Improving Hurricane Intensity and Streamflow Forecasts in Coupled Hydrometeorological Simulations by Analyzing Precipitation and Boundary Layer Schemes

MD MURAD HOSSAIN KHONDAKER[a] AND MOSTAFA MOMEN[a]

[a] *Department of Civil and Environmental Engineering, University of Houston, Houston, Texas*

(Manuscript received 7 September 2023, in final form 3 May 2024, accepted 3 June 2024)

ABSTRACT: Hurricanes have been the most destructive and expensive hydrometeorological event in U.S. history, causing catastrophic winds and floods. Hurricane dynamics can significantly impact the amount and spatial extent of storm precipitation. However, the complex interactions of hurricane intensity and precipitation and the impacts of improving hurricane dynamics on streamflow forecasts are not well established yet. This paper addresses these gaps by comprehensively characterizing the role of vertical diffusion in improving hurricane intensity and streamflow forecasts under different planetary boundary layer, microphysics, and cumulus parameterizations. To this end, the Weather Research and Forecasting (WRF) atmospheric model is coupled with the WRF hydrological (WRF-Hydro) model to simulate four major hurricanes landfalling in three hurricane-prone regions in the United States. First, a stepwise calibration is carried out in WRF-Hydro, which remarkably reduces streamflow forecast errors compared to the U.S. Geological Survey (USGS) gauges. Then, 60 coupled hydrometeorological simulations were conducted to evaluate the performance of current weather parameterizations. All schemes were shown to underestimate the observed intensity of the considered major hurricanes since their diffusion is overdissipative for hurricane flow simulations. By reducing the vertical diffusion, hurricane intensity forecasts were improved by ~39.5% on average compared to the default models. These intensified hurricanes generated more intense and localized precipitation forcing. This enhancement in intensity led to ~16% and ~34% improvements in hurricane streamflow bias and correlation forecasts, respectively. The research underscores the role of improved hurricane dynamics in enhancing flood predictions and provides new insights into the impacts of vertical diffusion on hurricane intensity and streamflow forecasts.

SIGNIFICANCE STATEMENT: Despite significant recent improvements, numerical weather prediction models struggle to accurately forecast hurricane intensity and track due to many reasons such as inaccurate physical parameterization for hurricane flows. Furthermore, the performance of existing physics schemes is not well studied for hurricane flood forecasting. This study bridges these knowledge gaps by extensively evaluating different physical parameterizations for hurricane track, intensity, and flood forecasts using an atmospheric model coupled with a hydrological model. Then, a reduced diffusion boundary layer scheme is developed, making remarkable improvements in hurricane intensity forecasts due to the overdissipative nature of the considered schemes for major hurricane simulations. This reduced diffusion model is shown to significantly enhance hurricane flood forecasts, indicating the significance of hurricane dynamics on its induced precipitation.

KEYWORDS: Boundary layer; Intensification; Diffusion; Hurricanes/typhoons; Hydrometeorology; Numerical weather prediction/forecasting

1. Introduction

Hurricanes have been the costliest natural disaster in the United States, inflicting more than $1 trillion in total estimated damage between 1980 and 2021 (Smith 2021; Chen 2022). A significant component of hurricane damage is caused by torrential rainfalls that induce catastrophic floods in coastal areas. For instance, ~1.25 million people experienced over 45 in. of rain in 7 days when Hurricane Harvey made landfall in Texas in 2017. The damage from this devastating hurricane event, which cost more than ~$100 billion, made it the second costliest natural catastrophe in U.S. history thus far (NOAA 2022; Blake and Zelinsky 2018). Future hurricanes may become even more intense (Mei and Xie 2016; Emanuel 2005; Cheikh and Momen 2020) and more frequent (Emanuel 2017) and induce more rainfalls (Shearer et al. 2022; Reed et al. 2022; Liu et al. 2019) due to climate change. It is thus crucial to accurately characterize hurricane dynamics and improve their forecasts to reduce some aspects of hurricane damage.

Numerical weather prediction (NWP) models play a crucial role in predicting these extreme weather events. The Weather Research and Forecasting (WRF) Model (Skamarock et al. 2019), the state-of-the-art mesoscale atmospheric NWP system, is often utilized in the United States for weather research and forecasts. The WRF model has been used in many studies to investigate different hurricane physics, which led to several

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JHM-D-23-0153.s1.

Corresponding author: Mostafa Momen, mmomen@uh.edu

DOI: 10.1175/JHM-D-23-0153.1

remarkable findings such as rainband–eyewall interactions (Davis et al. 2008; Nolan et al. 2021b, 2013; Yang et al. 2019; Jeworrek et al. 2019; Zhou and Wang 2009), turbulent mixing in hurricane flows (Zhu et al. 2019), and the impacts of precipitation physics on hurricane simulations (Nasrollahi et al. 2012).

Hurricane-induced flood forecasts strongly depend on hurricane intensity and track predictions. Despite the recent enhancements of model physics and cloud-permitting grid resolutions in NWP models like WRF for hurricane simulations, they still face many challenges to accurately predict hurricane intensity and track (Emanuel 2017; Xue et al. 2013; Romdhani et al. 2024). One of the primary causes of this disparity is that the current physical parameterizations are neither specifically designed nor evaluated for simulating hurricanes, which have unique dynamics (Zhang 2010; Momen et al. 2021; Sabet et al. 2022; Li et al. 2023). Hurricanes can alter the typical coherent structure of turbulent eddies in regular atmospheric flows. Strong rotation in hurricanes leads to the formation of smaller eddies with reduced turbulent length scales compared to the regular atmospheric boundary layers (Momen et al. 2021). Romdhani et al. (2022) demonstrated that the Advanced Research version of WRF's (ARW) default horizontal diffusion is overly diffusive and found that decreasing diffusion considerably improves hurricane intensity estimations. Previous studies also found that reducing the default vertical eddy diffusion in WRF leads to significantly improved hurricane structure and intensity forecasts compared to observations over the ocean (J. A. Zhang et al. 2017; Matak and Momen 2023; Gopalakrishnan et al. 2013; Zhang et al. 2015; Zhang and Pu 2017; F. Zhang et al. 2017). The reason for this eddy diffusion decrease is due to the strong rotational effects that exist in hurricane flows, which can suppress turbulence production similar to rotating shear flows (Tritton 1992; Cazalbou et al. 2005; Durbin 2011; Arolla and Durbin 2014). Such diffusion reductions also lead to shallower boundary layers, greater hurricane inflow, deeper updrafts, and enhanced boundary layer convergence to the eye of the hurricane (Ming et al. 2023; J. A. Zhang et al. 2017).

To forecast hurricane-induced flood inundation levels, the meteorological forecasts need to be coupled with a hydrological model. To this end, the WRF atmospheric model is recently coupled to the WRF hydrological (WRF-Hydro) model. WRF-Hydro is a leading-edge framework that bridges the atmospheric and hydrological modeling systems, allowing a physics-based, fully coupled surface hydrology-regional atmospheric modeling capability for hydrometeorological applications (Maidment 2017; Gochis et al. 2020). Since 2016, the National Water Model (NWM), which is based on the WRF-Hydro model architecture (Gochis et al. 2020), has become a fully operational hydrologic forecasting system providing real-time, high-resolution, distributed hydrologic forecasts for the contiguous United States (Lahmers et al. 2019). The hydrological parameters of this model (e.g., soil, runoff, groundwater, and vegetation) have been calibrated for streamflow predictions using available streamflow gauges in the United States (Tijerina-Kreuzer et al. 2021; Lahmers et al. 2021; Mascaro et al. 2023). The WRF-Hydro model has been widely used and evaluated for streamflow forecasts, especially for flash floods (Coelho et al. 2022; Viterbo et al. 2020; Senatore et al. 2020; Lin et al. 2018a; Zhang et al. 2020) and hurricane-induced flood predictions (Abbaszadeh et al. 2020; Kim et al. 2021; Jafarzadegan et al. 2021; Chen et al. 2021; El Gharamti et al. 2021; Yin et al. 2021, 2022a,b; Bao et al. 2022; Zhang et al. 2020; Katsafados et al. 2018; Shastry et al. 2019). Some recent developments of WRF-Hydro and NWM include introducing channel infiltration for semiarid regions (Lahmers et al. 2019, 2021), a vector-based channel network routing parameterization (Lin et al. 2018b), a new river boundary parameterization (Jafarzadegan et al. 2021), and a 2D groundwater scheme (Rummler et al. 2022).

Accurate precipitation forcing is essential for correct hurricane-induced flood forecasts. Several studies found that the input precipitation forcing data can significantly impact the accuracy of streamflow forecasts (Ma et al. 2021; Chao et al. 2021; Arnault et al. 2018). Precipitation is typically one of the most challenging and uncertain variables in NWP models (Nielsen-Gammon et al. 2005). The precipitation data can be ingested into WRF-Hydro in two ways. The WRF-Hydro system can either be coupled with the WRF atmospheric model (WRF-ARW) to receive precipitation forcing or employ such data in the uncoupled mode from observations including rain gauges, radar, and remote sensing (Gochis et al. 2020). The uncoupled simulations are useful for model calibration and spinup in the hindcasting mode, whereas coupled simulations are necessary for real-time and long-term forecasts.

The coupled ARW-WRF-Hydro model has been used in several hydrological and quantitative precipitation forecasting studies (Pal et al. 2021; Arnault et al. 2018; Lahmers et al. 2020; Verri et al. 2017; Senatore et al. 2020, 2015). For instance, Yucel et al. (2015) assessed the efficacy of WRF-derived precipitation in forecasting flood hydrograph characteristics before and after 3D data assimilation. Some studies also characterized the impacts of different planetary boundary layer (PBL) parameterizations on terrestrial water flow uncertainties (Arnault et al. 2018). However, most of these studies primarily focused on regular flood events or general watershed behavior, and extreme hurricane events have received less attention. Furthermore, the interacting effects of different PBL, microphysics, and cumulus parameterizations on hurricane-induced streamflow forecasts are not comprehensively established yet. Despite recent improvements in hurricane intensity predictions via eddy diffusion adjustments, it is not yet well known how such enhancements impact the accuracy of flood forecasts. The objective of this study is to address these knowledge gaps by conducting coupled ARW-WRF-Hydro simulations for four hurricanes in the United States. Unlike previous studies that only tested a single PBL type parameterization and used a different atmospheric dynamical core, we will use WRF-ARW and evaluate different PBL closures and adjustments. Moreover, unlike most prior efforts, we will couple the WRF-Hydro with WRF-ARW's precipitation forcing to directly determine the impacts of hurricane meteorological enhancements on hydrological forecasts. To this end, we will answer the following research questions:

1) How do different PBL, microphysics, and cumulus parameterizations affect the accuracy of real hurricane intensity, track, and flood forecasts?
2) What are the impacts of improvements in hurricane intensity predictions, via adjusting the default diffusion, on precipitation forcing and corresponding flood forecasts?

The paper addresses these questions as follows. Section 2 describes the selected hurricanes, the suite of stand-alone WRF-Hydro and coupled ARW-WRF-Hydro simulations, and evaluation metrics. Next, the results of WRF-Hydro calibrations and the impacts of different PBL, microphysics, and cumulus parameterizations on coupled simulations are discussed in section 3 (question 1). Then, the effects of adjusting the vertical diffusion on simulated hurricane intensity and precipitation, as well as the accuracy of streamflow forecasts, are characterized in this section (question 2). Finally, section 4 summarizes the main findings of this study.

2. Methods

a. Study areas and hurricane cases

In this study, three hurricane-prone regions in the United States were selected to conduct one-way coupled ARW-WRF-Hydro (denoted by "C-AWH" hereafter) simulations. Four category 4–5 hurricanes that made landfall in these regions are simulated to evaluate the impacts of different hydrometeorological parameterizations on hurricane intensity and precipitation. The three selected WRF-Hydro basins and the best-observed track for the simulated hurricanes, along with their WRF-ARW domain, are shown in Fig. 1. The details of the simulation periods and selected basins are summarized in Table 1. The most expensive hurricane in U.S. history, Katrina, made landfall in Louisiana in 2005 and caused an estimated damage of more than $150 billion (Knabb et al. 2023). Hurricane Harvey, the second costliest hurricane in the United States, hit Texas in 2017 and caused historically heavy rainfall and severe damage (Blake and Zelinsky 2018). In 2004, Hurricane Frances, along with three more hurricanes, swept across Florida within 6 weeks, resulting in an estimated $82 billion in damage. Hurricane Irma flooded low-lying neighborhoods and roadways in Florida in 2017 and caused more than $75 billion in damages (Cangialosi et al. 2021). In total, the four hurricanes caused a total of more than $400 billion in economic damages (Blake et al. 2011).

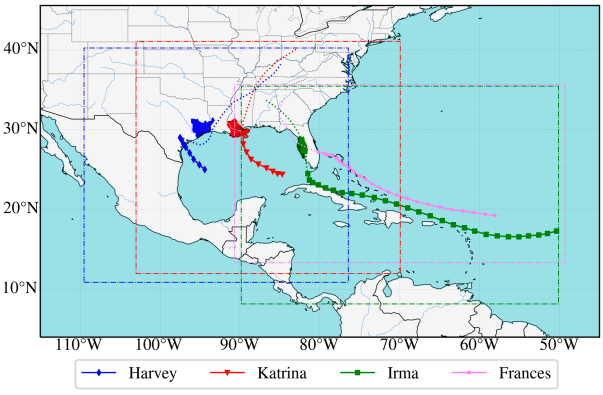

FIG. 1. The best-observed track of the four selected hurricanes. The solid lines show the tracks of the hurricanes during the simulation period with at least category one strength before landfall, and dotted lines denote their simulated trajectory 2–3 days after landfall. The dashed boxes show the simulated domain for each hurricane. The markers indicate the position of the hurricane's eye every 6 h. The shaded regions represent HUC-8 subbasins in Texas (blue), Louisiana (red), and Florida (green) for streamflow evaluations.

All these major hurricanes led to significant precipitation-induced flooding. To evaluate the simulated flood response of these hurricanes, we used 64 U.S. Geological Survey (USGS) streamflow gauges in the selected areas. The subbasins are selected based on the Hydrologic Unit Code 8 (HUC-8) in the hurricane-inflicted areas (Fig. 2). These subbasins are in hurricane-prone areas of the United States, with hurricane return periods ranging from 7 to 9 years (Blake et al. 2011). The details of the number of gauges and hydrological basins are shown in Table 1.

b. PBL schemes and reducing vertical exchange coefficient

The PBL is the lowest and most turbulent layer of the atmosphere, which is in direct contact with Earth's surface and

TABLE 1. Selected hurricanes and their simulation periods in C-AWH models and their corresponding hydrologic subbasins and the number of gauges for streamflow evaluation.

