Assimilating Airborne Doppler Radar Observations with an Ensemble Kalman Filter for Convection-permitting Hurricane Initialization and Prediction: Katrina (2005)

Yonghui Weng and Fuqing Zhang

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

Submitted to Monthly Weather Review for publication

Accepted, 9 Sepetember 2011

Corresponding author address: Prof. Fuqing Zhang, Department of Meteorology, The Pennsylvania State University, University Park, PA 16802. Email: fzhang@psu.edu

Abstract

Through a WRF-based ensemble Kalman filter (EnKF) data assimilation system, the impact of assimilating airborne radar observations for the convection-permitting analysis and prediction of Hurricane Katrina (2005) is examined in this study. A forecast initialized from EnKF analyses of airborne radar observations had substantially smaller hurricane track forecast errors than the NOAA's operational forecasts and a control forecast initialized from NCEP analysis data for lead times up to 120 h. Verifications against independent in-situ and remotely sensed observations show that EnKF analyses successfully depict the inner-core structure of the hurricane vortex in terms of both dynamic (wind) and thermodynamic (temperature and moisture) fields. In addition to the improved analyses and deterministic forecast, an ensemble of forecasts initiated from the EnKF analyses also provided forecast uncertainty estimates for the hurricane track and intensity.

Also documented here are the details of a series of data thinning and quality control procedures that were developed to generate super-observations from large volumes of airborne radial velocity measurements. These procedures have since been implemented operationally on the NOAA hurricane reconnaissance aircraft that allows for more efficient real-time transmission of airborne radar observations to the ground.

1. Introduction

Hurricanes are among the costliest and deadliest natural disasters. Hurricane forecast accuracy of dynamic models depends tremendously on how well the initial vortex specified in a numerical weather prediction (NWP) model represents the observed hurricane intensity, structure and internal dynamics. A lack of sufficient observations over the tropical oceans and the dynamically inconsistent depiction of tropical cyclones (TCs) by operational data assimilation systems may be one of the largest factors contributing to inadequate vortex initialization in NWP forecasting efforts (Rogers et al. 2006; Zhang et al. 2009, hereafter refer to as Z09).

Methods of hurricane initialization that have been applied both operationally and in research experiments can be classified into two categories with regard to methodology: 1) vortex bogussing and 2) data assimilation. A bogussing scheme attempts to artificially create a TC circulation in the model initial conditions by inserting a balanced idealized or realistic vortex. One approach of this scheme can be to insert a bogus vortex circulation into the initial data fields before model initialization. The bogus vortex circulation is generated from an analytical expression after the objective analysis or from a model forecast with a mature hurricane or from the previous forecast of the same model, which can then be adjusted to fit the targeted TC's size, location and intensity (Kurihara et al., 1993; Ueno, 1995). Another approach can be to produce bogus observations based on the observed storm's intensity and size, and then assimilate the synthetic observations to initialize the vortex circulation. One example of this method is the bogus data assimilation (BDA) method of Zou and Xiao (2000) that assimilated bogus sea-level pressure and its derived wind observations through a variational procedure in order to generate all other model variables in balance with an initial state adapted for the prediction model.

The bogussing schemes are the most commonly used initialization methods for the current-generation operational models for TC forecasting. For example, the Navy Operational Global Atmospheric Prediction System (NOGAPS) assimilates bogus observations into the model initial fields in order to correct the TC location and intensity[1]. The data assimilation system of the operational Global Forecast System (GFS) at the National Centers for Environmental Prediction (NCEP), three-Dimensional Variational (3DVar) Gridpoint Statistical Interpolation (GSI), has been used to assimilate tropical storm pseudo sea-level pressure observations for TC vortex initialization since 15 December 2009[2]. The relocation technique developed by Kurihara et al. (1995) for the Geophysical Fluid Dynamics Laboratory (GFDL) hurricane prediction model uses a vortex separation method to remove the initial vortex from the first guess field and reinserts a new tropical storm vortex integrated by the model into the guess field. The Hurricane Weather Research and Forecast (HWRF) model, the next generation regional-scale hurricane prediction model at NCEP, uses a three-Dimensional Variational (3DVar) Gridpoint Statistical Interpolation (GSI) scheme to update a first guess vortex from a previous model run for insertion into the model initial fields[3].

Bogussing schemes have been developed to generate a balanced synthetic vortex that is representative of an observed storm's size, intensities, and even shape, for merging with the model initial field. However, such methods cannot completely represent realistic vortex structure. Recent studies have shown that the inner-core flow asymmetries in a TC are random (e.g., Nguyen et al. 2008; Fang and Zhang 2010, 2011). Their work also indicated that the horizontal length scale of convectively induced vorticity anomalies in a vertical-plume-dominated regime was less than the storm scale (~100 km) but larger than the scales of

[1] More model information is documented online at http://www.srh.noaa.gov/ssd/nwpmodel/html/nhcmodel.htm
[2] More details refer to http://www.emc.ncep.noaa.gov/gmb/STATS/html/model_changes.html
[3] More details refer to http://www.nhc.noaa.gov/modelsummary.shtml

individual tropical cumulus clouds (~1 km). Such small-scale stochastic convective systems cannot be realistically represented by empirical bogussing schemes. Therefore, it is necessary to use an advanced data assimilation scheme to assimilate high-resolution observations in order to represent the realistic mesoscale structure of TCs.

As an alternative advanced approach to variational data assimilation, the ensemble Kalman filter (EnKF) uses an ensemble of forecasts to estimate flow-dependent background error covariance (Evensen 1994). The advantage of the EnKF over variational data assimilation methods is its ability to estimate model forecast uncertainties and its ability to assimilate observations from convective to regional scales (e.g., Snyder and Zhang 2003; Dowell et al. 2004; Zhang and Snyder 2007). Studies assimilating real observations with an EnKF have demonstrated significant improvements on hurricane analyses and forecasts. For example, Torn and Hakim (2009) assimilated observations from Automated Surface Observing System (ASOS) stations, ships, buoys, rawinsondes, the Aircraft Communications Addressing and Reporting System (ACARS) and cloud motion vectors into the Advanced Research Weather Research and Forecasting (ARW-WRF) model (Skamarock et al. 2005) over the lifespan of Hurricane Katrina (2005). Their results showed less initial imbalances in the EnKF analyses, in comparison to the GFS initial conditions. In addition, the EnKF may shorten the model spin-up time and also perform better for TC data assimilation than comparable bogussing methods. Z09 adapted the EnKF for a WRF-based regional-scale TC convective-permitting analysis and prediction system capable of assimilating high-resolution ground-based Doppler radar observations. In Z09, the EnKF system outperformed the 3DVar assimilation technique for the initialization and forecast of Hurricane Humberto (2007) with the assimilation of ground-based radar observations. Weng et al. (2011) further compared hurricane initialization methods by assimilating ground-based

Doppler radar observations with 3DVar, 4DVar and EnKF for Hurricane Katrina (2005) and concluded that the EnKF had the largest potential for hurricane initialization at convection-permitting scales.

Hurricanes spend most of their lifetimes over the ocean, where an insufficient number of conventional observations (e.g., buoy, ship and aircraft data) are available for resolving inner-core structures. However Doppler radar observations, both ground-based and airborne, may well cover the inner-core areas and can be used to initialize hurricanes. Research concerning the EnKF assimilation of ground-based radar observations has shown promising results, but is limited to cases where TCs track near the coast (Zhang et al., 2009; Weng et al., 2011). Therefore, a method for assimilating high-resolution in-situ and remotely sensed observations (such as those from satellites and airborne Doppler radars) could yield important benefits for hurricane initialization.

In the current study, the impact of assimilating airborne Doppler radar radial velocity observations for convection-permitting analysis and prediction of hurricanes will be explored in detail. The selected case study is for Hurricane Katrina (2005) (Knabb et al., 2006), which developed from a tropical depression that formed near the Turks and Caicos around 1800 UTC 22 August 2005. It reached tropical storm strength (>17.5 m s^{-1} maximum sustained winds) around 1200 UTC 24 August and attained hurricane strength (>35 m s^{-1} winds) just before making its first landfall on the southeastern coast of Florida during the afternoon of 25 August. The storm weakened slightly to a tropical storm immediately following landfall, but it quickly regained hurricane intensity after crossing into the southeastern Gulf of Mexico. Katrina intensified to a Category 3 hurricane on the Saffir-Simpson scale (> 50 m s^{-1} winds) by 1200 UTC 27 August, and intensified further to Category 5 (> 72.5 m s^{-1} winds) by 1200 UTC 28

August. The storm reached its peak intensity on the afternoon of 28 August with maximum sustained winds of 75 m s^{-1} and a central pressure of 902 hPa, which was the fourth lowest pressure ever recorded by 2010 for an Atlantic tropical cyclone. While inner-core structural changes weakened the hurricane to a strong Category 3 storm with about 920 hPa minimum central pressure as it neared shore, winds of this intensity are still powerful enough (about 57 m s^{-1}) to cause significant structural damage (Knabb et al., 2006). Katrina's extraordinary size contributed to devastating storm surge of exceeding 10 m peaks in several locations as Katrina made landfall in southeast Louisiana and finally along the Mississippi coast (Fritz wt al., 2007). In addition, 1464 lives from Louisiana were lost according to the report of Louisiana Department of Health and Hospitals[4], and total $81 billion property damage was estimated (Pielke et al., 2008).

