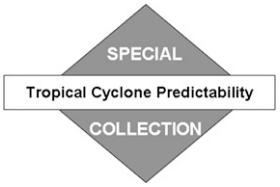

Cloud-Resolving Hurricane Initialization and Prediction through Assimilation of Doppler Radar Observations with an Ensemble Kalman Filter

FUQING ZHANG

Department of Meteorology, The Pennsylvania State University, University Park, Pennsylvania

YONGHUI WENG

Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, and Department of Atmospheric Sciences, Texas A&M University, College Station, Texas, and The Graduate School, Chinese Academy of Sciences, Beijing, China

JASON A. SIPPEL

Department of Atmospheric Sciences, Texas A&M University, College Station, Texas

ZHIYONG MENG

Department of Atmospheric Sciences, School of Physics, Peking University, Beijing, China

CRAIG H. BISHOP

Marine Meteorology Division, Naval Research Laboratory, Monterey, California

(Manuscript received 8 May 2008, in final form 20 January 2009)

ABSTRACT

This study explores the assimilation of Doppler radar radial velocity observations for cloud-resolving hurricane analysis, initialization, and prediction with an ensemble Kalman filter (EnKF). The case studied is Hurricane Humberto (2007), the first landfalling hurricane in the United States since the end of the 2005 hurricane season and the most rapidly intensifying near-landfall storm in U.S. history. The storm caused extensive damage along the southeast Texas coast but was poorly predicted by operational models and forecasters. It is found that the EnKF analysis, after assimilating radial velocity observations from three Weather Surveillance Radars-1988 Doppler (WSR-88Ds) along the Gulf coast, closely represents the best-track position and intensity of Humberto. Deterministic forecasts initialized from the EnKF analysis, despite displaying considerable variability with different lead times, are also capable of predicting the rapid formation and intensification of the hurricane. These forecasts are also superior to simulations without radar data assimilation or with a three-dimensional variational scheme assimilating the same radar observations. Moreover, nearly all members from the ensemble forecasts initialized with EnKF analysis perturbations predict rapid formation and intensification of the storm. However, the large ensemble spread of peak intensity, which ranges from a tropical storm to a category 2 hurricane, echoes limited predictability in deterministic forecasts of the storm and the potential of using ensembles for probabilistic forecasts of hurricanes.

1. Introduction

Landfalling hurricanes are among the deadliest and costliest natural hazards. Over the past decade, significant progress has been made in short-range track forecasts of tropical cyclones. The current-day average 48-h forecast position is as accurate as a 24-h track forecast was 10 yr ago (Franklin 2004). However, there is virtually no improvement in our ability to predict hurricane intensity in terms of minimum sea level pressure, maximum wind speed, or amount of precipitation (Houze et al. 2007). We thus have very limited skill in predicting tropical cyclone formation, rapid intensification,

Corresponding author address: Dr. Fuqing Zhang, Dept. of Meteorology, The Pennsylvania State University, University Park, PA 16802.
E-mail: fzhang@psu.edu

DOI: 10.1175/2009MWR2645.1

fluctuation, or decay (Elsberry et al. 2007). High-resolution cloud-resolving mesoscale models, along with better initialization of the initial vortex, may be necessary to faithfully represent the internal dynamics that is crucial for hurricane intensity forecasts (Houze et al. 2007; Chen et al. 2007; Davis et al. 2008).

Despite improvements in using advanced data assimilation methods with or without initial vortex bogussing, our ability to initialize a tropical cyclone with dynamically consistent structure and intensity remains limited, even with the assimilation of radar observations (e.g., Zou and Xiao 2000; Pu and Braun 2001; Xiao et al. 2007). Numerical weather prediction models also have known difficulties in their "spinup" of a tropical cyclone or hurricane vortex with appropriate moisture, diabatic, and divergence structures at the initial time. Part of the difficulty of hurricane initialization comes from the lack of routine four-dimensional observations with sufficient spatial and temporal resolution to represent the initial hurricane structure and intensity. Another part of the difficulty comes from the deficiency of the current generation of operational data assimilation systems, which use static background error covariance. The mostly balanced, isotropic, flow-independent background statistics derived from long-term averages of past short-term forecast error (Parrish and Derber 1992) are ill-suited for the highly flow-dependent background error covariances associated with tropical cyclones. In addition, operational models generally have insufficient model resolution to effectively incorporate high-resolution convective-scale observations (such as those from radars) for cloud-resolving hurricane prediction. Physical (diabatic) initializations using rainfall, radar, and/or satellite observations are a promising approach (Krishnamurti et al. 2001), though its effectiveness in spinning up a full hurricane vortex for cloud-resolving hurricane prediction remains to be fully explored.

The ensemble Kalman filter (EnKF) is a state-estimation technique that uses short-term ensemble forecasts to estimate flow-dependent background error covariance or other probabilistic aspects of the background forecast. It was first proposed by Evensen (1994) and has been adopted for data assimilation of many disciplines in the geosciences and beyond (Evensen 2003; Hamill 2006). For the past few years, the feasibility and performance of the EnKF have been demonstrated through both simulated and real-data observations ranging from convective scales using radar observations (e.g., Snyder and Zhang 2003; Zhang et al. 2004; Dowell et al. 2004; Tong and Xue 2005) to mesoscale and regional scales (e.g., Zhang et al. 2006a; Torn et al. 2006; Meng and Zhang 2007; Fujita et al. 2007; Meng and Zhang 2008a,b). The benefits of directly estimating forecast covariances are also likely to increase with the next generation of NWP models that resolve scales at which physical balances are not amenable to stationary, isotropic covariance models currently used for operational forecasts.

The current study explores for the first time the use of EnKF to directly assimilate Doppler radar radial velocity observations for cloud-resolving hurricane analysis and prediction, both deterministically and probabilistically. The case to be examined is Hurricane Humberto (2007), the first landfalling hurricane in the United States since the end of the 2005 hurricane season and the most rapidly intensifying near-landfall storm in U.S. history. Humberto strengthened from a 40 mile per hour (mph) depression at 1200 UTC 12 September 2007 to a 92-mph hurricane at 0700 UTC 13 September, a 52 mph increase in surface wind speed in only 19 h. The storm caused extensive damage along the southeast Texas coast and was poorly predicted by operational models and forecasters. The real-time forecast by the operational Global Forecast System (GFS) running at the National Centers for Environmental Prediction (NCEP) failed to capture the intensification and genesis of the storm (Fig. 1). The Weather Research and Forecasting (WRF) model also failed in postevent, 4.5-km, cloud-resolving simulations that were initialized with the GFS analyses (as in the control ensemble analysis and forecast detailed in subsequent sections) with lead times every 6 h from 6 to 48 h.

At different stages of formation and intensification, Humberto was within range of coastal Weather Surveillance Radars-1988 Dopplers (WSR-88Ds) at Corpus Christi (KCRP) and Houston–Galveston (KHGX) in Texas and Lake Charles (KLCH) in Louisiana. These radars provided valuable convective-scale observations of the storm but were not assimilated by real-time NCEP operational models. This study is among the first to apply the EnKF to assimilate real-data radar velocity observations for complex weather phenomena exhibiting different scales of motion [moving beyond a single supercell storm examined in Dowell et al. (2004)]. The following section will introduce the forecast model, the EnKF technique, and the processing of the observations to be assimilated. Section 3 presents the use of the EnKF for this storm in terms of analysis quality, and deterministic and ensemble forecasts. Comparison to the performance of data assimilation with a 3-dimensional variational method that assimilates the same radar observations as in the WRF model is given in section 4. Concluding remarks are given in section 5.

2. Methodology

a. The forecast model: WRF

The Advanced Research WRF (ARW) is used in this study. WRF is a fully compressible, nonhydrostatic

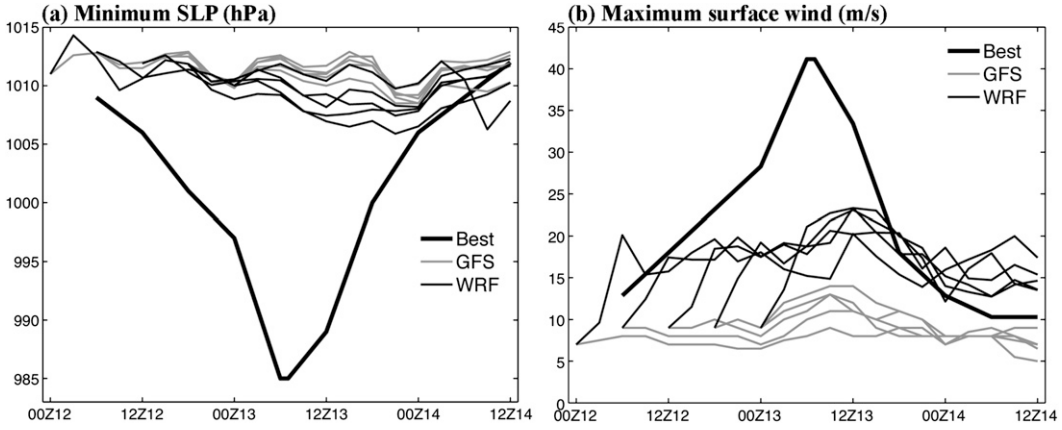

FIG. 1. Time evolution of (a) minSLP and (b) maxWSP forecasts by operational GFS forecasts starting every 6 h from 0000 UTC 12 Sep to 0000 UTC 12 Sep 2007 and by 4.5-km WRF forecasts starting from the operational GFS analyses in comparison to the NHC best-track estimate.

mesoscale model (Skamarock et al. 2005). The vertical coordinate follows the terrain using hydrostatic pressure, and the model uses an Arakawa C grid. Prognostic variables are the column mass of dry air, velocities (u, v, w), potential temperature, geopotential, and mixing ratios for water vapor, cloud, rain, ice, snow, and graupel.

In the control experiments, three model domains with two-way nesting are used (Fig. 2). The two coarse domains (D1 and D2) both have 160×121 grid points and grid spacings of 40.5 and 13.5 km, respectively. The innermost domain, D3, has 253×253 grid points and a grid spacing of 4.5 km. All model domains have 35 vertical layers, and the model top is set at 10 hPa. The physical parameterization schemes include the Grell–Devenyi cumulus scheme (Grell and Devenyi 2002), WRF single-moment six-class microphysics with graupel (Hong et al. 2004), and the Yonsei State University (YSU) scheme (Noh et al. 2003) for planetary boundary layer processes. The NCEP GFS operational analysis at 0000 UTC 12 September and its forecast are used to create the initial and boundary conditions. Data assimilation is performed for all domains but all verification is performed for D3.

b. The data assimilation method: EnKF

The EnKF implemented in the WRF model is the same as that in Meng and Zhang (2008a,b) except that no multischeme ensemble is used for this study. This version of the filter was originally implemented in the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5), which is documented in Zhang et al. (2006a). It uses the covariance relaxation of Zhang et al. (2004) to inflate the background error covariance. Unlike the standard inflation method (Anderson 2001), in which all points in the prior field are inflated, this relaxation method only inflates the covariance at updated points via a weighted average between the prior perturbation (denoted by superscript f) and the posterior perturbation (denoted by superscript a) as follows:

$$(\mathbf{x}^a_{\text{new}})' = (1 - \alpha)(\mathbf{x}^a)' + \alpha(\mathbf{x}^f)'. \quad (1)$$

The weighting coefficient, α, is set to 0.5 in the observing system simulation experiment (OSSE) studies of Zhang et al. (2004, 2006a) for the perfect-model experiments but a range from 0.7 to 0.8 is found to be necessary for imperfect-model experiments (Meng and Zhang 2007) or real-data applications (Whitaker et al. 2008; Meng and Zhang 2008a,b; Torn and Hakim 2008) due to unavoidable imperfections in the forecast model. A value of 0.8 is used for the present real-data study.

c. Ensemble initial and boundary perturbations

Although the optimum ensemble size for estimating the forecast uncertainty is still under active research, 30 members are used herein. An ensemble size of 20–50 was found to be affordable and reasonable based on previous studies (e.g., Houtekamer and Mitchell 2001; Anderson 2001; Snyder and Zhang 2003; Zhang 2005; Zhang et al. 2006a; Meng and Zhang 2007; Meng and Zhang 2008a,b). As in Zhang et al. (2006a), the initial ensemble is generated with the WRF's three-dimensional variational data assimilation (3DVAR) using the cv3 background error covariance option (Barker et al. 2004). To create a largely balanced perturbation, we first generate a set of random control vectors with a normal distribution (zero mean and unit standard deviation). Then, the control increment vector is transformed back to model space via an EOF transform, a recursive filter,