Hurricane (year)	C-AWH simulation period	Maximum wind intensity (m s^{-1})	Region (HUC-4, HUC-8)	No. of gauges
Harvey (2017)	0000 UTC 25 Aug–0000 UTC 5 Sep	59	HUC-4: 1204, HUC-8: 12020003,6-7, 12030202-3,12010005	31
Katrina (2005)	0000 UTC 27 Aug–0000 UTC 5 Sep	77.2	HUC-4: 0807, 0809, HUC-8: 03180004	8
Irma (2017)	0000 UTC 4 Sep–0000 UTC 21 Sep	79.7	HUC-4: 0310	25
Frances (2004)	1200 UTC 30 Aug–0000 UTC 13 Sep	64.3	HUC-4: 0310	25

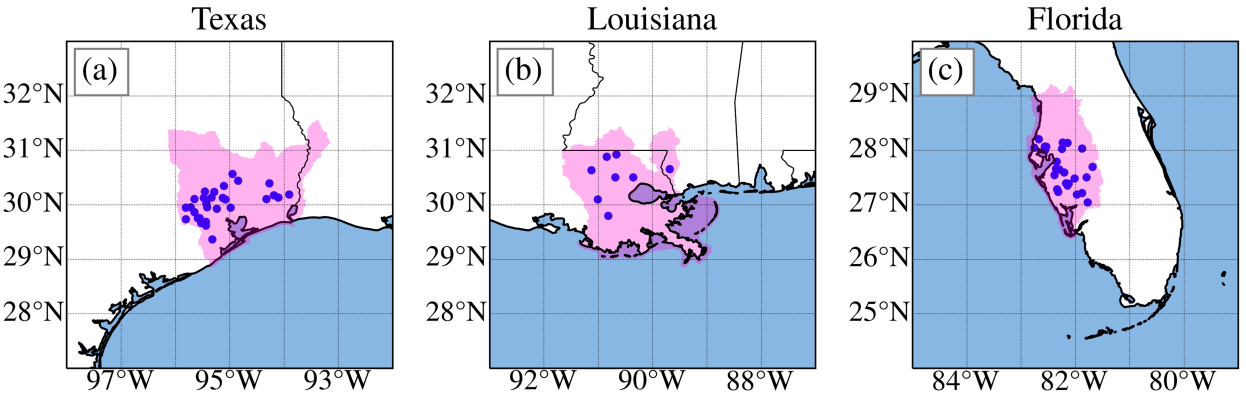

FIG. 2. Location of the USGS gauges (blue dots) within (a) Texas, (b) Louisiana, and (c) Florida for the considered subbasins of this study (magenta color).

extends up to ~1–2 km above the ground (Momen and Bou-Zeid 2016, 2017). The PBL schemes are used to parameterize the unresolved turbulent vertical fluxes of momentum, heat, and scalars (e.g., moisture) inside the PBL (Momen 2022; Momen et al. 2018) and throughout the atmosphere. In the current study, two widely used Yonsei University (YSU; Hong et al. 2006) and Mellor–Yamada–Janjić (MYJ; Janjić 1994) PBL schemes are employed for simulating hurricanes. The MYJ scheme is a local closure model that calculates vertical turbulent fluxes using average local atmospheric variables. On the other hand, the YSU model implicitly assesses the nonlocal fluxes through a parameterized nonlocal term (Hu et al. 2010).

The YSU scheme is recommended by the WRF user guide for hurricane simulations (Wang et al. 2018). In the YSU scheme, the momentum diffusivity coefficient K_m is defined as

$$K_m = \kappa w_s z\left(1 - \frac{z}{h}\right)^p, \tag{1}$$

where p is the profile shape exponent (=2 in YSU), κ is the von Kármán constant (=0.4), z is the height from the surface, and h is the height of the PBL. Using the Prandtl number relation, the eddy diffusivity for moisture and temperature is determined from K_m (Hong et al. 2006). The MYJ PBL scheme determines K_m from the prognostic turbulent kinetic energy (TKE) equations.

Due to the overdiffusive nature of current turbulence schemes in WRF for major hurricane simulations (Romdhani et al. 2022), the depth of the eddy diffusivity profile is adjusted in all these schemes to control the magnitude of the vertical diffusion and consequently the intensity of the simulated hurricanes. To this end, in the newly modified runs, the vertical eddy diffusivity is set to zero from a certain elevation to the topmost level of the domain following a method described in Matak and Momen (2023). In this method, a limiter is used to set the vertical eddy momentum diffusivity coefficient K_m in Eq. (1) to zero from a certain level above the surface z_{level_x} as $K_m(z > z_{\text{level}_x}) = 0$. This level is set to the fourth vertical level here, which is equivalent to ~260-m height above the surface. This limiter effectively decreases the hurricane boundary layer depth, making it more consistent with observations. For example, previous observational estimates showed that the mixed-layer depth decreases to ~250 m at the hurricane eyewall based on the composite dropsonde data (Zhang et al. 2011). This method is validated in Matak and Momen (2023), which showed that reducing the depth of the vertical diffusion in the YSU scheme had superior performance (~44% average intensity improvement) compared to reducing its magnitude (~25% average intensity improvement) in different tested major hurricane simulations.

This approach is called reduced diffusion (RD) modulation hereafter. This method is also consistent with previous studies (J. A. Zhang et al. 2017; Gopalakrishnan et al. 2013), where lower vertical eddy diffusivities are found to be more accurate for high-intensity hurricane simulations. The modified vertical momentum and scalar exchange coefficient profiles are shown in Figs. S1 and S2 in the online supplemental material, respectively. Note that since the intensification of hurricanes occurs over the ocean and they rapidly decay over land, this modification is only applied for hurricanes when they are over the ocean (Tang et al. 2018).

c. *The suite of SWH simulations*

To evaluate the hydrological performance in the selected domains, a thorough stepwise calibration is initially conducted for the WRF-Hydro model in uncoupled mode. The preliminary calibrated parameter values of NWM version 2.0 are used to initiate this calibration for the selected study areas. A total of 13 runoff, channel, vegetation, and soil parameters are considered for calibration runs. 201 standalone WRF-Hydro (denoted by "SWH" hereafter) simulations are performed to assess the sensitivity of these variables (please refer to Table S1 for a list of all the WRF-Hydro calibration parameters). For instance, "refkdt" is one of the important parameters calibrated in these basins. The parameter refkdt is a scaling parameter for surface runoff that controls the amount of runoff produced for a given volume of precipitation (Cerbelaud et al. 2022). It considerably impacts surface infiltration, hence partitioning total runoff into surface and subsurface runoff (Schaake et al. 1996). Increasing

TABLE 2. The suite of C-AWH simulations for different PBL, MP, and CU parameterizations. Total number of simulations = 9 × 4 = 36 using a CU scheme with 8-km grid resolution. For the grid sensitivity test, 24 No_CU simulations were conducted with 2-, 8-, and 32-km resolutions. The boldface texts highlight the selected PBL, MP, and CU schemes for running the reduced diffusion cases.

Case name	PBL	MP	CU	Diffusion
Base	YSU	WSM6	KF	Default
Base_RD	**YSU**	WSM6	KF	**Reduced**
PBL_MYJ	MYJ	WSM6	KF	Default
MP_Thom	YSU	Thompson	KF	Default
MP_Thom_RD	YSU	**Thompson**	KF	**Reduced**
MP_WDM	YSU	WDM6	KF	Default
CU_GF	YSU	WSM6	GF	Default
CU_GF_RD	YSU	WSM6	**GF**	**Reduced**
CU_MSKF	YSU	WSM6	Multiscale KF	Default
No_CU (grid sensitivity)	YSU	WSM6	No CU	Default and reduced

refkdt decreases the simulated volume of hydrographs (surface runoff) and vice versa. Prior studies have found refkdt to be a very sensitive parameter (Lahmers et al. 2019; Rummler et al. 2022), which should be optimized for realistic channel flow and soil moisture results.

The precipitation forcing of these uncoupled simulations is provided by the hourly North American Land Data Assimilation System phase 2 (NLDAS-2) dataset (Xia et al. 2014). The Earth System Modeling Framework (ESMF) regridding code is used to prepare the forcing data for the working resolution of the WRF-Hydro model. The WRF-Hydro system is configured based on the NWM configuration with 1-km terrain routing and 250-m channel routing resolution (Gochis et al. 2020). For all the cases, a spinup of 3 years is done in WRF-Hydro before simulating the hurricane events. The last year of the spinup period is also used to evaluate the calibration performance during the normal year period.

d. The suite of C-AWH simulations

In this study, a comprehensive sensitivity analysis of the simulated hurricanes and their induced floods is conducted by varying the PBL, microphysics, and cumulus schemes. Three sets of WRF-ARW simulations are performed with a grid resolution of 8 km. In the first set, the YSU and MYJ PBL schemes are used to assess the impacts of different PBL schemes on WRF-ARW and WRF-Hydro simulations. In the second suite of simulations, three different microphysics parameterization schemes, WRF single-moment 6-class (WSM6; Hong and Lim 2006), Thompson (Thompson et al. 2008), and WRF double-moment 6-class (WDM6; Lim and Hong 2010), are considered for evaluating the influence of microphysics parameterizations on hurricane-induced flood forecasts. Finally, in the last suite, three cumulus schemes, the Kain–Fritsch (KF) scheme (Kain 2004), the Grell–Freitas (GF) scheme (Grell and Freitas 2014), and the Multiscale Kain–Fritsch (MSKF) scheme (Zheng et al. 2016), along with a case with no cumulus scheme, are used to simulate hurricanes. A summary of all conducted C-AWH cases and their configurations is listed in Table 2.

In Table 2, the Base case includes the YSU PBL scheme (recommended by WRF for hurricane simulations), WSM6 microphysics scheme, and KF cumulus scheme, which is used in all cases for intercomparison. In section 3b, the results of the default WRF models will be presented. In section 3c, new suites of simulations will be shown for the reduced diffusion PBL scheme (shown as *_RD) for the three best cases in each set aiming for improvement in hurricane intensity predictions. In total, 4 × 9 = 36 coupled ARW-WRF-Hydro simulations are conducted using a cumulus scheme with 8-km grid resolution.

Finally, to corroborate the generality of the findings, a grid resolution sensitivity test is conducted with 2-, 8-, and 32-km grid sizes for all considered hurricanes, which are shown in supplemental material. This test is conducted by turning off the cumulus parameterization (No_CU) since cumulus schemes are not recommended for grid resolutions smaller than 4 km (Li and Bou-Zeid 2014; Jeworrek et al. 2019). We chose the 2-km fine resolution to ensure it does not fall in the gray zone for using the cumulus scheme (~3–10 km). For the four considered hurricanes, we did two diffusion cases (default and reduced) for each grid resolution, resulting in 4 (hurricanes) × 2 (diffusion) × 3 (resolutions) = 24 additional NO_CU simulations in total. Note that these eight 2-km grid resolution simulations consumed about 2 million CPU core hours and 14 TB of storage space, and thus, conducting all cases with such a fine resolution was out of our computational and storage resources. Hence, all other cases were run using an 8-km grid resolution, which will be shown to provide comparable and sufficient accuracy for the considered cases. A similar outcome was also demonstrated in prior studies (Matak and Momen 2023).

For all cases, the WRF-ARW results are regridded into WRF-Hydro forcing files with 1-km spatial resolution using the ESMF regridding tool and used in WRF-Hydro simulations. Restart files from the spinup period are used to initialize WRF-Hydro simulations with the initial model state. The forcing dataset for WRF-ARW simulations is generated using the NCEP final (FNL) operational global analysis dataset using the WRF preprocessing system (WPS).

e. Evaluation metrics

To evaluate the accuracy of simulated streamflow forecasts compared to the observation, three metrics will be used: Pearson's correlation coefficient R, Kling–Gupta efficiency (KGE), and mean absolute percentage error (MAPE). The Pearson correlation coefficient is used to assess the correlation

between the streamflow forecasts and observed data and is calculated as follows:

$$R = \frac{\sum_{i=1}^{N}(q_s^i - \overline{q_s})(q_o^i - \overline{q_o})}{\sqrt{\sum_{i=1}^{N}(q_s^i - \overline{q_s})^2 \sum_{i=1}^{N}(q_o^i - \overline{q_o})^2}}, \quad (2)$$

where q_s^i and q_o^i are streamflow from simulation and observation at time step i, respectively; the overbar denotes the sample mean, and N represents the number of samples for each gauge. A high value of $R > 0.5$ shows that the simulation and observation have a positive linear relationship, and $R \approx 0$ means there is no correlation. For streamflow observations, the USGS gauges described in Table 1 and Fig. 2 are employed.

KGE is another commonly used metric (Gupta et al. 2009; Jackson et al. 2019), which weighs correlation, bias, and variance errors equally and is calculated as follows:

$$\text{KGE} = 1 - \sqrt{(R-1)^2 + \left(\frac{\overline{q_s}}{\overline{q_o}} - 1\right)^2 + \left(\frac{\sigma_s}{\sigma_o} - 1\right)^2}, \quad (3)$$

where σ_s and σ_o denote the standard deviations of simulated and observed streamflow, respectively. When KGE is close to one, it is considered optimal, and negative values of KGE indicate poor performance.

The next metric that will be used for evaluating both streamflow and hurricane intensity forecasts is MAPE. For streamflow, MAPE_{SF} evaluates how well the simulations capture the total flow volume rate and is defined as

$$\text{MAPE}_{\text{SF}} = \frac{1}{N}\sum_{i=1}^{N}\frac{|q_s^i - q_o^i|}{\overline{q_o}} \times 100\%. \quad (4)$$

While this metric is sometimes called normalized mean absolute error in the hydrology literature, we refer to it as MAPE_{SF} in this paper for consistency with the intensity metric. For WRF-ARW simulations, $\text{MAPE}_{\text{Intensity}}$ represents how well the simulations predict the actual hurricane intensity as the following:

$$\text{MAPE}_{\text{Intensity}} = \frac{1}{N}\sum_{i=1}^{N}\frac{|\text{ws}_s^i - \text{ws}_o^i|}{\text{ws}_o^i} \times 100\%, \quad (5)$$

where ws_s^i and ws_o^i are the 10-m surface wind speeds from simulation and observation, respectively.

Finally, the predicted track of hurricanes is evaluated using the mean absolute error (MAE) since there is no universal length scale to nondimensionalize hurricane track (Romdhani et al. 2022). To this end, the distance between the best-observed hurricane center X_o and the simulated hurricane center X_s is used to calculate track error by establishing the following $\text{MAE}_{\text{Track}}$:

$$\text{MAE}_{\text{Track}} = \frac{1}{N}\sum_{i=1}^{N}|X_s^i - X_o^i|, \quad (6)$$

where the distance between the simulated and observed hurricanes is calculated using the haversine formula (Choudhury and Das 2017). For intensity and track observations, the best-observed reported data from the U.S. National Hurricane Center (NHC) are used (Kidder et al. 2000; Klotz and Uhlhorn 2014; Huffman et al. 2015).

3. Results

a. *Calibrating WRF-Hydro parameters for the selected regions*

In this study, the WRF-Hydro stand-alone model is calibrated using a stepwise approach. First, 13 parameters are chosen based on preliminary in-house tests and previous calibration studies (Lahmers et al. 2019). Then, a thorough sensitivity analysis is performed using the recommended range of parameter values in all the domains separately. The calibrations are evaluated for the considered hurricane events as well as one normal year before each hurricane period to improve the intrinsic streamflow pattern of the selected basins and extreme event forecasts. In total, 201 simulations were performed for each domain. A list of the 13 parameters along with the details on this calibration can be found in Table S1.

Among all these parameters, the surface runoff parameter (refkdt), the vegetation parameter (mp), Manning's roughness n, and saturation soil moisture content (smcmax) were found to be the most sensitive ones, consistent with previous studies (Kilicarslan et al. 2021; Lahmers et al. 2019; Abbaszadeh et al. 2020). After an extensive parameter calibration test, refkdt and mp appeared to be the main variables that their modifications could consistently outperform NWM's default models in the tested periods. Hence, a stepwise calibration is done for each subbasin domain of Fig. 2 by focusing on adjusting refkdt and mp values and maximizing the KGE in all the considered streamflow gauges in these regions. Table 3 summarizes the suite of 88 conducted simulations for calibrating these parameters, their adjustment types, and calibrated parameter values for each domain.

The increase in refkdt increases the ground infiltration of the overland flow and consequently decreases the channel streamflow. Increasing Manning's roughness increases the surface friction and decreases the streamflow. All these streamflow responses are shown as an example for Hurricane Harvey in Fig. S3. Figure 3 shows the hydrographs of the calibrated model for three gauges from three subbasins. It demonstrates that the calibrated model captures the timing and magnitude of the streamflow peaks better than both NWM models. NWM 2.0 overpredicts the streamflow prediction for Hurricane Katrina in Louisiana (Fig. 3b), while it underpredicts the discharge for Hurricane Irma in Florida (Fig. 3c) in the considered gauges. Our calibrated model agrees more with the observed USGS gauge measurement and works better than these two versions of the NWM model.

