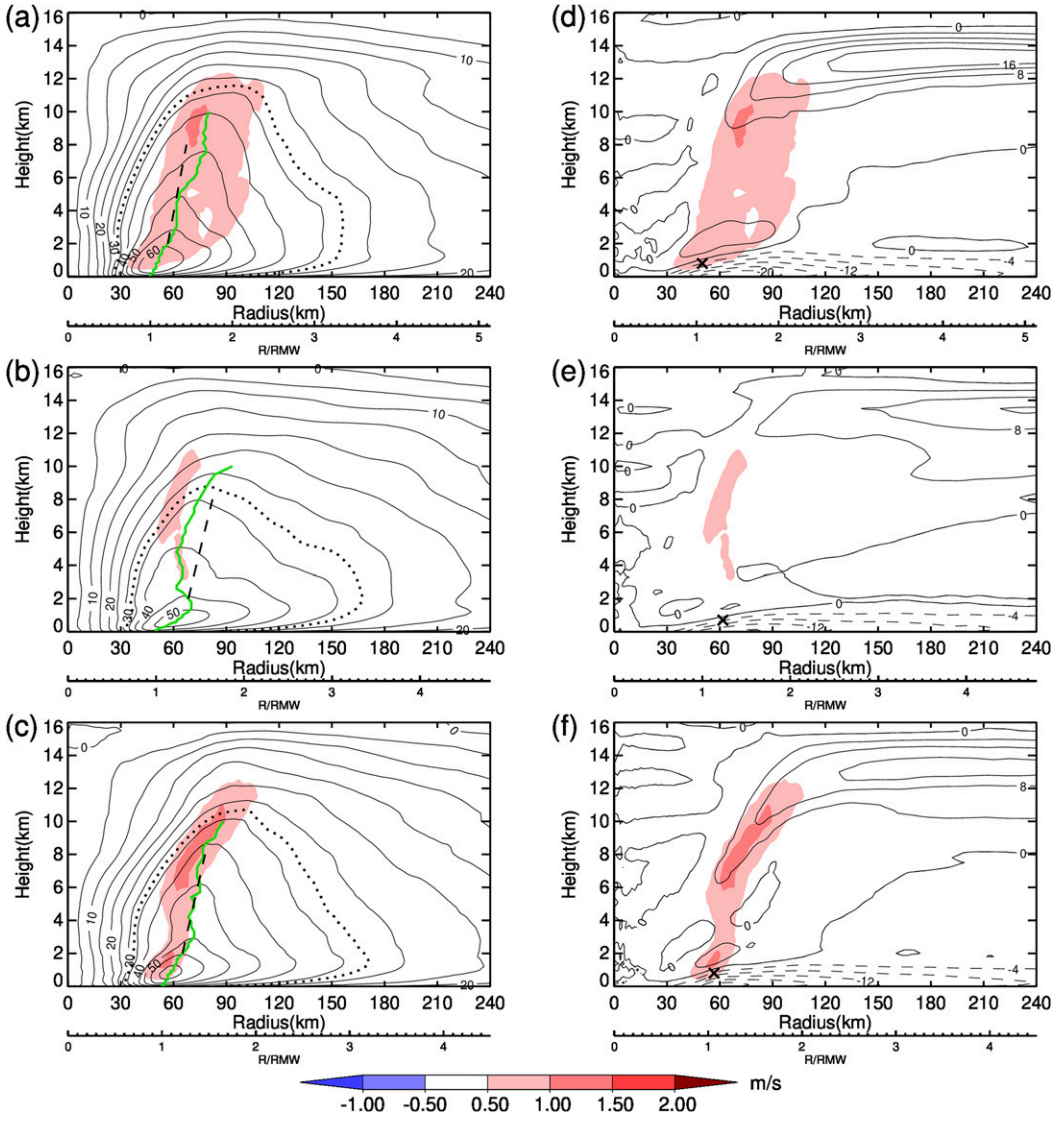


FIG. 5. (a) Radius–height cross section of distributions of azimuthally averaged tangential wind speed (contours), vertical velocity (shaded) of L750 averaged between 90 and 96 h. RMWs at different levels are shown by green lines. The dashed line denotes RMWs derived from the fitted function based on observations by Stern and Nolan (2009). The dotted line denotes the contour of hurricane-force wind (33 m $s^{-1}$). (b) L2KM and (c) LVAR. The second $x$ axis is scaled by RMW near the surface. (d)–(f) As in (a)–(c), but for radial wind (solid line for outflow and dashed line for inflow) and vertical velocity (shaded). The black X symbol denotes the location of the maximum tangential wind.

heights are shown by green lines. The tilt of the vortex eyewall of LVAR is closest to that of the observation-based diagnosis. The tilt of the eyewall of L750 is close to the diagnosis below 4 km, and greater above. Nevertheless, RMWs of the vortexes of both L750 and LVAR generally expand with height. That is, however, not the case for L2KM. The L2KM eyewall tilts outwards near the surface much more than above 2 km, with RMW sloping inwards between 1 and 4 km. This is consistent with the less structured vortex in L2KM.

Figures 5d–f suggest that the inflow layer depth (i.e., the contour of zero radial wind) is above or near the height of maximum tangential wind (denoted by black crosses) for all simulations. This is consistent with observational studies as well as other modeling studies (Zhang and Marks 2015). The heights of boundary layer jets in the LVAR and L750 simulations are located about 800 m above the surface, which is 100 m higher than in L2KM (Table 1). The maximum radial inflow wind near the surface is located approximately 100 m above the surface, with the L2KM simulation producing the weakest wind (Table 2). The inflow layer becomes shallower as a larger $L_h$ is used (Fig. 6). Outflow immediately above the inflow layer of both LVAR and L750 is more pronounced than

TABLE 1. Magnitudes and locations of the maximum azimuthally averaged tangential wind in three simulations.

| Simulation | Max tangential wind ($m\,s^{-1}$) | $R$ (km) | $Z$ (m) |
|---|---|---|---|
| L750 | 65 | 50 | 800 |
| L2KM | 52 | 62 | 700 |
| LVAR | 62 | 57 | 800 |

TABLE 2. Magnitudes and locations of the maximum azimuthally averaged radial inflow in three simulations.

| Simulation | Max radial inflow ($m\,s^{-1}$) | $R$ (km) | $Z$ (m) |
|---|---|---|---|
| L750 | 23 | 70 | 100 |
| L2KM | 16 | 80 | 100 |
| LVAR | 22 | 65 | 100 |

that of the L2KM simulation, while L750 has the strongest outflow jet in the upper level.

### b. *Real cases*

We select Hurricanes Harvey (2017) and Lane (2018) to test how the new formulation can affect the track and intensity forecasts by the HWRF Model.

Harvey was one of the costliest tropical cyclones on record. It originated from a tropical wave on 12 August 2017 and became a named storm on 17 August 2017. Harvey made landfall on 25 August at San Jose Island and Holiday Beach, Texas, at category-3 intensity, bringing heavy amounts of rain and causing catastrophic flooding. Operational models such as the NCEP HWRF, Hurricanes in a Multiscale Ocean-coupled Nonhydrostatic (HMON), and Global Forecast System (GFS) performed reasonably well in forecasting the location, intensity, and timing of the landfall of such a catastrophic storm 2–3 days before its landfall. They, however, have struggled to make accurate forecasts more than 3 days before landfall. Forecasting the location and timing of landfall as accurately and early as possible is critical for agencies to prepare to save lives and property.