In the following section, the WRF-EnKF system, the EnKF technique and the pre-processing of observations will be discussed. Section 3 provides a comparison between observations and hurricane forecasts initialized with the EnKF analyses. Independent in-situ and remotely sensed observations of the inner-core vortex structure are used to evaluate the performance of the EnKF analyses in section 4, and section 5 examines the analysis increments in comparison to a dual-Doppler analysis. Concluding remarks are given in section 6.

2. Data, methodology and experimental design

2.1 WRF and EnKF

Version 3.1 of the Advanced Research WRF (WRF-ARW) (Skamarock et al. 2005) was used in this study, which is an upgrade from the Version 2.2 used in Z09. Three two-way nested

[4] http://www.dhh.louisiana.gov/offices/page.asp?ID=192&Detail=5248

domains (Fig. 1a) are used for the analyses and forecasts. The coarsest domain (D1) has 202x181 horizontal grid points (at 40.5 km spacing); the second domain (D2) has 181x161 grid points (at 13.5 km spacing); and the innermost domain (D3) has 253x253 grid points (at 4.5 km spacing). After 24 h of model integration, the two inner domains are auto-movable centered on the storm's center using the WRF model's vortex-following algorithm. All model domains have 35 vertical layers with the domain top at 10 hPa. The physical parameterization schemes used for this study include the Grell–Devenyi cumulus scheme (Grell and Devenyi 2002, for D1 only), the WRF single-moment six-class microphysics with graupel scheme (Hong et al. 2004), the thermal diffusion scheme for land-surface, the Monin-Obukhov scheme for the surface layer and the Yonsei State University (YSU) scheme (Noh et al. 2003) for planetary boundary layer processes.

The initial ensemble is generated by the WRF data assimilation system (WRFDA) Version 3.1 using the "cv3" background error covariance option (Barker et al. 2004). The perturbation method for the ensemble initial and lateral boundary files is unchanged from Z09 except 60 ensemble members are used in this study (sensitivity experiments not presented here show that the 60-member EnKF improves both the analysis and forecast over using a 30-member EnKF as in Z09 while no apparent additional improvement is found in further doubling the size).

The EnKF implemented here follows Z09 except that the assimilated observations are NOAA P-3 aircraft airborne radar super-observations (SOs, details in the next subsection). Data assimilation is performed for all domains. The weighting coefficient, α, is set to 0.8 for the covariance relaxation following equation (5) of Zhang et al. (2004). The successive covariance localization (SCL) method proposed in Z09 is also used in this study to assimilate dense radar observations by using varying degrees of data thinning over different resolution domains. In this study, 1/9 of the total observations are randomly chosen and assimilated with a horizontal localization radius of influence (ROI) of 1215 km in all three domains in order to capture the large-scale background flow. Then another 2/9 of the observations are assimilated with a ROI of 405 km in the 13.5 and 4.5 km domains. Finally, the other 6/9 of observations are assimilated using a ROI of 135 km in the inner 4.5 km domain.

The SCL technique is designed to assimilate dense observations such as from radars that contain multi-scale information of the atmosphere (Z09), and to reduce computational costs and sampling errors. It uses a different ROI for different groups of observations by random sampling. The 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. 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). The decrease of error variance and length scale after more observations are assimilated is demonstrated in Zhang et al. (2006) for the mesoscale EnKF. The decrease of error length scale with more observations is also discussed in Daley (1991). The sensitivity experiments in Z09 (their Figure 15) showed the EnKF using the SCL method improved over using a fixed ROI for all observations.

2.2 Airborne radar observations and "superobbing"

Doppler radial velocity observations retrieved from the tail radar of a NOAA P-3 airplane are assimilated with the above WRF-EnKF system in this study. The airborne radar observations were collected from the NOAA N43RF flight mission[5] on 25 August 2005. The airplane took off at 1310 UTC from MacDill Air Force Base near Tampa Bay, FL and reached

[5] Information available at: http://www.aoml.noaa.gov/hrd/Storm_pages/katrina2005/20050825I1.html

the outer bands of Katrina by about 1400 UTC, where it started the first leg (a line crossing the storm center) of data collection. Then the plane made six legs of radar observations: 1401-1441, 1513-1553, 1558-1638, 1658-1738, 1827-1907 and 1940-2040 UTC and landed at 2115 UTC back at the MacDill Air Force Base in Tampa, Florida. Figure 1b shows the flight track and six legs along with the composite reflectivity from the Miami ground-based radar (KAMX) at 2000 UTC, which shows a clear eye just offshore of Fort Lauderdale.

The NOAA P-3 Doppler radar (Marks and Houze 1984, 1987; Marks et al. 1992) uses a forward and backward scanning technique to sweep out a 3-D radial velocity volume along the flight track. Antennas switch between the forward and backward directions when the beam reaches the top of the sweep cone, so one cycle of scanning includes two sweeps, and each beam has an angle of 70° from the airplane. The space along the flight track between two cycles is about 1.4 km, and the difference of time between two cycles is about 12 seconds. The primary data quality control, such as removal of noisy data, airplane velocity correction and unfolding, is made by the Hurricane Research Division (HRD) (Gamache et al. 2004)[6]. The data thinning and further data quality control applied here follows these steps:

1) Splitting the original file into one-sweep files. The original file provided by HRD includes all the observations of a leg so that the sweeps from different directions (forward and backward scans) are included in this file. Splitting the forward and backward sweeps is based on the earth-relative elevation from the horizon since the radar antenna switches between forward and backward sweeps at the top of the aircraft.
2) Creating a volume. The forward (backward) sweeps within 1 minute are then

[6] The quality controlled data with processing details is available online at the HRD's website: http://www.aoml.noaa.gov/hrd/Storm_pages/katrina2005/radar.html

combined into a single volume. For the NOAA P-3 radar, a forward (backward) volume usually includes five sweeps in each scanning direction. The length of one volume is about 5.6 km, and the observation time is about 1 minute.

3) Dividing the volume into smaller bins. The volume including the same direction sweeps is then divided into many partitions. Each partition has the same length as that in the volume and is composed of about five trapezoids. Each trapezoid is an interception from a sweep within 5 km in the radial direction and 5° in the azimuth direction. Figure 2 shows a sample of the trapezoid bins.

4) Observation selection. Any raw observations with values smaller than 2 m s^{-1} are inseparable from radar noises and will be discounted. Raw observations with values larger than 71 m s^{-1} will be also discounted since it is the maximum unambiguous radial velocity with the interlaced pulse repetition time (PRT) technique for the NOAA P-3 tail-mounted Doppler radar. We also remove all raw observations within 4 km of the radar to avoid large wind spread in a bin.

5) Quality control of raw observations. A raw radial velocity (V_r) observation will be discounted if its difference from the mean of all other raw observations within the bin exceeds twice of the standard deviations to reduce the spread of the bin.

6) Quality control of a bin. A bin will be discounted if the standard deviation of all available observations in the bin exceeds a value twice greater than the standard deviation of all observations in the volume.

7) Bin selection. There should be at least four valid raw Vr observations within an averaged bin.

8) Creating SOs. The final SO value of the bin will be the average of at most ten raw

observations, which are closest to the center of the bin. The position of the closest raw observation will mark the location of the SO.

9) Further thinning and randomly sorting. Depending on the resolution of the assimilating model, further thinning of the SOs may be necessary. Half of the SOs are chosen randomly in an arbitrary order for assimilation in the current study.

Following Dowell et al. (2004), the observation error for the SOs is assumed to be 3 m s^{-1} and the threshold value is set to 15 m s^{-1}. A SO is useable if the difference between the SO and the background is smaller than this threshold. The observation time for SOs in each leg will be rounded to the nearest 30 minutes, and their geographical positions will be adjusted according to the hurricane's motion in order to maintain the SOs' relative position around the hurricane center. Although the value of the observation error is rather empirically determined, our sensitivity EnKF experiments with values of 1.0 and 5.0, respectively, perform similarly to the control experiment presented herein in terms of both analysis and forecast (not shown). Nevertheless, it is acknowledged that most of the parameters and thresholds in the data thinning and quality control procedures for the airborne radar follow those used for ground-based radar data in Z09. More rigorous testing on the sensitivity of the EnKF performance to these empirical values will be needed in the future.

It is also worth noting that our desire to generate SOs is partially motivated by the limitations in bandwidth for transferring large volumes of raw radar observations from the aircrafts to the ground. Our SO generation technique is currently implemented on the NOAA aircrafts which allows for more efficient transmitting of the airborne Doppler radar to the ground in real-time.