To comprehensively analyze the performance of the calibrated model, the error statistics of all the considered USGS gauges were calculated. Figure 4 presents the average performance metric for all cases during the hurricane period and one normal year before that averaged for the four considered hurricane cases. As the figure indicates, the calibrated model

TABLE 3. Selected calibration parameters for WRF-Hydro, including their adjustment type, calibration ranges, final calibrated values, and the total number of simulations for all hurricane cases. Here, "—" means the parameter value is not calibrated by the indicated method type (replacement or multiplication) for that region, and the alternate method is used.

			Final calibrated value			
Parameter	Type	Tested values/multipliers	Texas	Louisiana	Florida	Number of simulations
refkdt	Multiplier	0.25, 0.33, 0.5, 0.66, 0.75, 1.5, 2	—	0.5	2	28
refkdt	Replace	0.25, 0.33, 0.5, 0.66, 0.75, 1, 2, 3	3	—	—	32
mp	Multiplier	0.5, 0.8, 0.9, 1.1, 1.25, 1.6, 2	1.6	NWM 2.0	1.25	28
		Total number of simulations				88

outperforms both NWM models on average in the considered regions. The calibrated model enhances the NWM 2.1's $MAPE_{SF}$ forecast by ~24.1% in the normal year before the hurricane (Fig. 4c) and by ~15.9% during the hurricane event (Fig. 4a). It also improves the NWM 2.0's $MAPE_{SF}$ and KGE forecasts by ~18.4% and ~80.9%, respectively, in the 1 year before the hurricane period (Figs. 4c,d) and by ~5.7% and ~37.8% during the hurricane period (Figs. 4a,b). Note that while the calibration has some low sensitivity to the considered rainfall scenario (e.g., normal or hurricane), we used a value that simultaneously improves all cases. The full calibration results for all the considered cases are shown in Table S2. Furthermore, to evaluate the generality of the calibrated parameters, we tested them for Hurricanes Ike and Jeanne, which are not included in the calibration process. It is found that this calibration outperforms both default NWM cases for these two separate hurricane events, validating its generality (see Fig. S4). Hence, this calibrated WRF-Hydro model will be used for the coupled C-AWH runs in the next section since it consistently reduces the forecasting errors of the default NWM models.

b. *Coupled ARW-WRF-Hydro default cases*

After calibrating the hydrological parameters in the selected regions, the impacts of meteorological parameterizations on hurricane-induced floods are characterized by coupling WRF-Hydro's precipitation forcing with the WRF-ARW atmospheric model. The coupling in this study is implemented one way where the meteorological and precipitation forcing data from the WRF-ARW are ingested into the WRF-Hydro model. The reason for this is that compared to one-way coupling, two-way coupled runs (feedback from surface fluxes of WRF-Hydro over land to WRF-ARW) are found to have a minimal influence on the total precipitation amount and streamflow forecasts (Givati et al. 2016; Senatore et al. 2015; Rudisill et al. 2022). We also demonstrated that two-way coupling had less than a 2% impact on hurricane intensity and track error forecasts in four additional two-way coupled test runs as shown in Fig. S5. Hence, given the considerable overhead computational costs and complexities of two-way coupled runs without significant impacts on the results, we conducted one-way coupled simulations similar to most previous studies. First, the performance of the default PBL, microphysics, and cumulus

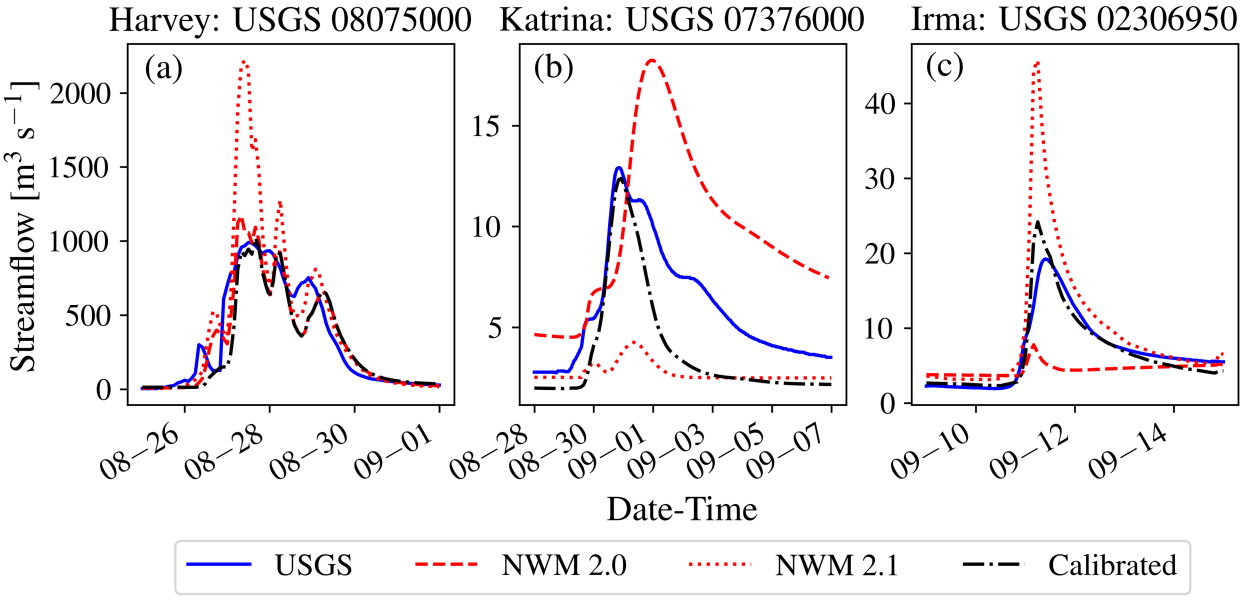

FIG. 3. Streamflow comparison between our calibrated model, NWM 2.0, NWM 2.1, and USGS gauge observation for (a) Hurricane Harvey, (b) Hurricane Katrina, and (c) Hurricane Irma.

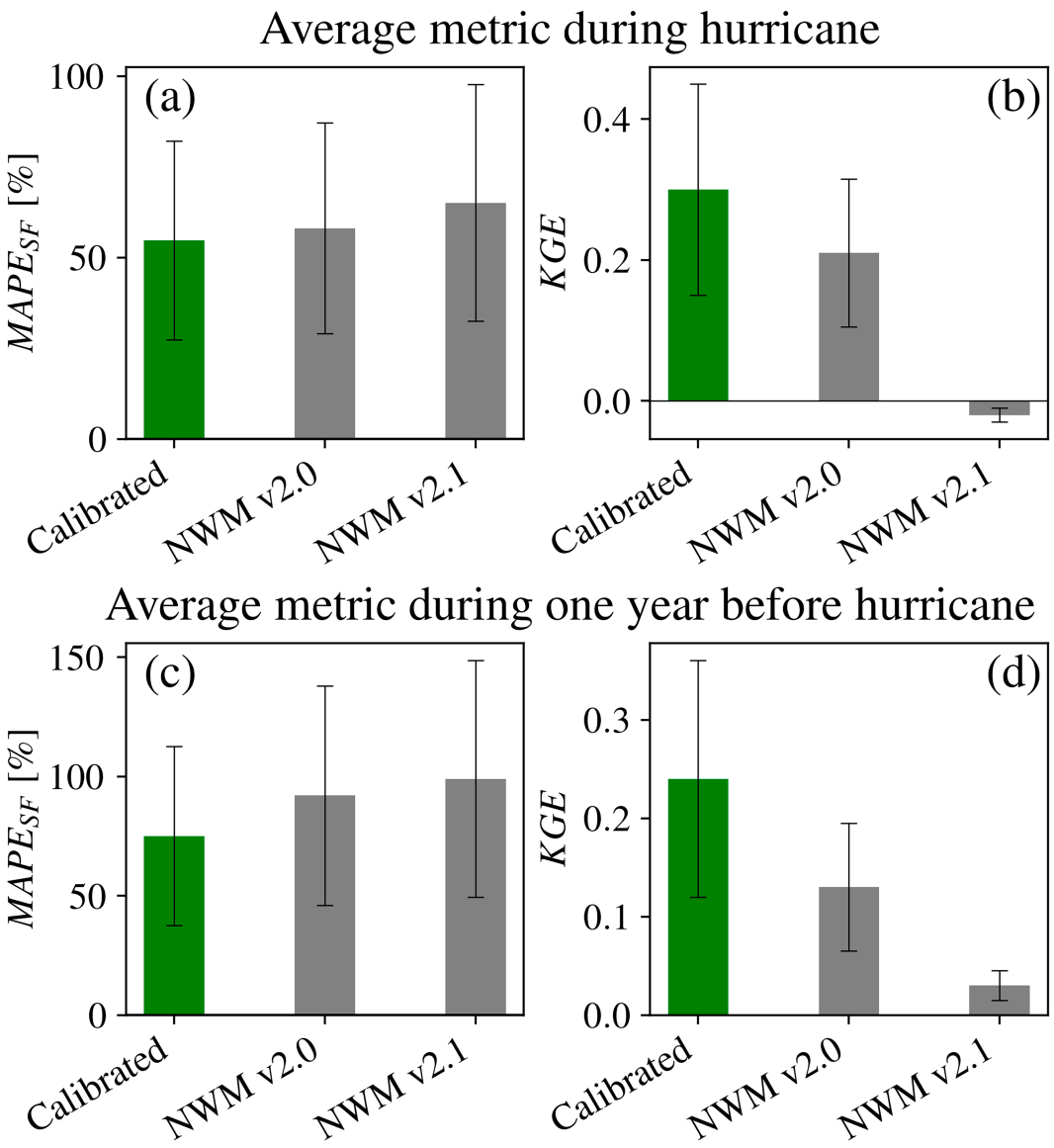

FIG. 4. Error bar plots showing the average calibrated model forecasting skills: (left) $MAPE_{SF}$ and (right) KGE during (a),(b) the hurricane period and (c),(d) one normal year before the hurricane. The calibration metrics are colored green and compared to the default NWM 2.0 and 2.1 results (gray).

schemes on hurricane intensity, track, and streamflow forecasts is determined. In the next section, the effects of adjusting the vertical diffusion on hurricane intensity and precipitation will be examined by conducting new runs.

Changing these schemes significantly impacts the simulated hurricane track. Figure 5 presents the hurricane track results of simulations with the base and CU_GF configurations. As the figure indicates, altering the cumulus scheme can improve or worsen the simulated track. For instance, the simulated Hurricane Frances track considerably improves by changing the cumulus scheme from KF [Fig. 5(a4)] to GF [Fig. 5(b4)]. Hence, these hydrometeorological schemes can distinctively modulate a hurricane's forecasted trajectory, which can significantly impact hurricane landfall location and thus hurricane-induced flood forecasts. Of note, while we refer to these WRF simulations as forecasts, they are not real forecasts (essentially hindcasts) since we do not use the actual operational forecast products as the forcing dataset in WRF, such as the outputs of the Global Forecast System (GFS; National Centers for Environmental Prediction et al. 2015). We used the NCEP FNL global analysis operational dataset to conduct these runs, which has more observational assimilation (~10%) compared to the GFS forecasts (Li et al. 2024), providing a more precise and realistic understanding of hurricane dynamics. Note that the term forecast is often used in previous studies when using these data to conduct retrospective

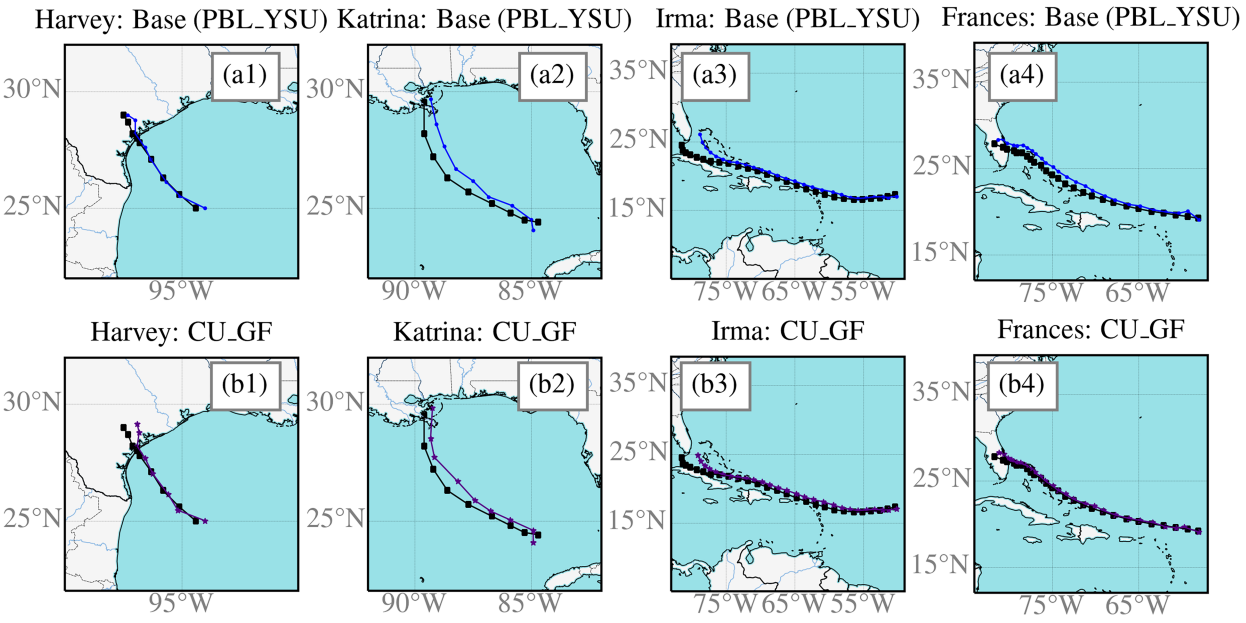

FIG. 5. The simulated hurricane tracks vs the best-observed track (black line) for all hurricanes with (a) Base (PBL_YSU) and (b) CU_GF schemes in WRF-ARW.

WRF simulations (Li and Chen 2022; Tien et al. 2013; Zhang and Pu 2017; Nolan et al. 2021a; Li and Pu 2021).

Altering the WRF-ARW's PBL, microphysics, and cumulus schemes will impact not only the hurricanes' track but also the intensity of the simulated hurricanes. Both can affect the location and intensity of hurricane-induced precipitation forcing and thus streamflow forecasts of the C-AWH run. To comprehensively analyze the impacts of these schemes on the accuracy of hurricane-induced floods, the overall statistics of hurricane intensity, track, and streamflow forecasts are calculated for the four considered hurricanes, and the average performance metrics are summarized in Fig. 6.

Among all the considered cases, CU_GF showed the best hurricane track performance (Fig. 6b). In terms of intensity predictions, the YSU scheme outperformed the MYJ, MP_Thom was better than MP_WDM, and CU_GF had superior performance compared to CU_MSKF and NO_CU schemes on average in the four considered cases (Fig. 6a). In general, the MP_Thom made the best hurricane intensity predictions in most cases, with an ~8% relative average improvement compared to the Base case. The hurricane intensity and track error metrics of each hurricane case, which are used to calculate the averages in Fig. 6, are presented in Table S3.

