

# A Multiple-Model Convection-Permitting Ensemble Examination of the Probabilistic Prediction of Tropical Cyclones: Hurricanes Sandy (2012) and Edouard (2014)

CHRISTOPHER MELHAUSER, FUQING ZHANG, AND YONGHUI WENG

*Department of Meteorology and Center for Advanced Data Assimilation and Predictability Techniques,  
The Pennsylvania State University, University Park, Pennsylvania*

YI JIN, HAO JIN, AND QINGYUN ZHAO

*Naval Research Laboratory, Monterey, California*

(Manuscript received 25 April 2016, in final form 20 December 2016)

## ABSTRACT

This study examines a multimodel comparison of regional-scale convection-permitting ensembles including both physics and initial condition uncertainties for the probabilistic prediction of Hurricanes Sandy (2012) and Edouard (2014). The model cores examined include COAMPS-TC, HWRF, and WRF-ARW. Two stochastic physics schemes were also applied using the WRF-ARW model. Each ensemble was initialized with the same initial condition uncertainties represented by the analysis perturbations from a WRF-ARW-based real-time cycling ensemble Kalman filter. It is found that single-core ensembles were capable of producing similar ensemble statistics for track and intensity for the first 36–48 h of model integration, with biases in the ensemble mean evident at longer forecast lead times along with increased variability in spread. The ensemble spread of a multicore ensemble with members sampled from single-core ensembles was generally as large or larger than any constituent model, especially at longer lead times. Systematically varying the physics parameterizations in the WRF-ARW ensemble can alter both the forecast ensemble mean and spread to resemble the ensemble performance using a different forecast model. Compared to the control WRF-ARW experiment, the application of the stochastic kinetic energy backscattering scheme had minimal impact on the ensemble spread of track and intensity for both cases, while the use of stochastic perturbed physics tendencies increased the ensemble spread in track for Sandy and in intensity for both cases. This case study suggests that it is important to include model physics uncertainties for probabilistic TC prediction. A single-core multi-physics ensemble can capture the ensemble mean and spread forecasted by a multicore ensemble for the presented case studies.

## 1. Introduction

The substantial impacts of tropical cyclones (TCs) on property and life make improvement of TC forecasts, particularly TC track, intensity, and inland flooding, crucial for public safety. Improving TC forecasts will likely require a combination of better understanding of TC dynamics and inner-core processes along with advanced dynamical weather models that incorporate remote and in situ observations with advanced data assimilation techniques.

In the United States, there are several regional-scale high-resolution dynamical models that are configured independently to forecast TC track and intensity in real

time. Each model has its own dynamic core (hereafter model core), including, but not limited to, the prognostic variables used to represent the atmospheric state, formulations, and numerical solvers of fundamental equations that govern atmospheric behavior, map projections, and the horizontal and vertical grid structure. The parameterizations of subgrid-scale processes also vary across models. These physical parameterizations include shortwave and longwave radiation, microphysical processes, cumulus parameterization for coarse model grids, planetary boundary layer (PBL) schemes, and surface physics representing moisture and energy fluxes; each has its own assumptions and complexity. Each choice of physical parameterization can have a substantial impact on TC structure, track, and/or intensity (Sundqvist 1970; Willoughby et al. 1984; Wang and Holland 1996; Braun and Tao 2000;

---

Corresponding author e-mail: Prof. Fuqing Zhang, fzhang@psu.edu

Green and Zhang 2013, 2014; Fovell and Su 2007; Fovell et al. 2009, 2010; Bu et al. 2014; Melhauser and Zhang 2014).

The use of ensembles may partially address model errors related to deficiencies in the model core and/or physical parameterizations. Generally, a consensus or ensemble mean forecast—either with the same model core, but varying initial conditions and/or model physics, or with a combination of multiple model cores—provides lower RMS error forecasts compared to forecasts by any individual member or component model (e.g., Goerss 2000; Sampson et al. 2008), primarily by canceling random forecast errors through ensemble averaging (Toth and Kalnay 1993). The ensemble spread provides information about forecast uncertainties (e.g., Tracton and Kalnay 1993; Palmer 2002).

For any ensemble prediction system, it is important that the spread be sufficient to cover all sources of forecast error; otherwise, the ensemble is underdispersive, a common issue with current operational and research ensemble systems (e.g., Novak et al. 2008; Torn 2016). Methods for mitigating deficiencies in the ensemble spread include 1) improving the representation of the initial condition uncertainty (e.g., Zhang et al. 2006; Hohenegger and Schär 2007; Houtekamer et al. 2009), 2) improving the representation of the lateral boundary condition uncertainty of regional models (e.g., Nutter et al. 2004b,a; Torn et al. 2006; Romine et al. 2013), and 3) accounting for forecast model uncertainties by using multicore ensembles, single-core multiphysics ensembles (e.g., Leslie and Fraedrich 1990; Krishnamurti 1999; Abernson 2001; Vijaya Kumar et al. 2003; Meng and Zhang 2007, 2008a,b; Johnson and Wang 2012; Qi et al. 2014), and/or stochastic physics ensembles (e.g., Shutts 2005; Palmer et al. 2009; Romine et al. 2014; Berner et al. 2015). For example, a study by Lang et al. (2012) found that using a combination of initial condition perturbations and stochastic physics can increase the TC track spread to match the average error in the European Centre for Medium-Range Weather Forecasts ensemble. Another study by Torn (2016) found that combinations of initial condition and physics uncertainties are needed to produce an ensemble spread that is closer to the mean error in the TC intensity forecasts.

The current study examines the impact of model error on ensemble forecasts of TC track and intensity, using an identical set of initial condition perturbations with the analysis uncertainties derived from an ensemble Kalman filter (EnKF) to initialize three state-of-the-art TC-configured regional convection-permitting models. Numerous ensemble experiments

with two selected TC events seek to elucidate 1) our understanding of the practical limits of the predictability of these events given realistic initial condition and forecast model uncertainties and 2) how to best capture realistic forecast uncertainties using a limited-size, convection-permitting, regional-scale ensemble. To the best of the authors' knowledge, no previous study has tried to independently quantify the evolution of model error using multicore ensembles while controlling for initial condition uncertainties between models. In addition to using multicore ensembles, this study also compares the impacts used to represent model uncertainties, including different physical parameterizations and stochastic physics when using the same forecast model (single core). More specifically, this study seeks to evaluate the following questions: 1) How does the evolution of the ensemble mean and spread with the same initial conditions compare using different models? 2) Is a single-core multiphysics ensemble sufficient for representing model uncertainties in TC prediction, or is there a benefit to using a multicore ensemble? 3) Can a single-core single-physics ensemble with stochastic physics and/or inflated initial condition uncertainty impact the ensemble mean and spread similarly to a multicore or multiphysics ensemble?

An overview of the two TC events is provided in section 2, along with different model and physics configurations, as well as the initial and boundary condition generation methodology. Section 3 presents results and discusses the ensemble performance between different physics and initial condition uncertainties for both TC case studies. Section 4 provides concluding remarks.

## 2. Methodology

### a. Study cases

Hurricanes Sandy (2012) and Edouard (2014) are chosen for this study because of the large divergence between the deterministic forecasts generated by different operational forecasting systems. Notably, Sandy had a more difficult track forecast and Edouard had a more difficult intensity forecast.

After developing in the northwest Caribbean Sea and traversing north over Jamaica and Cuba, Sandy restrengthened into a category 1 hurricane and turned northeast while moving north of the Bahamas. It paralleled the eastern seaboard and, subsequently, took a northwest turn by 1200 UTC 29 October, briefly intensifying back into a category 2 hurricane before weakening and making landfall in New Jersey (Blake et al. 2013). The official NHC track forecast

(OFCL) error for Sandy was below the OFCL 5-yr mean track forecast error at all lead times, with the Global Forecasting System (GFS) ensemble mean performing slightly better than OFCL for the first 48 h (Blake et al. 2013). In this study, forecasts were initialized at 0000 UTC 26 October when the storm was moving toward the northwest over the Bahamas. At this initialization time, operational and experimental ensemble forecasts had both landfalling and out-to-sea trajectories (refer to Fig. 1 in Munsell and Zhang 2014).

Edouard developed into a tropical depression at 1200 UTC 11 September and was subsequently named at 0000 UTC 12 September. The storm tracked around the southwestern side of a deep-layer subtropical ridge while embedded in an environment with favorable upper-level winds and sea surface temperatures, but drier-than-normal midlevel air (Stewart 2014). The storm slowly strengthened in this environment and underwent a period of rapid strengthening from 0600 UTC 14 September to 0600 UTC 15 September; its peak intensity of 105 kt (where 1 kt =  $0.51 \text{ m s}^{-1}$ ) was reached at 1200 UTC 16 September (Stewart 2014). In this study, forecasts were initialized at 1200 UTC 11 September when the storm was a  $15 \text{ m s}^{-1}$  tropical depression. At this initialization time, there are large uncertainties in the intensity forecast generated by the experimental real-time convection-permitting ensemble.

### b. Models

This study uses three state-of-the-art nonhydrostatic TC-configured regional models: the U.S. Navy's Coupled Ocean–Atmospheric Mesoscale Prediction System for Tropical Cyclones (COAMPS-TC; Hodur 1997; Doyle et al. 2012, 2014; Jin et al. 2014),<sup>1</sup> the Hurricane Weather Research and Forecasting Model (HWRF; Tallapragada et al. 2014), and the Advanced Research version of the WRF Model (WRF-ARW; Skamarock et al. 2008). The 2014 Hurricane Forecast Improvement Project (HFIP) configuration of COAMPS-TC, the 2013 pseudo-operational configuration of HWRF, and the 2014 Pennsylvania State University WRF-ARW ensemble Kalman filter (PSU-WRF-EnKF) real-time forecast system (Zhang et al. 2009, 2011; Weng and Zhang 2012, 2016) configuration of WRF-ARW are used as the control ensembles for each individual model. The configuration for each model's control ensemble reduces the sources of uncertainty by using only the atmospheric component of

the model. All ensembles use identical prescribed sea surface temperatures from GFS forecasts and are run at the same horizontal grid resolution. All models use three domains, which are two-way nested, with vortex-following inner domains. The domain configurations can be found in Table 1 with an example domain setup for Edouard shown in Fig. 1. Both WRF-ARW and COAMPS-TC use a Mercator projection and have a fixed outer domain while HWRF uses a rotated latitude–longitude projection with a movable outer domain that is centered on the TC during model initialization.

### c. Initial and boundary conditions

This study examines the evolution of forecast errors due to initial condition and model errors within and between models; thus, it is imperative that the initial conditions be identical and the boundary conditions be as close as possible between models, given the grid and domain limitations. To accomplish this task, each forecast is cold started using the model-specific initialization procedure [e.g., real (for WRF-ARW), real\_nmm (for HWRF), or coama (for COAMPS-TC)] to interpolate identical global model output ( $U, V, T$ , relative humidity, geopotential height, surface pressure, sea level pressure, soil moisture, soil temperature, surface skin temperature, and terrain height) onto the model-specific grids to generate boundary conditions.

To generate identical ensemble initial conditions, the NCEP Global Data Assimilation System (GDAS)  $0.5^\circ \times 0.5^\circ$  latitude–longitude surface and standard pressure level analysis at a given initialization time is merged with the 60-member PSU-WRF-EnKF (Weng and Zhang 2012) real-time 9- and 3-km WRF domains on a common  $0.1^\circ \times 0.1^\circ$  latitude–longitude horizontal grid<sup>2</sup> with standard vertical pressure levels. The PSU-WRF-EnKF and GDAS output are linearly averaged from only the PSU-WRF-EnKF output within a 300-km radius of the vortex center to only the GDAS output beyond the 600-km radius,<sup>3</sup> providing a smooth transition in the far-field TC environment. The PSU-WRF-EnKF generates 60 high-resolution flow-dependent perturbed TC vortices with the GDAS output providing identical synoptic environments away from the TC vortex in each ensemble. Given the grid and domain limitations of each

<sup>1</sup> COAMPS-TC is the registered trademark of the Naval Research Laboratory.

<sup>2</sup> The  $0.1^\circ$  horizontal spacing was chosen to reduce interpolation errors, but stay within computing resource constraints.

<sup>3</sup> The PSU-WRF-EnKF only assimilates observations within 800 km of the TC center.

TABLE 1. Model domain and physical parameterization settings for pseudo-operational versions of WRF-ARW (APSU), HWRF, and COAMPS-TC (COTC).

Model	Domain	Cumulus physics	Microphysics	PBL	Radiation	Surface options
APSU	D01: 27 km (379 × 244) D02: 9 km (304 × 304) D03: 3 km (304 × 304)  Vertical levels: 43 Model top: 10 hPa	Grell–Freitas <sup>a</sup> (D01 only)	WSM6 <sup>d</sup>	Yonsei University (YSU) scheme <sup>g</sup>	Dudhia shortwave <sup>j</sup> ; RRTM longwave <sup>k</sup>	Modified MM5 similarity (WRF option 91) PSU formulation surface TC flux <sup>n</sup>  Five-layer thermal diffusion land surface
HWRF	D01: 0.18° (216 × 432) D02: 0.06° (88 × 170) D03: 0.02° (180 × 324) Vertical levels: 43 Model top: 50 hPa	New SAS <sup>b</sup> (D01 and D02)	Tropical Ferrier <sup>e</sup>	Modified GFS <sup>h</sup>	GFDL shortwave and longwave <sup>l</sup>	HWRF surface physics GFDL hurricane slab model land surface
COTC	D01: 27 km (379 × 244) D02: 9 km (304 × 304) D03: 3 km (304 × 304) Vertical levels: 40 Model top: ~12 hPa	Kain–Fritsch <sup>c</sup> (D01 only)	COAMPS version 2 with drizzle <sup>f</sup>	Mellor–Yamada 2.5 scheme <sup>i</sup> with prognostic TKE	Fu–Liou <sup>m</sup>	COAMPS surface physics <sup>o</sup>

<sup>a</sup> See Grell and Freitas (2013).

<sup>b</sup> See Han and Pan (2011).

<sup>c</sup> See Kain and Fritsch (1993).

<sup>d</sup> See Hong and Lim (2006).

<sup>e</sup> See Ferrier (1994).

<sup>f</sup> See Rutledge and Hobbs (1983).

<sup>g</sup> See Hong et al. (2006).

<sup>h</sup> See Hong and Pan (1996).

<sup>i</sup> See Mellor and Yamada (1982).

<sup>j</sup> See Dudhia (1989).

<sup>k</sup> See Mlawer et al. (1997).

<sup>l</sup> See Fels and Schwarzkopf (1981).

<sup>m</sup> See Fu and Liou (1992).

<sup>n</sup> See Green and Zhang (2013).

<sup>o</sup> See Louis (1979).

dynamic core, the operational GFS forecast from 6 to 120 h is used for each ensemble member to help control for uncertainties when defining dynamic-core-specific boundary conditions. With this configuration, the three-dimensional TC circulation and near-TC environment (<600 km from TC center) are perturbed for each ensemble member while the larger-scale synoptic environment is identical. Uncertainties in large-scale initial conditions and physics are beyond the scope of the current study.