2.3 Experiment design

The NCEP GFS operational analysis at 0000 UTC 25 August and its forecast were used to generate the initial and boundary conditions respectively. A 60-member ensemble initiated by adding perturbations derived from an application of the WRF 3DVar default background error statistics to the GFS analysis was integrated to 1430 UTC 25 August. Half of the SOs of each leg were chosen randomly and assimilated by the EnKF at 1430, 1530, 1630, 1700, 1900 and 2000 UTC over all domains with the SCL method. No other observation was assimilated in this study whose primary objective is to examine the impact of assimilating airborne radar observations. The ensemble mean before assimilation will be called “prior”, while the mean after assimilation will be called “posterior”. A deterministic WRF forecast initialized with the GFS analysis at 0000 UTC 25 August without assimilating observations will be called “NoDA”.

The forecasts are initialized at 2000 UTC 25 August with fixed model domains, but the two inner domains become movable and follow the vortex center starting at 0000 UTC 26 August. The NCEP GFS operational forecast initialized at 0000 UTC 25 August is used as the boundary conditions. A deterministic forecast (denoted as “EnKF_DF”) is initialized from the EnKF mean (EnKF updates the ensemble mean first, then updates the perturbations to each member), while an ensemble of forecasts (denoted as “EnKF_EF”) is initialized with the EnKF analysis perturbations. The NHC official forecast (OFCL) and model forecast (GFDL) available at 0000 UTC 26 is used as baselines in this study. The WRF forecast initialized with the GFS operational analysis at 0000 UTC 25 (NoDA) is also used to examine the impacts of hurricane forecasts with airborne radar data assimilation.

3. Verification of the forecasts initialized from the EnKF analyses

3.1 Forecasts from the EnKF analyses

Generally speaking, the ensemble mean represents the best estimate of the state and contains fewer errors than most individual members (Leith 1974; Murphy 1988). Howerer, deterministic forecasts are widely assumed to represent the best available approximation of the future state of the atmosphere. To make a comparison with OFCL and GFDL operational forecasts, we first evaluate the deterministic hurricane track and intensity forecasts of EnKF_DF. As shown in Fig. 3a, hurricane tracks from OFCL and GFDL operational forecasts are located to the right of the NHC best track. Hurricanes from these forecasts made landfall at locations close to Panama City, Florida, while EnKF_DF closely followed the NHC best track, despite a 6-h delay. Despite an initial deviation to the left, the landfall position error without considering the lag from EnKF_DF was only ~30 km. The forecasts without EnKF assimilation of observations, NoDA, are in the middle of the operational forecasts and the EnKF deterministic forecast. From this result, it is encouraging that the track forecast for EnKF_DF, especially the landfall location, shows significant improvements over the control WRF forecast without EnkF assimilation (NoDA).

For intensity forecasts in terms of minimal sea-level pressure ("P_{min}", Fig. 3b) and 10-m maximum wind speed ("V_{max}", Fig. 3c), both OFCL and GFDL are weaker than the NHC best track. At 2000 UTC 25 August (initial time of EnKF forecasts) the P_{min} and V_{max} for EnKF_DF were much closer to the NHC best track than NODA. After 1800 UTC 27 August, P_{min} for EnKF_DF was lower than that of NoDA and the NHC best track, but V_{max} of EnKF_DF intensified quickly and reached its peak intensity at 1800 UTC 28 August. Shortly after, the observed Katrina became weaker while the WRF simulations with and without data assimilation continued to develop before making landfall. The simulated storm movement trails the observations, a result similar to previous studies using a mesoscale (Davis et al. 2008) or global (Shen et al. 2006) model.

Another advantage of the EnKF over variational data assimilation methods is that it naturally provides an ensemble of initial conditions. Ensemble track forecasts from the EnKF perturbations are also shown in Fig. 3 along with the deterministic forecasts. The large ensemble forecast spread indicates the large amount of uncertainty associated with the prediction of Hurricane Katrina at this initialization time. Ensemble means for track and intensity (EnKF_EF-mean) show the same improvement as the deterministic forecast EnKF_DF. It is worth noting that despite overall improved forecast by assimilating radar observations, the track forecasts of both EnKF_DF and ensemble mean have a substantial southward error compared to the NHC best track. This southward bias may be caused by the initial adjustment in the model after all observations are assimilated, the impacts of sampling/model error, and the inhomogeneity due to interactions with land at the earlier forecast hours.

3.2 Verification by NHC best-track estimates

Figure 4 shows the root-mean square errors (RMSEs) of track and V_{max} forecasts for OFCL, GFDL, NoDA, EnKF_DF and EnKF_EF-mean, as well as the standard deviations for EnKF_EF. At the initial time of 0000 UTC 26 August, the errors of operational track and intensity forecasts by OFCL and the GFDL model are smaller than those of the WRF forecasts including NoDA, EnKF_DF and EnKF_EF. This is due to the fact that the OFCL directly used the observations and the GFDL model was initialized with a balanced bogus vortex designed to closely fit the observed position and intensity. The advantage of the bogus method faded 24 h after initialization, as track forecast errors for the WRF cases became smaller than those of the operational forecasts. Overall, the track errors for the deterministic forecast initialized with the EnKF assimilation (EnKF_DF) are smaller than the track errors for the case without assimilation of airborne radar data (NoDA), and significant track forecast improvements for EnKF_DF and EnKF_EF occurred during the period after the hurricane strengthened quickly and before the second landfall (error bars of 0-96 h in Fig. 4a).

Track errors increase uniformly over time in the OFCL and GFDL forecasts; however, the EnKF-initialized forecasts all show peaks in position error at the 24-h lead time (Fig. 4a). Comparison with Fig. 3a shows that this track error is the result of a leftward-shift in the initial TC trajectory following assimilation. It is therefore possible that the later track improvements shown in Fig. 4a are a result of serendipitous errors that take the vortex further from the center of the eastward-moving upper-level anticyclone over the first 24 h of the integration. A series of simulations involving the blending of the EnKF-initialized vortex with the NoDA environment -- all of which result in the initial leftward track shift -- show that assimilated structures in the outer bands (600-900 km) lead to a change in the steering flow later in the integration, possibly through the maintenance of a stronger anticyclone over northern Florida. Results from these simulations are not shown here since they extend beyond the scope of this study; however, they suggest that the reductions in final landfall position errors in the EnKF-initialized integrations result from an improved representation of TC/environment interactions rather than from fortuitous initial trajectory errors.

At the initial time, errors of the intensity forecast for EnKF_DF and EnKF_EF are much smaller than that of NoDA, which is because the EnKF assimilation provides a more realistic vortex (more in next section), while the initial vortex directly interpolated from the global model is too weak. As is also shown in Fig.4, the ensemble spread of EnKF_EF has the same magnitude as the error of the ensemble mean forecast (EnKF_EF-mean), implying large uncertainties in this forecast.

It is worth noting that there are apparent discrepancy between the maximum surface wind and the minimum sea-level pressure forecasts among all WRF forecasts shown here, especially near the peak intensity of observed Katrina. The simulated storms generally appear to be stronger with a lower P_{min} but weaker with a smaller V_{max}, which was also the case in Z09. The exact reason for this discrepancy is beyond the scope of the current study but it is likely a combination of model error, and/or the use of a 4.5 km grid spacing, which is still too coarse to resolve the eyewall gradients.

4. Verification with independent in-flight measurements

The large ensemble hurricane position and intensity spread for forecasts shown in Fig. 3 may come from different inner-core structures or larger-scale environmental disturbances such as midlatitude trough, environmental shear and anticyclone (McTaggart-Cowan et al. 2007). The current study focuses on the impacts of assimilating airborne radar data on the initialization of hurricane inner-core structure. In this section we will compare the inner-core structures between the WRF-ARW forecast without data assimilation (NoDA), the short-term ensemble mean forecast before the EnKF analysis (named as prior) and after the EnKF analysis (named as posterior), and then verify each experiment with independent observations from the aircraft in-flight measurements and the SFMR-retrieved winds (SFMR refers to Stepped Frequency Microwave Radiometer).

Figure 5 shows the comparison between the wind speeds derived from the lowest model level (~35 m above sea level) of NoDA, prior, posterior and the SFMR-retrieved surface wind speeds in a storm-relative coordinate[4]. The SFMR data were retrieved from the NOAA P-3 flight mission on August 25 and organized into six stages based on the legs of the flight (taken at the same time as the airborne Doppler measurements), and the model-derived wind speeds are calculated at the same corresponding time of six legs of SOs (the SFMR retrieved surface winds were time-tagged using a similar method that was used for the airborne Doppler velocity SOs). The wind speed of NoDA (gray dashed curves in Fig. 5) fails to match the SFMR-derived measurements (gray curves in Fig. 5) for all six legs, and the intensity of the hurricane vortex during this period being significantly weaker. The prior valid at 1430 UTC before assimilating the first leg of airborne radar observations (gray dash-dot curve in Fig. 5a) is similar to that of NoDA, both of which are much weaker than that estimated by the SFMR-derived measurement in the first leg. Note that the prior at this time is the mean of the 60-member ensemble forecast, which is different from the single deterministic forecast provided by NoDA (refer to section 2.3). The posterior from the EnKF analysis valid at 1430 UTC, after assimilation of the first leg of radar observations, reproduces the observed wind structure along the flight track across the inner core, despite still weaker in intensity (gray dash-dot vs. black solid curves in Fig. 5a).