As in the idealized TC experiment analyzed above, we run the operational HWRF system (2017 version) using the three $L_h$ formulations for comparisons. The HWRF system is initialized every 6 h starting from 1200 UTC 20 August 2017 when the weakened Harvey began to redevelop and intensify before making landfall approximately 5 days later. Boundary conditions are derived from NCEP GFS forecast data. Initial conditions are derived from GFS analysis data, enhanced by a data assimilation system and vortex initialization process (Biswas et al. 2018). Each simulation produces 5-day forecasts. Our analyses will focus on a period on and before Harvey's landfall.

To assess the performance of the simulations, the intensity and track (on and before landfall) of all simulations are compared with the National Hurricane Center's (NHC) best track data. Results show that the simulation using the new flow-dependent $L_h$ (LVAR) substantially improves the track forecast, compared with simulations using the constant $L_h$ values (L750 and L2KM) for the lead times beyond 48 h (Fig. 7a). L750 also generates better track forecast than L2KM, but not as good as LVAR. For example, at day 5, the track error of LVAR [53.0 n mi (1 n mi = 1.852 km)] is reduced by approximately 58%, compared with L750 (127.6 n mi). The L2KM simulation produces the largest track error (208.6 n mi), which is 64% larger than that of L750. Figures 7b and 7c show the mean absolute track errors in the cross-track and along-track directions at different forecast hours, suggesting that the simulation of LVAR significantly reduces both along-track and cross-track errors, with the reduction of cross-track error contributing much more to the track improvement. Reducing $L_h$ from 2 km to 750 m improves the cross-track forecast, but degrades the along-track forecast slightly.

To further highlight the impact of different $L_h$ formulations on track forecasts, Fig. 8 shows the spaghetti plots for the storm tracks of the three simulations. Because the TC tracks of all simulations initialized within 3 days before landfall are very close, we only show the tracks of the runs initialized between 4 and 5 days before landfall in order to make clearer comparisons. In general, the tracks of L750 have a significant southward bias compared with the best track data, and the track forecast of L2KM is biased further to the south. Such a southward bias is significantly reduced in the simulation using the new flow-dependent length formulation (Fig. 8c). This suggests that the use of the new $L_h$ formulation has a positive impact on storm movement, though we still need to further investigate what underlying physical mechanisms may cause such an improvement.

The large-scale environmental wind is one of the important factors controlling TC tracks. Figure 9 illustrates the streamlines of the environmental wind fields at 850- and 500-hPa levels as well as a layer-mean wind field at the 42nd hour for each simulation initialized at 1200 UTC 20 August 2017. For reference, the environmental flows from the NCEP GFS analysis data at the same levels and time are also shown. The environmental wind field in the vicinity of the simulated Hurricane Harvey is obtained by removing local winds associated with the TC vortex using the method outlined in Galarneau and Davis (2013). The layer-mean wind field is calculated by averaging environmental wind vectors from 850- to 300-hPa levels in an

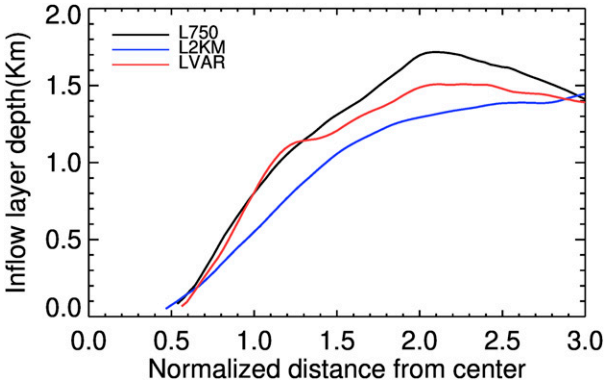


FIG. 6. The depths of the inflow layer as a function of distance from the vortex center normalized by RMW near the surface in three simulations.

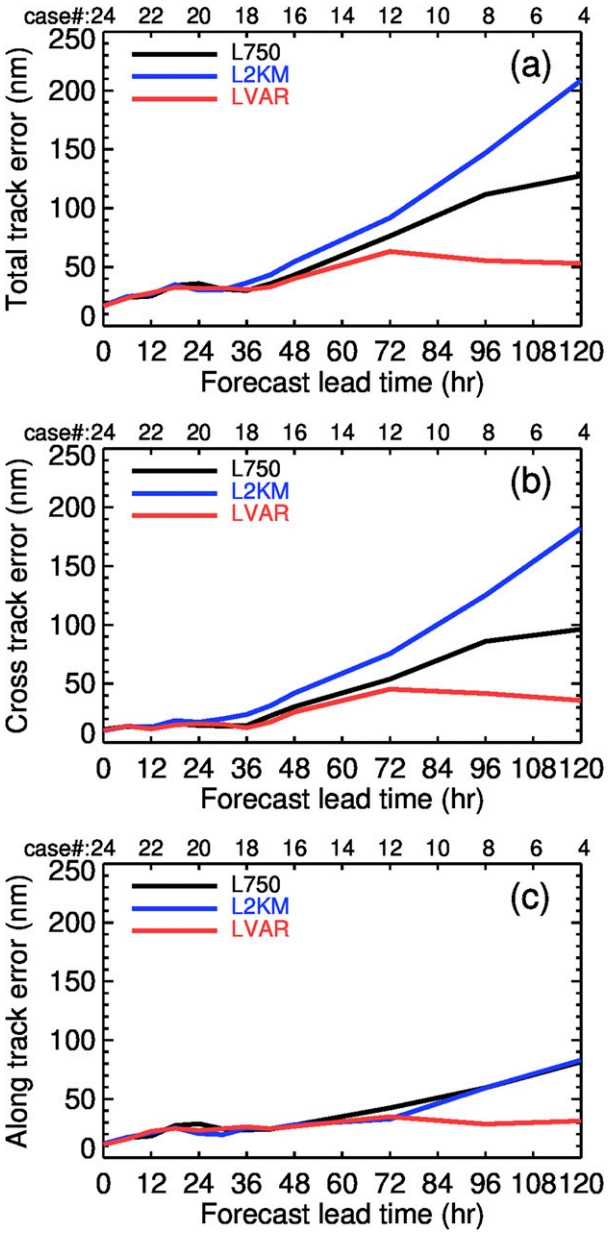


FIG. 7. Mean absolute error for track forecasts as a function of forecast lead time from three simulations: (a) total track, (b) cross-track component, and (c) along-track component. The upper *x* axis shows the number of verifiable cycles.