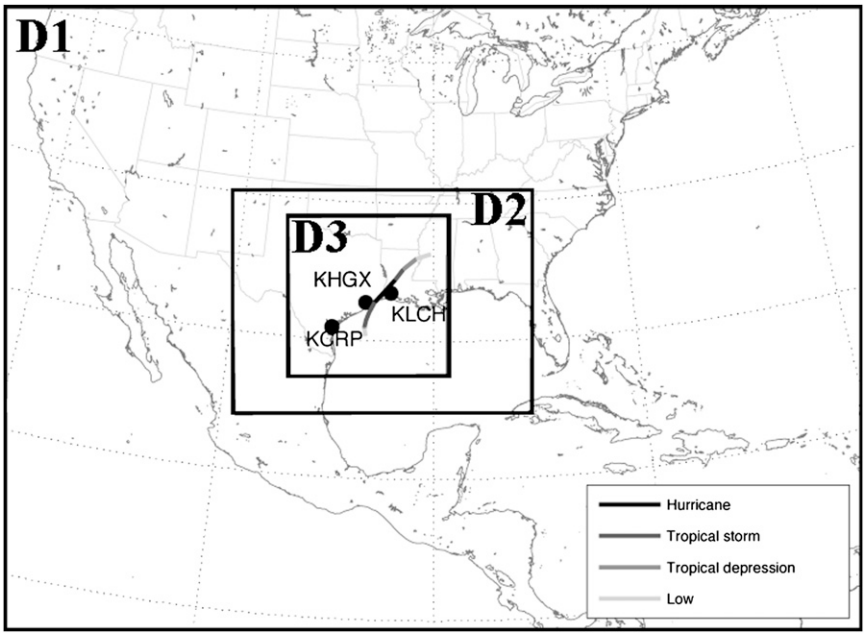

FIG. 2. Configuration of WRF model domains 1, 2, and 3 with horizontal grid spacings of 40.5, 13.5, and 4.5 km, respectively. Also depicted are the NHC best-track estimates of Humberto with intensity in gray scale and the three WSR-88D locations.

and physical transformation via balance equations. The perturbed variables include horizontal wind components, potential temperature, and mixing ratio for water vapor, and their error statistics are defined by the climatological background error covariance. Other prognostic variables such as vertical velocity (w) and mixing ratios for cloud water (q_c), rainwater (q_r), snow (q_s), and graupel (q_g) are not perturbed. The perturbation standard deviations thus generated are approximately 2 m s^{-1} for horizontal wind components (u and v), 0.8 K for temperature (T), 1 hPa for pressure perturbation (p'), and 0.8 g kg^{-1} for the water vapor mixing ratio (q). The 3DVAR perturbations are added to the GFS analysis to form an initial ensemble, which is then integrated for 9 h to develop an approximately realistic, flow-dependent background error covariance structure before the first observation is assimilated. Similar methods, using 3DVAR to generate the initial ensemble for the EnKF, are also employed in Houtekamer et al. (2005) and Barker (2005).

The simplest way to perturb lateral boundary conditions for a limited-area model is to use a global ensemble forecast with the correct size and resolution [which are usually unavailable; Chessa et al. (2004)]. Torn et al. (2006) examined several alternative boundary perturbation methods and concluded that the error originating from using different methods is limited to near the edges of the domain. In this paper, the GFS operational forecasts are used to create boundary conditions that are perturbed in the same manner as with the initial ensemble.

d. Superobservations and quality control

With large volumes of radar observations that are recorded at a much higher resolution than the forecast model grid spacing for the EnKF data assimilation, significant data thinning and quality control of observations become necessary. The process of combining multiple observations into one high-accuracy ''super'' observation (SO) is often referred to as ''superobbing.'' An SO for radar radial velocity is created through horizontal averaging in polar space of the raw polar volume of data (Lindskog et al. 2000, 2004; Alpert and Kumar 2007). To minimize horizontal correlations of the SOs, each pixel of the raw data is allowed to influence one SO only. To avoid the averaging of the radial velocity (Vr) in significantly different directions, the averaging bin is confined within 5 km in the radial direction and 5° in the azimuthal direction. Our selection of the bin size is within the range of values used in the literature (e.g., Lindskog et al. 2000; Alpert and Kumar 2007; Montmerle and Faccani 2009).

For quality control of the observations, we first use the NCAR radar editing software called SOLO [information online at http://www.eol.ucar.edu/rdp/solo/solo_home.html] to dealias the range-folded data and to

remove apparently erroneous observations. We then implement the following additional quality control measures in the SO generation for this study: 1) any raw observations with values smaller than 2 m s^{-1} or larger than 70 m s^{-1} or with distances to the radar smaller than 4 km will be discounted, 2) a raw Vr observation will be discounted if its deviation from the bin mean exceeds twice the standard deviation of all raw observations in the bin, 3) there shall be at least four valid Vr raw observations within an averaging bin, 4) there will be no SO for a bin whose standard deviation is twice the average of the standard deviations in all bins, and 5) the final SO value of the bin will be the average of at most 10 raw observations that are closest to the center of the bin. These quality control procedures are in place to minimize the impact of ground clutter and to correct the failures in the subjective dealiasing step; similar procedures were also employed in Xu et al. (2003) and Montmerle and Faccani (2009).

Additional quality controls are implemented in the processing of EnKF analyses. The observation errors of all SOs are assumed to be 3 m s^{-1} in this study, and an SO will also be discounted if the difference between this SO and the forecast prior is larger than 5 times the observation error. The 3 m s^{-1} observational error is consistent with the range used in Dowell et al. (2004) and Montmerle and Faccani (2009), and is also a conservative version of the 2.4 m s^{-1} estimated in Xu et al. (2003).

e. *Successive covariance localization*

A *successive covariance localization* (SCL) technique is designed to assimilate dense radar observations that contain information about the state of the atmosphere at a wide range of scales. The method is also designed to reduce computational costs and sampling errors. This technique uses the Gaspari and Cohn (1999) fifth-order correlation function for covariance localization, but a different localization radius of influence (ROI) is used for different groups of observations by random sampling. SCL assumes that both large- and small-scale errors are simultaneously present. First, one tries to remove dynamically important aspects of the large-scale error by assimilating a relatively small subset of observations with a large ROI. Next, the ROI is made smaller, and higher-density observations are used to constrain both smaller-scale errors and what remains of the large-scale error. The process is repeated until all scales resolved by the observational network have been adequately dealt with. The SCL method has some resemblance to the successive correction method used in some earlier empirical objective analysis schemes (e.g., Barnes 1964), though in the EnKF the same observation will not be used twice. Sensitivity experiments that demonstrate the benefits of using the SCL method over using single ROIs for all observations will be presented in section 3e. The use of SCL is partially motivated by the fact that with serial observation processing of the EnKF, the error correlation length scale decreases as the previously assimilated observations better define the large scales; hence, later observations should be assimilated with tighter localization (Bishop and Hodyss 2007).

In this particular case, we first use a horizontal ROI of 1215 km for 10% of all SOs to capture the large-scale background flow in all three domains (the numbers of SOs at each time from each radar are shown in Fig. 3). We then use an ROI of 405 km to assimilate another 20% of the total SOs to represent the mesoscale flow (i.e., tropical cyclone scale) for the 13.5- and 4.5-km domains (i.e., D2 and D3). Last, we use an ROI of 135 km to assimilate another 60% of the total SOs to capture even smaller-scale phenomena that include mesoscale vortices just in D3. In other words, we assimilate 10%, 30%, and 90% of the total SOs for D1, D2, and D3, respectively. While the partitioning of the observations into different bins remains subjective, we binned more observations to smaller scales since arguably there are larger degrees of freedom and lesser degrees of balance at smaller scales. The remaining 10% of the SOs that are not assimilated by any domain will be used for verifications.

The vertical ROI is 34 based on the number of vertical levels (as in Zhang et al. 2006a) for all three domains. The Gaspari and Cohn (1999) fifth-order correlation function is also used for the vertical covariance localization.

3. EnKF performance

a. *EnKF analyses*

The control EnKF experiment begins to assimilate the SOs from the KCRP and KHGX WSR-88Ds at 0900 UTC 12 September 2007, 6 h before the tropical depression status was declared by National Hurricane Center (NHC). The EnKF continues to assimilate SOs from these two radars every hour until 1800 UTC 12 September. After this time, the KCRP radar is too far from the storm to have any significant impact while KLCH radar begins to offer significant coverage of the storm. The KLCH radar observations are thus assimilated beginning at 1900 UTC 12 September. The EnKF assimilation continues every hour until 1200 UTC 13 September, a few hours after the storm in the NHC best track reaches its peak intensity and starts weakening over land. The numbers of SOs from the three radars by

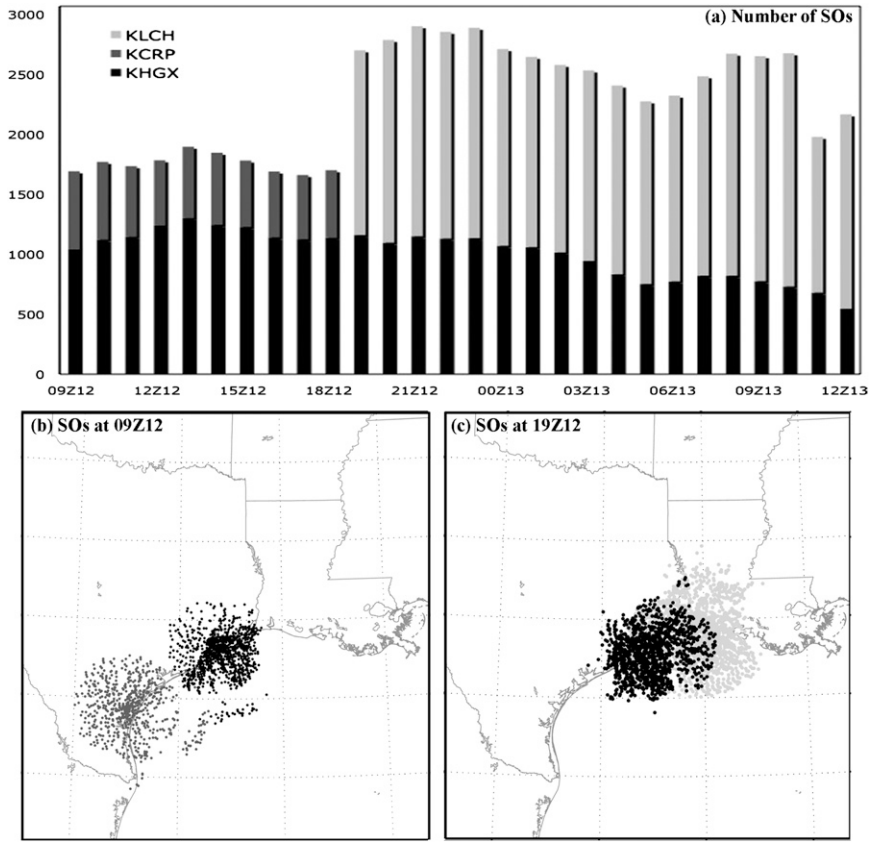

FIG. 3. (a) The number of SOs from each radar at different times by the control EnKF experiment and the exemplar distributions of SOs at (b) 0900 and (c) 1900 UTC 12 Sep 2007.

the EnKF at different times and example distributions of the SOs at 0900 and 1900 UTC 12 September are shown in Fig. 3.

Figure 4 shows the time evolution of the minimum sea level pressure (minSLP) and maximum surface wind speed (maxWSP) estimated from the posterior EnKF mean analysis field (gray; hereafter ''the EnKF analysis'') as well as the average (red) of the maximum–minimum values estimated from each ensemble member's posterior (green) in comparison to the NHC best-track analyses (black). Minimum SLP in the EnKF analysis (gray) agrees well with the average of the members' minimum values (red) due to collocation of the centers of most members. The analysis mean and some ensemble members also compare favorably with the NHC best-track analysis. However, due to the strong spatial and temporal variability of the maximum surface wind speed, maxWSP in the EnKF analysis (gray) is significantly smaller than that of the average (red) of the members' maxima. Yet, the average of each member's maxWSP (red) matches well with the NHC best-track estimate (black) but larger errors are seen after landfall. We therefore believe it is more appropriate to use the averages of maxWSP and minSLP from each member for verifying ensemble forecasts in terms of extreme values.

There is large ensemble spread of both minSLP and maxWSP among the analysis members (Fig. 4). Standard deviation of minSLP (maxWSP) increases from 1 to 2 hPa (1–3 m s^{-1}) during the first few assimilation cycles to 10.2 hPa (8.3 m s^{-1}) near the peak intensity time at 0800 UTC 13 September but drops quickly to 5.4 hPa (4.0 m s^{-1}) at 0900 UTC 13 September after the storm in most members makes its landfall (not shown). Likewise, the minSLP (maxWSP) at the peak intensity time varies from 992 (27 m s^{-1}) to 960 hPa (59 m s^{-1}). In terms of the corresponding intensity categories, this represents a range from a strong tropical storm to a major hurricane. Large disparities between ensemble members demonstrate significant EnKF analysis uncertainties during and after the rapid intensification of Humberto.