#### *d. Single-core multiphysics and stochastic physics sensitivity experiments*

To examine the impact of using a single-core ensemble with varying physical parameterizations, the PSU real-time configuration of WRF-ARW (APSU; Table 1) was systematically modified. The cumulus parameterization (CP), microphysics scheme (MP), planetary boundary scheme (PBL), shortwave and

longwave radiation schemes (RAD), and surface flux settings (SFC) were all changed incrementally from “APSU like” (APS1) to “HWRF like” (APS5); the physical parameterization differences from APSU are listed in Table 2. As a result of computing restraints, the physics were additively modified, for example, modifying SFC settings (APS1); modifying SFC settings and MP (APS2); modifying SFC settings, MP, and RAD (APS3), etc., to systematically sample a combination of physical parameterization schemes between the APSU-like and HWRF-like configurations.

To examine the impact of stochastic physics on ensemble mean and spread, this study modifies the APSU ensemble with spatially and temporally correlated perturbations using either the stochastic kinetic energy backscatter scheme (SKEBS; Shutts 2005; Berner et al. 2009) or stochastically perturbed parameterization tendencies (SPPT; Palmer et al. 2009;

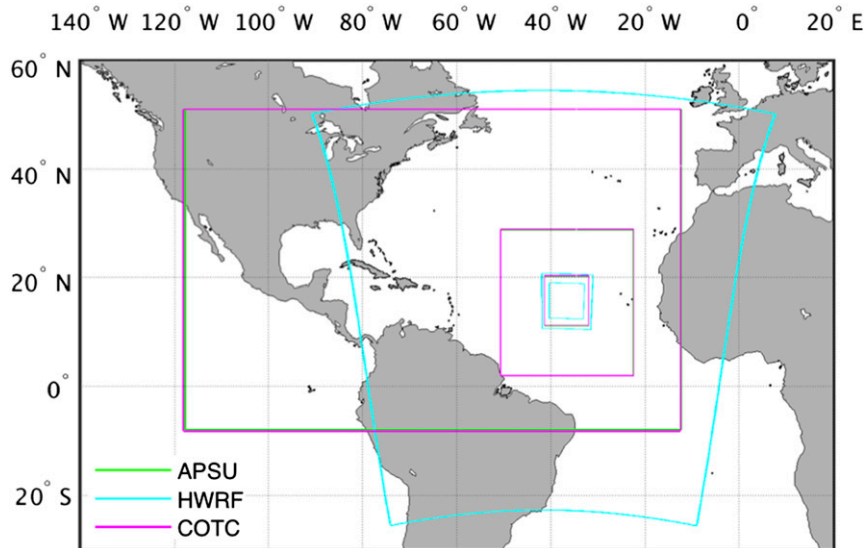


FIG. 1. Domain setup for the 1200 UTC 11 Sep 2014 initialization of Hurricane Edouard for WRF-ARW (APSU; green), HWRF (cyan), and COTC (magenta). The D01 domain (largest domain for each model) is globally fixed for APSU and COTC, but centers on the TC for HWRF at initialization and is fixed for the remainder of the forecast. Domains D02 and D03 (telescoped from D01 for each model) are two-way nested and vortex following for all models, centered on the TC at initialization. Note that the APSU and COTC domains are nearly identical and thus substantially overlap. Detailed domain settings can be found in Table 1.

Romine et al. 2014; Berner et al. 2015). SKEBS is an additive scheme, stochastically perturbing the model state, while SPPT is a multiplicative scheme, stochastically perturbing the total physical parameterization tendency.

SKEBS accounts for a model's shortcomings in reproducing the turbulent energy cascade from unresolved subgrid-scale processes. This scheme generates random temporal and spatial stochastic forcing patterns and adds the perturbations to the rotational  $u$ - and  $v$ -wind components and to potential temperature. For this study, the default configuration<sup>4</sup> in WRF-ARW v3.6 is used (Table 3) and the horizontal perturbations are assumed constant with height. The perturbations scales are in line with the results of Judt et al. (2016), who found large-scale SKEBS perturbations have the largest impact on the TC vortex and forecast uncertainty.

SPPT accounts for uncertainties in deterministic subgrid-scale physical parameterization schemes. It generates a probabilistic solution by multiplying the

subgrid-scale physical parameterization total tendency with a stochastic forcing pattern; the pattern generator is similar to SKEBS. The pertinent SPPT parameters are listed in Table 3. Like SKEBS, the horizontal perturbations are assumed constant with height.

### 3. Results and discussion

#### a. Intermodel comparison: Evolution of ensemble track and intensity

As a first step, it is important to see whether different regional TC ensembles with an identical set of initial perturbations (from the EnKF analysis uncertainty) produce similar TC forecasts. A 20-member random sample of the 5-day ensemble track and maximum 10-m wind speed (hereafter intensity) forecasts for APSU, HWRF, and COAMPS-TC (COTC) are shown in Figs. 2a and 3a for Sandy and Figs. 4a and 5a for Edouard. Each pseudo-operational model can generally capture the track and spread for both cases, but systematic errors become evident at longer lead times. Recall that each model is initialized from the same blended global and regional analysis; thus, differences in terms of both mean and spread between different ensembles can be attributed to the different physical parameterizations and/or dynamic-core configurations.

<sup>4</sup>Parameter tuning for both SKEBS and SPPT was not performed, which may lead to suboptimal results for tropical cyclones. The parameter values are similar to those of Romine et al. (2014) and Berner et al. (2015).

TABLE 2. Model physical parameterization differences from the APSU configuration for sensitivity experiments APS1–5. Dashed lines indicate the same scheme as in APSU is used.

Expt	Cumulus physics	Microphysics	PBL	Radiation	Surface options
APS1	—	—	—	—	Garratt surface flux (isftcflx = 2)
APS2	—	Eta (Ferrier) <sup>b</sup>	—	—	Garratt surface flux (isftcflx = 2)
APS3	—	Eta (Ferrier) <sup>b</sup>	—	GFDL shortwave–longwave <sup>d</sup>	Garratt surface flux (isftcflx = 2)
APS4	New SAS (D01 and D02) <sup>a</sup>	Eta (Ferrier) <sup>b</sup>	—	GFDL shortwave–longwave <sup>d</sup>	Garratt surface flux (isftcflx = 2)
APS5	New SAS (D01 and D02) <sup>a</sup>	Eta (Ferrier) <sup>b</sup>	GFS <sup>c</sup>	GFDL shortwave–longwave <sup>d</sup>	Garratt surface flux (isftcflx = 2)

<sup>a</sup> See Han and Pan (2011).

<sup>b</sup> See Rogers et al. (2001).

<sup>c</sup> See Hong and Pan (1996).

<sup>d</sup> See Fels and Schwarzkopf (1981).

The evolution of the ensemble tracks for Sandy in Fig. 2a is consistent across models during the first 60 h, but systematic divergence between the ensemble means (Fig. 2b) becomes evident at longer lead times as Sandy begins to curve back toward the northwest. The COTC ensemble members track farther east compared to the APSU members, while the HWRF ensemble members are mostly centered on the NHC best track. It is encouraging that all three models capture the complex track divergence of Sandy. The ensemble intensity forecasts are remarkably consistent between models before landfall (~96 h; Fig. 3a). Besides the low wind speeds of HWRF at initialization,<sup>5</sup> COTC generally has a lower intensity over the first 48 h compared with HWRF and APSU. At longer lead times, prior to landfall (~60–96 h), all three models capture the secondary reintensification of Sandy, evident in the ensemble means (Fig. 3b), although all of the models tend to intensify the TC 12 h too early.

The evolution of Hurricane Edouard may depend on the less predictable internal TC dynamics associated with convective processes. These highly nonlinear processes may be handled differently by different model cores and/or physical parameterizations. The APSU and COTC solutions (Fig. 5a) have a similar pattern of evolution throughout the 5-day forecast, strengthening the TC with similar mean magnitudes (Fig. 5b). The HWRF intensities have a much larger spread, with many cases not strengthening. The strongest HWRF ensemble member is weaker than the NHC best track and many of the APSU and COTC members. There appears to be a link between the intensity and track for Edouard; the weaker ensemble members track farther west (Fig. 4a), which is also

evident in the farther west track of the HWRF ensemble mean (Fig. 4b).

#### b. Intermodel comparison: Multicore ensemble

To objectively quantify the relationship between the single-core ensembles and a multicore ensemble, a 60-member ensemble (MCOR) is generated by randomly sampling 20 ensemble members without replacement from each of the APSU, HWRF, and COTC ensembles. The absolute error of the ensemble mean (hereafter error) and the ensemble spread of a multicore ensemble provide an objective measure of performance.

The track and intensity ensemble spread and error for each individual model and the MCOR ensemble are shown for Sandy in Figs. 6a,b and Edouard in Figs. 7a,b.<sup>6</sup> The NHC best track is used for verification and is assumed to be perfect. For Sandy, the intensities for all three models and MCOR are similar (Fig. 6b) with no systematic difference between MCOR and the individual model ensembles, except postlandfall. For track, MCOR generally has a lower error and larger spread than the individual model ensembles. At shorter lead times up to approximately 36 h, the ensemble spreads of APSU, HWRF, COTC, and MCOR are all similar. After this period, MCOR covers a larger solution space and has lower error as a result of the averaging of the systematically farther west (east) APSU (COTC) members. After APSU and HWRF make landfall, the inclusion of the easterly COTC tracks in MCOR shifts the ensemble mean east, increasing the error.

Examining the ensemble spread and error for Edouard, MCOR has the largest spread in intensity (Fig. 7b) after approximately 24 h compared with the

<sup>5</sup> HWRF does not use the diagnosed 10-m wind speeds from the initial conditions during cold-start initialization, unlike APSU and COTC.

<sup>6</sup> Only relative differences of the ensemble error and spread between experiments will be analyzed as a result of the small sample size of this study (Murphy 1988).

TABLE 3. Stochastic physics parameters settings.

Parameter	Value
SKEBS	
Total backscattered dissipation rate for streamfunction	$1 \times 10^{-5} \text{ m}^2 \text{ s}^{-3}$
Total backscattered dissipation rate for potential temp	$1 \times 10^{-6} \text{ m}^2 \text{ s}^{-3}$
Decorrelation time for streamfunction perturbations	6 h
Decorrelation time for temp perturbations	3 h
SPPT	
Gridpoint variance	0.1
Decorrelation time	6 h
Decorrelation length scale	150 km

individual model ensembles. The error and spread of APSU and COTC diverge from HWRF after 48 h, after which continued strengthening was observed in APSU and COTC, but not in HWRF. The MCOR ensemble performs well for the first 72 h with a smaller intensity error and larger intensity spread than any individual model ensembles. However, MCOR is systematically degraded after approximately 72 h as a result of the inclusion of HWRF.

The track ensemble spread for Edouard (Fig. 7a) is consistent for all individual model ensembles for approximately the first 48 h. By the end of the forecast, HWRF has the largest spread of the three individual model ensembles. For shorter lead times (<48 h), the MCOR ensemble has similar error and spread when compared with the individual model ensembles, but for times > 96 h it has the lowest error, and a spread similar to that of HWRF. Including the HWRF model in the MCOR ensemble substantially increases the MCOR track spread. The track biases of the individual model ensembles at longer lead times average in MCOR to reduce the track error relative to any individual component ensemble, similar to what was found with Sandy.

Overall, common between both case studies, the MCOR ensemble generally has the largest spread in both track and intensity compared with any individual model ensemble, but the error is case specific. An increased ensemble spread at all lead times is evident for the multicore ensemble.

### c. Intramodel comparison: Multiphysics ensembles

Model physical parameterization schemes can strongly influence forecasts of TC intensity (e.g., Lord et al. 1984; Wang 2002; McFarquhar et al. 2006; Zhu and Zhang 2006; Jin et al. 2007; Green and Zhang 2013, 2014) and track (e.g., Fovell et al. 2010; Bu et al. 2014). Given that the APSU, HWRF, and COTC ensembles have different

dynamic cores and unique physical parameterizations, it is worth exploring if a single-core ensemble with systematically varying physical parameterizations (limited by available physics parameterizations in WRF-ARW) can perform similarly to the MCOR ensemble. The WRF-ARW model is used to generate sensitivity experiments by varying the model physical parameterizations from “APSU like” (APS1) to “HWRF like” (APS5). The physical parameterization differences from the APSU configuration are outlined in Table 2. A single-core multiphysics 60-member ensemble (MPHY) is constructed by randomly sampling 12 members, without replacement, from APS1–5 for comparison with the MCOR ensemble.

The ensemble mean track forecasts of the APS1–5 sensitivity experiments (Fig. 2c for Sandy and Fig. 4c for Edouard, respectively) show the track sensitivity to the physical parameterizations. Focusing on Sandy (Fig. 2c), the APS3–5 ensembles have a mean eastward displacement after approximately 60 h relative to APSU, with APS5 resembling COTC rather than APSU at longer lead times. The APS1–5 ensemble track spread (Fig. 6c) has similar growth characteristics and magnitude as APSU for the first 72 h, before any substantial interaction with land. Comparing the MPHY track error for Sandy to the APSU, HWRF, and COTC ensembles (Fig. 6a), the MPHY mean follows APSU with consistent increased error at approximately 24- and 72-h lead times, but overall, the error is reduced relative to APSU as a result of the inclusion of eastward-tracking members APS4–5 in the MPHY ensemble. The spread of MPHY remains as large as or larger than APSU at longer lead times and, generally, is closer to the MCOR spread. Statistically testing<sup>7</sup> the track error distributions under the null hypothesis that MPHY and MCOR are drawn from the same underlying error distribution, MPHY and MCOR generally cannot be statistically differentiated for most lead times. The lead times when the null hypothesis can be rejected are indicated with gray asterisks (Fig. 6a).

For Edouard, both the track and intensity forecasts are sensitive to the physical parameterization configurations. The ensemble mean tracks for APS2–5 are generally farther west (Fig. 4c) at longer lead times compared to APSU. For this case, APS1–5 exhibit a nonlinear monotonic decrease in TC intensity (Fig. 5c) throughout a majority of the 5-day forecast, with a larger decrease found at longer lead times. As briefly discussed in section 3a, a track–intensity feedback may

<sup>7</sup> A 10 000-sample bootstrapped Kolmogorov–Smirnov two-sample test at the 95% significance level is performed between MPHY and MCOR at each 6-h lead time.