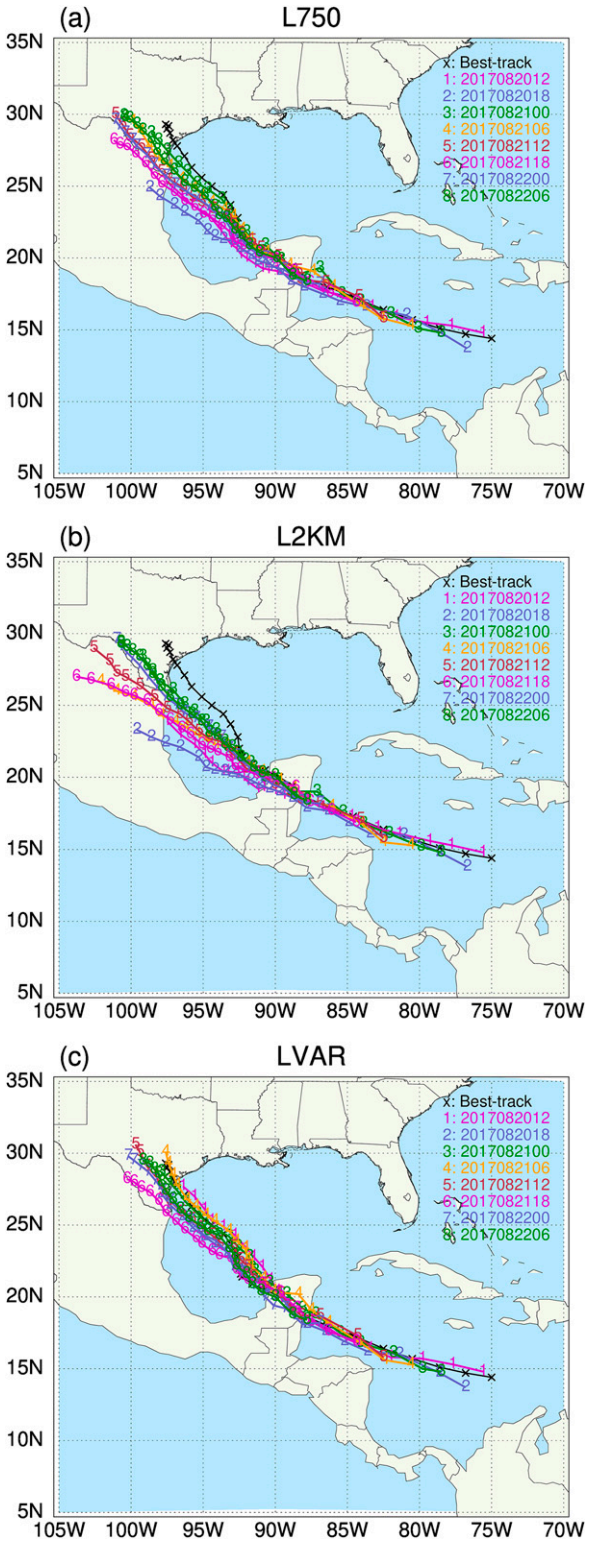


FIG. 8. (a) TC tracks of L750 simulations initialized every 6 h starting from 1200 UTC 20 Aug 2017 to 0600 UTC 22 Aug 2017 (i.e., 4 and 5 days before landfall). (b) L2KM. (c) LVAR.

interval of 50 hPa. The tracks of the three simulations are very close before the 42nd hour and become more visibly different after that. At this hour, the environmental wind fields of all simulations are consistent with those from the analysis data except for small differences in streamline curvatures. The TC movement of each simulation generally follows the streamline of the layer-mean flow. It is found that the simulated TC in LVAR follows closely the streamlines at both 850- and 500-hPa levels; consistent with the analysis (Figs. 9j,k). However, the simulated TC movements of L750 and L2KM follow the

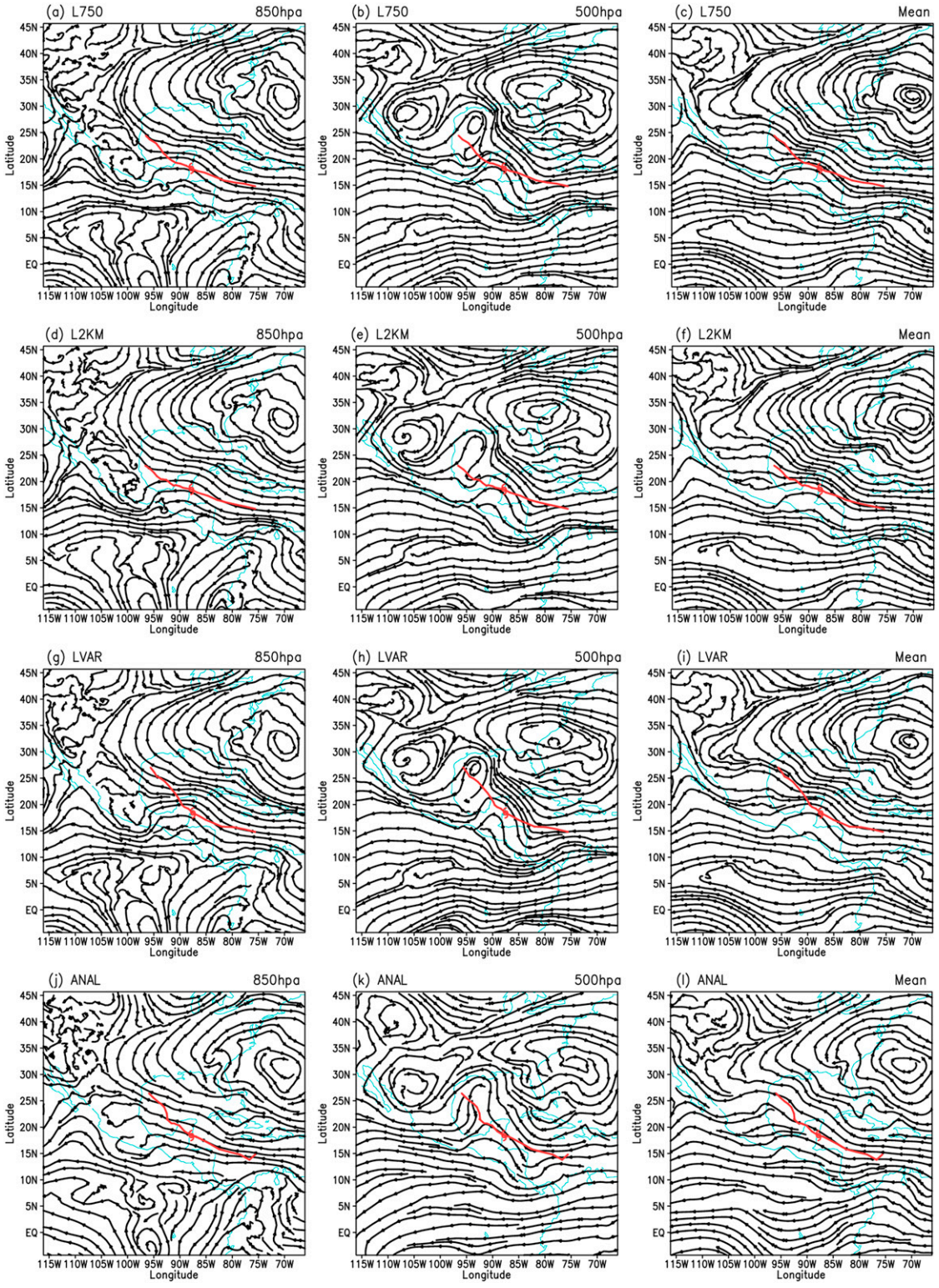


FIG. 9. Environmental wind streamlines at the 42nd hour of the L750 simulation at the (a) 850- and (b) 500-hPa levels and (c) the layer average between the 850- and 300-hPa levels. The red hurricane symbol denotes the vortex location at 42 h, and the red line is the simulated 5-day track. (d)–(f) L2KM, (g)–(i) LVAR, and (j)–(l) GFS analysis.