The ensemble forecast mean initialized with the prior at 1530 UTC before assimilation of the second-leg radar observations preserves much of the observed inner-core structure (according to the SFMR estimates), but the errors in the hurricane intensity remain quite large (gray dash-dot curve in Fig. 5b). However, the priors at the later assimilation times apparently benefit from the improved posteriors during the previous assimilation cycles (gray dash-dot and black solid curves in Fig.5c, e, f) in terms of the vortex structure. Similar results to those shown in Fig. 5 can be found from the verification with flight-level wind speeds detected by the NOAA

[4] Similar performance can be seen from plots similar to Figs. 5-8 but are verified in an earth-relative coordinate (not shown).

P-3 airplane in-flight measurements (Fig. 6).

Verification of the flight-level temperature (Fig. 7) with in-situ measurements on the airplane is also encouraging. Even though the analyses (posterior) of temperature cannot completely reproduce the sharp rise of temperature inside the eyewall (except for maybe after the last assimilation cycle), the temperature structure is improved by continuous assimilation of inner-core radar observations with the EnKF. Unlike the improvements for the wind and temperature fields, the assimilation of radar radial velocity observations appears to have a much smaller impact on the moisture field (Fig.8). This is possibly due to weaker correlations between winds and moisture than between winds and winds or between winds and temperature. The moisture field has relatively more spectral power for both total and error energy in smaller scales (as seen in Fig.9), which are less predictable and more random resulting in smaller correlations with the Doppler winds (also refer to Zhang et al. 2006).

To further evaluate the performance of the EnKF analysis, the RMSEs of the model-derived winds, verified against the SFMR surface wind speed, flight-level wind component u, v, wind speed, temperature and relative humidity observations along each leg of the flight track, are calculated in the earth-relative coordinate and shown in Fig. 9a-f, respectively. For SFMR surface wind speed (Fig. 9a), after assimilation of the first-leg radar data (1430 UTC), the RMSE of the posterior is ~5 m s^{-1}, which is less than half of that of the prior (10.5 m s^{-1}) and also significantly smaller than that of NoDA (~8.8 m s^{-1}). With improved initial conditions in each member and after 1-h short-term ensemble forecast, the RMSE of the prior (~9.2 m s^{-1}) becomes smaller than that of NoDA (~10 m s^{-1}) before assimilation of the second-leg radar data (valid at 1530 UTC), the relative error magnitude reverses from that at 1430 UTC which persists throughout the subsequent analysis and forecasts cycles. In comparing the RMSEs between the prior and posterior valid at the same times, it is evident that RMSEs of posterior are always smaller than those of prior at the same assimilation cycle, which means the EnKF assimilation decreases the error of analysis fields and acts to constrain errors (e.g., the posterior errors are lower than the prior). Ensemble standard deviations (STD) for the prior and posterior (Fig. 9a) are of the same order of magnitude; also the STD for posterior is almost equal in value to the RMSE for posterior during each cycle, which means the ensemble spread provides a good quantitative estimate of the forecast error. This result further supports the use of 0.8 as the relaxation (mixing) coefficient in this study (the same value as in Z09) for convection-permitting hurricane initialization by assimilation of high-resolution inner-core observations from Doppler radars.

The verification with flight-level wind speeds detected by the NOAA P-3 airplane in-flight measurements (Fig. 9b-d) is similar to the verification of SFMR surface wind speed (Fig. 9a). Partitioning the RMSEs of the wind speeds into zonal and meridional components further suggests that the EnKF not only improves the wind speed analysis and forecast but also the wind directions throughout the hurricane inner-core vortex. For the verification of flight-level temperature (Fig. 9e), reduction of the RMSEs at each assimilation time further demonstrates the EnKF's ability to estimate (and benefits from using) flow-dependent error covariance (between dynamic and thermodynamic variables) calculated from short-term ensemble forecasts in the hurricane inner-core region. Examinations of the dynamics, evolution and structure of such flow-dependent error covariance will be presented in detail in a separate study.

5. Analysis increments before and after the EnKF assimilation

To further evaluate the performance of the EnKF analyses, in this section the structure and magnitude of analysis increments for different variables sampled in two-dimensional (both horizontal and vertical) planes will be examined. Fig. 10a-c compares the forecast and analysis sea-level pressure (SLP), the lowest model level wind speed, and the storm center positions from experiment NoDA, and the prior and posterior at 1430 UTC 25 August 2005. The SLP of the prior has a similar structure and magnitude as NoDA with the exception of being smoother due to ensemble averaging. The P_{min}s of NoDA and the prior are both close to 1007 hPa, which is ~15 hPa higher than that of the NHC best track (which is performed only four times a day, so values are linearly interpolated to for times absent from the data). The lowest level winds of NoDA and the prior also show a weaker vortex with V_{max}s of about 19 and 13 m s^{-1}, respectively. The EnKF assimilation of the SOs increases the storm intensity with the posterior V_{max} reaching ~26 m s^{-1} in the northwestern quadrant (which is also much closer to the NHC best track V_{max} of ~33 m s^{-1}).

The storm's center for NoDA (blue dot in Fig. 10a) and the prior (green dot in Fig. 10b) are located at around (78.52W, 25.81N), and (78.54W, 26.19N), respectively, and are about 85 and 70 km away from the best track position of (79.25W, 26.20N) at this time (black dots in Fig. 10). After the assimilation of the first leg of radar data, the P_{min} for the EnKF posterior is about 1004 hPa, which is still higher than that of the best track. Despite this deficiency, the EnKF analysis reduces the P_{min} error by about 3 hPa. The storm's center for the EnKF posterior is located at about (79.34W, 26.21N), which is only about 9 km away from that of the best track, resulting in a position error reduction of nearly 90%. The storm's track spread for the prior ensemble forecasts (green circles in Fig. 10b) is very large, and the STD is ~53 km. After the assimilation of the first leg of radar SOs, the location of all members is incremented closer to the analyses and the STD of the posterior track is only ~3 km.

Figure 10d-i shows the increments (differences between the EnKF posterior and prior) of SLP and the lowest model level winds at the six assimilation times. The advantages of continuously assimilating observations are clearly demonstrated by the evolution of the P_{min} and the V_{max}, both of which approach the best track observations after two or three assimilation cycles (Fig. 11). However, given the lack of dense surface observations over the ocean, it is difficult to judge whether these increments are optimal. The structure and amplitude of these increments do provide further evidence of the need for flow-dependent error covariance both in time and space. Consistent with Zhang et al. (2006), there is also some evidence that the increments and error corrections in earlier assimilation cycles (over a broader area; Fig. 10d, e, g) have more larger-scale components than those at the later analysis times (more concentrated near the storm center; Fig. 10f, h, i). Another related reason might be the ensemble perturbations at earlier assimilation times retain a memory of the cv3 perturbations that were used to generate the initial ensemble which are large-scale in nature.

Figure 12 shows the horizontal winds at the 3-km height from the dual-Doppler radar wind analyses retrieved from the NOAA P-3 airborne radar observations (first column), compared with the EnKF prior (second column), posterior (third column) and their difference (analysis increment, last column) at the six assimilation times (from the first row to the last row). During this period, the NoDA forecast presents a very weak vortex (not shown). In the first cycle (1430 UTC 25 August 2005, the first row in Fig. 12), the winds for the prior (Fig. 12b) also present a very weak vortex, and the maximum wind speed is less than 7 m s^{-1}. After the first assimilation cycle, the EnKF posterior (Fig. 12c) shows a stronger cyclonic vortex at 3 km with a maximum of ~23 m s^{-1} in both the northwest and the southeast quadrants. Wind speeds in the center are weaker than 7 m s^{-1} and the center of the circulation is very close to what is shown in the dual-Doppler radar analyses (Fig. 12a). Although still weaker than the dual-Doppler analysis, the EnKF posterior wind structure and amplitude are much closer to the observations than the prior. The EnKF analysis increment (Fig. 12d) shows that the EnKF assimilation increases the wind speed around the observed vortex center and the maximum increment is up to 15 m s^{-1}. The cyclonic vortex of the prior (second column in Fig. 12) becomes increasingly close to the observed vortex center and size (first column) after two assimilation cycles in both center position and size, but still contains some biases in structure and intensity. This is shown at 1730 (fourth row in Fig. 12), 1900 (fifth row) and 2000 UTC (last row), where the areas of maximum wind speed are displaced from locations indicated by the dual-Doppler analyses. Starting from the second assimilation cycle (second row), the posterior (third column) improves the vortex intensity and the distribution of maximum wind speed. The analysis increment (last column) shows that at the first assimilation cycle, the positive increment spans most of the display domain (Fig. 12d). Starting from the third cycle, the area of positive increment becomes much smaller with large increments only around the vortex center. Nevertheless, the EnKF assimilation continuously increases the vortex intensity and improves the depiction of the vortex structure.