Despite the large spread, data assimilation with the EnKF is clearly beneficial. All analysis ensemble members capture the storm's formation and intensification when EnKF assimilates Vr observations. This is in

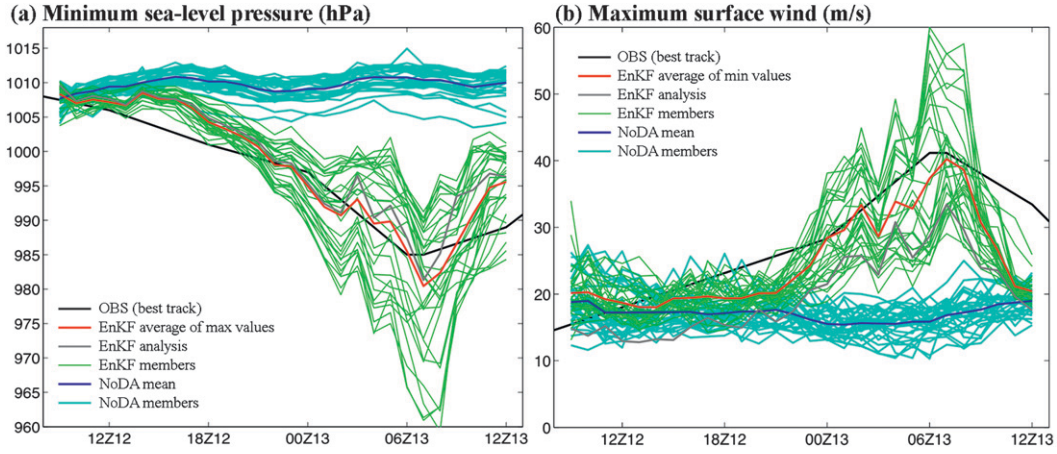

FIG. 4. Time evolutions of (a) minSLP and (b) maxWSP estimated from the EnKF analysis and NoDA ensemble forecast. Thin green (cyan) lines represent the maxWSP and minSLP estimated from each ensemble member of the EnKF analysis (NoDA forecast) with thick red (blue) lines representing the average of maxWSP and minSLP in each member. The gray line indicates the maxWSP–minSLP estimated from the EnKF analysis mean while the black curve is the NHC best-track estimate.

strong contrast to pure ensemble forecasts started with the same prior perturbations but without EnKF assimilation of Vr (hereafter refer to as "NoDA"). The NoDA ensemble neither captures the mean development nor the realistic uncertainties associated with the mean forecast (blue curves in Fig. 4).

Figure 5 compares the observed 0.5° base scan of the radial velocity from KHGX (only 1/10 of the observations plotted) with corresponding simulated values from both the EnKF analysis and the NoDA ensemble forecast mean valid at 0900 and 1800 UTC 12 September, and 0300 UTC 13 September. The benefit of assimilating Vr observations with EnKF is evident even after the first volume of observations is assimilated at 0900 UTC 12 September. The analyses at this time capture the coastal mesoscale circulation much better than the NoDA ensemble mean. Since NoDA at 0900 UTC 12 September is simply the EnKF prior estimate with no Vr observations assimilated, a comparison of Figs. 5b and 5c (verifying against Fig. 5a) shows the immediate benefit of the EnKF at the initial assimilation time. Subsequently, the EnKF well analyzes the cyclone vortex structure, intensity, and evolution (Fig. 5).

The quality of the EnKF analysis is also independently verified against the additional 10% of the total radial velocity SOs from both the KHGX and KLCH radars that have never been assimilated (refer to section 2e). Figure 6 shows the root-mean-square error (RMSE) of the EnKF analysis in comparison to the NoDA experiment plotted every 3 h (RMSE is the difference between the simulated radial velocity and the 10% SOs verified at the SO's locations). Except for at the first analysis time for the KLCH radar and consistent with the other measures of analysis quality discussed above, the EnKF analysis error maintained a remarkably smaller amplitude verifying against these independent radial velocity SOs. The mean analysis errors verified against both radars are similar to the assumed observational error of 3 m s^{-1}. Consistent with the increasing analysis ensemble spread in the maxWSP in Fig. 3c, there is a significant increase in the analysis error during the peak intensity times, especially when verifying against the KHGX radar SOs (Fig. 6a).

Figure 7 shows a comparison of the radar reflectivities from the observed composite, the EnKF analysis, and the NoDA ensemble forecast mean valid at 1200 UTC 12 September, and 0000 and 1200 UTC 13 September, respectively. The impact of Vr observation assimilation on the unobserved reflectivity variable is evident in the progressively better posterior estimates (analysis means) of reflectivity. At 1200 UTC 12 September, while the NoDA ensemble forecast mean simulates a broader area of light precipitation, the EnKF analysis begins to localize the precipitation into two primary bands. The first band is located along the Gulf coast of Texas and Louisiana to the east and northeast of Houston, and the other is farther west but far south of the display domain. This compares much more favorably to the observations (though the convection in the EnKF analysis is still weaker and broader partly due to the ensemble averaging effects discussed above).

At 0000 UTC 13 September, during the storm's rapid intensification, the EnKF analysis captures an impressive developing tropical cyclone south of KHGX with multiple spiral rainbands to the north and east

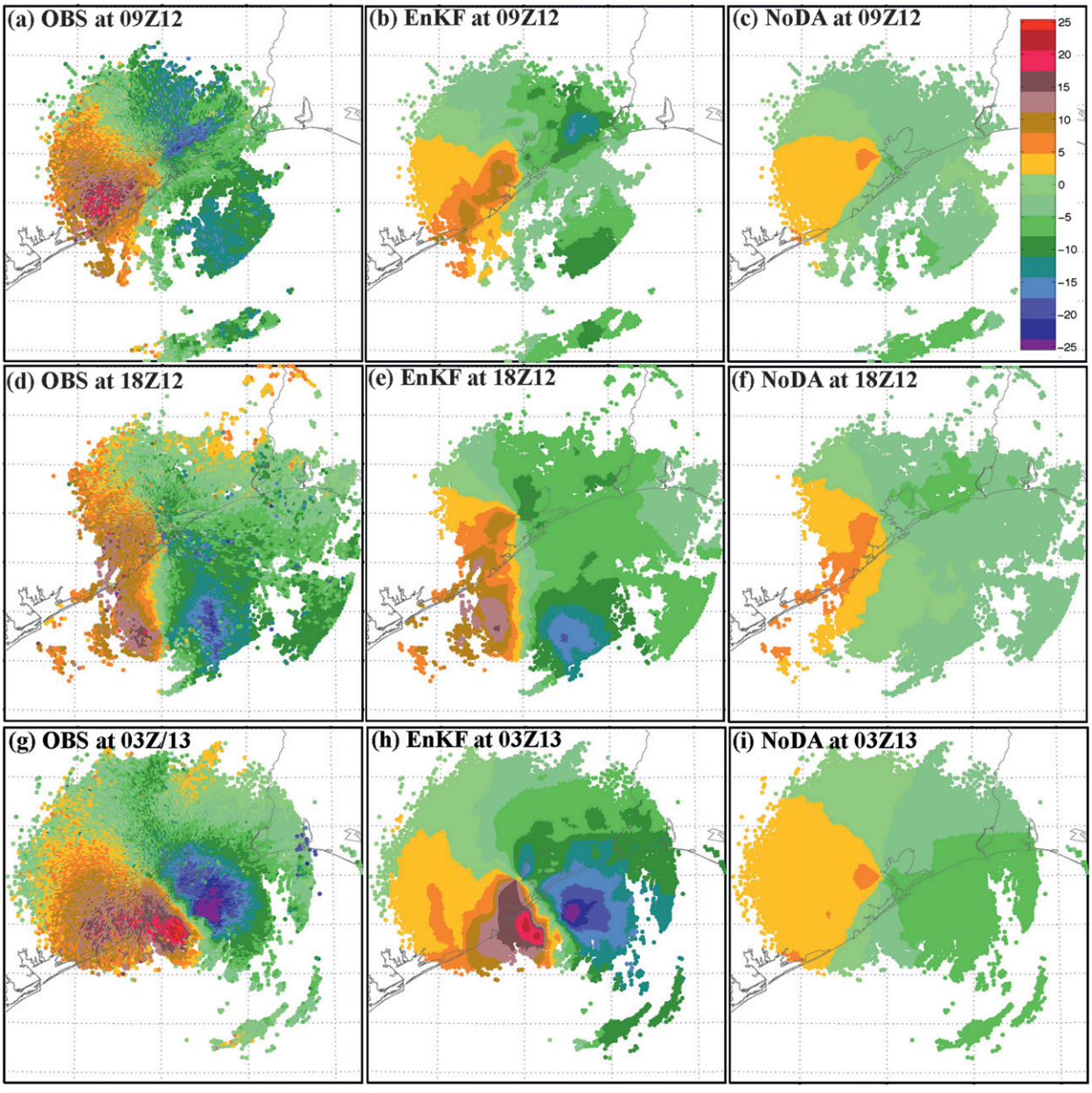

FIG. 5. The raw radial velocity observations from the (left) 0.5° base scan of the KHGX radar (OBS), (middle) corresponding EnKF analysis, and (right) NoDA ensemble forecast mean valid at 0900 and 1800 UTC 12 Sep, and 0300 UTC 13 Sep 2007, respectively.

quadrants fueled by warm moist air from the south. Except for a spurious onshore mesoscale rainband right across the Texas and Louisiana border, the position, intensity, and structure of the rainbands, including the developing eyewall in the EnKF analysis (Fig. 7e), compare remarkably well with the observed reflectivity (Fig. 7d). At 1200 UTC 13 September, the final analysis time after the storm begins its rapid weakening over land, the EnKF analysis correctly places the center of the storm over the Texas–Louisiana border. It also correctly analyzes the broad rainbands to the east of the storm (Figs. 7g and 7h). The NoDA ensemble forecast mean, on the other hand, does not simulate any tropical development (Figs. 7c, 7f, and 7i).

b. Deterministic forecasts from the EnKF analysis

Next, we examine the value that using EnKF to assimilate Vr adds to forecasts. Single, deterministic forecasts from the control EnKF mean analyses are performed every 3 h from 1200 UTC 12 September to 1200 UTC 13

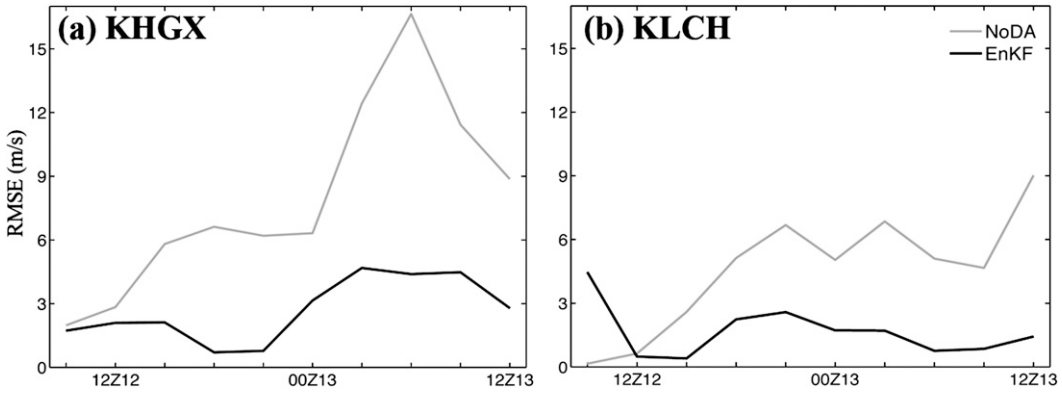

FIG. 6. The RMSE of the radial velocity of the EnKF analysis vs the NoDA experiment verified against the 10% independent radial velocity SOs from (a) KHGX and (b) KLCH.

September. All these simulations are integrated until 1200 UTC 14 September.

Figure 8 shows the simulated cyclone position, minSLP, and maxWSP in deterministic WRF forecasts initialized with EnKF analyses at different times. Despite a slight delay in peak intensity compared to the best-track estimates, all forecasts initialized with EnKF analyses simulate significant tropical cyclone formation and intensification if WSR-88D radial velocity observations are assimilated for 9 h or longer. On average, continuous assimilation through time and assimilating more observations produces both better analyses of the initial storm and better deterministic track and intensity forecasts, which is especially evident for forecasts initialized before 1800 UTC 12 September. These WRF forecasts from the EnKF analyses are in strong contrast to the nearly complete forecast failure of the WRF forecasts (with the same model configuration) cold started from the GFS analyses at all lead times (Fig. 1). This signifies the importance and potential of assimilating radar observations in improving cloud-resolving tropical cyclone initialization and prediction at all lead times.

