• Original Paper •

Evaluation of a Regional Ensemble Data Assimilation System for Typhoon Prediction※

Lili LEI[1], Yangjinxi GE[1], Zhe-Min TAN[*1], Yi ZHANG[1], Kekuan CHU[1], Xin QIU[1], and Qifeng QIAN[2]

[1]*Key Laboratory of Mesoscale Severe Weather/Ministry of Education*, *School of Atmospheric Sciences*,
Nanjing University, *Nanjing* 210063, *China*
[2]*National Meteorological Center*, *China Meteorological Administration*, *Beijing* 100081, *China*

ABSTRACT

An ensemble Kalman filter (EnKF) combined with the Advanced Research Weather Research and Forecasting model (WRF) is cycled and evaluated for western North Pacific (WNP) typhoons of year 2016. Conventional in situ data, radiance observations, and tropical cyclone (TC) minimum sea level pressure (SLP) are assimilated every 6 h using an 80-member ensemble. For all TC categories, the 6-h ensemble priors from the WRF/EnKF system have an appropriate amount of variance for TC tracks but have insufficient variance for TC intensity. The 6-h ensemble priors from the WRF/EnKF system tend to overestimate the intensity for weak storms but underestimate the intensity for strong storms. The 5-d deterministic forecasts launched from the ensemble mean analyses of WRF/EnKF are compared to the NCEP and ECMWF operational control forecasts. Results show that the WRF/EnKF forecasts generally have larger track errors than the NCEP and ECMWF forecasts for all TC categories because the regional simulation cannot represent the large-scale environment better than the global simulation. The WRF/EnKF forecasts produce smaller intensity errors and biases than the NCEP and ECMWF forecasts for typhoons, but the opposite is true for tropical storms and severe tropical storms. The 5-d ensemble forecasts from the WRF/EnKF system for seven typhoon cases show appropriate variance for TC track and intensity with short forecast lead times but have insufficient spread with long forecast lead times. The WRF/EnKF system provides better ensemble forecasts and higher predictability for TC intensity than the NCEP and ECMWF ensemble forecasts.

Key words: ensemble Kalman filter, typhoon prediction, ensemble forecast

Article Highlights:

- An 80-member cycled WRF/EnKF with 6-h data assimilation is operated and evaluated for western North Pacific typhoons of year 2016.
- The WRF/EnKF system has an appropriate amount of variance for 6-h track forecasts and better typhoon intensity forecasts than global models.
- The WRF/EnKF system provides better ensemble forecasts and higher predictability for typhoon intensity than global models.

---

1. Introduction

Over the last few decades, there has been a steady decrease in tropical cyclone (TC) track forecast errors; but there have been minimal changes in TC intensity forecast errors over the same period (e.g., Rogers et al., 2006; Rappaport et al., 2009). TC track and intensity forecasts from numerical weather prediction models are limited by initial condition errors that are associated with the large-scale environments and TC structures, as well as model errors introduced by grid resolution and physical parameterization schemes. While TC motion is mostly controlled by the large-scale environment, TC intensity depends on the large-scale kinematic and thermodynamic environment, the inner-core dynamics, and the lower boundary condition including the surface sea temperature, ocean heat content, and land surface (e.g., Wang and Wu, 2004).

One feature of TCs is the large gradients in the mass

---

and wind fields, which are often difficult to solve due to relatively coarse grid resolutions compared to the scale of the TCs. Therefore, several different techniques have been developed to relocate the storm to the observed position and construct initial conditions with more realistic TC structures. One kind of technique splits the forecast field into basic and disturbance fields and simply relocates the TC circulation to the observed position, then obtains the updated forecast filed with a correct TC position (e.g., Liu et al., 2000; Hsiao et al., 2010). Another kind of technique generates a more realistic TC initial structure by adding a TC-like vortex into the analysis field through data assimilation schemes (e.g., Kurihara et al., 1995; Zou and Xiao, 2000; Wang et al., 2008). To avoid the interference between a bogus vortex and actual observations, there are techniques that first implant a bogus vortex and then assimilate the profile data into the bogus vortex (e.g., Chou and Wu, 2008). The third kind of technique is dynamical initialization that incorporates the three-dimensional initial TC structure through the numerical forecast model (e.g., Kurihara et al., 1993; Cha and Wang, 2013; Hendricks et al., 2013; Liu and Tan, 2016). These techniques are somewhat limited by their assumptions, so they may not be optimal at all times.

Compared to the previously discussed TC initialization methods, advanced data assimilation methods, which are not specialized for TC initialization, have been shown to make great impacts on TC forecasts. Among the widely applied data assimilation methods, ensemble-based assimilation approaches, such as the ensemble Kalman filter (EnKF; Burgers et al., 1998), have shown great promise for atmospheric data assimilation in both global (e.g., Whitaker et al., 2008; Buehner et al., 2010a, b; Houtekamer et al., 2014) and regional applications (e.g., Dowell et al., 2004; Tong and Xue, 2005; Meng and Zhang 2008; Aksoy et al., 2009). As a Monte Carlo approximation to the Kalman filter (Kalman, 1960), the EnKF uses flow-dependent error statistics estimated from short-term ensemble forecasts to determine the weight given to observations relative to model forecasts and spread observation information to model state variables. The usage of flow-dependent error statistics should provide more effective assimilation of observations near TCs, and the set of ensemble analyses can naturally provide initial conditions for TC ensemble forecasting.

Previous work has shown that EnKF assimilation is able to provide dynamically consistent TC state estimation and improve TC track and intensity forecasts. Hamill et al. (2011) initialized the National Centers for Environmental Prediction (NCEP) Global Forecasting System (GFS) with EnKF analyses and obtained improved TC track forecasts compared to the operational ensemble data assimilation systems at the time (e.g., ensemble transform technique (Wei et al., 2008) and ensemble transform Kalman filter (Hunt et al., 2007)). Torn and Hakim (2009) cycled an EnKF over the lifetime of Hurricane Katrina (2005) and obtained a 50% reduction in TC track and intensity errors compared to the NCEP GFS and National Hurricane Center (NHC) official forecasts. Moreover, Zhang et al. (2009, 2011b) showed that

assimilating Doppler radar data from both land-based and reconnaissance aircraft platforms with an EnKF led to improved TC intensity forecasts. Assimilating microwave radiances with an EnKF was also found to be beneficial for TC predictions (Schwartz et al., 2012). Besides these individual case studies with relatively short periods over which observations are assimilated, Torn (2010) cycled an EnKF over the life cycle of multiple TCs ranging from marginal to intense TCs in the Atlantic basin and found that cycling with an EnKF system was particularly effective for weak TCs. Cavallo et al. (2013) evaluated a cycling EnKF for the 2009 North Atlantic hurricane season and obtained systematically reduced TC track and intensity errors by assimilating observations, except for strong TCs. Xue et al. (2013) systematically compared the EnKF and three-dimensional variational data assimilation (3DVAR; Kleist et al., 2009) for the 2010 North Atlantic hurricane season and found significantly improved TC intensity forecasts initialized from the EnKF compared to those initialized from the 3DVAR.

Although previous studies have been encouraging, systematic work on regional cycling EnKF systems for the western North Pacific (WNP) basin is limited. Besides the internal dynamical and physical processes that have important impacts on TC track and intensity, TCs over the WNP are strongly influenced by complex environmental conditions, including vertical wind shear, synoptic-scale weather systems, Asia monsoons, easterly waves, Rossby wave energy dispersion, and so on (e.g., Fudeyasu and Yoshida, 2018; Ma et al., 2019). The interactions across scales impose challenges on the predictability, data assimilation, and forecasts for typhoons. Thus, this study describes the results of a cycling mesoscale EnKF system combined with the Advanced Research Weather Research and Forecasting model (WRF) applied for most of the 2016 WNP typhoon season. The WRF/EnKF system produces an 80-member ensemble analysis every 6 h. For each storm during the experimental period, a 5-d deterministic forecast is launched from the ensemble mean analysis every 6 h; and a 5-d ensemble forecast is obtained from the 80-member ensemble analyses for seven typhoons whose intensities are underestimated. The application of the combined data assimilation and regional forecasting system over a nearly complete typhoon season provides a large sample for assessing its potential benefit to TC forecasts.

This paper is organized as follows. Section 2 describes the modeling and data assimilation system. An overview of the 21 storms studied for the 2016 WNP typhoon season is provided in section 3. Section 4 discusses verifications for the cycling data assimilation system, while section 5 provides verifications for the 5-d deterministic forecasts. The performances of ensemble forecasts are evaluated in section 6. A summary and conclusion are provided in section 7.

2. Experimental design

Ensemble analyses are generated every 6 h by cycling a WRF/EnKF system from 0000 UTC 1 July to 1200 UTC 21

October. During this period, 21 TCs are simulated, whose categories and starting and ending dates are given by Table 1 and tracks are shown in Fig. 1. Within the duration of each TC, a 5-d forecast is launched every 6 h from the ensemble mean analysis. Due to limited computational resources, 5-d ensemble forecasts are launched for the typhoons whose rapid intensification processes are not well captured by the WRF/EnKF system and the ensemble mean analysis of the minimum sea level pressure (SLP) is higher than the observed value. Table 2 lists the typhoons and initial times for the 5-d ensemble forecasts.

The simulation of the 2016 WNP typhoon season uses WRF V3.4 (Skamarock et al., 2008). The model's domain 1, shown in Fig. 1, covers most of China and the WNP basin.

Table 1. The name, category, and duration for each TC in the 2016 WNP typhoon season assimilation.

| TC name	Category	Duration
Nepartak	Typhoon	0600 UTC 3 Jul – 0600 UTC 9 Jul
Lupit	Tropical Storm	0000 UTC 21 Jul – 1200 UTC 24 Jul
Mirinae	Severe Tropical Storm	0600 UTC 24 Jul – 1200 UTC 28 Jul
Nida	Severe Tropical Storm	1200 UTC 29 Jul – 0600 UTC 2 Aug
Omais	Severe Tropical Storm	1800 UTC 4 Aug – 0000 UTC 9 Aug
Conson	Tropical Storm	0600 UTC 8 Aug – 1200 UTC 14 Aug
Chanthu	Severe Tropical Storm	1200 UTC 13 Aug – 1200 UTC 17 Aug
Mindulle	Typhoon	1800 UTC 17 Aug – 1800 UTC 22 Aug
Dianmu	Tropical Storm	0000 UTC 18 Aug – 0600 UTC 19 Aug
Lionrock	Typhoon	1800 UTC 19 Aug – 1800 UTC 30 Aug
Kompasu	Tropical Storm	1800 UTC 19 Aug – 0000 UTC 21 Aug
Namtheun	Typhoon	1200 UTC 31 Aug – 0600 UTC 5 Sep
Meranti	Typhoon	1800 UTC 8 Sep – 0000 UTC 15 Sep
Malakas	Typhoon	0000 UTC 12 Sep – 1200 UTC 20 Sep
Rai	Tropical Storm	0000 UTC 12 Sep – 0000 UTC 13 Sep
Megi	Typhoon	0000 UTC 23 Sep – 0600 UTC 28 Sep
Chaba	Typhoon	0000 UTC 28 Sep – 0600 UTC 5 Oct
Aere	Severe Tropical Storm	0600 UTC 6 Oct – 0000 UTC 10 Oct
Songda	Typhoon	1200 UTC 8 Oct – 0000 UTC 13 Oct
Sarika	Typhoon	1200 UTC 12 Oct – 1200 UTC 19 Oct
Haima	Typhoon	1800 UTC 14 Oct – 1200 UTC 21 Oct
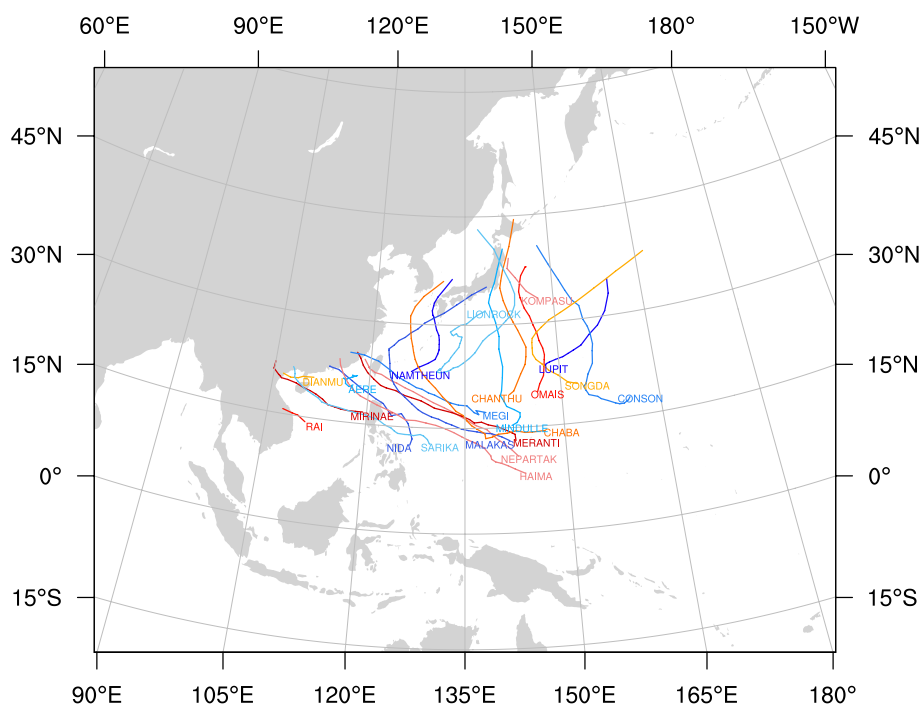
Fig. 1. Tropical cyclone tracks for each of the 21 TCs studied here. See Table 1 for a detailed list of storms. The colors are used to differentiate the TC tracks.