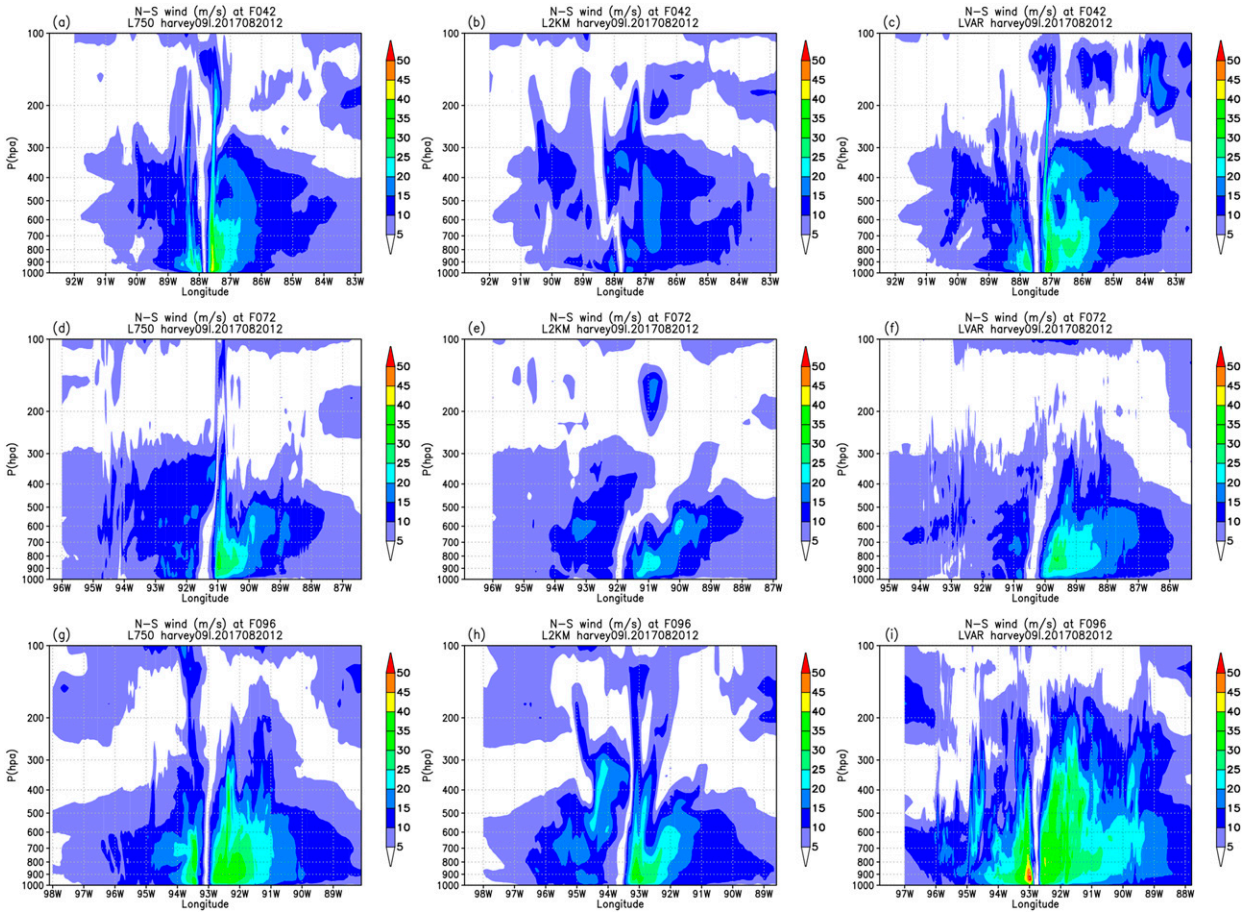


FIG. 10. East–west vertical cross section of tangential wind speed through the vortex center at the (a) 42nd, (d) 72nd, and (g) 96th hours of L750. (b),(e),(h) L2KM and (c),(f),(i) LVAR.

streamline at the 850-hPa level much more closely than that at the 500-hPa level. Such a difference might be related to TC structures such as depth and size of the three simulations and could partially explain why the tracks of the three simulations become more spread afterward due to different environmental wind directions at high and low levels. Figure 10 shows the evolution of the vertical distributions of tangential winds in the east–west cross section through the TC center at the 42nd, 72nd, and 90th hours of each simulation. In general, the vortex of LVAR extends to higher level in the vertical and wider in the horizontal (see the shaded area of wind speed greater than 20 m $s^{-1}$) than the other two simulations. L2KM produces the weakest and shallowest vortex due to the use of the largest horizontal mixing length. In the LVAR simulation, $L_h$ is small near the center and increase with the distance to the center; this may contribute to the simulated vortex stronger, deeper, and wider than the other two simulations. The vortex with a deeper structure could be steered more by the environmental wind at higher levels, as illustrated above.

Differences in the simulated track can have a significant impact on surface precipitation simulations. Figure 11 shows the accumulated surface precipitation distribution during each of the 5-day runs initialized at 1200 UTC 20 August 2017, along with the simulated and observed TC tracks. It is seen that the locations of surface precipitation are very different, in addition to the amount of precipitation. To verify the results, the simulated precipitations near the coastal area are compared with NCEP/EMC surface precipitation analysis Stage IV data that covers continental United States and coastal area (Du 2011). Figure 11d shows the equitable threat score (ETS) values varying with the precipitation amount greater than different thresholds for the three simulations. Broken and solid lines denote ETS values within 2° and 6° of the observed track, respectively. ETS measures the fraction of observed and forecasted precipitations that were correctly predicted, adjusted for hits associated with random chance. L2KM and L750 show no skill at all rainfall thresholds within 2° of the observed track, while LVAR shows skills for rainfall thresholds between 0.2 and 3 in. Within a larger area (6°), ETS values for all simulations increase. LVAR shows large skills for thresholds between 0.5 and 3 in., but L2KM and L750 only show very low skills at rainfall thresholds below 1 in. The very low skills of LVAR and L750 are mainly due to the large track errors, compared with LVAR.

Figure 12a suggests that LVAR produces the best forecast of intensity. Compared with L2KM, L750 also improves the