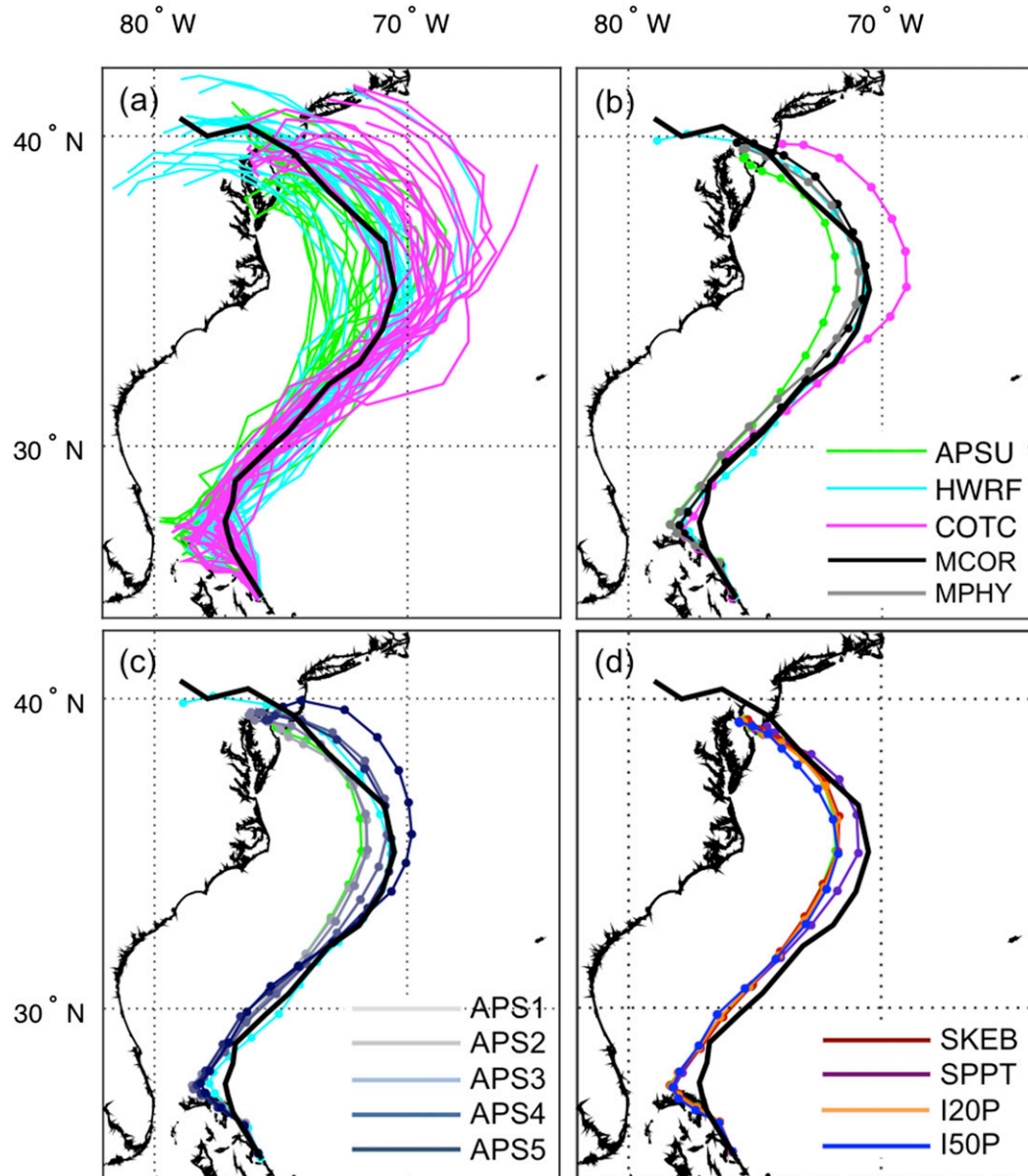


FIG. 2. Ensemble track for (a) APSU (green), HWRP (cyan), and COTC (magenta; 20 randomly selected ensemble members from each ensemble). Ensemble mean tracks for (b) APSU (green), HWRP (cyan), COTC (magenta), MPHY (gray), and MCOR (black); (c) APS1 (light gray) through APS5 (dark navy); and (d) stochastic physics experiments SKEB (maroon) and SPPT (purple), as well as inflated initial perturbation experiments I20P (yellow) and I50P (blue), for Hurricane Sandy initialized at 0000 UTC 26 Oct 2012. The NHC best-track data (black line) are plotted every 6 h. The solid dots in (b)–(d) indicate the position every 6 h.

be present, with the weaker ensemble members generally tracking farther westward. Statistically testing the track error distributions between MPHY and MCOR (gray asterisks in Fig. 7a) indicates that the differences between the error distributions at lead times  $> 80$  h are statistically significant. Focusing back on one of the component ensembles of MCOR, the COTC ensemble generally has farther eastward

tracking members at longer lead times. The physical parameterization packages used in the APS1–5 configuration are different from the COTC configuration and are not able to capture many of the COTC solutions. It is hypothesized that a well-designed multiphysics ensemble with a sampling of physical parameterizations that generate physically realistic representations of atmospheric processes will alleviate



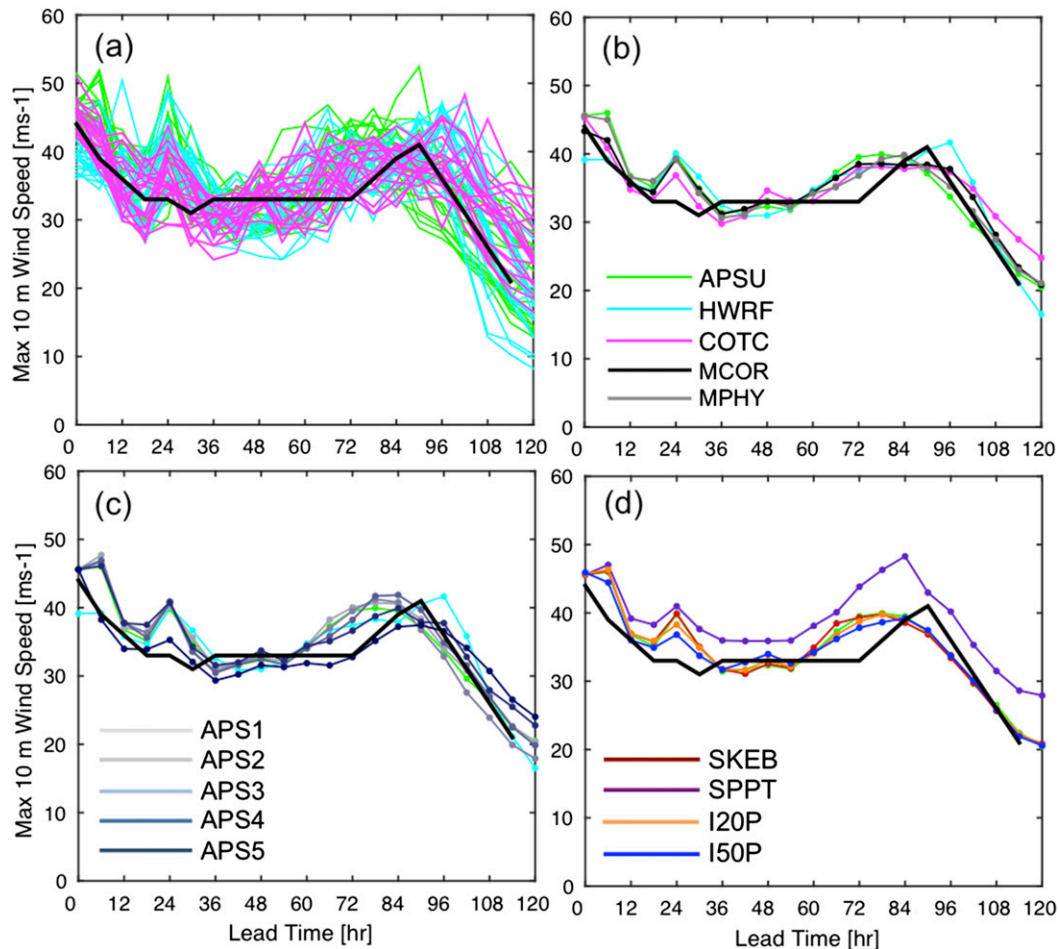


FIG. 3. As in Fig. 2, but for ensemble intensity ( $\text{m s}^{-1}$ ).

this issue.<sup>8</sup> The current MPHY is not sampling the entire solution space, only solutions similar to those of APSU and HWRF.

The intensity spread for MPHY behaves similarly to MCOR at shorter lead times, but is strongly influenced by the poorly performing members of APS4–5 at longer lead times. Generally, the single-core multiphysics ensemble error distributions of track and intensity cannot be statistically differentiated from a multicore ensemble for a majority of the lead times for both cases. It is hypothesized that the MPHY ensemble may be deficient in solution diversity compared to the MCOR ensemble as it is lacking solutions that resemble the COTC ensemble. Nevertheless, it is clear that a single-core ensemble can be configured to capture

TC uncertainties similarly to a different model-core ensemble by modifying the physical parameterizations.

*d. Intramodel comparison: Stochastic physics and inflated initial condition ensembles*

Other approaches commonly used to increase the ensemble spread include stochastic physics and modifying the initial perturbation magnitude in the initial conditions. The prior targets deficiencies in the forecast model, while the latter targets deficiencies in the initial conditions. Two different stochastic physics methodologies available in WRF-ARW—SKEBS, and SPPT, which target different model deficiencies—are used in this study to increase the spread of the APSU single-core single-physics ensemble. Additionally, a simple scalar multiplicative inflation factor of 1.2 (20%; I20P) or 1.5 (50%; I50P) is applied to the control APSU EnKF perturbations (directly impacting the transformed  $U$ ,  $V$ ,  $T$ , relative humidity, geopotential height, surface pressure, and sea level pressure on the interpolated global input grid).

<sup>8</sup>The COTC physical parameterizations have been developed independently from WRF-ARW and HWRF and have unique physics parameterizations based on different underlying assumptions (see Table 1), which are not available in WRF-ARW or HWRF.

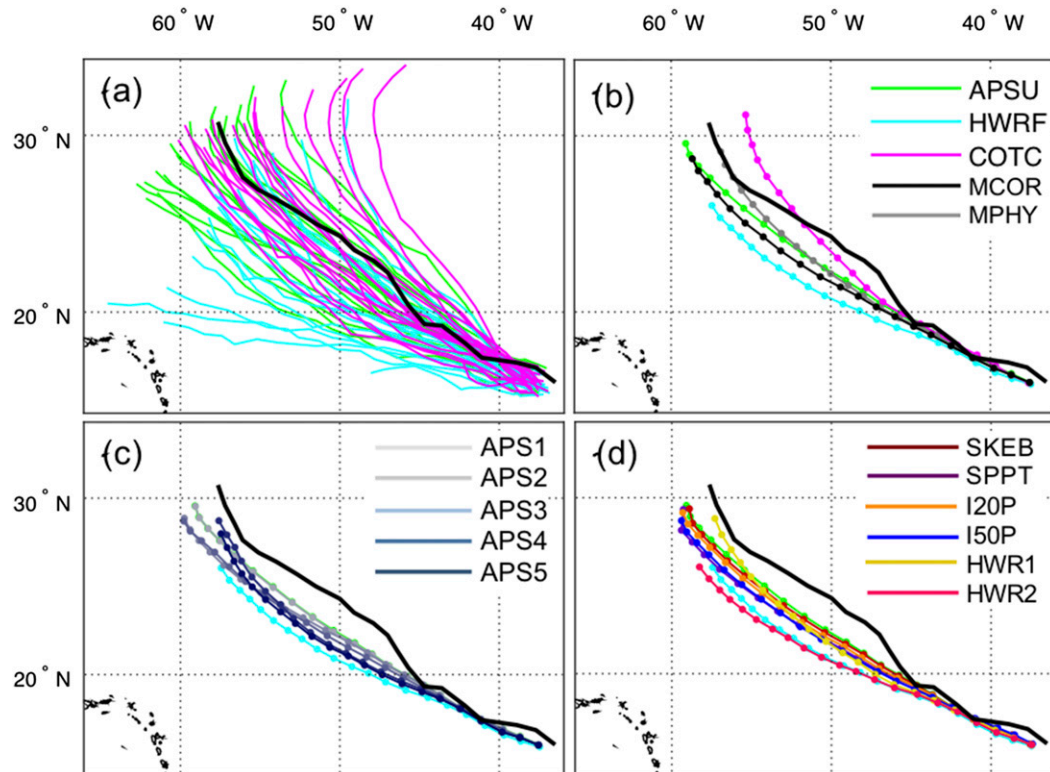


FIG. 4. As in Fig. 2, but for Hurricane Edouard (2014) initialized at 1200 UTC 11 Sep 2014 with (d) additional experiments HWR1 (yellow) and HWR2 (pink).

The SKEB, I20P, and I50P experiments yield minimal changes to the mean track (Fig. 2d) and intensity (Fig. 3d) of Sandy. The SPPT experiment shifts the ensemble mean eastward at lead times after 48 h, more similar to HWRF, and strengthens the system. Since SPPT impacts the physical parameterizations directly, areas sensitive to active physical processes will be heavily influenced. Focusing on the track ensemble spread (Fig. 6e), both stochastic physics and initial condition perturbations increase the ensemble spread compared to APSU. SKEB has minimal impact at all lead times, SPPT has the largest impact at longer lead times, and I50P increases the spread throughout the experiment. Statistically testing the track error distributions between the various experiments and MCOR (asterisks at the top of Fig. 6e) show an increased number of lead times in all experiments that can be statistically differentiated from MCOR. For the intensity (Fig. 6f), it is clear that SPPT substantially increases the spread at all lead times, but severely overestimates the intensity, resulting in larger errors. Tuning of the SPPT parameters may help reduce the intensity error, but the configurations used here are in line with previous convection-permitting ensemble studies (e.g., Romine et al. 2014; Berner et al. 2015). The I50P experiment shows a slight improvement in error as well as increased spread at a majority of lead times.

For Edouard, the inflated initial condition experiments have a larger ensemble track spread compared with the stochastic physics (Fig. 7e), with SKEB and SPPT yielding similar results. After approximately 36 h, the error in all of the sensitivity experiments increased relative to MCOR, with SKEB causing minimal degradation while SPPT and I50P substantially degraded the track forecast. All stochastic physics and inflated initial condition ensembles had lower ensemble spread compared to MCOR, missing eastward solutions generated by the COTC ensemble at longer lead times. The unique physical parameterizations within COTC represent the diversity in model physics needed to improve TC forecasts. The error distributions of track are statistically different (asterisks at the top of Fig. 7e) from MCOR, indicating these two experiments were not capable of covering the diversity of solutions found in the COTC ensemble. For intensity (Fig. 7f), I50P and SPPT had a relatively larger impact than SKEB on intensity ensemble spread and forecast error.