Table 2. Forecast initialization times for the typhoons whose ensemble mean analysis of the minimal sea level pressure is higher than the observation.

TC name	Forecast initialization times	Min. SLP
Nepartak	1200 UTC 3 Jul	911 hPa, 0600 UTC 6 Jul
Namtheun	1200 UTC 31 Aug	948 hPa, 0600 UTC 2 Sep
Meranti	0000 UTC 10 Sep	898 hPa, 0000 UTC 13 Sep
Chaba	0600 UTC 28 Sep	915 hPa, 0600 UTC 3 Oct
Songda	1200 UTC 8 Oct	926 hPa, 1800 UTC 11 Oct
Sarika	1200 UTC 12 Oct	944 hPa, 1200 UTC 17 Oct
Haima	1800 UTC 14 Oct	913 hPa, 1800 UTC 18 Oct

This domain could include the track of all storms in the WNP basin. 18-km horizontal grid spacing is used for domain 1, with 520 × 660 grid points. When the Joint Typhoon Warning Center (JTWC) announces an advisory TC position, a vortex following domain 2 of 6-km horizontal grid spacing and 180 × 180 grid points is set for 6-hourly cycling. When a 5-day forecast is produced for both the ensemble mean and the ensemble members, an additional high-resolution domain is set into the vortex following domain 2. This high-resolution domain has 2-km horizontal grid spacing and 300 × 300 grid points. There are 57 vertical levels, with the model top at 10 hPa. The implementation of WRF has the following components: the WRF 6-class microphysics scheme (Hong and Lim, 2006), the Yonsei University (YSU) planetary boundary layer (PBL) scheme (Hong et al., 2006), the Noah land surface model (Ek et al., 2003), the Rapid Radiative Transfer Model (RRTM) longwave scheme (Mlawer et al., 1997), and the RRTM shortwave scheme (Iacono et al., 2008). The cumulus parameterization uses the Tiedtke cumulus scheme (Tiedtke, 1989; Zhang et al., 2011a) and is only used for the 18-km domain 1.

The ensemble initial conditions (ICs) at the starting date 0000 UTC 1 July are generated from the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) analysis of 0.25° resolution using the fixed-covariance perturbation technique of Torn et al. (2006). This technique produces random perturbations that sample the NCEP background error covariance by use of the WRFDA-3DVAR (Barker et al., 2012), and the initial ensembles are generated by adding these random perturbations to the GFS analysis. Ensemble lateral boundary conditions (LBCs) are generated in a similar manner to the ensemble ICs, except that the NCEP GFS forecasts valid at the appropriate times are used. LBCs for the 5-d forecast launched from the ensemble mean analysis are produced using the NCEP GFS forecasts launched from the same analysis date but at the appropriate times.

Observations from the NCEP global data assimilation system (GDAS) are assimilated every 6 h, including conventional in situ observations, cloud-motion vectors, and remotely sensed satellite radiances from the Advanced Microwave Sounding Unit-A (AMSUA), High resolution IR Sounder (HIRS), Atmospheric IR sounder (AIRS), and Microwave Humidity Sounder (MHS). For the assimilated radiance observations, a thinning mesh of 60 km is used, at which the radiance observation errors are assumed to be uncorrelated (Lin et al., 2017). In addition, the JTWC advisory minimum SLP at the observed position (latitude and longitude of the lowest sea level pressure) are assimilated. The observation error variances are the same as those used in the NCEP GDAS.

An ensemble square-root filter (EnSRF; Whitaker et al., 2008) that is very similar to the NOAA operational EnKF for the NCEP GFS in 2016 (https://dtcenter.org/sites/default/files/community-code/enkf/docs/users-guide/EnKF_UserGuide_v1.3.pdf) is used to assimilate the observations. Note that, recently, a localized ensemble transform Kalman filter with model space localization (Hunt et al., 2007; Lei et al., 2018) has been implemented for the NCEP GFS. Ensemble size is 80. The observation prior Hx^b, where H is the observation forward operator and x^b is the model ensemble background or prior, is computed by the "observer" portion of the Gridpoint Statistical Interpolation system (GSI; Wu et al., 2002; Kleist et al., 2009). The "observer" runs separately for the ensemble mean and the 80 ensemble members, and the obtained observation ensemble priors are used by the EnKF. Bias correction of radiance observations is computed using the EnKF based on iterated analysis error covariance (Miyoshi et al., 2010).

The EnKF system requires additional steps designed to overcome sampling errors that result from using a limited ensemble size and also account for model error. Sample covariances derived from the ensembles are localized by the Gaspari and Cohn (1999) localization function that is an approximately Gaussian fifth-order piecewise continuous polynomial function. Observation space localization (Houtekamer and Mitchell, 1998) is applied, by which observation impact is tapered to 0 at 1000 km in the horizontal and 1.5 ln (hPa) in the vertical. The vertical locations of the non-local radiance observations are assigned to the vertical level at which the weighting function maximizes. At each assimilation time, the ensemble deviations from the ensemble mean are inflated posterior to assimilation using the relaxation posterior ensemble spread to prior ensemble spread (relaxation-to-prior spread; Whitaker and Hamill, 2012) with a relaxation coefficient of 1.15. To maintain an appropriate ensemble spread, a relaxation coefficient larger than 1.0 is necessary, which forces posterior ensemble spread to be larger than prior ensemble spread (Schwartz and Liu, 2014; Schwartz, 2016). During a 10-day test period, a group of sen-

sitivity experiments with varying localization and inflation parameters are conducted. The localization and inflation parameters are chosen based on the sensitivity experiments that provide the smallest 6-h prior errors comparing to the conventional observations and JTWC advisory TC data.

To overcome the underestimation of TC initial intensity given an 18-km horizontal resolution (Torn, 2010) and avoid the technical challenges associated with moving nested domains for each ensemble member (like each ensemble member having a different nest location and interpolation from the coarse grid), the non-cycled nesting procedure of Cavallo et al. (2013) is adopted. Given our domain configuration, the EnKF data assimilation is done on the 18-km resolution domain 1, but not on the 6-km vortex-following domain 2. However, the 18-km resolution domain 1 benefits from the increased resolution for storm intensity, since the 6-h forecast on the 18-km grid is averaged from the 6-km grid where the 6-km grid exists. Cavallo et al. (2013) noted a 22% error reduction in the 6-h prior of TC minimum SLP and a moderate error reduction in the 6-h prior of TC maximum wind speed when using non-cycled nests within the data assimilation compared to no nests within the data assimilation.

3.   Overview of cases

A short review of the 21 TCs in the 2016 WNP typhoon season during the assimilation period is provided in this section (Fig. 1). Most of the TCs formed in the main development region, while a few formed at relatively western longitudes. More than half of the TCs moved westward, and some of them turned poleward at some point in their lives, while several TCs directly moved poleward. Besides these TC tracks, Typhoon Lionrock first moved southwest, then turned northeast, and at last turned towards the northwest due to a high pressure system located east of Japan.

The lifetime durations and categories of the 21 TCs are given in Table 1. The lifetime durations of the storms ranged from 24 h (Rai) to 11 days (Lionrock). There were short-lived storms, such as Rai, Nida, Lupit, and Aere, and long-lived storms, such as Sarika, Chaba, Malakas, and Lionrock. According to the Regional Specialized Meteorological Centre (RSMC) Tokyo's tropical cyclone intensity scale (Typhoon Committee, 2015), there were five tropical storms (TS) (Lupit, Conson, Dianmu, Kompasu, and Rai), whose maximum wind speeds were less than 24.2 m s^{-1}; and there were five severe tropical storms (STS) (Mirinae, Nida, Omais, Chanthu, and Aere), whose maximum wind speeds were less than 32.4 m s^{-1}. The remaining storms reached typhoon (TY) scale, with maximum wind speeds larger than 32.4 m s^{-1}. Based on the Kaplan-DeMaria criteria (Kaplan and DeMaria, 2003), all the typhoons were characterized by at least one instance of rapid intensification (maximum wind speed increasing more than 15.4 m s^{-1} per 24 h). Both Malakas and Songda met the rapid intensification criteria twice during their lifetimes. A wide variety of TC intensities are captured in this study.