Among the simulations initialized from the mean EnKF analyses, forecasts from 1800 and 2100 UTC 12 September are the most remarkable. The peak intensity of both simulations (which have significant lead times) is within 2 hPa of the observed (best track) peak intensity. Also, the maximum surface winds in both runs reach or nearly reach category 1 hurricane intensity, which is less than 5 m s^{-1} different from the best-track estimates. Both forecasts are considered to be quite successful given the uncertainties in the best-track estimate and since a relatively small number of model outputs (every 3 h) is used for determining the simulated intensity. The slight delay in the peak intensity for all forecasts reflects a slight lag in the simulated landfall time compared to the best-track observations. Broadly speaking, the initial positions and subsequent track errors in the EnKF analyses become smaller as the radar data assimilation cycle proceeds.

Figure 9 compares the observed composite radar reflectivity with the corresponding simulated reflectivity in deterministic forecasts valid at 2100 UTC 12 September and 0300 and 0900 UTC 13 September. Experiment NoDA is initialized at 0900 UTC 12 September from the ensemble-mean forecast, and experiment EnKF is initialized at 1800 UTC 12 September after 9 h of EnKF analysis. Consistent with the track and intensity forecasts shown in Fig. 8, the deterministic forecast from the EnKF analysis mean compares favorably with radar observations in terms of the structure and placement of rainbands and the formation of the eyewall right before moving over the coast. These phenomena are nonexistent in the forecast without EnKF assimilation of Vr observations (Figs. 9c, 9f, and 9i). The forecast without data assimilation also develops widespread convection over the Gulf region across the display domain, but the convection is highly disorganized with only weak cyclonic circulation in the center. While the above forecast from the EnKF analysis is far better, it is also far from perfect. For example, the outer spiral rainband to the far east of the storm is not well captured, partly because this forecast is initialized at 1800 UTC 12 September, before the KLCH radar observations are assimilated.

c. Deterministic forecast with a 1.5-km movable nested domain

Since the 4.5-km control experiments only marginally resolve moist convection, we perform another high-resolution experiment that nests an additional domain with 1.5-km horizontal grid spacing (1.5KM). Experiment 1.5KM has a fourth domain of 253 × 253 grid points centered on the maximum vorticity and moves

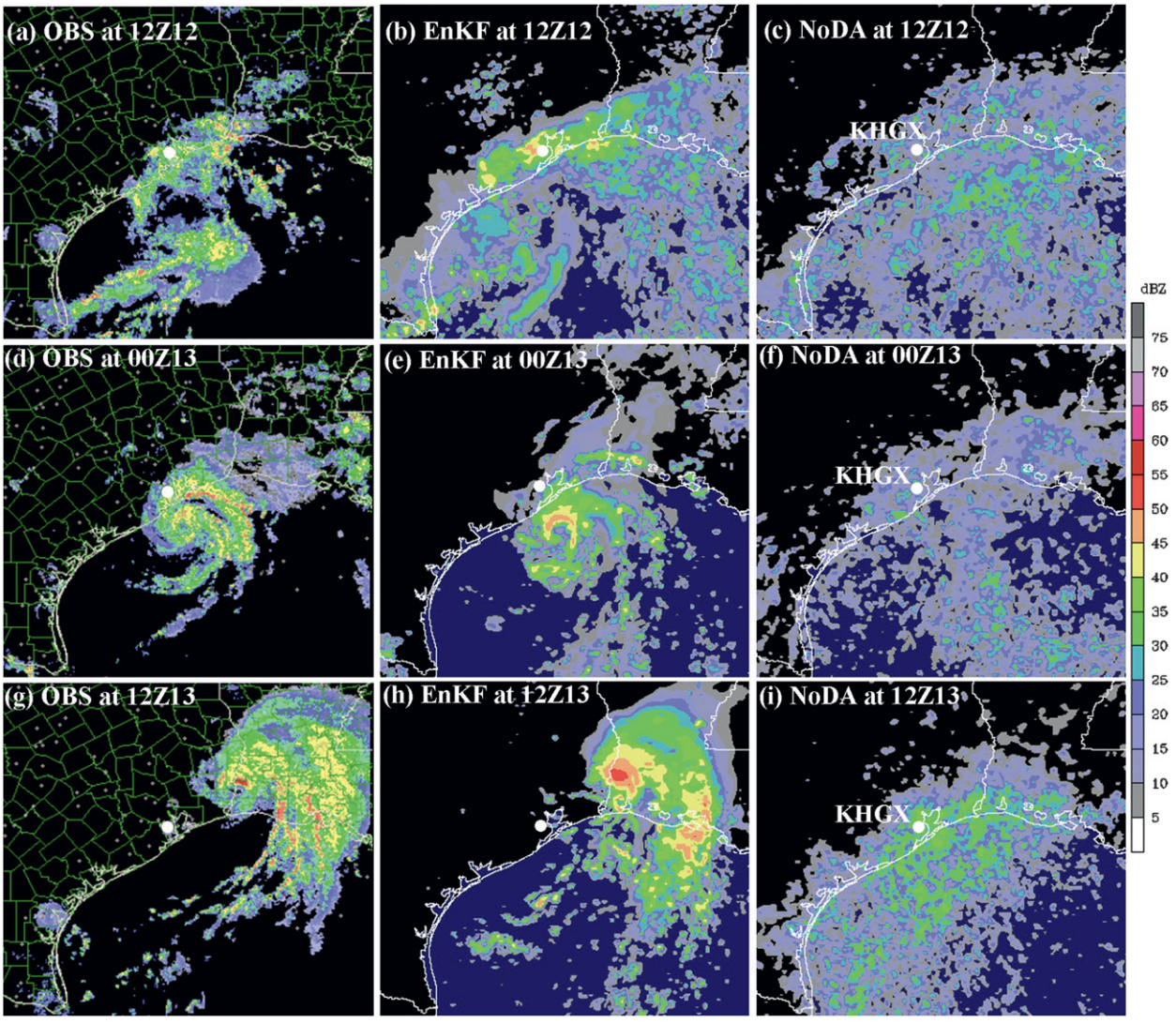

FIG. 7. Comparison of the (left) radar reflectivity (dB*Z*) from the observational composite mosaic (OBS), (middle) corresponding EnKF analysis, and (right) NoDA ensemble forecast mean valid at 1200 UTC 12 Sep, and 0000 and 1200 UTC 12 Sep 2007, respectively.

with the tropical cyclone using a vortex-tracking method (Chen et al. 2007) and two-way nesting. Experiment 1.5KM is initialized at 1800 UTC 12 September with the same EnKF analysis from the 4.5-km domain discussed above. Figure 10 shows the 1.5KM-simulated minSLP and maxWSP in comparison to both the best-track estimate and the 4.5-km control experiment starting at the same time. Aside from having slightly slower movement, the simulated track in this 1.5-km forecast is nearly identical to that of the 4.5-km control forecast initialized at the same time (Fig. 10a). The peak intensity is 5 hPa weaker in terms of minSLP (Fig. 10b) and is comparable in terms of maxWSP (Fig. 10c). The difference is well within the uncertainties of the EnKF analysis (Fig. 4), and the ensemble forecasts started at the same time with the EnKF analysis perturbations (next subsection).

Although the storm in 1.5KM is slightly weaker in terms of peak intensity, Fig. 11 shows that the higher-resolution simulation captures a much more realistic detailed mesoscale structure (e.g., 0000 and 0600 UTC 13 September) of the cyclone. This additional structure compares more favorably to the observed radar reflectivity at different stages of the cyclone development than that of the 4.5-km control forecast. In particular, it is only the 1.5KM run that captures the strong asymmetry of the observed reflectivity at 0600 UTC 13 September. Future studies will perform high-resolution forecasts at different times and/or utilize the EnKF analysis on the 1.5-km model domain.

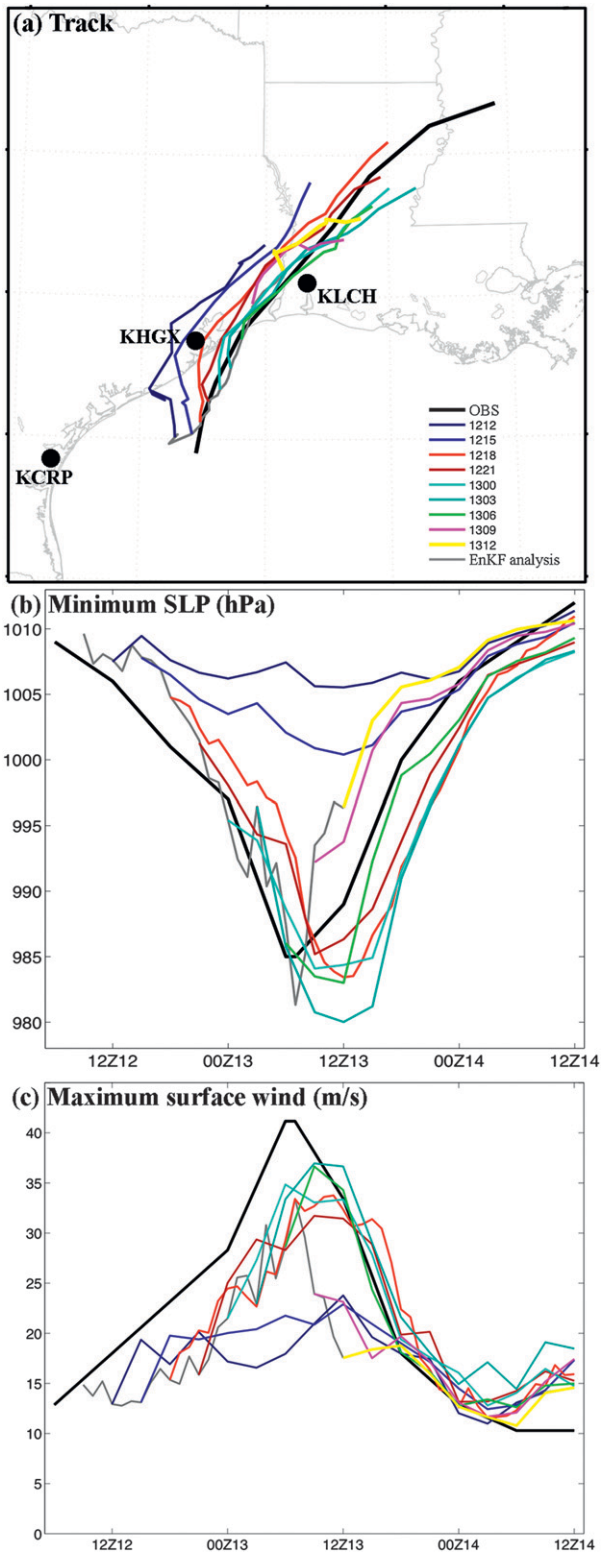

FIG. 8. The simulated (a) positions, (b) minSLP, and (c) maxWSP of Humberto in the deterministic WRF forecasts (color curves) initialized with the EnKF analyses every 3 h from 1200 UTC 12 Sep to 1200 UTC 13 Sep 2007 in comparison with the NHC best-track estimate (black).

d. Ensemble forecasts from EnKF analysis

Large variations between deterministic forecasts initialized with the EnKF mean analysis at different times and the difficulty in real-time operational forecasting suggest that the predictability of this storm is rather limited. The EnKF analysis helps to understand this predictability problem because it provides consistent, flow-dependent uncertainties that are used for initializing ensembles for probabilistic prediction.

A 30-h ensemble forecast is initiated with the control EnKF analysis and perturbations at 1800 UTC 12 September. Consistent with the large variations between deterministic forecasts starting with EnKFs at different times (Fig. 8), there is also large spread among different members from the EnKF-initialized ensemble. This is shown in Fig. 12, which plots the evolution of the position, minSLP, and maxWSP as simulated by all members and the deterministic forecast initialized from the EnKF analysis (mean). The spread of minSLP triples over 30 h from less than 1.5 hPa at 1800 UTC 12 September (the initial time) to 4.5 hPa at 0900 UTC 13 September, while the spread of maxWSP grows from 1.6 to 6.3 m s^{-1} during the same period (even larger ensemble spreads are observed at 0700 and 0800 UTC 13 September; not shown).

The large forecast uncertainty and sensitivity to initial perturbation can also be clearly seen in Fig. 13, which shows a scatterplot of minSLP in the ensemble members at 1800 UTC 12 September (i.e., the initial time for the ensemble) versus minSLP at 0900 UTC 13 September (i.e., a 15-h forecast near the peak intensity time). The strong correlation (~0.7) between the initial minSLP and minSLP at the time of peak intensity highlights the importance of the initial analysis accuracy. The initial ensemble spread at 1800 UTC 12 September is comparable to or even smaller than the typical errors in the best-track estimates, which further demonstrates the limited predictability of deterministic forecasts even after radar observations are assimilated. Thus, the need for probabilistic/ensemble forecasting of tropical cyclones is clear.