#### e. Initial condition versus model uncertainties: Track and intensity measures

The pairwise root-mean-squared difference (RMSD) between the same ensemble member of

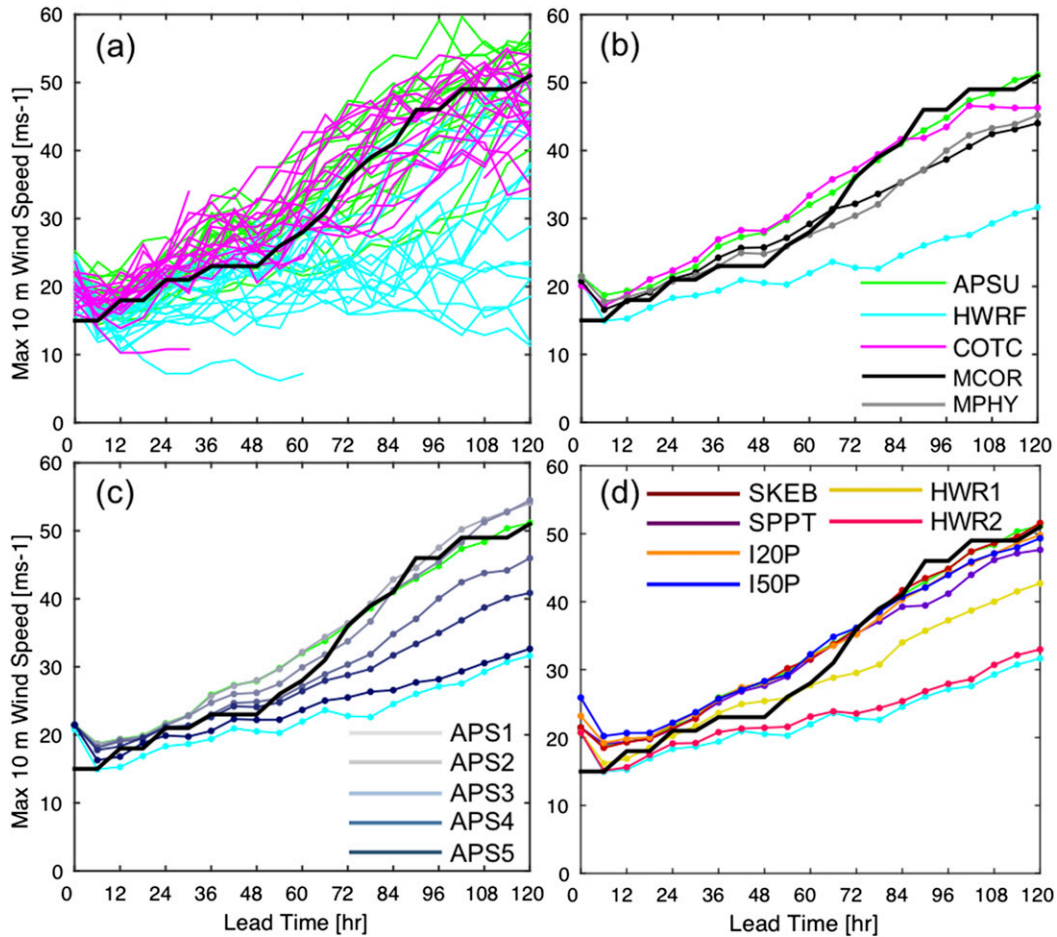


FIG. 5. As in Fig. 3, but for Hurricane Edouard (2014) initialized at 1200 UTC 11 Sep 2014 with (d) additional experiments HWR1 (yellow) and HWR2 (pink).

different model configurations (dynamic core and/or physical parameterizations) provides information about the relative error growth between different model configurations. The track RMSD for Sandy (Fig. 8a) and both track and intensity for Edouard (Figs. 8c,d) suggest that the systematic RMSD differences between APS1–5 and APSU and between APSU, HWR1, and COTC can be reproduced by using a single-core ensemble with different physical parameterizations. There is no clear systematic difference in intensity for Sandy (Fig. 8b). By modifying the physical parameterizations, the APSU ensemble can be configured to alter both the error growth rate and magnitude, comparable to the uncertainties found when using two different forecast models. There is a general monotonic increase in pairwise RMSD for both track (Fig. 8c) and intensity (Fig. 8d) for Edouard from APS1 to APS5, with the pairwise RMSD magnitude and growth rate proportional to the forecast length.

*f. Initial conditions versus model uncertainties: Domain-integrated measures*

The domain-integrated measure of the difference between ensembles can be accomplished by examining the evolution of the difference between kinetic and thermal energy per unit mass at each model grid point between experiments, called the difference total energy (DTE; Zhang et al. 2002, 2004), which is calculated as

$$DTE_{i,j,k} = \frac{1}{2} \sum_{n=1}^N (U_{i,j,k,n}'^2 + V_{i,j,k,n}'^2 + \kappa T_{i,j,k,n}'^2), \quad (1)$$

where  $U$  and  $V$  are the zonal and meridional wind components, respectively;  $T$  is the air temperature; the primes denote the difference between two experiments;  $i, j,$  and  $k$  correspond to the Cartesian model grid indices of the model domain;  $n$  is the ensemble index,  $N$  is the total number of ensemble members, and  $\kappa = C_p/T_0$  (where  $C_p = 1005 \text{ J kg}^{-2} \text{ K}^{-1}$ , the specific heat capacity of dry air at 273 K, and  $T_0 = 273 \text{ K}$  is a reference temperature used to

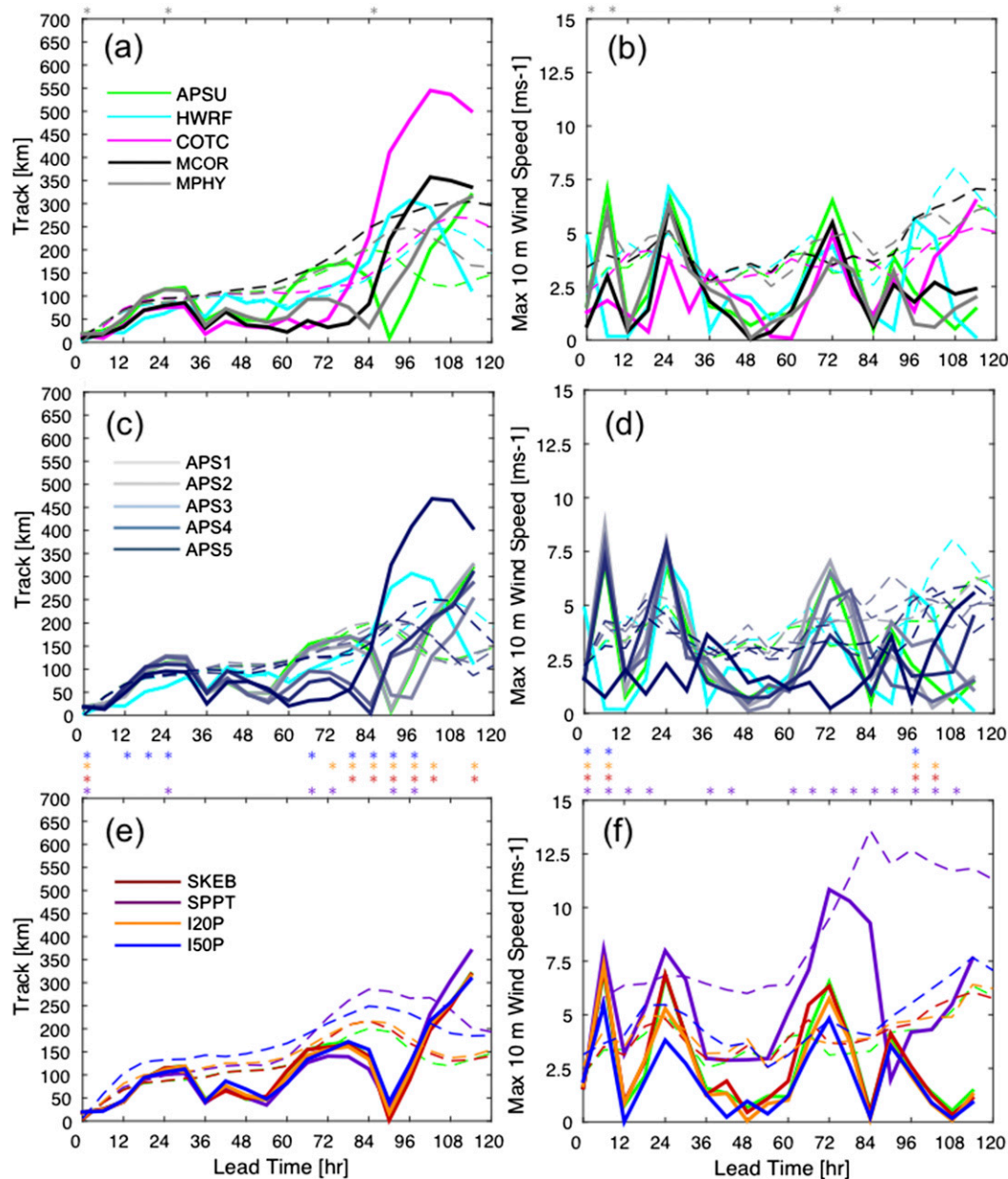


FIG. 6. RMSE verified with NHC best-track data (solid line) and the ensemble spread ( $\text{m s}^{-1}$ ; dashed line) at 6-hourly lead times for (left) track (km) and (right) intensity ( $\text{m s}^{-1}$ ) of (a),(b) APSU (green), HWRf (cyan), COTC (magenta), MPHY (gray), and MCOR (black); (c),(d) APS1 (light gray) through APS5 (dark navy); and (e),(f) stochastic physics experiments SKEB (maroon) and SPPT (purple), as well as inflated initial perturbation experiments I20P (yellow) and I50P (blue) for Hurricane Sandy initialized at 0000 UTC 26 Oct 2012. The asterisks above panels (a),(b),(e), and (f) corresponding to the aforementioned ensembles indicate the lead times when the ensemble spread is statistically different from the MCOR ensemble spread at the 95% significance level for MPHY in (a) and (b) and SKEB, SPPT, I20P, and I50P in (e) and (f).

define the reference state). The intraensemble DTE (ensemble spread) is calculated by differencing each ensemble member and the ensemble mean within each ensemble (e.g., for the  $u$ -wind component,  $U'_n = U_n^{\text{APSU}} - \overline{U}^{\text{APSU}}$ ) and the inter-ensemble DTE (pairwise difference) is calculated by differencing corresponding members between

different ensembles (e.g., for the  $u$ -wind component,  $U'_n = U_n^{\text{APSU}} - U_n^{\text{HWRf}}$ ). Since each forecast model has its own domain configuration (Fig. 1), all domain 1 (D01;  $\sim 27$  km) analyses and forecasts are interpolated onto a  $0.25^\circ \times 0.25^\circ$  grid with a common  $50^\circ \times 50^\circ$  static domain, where all three model domains overlap. The common

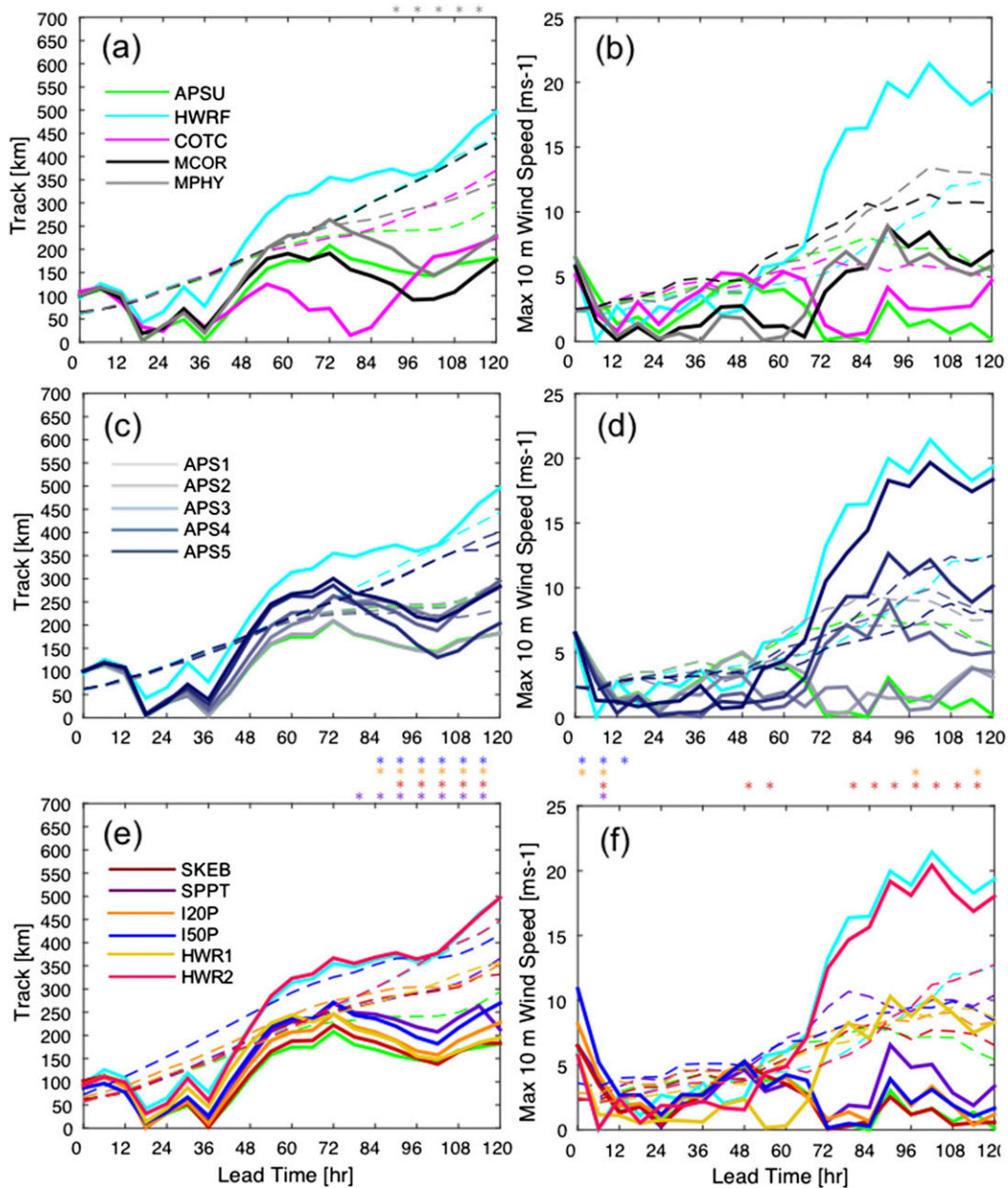


FIG. 7. As in Fig. 6, but for Hurricane Edouard (2014) initialized at 1200 UTC 11 Sep 2014 with (e),(f) experiments HWR1 (yellow) and HWR2 (pink).

domain encompasses each respective TC and its environment for the duration of the 5-day simulations, allowing for consistent inter- and intra-ensemble DTE calculations.