The WRF-ARW precipitation forcing outputs are then used to conduct streamflow predictions with WRF-Hydro. Figure 6 indicates that the Base (PBL_YSU), MP_Thom, and CU_GF cases outperform their corresponding PBL, microphysics, and cumulus parameterizations based on average KGE and $MAPE_{SF}$ in the considered cases. Table S4 presents the streamflow statistics of all the default C-AWH simulations. Based on these results, we chose the best PBL, microphysics, and cumulus schemes that had the best performance first in intensity (Fig. 6a) and then in streamflow error $MAPE_{SF}$ (Fig. 6c) metrics, which were YSU (Base), MP_Thom, and CU_GF, respectively. These cases will then be modified in the new runs of the next section to further improve their forecasts. The rationale for choosing the best intensity predictions is that our newly introduced adjustment (vertical diffusion) will significantly impact hurricane intensity forecasts. Hence, by choosing the best intensity cases, we would assess our new modifications in the best existing parameterizations in WRF-ARW and seek to further enhance their forecasts by introducing our new adjustments. We note these schemes performed best for the four considered hurricanes, and the results may not be general for all tropical cyclone simulations. Furthermore, the PBL_YSU (Base) case is the WRF's default configuration, which uses the recommended PBL and microphysics schemes for hurricane simulations (Wang et al. 2018). In this study, we will focus on improving PBL parameterizations and their impacts on streamflow forecasts since it is established in prior studies that diffusion in existing PBL schemes is overly dissipative for rotating hurricane flows (Momen et al. 2021; Romdhani et al. 2022). We will follow the new approach introduced in Matak and Momen (2023) and characterize its impacts on hurricane precipitation and streamflow forecasts in coupled hydrometeorological simulations.

c. *The impacts of the new vertical diffusion adjustments on simulated hurricane intensity*

In this section, we present the impacts of adjusting the default vertical diffusion parameterization in WRF-ARW on hurricane wind intensity forecasts. The newly applied diffusion adjustments significantly improved the hurricane intensity predictions compared to the default WRF-ARW runs. Figure 7 shows the hurricane intensity time series for the Base and CU_GF cases for the default and reduced diffusion

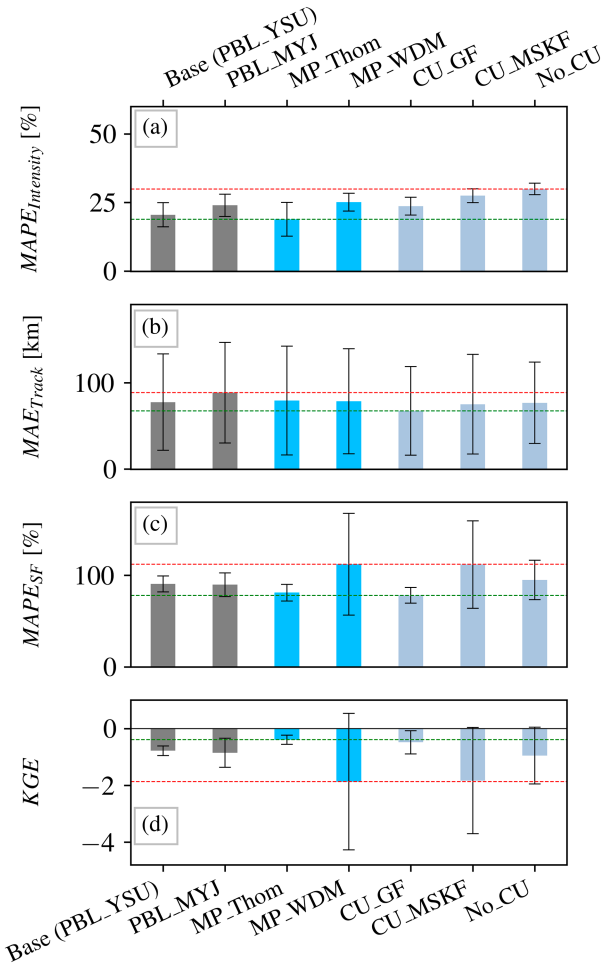

FIG. 6. The error performance metrics for all default C-AWH simulations with different PBL, MP, and CU parameterizations, averaged over four considered hurricanes. The error bar denotes the 10th–90th percentile of performance metrics for each case. The red and green dashed lines represent the worst and best values of the corresponding metrics, respectively.

configurations. The figure clearly shows that reducing the default vertical diffusion (green lines in Fig. 7) remarkably improved the hurricane wind intensity forecasts in all the considered cases compared to the default runs (gray lines in Fig. 7). For all considered hurricanes, this vertical diffusion adjustment intensifies the simulated hurricanes and hence can capture the best-observed maximum wind speed data (black lines in Fig. 7) better than the default schemes.

Overall, this reduced diffusion adjustment improved the hurricane intensity forecasts by reducing $MAPE_{Intensity}$ by ~21%–37% on average for all hurricanes. This result is consistent with previous studies which have demonstrated using other PBL codes that most of the current PBL schemes are overly diffusive for simulating major hurricanes (Gopalakrishnan et al. 2013; Zhang and Pu 2017; Tang et al. 2018; Zhang et al. 2015). The reason for this overestimation is that current turbulence models typically do not consider the turbulence suppression effects due to strong rotation in hurricanes. These remarkably improved intensity forecasts can highly influence hurricane-induced precipitation and thus streamflow forecasts in WRF-Hydro.

d. *The impacts of the new vertical diffusion changes on simulated hurricane vortex size*

Altering the vertical diffusion will also impact the simulated hurricane vortex size. To evaluate this effect, the 10-m wind is averaged radially and is plotted for each hurricane in Fig. 8. The figure indicates that the radius of maximum winds (RMW) decreases when the vertical diffusion is reduced (green lines) compared to the default cases (gray lines). This is consistent with prior studies that showed a similar effect that when the storms intensified their vortex size reduced (Gopalakrishnan et al. 2013; Zhang and Pu 2017; Tang et al. 2018; Zhang et al. 2015). To evaluate the accuracy of these radial wind profiles, we compared them with the International Best Track Archive for Climate Stewardship (IBTrACS) data (Knapp et al. 2010) shown as black dots in Figs. 8a–d. The average RMW from reduced diffusion, default, and observation-based data is shown in Fig. 8e. The figure shows that the default cases significantly overestimate the size of the RMW (~57% relative error), while the reduced diffusion cases remarkably improve the vortex size forecasts (~27% relative error). Moreover, as Fig. 8f indicates the average maximum wind intensity of the reduced diffusion cases has a much lower error (~4%) than the default cases (~33%) when compared to observations.

e. *Storm intensity improvement and its relationship with the precipitation*

Previous studies have indicated that hurricane intensity and precipitation amount are related (Kanada et al. 2012). This is significant for the streamflow response as it is shown to rely more on the precipitation input during extreme hurricane events than during regular seasons (Abbaszadeh et al. 2020). Here, we characterize the relationship between simulated precipitation and hurricane intensity in default and reduced diffusion PBL schemes. Figure 9 shows the simulated intensity and precipitation distribution for Hurricane Katrina at 1200 UTC 29 August 2005. The darker red colors of Fig. 9b compared to the default run of Fig. 9a indicate that reducing the vertical diffusion intensifies the hurricane winds. Furthermore, it influences the size of hurricanes by reducing the hurricane vortex, consistent with previous studies (Matak and Momen 2023). Increasing the intensity of the hurricane has substantial impacts on the precipitation amount and distribution. Figure 9d shows that when the hurricane intensifies, the maximum rate of precipitation increases (cf. darker colors of Fig. 9d with Fig. 9c), and the areal extent of precipitation decreases. A similar wind intensity–precipitation relationship is also apparent in Hurricane Harvey which is shown in Fig. S6. We also compared the generated wind and precipitation fields with available satellite data from Hurricane Katrina and Harvey. It is found that the intensity and size of the reduced diffusion cases agree more with satellite observations (see Figs. S7 and S8). In the next section,

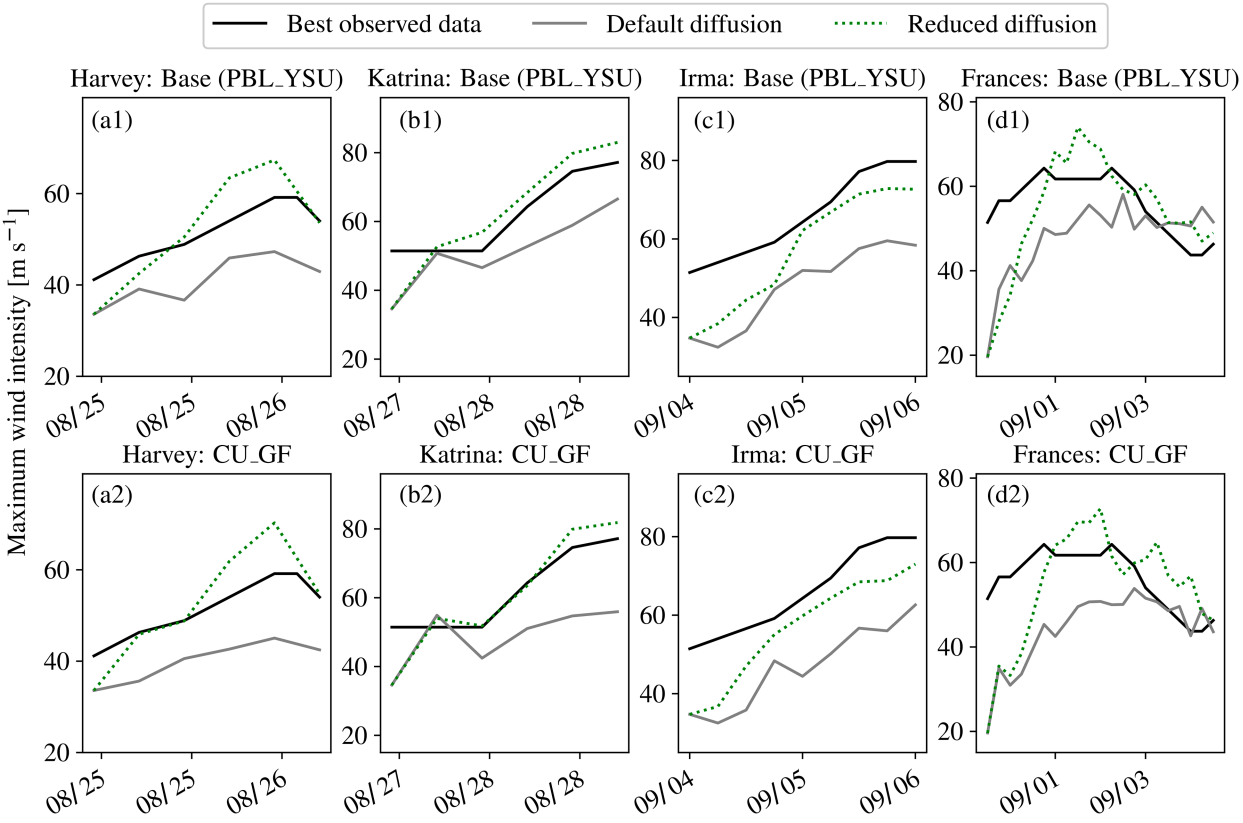

FIG. 7. The simulated hurricane wind intensity (maximum 10-m wind speed) time series vs the best-observed wind intensity for all hurricanes with (top) Base (PBL_YSU) and (bottom) CU_GF WRF-ARW configurations. The black solid line, gray solid line, and green dotted line denote hurricane intensity from observation, default diffusion, and RD schemes, respectively.

the accuracy of these simulations will be extensively evaluated using observations.

This hurricane intensity–precipitation relation can be quantitatively characterized. Figure 10 depicts their correlation by presenting the maximum hourly intensity for all four simulated hurricanes against maximum hourly precipitation. The positive intensity–precipitation relationship is apparent in this figure as the higher precipitation occurs at greater hurricane intensity. These results are also consistent with Cerveny and Newman (2000), who showed that the average rainfall amount rises with each category of hurricane wind speed. Furthermore, the maximum intensity and precipitation amounts for reduced diffusion schemes are greater than the default diffusion cases (cf. green and gray dots in Fig. 10).

To further examine the relationship between the maximum precipitation and intensity, a linear regression analysis is conducted between these two variables. Table 4 denotes the linear regression parameters, including the Pearson correlation coefficient R, slope, p value, and standard error. The average R value of this linear regression for all the cases is 0.79 with a low p value, which corroborates a high correlation between these two variables. Furthermore, this table shows that as the vertical diffusion is reduced, the maximum simulated intensity and precipitation increase consistently in the considered hurricane cases. We also examined this relationship separately for the inner and outer rainbands. We found that this linear relationship holds for the inner rainband (within 3 × RMW, see Fig. S9); however, it does not necessarily hold for the outer rainbands (not shown).

To evaluate the impacts of the reduced diffusion adjustments on precipitation forcing, the average total rainfall for all hurricanes is depicted in Fig. 11. The total precipitation is calculated from the sixth hour of the simulation to the dissipation of each hurricane. Figure 11a shows the average total precipitation inside the 75-km distance with the hurricane eye, which is ~2 RMW in different cases and represents the inner rainband (Li et al. 2017; Li and Wang 2012; Wang 2009; Gao et al. 2020). This figure indicates that the total precipitation in the inner rainband of the hurricanes is increased for reduced diffusion cases compared to the default WRF runs. However, in the outer rainband, the total precipitation of the default cases is higher than in the reduced diffusion cases (Fig. 11b). This leads to a decrease in the total precipitation of the reduced diffusion cases (Fig. 11c), which we will show agrees more with observations. The reason for this total decrease is related to the decreased hurricane size for the reduced diffusion cases (Figs. 8 and 9), which leads to a lower overall generated precipitation.

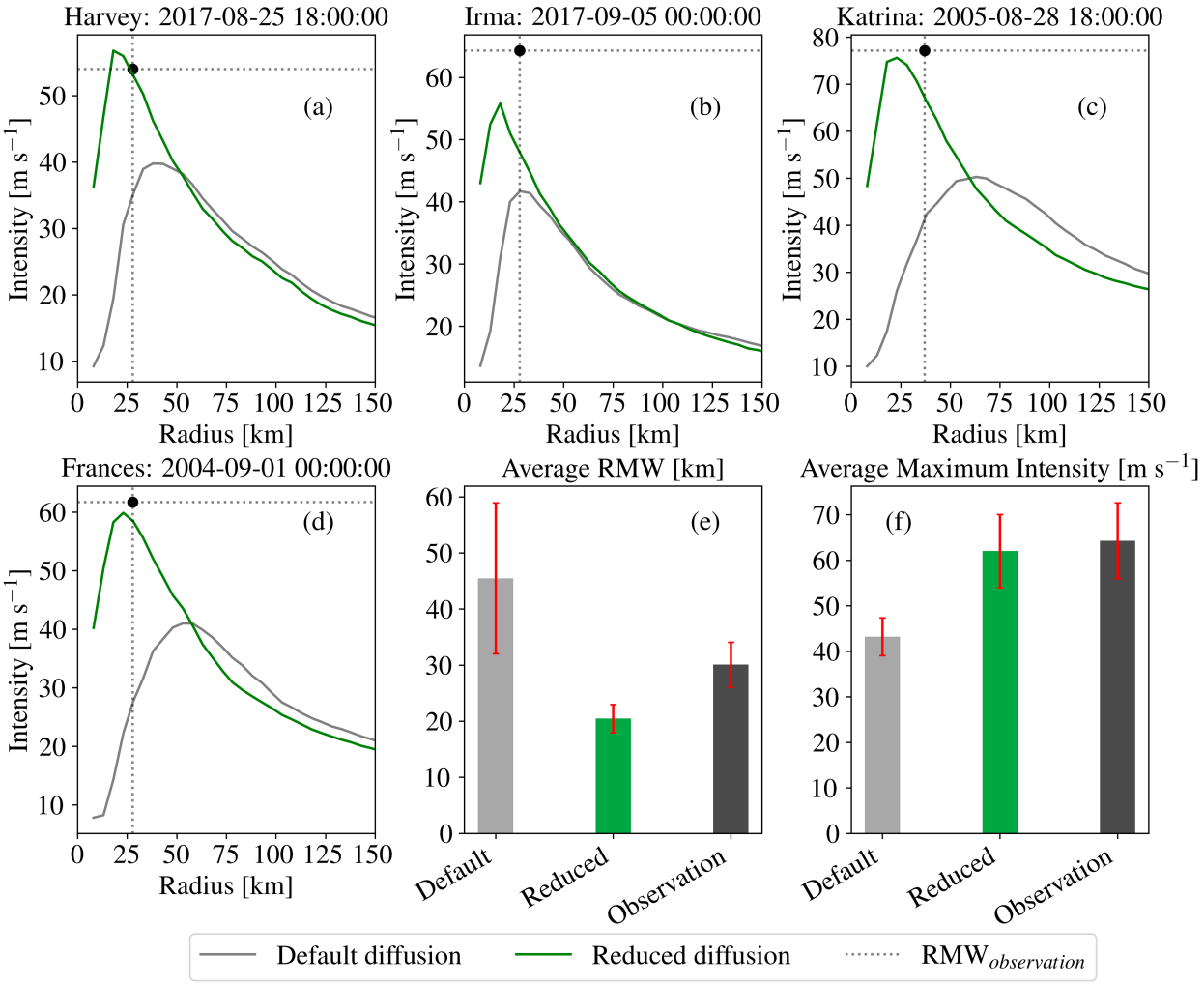

FIG. 8. (a)–(d) The radial profiles of the average surface wind speed for all hurricanes, (e) average RMW, and (f) average surface maximum wind speed compared to observation-based data (IBTrACS).