To further exemplify error growth between ensemble members, Fig. 14 shows the simulated maximum reflectivity from one of the weakest members and one of the strongest members at 1800 UTC 12 September and 0900 UTC 13 September (based on minSLP at 0900 UTC 13 September; marked in Fig. 13). Despite having apparently similar structures and strengths at 1800 UTC 12 September, the two members diverge tremendously by 0900 UTC 13 September, again signifying large uncertainties in the deterministic forecasts of hurricanes. Ongoing studies are currently investigating both the

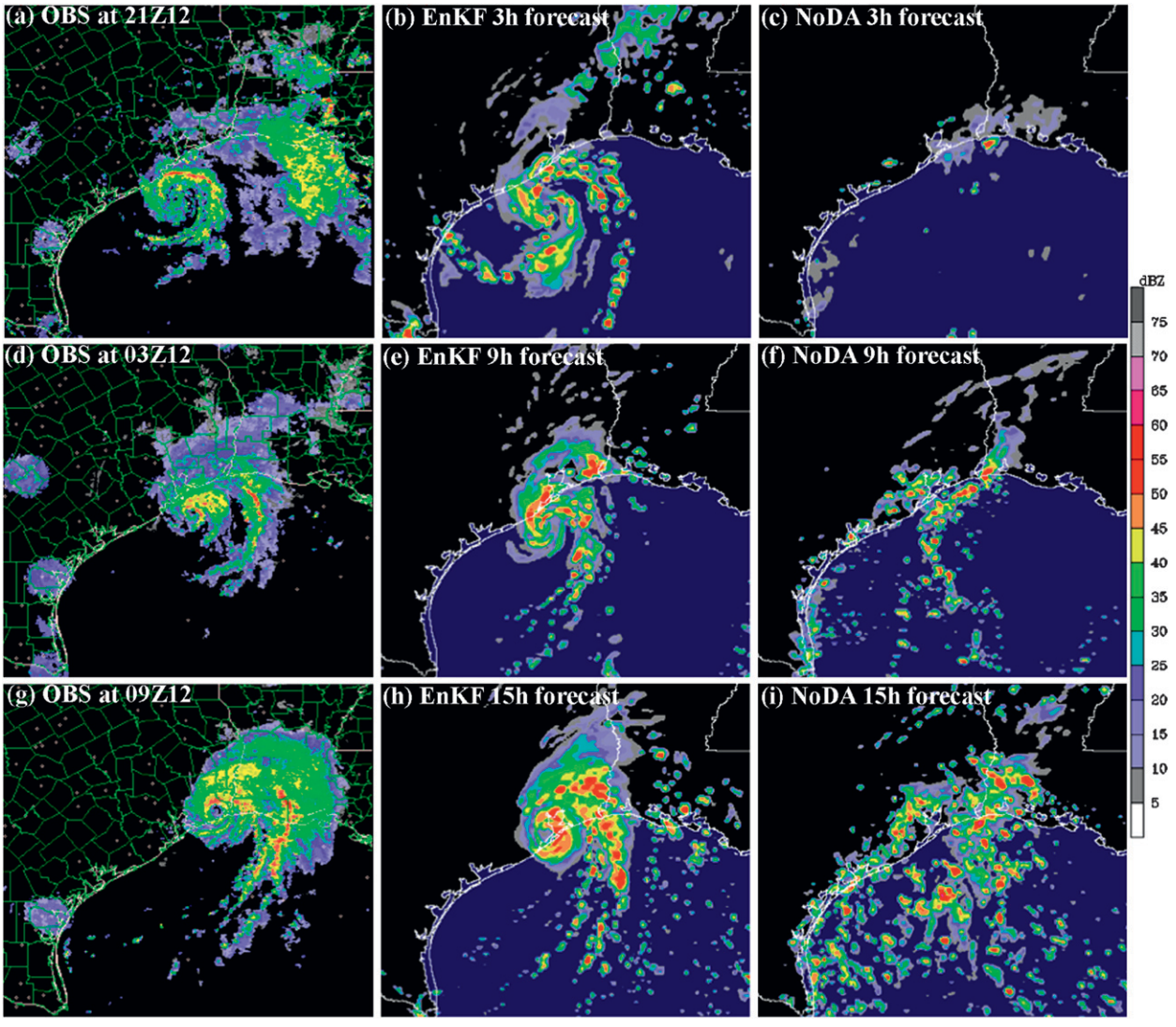

FIG. 9. Comparison of radar reflectivity (dBZ) from (left) the observational composite mosaics (OBS), (middle) the deterministic forecast initialized with the EnKF analysis at 1800 UTC 12 Sep 2007, and (right) with the NoDA ensemble forecast mean valid at 2100 UTC 12 Sep, and 0300 and 0900 UTC 13 Sep 2007, respectively.

mechanism leading to rapid tropical cyclone formation and intensification and the dynamics that lead to the rapid error growth for this event. Recently, studies revealed that upscale growth of moist convection, such as in the form of vertical hot towers, may play a critical role in the internal dynamics (Krishnamurti et al. 2005; Montgomery et al. 2006). Limited predictability of moist convection could ultimately limit the predictability of tropical cyclones (Sippel and Zhang 2008; Zhang and Sippel 2009), as is the case for extratropical cyclones (Zhang et al. 2002, 2003, 2007; Tan et al. 2004) or continental warm season mesoscale convective systems (Zhang et al. 2006b; Hawblitzel et al. 2007; Bei and Zhang 2007).

e. Sensitivity experiments

In the control EnKF analysis and forecasts discussed above, we use the SCL technique that is designed to assimilate dense radar observations that contain information about the state of the atmosphere at a wide range of scales and also to reduce the computational costs and sampling errors (refer to section 2e). Four sensitivity experiments are performed to examine the effectiveness of SCL in comparison to typical covariance localization. Experiments FIX1, FIX2, and FIX3 assimilate the same SOs as in the control EnKF analysis except that fixed ROIs of 1215, 405, and 135 km, respectively, are used for all domains. Experiment DX30 uses a fixed

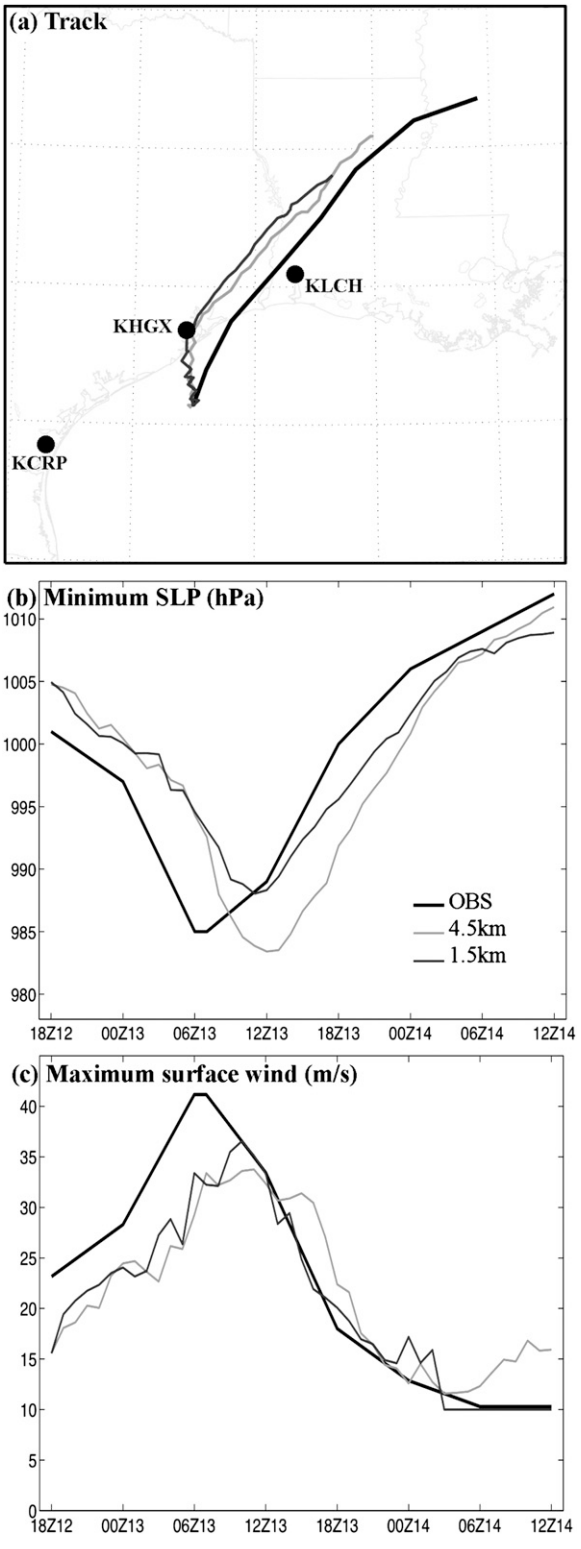

FIG. 10. The simulated (a) position, (b) minSLP, and (c) maxWSP of Humberto from the 1.5-km deterministic forecast initialized with the EnKF analysis at 1800 UTC 12 Sep 2007 (1.5KM) in comparison with the 4.5-km forecast and the best-track estimate.

ROI in the unit of horizontal grid spacings in each domain (30 grid points) but will have varying ROIs in physical distance (1215 km for D1, 405 km for D2, and 135 km for D3). DX30 is somewhat similar to the control EnKF analysis with SCL except that the ROIs will not change in a given domain.

The forecast performance of the WRF simulations initialized at 1800 UTC 12 September with analyses from these four experiments in comparison to the control EnKF analysis is displayed in Fig. 15. Neither FIX1 or FIX2 with a fixed ROI captures the rapid development of the storm while DX30, which uses different ROIs for different domains, has decent tropical development, though it is significantly weaker than using the control EnKF analysis with SCL. FIX3 also simulates noticeably stronger development of the storm than do either FIX1 or FIX2 (implying the benefits and need for tighter covariance localization for the convective-scale radar observations) but its results are progressively weaker than those of DX30 and the CNTL analysis (implying the benefit of updating larger scales with convective-scale radar observations using broader covariance localization in addition to the tighter localization for the convective scales). The track forecasts from all four sensitivity experiments are significantly worse than that using the SCL technique. Nevertheless, we acknowledge that the assignment of the numbers of SOs to different bins and the selections of different ROIs for different groups of SOs are still empirical at present and deserve more in-depth study in the future.

Since we use the relaxation method as in Eq. (1) to inflate the covariance in order to avoid filter divergence, a weighting coefficient, α, set to 0.8 (similar to Meng and Zhang 2008a,b; Torn and Hakim 2008) implies that an overwhelmingly large portion of the final variance comes from the prior. A large value of α is found to be necessary for imperfect-model experiments (Meng and Zhang 2007) to compensate for unavoidable imperfections in the forecast model. However, in the presence of significant model error (and with limited ensemble size and different covariance localization), it is unclear whether the data assimilation configurations are optimal. We performed two experiments (MIX1 and MIX2) with α values of 0.5 and 0.65, respectively, to examine the sensitivity of the EnKF's performance to the choice of inflation factor (i.e., the relaxation coefficient α). Neither experiment gave better analyses or forecasts in terms of both track and intensity (Fig. 16) compared to the CNTL experiment (with weighting coefficient $\alpha = 0.8$), despite better agreement between the mean-square error and the predicted innovation variance (not shown), which should be more desirable (Houtekamer et al. 2005). Results from these two sensitivity experiments