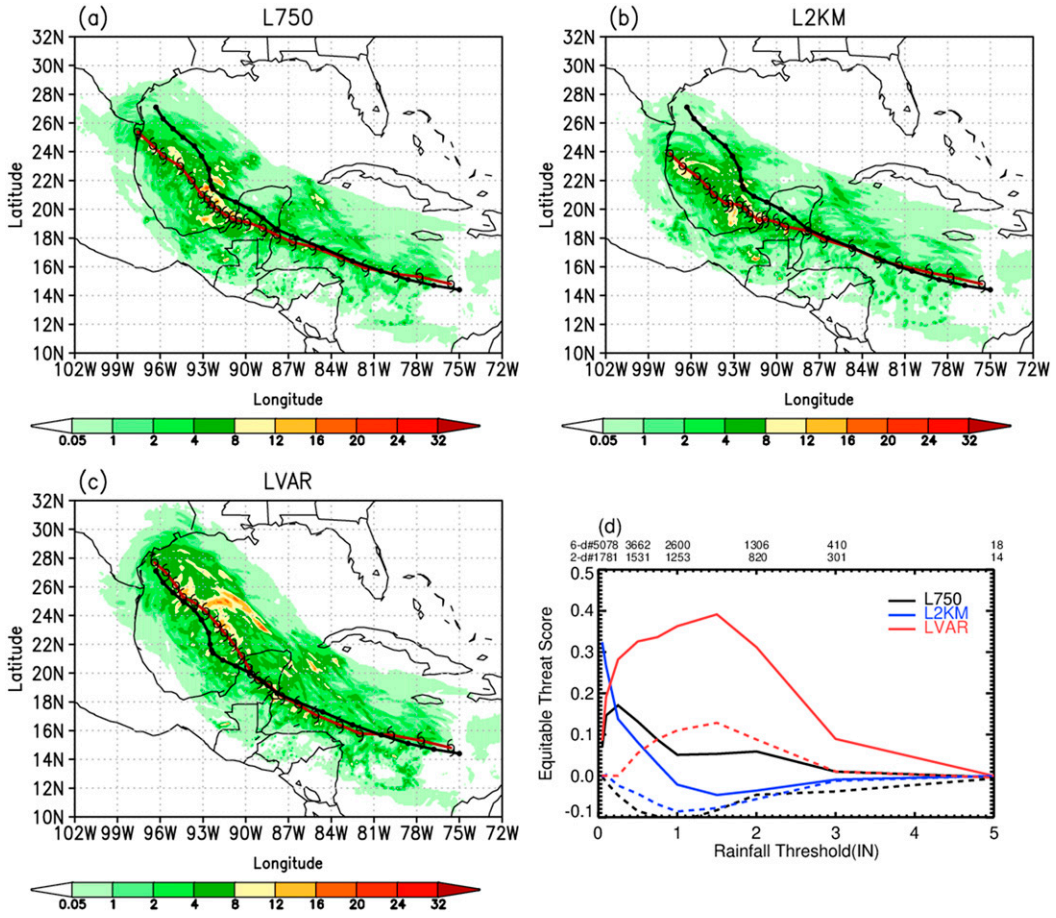


FIG. 11. (a) The 5-day accumulated surface precipitation for the L750 simulation initialized at 1200 UTC 20 Aug 2017. Hurricane symbols denote the simulated track, and filled black circles denote the observed hurricane track. (b) L2KM and (c) LVAR. (d) Equitable threat scores for 5-day accumulated surface precipitations simulated by L750 (black), L2KM (blue), and LVAR (red) varying with precipitation amount greater than different thresholds. Solid and broken lines denote for the scores calculated within 6° and 2° of the observed track, respectively. Both model output and observational data are interpolated to a common domain with 0.1° resolution. The upper $x$ axis in (d) shows the numbers of grid points with available observation and model data for LVAR within 2° and 6° of the observed track.

intensity forecast, but less significantly than LVAR. For 0–48-h lead times, the intensity errors of the three simulations are statistically close. LVAR yields the smallest intensity errors for the lead times beyond 48 h, which is better than both runs using constant $L_h$ values (L750 and L2KM). At 120 h, the intensity error of LVAR [27 kt (1 kt $\approx$ 0.51 m s$^{-1}$)] is 10% and 48% smaller than L750 (30 kt) and L2KM (52 kt), respectively. Such a large improvement is likely due to the improvement in track. Figure 12b shows the averaged values of intensity bias at different forecast lead times. In general, LVAR reduces the negative bias beyond 48 h when compared with L2KM and L750. At 120 h, the biases are −52, −28, and −18 kt for L2KM, L750, and LVAR, respectively. The use of the flow-dependent $L_h$ improves the bias by 36%, compared with that using the $L_h$ value of 750 m in the current operational HWRF system. Figures 12c and 12d show the time series of the averaged values of absolute error and bias of minimum center pressure. Consistent with the averaged intensity error, LVAR produces the best result, followed by L750.

Unlike Hurricane Harvey (2017), Hurricane Lane (2018) travels over the eastern Pacific Ocean without too much direct interaction with land. Similar to the experimental simulations for Hurricane Harvey, the operational HWRF Model is run with the same configurations (i.e., L750, L2KM, and LVAR) to simulate the forecast cycles of Hurricane Lane initialized every 6 h starting from 0000 UTC 15 August 2018. Figure 13 compares the mean absolute errors of the track and intensity of the three simulations. Track errors among the three simulations are close, with LVAR slightly better before 96 h (Fig. 13a). There are large differences in intensity errors, with LVAR producing the best result. Compared with that of L750, the intensity error of LVAR is reduced at nearly all lead times, with the reduction of 25% at day 5. L2KM significantly underpredicts the intensity and has the largest errors due to too

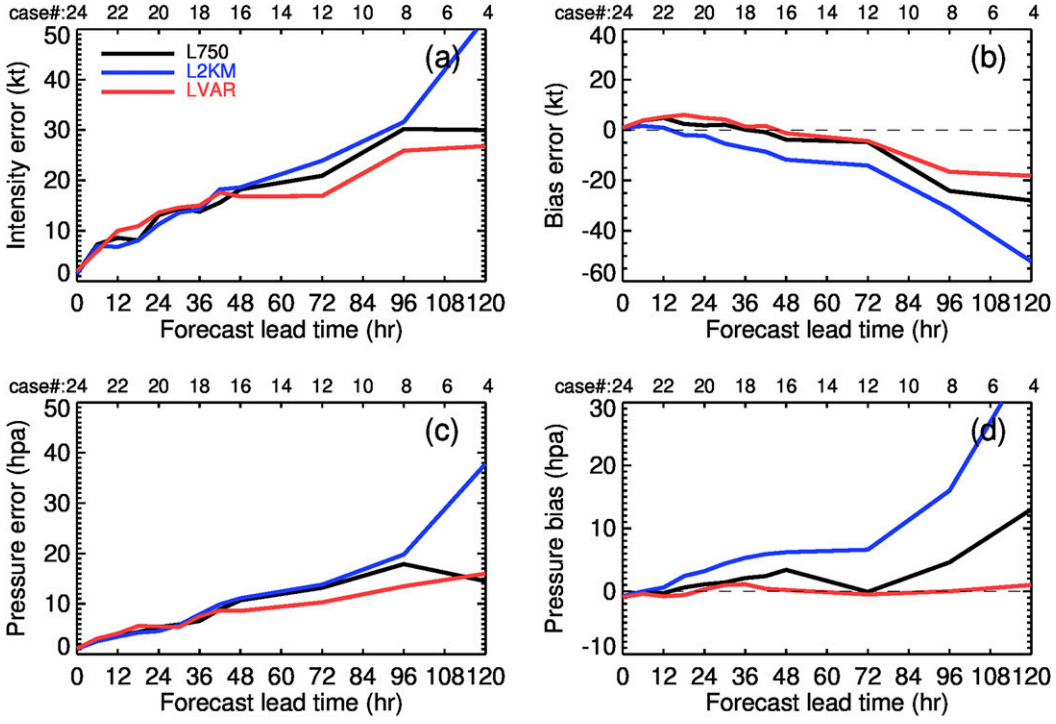


FIG. 12. (a) Mean intensity errors (i.e., mean absolute errors) from three simulations using different $L_h$ formulations. (b) Mean intensity biases (i.e., mean differences). (c) Mean minimum center pressure errors. (d) Mean center pressure bias. The upper $x$ axis shows the number of verifiable cycles.

strong horizontal diffusion caused by the unrealistic horizontal mixing length. To assess the sensitivity of the results to the scaling factor $s$ or $\alpha$ in Eq. (8), we redo the LVAR experiments by setting $\alpha$ values to 1.25 (experiment E125) and 1.35 (experiment E135), respectively. Results are also shown in Fig. 13 with purple and green lines. In general, E125 yields the largest errors in track and intensity among E125, E135, and LVAR, followed by E135. Relative differences in intensity and track errors can reach 10%–20% for a 7% change in parameter $\alpha$ or a 25% change in factor $s$, indicating that the simulated hurricane track and intensity is sensitive to the horizontal mixing length as have been reported in the literature. The sensitive test also suggests that it is important to carefully adjust those parameters to have a more realistic simulation.

### c. *Retrospective runs*

Retrospective simulations of the forecast cycles of dozens of hurricanes using the latest operational HWRF system are performed to further test the impact of the new formulation. Two sets of retrospective simulations are designed. One uses the default operational configuration as a control experiment (CNTL), in which the horizontal mixing length is a constant fraction of horizontal grid spacing over each domain. The other is the same as CNTL except that the new $L_h$ formulation is used

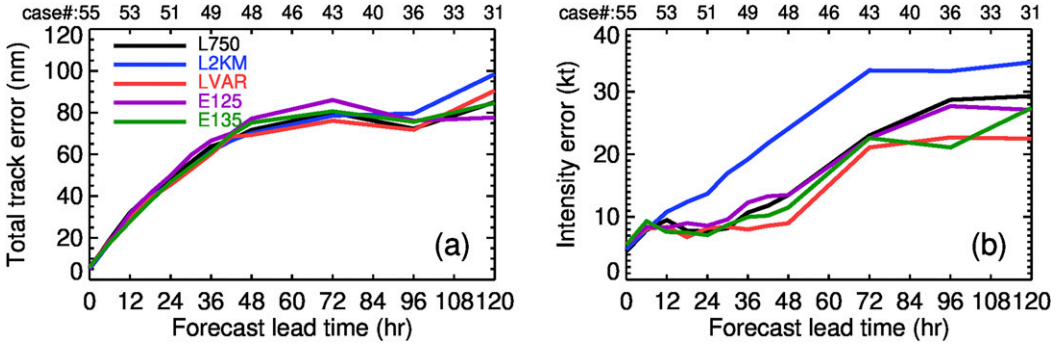


FIG. 13. (a) Mean absolute errors of tracks simulated by L750, L2KM, and LVAR for Hurricane Lane (2018) over the EPAC basin. E125 and E135 are results from the simulations same as LVAR but with different $\alpha$ values in Eq. (8). (b) As in (a), but for intensity.