f. The impacts of changing the vertical diffusion on accumulated precipitation

To characterize the effects of the new reduced diffusion runs on the magnitude and spatial distribution of precipitation, we compared our simulation's precipitation forecasts with the Stage IV analysis of observational data. To this end, the average total precipitations of the considered cases over land are calculated for both methods versus the Stage IV data in Fig. 12a. In this figure, we included all three best considered PBL, microphysics (MP), and cumulus (CU) schemes for Hurricanes Harvey, Katrina, and Frances to obtain an ensemble average of the accumulated precipitation during 1 day over land. Hurricane Irma was not included in this figure since the default diffusion cases do not properly make landfall and have significant track errors (Table S5), which can make this analysis more prone to track errors rather than size and intensity impacts (a separate analysis of Irma will be shown after this figure). The figure indicates that the default cases overestimate the mean total precipitation over land compared to the Stage IV data likely due to their overestimated vortex size (Fig. 8e). On the other hand, our reduced diffusion case (green bar) agrees more with the observed data compared to the default case. To quantify the improvements, $MAPE_{Precipitation}$ is introduced to calculate the average of total precipitation error from Fig. 12a (see Table S6). Figure 12b indicates that the reduced diffusion improves the total precipitation forecasts of the default diffusion cases by ~29% on average. Finally, to evaluate the spatial distribution of the hurricane precipitation we calculated the grid-wise MAE between the generated precipitation forcing maps of the simulations and Stage IV observational data. This metric is named MAE_{Grid} and is shown in Fig. 12c. The figure shows that the reduced diffusion cases also improve the spatial distribution of the generated precipitation forcing by ~5% on average compared to the default cases. Therefore, the generated precipitations from the reduced diffusion cases agree more with the Stage IV observational data compared to the default cases.

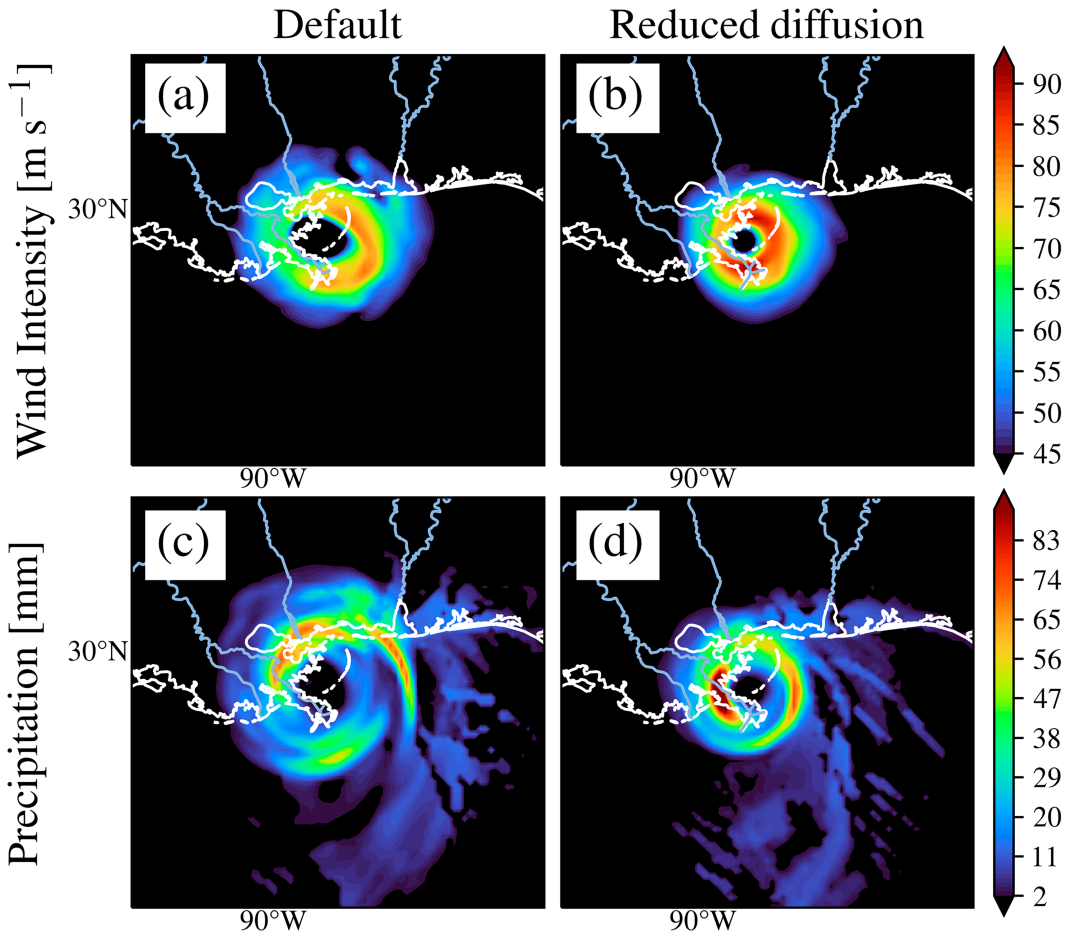

FIG. 9. (a),(b) Wind intensity at 500-m altitude and (c),(d) induced precipitation for Hurricane Katrina at 1200 UTC 29 Aug 2005 with (left) default and (right) RD WRF-ARW configuration.

The precipitation forecast accuracy and the predicted streamflow response for all hurricanes are nonlinearly dependent on hurricane track and intensity. Changing the vertical diffusion influences both the intensity and track of the simulated hurricanes. Hence, the reduced vertical diffusion parameterization nonlinearly influences precipitation data due to the complex interplay of hurricane track and intensity in rainfall location and peaks. For instance, Fig. 13 compares the hourly accumulated precipitation amount simulated for Hurricane Irma (Base) using the default and reduced diffusion schemes with the Stage IV analysis dataset (Du 2023). This figure shows that reducing the vertical diffusion considerably improved Hurricane Irma's accumulated precipitation extent (Fig. 13d) when compared to the Stage IV dataset (Fig. 13b). The reason that the default PBL scheme run (Fig. 13c) does not predict the precipitation extent well is mainly due to its inaccurate track prediction [Fig. 5(a3)], which does not make landfall correctly perhaps because its intensity was underestimated and affected its trajectory. On the other hand, the predicted track of Hurricane Irma in the reduced diffusion case outperformed the default base case by ~22.8% (see Table S5), leading to improved precipitation extent. This is also evident from one USGS gauge data shown in Fig. 13a in which the default diffusion case (gray) completely misses the streamflow peak induced by Hurricane Irma, while the reduced diffusion case (green) better agrees with the USGS data (blue).

Hurricane intensity enhancements can also influence the spatial precipitation extent by altering the maximum precipitation rate and size as described in the previous section. For instance, while the track is not significantly improved in some cases (e.g., CU_GF_RD of Frances), the streamflow prediction of their reduced diffusion case outperforms the default WRF-ARW case (e.g., ~11% improvement in $MAPE_{SF}$ for Frances). This improvement in the streamflow forecasts is likely due to the improvement in the hurricane intensity (Fig. 7) of the reduced diffusion case that outperforms the default case and captures the intensification of this hurricane before landfall (see Table S5 for more details). Another similar example is shown for the Hurricane Harvey CU_GF case in Fig. S10. While the track improvement is minor in this case, the intensity forecast improved by ~57%, which is likely the reason behind the improved streamflow forecasts (Table S7). Therefore, both the intensity

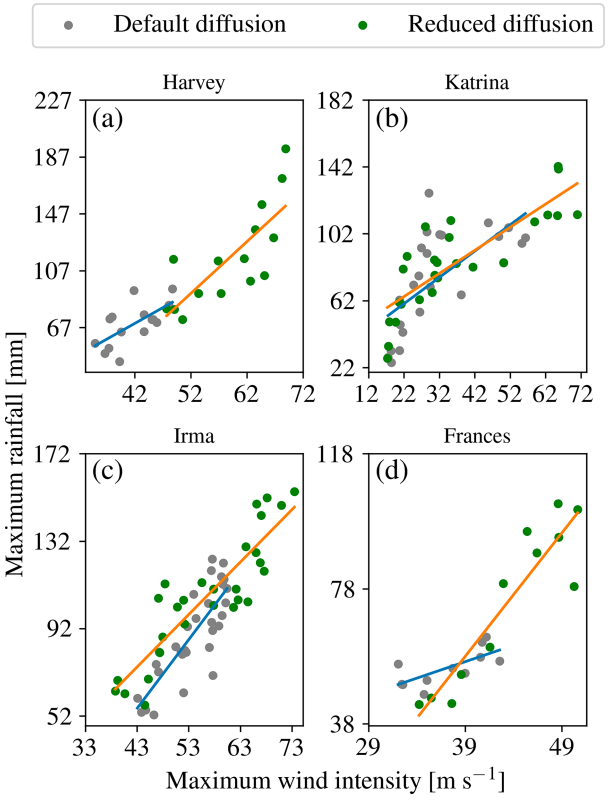

FIG. 10. Scatterplots of maximum hourly intensity and rainfall of the Base case (PBL_YSU) within 500 km of the hurricane eye for each case. The gray and green dots with blue and orange best-fitting straight lines represent the default and RD simulation, respectively.

and track of the simulated hurricanes can interplay in forecasting hurricane precipitation. Furthermore, since the new reduced diffusion cases alter both the intensity and track of hurricanes, the impacts of the new diffusion adjustments on the accumulated precipitation data are nonlinear and result from complex interactions of both parameters. Thus, to further elucidate the accuracy of the precipitation forcing of the newly adjusted reduced diffusion cases compared to the default cases, an extensive statistical analysis of the streamflow prediction errors is conducted in the next section.

g. The impacts of altering the vertical diffusion on the accuracy of hurricane flood forecasts

To comprehensively determine the effects of vertical diffusion on precipitation forcing, we calculated the statistics of all three considered cases. Figure 14 shows the average performance metrics of all the default and reduced diffusion cases. Figures 14(a1)–(a3) indicate that reducing the default vertical diffusion in different configurations consistently improves the intensity of the simulated hurricanes as expected (cf. green and gray bars in Fig. 14). The reduced diffusion cases have improved $\text{MAPE}_{\text{Intensity}}$ on average by ~39.5% compared to their default diffusion schemes. A summary of the obtained improvements $[(\text{Error}_{\text{Default}} - \text{Error}_{\text{RD}}) \times 100\%/\text{Error}_{\text{Default}}]$ in each configuration is listed in Table 5. The maximum observed intensity improvement was for Hurricane Katrina with ~51% improvement, and the lowest improvement was for Hurricane Frances by ~18%. Note that these considerably improved intensity forecasts are obtained for the best-performing default PBL, cumulus, and microphysics schemes (cf. gray and brown bars in Fig. 14).

In terms of hurricane trajectory forecasts, the overall track error $\text{MAE}_{\text{Track}}$ of the reduced diffusion cases slightly improved their corresponding default cases by ~5.2% on average for all the considered cases. However, the track forecasts do not consistently improve in all tested hurricanes and configurations. The reason for this is that improving hurricane track forecasts is generally a more challenging task since it is not only influenced by the hurricane vortex but also influenced by the environmental wind field (Velden et al. 1992) and global scale weather processes (Fierro et al. 2009), which are provided as the boundary condition data of our runs.

These improved hurricane simulations are then coupled with WRF-Hydro to evaluate the performance of streamflow forecasts compared to the USGS gauges. Figures 14(c1)–(c3) indicate that reducing the vertical diffusion (green bars) consistently improves the average MAPE_{SF} compared to the default schemes (gray bars) in the three considered configurations. Table 5 shows MAPE_{SF} was improved by ~15.8% on average in all the considered cases. The lowest average MAPE_{SF} improvement was for Hurricane Katrina by ~10%, and the highest observed improvement was for Hurricane Irma by ~30% compared to the default diffusion cases. The detailed

TABLE 4. The linear regression analysis parameters for the default and reduced vertical diffusion cases of the hurricane simulations, evaluating the relationship between maximum wind intensity and precipitation.

Hurricane	Diffusion	Max intensity (m s^{-1})	Max precipitation (mm)	Slope	R	p value	Standard error
Harvey	Default	48.7	94.2	2.22	0.66	7.9×10^{-3}	0.71
	Reduced	68.8	192.9	3.65	0.78	5.8×10^{-4}	0.81
Katrina	Default	56.3	126.4	1.57	0.66	3.8×10^{-4}	0.37
	Reduced	70.9	142.5	1.38	0.83	1.5×10^{-7}	0.19
Irma	Default	60.4	123.9	3.15	0.83	1.3×10^{-8}	0.40
	Reduced	73.5	154.7	2.40	0.89	2.4×10^{-10}	0.24
Frances	Default	42.5	63.8	0.97	0.72	1.3×10^{-2}	0.31
	Reduced	50.6	103.3	3.65	0.92	2.4×10^{-5}	0.50

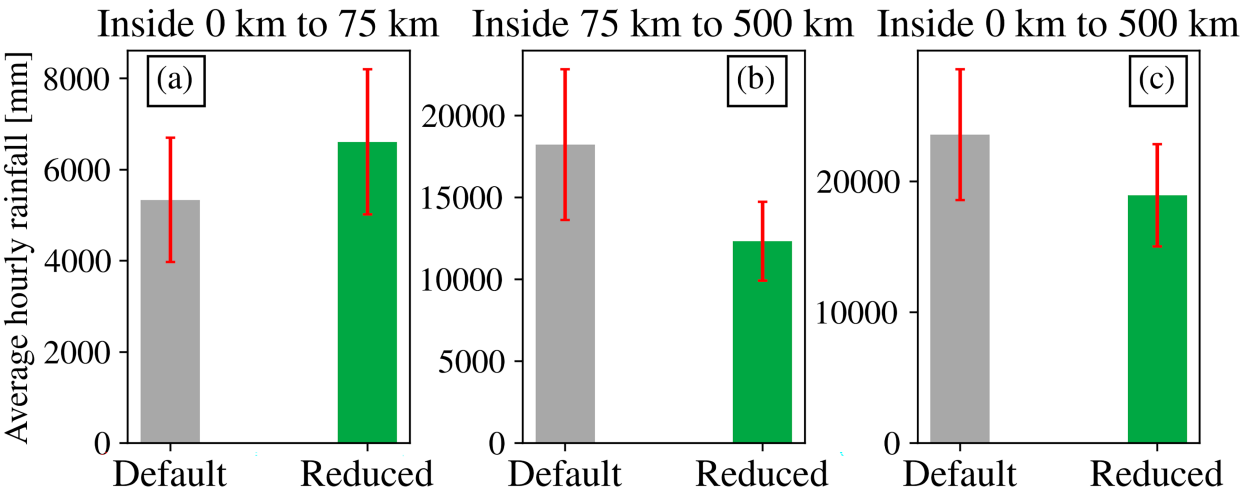

FIG. 11. The total hourly rainfall (mm) averaged for all hurricanes inside (a) 0–75, (b) 75–500, and (c) 0–500 km. The 75-km radius from the hurricane eye is ≈1.5–2.5 × RMW for the default and RD cases.

streamflow statistics for all the considered cases are presented in Table S7.