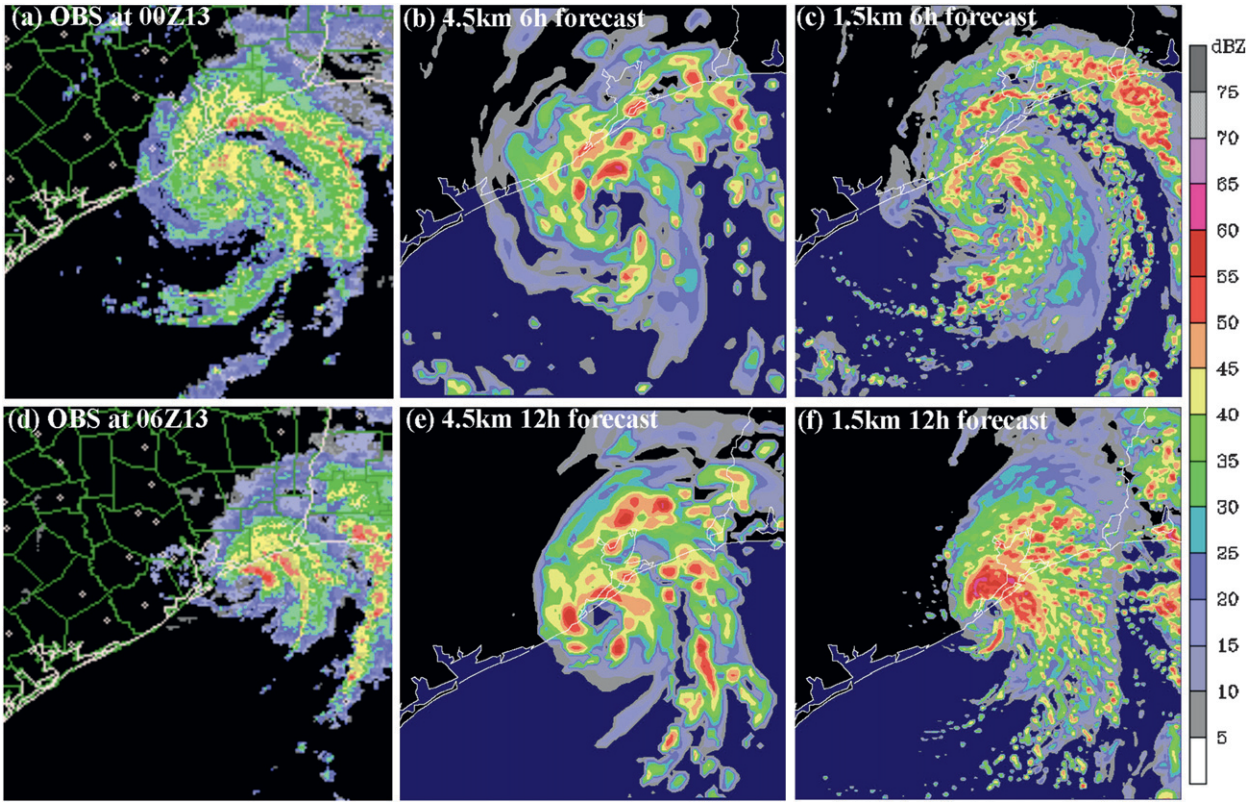

FIG. 11. Comparison of radar reflectivity (dBZ) from (left) observational composite mosaics (OBS) with (middle) those derived from the 4.5-km control forecast, and (right) the 1.5-km forecast valid at 0000 and 0600 UTC 13 Sep 2007, respectively.

are consistent with our recent studies (Meng and Zhang 2008a,b) for regional-scale real-data applications. Several factors may have contributed to this mismatch between the mean-square error and predicted innovation variance: 1) the observational uncertainty is usually inflated to be conservative; 2) the consistency assumption is based on perfect-model and unbiased statistics, which are unattainable in real-data experiments; and 3) there are some decaying modes (such as numerical noise) in the ensemble variance that may not be projected well onto the ensemble mean error.

We also performed experiments examining the impacts of assimilating conventional observations in comparison to the radar observations used in the CNTL analysis (Fig. 17). Experiment OBS1 assimilates all of the conventional observations archived in the Meteorological Assimilation Data Ingest System (MADIS) dataset hourly from 0900 to 1800 UTC 12 September, including radiosonde, wind profiler, mesonet, METAR (routine aviation weather report), maritime, and satellite-cloud-derived wind observations. Forecasts initialized from the EnKF analysis in OBS1 at 1800 UTC 12 September are not capable of capturing the rapid development of Humberto (Figs. 16c and 16d). On the other hand, experiment OBS2 assimilates hourly all conventional observations archived in MADIS along with the same radar observations used in the CNTL EnKF analysis. Not surprisingly, forecasts initialized with the EnKF analysis in OBS2 simulate the rapid development of the storm. However, at least for the relatively shorter time scales considered in this case, there is no apparent improvement in the analyses and forecasts by assimilating conventional observations in addition to the radar observations used in the CNTL experiment. The minimum SLP from OBS2 is not as low as in the CNTL run but it is well within the range of uncertainty shown in the control ensemble forecasts (Fig. 12). These additional sensitivity experiments further demonstrate the importance of the convective-scale radial velocity observations provided by the coastal WSR-88Ds. The success of the deterministic forecasts from the EnKF analyses of CNTL and OBS2 suggests that the larger-scale flow may be good enough not to adversely impact the storm development with the 24–48-h periods that we examined.

4. Comparison with WRF-3DVAR

Since the data assimilation schemes used in operational forecast models at NCEP at the time of Humberto

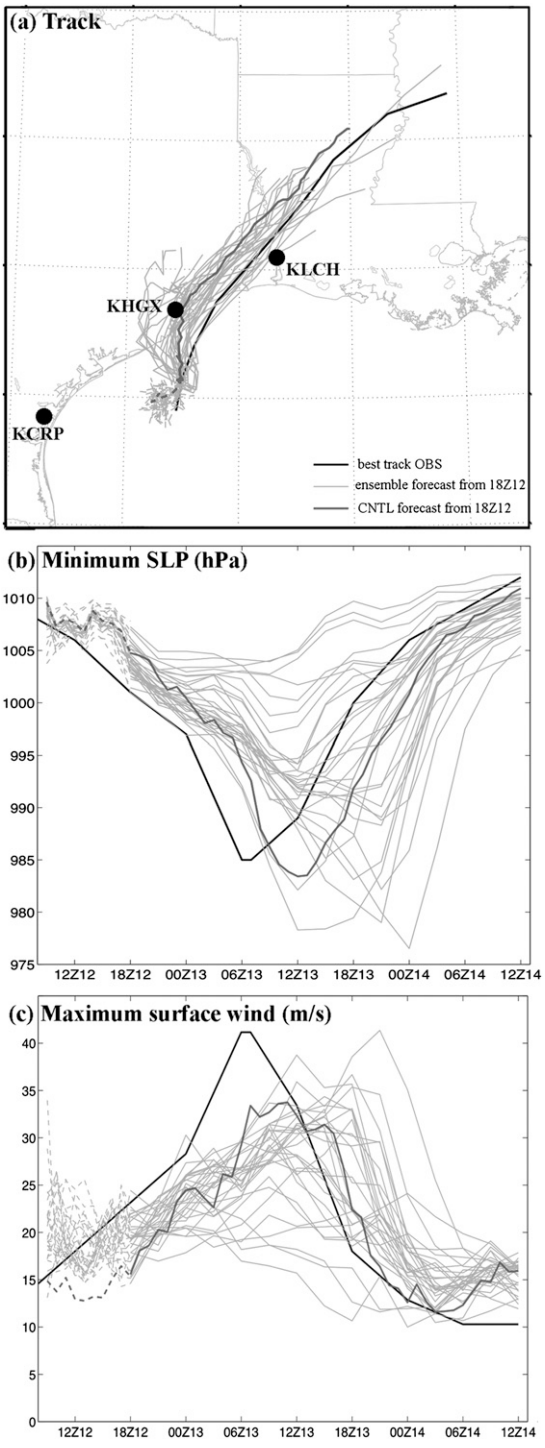

FIG. 12. The simulated (a) position, (b) minSLP, and (c) maxWSP of Humberto by the ensemble forecast initialized with the EnKF perturbations at 1800 UTC 12 Sep 2007 (light gray) in comparison to the deterministic forecast initialized from the EnKF mean analysis (dark gray) and the NHC best-track estimate (black). Analysis and uncertainty until 1800 UTC 12 Sep 2007 are also shown in dashed curves.

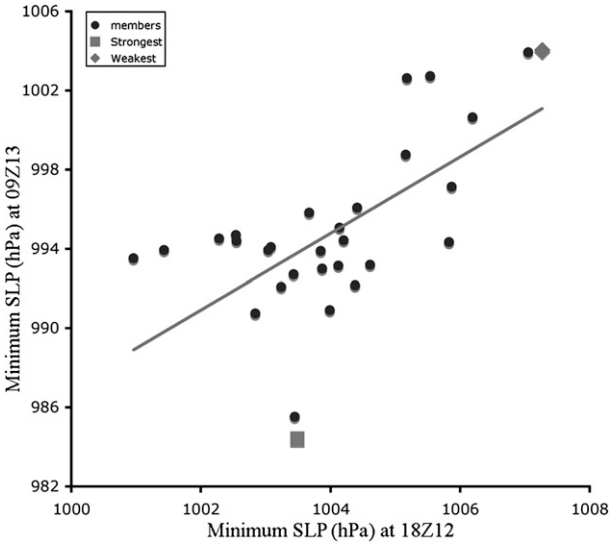

FIG. 13. The scatterplot of the forecasted minimum SLP at 1800 UTC 12 Sep (x axis; 0 h) and 0900 UTC 13 Sep 2007 (y axis; 15 h) in different ensemble members with the weakest and strongest members at 0900 UTC 13 Sep highlighted.

were based on the 3DVAR methodology, here we use the WRF model and its 3DVAR system to assimilate exactly the same observations for a comparison with the EnKF analysis discussed above.

The WRF-3DVAR method used here was developed primarily at NCAR and is now operational at the Air Force Weather Agency (Barker et al. 2004). Its configuration is based on an incremental formulation, producing a multivariate analysis in the model space. Its incremental cost function is minimized in a preconditioned control variable space where the errors of different control variables are largely uncorrelated. As in any other variational data assimilation technique, the structure of the background error covariance may play a very important role in the 3DVAR system. Experiment ''3DVAR1'' uses the WRF-3DVAR default background error statistic (its ''cv'' option 3 originated from an earlier version of the NCEP GFS system). The first guess comes from the 9-h WRF 4.5-km (single, deterministic) forecast initialized with the GFS Final Global Data Assimilation System (FNL) analysis at 0000 UTC 12 September. WRF/3DVAR begins assimilation at 0900 UTC 12 September and continues to assimilate exactly the same radar observations as in the EnKF analysis hourly until 1800 UTC 12 September, after which time a 30-h forecast without further data assimilation is performed. Experiment 3DVAR2 is performed in the same manner as 3DVAR1 except that a 30-member 12-h short-term ensemble forecast initialized at 0000 UTC 12 September (the same as in the initial ensemble

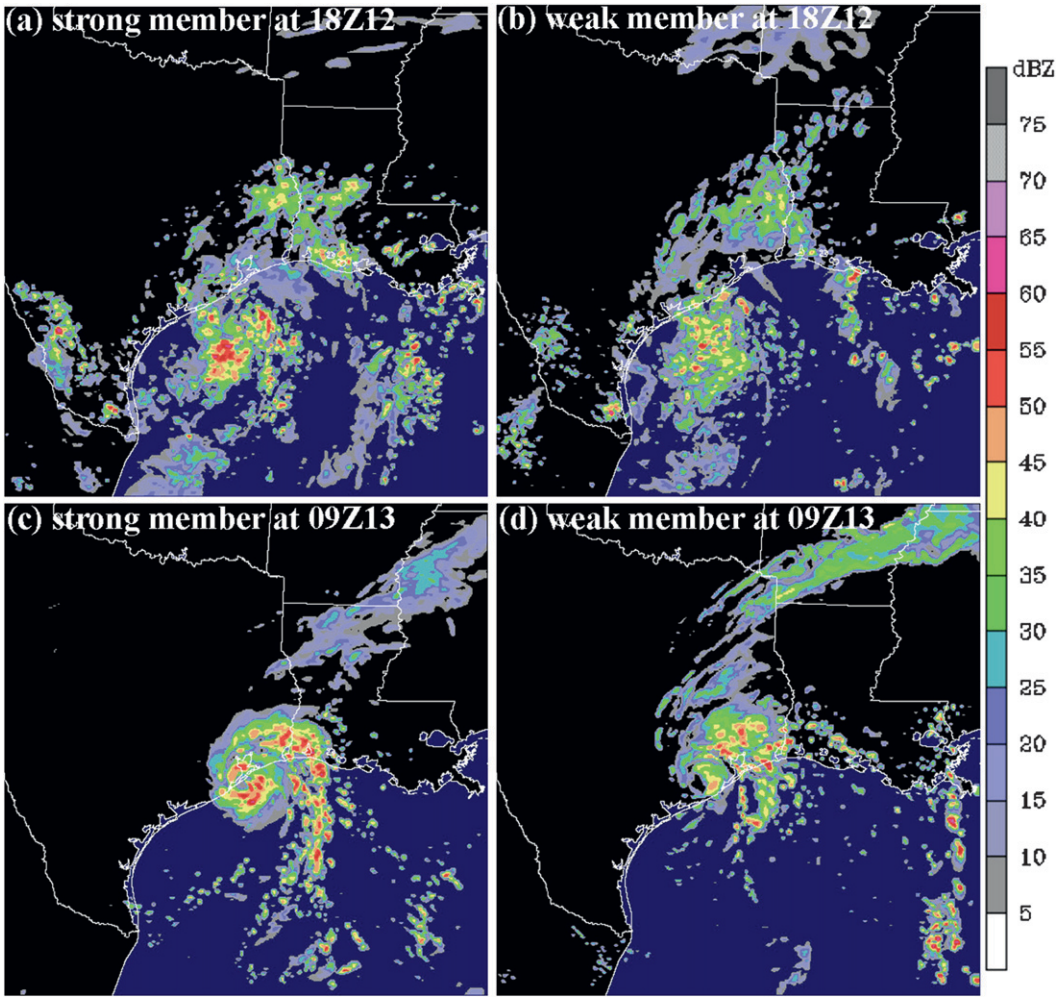

FIG. 14. The comparison of the simulated maximum reflectivity (dBZ) derived from the (left) strongest member and (right) weakest member of the ensemble valid at 1800 UTC 12 Sep and 0900 UTC 13 Sep 2007, respectively.

for the EnKF analysis) is used to recalculate the WRF/3DVAR background error statistics [its "cv" option 5, as in Meng and Zhang (2008a)].