4.   Cycling verification

In this section, the cycling WRF/EnKF system is evaluated. Ensemble priors, i.e., 6-h ensemble forecasts, are verified against the JTWC advisory TC track position and intensity estimates that are the same as the assimilated quantities. The simulated TC position is given by the location of the minimum SLP. Figure 2 shows the TC track RMS errors, ensemble spread, and biases (observation minus forecast) for tropical storms (TS), severe tropical storms (STS), typhoons (TY), and all TCs (ALL), respectively. Figure 3 represents the RMS errors and biases of TC minimum SLP and maximum wind speed for different TC categories. The statistical
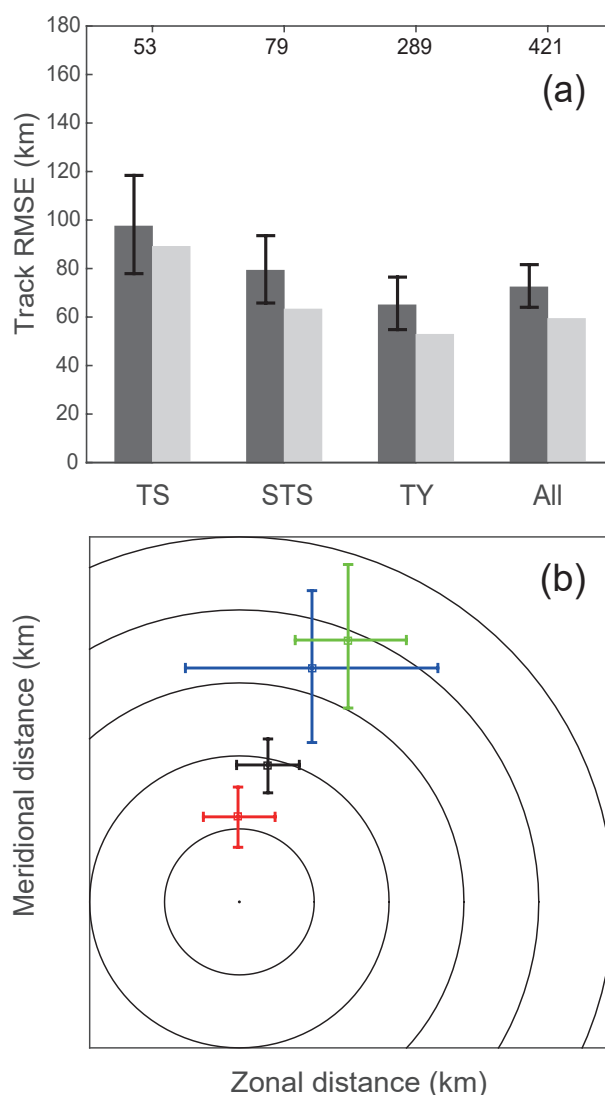

Fig. 2. (a) 6-h forecast ensemble mean RMS error (dark gray bar) and ensemble spread (light gray bar) of TC positions as a function of TC intensity. The number of verification times is given along the top. (b) 6-h forecast ensemble mean bias of TC positions for tropical storm (TS, blue line), severe tropical storm (STS, green line), typhoon (TY, red line), and all storms (ALL, black line). The range rings denote 10-km intervals. Error bars denote the 5% and 95% percentiles determined from bootstrap resampling.

significances of RMS errors and biases are determined using a bootstrap resampling with replacement method (Efron and Tibshirani, 1993). Each error distribution is resampled 1000 times with a bootstrap process. Error bars denote 5% and 95% percentiles; thus, only significance levels greater than 90% are considered statistically significant.

For all TC categories, the ensemble spread of TC position is slightly smaller than the RMS error. As shown by Anderson (1996) and Hamill (2001), verification of an ensemble forecast can be affected by the used imperfect observations. The ensemble spread here does not take into account the observation uncertainties; thus, the generally consistent TC position error with ensemble spread indicates a well calibrated ensemble being provided by the WRF/EnKF system (Houtekamer et al., 2005). The TC position error and spread of ALL are dominated by those of TY, because TY has many more samples than TS and STS. The position error and spread are inversely proportional to the TC intensity. This may be a result of poorly defined TC centers for weak TCs. Note that the 6-h track errors are larger than the operational forecast errors because there are ensemble members with large track errors, especially when the vortex is weak (figures are not shown). The zonal and meridional position biases of TY are also much smaller than those of TS and STS. For all instances, there are larger meridional position biases than zonal position biases, which is consistent with the persistently positive v-wind biases (figures are not shown).

Unlike for TC position, the ensemble spread of minimum SLP is insufficient to present the ensemble mean error, espe-cially for TY. Similarly, the magnitude of TC maximum wind speed suffers from a large mismatch between the ensemble mean error and ensemble spread. Thus, the ensemble spread is not calibrated well for TC minimum SLP and maximum wind speed. TY has a much larger minimum SLP error than TS and STS, but the ensemble spread does not increase much from TS/STS to TY. Compared to TS and STS, the larger minimum SLP error of TY is also accompanied by a larger bias. TS and STS have similar errors of approximately 7 hPa, and similar biases of approximately +/− 0.75 hPa; but TY has an increased error of 22.21 hPa and an increased bias of −13.12 hPa. Although TS has a similar minimum SLP error and absolute bias to STS, TS has a larger maximum wind speed error and bias than STS. The error (bias) of maximum wind speed is 12.59 m s^{-1} (−9.06 m s^{-1}) for TS, 7.37 m s^{-1} (−3.24 m s^{-1}) for STS, and 14.61 m s^{-1} (9.5 m s^{-1}) for TY. Therefore, given a 6-km horizontal grid spacing, the WRF model is still unable to resolve the large gradients of TY's wind and mass fields. The WRF/EnKF system overestimates the maximum wind speed for weak storms like TS and STS, but underestimates the minimum SLP and maximum wind speed for strong storms like TY.

5. Forecast verification

In this section, WRF/EnKF TC track and intensity deterministic forecasts launched from the ensemble mean analyses every 6 h are evaluated against the JTWC advisory TC track
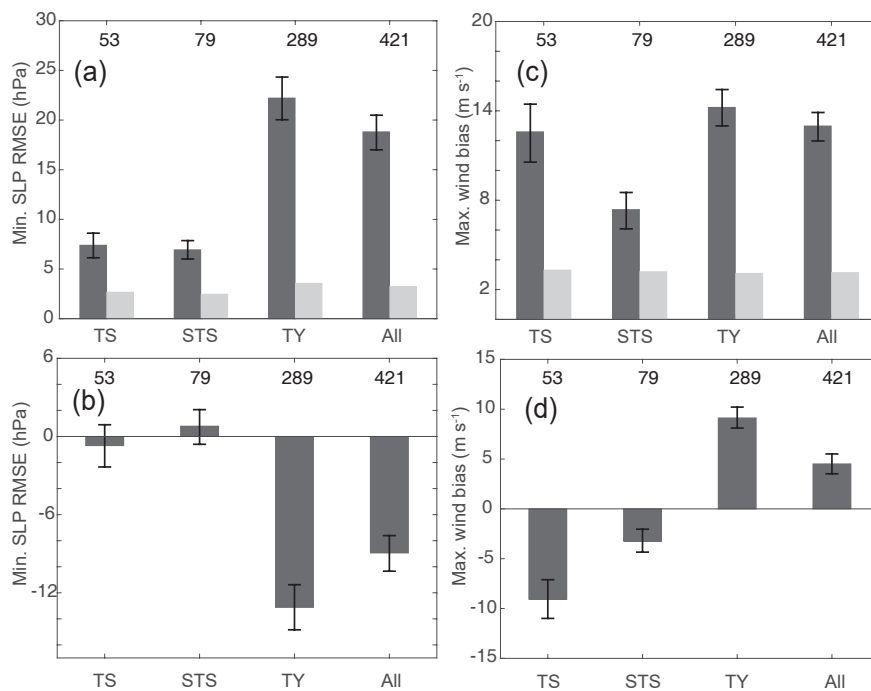

Fig. 3. (a) 6-h forecast ensemble mean RMS error (dark gray bar) and ensemble spread (light gray bar) of TC minimum SLP as a function of TC intensity. The number of verification times is given along the top. (b) 6-h forecast ensemble mean bias of TC minimum SLP as a function of TC intensity. (c) As in (a), but for the TC maximum wind speed. (d) As in (b), but for the TC maximum wind speed. Error bars denote the 5% and 95% percentiles determined from bootstrap resampling.

position and intensity estimates. For comparison, errors of the NCEP Global Ensemble Forecast System (GEFS) forecast (Zhou et al., 2017) and European Centre for Medium-Range Weather Forecasts (ECMWF) Ensemble Prediction System (EPS) global forecast (details are available online at https://www.ecmwf.int/sites/default/files/elibrary/2012/14557-ecmwf-ensemble-prediction-system.pdf) are also computed at the equivalent time. The NCEP GEFS has one control member and 20 ensemble members, and the 21 total ensemble members are integrated with the GFS model at approximately 34-km horizontal grid spacing for the first 8-d forecasts. The ECMWF EPS contains one control member and 50 ensemble members, and the 51 total ensemble members have a grid spacing of about 32 km through the 10-d forecast. Since WRF/EnKF uses a single forecast from the ensemble mean analysis, deterministic forecasts from the control members of the NCEP GEFS and ECMWF EPS are used for comparison in this section. Please note that WRF/EnKF has much higher horizontal grid resolutions (18 km, 6 km, and 2 km) than NCEP and ECMWF for the vortexes.

Instead of launching a deterministic forecast from the ensemble mean analysis, Torn (2010) and Cavallo et al. (2013) used a single member of the WRF/EnKF analysis

ensemble for the deterministic forecast in order to avoid the overly smoothed depiction of the TC mass and wind fields which result from ensemble averaging. The single-member initial condition is chosen as the closest member to the ensemble mean analysis based on normalized latitude, longitude, and minimum SLP differences. To show the differences of a single forecast from the ensemble mean analysis or from a single ensemble member, a 5-d deterministic forecast is run from the ensemble mean analysis and a randomly chosen ensemble member every 6 h for 5 TS, 5 STS, and 8 TY. Results show that the deterministic forecast from the ensemble mean analysis and that from a single ensemble member have similar errors of minimum SLP and maximum wind speed through the 5-d forecast period (figures are not shown). The forecast from the ensemble mean analysis has track RMS errors similar to those from a single ensemble member within 48 h, and the former has smaller track RMS errors than the latter beyond 48 h, although the differences are not statistically significant at a 90% confidence interval. Therefore, the deterministic forecast from the WRF/EnKF ensemble mean analysis is used here and compared to the control forecast of NCEP GEFS and ECMWF EPS.

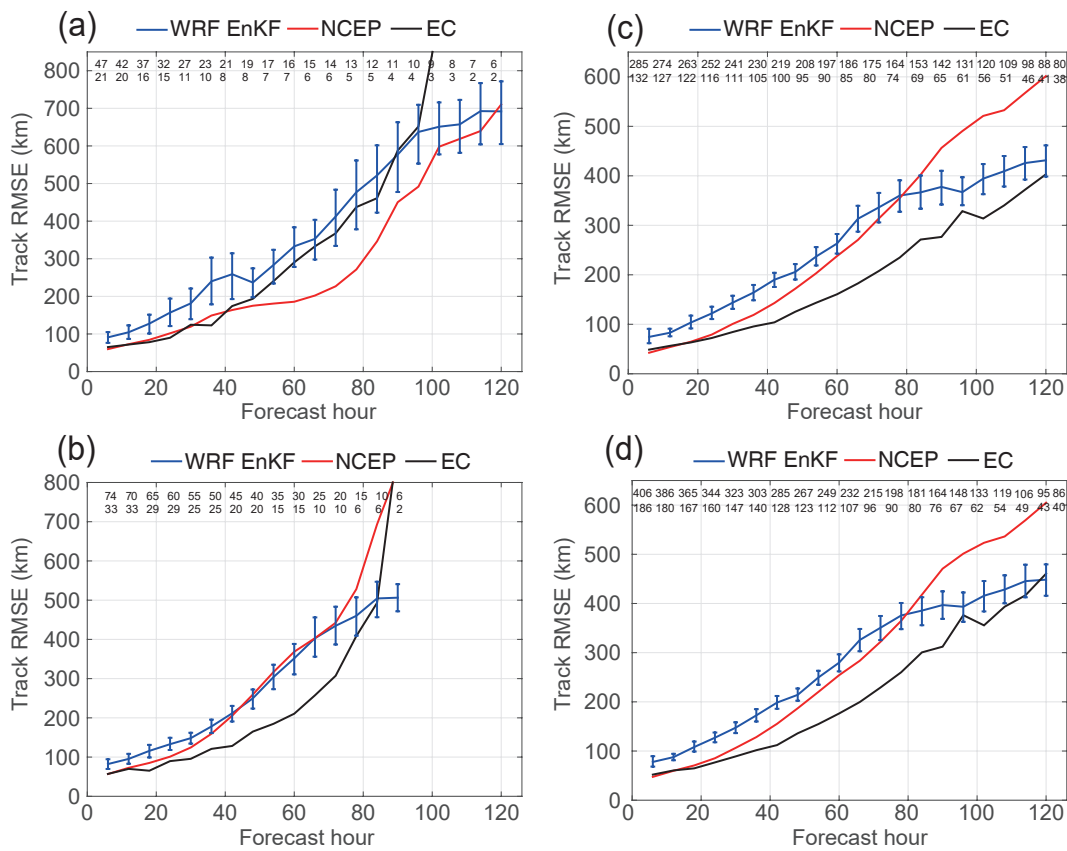Figure 4 shows the track RMS errors as a function of