Vertical cross sections of horizontal wind speeds along the flight tracks of each leg marked in the first column of Fig. 12 are presented in Fig. 13. A clear storm eye and a strong wind band around the eye are shown in the wind speed field derived by HRD from the dual Doppler radar analysis (Fig. 13a) for all six legs. Similar to the result of the 3-km horizontal wind analysis, the NoDA forecast does not simulate the storm structures for any of the six legs (not shown). Neither does the EnKF prior at 1430 UTC before assimilating the first-leg radar data show the storm structure (Fig. 13b). After the first-leg airborne radar observations are assimilated, the storm center position and the storm structure (Fig. 13c) are considerably closer to the dual-Doppler analyses (Fig. 13a); however, the storm intensity of posterior (Fig. 13c) is still weaker than that estimated by the dual-Doppler analysis (Fig. 13a). At the second assimilation time, the EnKF prior (Fig. 13f) shows a very weak vortex under 3 km in the same location as the analyses, while the posterior (Fig. 13g) strengthens the storm and builds the storm vortex up to 6 km. From the third assimilation cycle onward, the prior (Figs.13j, n, r and v) catches the storm's location and structure well, and the EnKF posterior (Figs. 13k, o, s and w) continues to converge towards the dual-Doppler analyses (Figs. 13i, m, q and u).

6. Summary and conclusions

Through a WRF-based EnKF data assimilation system, this study examines the impacts of assimilating airborne radar observations for the convection-permitting analysis and prediction of Hurricane Katrina (2005). In this study, a series of data thinning and quality control procedures to generate super-observations (SOs) of airborne radial velocity measurements are developed and implemented. These procedures allow sub-sampling of large volumes of radar measurements to higher-quality SOs with resolutions more compatible with the assimilating NWP models and reduce sampling errors as well as computational costs. Moreover, the procedures developed here have since been implemented operationally on the NOAA hurricane reconnaissance aircrafts to enable real-time transmission of airborne radar observations to the ground in the form of SOs.

Verifications against independent radar, in-situ and remotely-sensed measurements show that the EnKF analyses with assimilation of airborne radar observations successfully depict the inner-core structure of the hurricane vortex at the analysis time in terms of both dynamic (wind) and thermodynamic (temperature) fields. Despite an initial trajectory error, the forecasts initialized with the EnKF analyses of airborne radar observations improved the hurricane track forecasts at lead times beyond 12 h compared to NOAA's operational forecasts and the forecast without assimilating airborne radar observations. The error of the landfall forecast was about 30 km with a 6 h delay. Besides the improved deterministic analysis and forecasts, the ensemble forecasts initiated from the EnKF analyses also indicate large forecast uncertainty at this lead-time for the hurricane track and intensity forecasts.

Although the EnKF system successfully depicts the inner-core structure and makes improvements to the hurricane forecast, there are still many issues that are unclear. For example, how many ensemble members are suitable and acceptable for convection-permitting hurricane assimilation, especially for a real-time analysis and forecast system? What is the correlation between radial velocity and thermodynamic fields? A recent study of Poterjoy and Zhang (2011) examined the sample covariance structures estimated from ensemble forecasts initiated with the EnKF perturbations presented herein. Follow-up studies will continue to examine the sensitivity of the EnKF analysis to the ensemble size, the correlations between the assimilated variables and the continually updated fields, and the filter configurations (relaxation coefficient, and radius of influence). Moreover, since the airborne radar observations only cover the inner-core area, it remains unclear how much influence the radar observations have on improving the depiction of the storm's environment that subsequently improves predictability (Sippel and Zhang 2008, 2010; Zhang and Sippel 2009). Future studies will also assimilate more types of observations, such as those from conventional networks, dropsondes, flight-level in-situ measurements as well as other remotely-sensed observations, such as from SFMRs and satellites.

It is acknowledged that the above findings and conclusions are results from a single case in hindcast. More systematic experiments with this analysis and prediction system with many more cases are needed to assess the impacts of high-resolution inner-core observations from the airborne radars (as well as from in-situ or other remotely-sensed observations such as by satellite). Nevertheless, this same ensemble assimilation and forecast system was successfully implemented at the Texas Advanced Computing Center (TACC) high-performance computing facility, and has been used in real-time to ingest and assimilate airborne Doppler radar observations from the NOAA P-3 aircrafts since 2008. A summary performance of these real-time and post-event experiments over all applicable airborne Doppler missions during 2008-2010 is presented in Zhang et al. (2011). Both the historical hindcast of Katrina (2005) presented here and the ongoing real-time experiments with this convection-permitting ensemble analysis and forecast system provide a potential pathway for future improvements in hurricane prediction, especially focused on the intensity guidance (Zhang et al. 2011).

Acknowledgments: We benefited greatly from discussions with Frank Marks, John Gamache, Yongsheng Chen, Chris Snyder, Jon Poterjoy, Sim Aberson, Altug Aksoy, Chanh Kieu, Jeff Whitaker and Bob Gall. Editing and proofreading by Jon Poterjoy and Erin Munsell are greatly appreciated. We also benefited greatly from review comments by two anonymous reviewers and Ron McTaggart-Cowan on an earlier version of the manuscript. This work was supported in part by the NOAA Hurricane Forecast Improvement Project (HFIP), Office of Naval Research Grants N000140410471 and N000140910526, and the National Science Foundation Grant ATM-0840651. The computing for this study was performed at the Texas Advanced Computing Center.

6. References

Barker, D. M., 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.

Bishop, C. H. and D. Hodyess, 2007: Flow-adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. *Quarterly Journal of Royal Meteorological Society,* **133**, 2029-2044.

Daley, R., 1991: *Atmospheric data assimilation*. Cambridge University Press. 457pp.

Davis, C. A., Sarah C. J. and Michael R., 2008: Hurricane Vortex Dynamics during Atlantic Extratropical Transition. *J. Atm. Sci.*, **65**, 714-736.

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.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10143-10162.

Fang, J and F. Zhang, 2010: Initial development and genesis of Hurricane Dolly (2008).*Journal of the Atmospheric Sciences*, **63**, 655-672.

Fang, J and F. Zhang, 2011: Evolution of Multi-scale Vortices in the Development of Hurricane Dolly (2008). *Journal of the Atmospheric Sciences*, **68**, 103-122.

Fritz, H. M., and Coauthors, 2007: Hurricane Katrina storm surge distribution and field observations on the Mississippi barrier islands. *Estuarine Coastal Shelf Sci.,* **74**,12–20.

Gamache, J. F., 2005: Final report on JHT project entitled: Real-Time Dissemination of Hurricane Wind Fields Determined from Airborne Doppler Radar.

http://www.nhc.noaa.gov/jht/2003-2005reports/DOPLRgamache_JHTfinalreport.pdf

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

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.

Knabb, R. D.; J. R. Rhome, and D. P. Brown, 2006: Tropical Cyclone Report: Hurricane Katrina: 23–30 August 2005. National Hurricane Center, Miami, Florida. Online resource is available at http://www.nhc.noaa.gov/pdf/TCR-AL122005_Katrina.pdf.

Kurihara, M. A. Bender M. A. and R. J. Ross, 1993: An initialization scheme of hurricane models by vortex specification. *Mon. Wea. Rev.,* **121**, 2030-2045.

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.

Leith, C. E., 1974: Theoretical Skill of Monte Carlo Forecasts. *Mon. Wea. Rev.*, **102**, 409-418.

Marks, F. D., and R. A. Houze, 1984: Airborne Doppler radar observations in Hurricane Debby. *Bull. Amer. Meteor. Soc.*, **65**, 569-582.

Marks, F. D., and R. A. Houze, 1987: Inner core structure of Hurricane Alicia from Doppler radar observations. *J. Atmos. Sci.*, **44**, 1296—1317.

Marks, F. D., R. A. Houze, and J. F. Gamache, 1992: Dual-aircraft investigation of the inner core of Hurricane Norbert. Part 1: Kinematic structure. *J. Atmos. Sci.,* **49**, 919-942.

McTaggart-Cowan, R., L. F. Bosart, J. R. Gyakum, and E. H. Atallah, 2007: Hurricane Katrina (2005). Part II: Evolution and hemispheric impacts of a diabatically generated warm pool. *Mon. Wea. Rev.*, **135,** 3927–3949.

Murphy, J. M., 1988: The impact of ensemble forecasts on predictability. *Quart. J. Roy. Meteor. Soc.*,**114**, 463–493.

Nguyen, S.V., Smith, R.K., Montgomery, M.T. , 2008: Tropical-cyclone intensification and predictability in three dimensions. *Q. J. R. Meteorol. Soc.* **134**, 563 – 582.

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.

Pielke, R. A., Jr., J. Gratz, C.W. Landsea, D. Collins, M. A. Saunders, and R. Musulin, 2008: Normalized Hurricane Damages in the United States. *Natural Hazards Review*, **9**, 29-42.