To quantify the temporal prediction of the streamflow forecasts in addition to their bias error ($MAPE_{SF}$), the KGE and correlation coefficient metrics of the simulations are calculated. It is found that the reduced diffusion cases improve the KGE even more than the bias by an average of ~67.4% (Table 5) compared to the default case. They consistently increase KGE and even make it positive for some cases [cf. green and gray bars in Fig. 14(d3)]. The KGE was increased on average between 0.06 and 0.82 in different simulated cases (Table S7). It has been shown that the KGE value above −0.4 means that the model improves the mean flow benchmark (Knoben et al. 2019). Reducing the vertical diffusion increases the KGE values above this threshold in most of our considered cases. Finally, the correlation coefficient of the forecasted streamflow was improved on average when reducing the default vertical diffusion [cf. green and gray bars in Figs. 14(e1)–(e3)]. Table 5 shows that the reduced diffusion cases increased the correlation of streamflow forecasts with the USGS data by ~34% on average compared to the default cases.

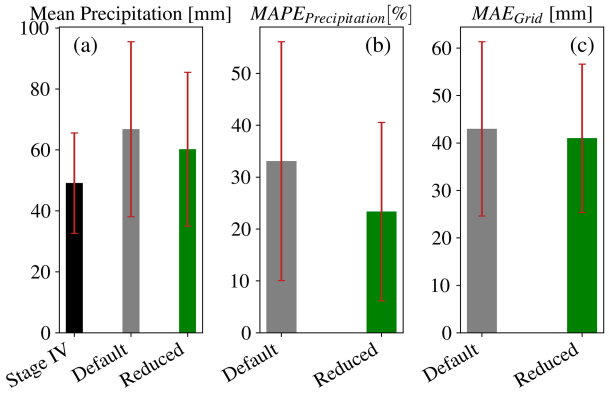

FIG. 12. Evaluation of the default and RD cases using stage IV observation-based dataset: (a) mean precipitation, (b) average $MAPE_{Precipitation}$, and (c) MAE_{Grid} to compare their mean, absolute bias, and spatial distribution performance with Stage IV analysis datasets.

The obtained results are general and grid resolution independent. To corroborate the generality of the findings, we performed a grid sensitivity test by varying the resolution of WRF-ARW from 2 to 32 km (see Fig. S11). Both hurricane intensity error and its simulated WRF-Hydro streamflow improved with increasing grid resolution as expected. The average $MAPE_{Intensity}$ and $MAPE_{SF}$ reduced by ~7%–20% and ~17%–30%, respectively, when increasing the grid resolution from 32 to 8 km and from 8 to 2 km. Furthermore, for all the considered grid resolutions, the reduced diffusion model gives a more accurate prediction in terms of both hurricane intensity and flood forecasts compared to the default diffusion cases (see Text S8). Finally, we also investigated the forecasting performance of this reduced diffusion model with the actual GFS forecast forcing dataset rather than historical analysis data. To this end, we performed eight additional simulations using four hurricanes and PBL_YSU (base) and PBL_YSU_RD cases. We observed a similar trend across all hydrometeorological evaluation metrics for the considered hurricane cases (see Fig. S12). The reduced diffusion model outperformed the default WRF model's average $MAPE_{Intensity}$ and KGE by ~31% and ~53%, respectively, confirming the generality and applicability of our findings for operational frameworks.

Our results indicate that inaccurate predictions of hurricane intensity will not only impact meteorological predictions but can also significantly influence streamflow flood forecasts. Previous studies have shown that current turbulence parameterizations in NWPs typically underestimate the intensity of major hurricanes (Matak and Momen 2023; Romdhani et al. 2022; Gopalakrishnan et al. 2021) since they do not account for turbulence suppression effects due to strong rotation in hurricanes. In this study, we adjusted the vertical diffusion in

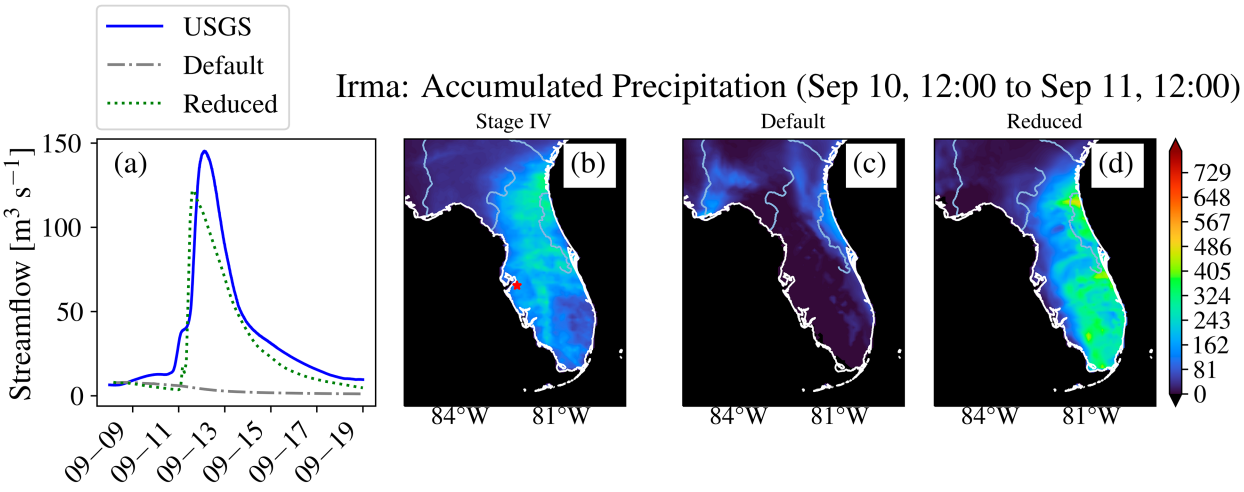

FIG. 13. Accumulated precipitation (mm) for Hurricane Irma from WRF-ARW simulations. (a) The streamflow prediction of default and RD compared with USGS 02300500 gauge observation [star in (b)] is presented for Hurricane Irma. (c) Default and (d) RD WRF-ARW simulations for Base cases are compared with (b) Stage IV observation-based data.

WRF-ARW following Matak and Momen (2023) to improve hurricane intensity predictions and showed how this change can enhance streamflow forecasts. The findings indicate the significance of PBL parameterizations on hurricane-induced flood forecasts and motivate the development of more accurate generalized schemes, which are specifically designed and evaluated for extreme weather events such as hurricanes. This will be particularly important since floods are one of the major damages of hurricanes and improved streamflow forecasts will be critical for precautionary actions.

4. Conclusions

Different NWP parameterizations can uniquely modulate and influence hurricane-induced flood forecasts. In this paper, we comprehensively characterized the impacts of various existing and adjusted PBL, microphysics, and cumulus schemes on hurricane intensity, track, and streamflow predictions. This was achieved by conducting 60 coupled ARW-WRF-Hydro simulations for four category 4–5 hurricanes (Harvey, Katrina, Irma, and Frances) in three U.S. coastal regions (Texas, Louisiana, and Florida). Before running the coupled simulations, 201 stand-alone WRF-Hydro simulations were initially conducted to comprehensively analyze and calibrate various hydrological model parameters using a stepwise approach. Next, the first suite of coupled simulations was performed by running 28 cases to evaluate WRF's seven different default schemes. Then, a second suite of new simulations was conducted in which the default vertical diffusion was reduced to improve the intensity of hurricane forecasts. To this end, the best-performing schemes were selected to run 12 more new reduced diffusion cases. Finally, a grid sensitivity test was conducted to corroborate the generality of the findings (20 runs). In summary, the key findings of this study are as follows:

1) The hydrological model parameters are typically calibrated for normal hydrometeorological events. Hence, the validity of these models needs to be revisited for extreme events such as hurricanes. Here, we examined the sensitivity of 13 parameters and were able to improve the NWM by calibrating the surface runoff parameter (refkdt) and vegetation parameter (mp) in the three considered regions. Our calibrated WRF-Hydro model outperformed NWM 2.0 (NWM 2.1) in KGE and $MAPE_{SF}$ during the hurricane period by ~38% (>100%) and ~6% (~16%) and in 1 year before the hurricane by ~81% (>100%) and ~19% (~24%), respectively.
2) The WRF-ARW atmospheric model was then coupled with the calibrated WRF-Hydro model to simulate the impacts of various WRF's PBL, cumulus, and microphysics parameterizations on hurricane-induced streamflow forecasts. The results of a comprehensive sensitivity analysis of seven considered configurations indicated that the YSU PBL, the Thompson microphysics, and the GF cumulus schemes had a superior performance in terms of hurricane intensity and streamflow forecasts for 8-km grid resolution simulations.
3) All the default WRF-ARW models' schemes underestimated the best-observed intensities of the considered major hurricanes. Hence, new adjustments were applied to the vertical diffusion parameterization in WRF-ARW by reducing the effective depth of the vertical eddy diffusivity in the best-performing parameterizations. The reduced vertical diffusion cases significantly improved the hurricane intensity forecasts ($MAPE_{Intensity}$) in all cases by ~39.5% on average. A decrease in diffusion causes a decrease in angular momentum dissipation in the PBL and leads to an increase in the hurricane's inflow, spinup, moisture convergence, and convection in the eyewall. Hence, by effectively reducing the friction, this modification is able to intensify hurricanes and enhance their forecasts. Furthermore, while hurricane track predictions are more challenging due to their dependence on the

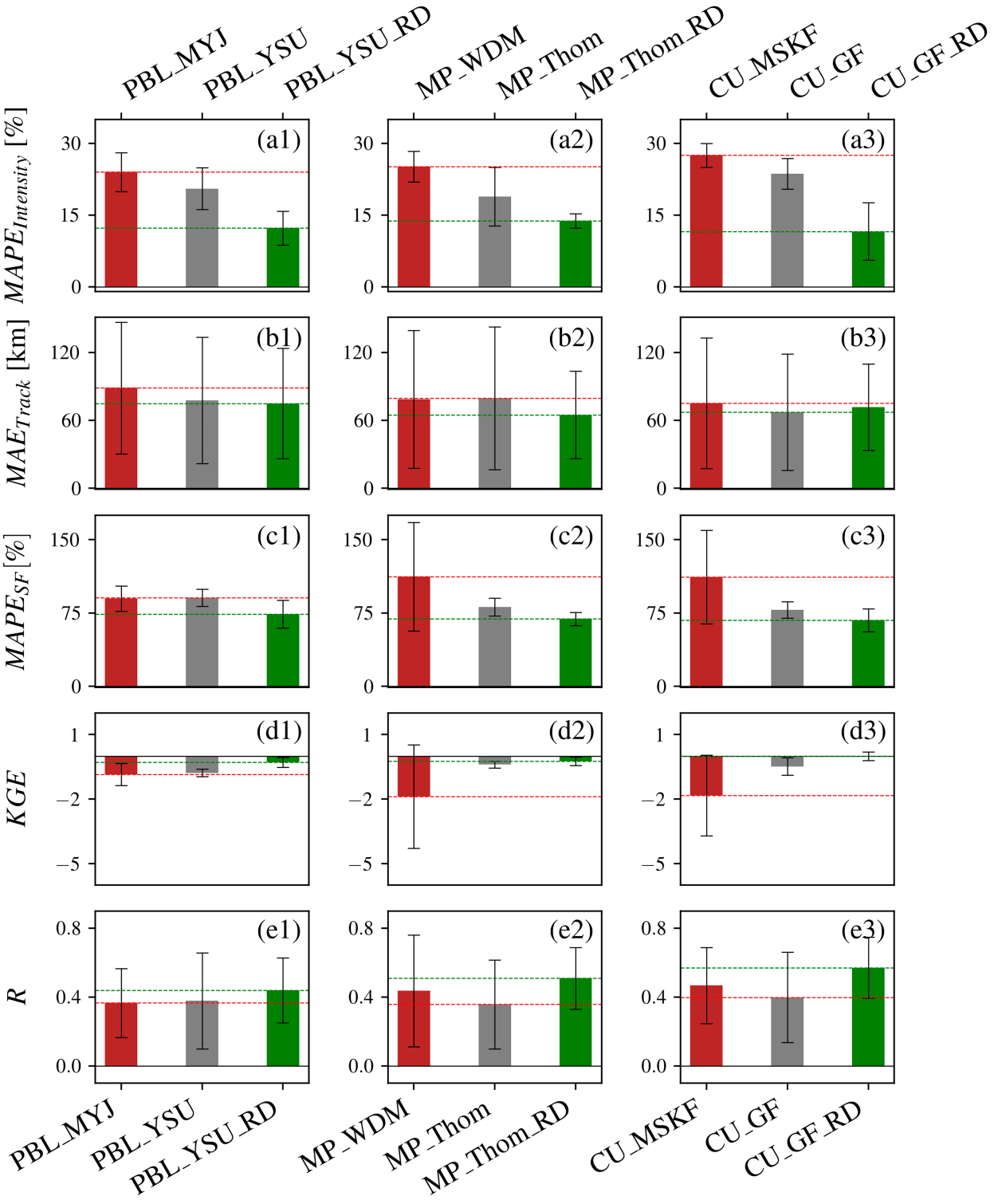

FIG. 14. Bar plots showing the average enhanced model performance skills: (a) $MAPE_{Intensity}$, (b) MAE_{Track}, (c) $MAPE_{SF}$, (d) KGE, and (e) Pearson's correlation coefficient R for different (left) PBL, (middle) MP, and (right) CU parameterizations. The gray bar represents the default diffusion, and the green bar shows the performance of the RD parameterization.

TABLE 5. Summary of hydrometeorological improvements of the RD hurricane simulations.

Metric/RD improvement	PBL_YSU (%)	MP_Thomp (%)	CU_GF (%)	Average of all 12 cases (%)
$MAPE_{Intensity}$	40.23	27.12	51.03	39.5
$MAPE_{SF}$	18.68	14.96	13.66	15.8
KGE	63.28	41.34	97.55	67.4
R	16.03	42.36	43.14	33.8

environmental global wind field, the reduced diffusion adjustment also slightly improved hurricane track forecasts by ~5.2% on average.

4) The implications of this hurricane intensity improvement on precipitation and streamflow forecasts were then characterized. It was shown that simulated hurricane intensity and maximum precipitation are correlated and that reducing the diffusion decreases the size of hurricanes and leads to more intense local precipitations. This maximum precipitation–intensity relationship is statistically quantified by a linear regression model. The average *R* value of this linear regression for all the considered cases is found to be 0.79, indicating a high correlation between these variables in major hurricanes. Compared to the Stage IV observation-based rainfall data, the reduced diffusion showed an average of ~29% improvement over default schemes for predicting the total accumulated precipitation.
5) Finally, the impacts of hurricane intensity enhancements on streamflow forecasts of WRF-Hydro were determined. A comprehensive error analysis of the conducted runs indicated that the streamflow performance metrics remarkably improved in the reduced diffusion cases by ~15.8% in $MAPE_{SF}$, ~67.4% in KGE, and ~33.8% in *R* on average. Therefore, enhancing hurricane intensity forecasts has a remarkable impact on hurricane streamflow predictions and needs to be carefully considered for accurate hurricane-induced flood forecasts.

Our results underscore the significance of turbulent diffusion parameterizations not only on hurricane intensity dynamics but also on hurricane-induced flood forecasts. These findings and insights are useful for enhancing hurricane and streamflow forecasts in hydrometeorological NWP models. Accurate hurricane forecasts are essential for proper evacuation orders and can reduce some damage from these extreme weather events.

Acknowledgments. The authors acknowledge financial support from the High Priority Area Research Seed Grant I0506984 of the University of Houston. Furthermore, the authors acknowledge support from the Physical and Dynamic Meteorology Program of the National Science Foundation under Grant AGS-2228299. Most of the simulations were performed on the University of Houston's computing clusters (Carya and Sabine). The authors also acknowledge computing cluster support from the National Center for Atmospheric Research (NCAR) under Project Number UHOU0002 as well as Bridges-2 at Pittsburgh Supercomputing Center through allocation EES230054 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by NSF.