The intra-ensemble domain-integrated DTE of the single-core single-physics ensembles for Sandy (Figs. 9a,b) and Edouard (Figs. 9c,d) reveals error growth characteristics similar to those of the scalar track and intensity metrics in Fig. 8. Initially, HWRF, APS4, and APS5 have a noticeable reduction in DTE for both Sandy and Edouard, remaining the low for the first 72 h for Sandy and during the

entire forecast period for Edouard. These three ensembles use similar convective parameterization schemes in the two coarse-resolution domains (D01 and D02), but without cumulus parameterization in the inner-most domain (D03). Analyses show that the convective parameterization may be more easily triggered by moist instabilities in the 9-km D02 than the ensembles without cumulus parameterization in this intermediate-resolution domain. The convective parameterization in D02 can stabilize the atmosphere, which in turn likely reduces the differences among

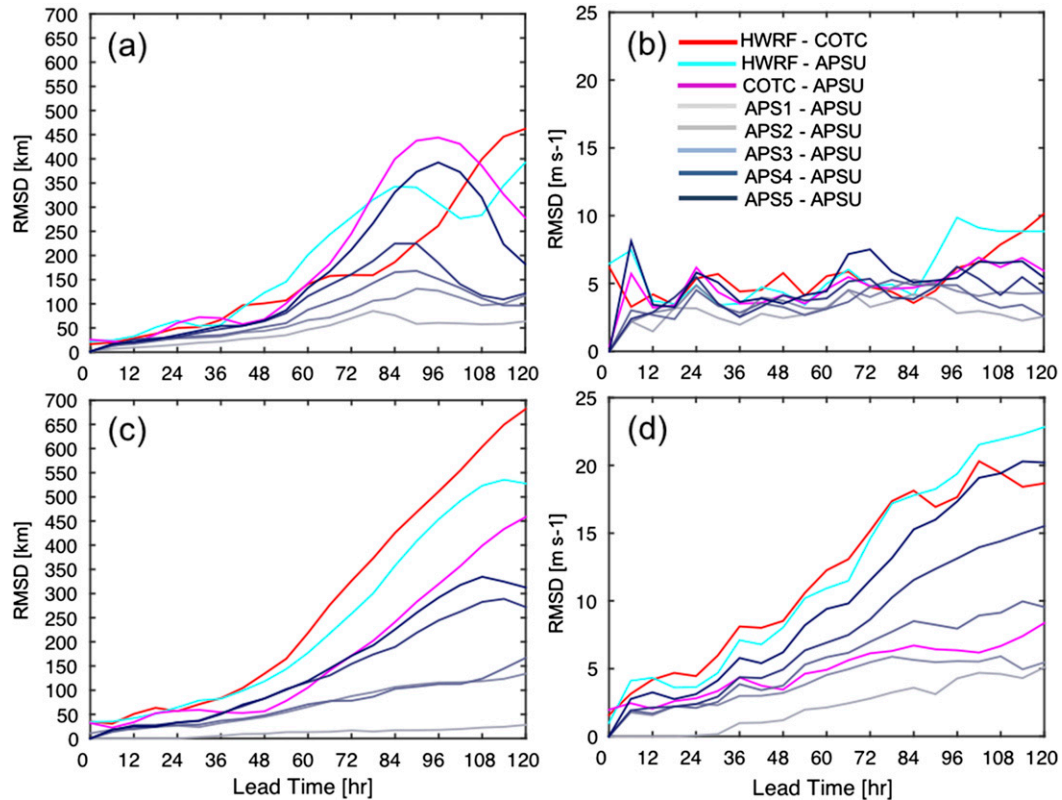


FIG. 8. Pairwise RMSDs of (a),(c) track (km) and (b),(d) maximum 10-m wind speed ( $\text{m s}^{-1}$ ) at 6-hourly lead times between HWRf and COTC (red), HWRf and APSU (cyan), COTC and APSU (magenta), and APS1 and APSU (light gray) through APS5 and APSU (navy). (top) Hurricane Sandy initialized at 0000 UTC 26 Oct 2012 and (bottom) Hurricane Edouard at 1200 UTC 11 Sep 2014.

ensemble members. Using a single-core ensemble, the SKEB and SPPT sensitivity experiments increase the intraensemble DTE due to the use of spatially and temporally correlated noise that differ from member to member, substantially increasing the differences among ensemble members.

Integrating the DTE over the ensemble  $n$  and vertical  $k$  dimensions and taking the root mean generates a two-dimensional spatial distribution of RMDTE, which is useful for spatially examining the evolution of the DTE. Figures 10a–d (Sandy) and later in Figs. 12a–d (Edouard) show the spatial evolution of the intramodel RMDTE due to IC perturbations of the ensemble surrounding the TC vortex. The intermodel differences in Figs. 10e–h (and Figs. 12e–h) indicate there is error growth associated with the large-scale environment besides the inner-core uncertainties.

Coherent error structures are evident for the intermodel RMDTE for Sandy (Figs. 10e–h) by 12 h, including the midlatitude trough over the central United States, the TC itself northeast of the Bahamas, and a warm front and upper-level ridge building to the northeast of the TC vortex. The intramodel RMDTE (Figs. 10a–d) at 12 h indicates systematic differences

associated with the TC and the surrounding environment, but they are limited to the spatial extent of the initial perturbations. By 36 h (Figs. 11a–d), the intramodel initial condition errors associated with the TC have also induced environmental differences east of the midlatitude trough over the U.S. eastern seaboard and the warm front and ridge to the west. A substantial difference in the midlatitude trough for HWRf over the eastern seaboard can be seen in Figs. 11e,g when compared to APSU and COTC. The same intra- and intermodel relationships seen in Sandy can be found with Edouard. Large-scale environmental error growth with coherent error structures in the Caribbean Sea associated with a tropical wave, the ITCZ off the coast of NE South America, and a tropical wave to the south of the Cape Verde Islands are evident by 12 h (Figs. 12e–h) and become more pronounced by 36 h (Figs. 13e–h). The spatial RMDTE analysis indicates model physical parameterizations can substantially impact the evolution of the initial condition uncertainties by as little as 12 h, despite using identical initial large-scale TC environments. Comparing APSU to APS5 provides further evidence that modifying the physical parameterizations

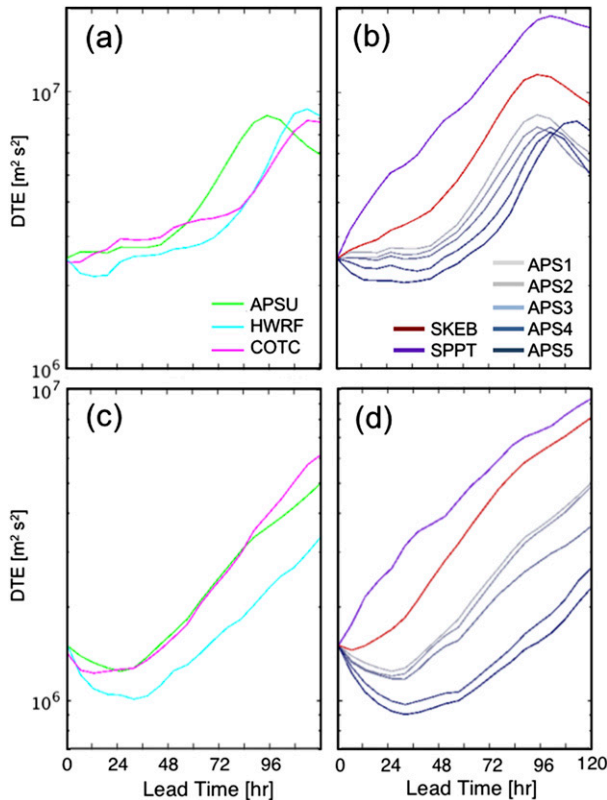


FIG. 9. Domain-integrated DTE ( $\text{m}^2 \text{s}^{-2}$ ) for (top) Hurricane Sandy initialized at 0000 UTC 26 Sep 2012 and (bottom) Hurricane Edouard initialized at 1200 UTC 11 Sep 2014 for (a),(c) APSU (green), HWRF (cyan), and COTC (magenta), and (b),(d) APS1–5 (gray to navy), SKEB (maroon), and SPPT (purple).

can alter the evolution of the ensemble similarly to changing the model core.

#### g. Ensemble composite and correlation analysis for Hurricane Edouard (2014)

A lingering question is why the HWRF ensemble initiated with the same initial conditions as the APSU and COTC ensembles performs so differently for both track (Fig. 4b) and intensity (Fig. 5b) for Edouard. From Fig. 5a, it is clear that the HWRF ensemble struggles to develop Edouard, with most members strengthening only marginally after 48 h while both the APSU and COTC ensembles begin a period of continual strengthening.

To investigate this issue, ensemble composites are constructed of relevant dynamic and thermodynamic fields important for TC development and maintenance. Relative humidity (e.g., Gray et al. 1975; Gray 1977) and environmental wind shear (e.g., Simpson and Riehl 1958; Gray 1968; Zhang and Tao 2013; Munsell et al. 2017) have been shown in past studies to impact TC genesis and maintenance. Since all members in both the

APSU and COTC ensembles develop and strengthen over the 5-day forecast period (Fig. 5a), the 10 worst-performing HWRF ensemble members (hereafter HWRF-poor, in terms of maximum 10-m wind speed at 96 h) are selected for TC-centered composites from D01. The same 10 subset ensemble members that are used to define the HWRF-poor composite are selected from the APSU and COTC ensembles for direct comparison. Examining these three subset ensemble composites of simulated maximum radar reflectivity,<sup>9</sup> mean sea level pressure (MSLP), and tilt (magenta arrow) and shear (red arrow) vectors<sup>10</sup> beginning at 24 h (Figs. 14a,d,g), it is clear that the HWRF subset struggles to generate strong convection down shear of the TC center. By 72 h (Figs. 14c,f,i), convection in HWRF-poor remains weak. The low- and midlevel circulation centers are misaligned, a sign of a moderately sheared system (Corbosiero and Molinari 2002; Rogers et al. 2003). The tilt magnitude continually increase for HWRF-poor throughout the simulation, while the APSU and COTC subset composites become steadily aligned (Zhang and Tao 2013; Tao and Zhang 2014) prior to intensification. However, it remains unclear whether and how the difference in the wind shear leads to the large divergence between the APSU–COTC ensembles and the HWRF ensemble. The presence of moderate wind shear can potentially lead to reduced TC predictability (e.g., Zhang and Tao 2013), while the difference in shear could lead to differences in the development of the stationary rainband to the northeast of the TC center.

This rainband is associated with a horizontal wind shear region between the subtropical high pressure that stretched across the northern Atlantic and the southeasterly winds induced by the large-scale TC circulation pattern. A difference between the HWRF-poor composite (Fig. 14h) and the composites of the APSU (Fig. 14b) and COTC (Fig. 14e) ensembles is clearly evident by 48 h. D02 of HWRF is smaller than those of APSU and COTC (Fig. 1), as such, the stationary rainband lies outside of the high-resolution nest. Therefore, the rainband is simulated with a lower-resolution grid that uses convective parameterization. It is hypothesized

<sup>9</sup> To generate comparable simulated radar reflectivities between models, only the cloud water and rainwater mixing ratios were used because of differences in the microphysics schemes. This will impact the simulated radar reflectivity intensity, but the relative difference between models is the focus.

<sup>10</sup> The TC-centered mean environmental wind shear is calculated within a  $3^\circ$  ring from  $2^\circ$  to  $5^\circ$  from the TC center, and the tilt vector is calculated as the displacement between the 500- and 850-hPa maximum potential vorticity centers.

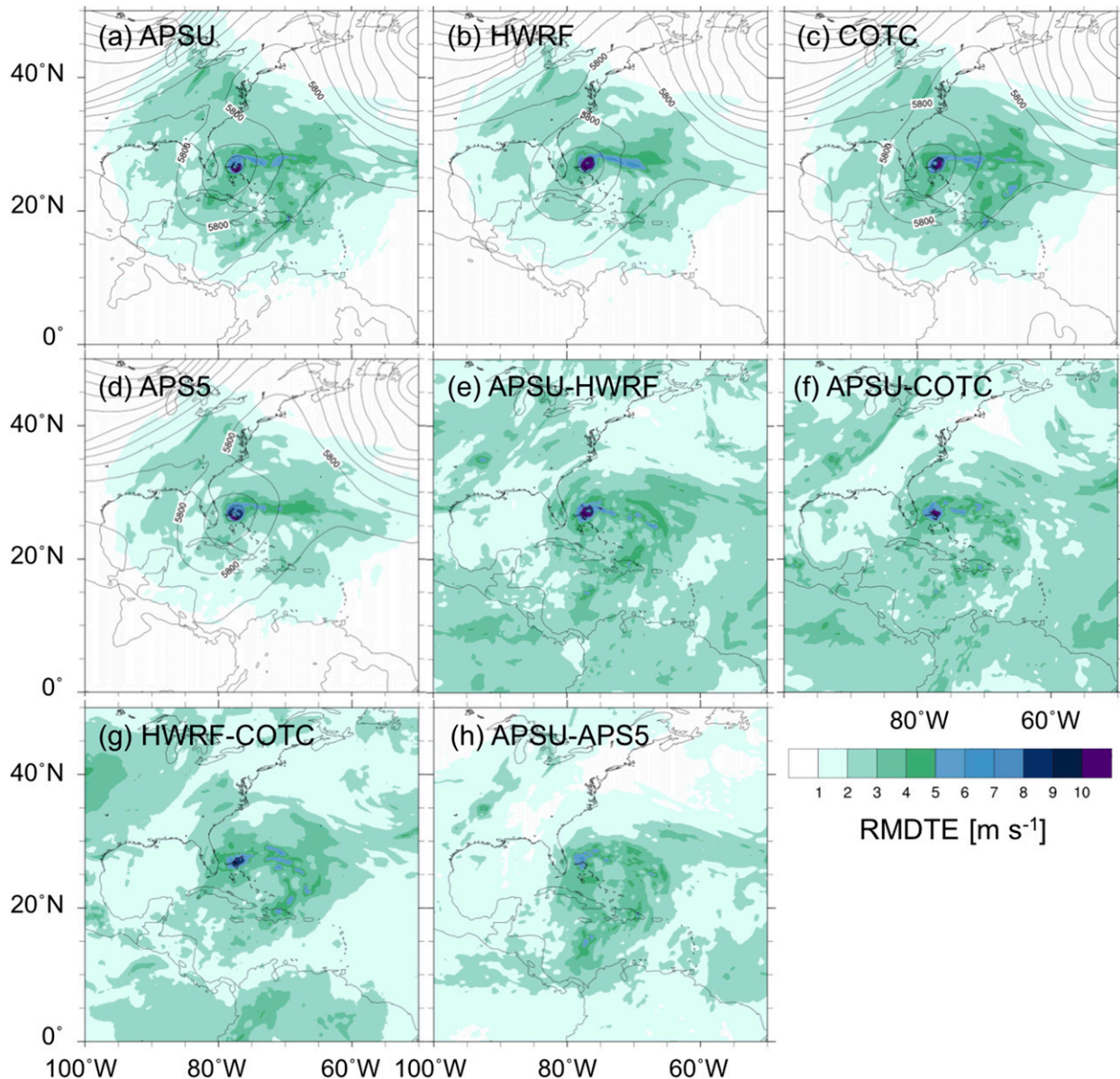


FIG. 10. Vertically integrated RMDTE ( $\text{m s}^{-1}$ ) for Hurricane Sandy (2012) at 12-h lead time for the simulation initialized at 0000 UTC 26 Oct 2012. Shown are the RMDTE of the ensemble spread for (a) APSU, (b) HWRf, (c) COTC, and (d) APS5; the RMDTE between multicore ensembles (e) APSU and HWRf, (f) APSU and COTC, and (g) HWRf and COTC; and (h) the RMDTE between APS5 and APSU. Panels (a)–(d) include the ensemble mean 500-hPa geopotential height contoured every 50 m. All panels show the static  $50^\circ \times 50^\circ$  common domain at  $0.25^\circ \times 0.25^\circ$  horizontal resolution for which each model core's 27-km D01 is interpolated to calculate all DTE analyses.

that the convective parameterization is a factor in modulating the convection associated with the stationary rainband in the HWRf ensemble.