Poterjoy, J. and F. Zhang, 2011: Dynamics and structures of error covariance in the core of a developing hurricane. *Journal of the Atmospheric Sciences*, in press.

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.

Shen, B., and coauthors, 2006: Hurricane Forecasts with a Global Mesoscale-Resolving Model: Preliminary Results with Hurricane Katrina (2005). *Geophys. Res. Lett.*, **33**, L13813, doi:10.1029/2006GL026143.

Sippel, J. A. and F. Zhang, 2008: A probabilistic analysis of the dynamics and predictability of tropical cyclogenesis. *Journal of the Atmospheric Sciences*, 65, 3440-3459.

Sippel, J. A and F. Zhang, 2010: Factors affecting the predictability of hurricane Humberto (2007). *Journal of the Atmospheric Sciences*, 63, 1759-1778.

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 Technical Note*, NCAR/TN-468+STR.

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

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.

Ueno, M., 1995: A study on the impact of asymmetric components around tropical cyclone center on the accuracy of bogus data and the track forecast. *Meteorol. Atmos. Phys.,* **56**, 125-134.

Weng, Y., Zhang, M. and Zhang, F., 2011: Advanced data assimilation for cloud-resolving hurricane initialization and prediction. *Computing in Science and Engineering,* **13**, 40-49.

Zou, X. and Xiao, Q., 2000: Studies on the Initialization and Simulation of a Mature Hurricane Using a Variational Bogus Data Assimilation Scheme. *J. Atmos. Sci.,* **57**, 836-860.

Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238-1253.

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

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

Zhang, F., and C. Snyder, 2007: Ensemble-based data assimilation. Bull. Amer. Meteor. Soc, 88, 565-568.

Zhang, F., Weng, Y., Sippel, J. A., Meng, Z., and Bishop, C. H., 2009: Convection-permitting Hurricane Initialization and Prediction through Assimilation of Doppler Radar Observations with an Ensemble Kalman Filter: Humberto (2007). Mon. Wea. Rev., 137, 2105-2125.

Zhang, F., Y. Weng, J. F. Gamache, and F. D. Marks, 2011: Performance of Cloud-resolving Hurricane Initialization and Prediction during 2008-2010 with Ensemble Data Assimilation of Inner-core Airborne Doppler Radar Observations. *Geophys. Res. Lett.,* 38, doi:10.1029/2011GL048469.

LIST OF FIGURE CAPTIONS

Figure 1. (a) Model domain configurations. The two inner domains are movable by following the vortex center 24h later at 0000 UTC 26 August 2005. Also plotted is the NHC Best Track of Katrina. (b) The Miami ground-base radar (KAMX) composite reflectivity (dBZ) at 2000 UTC 25 Aug 2005 (gray shaded) and the flight track for the P3 mission (gray thin solid). The six flight legs are numbered and marked with dark thick curves.

Figure 2. Geometry and schematics of an example SO bin selected for airborne Doppler radar quality control and data thinning.

Figure 3. Hurricane track (a), minimum sea level pressure (b) and maximum 10m wind speed (c) forecasts for The Best Track (black), OFCL (cyan), GFDL (sky blue), NoDA (blue), EnKF_DF (red), EnKF_EF (green) and EnKF_EF-mean (magenta).

Figure 4. Root-mean square errors (RMSEs) and ensemble standard deviations (STDs) of hurricane position (a) and intensity (b) forecasts. The RMSEs are for OFCL, GFDL, NoDA, EnKF_DF and EnKF_EF-mean, while the STD is the ensemble standard deviation of EnKF_EF.

Figure 5. Surface wind speed retrieved by SFMR (gray solid) and the lowest model level wind speed derived from NoDA (gray dashed), prior (gray dash-dot) and posterior (solid) in storm-relative space for the six legs (a-f) shown in Fig. 1b. The x-axis is the horizontal distance (km) between the flight and the storm center of the NHC operational analyses; the y-axis is the wind speed (m s^{-1}).

Figure 6. Same as that of Figure 5, but for flight-level wind speed ($m \cdot s^{-1}$).

Figure 7. Same as that of Figure 5, but for flight-level temperature (°C).

Figure 8. Same as that of Figure 5, but for flight-level relative humidity (%).

Figure 9. RMSEs of NoDA (blue), prior (green) and posterior (red) and ensemble STDs of prior (magenta next to green) and Posterior (magenta next to red) for (a) SFMR surface wind speed, (b) flight level u, (c) flight level v, (d) flight level wind speed ($m \cdot s^{-1}$), (e) flight level temperature (°C) and (f) flight level relative humidity (%), respectively.

Figure 10. Sea level pressure (SLP) (shaded), the lowest level winds (arrows) and wind speed (contours started at 10 m s^{-1} every 3 m s^{-1}) for (a) NoDA, (b) prior and (c) posterior at 1430UTC 25 Aug 2005; and increments of SLP (shaded), lowest model level winds (arrows) and 3 m s^{-1} wind speed (red contour) between posterior and prior after assimilating the airborne radar SOs of 1-6 legs at (d) 1430, (e) 1530, (f) 1630, (g) 1730, (h) 1900 and (i) 2000 UTC, respectively. Black dot is the storm center position linearly interpolated from the NHC Best Track; blue dot is the storm center of NoDA; green dot is the storm center position calculated from the ensemble mean prior; red dot is that of posterior that assimilated the first P3 leg radar radial velocity; green circles are the centers of the ensemble members of prior; and red circles (overlapped by red dots) are the centers of the ensemble members of posterior.

Figure 11. Evolution of (a) minimum sea level pressure (hPa) and (b) maximum surface wind speed (m s^{-1}) for OBS (gray solid) that is linearly interpolated from the NHC Best Track, NoDA (dash) and EnKF (solid).

Figure 12. Winds and wind speed at 3 km height. The columns from left to right are the dual-Doppler radar analyses, prior, posterior and EnKF increments (posterior minus prior), respectively. The rows from the top to the bottom are for 1430, 1530 1630, 1730, 1900 and 2000 UTC, respectively. The gray lines in the 1st column are tracks of the NOAA P-3 flight. The circled shaded areas in the last column indicate the increments are less than 1 m s^{-1}.

Figure 13. Track-height wind speed sections along the leg tracks shown in the first column of Fig. 12. Other information is the same as that in Fig. 12.

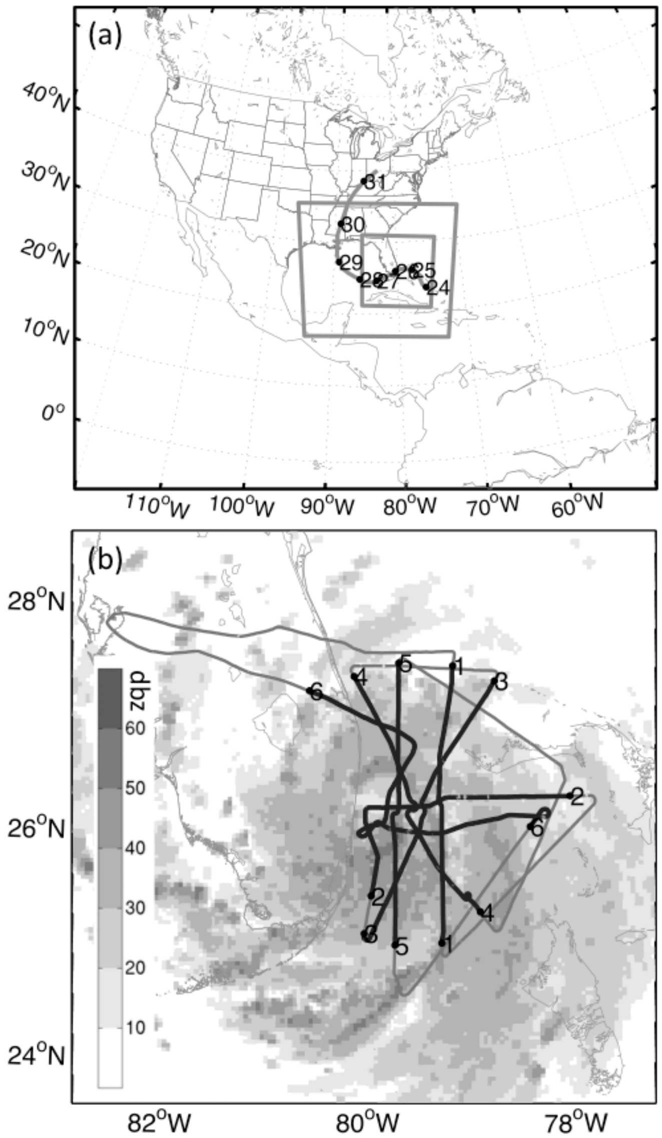

Figure 1. (a) Model domain configurations. The two inner domains are movable by following the vortex center 24h later at 0000 UTC 26 August 2005. Also plotted is the NHC Best Track of Katrina. (b) The Miami ground-base radar (KAMX) composite reflectivity (dBZ) at 2000 UTC 25 Aug 2005 (gray shaded) and the flight track for the P3 mission (gray thin solid). The six flight legs are numbered and marked with dark thick curves.