Figure 18 shows the evolution of minSLP and maxWSP from both WRF/3DVAR experiments in comparison to the corresponding EnKF analysis and forecast. Although both 3DVAR and EnKF assimilate exactly the same observations with the same model resolution, both 3DVAR experiments fail to perform satisfactorily. This result is in spite of the fact that the 3DVAR analyses may have a better fit of maximum wind speed at 1800 UTC 12 September before the pure forecast starts. We acknowledge that the performance of 3DVAR may be further improved through further tuning of the background error covariance (some improvement is seen when using the ensemble to generate the background error covariance in 3DVAR2) and/or with the addition of initial vortex bogussing (Q. Xiao, NCAR, 2008, personal communication), but the sensitivity of the performance of 3DVAR to different configurations and representations of background error statistics is beyond the scope of this study. On the other hand, there may still be room to improve the performance of EnKF through further tuning, which is also beyond the scope of the current study.

5. Summary and conclusions

This study explores the uses of Doppler radar observations for cloud-resolving hurricane analysis, initialization, and prediction with an ensemble Kalman filter (EnKF). The case studied is Hurricane Humberto (2007), the first landfalling hurricane in the United States since the end of the 2005 hurricane season and the most rapidly intensifying near-landfall storm in U.S. history. The storm caused extensive damage along the southeast

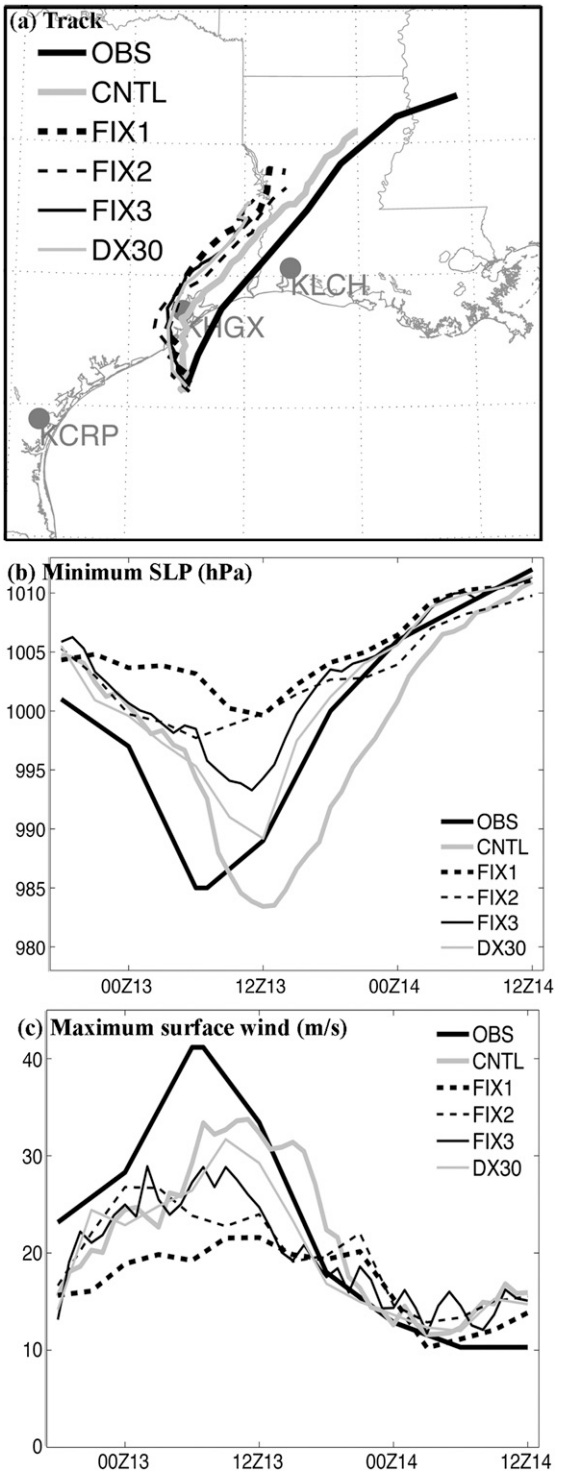

FIG. 15. The simulated (a) position, (b) minSLP, and (c) maxWSP of Humberto from four deterministic forecasts initialized with the EnKF analyses using fixed radii of influence (1215 km in FIX1, 405 km in FIX2, 135 km in FIX3, and 30 grid points for each domain in DX30) in comparison with the 4.5-km forecast from the control EnKF analysis using SCL and the best-track estimate.

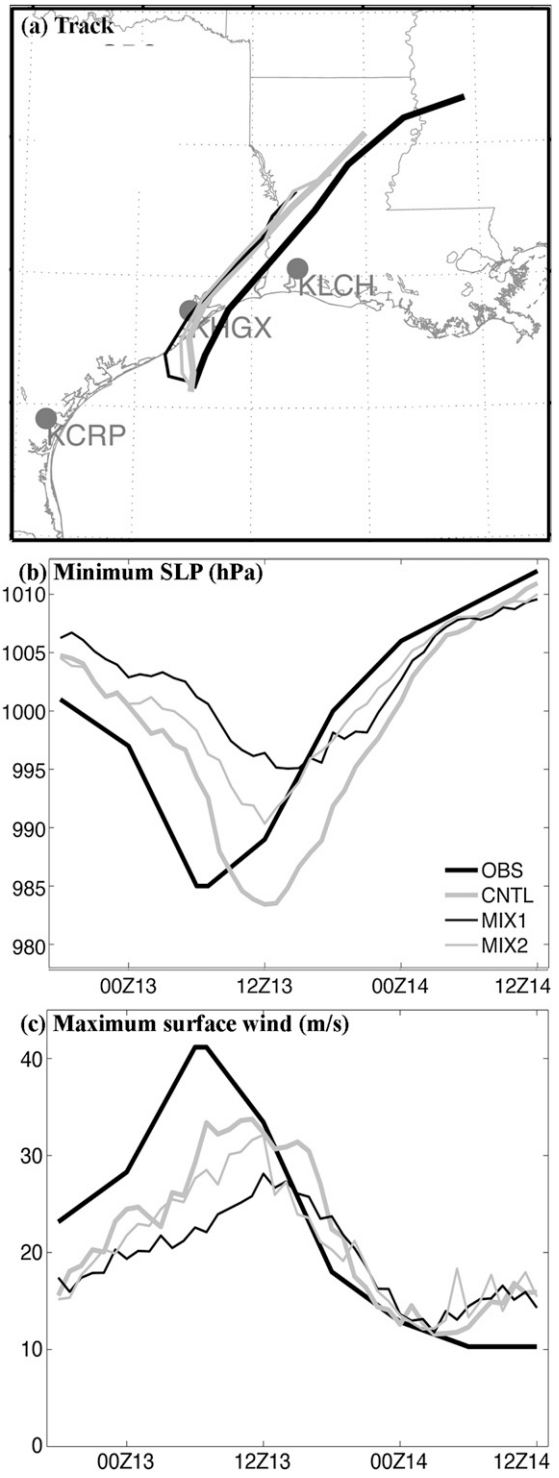

FIG. 16. The simulated (a) position, (b) minSLP, and (c) maxWSP of Humberto from two deterministic forecasts initialized with the EnKF analyses using different covariance relaxation coefficients ($\alpha = 0.5$ for MIX1 and $\alpha = 0.65$ for MIX2) in comparison to the 4.5-km forecast from the control EnKF analysis using SCL and the best-track estimate.

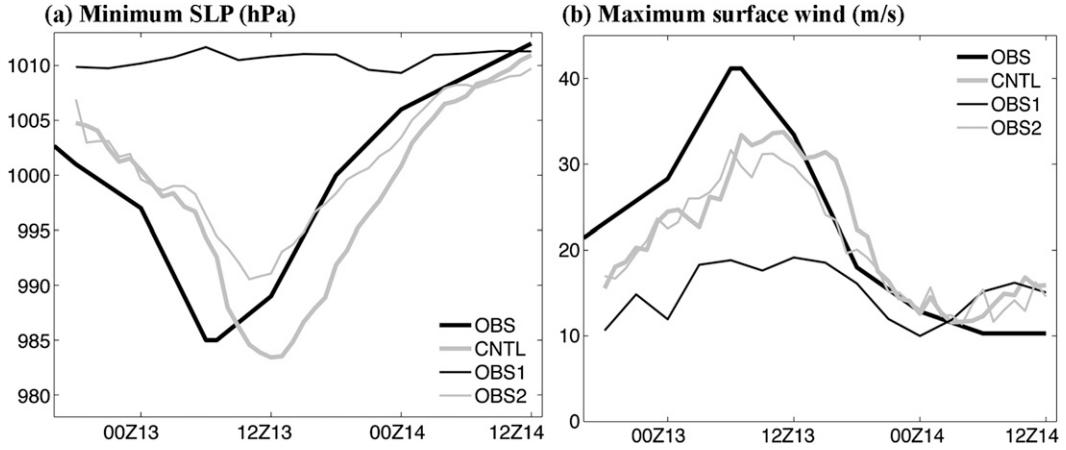

FIG. 17. The simulated (a) minSLP and (b) maxWSP of Humberto from the deterministic forecasts initialized with two different EnKF analyses (OBS1 and OBS2) at 1800 UTC 12 Sep 2007 in comparison with the 4.5-km forecast from the EnKF analysis and the best-track estimate. OBS1 assimilates only conventional observations while OBS2 assimilates both conventional observations and the radar observations.

Texas coast but was poorly predicted by operational models and forecasters. It is found that the EnKF analysis, after assimilating radial velocity observations from three WSR-88Ds along the Gulf coast, closely represents the best-track position and intensity of Humberto. Deterministic forecasts initialized from the EnKF analysis, despite having considerable variability with different lead times, are also capable of predicting the rapid formation and intensification of the hurricane. These forecasts are superior to operational forecasts, simulations without radar data assimilation, and forecasts initialized with the assimilation of the same observations with a three-dimensional variational method implemented with the same forecast model. Moreover, ensemble forecasts initialized with EnKF analysis perturbations before the rapid intensification show large spread among ensemble members. Such large spread further exemplifies the significant uncertainties in the deterministic prediction of hurricanes, especially the intensity forecasts.

In this study, EnKF demonstrates great promise in assimilating Doppler radar observations to initialize hurricanes with detailed, accurate mesoscale structures. Even though ground-based radar may only have

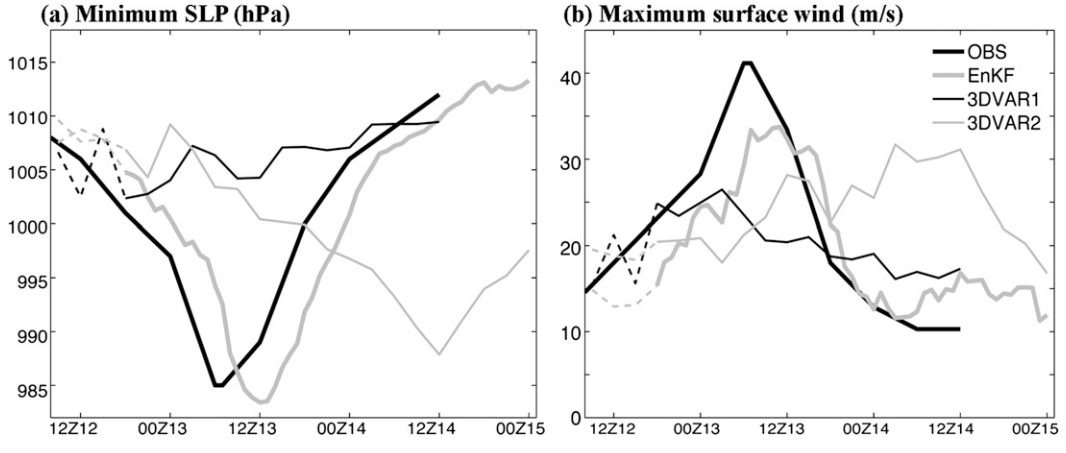

FIG. 18. The simulated (a) minSLP and (b) maxWSP of Humberto from the deterministic forecasts initialized with two different WRF-3DVAR analyses (3DVAR1 and 3DVAR2) at 1800 UTC 12 Sep 2007 in comparison with the 4.5-km forecast from the EnKF analysis and the best-track estimate. The WRF-3DVAR default background error covariance statistics is used in 3DVAR1 while a 12-h ensemble is used to generate the background statistics in 3DVAR2.

limited range in providing observations for hurricane prediction with lead times beyond 24 h, this may be complemented with airborne radar with Doppler observations and longer lead times that will soon become available for all hurricane aircraft reconnaissance missions (F. Marks, NOAA/HRD, 2008, personal communications). However, since only radial winds from three radars are assimilated in this study, the setup in these experiments may only be most effective for short periods of time. Beyond a couple of days, better constraints by observations at the synoptic scales will be needed to further improve hurricane track and intensity forecasts.