It has been shown in idealized and case studies that asymmetric convection of forming TC circulation patterns can be detrimental to the mean symmetric TC circulation and thus weaken the TC vortex (e.g., [Montgomery and Kallenbach 1997](#); [Nolan and Grasso 2003](#); [Nolan et al. 2007](#)). For the HWRf and APS5

experiments, it is hypothesized that the convection and latent heating associated with the enhanced convection along the stationary rainband to the northeast of the TC center evolve differently with the use of cumulus parameterization and the GFS PBL in HWRf and APS5 over the first 48 h. Such differences may alter the available energy to the symmetric TC circulation and keep the TC circulation highly tilted. This is a potential area for future study.



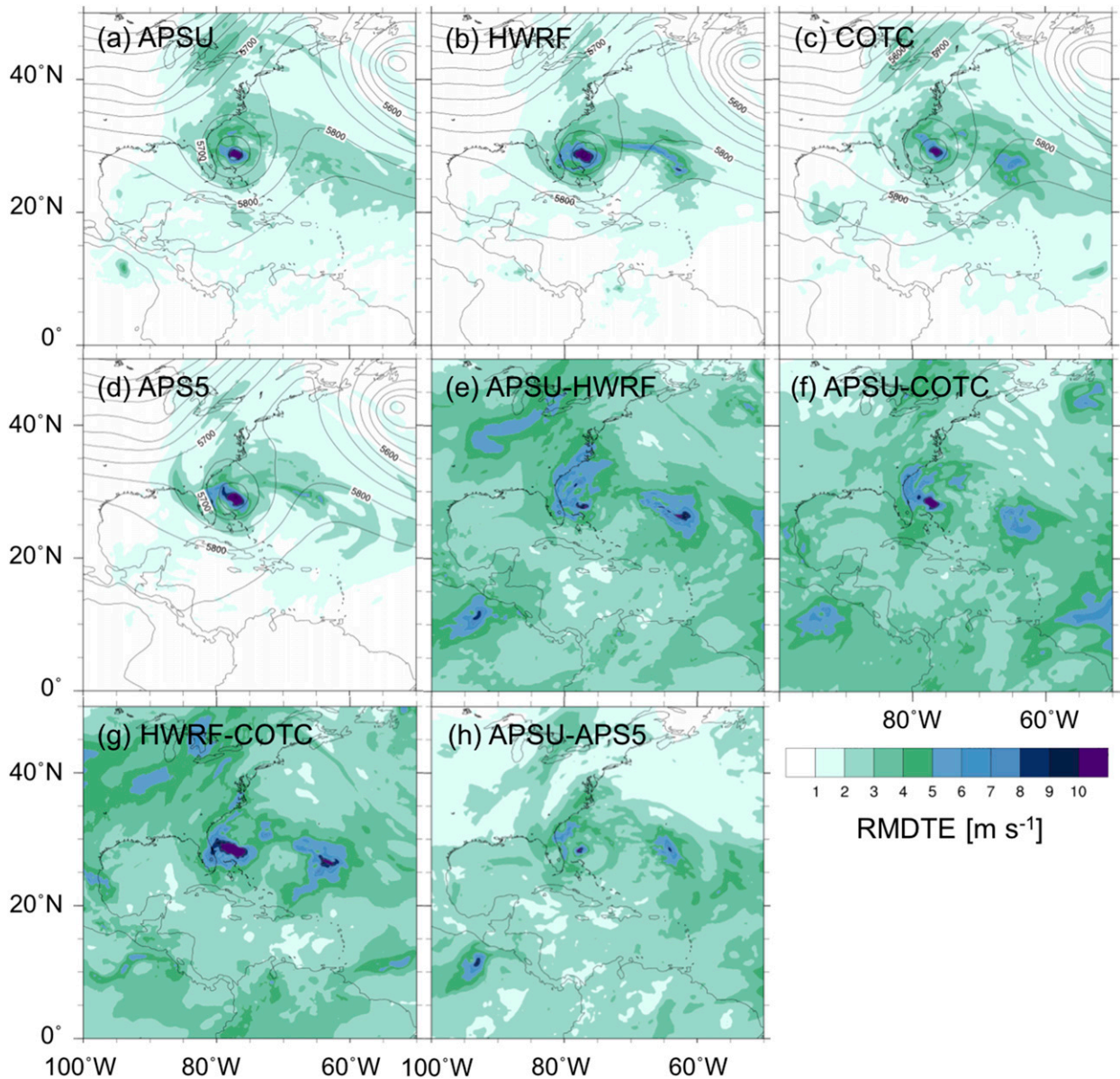


FIG. 11. As in Fig. 10, but for 36-h lead time initialized at 0000 UTC 26 Oct 2012.

To test the impact of convection parameterization surrounding the TC circulation in HWRF, the HWRF D02 is enlarged to match APSU–COTC and convective parameterization is either disabled (HWR1; Fig. 14k) or enabled (HWR2; Fig. 14n). HWR1 shows an increase in convection compared with HWR2. The track (Fig. 4d) and intensity (Fig. 5d) of HWR1 are improved compared with HWRF and HWR2. The use of convective parameterization in the local TC environment appears to have a considerable impact on the TC evolution that subsequently impacts both the track and intensity. The inherent uncertainties in representing moist convection (as well as in all other physical parameterizations) will

need to be included in the design of any future ensemble prediction systems.

#### 4. Summary and conclusions

This study has examined the mean and spread of ensembles using multiple TC-configured regional-scale convection-permitting models, including COAMPS-TC, HWRF, and WRF-ARW. Each model’s dynamic core and physical parameterization configurations were set using their 2014 “pseudo operational” HFIP configurations. The ensembles of each model were initialized with the same set of initial condition uncertainties

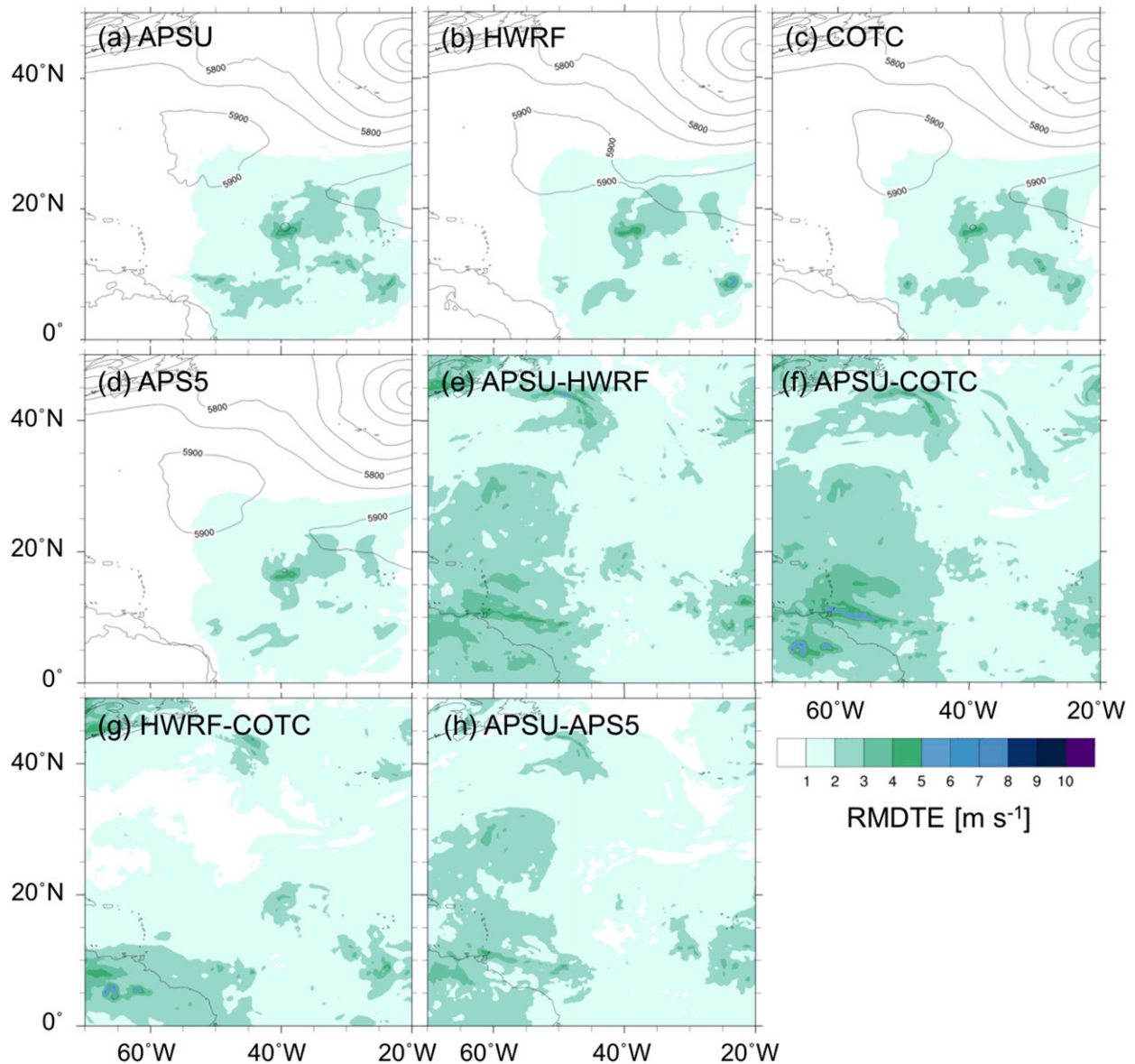


FIG. 12. As in Fig. 10, but for Hurricane Edouard (2014) initialized at 1200 UTC 11 Sep 2014.

derived from a WRF-ARW-based real-time cycling ensemble Kalman filter analysis perturbations that were blended with the Global Data Assimilation System analysis. Comparisons were made among ensembles with different model dynamical cores (multicore), as well as among the WRF-ARW ensembles using the same dynamical core (single core), but with multiphysics settings, stochastic physics, and inflated initial condition perturbations for two selected events: Hurricane Sandy (2012) and Hurricane Edouard (2014).

Comparing the track ensemble mean and spread showed that each of the TC-configured regional ensembles was capable of producing a similar ensemble

mean track and spread for shorter lead times (<36 h), but the ensemble mean error increased for each ensemble at longer lead times (>48 h). For longer lead times, the ensemble spread for both track and intensity from the multicore ensemble was larger than any single-core single-physics ensemble. For the WRF-ARW single-core ensembles with systematically varying physical parameterization experiments, the track and intensity ensemble mean could be systematically altered by modifying the physical parameterization configurations.

For the two cases studied, a single-core multiphysics ensemble randomly sampled from the WRF-ARW

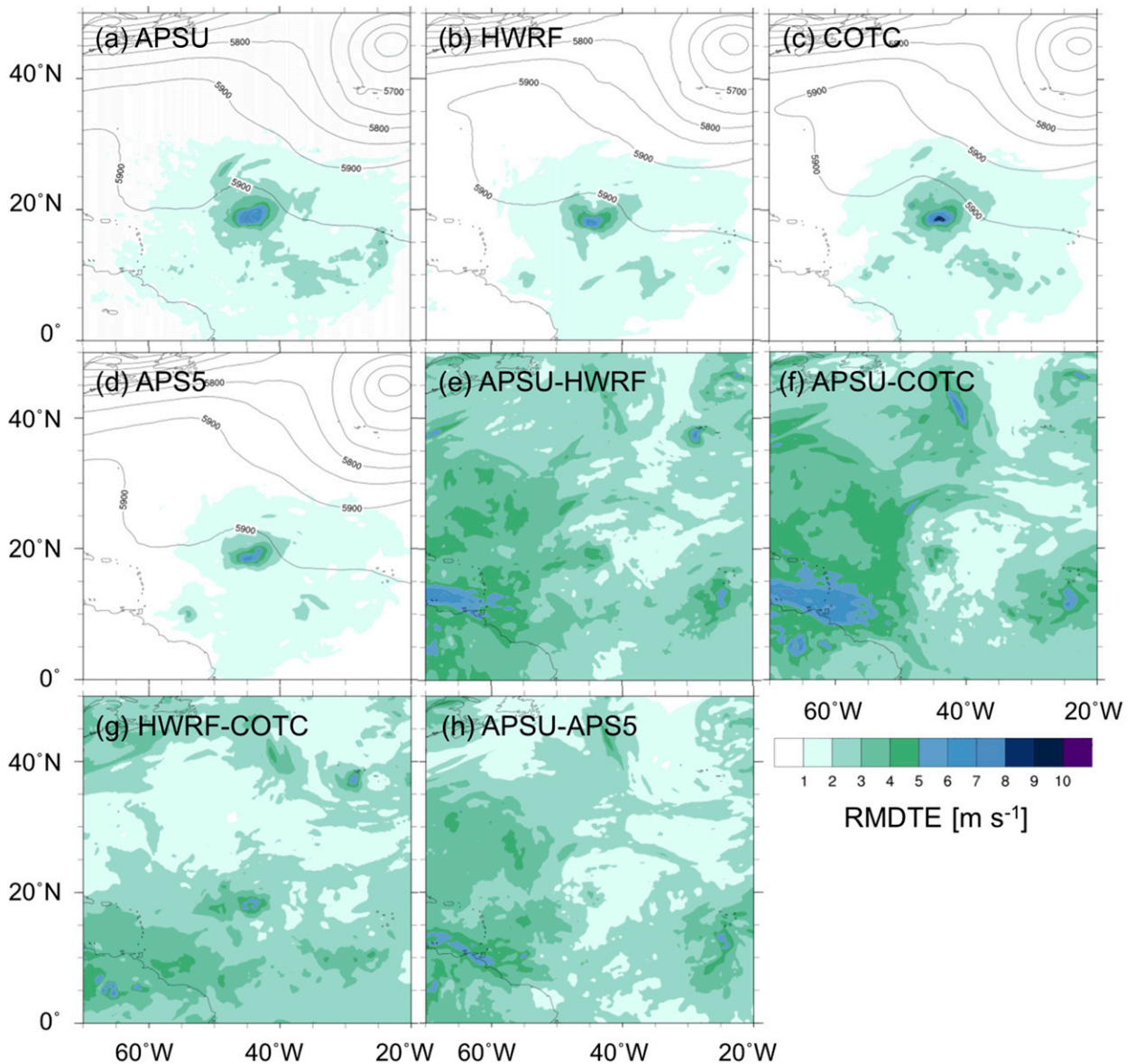


FIG. 13. As in Fig. 10, but for Hurricane Edouard (2014) at 36-h lead time initialized at 1200 UTC 11 Sep 2014.

single-core single-physics ensembles could generally resemble the performance of the multicore ensemble, except when the single-core multiphysics ensemble was deficient in the model physical parameterizations. This occurred for the Edouard case study and resulted in a reduced track spread. The single-core multiphysics ensemble benefited from a diversity of TC-tuned physical parameterizations to maximize the number of independent forecast solutions.