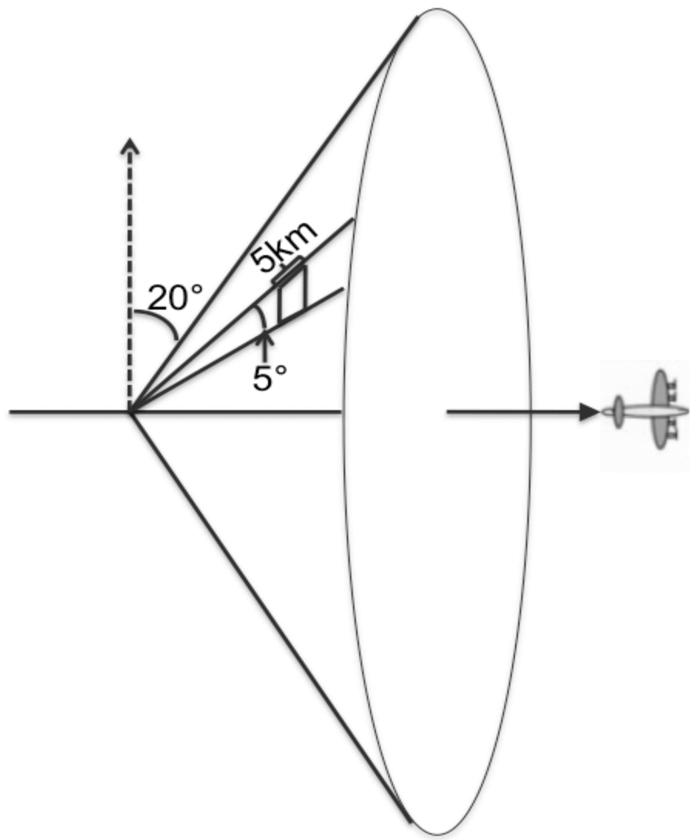

Figure 2. Geometry and schematics of an example SO bin selected for airborne Doppler radar quality control and data thinning.

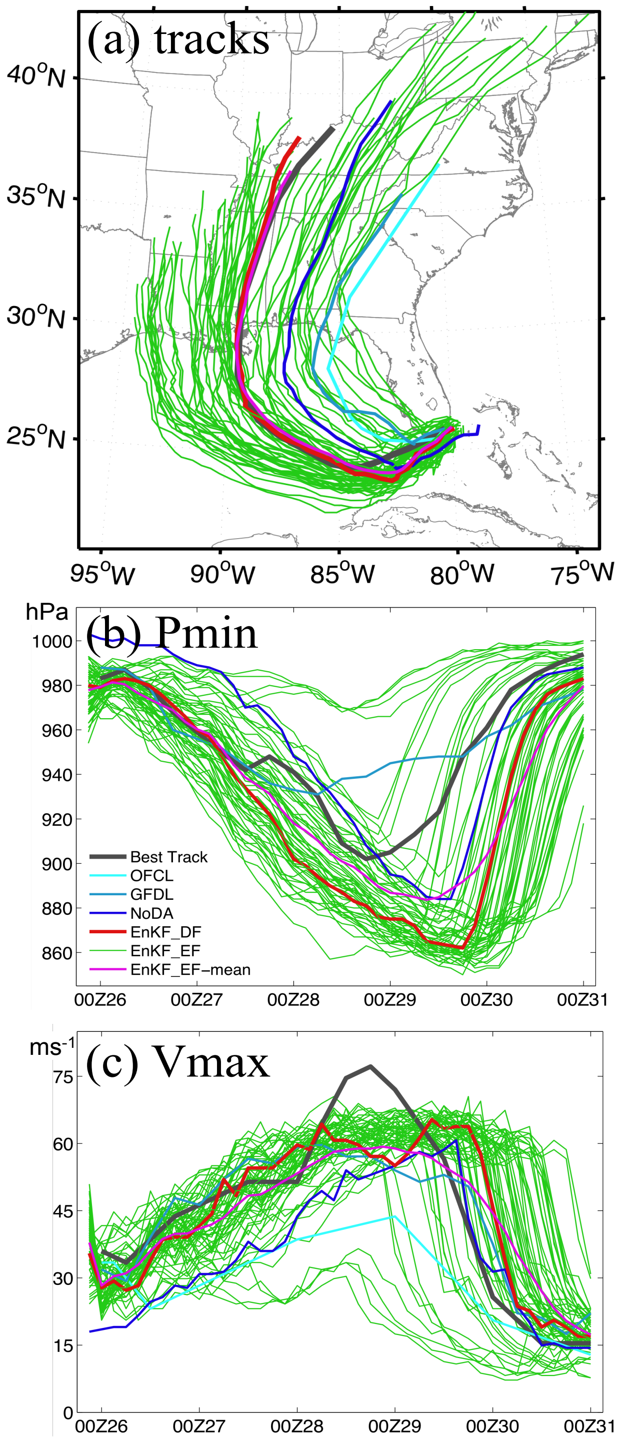

Figure 3. Hurricane track (a), minimum sea level pressure (b) and maximum 10m wind speed (c) forecasts for The Best Track (black), OFCL (cyan), GFDL (sky blue), NoDA (blue), EnKF_DF (red), EnKF_EF (green) and EnKF_EF-mean (magenta).

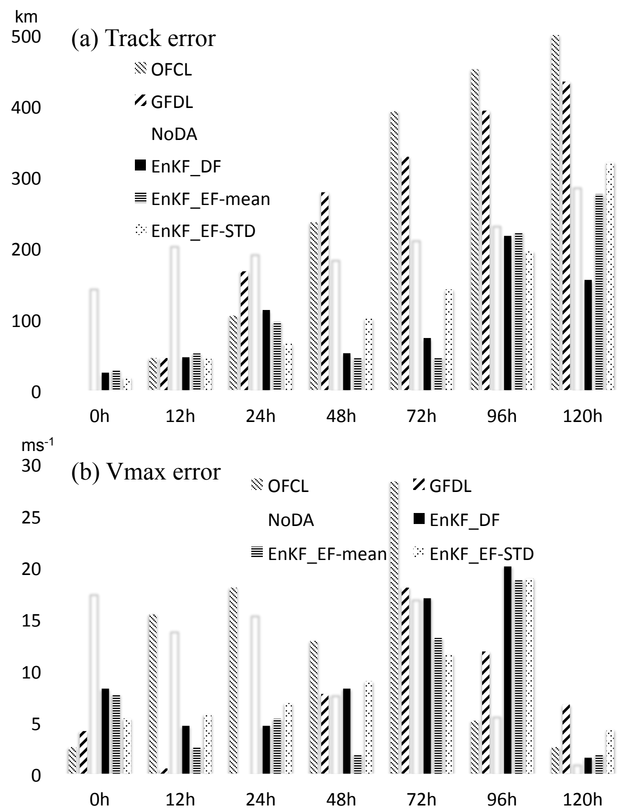

Figure 4. Root-mean square errors (RMSEs) and ensemble standard deviations (STDs) of hurricane position (a) and intensity (b) forecasts. The RMSEs are for OFCL, GFDL, NoDA, EnKF_DF and EnKF_EF-mean, while the STD is the ensemble standard deviation of EnKF_EF.