Fig. 4. RMS error of TC positions for 5-d forecast as a function of forecast hour for (a) TS, (b) STS, (c) TY, and (d) ALL. The blue solid line denotes the WRF/EnKF forecast launched from ensemble mean analysis, and the red and black solid lines are the NCEP GFS forecast and ECMWF forecast, respectively. The number of verification times for WRF/EnKF and NCEP is given by the first row along the top, and the number of verification times for ECMWF is given by the second row along the top. Error bars denote the 5% and 95% percentiles determined from bootstrap resampling.

forecast hours for TS, STS, TY, and ALL, respectively. The statistical significance of RMS error is denoted by the error bars showing 5% and 95% percentiles. The number of verification times is denoted on the top of the figure. The sample sizes of WRF/EnKF and NCEP are larger than that of ECMWF because WRF/EnKF and NCEP GEFS forecasts are computed every 6 h while the ECMWF EPS forecast is computed every 12 h. For TS, STS, TY, and ALL, WRF/EnKF deterministic forecasts generally have significantly larger track RMS errors than NCEP and ECMWF forecasts at short forecast lead times. The large initial track errors of WRF/EnKF are mainly resulted from the ensemble mean initial conditions, which have degraded initial position due to the ensemble averaging and have a subsequently fast error growth rate for the first 6 h due to imbalanced mass and wind fields of the ensemble mean. WRF/EnKF has insignificantly larger track errors or significantly smaller track errors than NCEP and ECMWF at long forecast lead times. Moreover, the meridional and zonal track biases of WRF/EnKF,

NCEP, and ECMWF generally increase with forecast lead times, while the meridional track biases are larger than the according zonal track biases, especially at longer forecast lead times for TS and STS (Fig. 5). The meridional and zonal position error differences at longer forecast lead times are consistent with those from 6-h priors, which is possibly due to the persistently positive v-wind biases. Consistent with track RMS errors, WRF/EnKF deterministic forecasts in general have smaller track biases than NCEP and ECMWF at long forecast lead times. Thus, the regional simulation cannot better represent the large-scale environment compared to the global simulation, especially at short forecast lead times.

Figures 6 and 7 display RMS errors and biases of minimum SLP as a function of forecast hour for different categories, respectively. For TS, WRF/EnKF deterministic forecasts produce larger RMS errors of minimum SLP than NCEP and ECMWF until about 60 h, and the error differences between WRF/EnKF and NCEP/ECMWF are not statistically
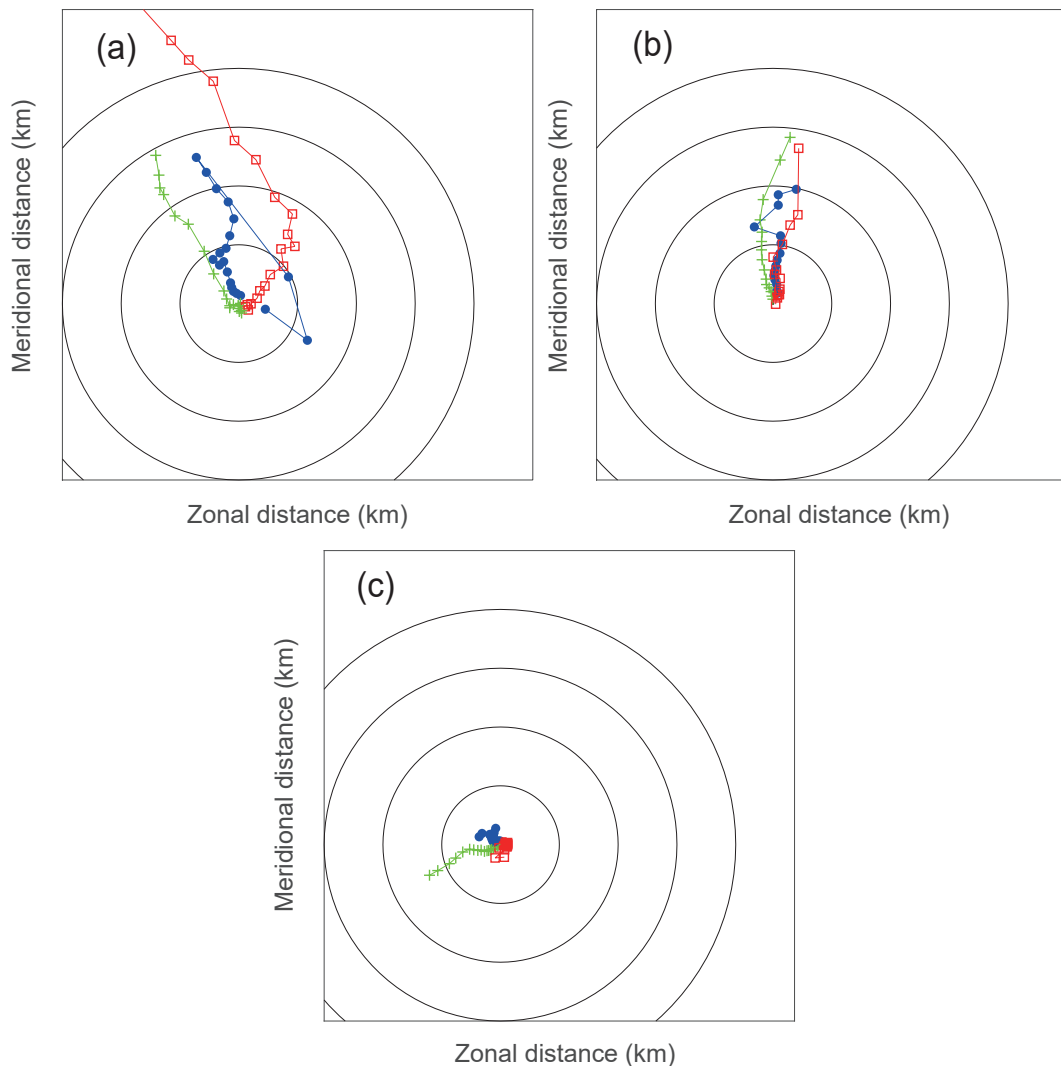

Fig. 5. Biases of TC positions from 5-d forecasts for (a) TS, (b) STS, and (c) TY. The blue dot, green plus, and red square denote WRF/EnKF forecasts launched from the ensemble mean analyses, NCEP GEFS control forecasts, and ECMWF EPS control forecasts, respectively. The range rings denote 200-km intervals.
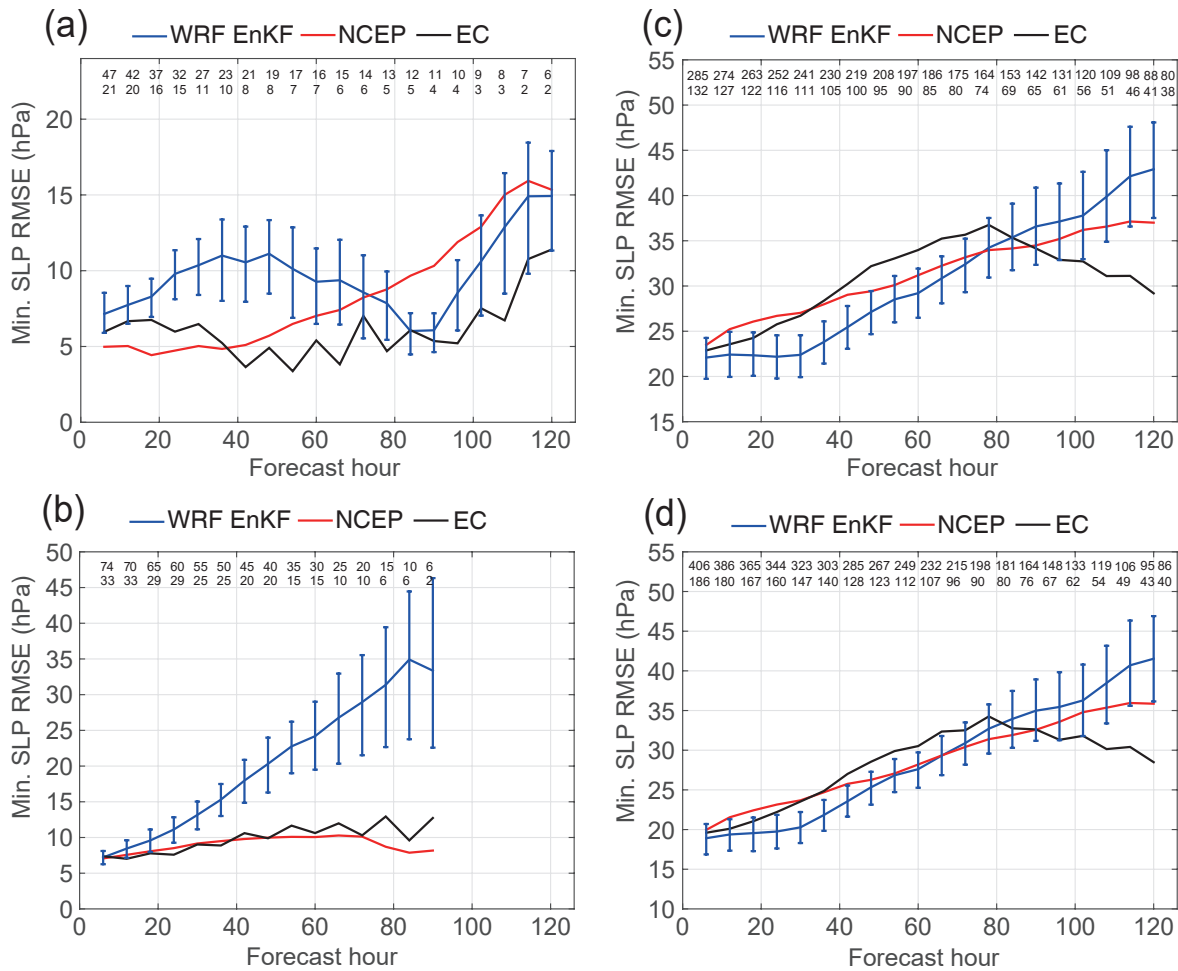
Fig. 6. Same as Fig. 4, except for RMS error of TC minimum SLP.

significant at long forecast lead times. WRF/EnKF and NCEP (ECMWF) forecasts have negative (positive) biases of minimum SLP at short forecast lead times, while WRF/EnKF and ECMWF (NCEP) forecasts have negative (positive) biases for forecast lead times longer than 48 h. For STS, NCEP and ECMWF forecasts obtain similar RMS errors of minimum SLP, which are smaller than those of WRF/EnKF. Consistently, NCEP and ECMWF forecasts have similar biases of minimum SLP, which have smaller magnitudes than the positive biases of WRF/EnKF. For TY, WRF/EnKF deterministic forecasts generally have significantly smaller RMS errors of minimum SLP than NCEP and ECMWF until forecast lead times of 42 h and 72 h, respectively. Consistent with the RMS errors, NCEP and ECMWF forecasts have similar negative biases of minimum SLP, which have much larger magnitudes than the biases of WRF/EnKF. The RMS errors and biases of minimum SLP for ALL are similar to those of TY since the samples from TY are larger than those of TS and STS and thus the features of TY dominate.