The SKEBS and SPPT stochastic physics algorithms in WRF-ARW were independently used to examine current methods of accounting for model uncertainties. Both increased the ensemble track and intensity spread

despite using a single-core single-physics ensemble. The SPPT algorithm had a more pronounced impact on both the ensemble mean and spread. Future work is needed to further tune the empirical parameters within the SKEBS and SPPT algorithms that could potentially lead to reduced ensemble mean error while maintaining consistency with ensemble spread.

A set of inflated initial condition perturbation experiments were performed with the WRF-ARW ensemble to account for initial condition uncertainties. These inflated initial perturbations increased both the track and intensity ensemble spread, with the magnitude of this increase being case specific. The larger

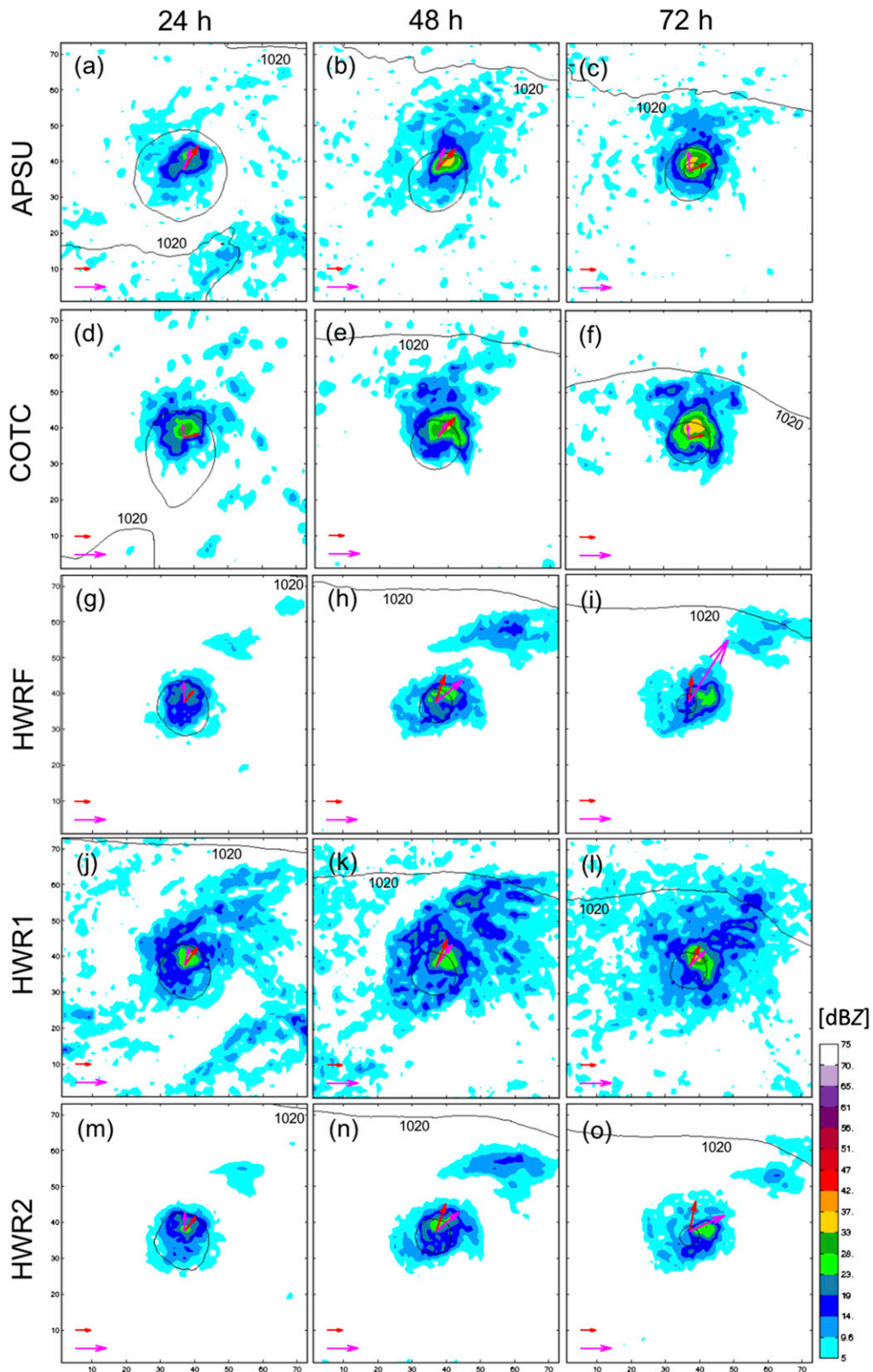


FIG. 14. HWRP-poor ensemble members D01 composite simulated maximum radar reflectivity (cloud water and rainwater only) and MSLP for (a)–(c) APSU, (d)–(f) COTC, (g)–(i) HWRP, (j)–(l) HWR1, and (m)–(o) HWR2 at lead times of (left) 24 h, (center) 48 h, and (right) 72 h initialized at 1200 UTC 11 Sep 2014. The deep-layer (200–850 hPa) shear vector (red arrow) and (850–500 hPa) tilt vector (magenta arrow) are overlaid in each panel. Reference shear vector (5 m s<sup>-1</sup>; red) and tilt vector (200 km; magenta) are shown in the bottom-left corner of each panel.

increase in the initial condition perturbation increased the ensemble spread similar to the single-core multiphysics and multicore ensembles.

For the two case studies presented, an ensemble could be constructed using a single core in combination with 1) varying physical parameterizations, 2) stochastic physics algorithms, or 3) inflated initial perturbations, to produce track and intensity forecast uncertainties similar to those produced by a multicore ensemble. The results suggest that model physical parameterizations (other than differences in the dynamic cores) for these two case studies may play an important role in evolving the case-specific TC track and intensity uncertainties. For Edouard, it is hypothesized that an increased number of TC-configured physical parameterizations used in the single-core multiphysics ensemble in WRF-ARW may have provided a larger sample of possible solutions and a better representation of the track forecast uncertainties. Examining the differences between APSU-COTC and HWRF track and intensity spread for Edouard provide insight into how model resolution and domain size may alter the evolution of TC forecast uncertainties.

The current study only uses two hurricane events to exemplify how the choice of the model dynamics core, model physical parameterizations, stochastic physics algorithms, and initial condition uncertainties impact the practical limits of predictability. Further study is needed using multiple seasons of ensemble TC forecasts to identify the impact on ensemble performance.

*Acknowledgments.* The authors thank Judith Berner at NCAR for her help implementing SKEBS and SPPT within the WRF-ARW model. We benefited greatly from the anonymous reviewers' comments to the earlier version of our manuscript. Proofreading by Robert Nystrom and Alex Kowaleski is greatly appreciated. The computing was performed at supercomputing facilities at Texas Advanced Computing Center and NOAA/ESRL. This work was supported by Office of Naval Research Grant N000140910526, National Science Foundation Grant AGS-1305798, NASA Grant NNX12AJ79G, and NOAA/HFIP.

#### REFERENCES

- Aberson, S. D., 2001: The ensemble of tropical cyclone track forecasting models in the North Atlantic basin (1976–2000). *Bull. Amer. Meteor. Soc.*, **82**, 1895–1904, doi:10.1175/1520-0477(2001)082<0000:TEOTCT>2.3.CO;2.
- Berner, J., G. J. Shutts, M. Leutbecher, and T. N. Palmer, 2009: A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF Ensemble Prediction System. *J. Atmos. Sci.*, **66**, 603–626, doi:10.1175/2008JAS2677.1.
- , K. R. Fossell, S.-Y. Ha, J. P. Hacker, and C. Snyder, 2015: Increasing the skill of probabilistic forecasts: Understanding performance improvements from model-error representations. *Mon. Wea. Rev.*, **143**, 1295–1320, doi:10.1175/MWR-D-14-00091.1.
- Blake, E. S., T. B. Kimberlain, R. J. Berg, J. P. Cangialosi, and J. L. B. II, 2013: Tropical cyclone report: Hurricane Sandy. National Hurricane Center, 157 pp. [Available online at [http://www.nhc.noaa.gov/data/tcr/AL182012\\_Sandy.pdf](http://www.nhc.noaa.gov/data/tcr/AL182012_Sandy.pdf).]
- Braun, S. A., and W.-K. Tao, 2000: Sensitivity of high-resolution simulations of Hurricane Bob (1991) to planetary boundary layer parameterizations. *Mon. Wea. Rev.*, **128**, 3941–3961, doi:10.1175/1520-0493(2000)129<3941:SOHRSO>2.0.CO;2.
- Bu, Y. P., R. G. Fovell, and K. L. Corbosiero, 2014: Influence of cloud–radiative forcing on tropical cyclone structure. *J. Atmos. Sci.*, **71**, 1644–1662, doi:10.1175/JAS-D-13-0265.1.
- Corbosiero, K. L., and J. Molinari, 2002: The effects of vertical wind shear on the distribution of convection in tropical cyclones. *Mon. Wea. Rev.*, **130**, 2110–2123, doi:10.1175/1520-0493(2002)130<2110:TEOVWS>2.0.CO;2.
- Doyle, J. D., and Coauthors, 2012: Real-time tropical cyclone prediction using COAMPS-TC. *Adv. Geosci.*, **28**, 15–28, doi:10.1142/9789814405683\_0002.
- , and Coauthors, 2014: Tropical Cyclone Prediction Using COAMPS-TC. *Oceanography*, **27**, 104–115, doi:10.5670/oceanog.2014.72.
- Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, doi:10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.
- Fels, S. B., and M. D. Schwarzkopf, 1981: An efficient, accurate algorithm for calculating CO<sub>2</sub> 15 μm band cooling rates. *J. Geophys. Res.*, **86**, 1205–1232, doi:10.1029/JC086iC02p01205.
- Ferrier, B. S., 1994: A double-moment multiple-phase four-class bulk ice scheme. Part I: Description. *J. Atmos. Sci.*, **51**, 249–280, doi:10.1175/1520-0469(1994)051<0249:ADMMPF>2.0.CO;2.
- Fovell, R. G., and H. Su, 2007: Impact of cloud microphysics on hurricane track forecasts. *Geophys. Res. Lett.*, **34**, L24810, doi:10.1029/2007GL031723.
- , K. L. Corbosiero, and H.-C. Kuo, 2009: Cloud microphysics impact on hurricane track as revealed in idealized experiments. *J. Atmos. Sci.*, **66**, 1764–1778, doi:10.1175/2008JAS2874.1.
- , —, A. Seifert, and K.-N. Liou, 2010: Impact of cloud-radiative processes on hurricane track. *Geophys. Res. Lett.*, **37**, L07808, doi:10.1029/2010GL042691.
- Fu, Q., and K. N. Liou, 1992: On the correlated *k*-distribution method for radiative transfer in nonhomogeneous atmospheres. *J. Atmos. Sci.*, **49**, 2139–2156, doi:10.1175/1520-0469(1992)049<2139:OTCDMF>2.0.CO;2.
- Goerss, J. S., 2000: Tropical cyclone track forecasts using an ensemble of dynamical models. *Mon. Wea. Rev.*, **128**, 1187–1193, doi:10.1175/1520-0493(2000)128<1187:TCTFUA>2.0.CO;2.
- Gray, W. M., 1968: Global view of the origin of tropical disturbances and storms. *Mon. Wea. Rev.*, **96**, 669–700, doi:10.1175/1520-0493(1968)096<0669:GVOTOO>2.0.CO;2.
- , 1977: Tropical cyclone genesis in the western North Pacific. *J. Meteor. Soc. Japan*, **55**, 465–482.
- , E. Ruprecht, and R. Phelps, 1975: Relative humidity in tropical weather systems. *Mon. Wea. Rev.*, **103**, 685–690, doi:10.1175/1520-0493(1975)103<0685:RHITWS>2.0.CO;2.
- Green, B. W., and F. Zhang, 2013: Impacts of air–sea flux parameterizations on the intensity and structure of tropical