Future studies are planned to examine the dynamics and predictability of Humberto with the EnKF analysis and forecasts. It remains unclear what observations are necessary and sufficient to define the initial tropical cyclone vortex and large-scale environment. Answers to these questions have strong implications related to how society might better distribute resources to cope with future hurricane-related disasters. This is extremely important given that the number of hurricanes and their intensity/destructiveness are reportedly on the rise with the warming climate (Emanuel 2005; Webster et al. 2005).

Despite the promising performance of both deterministic and probabilistic forecasts from the EnKF analysis, the intrinsic limit of hurricane predictability (i.e., in the face of nearly perfect observations and initialization) remains unclear in terms of both track and intensity forecasts. The limit of formation–intensity predictability given realistic initial conditions and model errors (which are still large at present) in numerical weather prediction models may be alleviated through improving our understanding of their dynamics and physics, the development of better numerical models, and improved data coverage and assimilation techniques. However, there always will be forecast errors due to the inherent limit of predictability arising from initial errors with amplitudes far smaller than any observation or analysis system (e.g., Zhang and Sippel 2009); these are errors that society will always have to cope with (Pielke 1997).

Such inherent uncertainties in hurricane forecasts highlight the need for developing advanced ensemble prediction systems to provide event-dependent probabilistic forecasts and risk assessments. In practice, despite an increasing role and the demonstrated benefits of using ensembles in aiding deterministic hurricane forecasting (Krishnamurti et al. 1999), the uncertainty involved with today's operational hurricane forecasts is still based on averaged climatological errors and thus is not case dependent. This case clearly demonstrates that uncertainty can be quite large at some times (e.g., Sippel and Zhang 2008) and having access to such information in the operational environment would serve forecasters well.

Acknowledgments. The authors benefited from discussions with Kerry Emanuel, Chris Snyder, Chris Davis, Jim Hansen, Jim Doyle, Pete Black, Tim Dunkerton, Shuyi Chen, and Dave Nolan. We thank two anonymous reviewers and the editor Tom Hamill for their valuable review comments. This research is supported by NSF Grant ATM-0205599 and by the Office of Navy Research under the Young Investigator Program (Award N000140410471).

REFERENCES

Alpert, J. C., and V. K. Kumar, 2007: Radial wind super-obs from the WSR-88D radars in the NCEP operational assimilation system. *Mon. Wea. Rev.,* **135,** 1090–1108.

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.,* **129,** 2884–2903.

Barker, D. M., 2005: Southern high-latitude ensemble data assimilation in the Antarctic mesoscale prediction system. *Mon. Wea. Rev.,* **133,** 3431–3449.

——, W. Huang, Y.-R. Guo, A. J. Bourgeois, and Q. N. Xiao, 2004: A three-dimensional variational data assimilation system for MM5: Implementation and initial results. *Mon. Wea. Rev.,* **132,** 897–914.

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.,* **3,** 396–409.

Bei, N., and F. Zhang, 2007: Mesoscale predictability of the torrential rainfall along the mei-yu front of China. *Quart. J. Roy. Meteor. Soc.,* **133,** 83–99.

Bishop, C. H., and D. Hodyss, 2007: Flow-adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. *Quart. J. Roy. Meteor. Soc.,* **133,** 2029–2044.

Chen, S. S., J. F. Price, W. Zhao, M. A. Donelan, and E. J. Walsh, 2007: The CBLAST-Hurricane Program and the next-generation fully coupled atmosphere–wave–ocean models for hurricane research and prediction. *Bull. Amer. Meteor. Soc.,* **88,** 311–317.

Chessa, P. A., G. Ficca, M. Marrocu, and R. Buizza, 2004: Application of a limited-area short-range ensemble forecast system to a case of heavy rainfall in the Mediterranean region. *Wea. Forecasting,* **19,** 566–581.

Davis, C. A., and Coauthors, 2008: Prediction of landfalling hurricanes with the advanced hurricane WRF model. *Mon. Wea. Rev.,* **136,** 1990–2005.

Dowell, D. C., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and thermodynamic retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. *Mon. Wea. Rev.,* **132,** 1982–2005.

Elsberry, R. L., T. D. B. Lambert, and M. A. Boothe, 2007: Accuracy of Atlantic and eastern North Pacific tropical cyclone intensity forecast guidance. *Wea. Forecasting,* **22,** 747–762.

Emanuel, K., 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature,* **436,** 686–688.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99,** 10 143–10 162.

——, 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.,* **53,** 343–367.

Franklin, J. L., 2004: National Hurricane Center verification report. Preprints, *57th Interdepartmental Hurricane Conf.*, Miami, FL, Office of the Federal Coordinator for Meteorological Services and Supporting Research. [Updates available online at www.nhc.noaa.gov/verification.]

Fujita, T., D. J. Stensrud, and D. C. Dowell, 2007: Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physics uncertainties. *Mon. Wea. Rev.,* **135,** 1846–1868.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.,* **125,** 723–757.

Grell, G. A., and D. Devenyi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.,* **29,** 1693, doi:10.1029/2002GL015311.

Hamill, T. M., 2006: Ensemble-based atmospheric data assimilation. *Predictability of Weather and Climate*, T. Palmer and R. Hagedorn, Eds., Cambridge University Press, 124–156.

Hawblitzel, D., F. Zhang, Z. Meng, and C. A. Davis, 2007: Probabilistic evaluation of the dynamics and predictability of mesoscale convective vortex event of 10–13 June 2003. *Mon. Wea. Rev.,* **135,** 1544–1563.

Hong, S.-Y., J. Dudhia, and S.-H. Chen, 2004: A revised approach to ice-microphysical processes for the bulk parameterization of cloud and precipitation. *Mon. Wea. Rev.,* **132,** 103–120.

Houtekamer, P. L., and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.,* **129,** 123–137.

——, ——, G. Pellerin, M. Buehner, M. Charron, L. Speak, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.,* **133,** 604–620.

Houze, R. A., S. S. Chen, B. F. Smull, W.-C. Lee, and M. M. Bell, 2007: Hurricane intensity and eyewall replacement. *Science,* **315,** 1235–1238.

Krishnamurti, T. N., C. M. Kishtawal, T. LaRow, D. Bachiochi, Z. Zhang, C. E. Williford, S. Gadhil, and S. Surendran, 1999: Improved weather and seasonal climate forecasts from multimodel superensemble. *Science,* **285,** 1548–1550.

——, and Coauthors, 2001: Real-time multianalysis–multimodel superensemble forecasts of precipitation using TRMM and SSM/I products. *Mon. Wea. Rev.,* **129,** 2861–2883.

——, and Coauthors, 2005: The hurricane intensity issue. *Mon. Wea. Rev.,* **133,** 1886–1912.

Lindskog, M., H. Järvinen, and D. B. Michelson, 2000: Assimilation of radar radial winds in the HIRLAM 3D-Var. *Phys. Chem. Earth,* **25B,** 1243–1249.

——, K. Salonen, H. Jarvinen, and D. B. Michelson, 2004: Doppler radar wind data assimilation with HIRLAM 3DVAR. *Mon. Wea. Rev.,* **132,** 1081–1092.

Meng, Z., and F. Zhang, 2007: Test of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect-model experiments. *Mon. Wea. Rev.,* **135,** 1403–1423.

——, and ——, 2008a: Test 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.

——, and ——, 2008b: Test of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part IV: Performance over a warm-season month of June 2003. *Mon. Wea. Rev.,* **136,** 3671–3682.

Montgomery, M. T., M. E. Nicholls, T. A. Cram, and A. B. Saunders, 2006: A vortical hot tower route to tropical cyclogenesis. *J. Atmos. Sci.,* **63,** 355–386.

Montmerle, T., and C. Faccani, 2009: Mesoscale assimilation of radial velocities from Doppler radars in a preoperational framework. *Mon. Wea. Rev.,* **137,** 1939–1953.

Noh, Y., W.-G. Cheon, and S.-Y. Hong, 2003: Improvement of the K-profile model for the planetary boundary layer based on large eddy simulation data. *Bound.-Layer Meteor.,* **107,** 401–427.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.,* **120,** 1747–1763.

Pielke, R. A., 1997: Reframing the U.S. hurricane problem. *Soc. Nat. Resour.,* **10,** 485–499.

Pu, Z. X., and S. A. Braun, 2001: Evaluation of bogus vortex techniques with four-dimensional variational data assimilation. *Mon. Wea. Rev.,* **129,** 2023–2039.

Sippel, J. A., and F. Zhang, 2008: Probabilistic evaluation of the dynamics and predictability of tropical cyclogenesis. *J. Atmos. Sci.,* **65,** 3440–3459.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the Advanced Research WRF version 2. NCAR Tech. Note NCAR/TN-468+STR, 88 pp.

Snyder, C., and F. Zhang, 2003: Tests of an ensemble Kalman filter for convective-scale data assimilation. *Mon. Wea. Rev.,* **131,** 1663–1677.

Tan, Z., F. Zhang, R. Rotunno, and C. Snyder, 2004: Mesoscale predictability of moist baroclinic waves: Experiments with parameterized moist convection. *J. Atmos. Sci.,* **61,** 1794–1804.

Tong, M., 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.

Torn, R. D., and G. J. Hakim, 2008: Performance characteristics of a pseudo-operational ensemble Kalman filter. *Mon. Wea. Rev.,* **136,** 3947–3963.

——, ——, and C. Snyder, 2006: Boundary conditions for a limited-area ensemble Kalman filter. *Mon. Wea. Rev.,* **134,** 2490–2502.

Webster, P. J., G. J. Holland, J. A. Curry, and H.-R. Chang, 2005: Changes in tropical cyclone number, duration, and intensity in a warming environment. *Science,* **309,** 1844–1846.

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. *Mon. Wea. Rev.,* **136,** 463–482.

Xiao, Q., Y. H. Kuo, J. Sun, W. C. Lee, D. M. Barker, and E. Lim, 2007: An approach of radar reflectivity data assimilation and its assessment with the inland QPF of Typhoon Rusa (2002) at Landfall. *J. Appl. Meteor. Climatol.,* **46,** 14–22.

Xu, Q., and J. Gong, 2003: Background error covariance functions for Doppler radial-wind analysis. *Quart. J. Roy. Meteor. Soc.,* **129,** 1703–1720.

Zhang, F., 2005: Dynamics and structure of mesoscale error covariance of a winter cyclone estimated through short-range ensemble forecasts. *Mon. Wea. Rev.,* **133,** 2876–2893.

——, and J. A. Sippel, 2009: Effects of moist convection on hurricane predictability. *J. Atmos. Sci.,* **66,** 1944–1961.

——, C. Snyder, and R. Rotunno, 2002: Mesoscale predictability of the ''surprise'' snowstorm of 24–25 January 2000. *Mon. Wea. Rev.,* **130,** 1617–1632.

——, ——, and ——, 2003: Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.,* **60,** 1173–1185.

——, ——, and J. Sun, 2004: Tests of an ensemble Kalman filter for convective-scale data assimilation: Impact of initial estimate and observations. *Mon. Wea. Rev.,* **132,** 1238–1253.

——, Z. Meng, and A. Aksoy, 2006a: Test of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part I: Perfect-model experiments. *Mon. Wea. Rev.,* **134,** 722–736.

——, A. Odins, and J. W. Nielsen-Gammon, 2006b: Mesoscale predictability of an extreme warm-season rainfall event. *Wea. Forecasting,* **21,** 149–166.

——, N. Bei, R. Rotunno, C. Snyder, and C. C. Epifanio, 2007: Mesoscale predictability of moist baroclinic waves: Cloud-resolving experiments and multistage error growth dynamics. *J. Atmos. Sci.,* **64,** 3579–3594.

Zou, X., and Q. 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.