The RMS errors and biases of maximum wind speed at 10 m as a function of forecast hour for different categories are shown in Figs. 8 and 9, respectively. For TS, WRF/

EnKF deterministic forecasts have similar RMS errors of maximum wind speed to NCEP and ECMWF within 36 h but larger RMS errors than NCEP and ECMWF beyond 48 h. Similar biases of maximum wind speed are obtained for WRF/EnKF, NCEP, and ECMWF forecasts. For STS, WRF/EnKF deterministic forecasts have larger RMS errors of maximum wind speed than NCEP and ECMWF. Meanwhile, WRF/EnKF deterministic forecasts have negative biases of maximum wind speed, and the magnitudes are larger than those from NCEP and ECMWF. Different from the weak storms, WRF/EnKF deterministic forecasts produce significantly smaller RMS errors of maximum wind speed than NCEP and ECMWF for TY. The error differences among WRF/EnKF, NCEP, and ECMWF beyond 90 h are generally not statistically significant. Consistently, WRF/EnKF deterministic forecasts produce positive biases of maximum wind speed, and the magnitudes are much less than those of NCEP and ECMWF. Similar results to the comparisons for TY are obtained for ALL.

Therefore, although WRF/EnKF deterministic forecasts generally have larger RMS errors of minimum SLP and maximum wind speed than NCEP and ECMWF for weak storms like TS and STS, WRF/EnKF deterministic forecasts produce
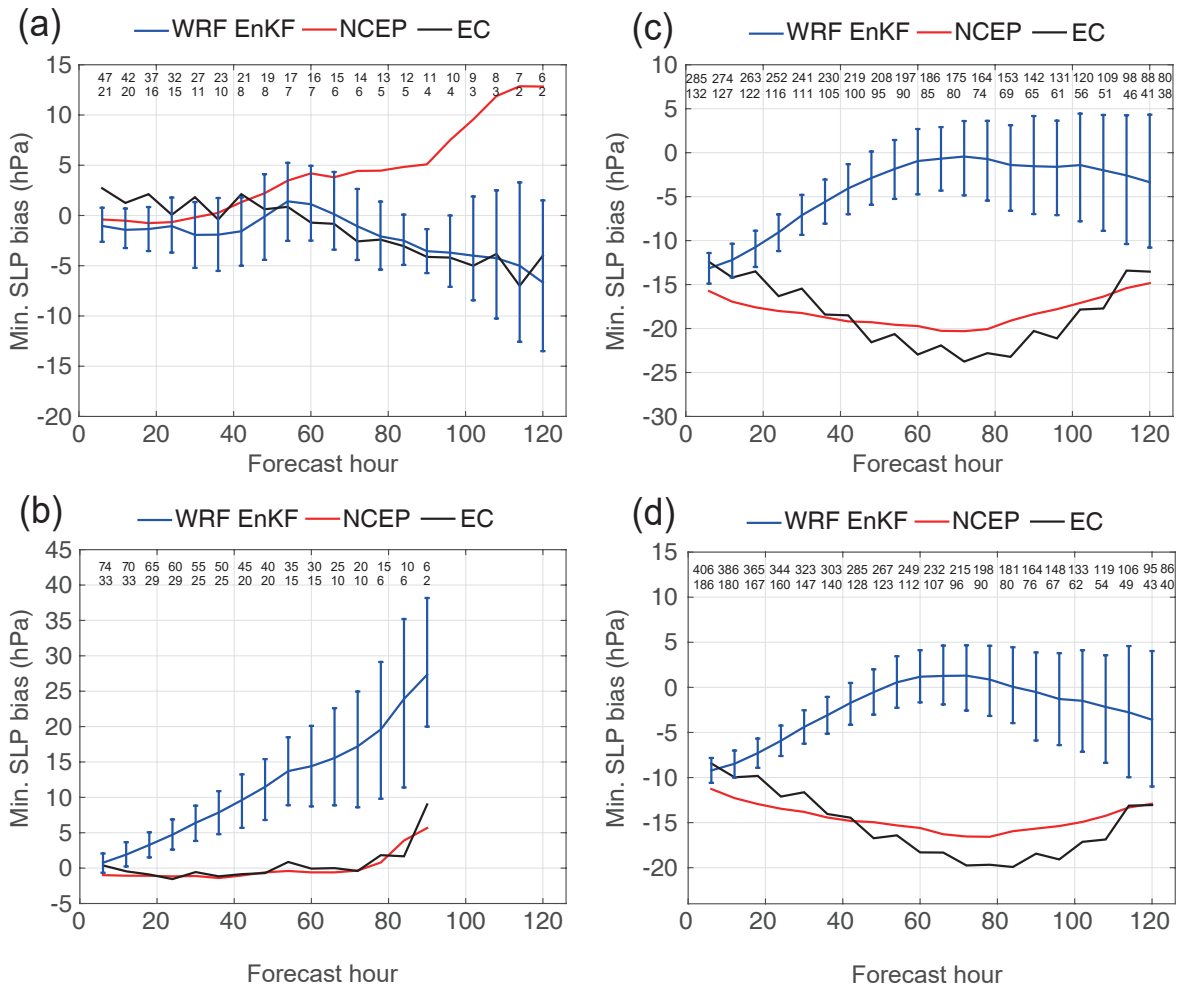
Fig. 7. Same as Fig. 4, except for bias of TC minimum SLP.

smaller RMS errors of intensity than NCEP and ECMWF for strong storms like TY. For weak storms, WRF/EnKF deterministic forecasts often have positive biases of minimum SLP and negative biases of maximum wind speed, which indicates overestimation of the intensity. For strong storms, WRF/EnKF deterministic forecasts often have negative biases of minimum SLP and positive biases of maximum wind speed, which indicates underestimation of the intensity. The underestimation of TY intensity is much better mitigated for WRF/EnKF compared to NCEP and ECMWF. There are environmental factors that may have influences on the vortex intensity, including temperature, specific humidity, vertical wind shear, etc. Figure 10 shows the profiles of differences for the mean specific humidity between WRF/EnKF and ECMWF forecasts. For each forecast, the mean specific humidity is averaged over an annulus that is centered around each vortex location with outer and inner circle radii of 5° and 2°, respectively. For weak storms, WRF/EnKF produces larger values of specific humidity below 500 hPa than ECMWF, which contributes to the intensity overestimation of weak vortexes for the WRF/EnKF. For strong storms, WRF/EnKF also has moister conditions than ECMWF, which explains the much better mitigated underestimation

of TY intensity of WRF/EnKF compared to ECMWF.

6. Ensemble forecast performance

Performances of the ensemble forecasts for the typhoons listed in Table 2 are evaluated in this section. The mean absolute errors of the ensemble-mean position and intensity forecasts are compared with the mean ensemble spread. The consistency between these two quantities indicates that the ensemble contains an appropriate amount of variance, while the inconsistency between these two quantities indicates a lack of growing modes or reflecting model biases. For comparison, the mean absolute errors and ensemble spreads of NCEP GEFS and ECMWF EPS forecasts are also computed, as shown in Fig. 11.

For TC track, the mean absolute errors and ensemble spreads of WRF/EnKF, NCEP, and ECMWF ensemble forecasts are comparable within 30 h. Beyond that, the mean absolute errors of WRF/EnKF and NCEP, especially NCEP, increase at a higher rate than the ensemble spread; but the ensemble track forecasts of ECMWF have an even larger ensemble spread than the error. Thus, the ensemble track forecasts of WRF/EnKF and NCEP are underdispersive, while
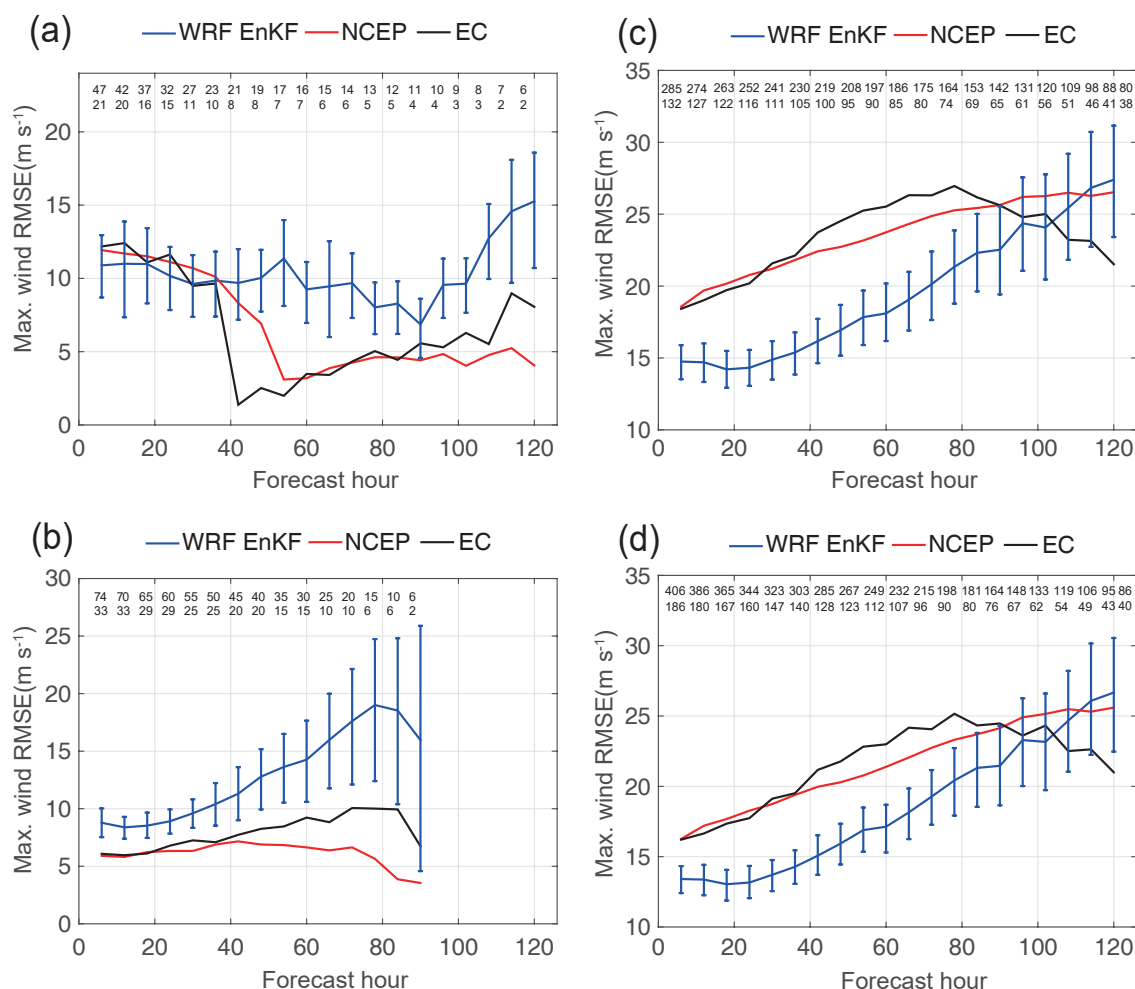
Fig. 8. Same as Fig. 4, except for RMS error of TC maximum wind speed.

the ensemble track forecasts of ECMWF are overdispersive. For minimum SLP and maximum wind speed, WRF/EnKF, NCEP, and ECMWF ensemble forecasts contain appropriate variance when the forecast hour is less than 18 h. Beyond that, all three ensemble forecasts have spread deficiency. Compared to the ensemble forecasts of NCEP and ECMWF, WRF/EnKF ensemble forecasts have smaller intensity errors but contain larger ensemble spread. Therefore, WRF/EnKF ensemble forecasts lack variance for TC track (Puri et al., 2001; Magnusson et al., 2008; Torn, 2010), but WRF/EnKF provides better intensity ensemble forecasts than NCEP and ECMWF. Numerous factors have impacts on the ensemble forecasts, like the ensemble initial conditions, numerical model and parameterization schemes for different physical processes, model error representations, etc. A systematic investigation for the ensemble forecasts of different ensemble systems will be reported in a future study.