Data availability statement. The data for the observed best tracks are already presented in the paper. The data for Hurricane Harvey with 8-km horizontal resolution are archived using Harvard Dataverse (Khondaker and Momen 2024). Additional data will be provided upon request.

REFERENCES

Abbaszadeh, P., K. Gavahi, and H. Moradkhani, 2020: Multivariate remotely sensed and in-situ data assimilation for enhancing community WRF-Hydro model forecasting. *Adv. Water Resour.*, **145**, 103721, https://doi.org/10.1016/j.advwatres.2020.103721.

Arnault, J., and Coauthors, 2018: Precipitation sensitivity to the uncertainty of terrestrial water flow in WRF-Hydro: An ensemble analysis for central Europe. *J. Hydrometeor.*, **19**, 1007–1025, https://doi.org/10.1175/JHM-D-17-0042.1.

Arolla, S. K., and P. A. Durbin, 2014: A rotation/curvature correction for turbulence models for applied CFD. *Prog. Comput. Fluid Dyn. Int. J.*, **14**, 341–351, https://doi.org/10.1504/PCFD.2014.065472.

Bao, D., Z. G. Xue, J. C. Warner, M. Moulton, D. Yin, C. A. Hegermiller, J. B. Zambon, and R. He, 2022: A numerical investigation of hurricane Florence-induced compound flooding in the cape fear estuary using a dynamically coupled hydrological-ocean model. *J. Adv. Model. Earth Syst.*, **14**, e2022MS003131, https://doi.org/10.1029/2022MS003131.

Blake, E. S., and D. A. Zelinsky, 2018: National hurricane center tropical cyclone report, Hurricane Harvey. NOAA Tech. Rep. AL092017, 77 pp., https://www.nhc.noaa.gov/data/tcr/AL092017_Harvey.pdf.

——, C. W. Landsea, and E. J. Gibney, 2011: The deadliest, costliest, and most intense United States tropical cyclones from 1851 to 2010 (and other frequently requested hurricane facts). NOAA Tech. Memo. NWS NHC-6, 49 pp., https://www.nhc.noaa.gov/pdf/nws-nhc-6.pdf.

Cangialosi, J. P., A. S. Latto, and R. Berg, 2021: National hurricane center tropical cyclone report, Hurricane Irma. NOAA Tech. Rep. AL112017, 111 pp., https://www.nhc.noaa.gov/data/tcr/AL112017_Irma.pdf.

Cazalbou, J.-B., P. Chassaing, G. Dufour, and X. Carbonneau, 2005: Two-equation modeling of turbulent rotating flows. *Phys. Fluids*, **17**, 055110, https://doi.org/10.1063/1.1920630.

Cerbelaud, A., J. Lefèvre, P. Genthon, and C. Menkes, 2022: Assessment of the WRF-Hydro uncoupled hydro-meteorological model on flashy watersheds of the Grande Terre tropical island of New Caledonia (South-West Pacific). *J. Hydrol.*, **40**, 101003, https://doi.org/10.1016/j.ejrh.2022.101003.

Cerveny, R. S., and L. E. Newman, 2000: Climatological relationships between tropical cyclones and rainfall. *Mon. Wea. Rev.*, **128**, 3329–3336, https://doi.org/10.1175/1520-0493(2000)128<3329:CRBTCA>2.0.CO;2.

Chao, L., K. Zhang, Z.-L. Yang, J. Wang, P. Lin, J. Liang, Z. Li, and Z. Gu, 2021: Improving flood simulation capability of the WRF-Hydro-RAPID model using a multi-source precipitation merging method. *J. Hydrol.*, **592**, 125814, https://doi.org/10.1016/j.jhydrol.2020.125814.

Cheikh, M. I., and M. Momen, 2020: The interacting effects of storm surge intensification and sea-level rise on coastal resiliency: A high-resolution turbulence resolving case study. *Environ. Res. Commun.*, **2**, 115002, https://doi.org/10.1088/2515-7620/abc39e.

Chen, M., and Coauthors, 2021: A comprehensive flood inundation mapping for Hurricane Harvey using an integrated hydrological and hydraulic model. *J. Hydrometeor.*, **22**, 1713–1726, https://doi.org/10.1175/JHM-D-20-0218.1.

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

Choudhury, D., and S. Das, 2017: The sensitivity to microphysical schemes on the skill of forecasting the track and intensity of tropical cyclones using WRF-ARW model. *J. Earth Syst. Sci.*, **126**, 57, https://doi.org/10.1007/s12040-017-0830-2.

Coelho, G. d. A., C. M. Ferreira, and J. L. Kinter III, 2022: Multiscale and multi event evaluation of short-range real-time flood forecasting in large metropolitan areas. *J. Hydrol.*, **612**, 128212, https://doi.org/10.1016/j.jhydrol.2022.128212.

Davis, C., and Coauthors, 2008: Prediction of landfalling hurricanes with the advanced hurricane WRF Model. *Mon. Wea. Rev.*, **136**, 1990–2005, https://doi.org/10.1175/2007MWR2085.1.

Du, J., 2023: NCEP/EMC 4KM gridded data (GRIB) stage IV data version 1.0. UCAR/NCAR – Earth Observing Laboratory, accessed 19 November 2023, https://doi.org/10.5065/D6PG1QDD.

Durbin, P., 2011: Review: Adapting scalar turbulence closure models for rotation and curvature. *J. Fluids Eng.*, **133**, 061205, https://doi.org/10.1115/1.4004150.

El Gharamti, M., J. L. McCreight, S. J. Noh, T. J. Hoar, A. Rafieeinasab, and B. K. Johnson, 2021: Ensemble streamflow data assimilation using WRF-Hydro and DART: Novel localization and inflation techniques applied to hurricane Florence flooding. *Hydrol. Earth Syst. Sci.*, **25**, 5315–5336, https://doi.org/10.5194/hess-25-5315-2021.

Emanuel, K., 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, **436**, 686–688, https://doi.org/10.1038/nature03906.

——, 2017: Will global warming make hurricane forecasting more difficult? *Bull. Amer. Meteor. Soc.*, **98**, 495–501, https://doi.org/10.1175/BAMS-D-16-0134.1.

Fierro, A. O., R. F. Rogers, F. D. Marks, and D. S. Nolan, 2009: The impact of horizontal grid spacing on the microphysical and kinematic structures of strong tropical cyclones simulated with the WRF-ARW model. *Mon. Wea. Rev.*, **137**, 3717–3743, https://doi.org/10.1175/2009MWR2946.1.

Gao, Q., Q. Li, and Y. Dai, 2020: Characteristics of the outer rainband stratiform sector in numerically simulated tropical cyclones: Lower-layer shear versus upper-layer shear. *Adv. Atmos. Sci.*, **37**, 399–419, https://doi.org/10.1007/s00376-020-9202-y.

Givati, A., D. Gochis, T. Rummler, and H. Kunstmann, 2016: Comparing one-way and two-way coupled hydrometeorological forecasting systems for flood forecasting in the Mediterranean region. *Hydrology*, **3**, 19, https://doi.org/10.3390/hydrology3020019.

Gochis, D. J., and Coauthors, 2020: The WRF-hydro modeling system technical description, version 5.2.0. NCAR Tech. Note, 108 pp., https://ral.ucar.edu/sites/default/files/public/projects/wrf-hydro/technical-description-user-guide/wrf-hydrov5.2technicaldescription.pdf.

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

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

Grell, G. A., and S. R. Freitas, 2014: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmos. Chem. Phys.*, **14**, 5233–5250, https://doi.org/10.5194/acp-14-5233-2014.

Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez, 2009: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.*, **377**, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003.

Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *Asia-Pac. J. Atmos. Sci.*, **42**, 129–151.

——, Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, https://doi.org/10.1175/MWR3199.1.

Hu, X.-M., J. W. Nielsen-Gammon, and F. Zhang, 2010: Evaluation of three planetary boundary layer schemes in the WRF model. *J. Appl. Meteor. Climatol.*, **49**, 1831–1844, https://doi.org/10.1175/2010JAMC2432.1.

Huffman, G. J., D. T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, P. Xie, and S.-H. Yoo, 2015: NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). Algorithm Theoretical Basis Doc., version 4, 26 pp, https://gpm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V4.5.pdf.

Jackson, E. K., W. Roberts, B. Nelsen, G. P. Williams, E. J. Nelson, and D. P. Ames, 2019: Introductory overview: Error metrics for hydrologic modelling – A review of common practices and an open source library to facilitate use and adoption. *Environ. Modell. Software*, **119**, 32–48, https://doi.org/10.1016/j.envsoft.2019.05.001.

Jafarzadegan, K., A. Alipour, K. Gavahi, H. Moftakhari, and H. Moradkhani, 2021: Toward improved river boundary conditioning for simulation of extreme floods. *Adv. Water Resour.*, **158**, 104059, https://doi.org/10.1016/j.advwatres.2021.104059.

Janjić, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927–945, https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2.

Jeworrek, J., G. West, and R. Stull, 2019: Evaluation of cumulus and microphysics parameterizations in WRF across the convective gray zone. *Wea. Forecasting*, **34**, 1097–1115, https://doi.org/10.1175/WAF-D-18-0178.1.

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.*, **43**, 170–181, https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

Kanada, S., A. Wada, M. Nakano, and T. Kato, 2012: Effect of planetary boundary layer schemes on the development of intense tropical cyclones using a cloud-resolving model. *J. Geophys. Res.*, **117**, D03107, https://doi.org/10.1029/2011JD016582.

Katsafados, P., G. Varlas, A. Papadopoulos, C. Spyrou, and G. Korres, 2018: Assessing the implicit rain impact on sea state during Hurricane Sandy (2012). *Geophys. Res. Lett.*, **45**, 12 015–12 022, https://doi.org/10.1029/2018GL078673.

Khondaker, M. H., and M. Momen, 2024: Improving hurricane intensity and streamflow forecasts in coupled hydro-meteorological simulations by analyzing precipitation and boundary layer schemes, Dataset, version 1. Harvard Dataverse, accessed 29 July 2024, https://doi.org/10.7910/DVN/PVZ9C7.

Kidder, S. Q., M. D. Goldberg, R. M. Zehr, M. DeMaria, J. F. W. Purdom, C. S. Velden, N. C. Grody, and S. J. Kusselson, 2000: Satellite analysis of tropical cyclones using the Advanced Microwave Sounding Unit (AMSU). *Bull. Amer. Meteor. Soc.*, **81**, 1241–1260, https://doi.org/10.1175/1520-0477(2000)081<1241:SAOTCU>2.3.CO;2.

Kilicarslan, B. M., I. Yucel, H. Pilatin, E. Duzenli, and M. T. Yilmaz, 2021: Improving WRF-Hydro runoff simulations of heavy floods through the sea surface temperature fields with higher spatio-temporal resolution. *Hydrol. Processes*, **35**, e14338, https://doi.org/10.1002/hyp.14338.

Kim, K. Y., W.-Y. Wu, E. Kutanoglu, J. J. Hasenbein, and Z.-L. Yang, 2021: Hurricane scenario generation for uncertainty modeling of coastal and inland flooding. *Front. Climate*, **3**, 610680, https://doi.org/10.3389/fclim.2021.610680.

Klotz, B. W., and E. W. Uhlhorn, 2014: Improved stepped frequency microwave radiometer tropical cyclone surface winds in heavy precipitation. *J. Atmos. Oceanic Technol.*, **31**, 2392–2408, https://doi.org/10.1175/JTECH-D-14-00028.1.

Knabb, R. D., J. R. Rhome Jamie, and D. P. Brown, 2023: Tropical cyclone report Hurricane Katrina, 23–30 August 2005, 43 pp., https://www.nhc.noaa.gov/data/tcr/AL122005_Katrina.pdf.

Knapp, K. R., M. C. Kruk, D. H. Levinson, H. J. Diamond, and C. J. Neumann, 2010: The International Best Track Archive for Climate Stewardship (IBTrACS). *Bull. Amer. Meteor. Soc.*, **91**, 363–376, https://doi.org/10.1175/2009BAMS2755.1.

Knoben, W. J. M., J. E. Freer, and R. A. Woods, 2019: Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrol. Earth Syst. Sci.*, **23**, 4323–4331, https://doi.org/10.5194/hess-23-4323-2019.

Lahmers, T. M., H. Gupta, C. L. Castro, D. J. Gochis, D. Yates, A. Dugger, D. Goodrich, and P. Hazenberg, 2019: Enhancing the structure of the WRF-Hydro hydrologic model for semiarid environments. *J. Hydrometeor.*, **20**, 691–714, https://doi.org/10.1175/JHM-D-18-0064.1.

——, C. L. Castro, and P. Hazenberg, 2020: Effects of lateral flow on the convective environment in a coupled hydrometeorological modeling system in a semiarid environment. *J. Hydrometeor.*, **21**, 615–642, https://doi.org/10.1175/JHM-D-19-0100.1.

——, and Coauthors, 2021: Evaluation of NOAA national water model parameter calibration in semi-arid environments prone to channel infiltration. *J. Hydrometeor.*, **22**, 2939–2969, https://doi.org/10.1175/JHM-D-20-0198.1.

Li, D., and E. Bou-Zeid, 2014: Quality and sensitivity of high-resolution numerical simulation of urban heat islands. *Environ. Res. Lett.*, **9**, 055001, https://doi.org/10.1088/1748-9326/9/5/055001.

Li, M., J. A. Zhang, L. Matak, and M. Momen, 2023: The impacts of adjusting momentum roughness length on strong and weak hurricane forecasts: A comprehensive analysis of weather simulations and observations. *Mon. Wea. Rev.*, **151**, 1287–1302, https://doi.org/10.1175/MWR-D-22-0191.1.

Li, Q., and Y. Wang, 2012: Formation and quasi-periodic behavior of outer spiral rainbands in a numerically simulated tropical cyclone. *J. Atmos. Sci.*, **69**, 997–1020, https://doi.org/10.1175/2011JAS3690.1.

——, ——, and Y. Duan, 2017: A numerical study of outer rainband formation in a sheared tropical cyclone. *J. Atmos. Sci.*, **74**, 203–227, https://doi.org/10.1175/JAS-D-16-0123.1.

Li, S., and C. Chen, 2022: Air-sea interaction processes during hurricane Sandy: Coupled WRF-FVCOM model simulations. *Prog. Oceanogr.*, **206**, 102855, https://doi.org/10.1016/j.pocean.2022.102855.

Li, X., and Z. Pu, 2021: Vertical eddy diffusivity parameterization based on a large-eddy simulation and its impact on prediction of hurricane landfall. *Geophys. Res. Lett.*, **48**, e2020GL090703, https://doi.org/10.1029/2020GL090703.

Li, Z., J. Peng, L. Zhang, and J. Guan, 2024: Exploring the differences in kinetic energy spectra between the NCEP FNL and ERA5 datasets. *J. Atmos. Sci.*, **81**, 363–380, https://doi.org/10.1175/JAS-D-23-0043.1.

Lim, K.-S. S., and S.-Y. Hong, 2010: Development of an effective double-moment cloud microphysics scheme with prognostic Cloud Condensation Nuclei (CCN) for weather and climate models. *Mon. Wea. Rev.*, **138**, 1587–1612, https://doi.org/10.1175/2009MWR2968.1.

Lin, P., L. J. Hopper Jr., Z. L. Yang, M. Lenz, and J. W. Zeitler, 2018a: Insights into hydrometeorological factors constraining flood prediction skill during the May and October 2015 Texas hill country flood events. *J. Hydrometeor.*, **19**, 1339–1361, https://doi.org/10.1175/JHM-D-18-0038.1.

——, Z.-L. Yang, D. J. Gochis, W. Yu, D. R. Maidment, M. A. Somos-Valenzuela, and C. H. David, 2018b: Implementation of a vector-based river network routing scheme in the community WRF-Hydro modeling framework for flood discharge simulation. *Environ. Modell. Software*, **107**, 1–11, https://doi.org/10.1016/j.envsoft.2018.05.018.