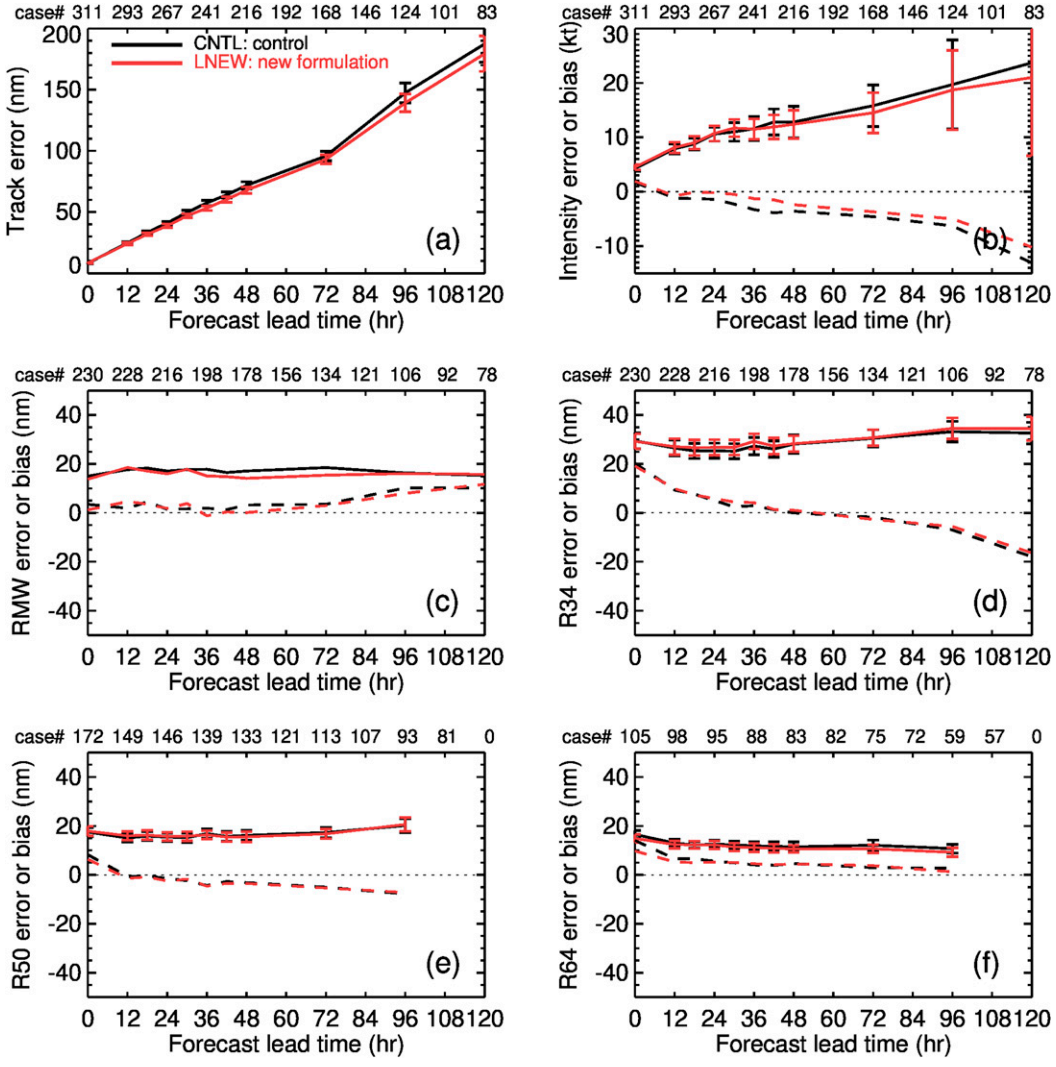


FIG. 14. (a) Mean absolute errors of tracks from retrospective runs for hurricanes in the NATL basin. The upper *x* axis shows the number of verifiable cycles in the calculation. (b) As in (a), but for intensity; broken lines denote mean biases. (c) RMW. Radius of (d) 34-, (e) 50-, and (f) 64-kt winds.

over three domains, with the empirical $\alpha$ values in Eq. (8) chosen to be 1.45, 1.16, and 0.94 (corresponding to the scaling factor $s$ values of 0.0197, 0.0120, and 0.0098) for the parent and two nested domains, respectively. This set of simulations is referred to as LNEW hereafter. A total of 24 hurricanes are selected in the retrospective tests, including a range of different TC characteristics in terms of strength, lifespan, and size in the NATL (Dorian 05L, Humberto 09L, Jerry 10L, Lorenzo 13L, Pablo 18L, Rebekah 19L in 2019 and Hanna 08L, Isaias 09L, Josephine 11L, Marco 14L, Omar 15L, Sally 19L in 2020) and EPAC (Erick 06E, Juliette 11E, Akoni 12E, Mario 14E, Lorena15E in 2019 and Boris 03E, Douglas 08E, Elida 09E, Ten10E, Fausto 11E, Hernan13E, Iselle 14E in 2020) basins.

To assess the impact of the new formulation on operational products, the simulated intensity, track, and vortex size are compared with the NHC's best track data using NHC's verification package. Figure 14 presents the mean (absolute) errors or biases of the intensity, track, and size produced by CNTL and LNEW simulations for the NATL hurricanes varying with the forecast lead time. The verification includes 311 verifiable cycles for track and intensity. Both track and intensity forecasts are improved by using the new formulation at nearly all forecast lead times. The absolute errors of track and intensity of LNEW are approximately 5%–10% smaller than those of CNTL. The mean intensity bias of LNEW is also reduced, though both CNTL and LNEW underpredict the intensity. In terms of the size of the simulated vortexes, LNEW produces a better RMW simulation than CNTL, with the mean error and bias being reduced approximately by 10% for the lead times between 36 and 84 h, and close for the other lead times. In general, the mean errors and biases of the radii of 34-kt (R34), 50-kt (R50), and 64-kt (R64) winds of the two simulations are very close. LNEW slightly improves the R64 forecast but