To illustrate the ensemble performance in detail, ensemble forecasts for typhoon Meranti from 0000 UTC 10 September and typhoon Sarika from 1200 UTC 12 October by the WRF/EnKF, NCEP, and ECMWF are shown in Figs. 12 and 13, respectively. The track ensemble forecasts for typhoon Meranti by WRF/EnKF, NCEP, and ECMWF all

have the ensemble plumes centered on the observed value; thus, typhoon Meranti appears predictable. The ensemble forecasts of the minimum SLP and maximum wind speed by the NCEP and ECMWF show some degree of confidence, but they are far from the observed intensity. The intensity ensemble forecasts by WRF/EnKF are much closer to the observed values than the ensemble forecasts by NCEP and ECMWF; WRF/EnKF predicts a central pressure that is approximately 40 hPa lower than NCEP and ECMWF and a wind speed that is about 20 m s^{-1} stronger than NCEP and ECMWF. Typhoon Sarika also appears predictable from the track ensemble forecasts, although WRF/EnKF shows less predictability than NCEP and ECMWF. Similar to typhoon Meranti, the intensity ensemble forecasts for typhoon Sarika by NCEP and ECMWF are far from the observed value, and all ensemble members fail to capture the re-intensification process. However, the intensity ensemble forecasts by WRF/EnKF have the ensemble plume around the observed intensity, and most ensemble members predict the re-intensification process. At forecast lead times from 12 h to 120 h, WRF/EnKF ensemble members for typhoons Meranti and Sarika have moister conditions than ECMWF at low levels (figures are not shown), which could explain the stronger vor-
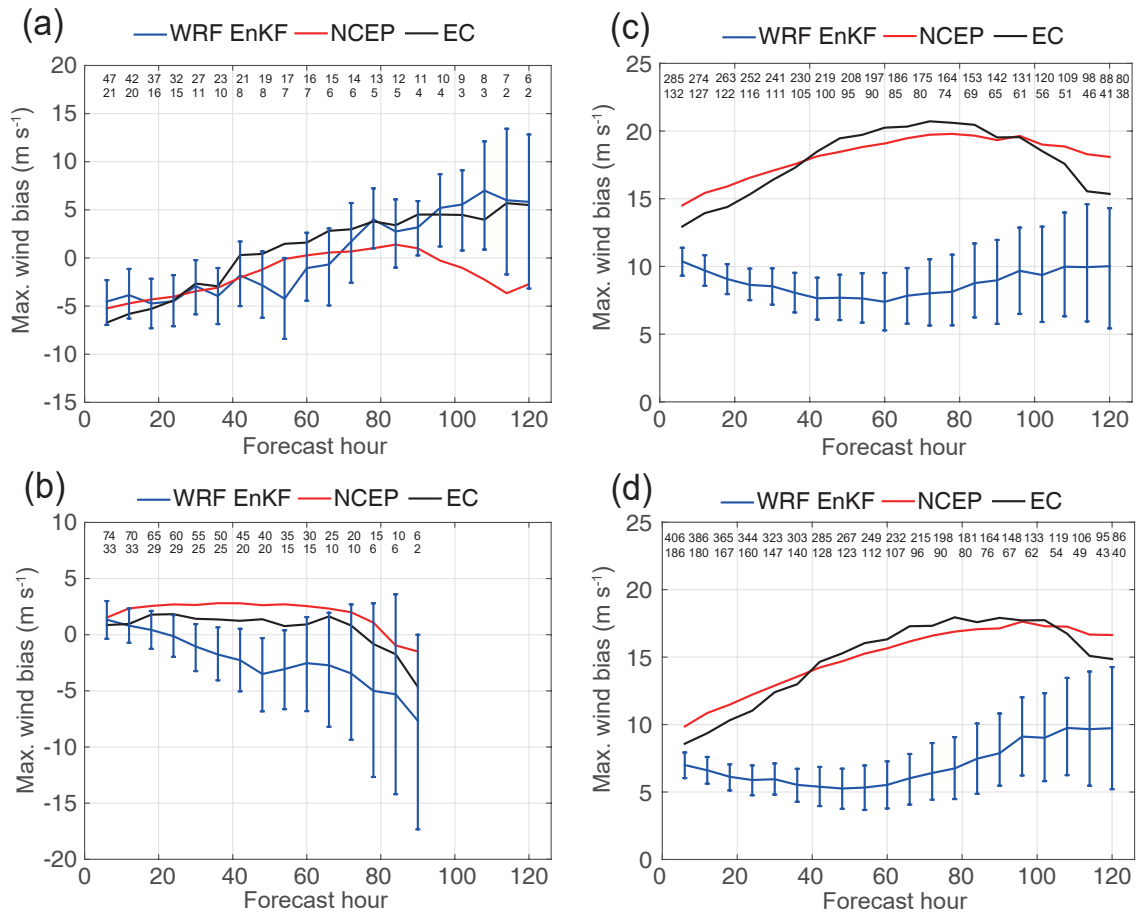
Fig. 9. Same as Fig. 4, except for bias of TC maximum wind speed.
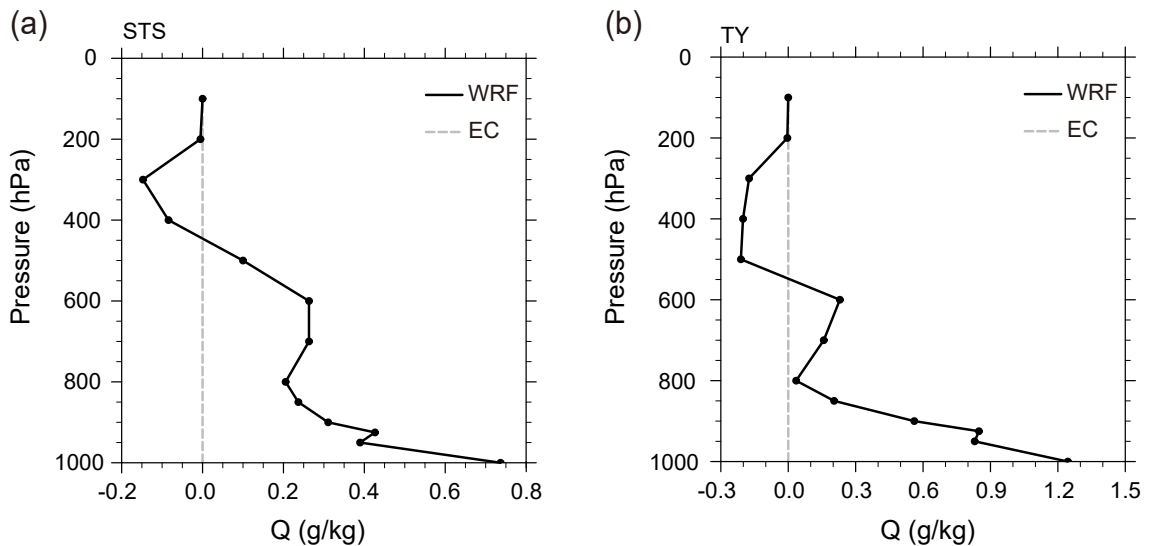

Fig. 10. Profiles of the differences of the mean specific humidity between WRF/EnKF and ECMWF 48-h forecasts for (a) STS and (b) TY. For each forecast, the mean specific humidity is the averaged specific humidity over an outer circle centered around each vortex with a 5° radius minus that over an inner circle with a 2° radius.

texes predicted by WRF/EnKF. Therefore, consistent with previous error statistics based on the mean absolute error and ensemble spread, WRF/EnKF has better intensity predictability than NCEP and ECMWF.

7. Conclusions

This study describes the performance of a cycling WRF /EnKF system during most of the 2016 WNP typhoon sea-

son. Conventional in situ data, radiance observations, and TC minimum SLP are assimilated every 6 h using an 80-member ensemble. For the 21 storms during the experimental

period, a 5-d deterministic forecast is launched from the ensemble mean analysis every 6 h within the duration of each storm; and a 5-d ensemble forecast is produced from the ensemble analyses for 7 typhoons whose intensities are underestimated. The forecast errors are compared to the ECMWF and NCEP operational models.
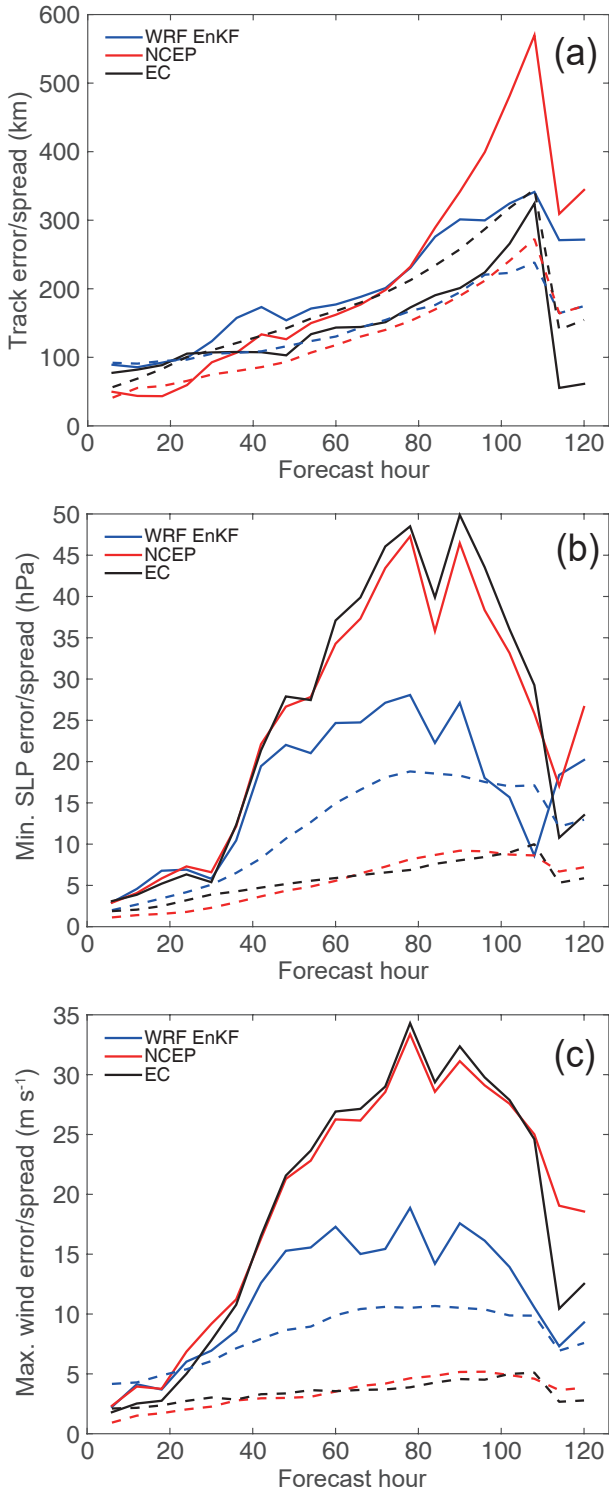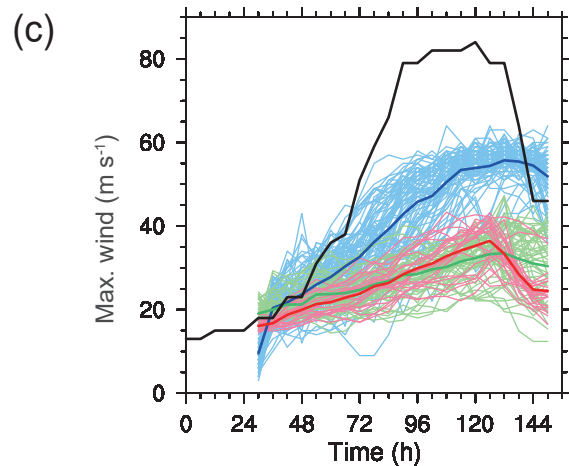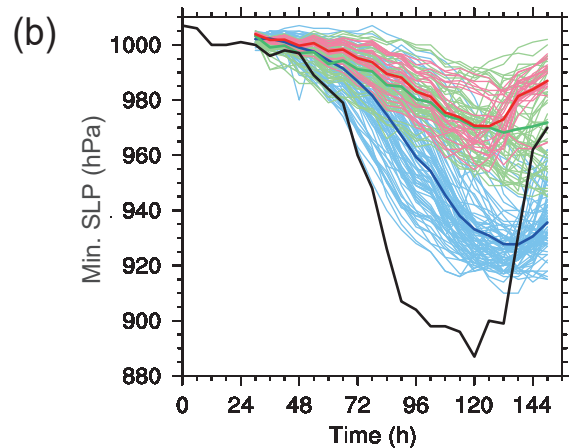