Liu, M., G. A. Vecchi, J. A. Smith, and T. R. Knutson, 2019: Causes of large projected increases in hurricane precipitation rates with global warming. *npj Climate Atmos. Sci.*, **2**, 38, https://doi.org/10.1038/s41612-019-0095-3.

Ma, Y., V. Chandrasekar, H. Chen, and R. Cifelli, 2021: Quantifying the potential of AQPI gap-filling radar network for streamflow simulation through a WRF-Hydro experiment. *J. Hydrometeor.*, **22**, 1869–1882, https://doi.org/10.1175/JHM-D-20-0122.1.

Maidment, D. R., 2017: Conceptual framework for the national flood interoperability experiment. *J. Amer. Water Resour. Assoc.*, **53**, 245–257, https://doi.org/10.1111/1752-1688.12474.

Mascaro, G., A. Hussein, A. Dugger, and D. J. Gochis, 2023: Process-based calibration of WRF-Hydro in a mountainous basin in southwestern U.S. *J. Amer. Water Resour. Assoc.*, **59**, 49–70, https://doi.org/10.1111/1752-1688.13076.

Matak, L., and M. Momen, 2023: The role of vertical diffusion parameterizations in the dynamics and accuracy of simulated intensifying hurricanes. *Bound.-Layer Meteor.*, **188**, 389–418, https://doi.org/10.1007/s10546-023-00818-w.

Mei, W., and S.-P. Xie, 2016: Intensification of landfalling typhoons over the Northwest Pacific since the late 1970s. *Nat. Geosci.*, **9**, 753–757, https://doi.org/10.1038/ngeo2792.

Ming, J., J. A. Zhang, X. Li, Z. Pu, and M. Momen, 2023: Observational estimates of turbulence parameters in the atmospheric surface layer of landfalling tropical cyclones. *J. Geophys. Res. Atmos.*, **128**, e2022JD037768, https://doi.org/10.1029/2022JD037768.

Momen, M., 2022: Baroclinicity in stable atmospheric boundary layers: Characterizing turbulence structures and collapsing wind profiles via reduced models and large-eddy simulations. *Quart. J. Roy. Meteor. Soc.*, **148**, 76–96, https://doi.org/10.1002/qj.4193.

——, and E. Bou-Zeid, 2016: Large-eddy simulations and damped-oscillator models of the unsteady Ekman boundary layer. *J. Atmos. Sci.*, **73**, 25–40, https://doi.org/10.1175/JAS-D-15-0038.1.

——, and ——, 2017: Analytical reduced models for the nonstationary diabatic atmospheric boundary layer. *Bound.-Layer Meteor.*, **164**, 383–399, https://doi.org/10.1007/s10546-017-0247-0.

——, ——, M. B. Parlange, and M. Giometto, 2018: Modulation of mean wind and turbulence in the atmospheric boundary layer by baroclinicity. *J. Atmos. Sci.*, **75**, 3797–3821, https://doi.org/10.1175/JAS-D-18-0159.1.

——, M. B. Parlange, and M. G. Giometto, 2021: Scrambling and reorientation of classical atmospheric boundary layer turbulence in hurricane winds. *Geophys. Res. Lett.*, **48**, e2020GL091695, https://doi.org/10.1029/2020GL091695.

Nasrollahi, N., A. Aghakouchak, J. Li, X. Gao, K. Hsu, and S. Sorooshian, 2012: Assessing the impacts of different WRF precipitation physics in hurricane simulations. *Wea. Forecasting*, **27**, 1003–1016, https://doi.org/10.1175/WAF-D-10-05000.1.

National Centers for Environmental Prediction/National Weather Service/NOAA/U. S. Department of Commerce, 2015: NCEP GFS 0.25 degree global forecast grids historical archive, https://doi.org/10.5065/D65D8PWK.

Nielsen-Gammon, J. W., F. Zhang, A. M. Odins, and B. Myoung, 2005: Extreme rainfall in Texas: Patterns and predictability. *Phys. Geogr.*, **26**, 340–364, https://doi.org/10.2747/0272-3646.26.5.340.

NOAA, 2022: U.S. billion-dollar weather and climate disasters, https://doi.org/10.25921/stkw-7w73.

Nolan, D. S., R. Atlas, K. T. Bhatia, and L. R. Bucci, 2013: Development and validation of a hurricane nature run using the joint OSSE nature run and the WRF model. *J. Adv. Model. Earth Syst.*, **5**, 382–405, https://doi.org/10.1002/jame.20031.

——, B. D. McNoldy, and J. Yunge, 2021a: Evaluation of the surface wind field over land in WRF simulations of Hurricane Wilma (2005). Part I: Model initialization and simulation validation. *Mon. Wea. Rev.*, **149**, 679–695, https://doi.org/10.1175/MWR-D-20-0199.1.

——, ——, ——, F. J. Masters, and I. M. Giammanco, 2021b: Evaluation of the surface wind field over land in WRF simulations of Hurricane Wilma (2005). Part II: Surface winds, inflow angles, and boundary layer profiles. *Mon. Wea. Rev.*, **149**, 697–713, https://doi.org/10.1175/MWR-D-20-0201.1.

Pal, S., F. Dominguez, M. E. Dillon, J. Alvarez, C. M. Garcia, S. W. Nesbitt, and D. Gochis, 2021: Hydrometeorological observations and modeling of an extreme rainfall event using WRF and WRF-Hydro during the RELAMPAGO field campaign in Argentina. *J. Hydrometeor.*, **22**, 331–351, https://doi.org/10.1175/JHM-D-20-0133.1.

Reed, K. A., M. F. Wehner, and C. M. Zarzycki, 2022: Attribution of 2020 hurricane season extreme rainfall to human-induced climate change. *Nat. Commun.*, **13**, 1905, https://doi.org/10.1038/s41467-022-29379-1.

Romdhani, O., J. A. Zhang, and M. Momen, 2022: Characterizing the impacts of turbulence closures on real hurricane forecasts: A comprehensive joint assessment of grid resolution, horizontal turbulence models, and horizontal mixing length. *J. Adv. Model. Earth Syst.*, **14**, e2021MS002796, https://doi.org/10.1029/2021MS002796.

——, L. Matak, and M. Momen, 2024: Hurricane track trends and environmental flow patterns under surface temperature changes and roughness length variations. *Wea. Climate Extremes*, **43**, 100645, https://doi.org/10.1016/j.wace.2024.100645.

Rudisill, W., K. E. Kaiser, and A. N. Flores, 2022: Evaluating long-term one-way atmosphere-hydrology simulations and the impacts of two-way coupling in four mountain watersheds. *Hydrol. Processes*, **36**, e14578, https://doi.org/10.1002/hyp.14578.

Rummler, T., A. Wagner, J. Arnault, and H. Kunstmann, 2022: Lateral terrestrial water fluxes in the LSM of WRF-Hydro: Benefits of a 2D groundwater representation. *Hydrol. Processes*, **36**, e14510, https://doi.org/10.1002/hyp.14510.

Sabet, F., Y. R. Yi, L. Thomas, and M. Momen, 2022: Characterizing mean and turbulent structures of hurricane winds via large-eddy simulations. *CTR Proc. of the Summer Program 2022*, Stanford, CA, National Science Foundation, 311–322, https://web.stanford.edu/group/ctr/ctrsp22/v01_Sabet.pdf.

Schaake, J. C., V. I. Koren, Q.-Y. Duan, K. Mitchell, and F. Chen, 1996: Simple water balance model for estimating runoff at different spatial and temporal scales. *J. Geophys. Res.*, **101**, 7461–7475, https://doi.org/10.1029/95JD02892.

Senatore, A., G. Mendicino, D. J. Gochis, W. Yu, D. N. Yates, and H. Kunstmann, 2015: Fully coupled atmosphere-hydrology simulations for the central Mediterranean: Impact of enhanced hydrological parameterization for short and long time scales. *J. Adv. Model. Earth Syst.*, **7**, 1693–1715, https://doi.org/10.1002/2015MS000510.

——, L. Furnari, and G. Mendicino, 2020: Impact of high-resolution sea surface temperature representation on the forecast of small Mediterranean catchments' hydrological responses to heavy precipitation. *Hydrol. Earth Syst. Sci.*, **24**, 269–291, https://doi.org/10.5194/hess-24-269-2020.

Shastry, A., R. Egbert, F. Aristizabal, C. Luo, C.-W. Yu, and S. Praskievicz, 2019: Using steady-state backwater analysis to predict inundated area from national water model streamflow simulations. *J. Amer. Water Resour. Assoc.*, **55**, 940–951, https://doi.org/10.1111/1752-1688.12785.

Shearer, E. J., V. Afzali Gorooh, P. Nguyen, K.-L. Hsu, and S. Sorooshian, 2022: Unveiling four decades of intensifying precipitation from tropical cyclones using satellite measurements. *Sci. Rep.*, **12**, 13569, https://doi.org/10.1038/s41598-022-17640-y.

Skamarock, W. C., and Coauthors, 2019: A description of the Advanced Research WRF Model version 4.3. NCAR Tech. Note NCAR/TN-556+STR, 145 pp., https://doi.org/10.5065/1dfh-6p97.

Smith, A. B., 2021: 2021 U.S. billion-dollar weather and climate disasters in historical context-hazard and socioeconomic risk mapping. 34 pp., https://www.ncei.noaa.gov/monitoring-content/billions/docs/billions-risk-mapping-2021-ams-forum.pdf.

Tang, J., J. A. Zhang, S. D. Aberson, F. D. Marks, and X. Lei, 2018: Multilevel tower observations of vertical eddy diffusivity and mixing length in the tropical cyclone boundary layer

during landfalls. *J. Atmos. Sci.*, **75**, 3159–3168, https://doi.org/10.1175/JAS-D-17-0353.1.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, https://doi.org/10.1175/2008MWR2387.1.

Tien, D. D., T. Ngo-Duc, H. T. Mai, and C. Kieu, 2013: A study of the connection between tropical cyclone track and intensity errors in the WRF model. *Meteor. Atmos. Phys.*, **122**, 55–64, https://doi.org/10.1007/s00703-013-0278-0.

Tijerina-Kreuzer, D., L. Condon, K. FitzGerald, A. Dugger, M. M. O'Neill, K. Sampson, D. Gochis, and R. Maxwell, 2021: Continental hydrologic intercomparison project, phase 1: A large-scale hydrologic model comparison over the continental United States. *Water Resour. Res.*, **57**, e2020WR028931, https://doi.org/10.1029/2020WR028931.

Tritton, D. J., 1992: Stabilization and destabilization of turbulent shear flow in a rotating fluid. *J. Fluid Mech.*, **241**, 503–523, https://doi.org/10.1017/S0022112092002131.

Velden, C. S., C. M. Hayden, W. Paul Menzel, J. L. Franklin, and J. S. Lynch, 1992: The impact of satellite-derived winds on numerical hurricane track forecasting. *Wea. Forecasting*, **7**, 107–118, https://doi.org/10.1175/1520-0434(1992)007<0107:TIOSDW>2.0.CO;2.

Verri, G., N. Pinardi, D. Gochis, J. Tribbia, A. Navarra, G. Coppini, and T. Vukicevic, 2017: A meteo-hydrological modelling system for the reconstruction of river runoff: The case of the Ofanto river catchment. *Nat. Hazards Earth Syst. Sci.*, **17**, 1741–1761, https://doi.org/10.5194/nhess-17-1741-2017.

Viterbo, F., and Coauthors, 2020: A multiscale, hydrometeorological forecast evaluation of national water model forecasts of the May 2018 Ellicott City, Maryland, flood. *J. Hydrometeor.*, **21**, 475–499, https://doi.org/10.1175/JHM-D-19-0125.1.

Wang, W., and Coauthors, 2018: User's guide for the Advanced Research WRF (ARW) modeling system, version 4.2, 464 pp., http://www2.mmm.ucar.edu/wrf/users/docs/user_guide_v4/v4.2/WRFUsersGuide_v42.pdf.

Wang, Y., 2009: How do outer spiral rainbands affect tropical cyclone structure and intensity? *J. Atmos. Sci.*, **66**, 1250–1273, https://doi.org/10.1175/2008JAS2737.1.

Xia, Y., J. Sheffield, M. B. Ek, J. Dong, N. Chaney, H. Wei, J. Meng, and E. F. Wood, 2014: Evaluation of multi-model simulated soil moisture in NLDAS-2. *J. Hydrol.*, **512**, 107–125, https://doi.org/10.1016/j.jhydrol.2014.02.027.

Xue, M., J. Schleif, F. Kong, K. W. Thomas, Y. Wang, and K. Zhu, 2013: Track and intensity forecasting of hurricanes: Impact of convection-permitting resolution and global ensemble Kalman filter analysis on 2010 Atlantic season forecasts. *Wea. Forecasting*, **28**, 1366–1384, https://doi.org/10.1175/WAF-D-12-00063.1.

Yang, L., J. Smith, M. Liu, and M. L. Baeck, 2019: Extreme rainfall from Hurricane Harvey (2017): Empirical intercomparisons of WRF simulations and polarimetric radar fields. *Atmos. Res.*, **223**, 114–131, https://doi.org/10.1016/j.atmosres.2019.03.004.

Yin, D., Z. G. Xue, J. C. Warner, D. Bao, Y. Huang, and W. Yu, 2021: Hydrometeorology and hydrology of flooding in Cape Fear River basin during hurricane Florence in 2018. *J. Hydrol.*, **603**, 127139, https://doi.org/10.1016/j.jhydrol.2021.127139.

——, D. F. Muñoz, R. Bakhtyar, Z. G. Xue, H. Moftakhari, C. Ferreira, and K. Mandli, 2022a: Extreme water level simulation and component analysis in Delaware estuary during Hurricane Isabel. *J. Amer. Water Resour. Assoc.*, **58**, 19–33, https://doi.org/10.1111/1752-1688.12947.

——, Z. G. Xue, D. Bao, A. RafieeiNasab, Y. Huang, M. Morales, and J. C. Warner, 2022b: Understanding the role of initial soil moisture and precipitation magnitude in flood forecast using a hydrometeorological modelling system. *Hydrol. Processes*, **36**, e14710, https://doi.org/10.1002/hyp.14710.

Yucel, I., A. Onen, K. K. Yilmaz, and D. J. Gochis, 2015: Calibration and evaluation of a flood forecasting system: Utility of numerical weather prediction model, data assimilation and satellite-based rainfall. *J. Hydrol.*, **523**, 49–66, https://doi.org/10.1016/j.jhydrol.2015.01.042.

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

——, ——, and C. Wang, 2017: Effects of boundary layer vertical mixing on the evolution of hurricanes over land. *Mon. Wea. Rev.*, **145**, 2343–2361, https://doi.org/10.1175/MWR-D-16-0421.1.

Zhang, J., P. Lin, S. Gao, and Z. Fang, 2020: Understanding the re-infiltration process to simulating streamflow in north central Texas using the WRF-Hydro modeling system. *J. Hydrol.*, **587**, 124902, https://doi.org/10.1016/j.jhydrol.2020.124902.

Zhang, J. A., 2010: Spectral characteristics of turbulence in the hurricane boundary layer over the ocean between the outer rain bands. *Quart. J. Roy. Meteor. Soc.*, **136**, 918–926, https://doi.org/10.1002/qj.610.

——, R. F. Rogers, D. S. Nolan, and F. D. Marks Jr., 2011: 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.

——, 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.

Zheng, Y., K. Alapaty, J. A. Herwehe, A. D. Del Genio, and D. Niyogi, 2016: Improving high-resolution weather forecasts using the Weather Research and Forecasting (WRF) model with an updated Kain–Fritsch scheme. *Mon. Wea. Rev.*, **144**, 833–860, https://doi.org/10.1175/MWR-D-15-0005.1.

Zhou, X., and B. Wang, 2009: From concentric eyewall to annular hurricane: A numerical study with the cloud-resolved WRF model. *Geophys. Res. Lett.*, **36**, L03802, https://doi.org/10.1029/2008GL036854.

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