- cyclones. *Mon. Wea. Rev.*, **141**, 2308–2324, doi:[10.1175/MWR-D-12-00274.1](https://doi.org/10.1175/MWR-D-12-00274.1).
- , and —, 2014: Sensitivity of tropical cyclone simulations to parametric uncertainties in air–sea fluxes and implications for parameter estimation. *Mon. Wea. Rev.*, **142**, 2290–2308, doi:[10.1175/MWR-D-13-00208.1](https://doi.org/10.1175/MWR-D-13-00208.1).
- Grell, G. A., and S. R. Freitas, 2013: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmos. Chem. Phys.*, **14**, 5233–5250, doi:[10.5194/acp-14-5233-2014](https://doi.org/10.5194/acp-14-5233-2014).
- Han, J., and H.-L. Pan, 2011: Revision of convection and vertical diffusion schemes in the NCEP Global Forecast System. *Wea. Forecasting*, **26**, 520–533, doi:[10.1175/WAF-D-10-05038.1](https://doi.org/10.1175/WAF-D-10-05038.1).
- Hodur, R. M., 1997: The Naval Research Laboratory's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS). *Mon. Wea. Rev.*, **125**, 1414–1430, doi:[10.1175/1520-0493\(1997\)125<1414:TNRLSC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<1414:TNRLSC>2.0.CO;2).
- Hohenegger, C., and C. Schär, 2007: Predictability and error growth dynamics in cloud-resolving models. *J. Atmos. Sci.*, **64**, 4467–4478, doi:[10.1175/2007JAS2143.1](https://doi.org/10.1175/2007JAS2143.1).
- Hong, S.-Y., and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.*, **124**, 2322–2339, doi:[10.1175/1520-0493\(1996\)124<2322:NBLVDI>2.0.CO;2](https://doi.org/10.1175/1520-0493(1996)124<2322:NBLVDI>2.0.CO;2).
- , and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **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, doi:[10.1175/MWR3199.1](https://doi.org/10.1175/MWR3199.1).
- Houtekamer, P. L., H. L. Mitchell, and X. Deng, 2009: Model error representation in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2126–2143, doi:[10.1175/2008MWR2737.1](https://doi.org/10.1175/2008MWR2737.1).
- Jin, H., M. S. Peng, Y. Jin, and J. D. Doyle, 2014: An evaluation of the impact of horizontal resolution on tropical cyclone predictions using COAMPS-TC. *Wea. Forecasting*, **29**, 252–270, doi:[10.1175/WAF-D-13-00054.1](https://doi.org/10.1175/WAF-D-13-00054.1).
- Jin, Y., W. T. Thompson, S. Wang, and C.-S. Liou, 2007: A numerical study of the effect of dissipative heating on tropical cyclone intensity. *Wea. Forecasting*, **22**, 950–966, doi:[10.1175/WAF1028.1](https://doi.org/10.1175/WAF1028.1).
- Johnson, A., and X. Wang, 2012: Verification and calibration of neighborhood and object-based probabilistic precipitation forecasts from a multimodel convection-allowing ensemble. *Mon. Wea. Rev.*, **140**, 3054–3077, doi:[10.1175/MWR-D-11-00356.1](https://doi.org/10.1175/MWR-D-11-00356.1).
- Judt, F., S. S. Chen, and J. Berner, 2016: Predictability of tropical cyclone intensity: Scale-dependent forecast error growth in high-resolution stochastic kinetic-energy backscatter ensembles. *Quart. J. Roy. Meteor. Soc.*, **142**, 43–57, doi:[10.1002/qj.2626](https://doi.org/10.1002/qj.2626).
- Kain, J. S., and J. M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain–Fritsch scheme. *The Representation of Cumulus Convection in Numerical Models*, Meteor. Monogr., No. 46, Amer. Meteor. Soc., 165–170.
- Krishnamurti, T. N., 1999: Improved weather and seasonal climate forecasts from multimodel superensemble. *Science*, **285**, 1548–1550, doi:[10.1126/science.285.5433.1548](https://doi.org/10.1126/science.285.5433.1548).
- Lang, S. T. K., M. Leutbecher, and S. C. Jones, 2012: Impact of perturbation methods in the ECMWF Ensemble Prediction System on tropical cyclone forecasts. *Quart. J. Roy. Meteor. Soc.*, **138**, 2030–2046, doi:[10.1002/qj.1942](https://doi.org/10.1002/qj.1942).
- Leslie, L. M., and K. Fraedrich, 1990: Reduction of tropical cyclone position errors using an optimal combination of independent forecasts. *Wea. Forecasting*, **5**, 158–161, doi:[10.1175/1520-0434\(1990\)005<0158:ROTCPE>2.0.CO;2](https://doi.org/10.1175/1520-0434(1990)005<0158:ROTCPE>2.0.CO;2).
- Lord, S. J., H. E. Willoughby, and J. M. Piotrowicz, 1984: Role of a parameterized ice-phase microphysics in an axisymmetric, non-hydrostatic tropical cyclone model. *J. Atmos. Sci.*, **41**, 2836–2848, doi:[10.1175/1520-0469\(1984\)041<2836:ROAPIP>2.0.CO;2](https://doi.org/10.1175/1520-0469(1984)041<2836:ROAPIP>2.0.CO;2).
- Louis, J.-F., 1979: A parametric model of vertical eddy fluxes in the atmosphere. *Bound.-Layer Meteor.*, **17**, 187–202, doi:[10.1007/BF00117978](https://doi.org/10.1007/BF00117978).
- McFarquhar, G. M., H. Zhang, G. Heymsfield, J. B. Halverson, R. Hood, J. Dudhia, and F. Marks, 2006: Factors affecting the evolution of Hurricane Erin (2001) and the distributions of hydrometeors: Role of microphysical processes. *J. Atmos. Sci.*, **63**, 127–150, doi:[10.1175/JAS3590.1](https://doi.org/10.1175/JAS3590.1).
- Melhauser, C., and F. Zhang, 2014: Diurnal radiation cycle impact on the pregenesis environment of Hurricane Karl (2010). *J. Atmos. Sci.*, **71**, 1241–1259, doi:[10.1175/JAS-D-13-0116.1](https://doi.org/10.1175/JAS-D-13-0116.1).
- Mellor, G., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys.*, **20**, 851–875, doi:[10.1029/RG020i004p00851](https://doi.org/10.1029/RG020i004p00851).
- Meng, Z., and F. Zhang, 2007: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect model experiments. *Mon. Wea. Rev.*, **135**, 1403–1423, doi:[10.1175/MWR3352.1](https://doi.org/10.1175/MWR3352.1).
- , and —, 2008a: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part III: Comparison with 3DVAR in a real-data case study. *Mon. Wea. Rev.*, **136**, 522–540, doi:[10.1175/2007MWR2106.1](https://doi.org/10.1175/2007MWR2106.1).
- , and —, 2008b: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part IV: Comparison with 3DVAR in a month-long experiment. *Mon. Wea. Rev.*, **136**, 3671–3682, doi:[10.1175/2008MWR2270.1](https://doi.org/10.1175/2008MWR2270.1).
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16 663–16 682, doi:[10.1029/97JD00237](https://doi.org/10.1029/97JD00237).
- Montgomery, M. T., and R. J. Kallenbach, 1997: A theory for vortex Rossby-waves and its application to spiral bands and intensity changes in hurricanes. *Quart. J. Roy. Meteor. Soc.*, **123**, 435–465, doi:[10.1002/qj.49712353810](https://doi.org/10.1002/qj.49712353810).
- Munsell, E. B., and F. Zhang, 2014: Prediction and uncertainty of Hurricane Sandy (2012) explored through a real-time cloud-permitting ensemble analysis and forecast system assimilating airborne Doppler radar observations. *J. Adv. Model. Earth Syst.*, **6**, 38–58, doi:[10.1002/2013MS000297](https://doi.org/10.1002/2013MS000297).
- , —, J. A. Sippel, and S. A. Braun, 2017: Dynamics and predictability of the intensification of Hurricane Edouard (2014). *J. Atmos. Sci.*, **74**, 573–595, doi:[10.1175/JAS-D-16-0018.1](https://doi.org/10.1175/JAS-D-16-0018.1).
- Murphy, J. M., 1988: The impact of ensemble forecasts on predictability. *Quart. J. Roy. Meteor. Soc.*, **114**, 463–493, doi:[10.1002/qj.49711448010](https://doi.org/10.1002/qj.49711448010).
- Nolan, D. S., and L. D. Grasso, 2003: Nonhydrostatic, three-dimensional perturbations to balanced, hurricane-like vortices. Part II: Symmetric response and nonlinear simulations. *J. Atmos. Sci.*, **60**, 2717–2745, doi:[10.1175/1520-0469\(2003\)060<2717:NTPTBH>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<2717:NTPTBH>2.0.CO;2).
- , Y. Moon, and D. P. Stern, 2007: Tropical cyclone intensification from asymmetric convection: Energetics and efficiency. *J. Atmos. Sci.*, **64**, 3377–3405, doi:[10.1175/JAS3988.1](https://doi.org/10.1175/JAS3988.1).
- Novak, D. R., D. R. Bright, and M. J. Brennan, 2008: Operational forecaster uncertainty needs and future roles. *Wea. Forecasting*, **23**, 1069–1084, doi:[10.1175/2008WAF2222142.1](https://doi.org/10.1175/2008WAF2222142.1).

- Nutter, P., D. Stensrud, and M. Xue, 2004a: Effects of coarsely resolved and temporally interpolated lateral boundary conditions on the dispersion of limited-area ensemble forecasts. *Mon. Wea. Rev.*, **132**, 2358–2377, doi:10.1175/1520-0493(2004)132<2358:EOCRAT>2.0.CO;2.
- , M. Xue, and D. Stensrud, 2004b: Application of lateral boundary condition perturbations to help restore dispersion in limited-area ensemble forecasts. *Mon. Wea. Rev.*, **132**, 2378–2390, doi:10.1175/1520-0493(2004)132<2378:AOLBCP>2.0.CO;2.
- Palmer, T. N., 2002: The economic value of ensemble forecasts as a tool for risk assessment: From days to decades. *Quart. J. Roy. Meteor. Soc.*, **128**, 747–774, doi:10.1256/0035900021643593.
- , R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G. Shutts, M. Steinheimer, and A. Weisheimer, 2009: Stochastic parametrization and model uncertainty. ECMWF Tech. Memo. 598, 42 pp. [Available online at <http://www.ecmwf.int/sites/default/files/elibrary/2009/11577-stochastic-parametrization-and-model-uncertainty.pdf>.]
- Qi, L., H. Yu, and P. Chen, 2014: Selective ensemble-mean technique for tropical cyclone track forecast by using ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.*, **140**, 805–813, doi:10.1002/qj.2196.
- Rogers, E., T. Black, B. Ferrier, Y. Lin, D. Parrish, and G. DiMego, 2001: Changes to the NCEP Meso Eta Analysis and Forecast System: Increase in resolution, new cloud microphysics, modified precipitation assimilation, modified 3DVAR analysis. NWS Tech. Procedures Bull. 488, 15 pp.
- Rogers, R., S. Chen, J. Tenerelli, and H. Willoughby, 2003: A numerical study of the impact of vertical shear on the distribution of rainfall in Hurricane Bonnie (1998). *Mon. Wea. Rev.*, **131**, 1577–1599, doi:10.1175//2546.1.
- Romine, G. S., C. S. Schwartz, C. Snyder, J. L. Anderson, and M. L. Weisman, 2013: Model bias in a continuously cycled assimilation system and its influence on convection-permitting forecasts. *Mon. Wea. Rev.*, **141**, 1263–1284, doi:10.1175/MWR-D-12-00112.1.
- , —, J. Berner, K. R. Fossell, C. Snyder, J. L. Anderson, and M. L. Weisman, 2014: Representing forecast error in a convection-permitting ensemble system. *Mon. Wea. Rev.*, **142**, 4519–4541, doi:10.1175/MWR-D-14-00100.1.
- Rutledge, S. A., and P. Hobbs, 1983: The mesoscale and microscale structure and organization of clouds and precipitation in mid-latitude cyclones. VIII: A model for the “seeder-feeder” process in warm-frontal rainbands. *J. Atmos. Sci.*, **40**, 1185–1206, doi:10.1175/1520-0469(1983)040<1185:TMAMSA>2.0.CO;2.
- Sampson, C. R., J. L. Franklin, J. A. Knaff, and M. DeMaria, 2008: Experiments with a simple tropical cyclone intensity consensus. *Wea. Forecasting*, **23**, 304–312, doi:10.1175/2007WAF2007028.1.
- Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.*, **131**, 3079–3102, doi:10.1256/qj.04.106.
- Simpson, R. H., and H. Riehl, 1958: Mid-tropospheric ventilation as a constraint on hurricane development and maintenance. *Proc. Tech. Conf. on Hurricanes*, Miami Beach, FL, Amer. Meteor. Soc., D4.1–D4.10.
- Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., doi:10.5065/D68S4MVH.
- Stewart, S. R., 2014: Hurricane Edouard (2014). National Hurricane Center Tropical Cyclone Rep., 19 pp. [Available online at [http://www.nhc.noaa.gov/data/tcr/AL062014\\_Edouard.pdf](http://www.nhc.noaa.gov/data/tcr/AL062014_Edouard.pdf).]
- Sundqvist, H., 1970: Numerical simulation of the development of tropical cyclones with a ten-level model. Part II. *Tellus*, **22**, 504–510, doi:10.1111/j.2153-3490.1970.tb00516.x.
- Tallapragada, V., and Coauthors, 2014: Hurricane Weather Research and Forecasting (HWRF) Model: 2013 scientific documentation. HWRF Development Testbed Center Tech. Rep., 113 pp. [Available online at [http://www.dtcenter.org/HurrWRF/users/docs/scientific\\_documents/HWRF\\_v3.7a\\_SD.pdf](http://www.dtcenter.org/HurrWRF/users/docs/scientific_documents/HWRF_v3.7a_SD.pdf).]
- Tao, D., and F. Zhang, 2014: Effect of environmental shear, sea-surface temperature, and ambient moisture on the formation and predictability of tropical cyclones: An ensemble-mean perspective. *J. Adv. Model. Earth Syst.*, **6**, 384–404, doi:10.1002/2014MS000314.
- Torn, R. D., 2016: Evaluation of atmosphere and ocean initial condition uncertainty and stochastic exchange coefficients on ensemble tropical cyclone intensity forecasts. *Mon. Wea. Rev.*, **144**, 3487–3506, doi:10.1175/MWR-D-16-0108.1.
- , G. J. Hakim, and C. Snyder, 2006: Boundary conditions for limited-area ensemble Kalman filters. *Mon. Wea. Rev.*, **134**, 2490–2502, doi:10.1175/MWR3187.1.
- Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330, doi:10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2.
- Tracton, M. S., and E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Wea. Forecasting*, **8**, 379–398, doi:10.1175/1520-0434(1993)008<0379:OEPATN>2.0.CO;2.
- Vijaya Kumar, T. S. V., T. N. Krishnamurti, M. Fiorino, and M. Nagata, 2003: Multimodel superensemble forecasting of tropical cyclones in the Pacific. *Mon. Wea. Rev.*, **131**, 574–583, doi:10.1175/1520-0493(2003)131<0574:MSFOTC>2.0.CO;2.
- Wang, Y., 2002: An explicit simulation of tropical cyclones with a triply nested movable mesh primitive equation model: TCM3. Part II: Model refinements and sensitivity to cloud microphysics parameterization. *Mon. Wea. Rev.*, **130**, 3022–3036, doi:10.1175/1520-0493(2002)130<3022:AESOTC>2.0.CO;2.
- , and G. J. Holland, 1996: The beta drift of baroclinic vortices. Part II: Diabatic vortices. *J. Atmos. Sci.*, **53**, 3737–3756, doi:10.1175/1520-0469(1996)053<3737:TBD0BV>2.0.CO;2.
- Weng, Y., and F. Zhang, 2012: Assimilating airborne Doppler radar observations with an ensemble Kalman filter for convection-permitting hurricane initialization and prediction: Katrina (2005). *Mon. Wea. Rev.*, **140**, 841–859, doi:10.1175/2011MWR3602.1.
- , and —, 2016: Advances in convection-permitting tropical cyclone analysis and prediction through EnKF assimilation of reconnaissance aircraft observations. *J. Meteor. Soc. Japan*, **94**, 345–358, doi:10.2151/jmsj.2016-018.
- Willoughby, H. E., H.-L. Jin, S. J. Lord, and J. M. Piotrowicz, 1984: Hurricane structure and evolution as simulated by an axisymmetric, nonhydrostatic numerical model. *J. Atmos. Sci.*, **41**, 1169–1186, doi:10.1175/1520-0469(1984)041<1169:HSAEAS>2.0.CO;2.
- Zhang, F., and D. Tao, 2013: Effects of vertical wind shear on the predictability of tropical cyclones. *J. Atmos. Sci.*, **70**, 975–983, doi:10.1175/JAS-D-12-0133.1.
- , C. Snyder, and R. Rotunno, 2002: Mesoscale predictability of the “surprise” snowstorm of 24–25 January 2000. *Mon. Wea. Rev.*, **130**, 1617–1632, doi:10.1175/1520-0493(2002)130<1617:MPOTSS>2.0.CO;2.
- , —, and J. Sun, 2004: Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238–1253, doi:10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2.

- , Z. Meng, and A. Aksoy, 2006: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part I: Perfect model experiments. *Mon. Wea. Rev.*, **134**, 722–736, doi:[10.1175/MWR3101.1](https://doi.org/10.1175/MWR3101.1).
- , Y. Weng, and J. A. Sippel, 2009: Cloud-resolving hurricane initialization and prediction through assimilation of Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2105–2125, doi:[10.1175/2009MWR2645.1](https://doi.org/10.1175/2009MWR2645.1).
- , —, J. F. Gamache, and F. D. Marks, 2011: Performance of convection permitting hurricane initialization and prediction during 2008–2010 with ensemble data assimilation of inner-core airborne Doppler radar observations. *Geophys. Res. Lett.*, **38**, L15810, doi:[10.1029/2011GL048469](https://doi.org/10.1029/2011GL048469).
- Zhu, T., and D.-L. Zhang, 2006: Numerical simulation of Hurricane Bonnie (1998). Part II: Sensitivity to varying cloud microphysical processes. *J. Atmos. Sci.*, **63**, 109–126, doi:[10.1175/JAS3599.1](https://doi.org/10.1175/JAS3599.1).