Fig. 11. The mean absolute error (solid) and ensemble spread (dashed) from ensemble forecasts of typhoons listed in Table 2 as a function of forecast hour for (a) track, (b) minimum SLP, and (c) maximum wind speed. The blue lines denote the WRF/ EnKF forecast, and the red and black lines displays the NCEP GEFS and ECMWF EPS forecasts, respectively.
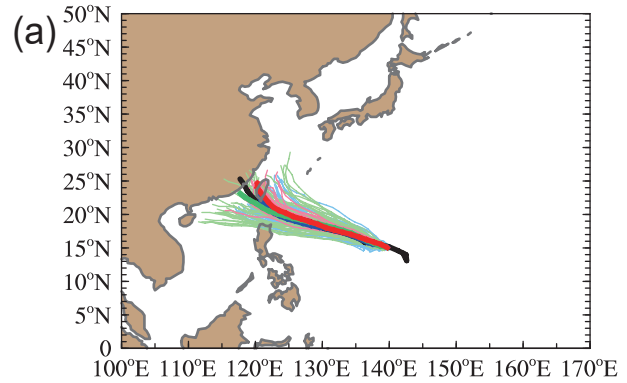
Fig. 12. 5-d ensemble forecast (a) tracks, (b) minimum SLP, and (c) maximum wind speed for typhoon Meranti from 0000 UTC 10 September. The thin blue, red, and green lines show the forecast values of the WRF/EnKF, NCEP, and ECMWF ensemble forecasts, respectively; and the thick lines denote the according ensemble mean. The black line denotes the observed value.

For all TC categories, the 6-h ensemble prior estimates of TC position from the WRF/EnKF system contain an appropriate amount of variance, while the TC intensity estimates are variance-deficient for all intensities, especially for TY. The TC position RMS error and spread are inversely proportional to the TC intensity, which may be a result of poorly defined TC centers for weak TCs. All TC instances are characterized by larger meridional position bias than zonal position

Fig. 13. Same as Fig. 12, except for typhoon Sarika from 1200 UTC 12 October.

bias. Category TY has much larger minimum SLP errors and biases than categories TS and STS, which indicates that a 6-km horizontal grid spacing is still unable to resolve the large gradients of TC wind and mass fields. Maximum wind speed errors and biases indicate that the WRF/EnKF system tends to overestimate maximum wind speed for TS and STS, but underestimate maximum wind speed for TY.

Compared to the NCEP and ECMWF operational control forecasts, the WRF/EnKF deterministic forecasts from the ensemble mean analyses often has larger TC track errors for all categories because the regional simulation cannot better represent the large-scale environment compared to the global simulation. A blending method that merges the analyses of global and regional models can be beneficial for TC track forecasting (Hsiao et al., 2015). The meridional track biases of WRF/EnKF, NCEP, and ECMWF are generally larger than the according zonal track biases for TS and STS, while WRF/EnKF and ECMWF produce much smaller zonal and meridional track biases than NCEP for TY. The WRF/EnKF deterministic forecasts exhibit smaller TC intensity errors for TY than the NCEP and ECMWF control forecasts, which is due to the higher grid resolution of the WRF/EnKF system; but the WRF/EnKF forecasts have larger TC intensity errors for TS and STS. The WRF/EnKF deterministic forecasts often have positive biases of minimum SLP and positive biases of maximum wind speed for weak storms, which means overestimation of the intensity. The NCEP and ECMWF control forecasts have negative biases of minimum SLP and positive biases of maximum wind speed for strong storms, which means underestimation of the intensity, but the WRF/EnKF deterministic forecasts produce smaller intensity biases than the NCEP and ECMWF control forecasts. Profiles of specific humidity differences between WRF/EnKF and ECMWF show that WRF/EnKF produces moister conditions than ECMWF for both weak and strong storms, which explains how WRF/EnKF overestimates intensity for weak storms and why WRF/EnKF has better mitigated underestimation of intensity for strong storms.

The ensemble forecasts from the WRF/EnKF system contain appropriate variance for TC track and intensity with short forecast lead times. With long forecast lead times, the ensemble forecasts of WRF/EnKF and NCEP are underdispersive for TC track, while the ensemble forecasts of ECMWF are overdispersive for TC track. The WRF/EnKF ensemble forecasts have smaller intensity errors but larger ensemble spread than the NCEP and ECMWF ensemble forecasts; thus, the WRF/EnKF system provides better TC intensity ensemble forecasts than NCEP and ECMWF, in terms of the comparison between amount of ensemble spread and mean absolute error. Moreover, the ensemble forecasts of WRF/EnKF can better capture the detailed intensity evolution than those of NCEP and ECMWF; thus, WRF/EnKF shows better intensity predictability than NCEP and ECMWF.

The large initial track errors of WRF/EnKF are possibly a result of fast error growth due to imbalances caused by data assimilation, which could be mitigated by appropriate ini-
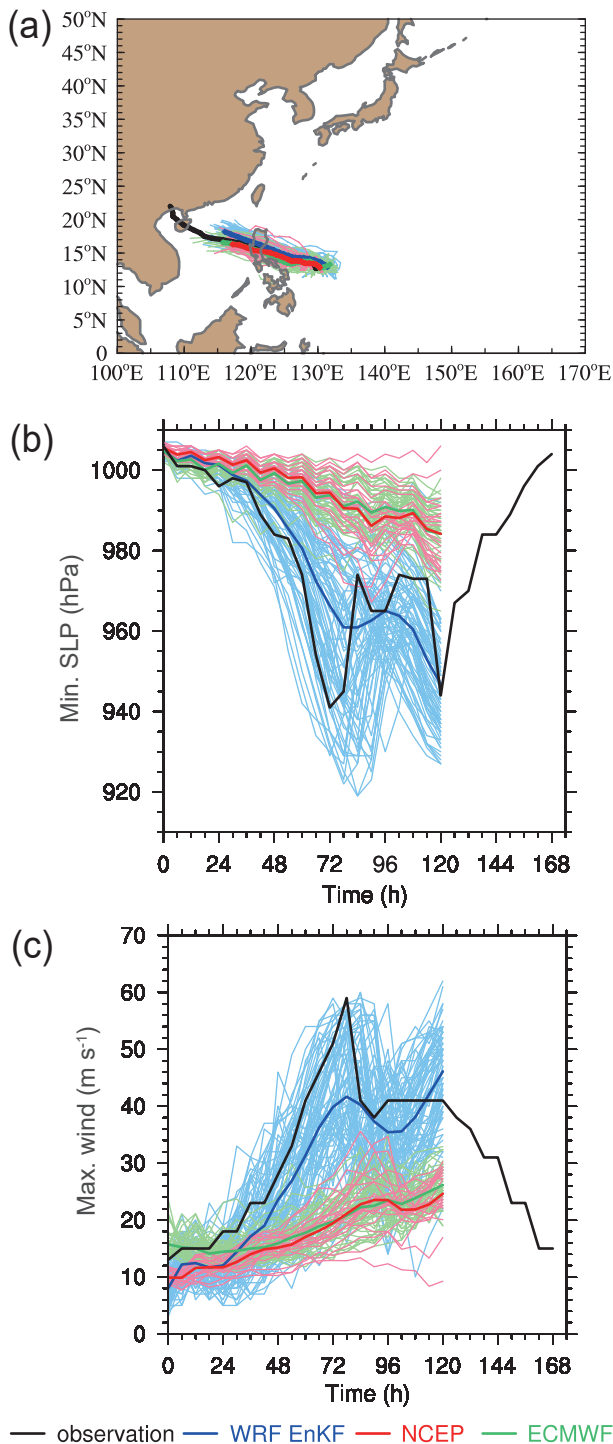
tialization methods and will be reported in a separate study. To improve the WRF/EnKF ensemble forecasts with enlarged ensemble spread, advanced data assimilation that updates ensemble perturbations with hybrid background error covariance (Lei et al., 2021) and additive inflation that can represent model uncertainties (Whitaker and Hamill, 2012) need be further studied. Moreover, the cycling ensembles and 5-d ensemble forecasts here provide a unique dataset for studying TC structure, dynamics, genesis, and predictability. Ensemble sensitivity analysis (e.g., Torn and Hakim, 2008; Lei and Hacker, 2015) using this output to understand the dynamical processes that limit the predictability of TC track, intensity, and structure will be presented in a future study. Data assimilation algorithms that can capture multiscale features of TCs and improve TC predictability will also be investigated.

REFERENCES

Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.*, **137**, 1805−1824, https://doi.org/10.1175/2008MWR2691.1.

Anderson, J. L., 1996: A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *J. Climate*, **9**, 1518−1530, https://doi.org/10.1175/1520-0442(1996)009<1518:AMFPAE>2.0.CO;2.

Barker, D., and Coauthors, 2012: The Weather Research and Forecasting model's community variational/ensemble data assimilation system: WRFDA. *Bull. Amer. Meteor. Soc.*, **93**, 831−843, https://doi.org/10.1175/BAMS-D-11-00167.1.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010a: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experi-

ments. *Mon. Wea. Rev.*, **138**, 1550−1566, https://doi.org/10.1175/2009MWR3157.1.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010b: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations. *Mon. Wea. Rev.*, **138**, 1567−1586, https://doi.org/10.1175/2009MWR3158.1.

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719−1724, https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Cavallo, S. M., R. D. Torn, C. Snyder, C. Davis, W. Wang, and J. Done, 2013: Evaluation of the advanced hurricane WRF data assimilation system for the 2009 Atlantic hurricane season. *Mon. Wea. Rev.*, **141**, 523−541, https://doi.org/10.1175/MWR-D-12-00139.1.

Cha, D.-H., and Y. Q. Wang, 2013: A dynamical initialization scheme for real-time forecasts of tropical cyclones using the WRF Model. *Mon. Wea. Rev.*, **141**, 964−986, https://doi.org/10.1175/MWR-D-12-00077.1.

Chou, K.-H., and C.-C. Wu, 2008: Typhoon initialization in a mesoscale model—Combination of the bogused vortex and the dropwindsonde data in DOTSTAR. *Mon. Wea. Rev.*, **136**, 865−879, https://doi.org/10.1175/2007MWR2141.1.

Dowell, D. C., F. Q. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. *Mon. Wea. Rev.*, **132**, 1982−2005, https://doi.org/10.1175/1520-0493(2004)132<1982:WATRIT>2.0.CO;2.

Efron, B., and R. J. Tibshirani, 1993: *An Introduction to the Bootstrap*. Chapman & Hall, 436 pp.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.*, **108**, 8851, https://doi.org/10.1029/2002JD003296.

Fudeyasu, H., and R. Yoshida, 2018: Western North Pacific tropical cyclone characteristics stratified by genesis environment. *Mon. Wea. Rev.*, **146**, 435−446, https://doi.org/10.1175/MWR-D-17-0110.1.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723−757, https://doi.org/10.1002/qj.49712555417.