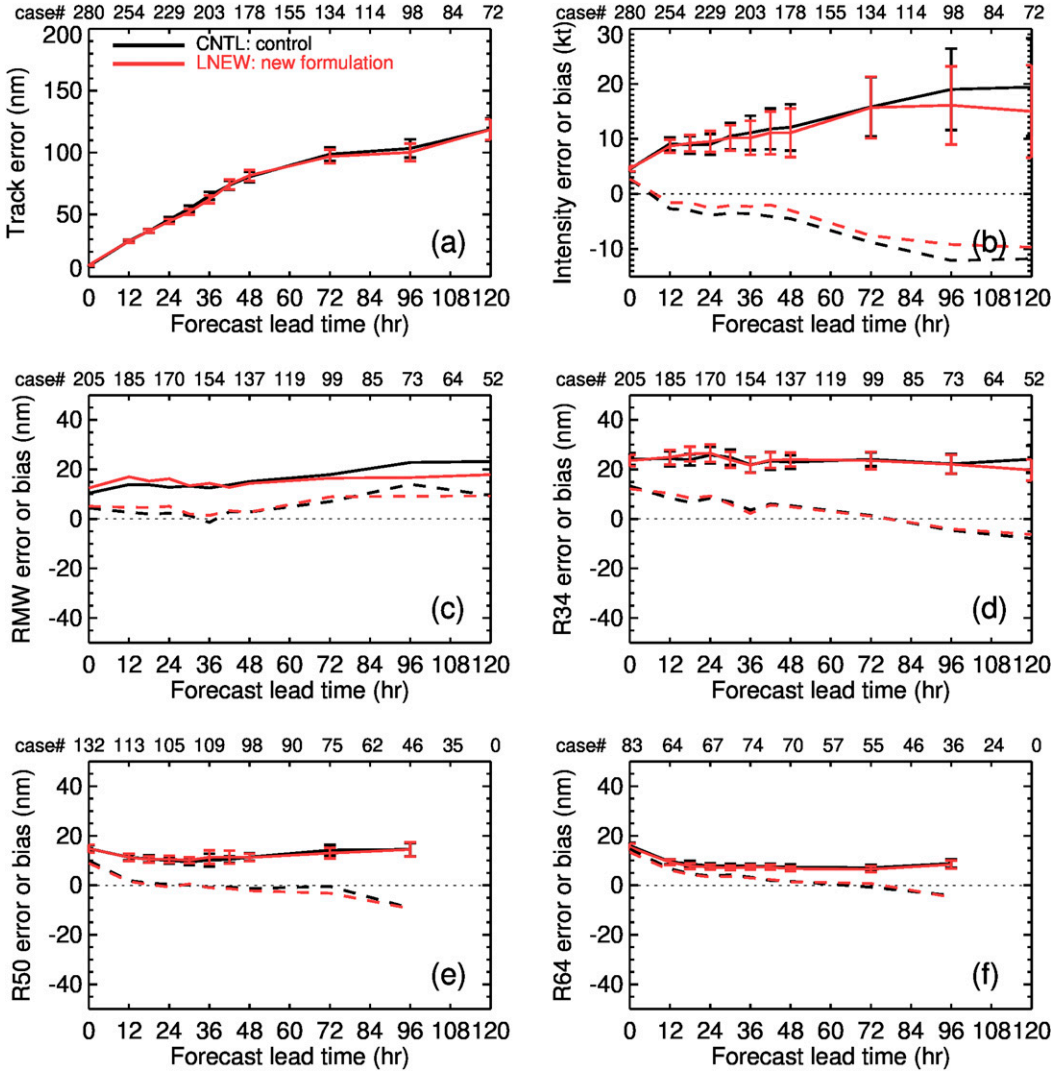


FIG. 15. As in Fig. 14, but for the EPAC basin.

degrades the R34 forecast. Given that the overall magnitudes of the horizontal mixing length of both simulations are close, those improvements of LNEW are associated with the spatial variation of the horizontal mixing length developed in the new formulation, particularly over the area within the distance of a few RMWs from the vortex center, producing more realistic simulations.

We also conduct a stratified verification analysis by grouping all simulations into 142 weak and 169 strong cycles based on the intensity at the initial time with a threshold of 50 kt. For strong cycles, LNEW slightly improves the track simulation for the lead times earlier than 72 h, with the mean track error being reduced by 3%–5%. However, LNEW improves the intensity simulation at nearly all lead times, with the mean intensity error and bias being reduced by 10%. The RMW error and bias of LNEW are also reduced for the lead times through 96 h. R34, R50, and R64 errors are nearly the same for both simulations. The performance of LNEW for weak cycles is a little different. For weak cycles, the mean track errors are reduced by 10% for all lead times, but the mean intensity errors are increased and reduced approximately by 10% for lead times before and after 48 h, respectively. LNEW produces the mean error of RMW smaller than CNTL for the lead times of 24–96 h, with nearly the same mean bias for all lead times. The mean errors of R34, R50, and R64 of LNEW are close to CNTL, with a slight degradation at some forecast lead times. Therefore, the new $L_h$ formulation could help the HWRF Model in track forecasts more for weak cycles than for strong cycles of NATL hurricanes. This is opposite to intensity forecasts.

Figure 15 presents the verification plots for the EPAC hurricanes. The verification includes 280 verifiable cycles. The track errors of both LNEW and CNTL are close, with the error of LNEW being only 3% smaller than CNTL. But the intensity forecast of LNEW is improved for nearly all lead times, with the mean error being reduced by 10% and the maximum

reduction reaching 20% at 120 h. The intensity bias is also significantly reduced for all lead times approximately by 50% in the early lead times and by 10%–20% in the late lead times. The mean error and bias of RMW are increased in the early lead times but reduced significantly in the late lead times. LNEW produces nearly the same mean errors of R34, R50, and R64 as CNTL does. We conduct the same stratified verification for the simulations of the EPAC hurricanes. The verification analysis for 132 strong cycles suggests that LNEW significantly improves the intensity forecast, with the mean error and bias being reduced by more than 10%, but degrade the track forecast slightly for the lead times after 36 h. The vortex size errors of the strong cycles of both simulations are nearly the same except that LNEW produces better RMW and R34 simulations in the late lead times. The verification for 148 weak cycles suggests that LNEW reduces the mean track error approximately by 5%–12% at all lead times after 24 h through 120 h. But LNEW increases the mean intensity error by 10% for the lead times before 84 h and reduced by 16% after that. The intensity biases of both simulations are very close except that LNEW reduces the bias after 60 h. Regarding the size of the simulated vortexes, LNEW degrades the RMW forecast in the lead time before 48 h and improves after that. The R50 error of LNEW is reduced and the R34 and R64 errors are close to those of CNTL. Therefore, the use of the new $L_h$ formulation in the HWRF Model could improve track forecasts for weak cycles, and intensity forecasts for strong cycles of EPAC hurricanes, which is the same as the conclusion for NATL hurricanes.

## 5. Summary and discussion

This study introduces a new formulation of horizontal mixing length for the parameterization of horizontal diffusion in TC simulations, which is expressed as a function of local wind and its horizontal gradients. In many current numerical models, the horizontal mixing length is simply formulated as a fixed fraction of grid size or a constant value over a given domain. The new formulation produces the horizontal mixing length and eddy diffusivity much closer to observations than does the currently used formulation. To test and understand the impact of the new formulation on TC simulations, the HWRF Model is configured to simulate the evolution of an idealized vortex in three experimental runs using the new formulation and constant values (750 m and 2 km) of horizontal mixing length. In the vortex region, there is a pattern that the mixing-length value near the vortex center is smaller than that outside the eyewall. The maximum value of horizontal eddy diffusivity appears near the surface in the area outside the eyewall. This is consistent with the results derived from observations, which indicates that both horizontal mixing length and eddy diffusivity do not peak in the eyewall area containing TC's most powerful winds. In the simulations using a constant value of horizontal mixing length, the horizontal eddy diffusivity in the vortex area generally increases with wind speed, which is not realistic according to observations. The vortex that is simulated using the constant value of 2 km exhibits the weakest inner core and the fastest broadening near the surface. The simulation using the new formulation produces the vortex intensity close to that using the constant value of 750 m, with the vortex expansion close to that using the constant value of 2 km. The HWRF Model with the new formulation simulates a more compact vortex, with the tilt of the vortex eyewall closer to the observation-based diagnosis. In real-case experiments, the new formulation significantly improves the 5-day forecasts of track and intensity of Hurricanes Harvey (2017) and Lane (2018). Retrospective simulations of 591 forecast cycles of 24 hurricanes in the NATL and EPAC basins using the latest HWRF system indicate that the new formulation can improve the forecasts of track, intensity, and size. A stratified verification suggests that the new formulation could help to improve the track forecast of weak cycles and the intensity forecast of strong cycles; this is consistent with that of a similar retrospective test (Wang et al. 2020) in the HMON model, another NCEP hurricane model (Mehra et al. 2018). These promising results give us confidence to implement the new formulation in the operational HWRF Model to further improve TC forecasts.

The horizontal mixing length in the new formulation is not a constant and varies with flow characteristics. It is especially useful and needed for numerical models to simulate a complex flow system such as TCs where horizontal mixing length varies significantly with location. In this case, treating the horizontal mixing length as a fixed fraction of grid spacing or a constant may misrepresent the effect of horizontal diffusion, especially when the model grid spacing is much larger than the actual mixing-length scale. Only when model resolution is very high compared to the actual mixing-length scale (e.g., in large-eddy simulations), using a fixed fraction of grid spacing may be sufficient to represent a subgrid horizontal diffusion contribution that is relatively small.

The new formulation can be applied to multiscale simulations over multiple domains. However, it should be noted that the scaling factor (or $\alpha$) in the new formulation is dependent on grid spacing as well as the algorithm of finite difference approximating derivatives in different numerical models. As a result, the best choice for the scaling factor could vary not only with grid spacing but also with different models. Further investigations on this dependence are warranted to facilitate the implementation of the new formulation. In addition, the scaling factor is determined empirically by comparing the distribution of the resulting horizontal length with that from observations. In our application of TC simulations, the scaling factor is chosen to make sure the resulting horizontal mixing length is around 500–800 m in the eyewall area, based on results derived from composite observations. With this approach, our tests indicate that the scaling factor $s$ values in the HWRF Model can be 0.0197, 0.0120, and 0.0098 (corresponding to $\alpha$ values of 1.45, 1.16, and 0.94) over the domains with grid spacing values of 13.5, 4.5, and 1.5 km, respectively. It should be kept in mind that those choices as well as the observed horizontal length scale are still uncertain due to turbulence observations for TCs being very limited at this time. The observed horizontal length scale could also depend on height and different TC scenarios. In this regard, more observations under a wide range of TC scenarios at different vertical

levels on the turbulence scale will benefit the parameterization representing subgrid-scale fluxes. The formulation can be further improved, e.g., by including other physical factors affecting horizontal diffusion such as turbulence kinetic energy and thermal structure.

Finally, previous studies have suggested that the intensity of simulated TCs increases with decreasing horizontal mixing length. In our real-case tests, it is also found that the track of simulated TCs is also significantly impacted by different horizontal mixing-length formulations. It is warranted to investigate how large-scale fields and TC structure respond to horizontal diffusion parameterizations and their impacts on track and intensity forecasts in subsequent studies.

*Acknowledgments.* This work is supported by NOAA's Hurricane Forecast Improvement Project. Thanks are due to Drs. Jili Dong and Hyun-Sook Kim for providing a thorough internal review and insightful comments and three anonymous reviewers for their constructive comments and suggestions, which helped us to improve the manuscript. We thank Mary Hart at EMC for editorial support.

*Data availability statement.* All datasets and results produced by the HWRF Model used in this study are available for public release upon request. Observational datasets analyzed in this study were from existing published data, which are included in Zhang and Montgomery (2012) and Stern and Nolan (2009).

## REFERENCES

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

Biswas, M. K., and Coauthors, 2018: Hurricane Weather Research and Forecasting (HWRF) Model: 2017 scientific documentation. NCAR Tech. Note NCAR/TN-544+STR, https://doi.org/10.5065/D6MK6BPR.

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, 2009a: 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 ——, 2009b: The effects of small-scale turbulence on maximum hurricane intensity. *13th Conf. on Mesoscale Processes*, Salt Lake City, UT, Amer. Meteor. Soc., 14.2, https://ams.confex.com/ams/pdfpapers/154827.pdf.

——, ——, 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.

Chan, J. C. L., and R. T. Williams, 1987: Analytical and numerical studies of the beta-effect in tropical cyclone motion. Part I: Zero mean flow. *J. Atmos. Sci.*, **44**, 1257–1265, https://doi.org/10.1175/1520-0469(1987)044<1257:AANSOT>2.0.CO;2.

Du, J., 2011: GCIP/EOP Surface: Precipitation NCEP/EMC 4KM Gridded Data (GRIB) Stage IV data, version 1.0. UCAR/NCAR–Earth Observing Laboratory, accessed 1 December 2020, https://doi.org/10.5065/D6PG1QDD.

Ferrier, B. S., Y. Jin, Y. Lin, T. Black, E. Rogers, and G. DiMego, 2002: Implementation of a new grid-scale cloud and precipitation scheme in the NCEP Eta model. *19th Conf. on Weather Analysis and Forecasting/15th Conf. on Numerical Weather Prediction*, San Antonio, TX, Amer. Meteor. Soc., 10.11, https://ams.confex.com/ams/SLS_WAF_NWP/techprogram/paper_47241.htm.

Galarneau, T. J., Jr., and C. A. Davis, 2013: Diagnosing forecast errors in tropical cyclone motion. *Mon. Wea. Rev.*, **141**, 405–430, https://doi.org/10.1175/MWR-D-12-00071.1.

Han, J., and H.-L. Pan, 2011: Revision of convection and vertical diffusion schemes in the NCEP global forecast system. *Wea. Forecasting*, **26**, 520–533, https://doi.org/10.1175/WAF-D-10-05038.1.

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

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.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1644, https://doi.org/10.1175/BAMS-83-11-1631.

Liu, B., H. Liu, L. Xie, C. Guan, and D. Zhao, 2011: A coupled atmosphere–wave–ocean modeling system: Simulation of the intensity of an idealized tropical cyclone. *Mon. Wea. Rev.*, **139**, 132–152, https://doi.org/10.1175/2010MWR3396.1.

Mehra, A., V. Tallapragada, Z. Zhang, B. Liu, L. Zhu, W. Wang, and H. Kim, 2018: Advancing the state of the art in operational tropical cyclone forecasting at NCEP. *Trop. Cyclone Res. Rev.*, **7**, 51–56, https://doi.org/10.6057/2018TCRR01.06.

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.

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.

Tang, J., J. Zhang, C. Kieu, and F. D. Marks, 2018: Sensitivity of hurricane intensity and structure to two types of planetary boundary layer parameterization schemes in idealized HWRF simulations. *Trop. Cyclone Res. Rev.*, **7**, 201–211, https://doi.org/10.6057/2018TCRR04.01.

Wang, W., X. Shen, and W. Huang, 2016: A comparison of boundary-layer characteristics simulated using different parametrization schemes. *Bound.-Layer Meteor.*, **161**, 375–403, https://doi.org/10.1007/s10546-016-0175-4.

——, J. 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.

——, and Coauthers, 2020: Testing a new horizontal mixing-length formulation in HMON. *Research Activities in Earth System Modelling*, Working Group on Numerical Experimentation, E. Astakhova, Ed., World Meteorological Organization, 4–13.

Zhang, J. A., 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.

——, ——, J. A. Sippel, R. F. Rogers, X. Zhang, S. G. Gopalakrishnan, RZ. Zhang, and V. Tallapragada, 2018: Evaluating the impact of improvement in the horizontal diffusion parameterization on hurricane prediction in the operational Hurricane Weather Research and Forecast (HWRF) Model. *Wea. Forecasting*, **33**, 317–329, https://doi.org/10.1175/WAF-D-17-0097.1.