Hamill, T. M., 2001: Interpretation of rank histograms for verifying ensemble forecasts. *Mon. Wea. Rev.*, **129**, 550−560, https://doi.org/10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2.

Hamill, T. M., J. S. Whitaker, M. Fiorino, and S. G. Benjamin, 2011: Global ensemble predictions of 2009's tropical cyclones initialized with an ensemble Kalman filter. *Mon. Wea. Rev.*, **139**, 668−688, https://doi.org/10.1175/2010MWR3456.1.

Hendricks, E. A., M. S. Peng, and T. Li, 2013: Evaluation of multiple dynamic initialization schemes for tropical cyclone prediction. *Mon. Wea. Rev.*, **141**, 4028−4048, https://doi.org/10.1175/MWR-D-12-00329.1.

Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *Journal of the Korean*

Meteorological Society, **42**, 129–151.

Hong, S.-Y., 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.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811, https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.*, **133**, 604–620, https://doi.org/10.1175/MWR-2864.1.

Houtekamer, P. L., X. X. Deng, H. L. Mitchell, S.-J. Baek, and N. Gagnon, 2014: Higher resolution in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **142**, 1143–1162, https://doi.org/10.1175/MWR-D-13-00138.1.

Hsiao, L.-F., C.-S. Liou, T.-C. Yeh, Y.-R. Guo, D.-S. Chen, K.-N. Huang, C.-T. Terng, and J.-H. Chen, 2010: A vortex relocation scheme for tropical cyclone initialization in Advanced Research WRF. *Mon. Wea. Rev.*, **138**, 3298–3315, https://doi.org/10.1175/2010MWR3275.1.

Hsiao, L.-F., and Coauthors, 2015: Blending of global and regional analyses with a spatial filter: Application to typhoon prediction over the western north Pacific Ocean. *Wea. Forecasting*, **30**, 754–770, https://doi.org/10.1175/WAF-D-14-00047.1.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, **230**, 112–126, https://doi.org/10.1016/j.physd.2006.11.008.

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

Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, **82**, 35–45, https://doi.org/10.1115/1.3662552.

Kaplan, J., and DeMaria M., 2003: Large-scale characteristics of rapidly intensifying tropical cyclones in the North Atlantic basin. *Wea. Forecasting*, **18**, 1093–1108, https://doi.org/10.1175/1520-0434(2003)018<1093:LCORIT>2.0.CO;2.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. H. Yang, 2009: Improving incremental balance in the GSI 3DVar analysis system. *Mon. Wea. Rev.*, **137**, 1046–1060, https://doi.org/10.1175/2008MWR2623.1.

Kurihara, Y., M. A. Bender, and R. J. Ross, 1993: An initialization scheme of hurricane models by vortex specification. *Mon. Wea. Rev.*, **121**, 2030–2045, https://doi.org/10.1175/1520-0493(1993)121<2030:AISOHM>2.0.CO;2.

Kurihara, Y., M. A. Bender, R. E. Tuleya, and R. J. Ross, 1995: Improvements in the GFDL hurricane prediction system. *Mon. Wea. Rev.*, **123**, 2791–2801, https://doi.org/10.1175/1520-0493(1995)123<2791:IITGHP>2.0.CO;2.

Lei, L. L., and J. P. Hacker, 2015: Nudging, ensemble, and nudging ensembles for data assimilation in the presence of model error. *Mon. Wea. Rev.*, **143**, 2600–2610, https://doi.org/10.1175/MWR-D-14-00295.1.

Lei, L. L., J. S. Whitaker, and C. Bishop, 2018: Improving assimilation of radiance observations by implementing model space localization in an ensemble Kalman filter. *Journal of Advances in Modeling Earth Systems*, **10**, 3221–3232, https://doi.org/10.1029/2018MS001468.

Lei, L. L., Z. R. Wang, Z.-M. Tan, 2021: Integrated hybrid data assimilation for an ensemble Kalman filter. *Mon. Wea. Rev.*, **149**, 4091–4105, https://doi.org/10.1175/MWR-D-21-0002.1.

Lin, H. D., S. S. Weygandt, S. G. Benjamin, and M. Hu, 2017: Satellite radiance data assimilation within the hourly updated rapid refresh. *Wea. Forecasting*, **32**, 1273–1287, https://doi.org/10.1175/WAF-D-16-0215.1.

Liu, H.-Y., and Z.-M., Tan, 2016: A dynamical initialization scheme for binary tropical cyclones. *Mon. Wea. Rev.*, **144**, 4787–4803, https://doi.org/10.1175/MWR-D-16-0176.1.

Liu, Q. F., T. Marchok, H.-L. Pan, M. Bender, and S. Lord, 2000: Improvements in hurricane initialization and forecasting at NCEP with global and regional (GFDL) models. NOAA Tech. Procedures Bulletin 472, 7 pp.

Ma, C., M. Peng, T. Li, Y. Sun, J. Liu, and M. Y. Bi, 2019: Effects of background state on tropical cyclone size over the Western North Pacific and Northern Atlantic. *Climate Dyn.*, **52**, 4143–4156, https://doi.org/10.1007/s00382-018-4372-3.

Magnusson, L., M. Leutbecher, and E. Källén, 2008: Comparison between singular vectors and breeding vectors as initial perturbations for the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.*, **136**, 4092–4104, https://doi.org/10.1175/2008MWR2498.1.

Meng, Z. Y., and F. Q. Zhang, 2008: 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, https://doi.org/10.1175/2008MWR2270.1.

Miyoshi, T., Y. Sato, and T. Kadowaki, 2010: Ensemble Kalman filter and 4D-VAR intercomparison with the Japanese operational global analysis and prediction system. *Mon. Wea. Rev.*, **138**, 2846–2866, https://doi.org/10.1175/2010MWR3209.1.

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**(D14), 16 663–16 682, https://doi.org/10.1029/97JD00237.

Puri, K., J. Barkmeijer, and T. N. Palmer, 2001: Ensemble prediction of tropical cyclones using targeted diabatic singular vectors. *Quart. J. Roy. Meteor. Soc.*, **127**, 709–731, https://doi.org/10.1002/qj.49712757222.

Rappaport, E. N., and Coauthors, 2009: Advances and challenges at the National Hurricane Center. *Wea. Forecasting*, **24**, 395–419, https://doi.org/10.1175/2008WAF2222128.1.

Rogers, R., and Coauthors, 2006: The intensity forecasting experiment: A NOAA multiyear field program for improving tropical cyclone intensity forecasts. *Bull. Amer. Meteor. Soc.*, **87**, 1523–1537, https://doi.org/10.1175/BAMS-87-11-1523.

Schwartz, C. S., 2016: Improving large-domain convection-allowing forecasts with high-resolution analyses and ensemble data assimilation. *Mon. Wea. Rev.*, **144**, 1777–1803, https://doi.org/10.1175/MWR-D-15-0286.1.

Schwartz, C. S., and Z. Q. Liu, 2014: Convection-permitting forecasts initialized with continuously cycling limited-area 3DVAR, ensemble Kalman filter, and "hybrid" variational–ensemble data assimilation systems. *Mon. Wea. Rev.*, **142**, 716–738, https://doi.org/10.1175/MWR-D-13-00100.1.

Schwartz, C. S., Z. Q. Liu, Y. S. Chen, and X.-Y. Huang, 2012:

Impact of assimilating microwave radiances with a limited-area ensemble data assimilation system on forecasts of typhoon Morakot. *Wea. Forecasting*, **27**, 424–437, https://doi.org/10.1175/WAF-D-11-00033.1.

Skamarock, W. C., and Coauthors, 2008: A description of the advanced research WRF version 3. NCAR/TN-475+STR, 125 pp, https://doi.org/10.5065/D68S4MVH.

Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea. Rev.*, **117**, 1779–1800, https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2.

Tong, M. J., and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Wea. Rev.*, **133**, 1789–1807, https://doi.org/10.1175/MWR2898.1.

Torn, R. D., 2010: Performance of a mesoscale ensemble Kalman filter (EnKF) during the NOAA High-Resolution Hurricane Test. *Mon. Wea. Rev.*, **138**, 4375–4392, https://doi.org/10.1175/2010MWR3361.1.

Torn, R. D., and G. J. Hakim, 2008: Ensemble-based sensitivity analysis. *Mon. Wea. Rev.*, **136**, 663–677, https://doi.org/10.1175/2007MWR2132.1.

Torn, R. D., and G. J. Hakim, 2009: Ensemble data assimilation applied to RAINEX observations of Hurricane Katrina (2005). *Mon. Wea. Rev.*, **137**, 2817–2829, https://doi.org/10.1175/2009MWR2656.1.

Torn, R. D., G. J. Hakim, and C. Synder, 2006: Boundary conditions for limited-area ensemble Kalman filters. *Mon. Wea. Rev.*, **134**, 2490–2502, https://doi.org/10.1175/MWR3187.1.

Typhoon Committee, 2015: Typhoon committee operational manual 2015 (report). World Meteorological Organization. Retrieved November 13, 2015. [Available online from https://library.wmo.int/doc_num.php?explnum_id=7633]

Wang, D. L., X. D. Liang, Y. Zhao, and B. Wang, 2008: A comparison of two tropical cyclone bogussing schemes. *Wea. Forecasting*, **23**, 194–204, https://doi.org/10.1175/2007WAF2006094.1.

Wang, Y., and C.-C. Wu, 2004: Current understanding of tropical cyclone structure and intensity changes—A review. *Meteorol. Atmos. Phys.*, **87**, 257–278, https://doi.org/10.1007/s00703-003-0055-6.

Wei, M. Z., Z. Toth, R. Wobus, and Y. J. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus A*, **60**, 62–79, https://doi.org/10.1111/j.1600-0870.2007.00273.x.

Whitaker, J. S., and T. M. Hamill, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Wea. Rev.*, **140**, 3078–3089, https://doi.org/10.1175/MWR-D-11-00276.1.

Whitaker, J. S., T. M. Hamill, X. Wei, Y. C. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. *Mon. Wea. Rev.*, **136**, 463–482, https://doi.org/10.1175/2007MWR2018.1.

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916, https://doi.org/10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2.

Xue, M., J. Schleif, F. Y. Kong, K. W. Thomas, Y. H. Wang, and K. F. 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.

Zhang, C. X., Y. Q. Wang, and K. Hamilton, 2011a: Improved representation of boundary layer clouds over the southeast pacific in ARW-WRF using a modified Tiedtke cumulus parameterization scheme. *Mon. Wea. Rev.*, **139**, 3489–3513, https://doi.org/10.1175/MWR-D-10-05091.1.

Zhang, F. Q., Y. H. Weng, J. A. Sippel, Z. Y. Meng, and C. H. Bishop, 2009: Cloud-resolving hurricane initialization and prediction through assimilation of Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2105–2125, https://doi.org/10.1175/2009MWR2645.1.

Zhang, F. Q., Y. H. Weng, J. F. Gamache, and F. D. Marks, 2011b: 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, https://doi.org/10.1029/2011GL048469.

Zhou, X. Q., Y. J. Zhu, D. C. Hou, Y. Luo, J. Y. Peng, and R. Wobus, 2017: Performance of the new NCEP global ensemble forecast system in a parallel experiment. *Mon. Wea. Rev.*, **32**, 1989–2004, https://doi.org/10.1175/WAF-D-17-0023.1.

Zou, X. L., and Q. N. Xiao, 2000: Studies on the initialization and simulation of a mature hurricane using a variational bogus data assimilation scheme. *J. Atmos. Sci.*, **57**, 836–860, https://doi.org/10.1175/1520-0469(2000)057<0836:SOTIAS>2.0.CO;2.