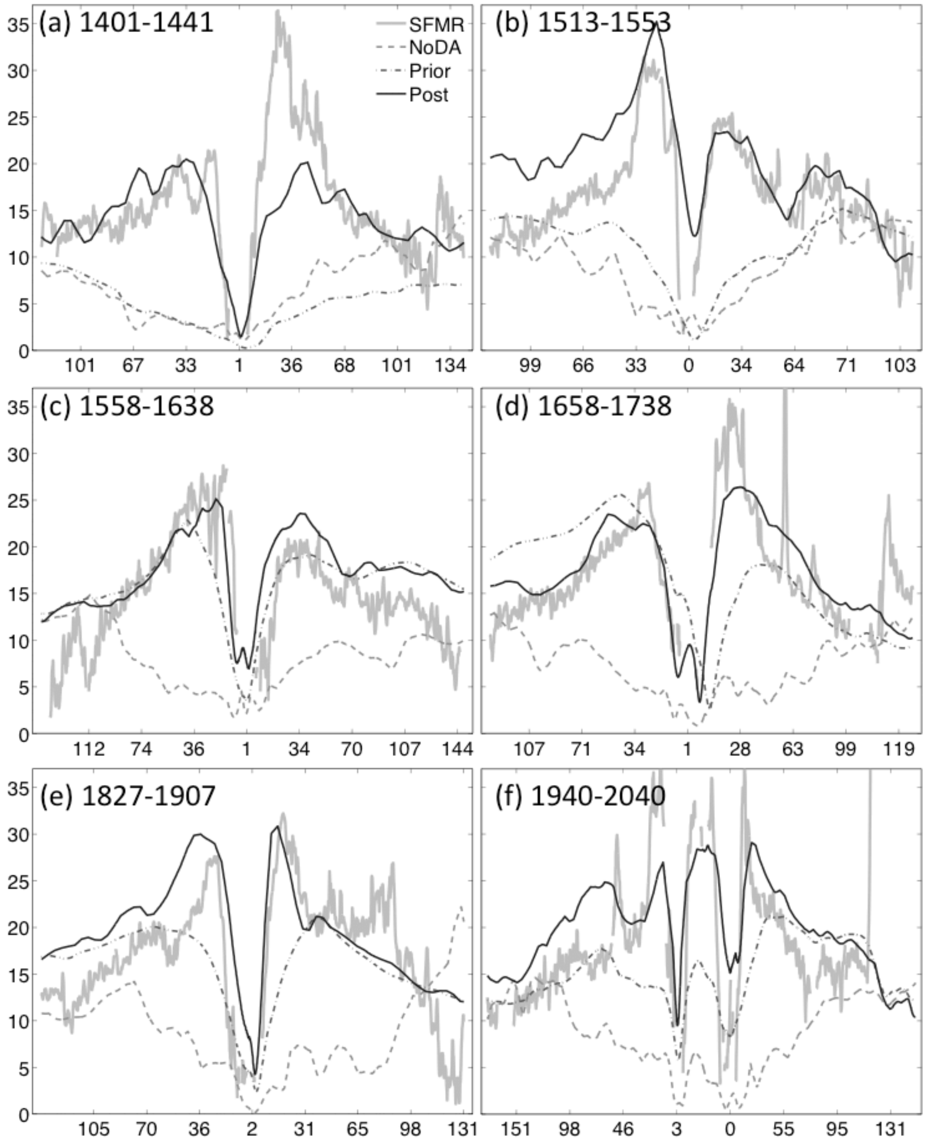

Figure 5. Surface wind speed retrieved by SFMR (gray solid) and the lowest model level wind speed derived from NoDA (gray dashed), prior (gray dash-dot) and posterior (solid) in storm-relative space for the six legs (a-f) shown in Fig. 1b. The x-axis is the horizontal distance (km) between the flight and the storm center of the NHC operational analyses; the y-axis is the wind speed (m s^{-1}).

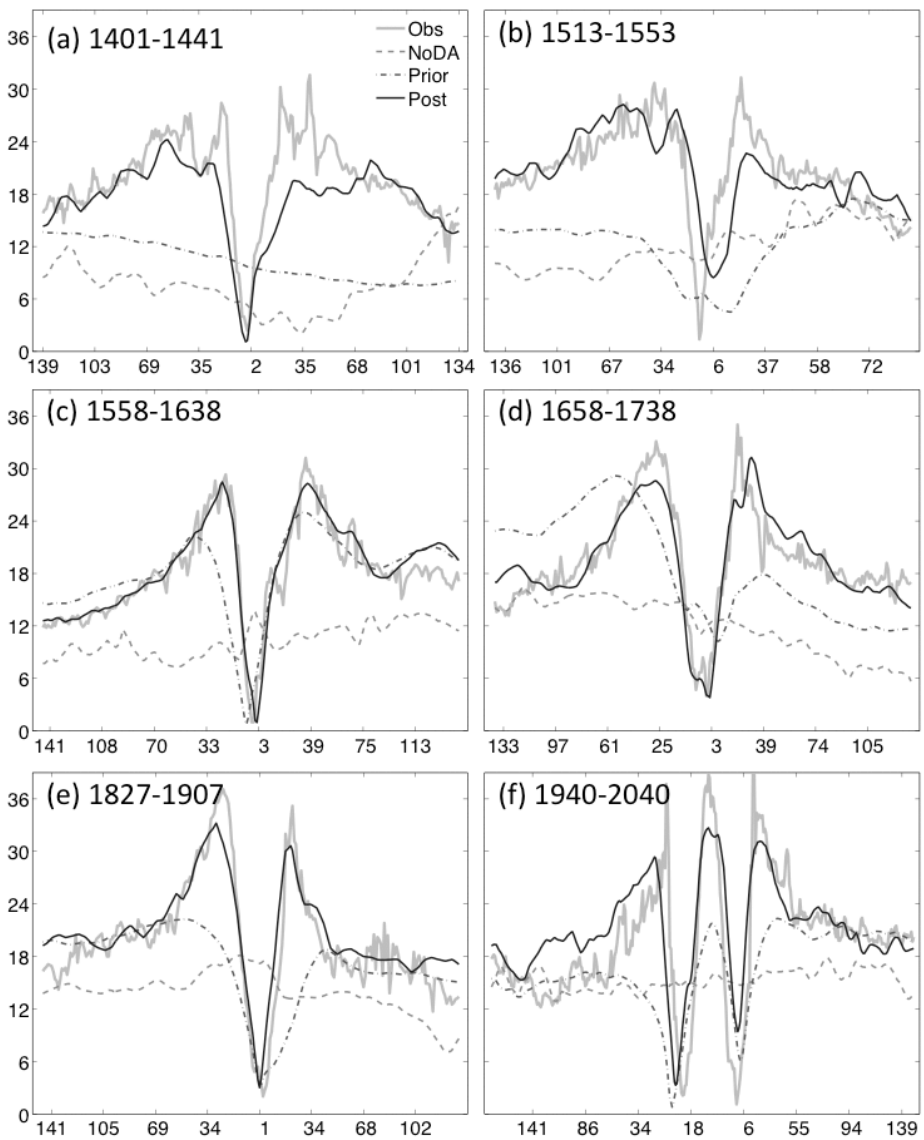

Figure 6. Same as that of Figure 5, but for flight-level wind speed ($m \cdot s^{-1}$).

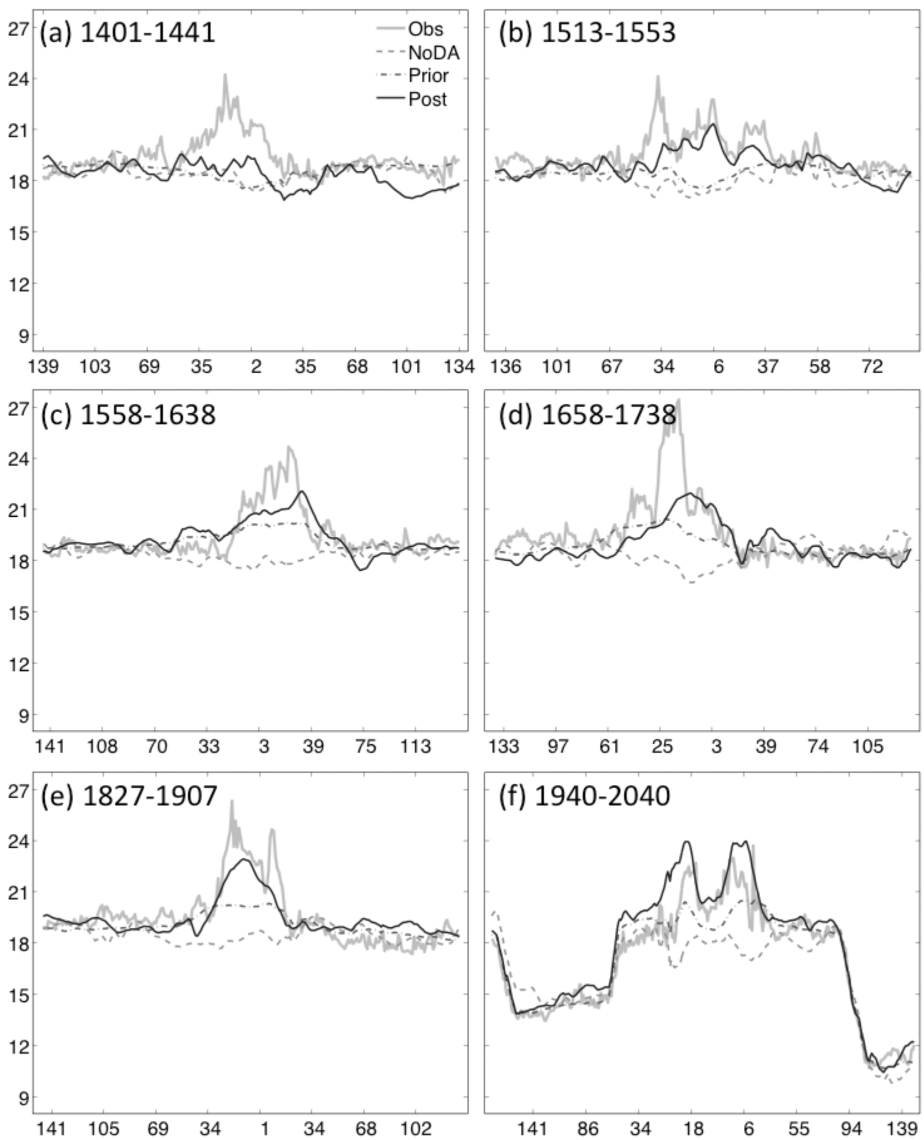

Figure 7. Same as that of Figure 5, but for flight-level temperature (°C).

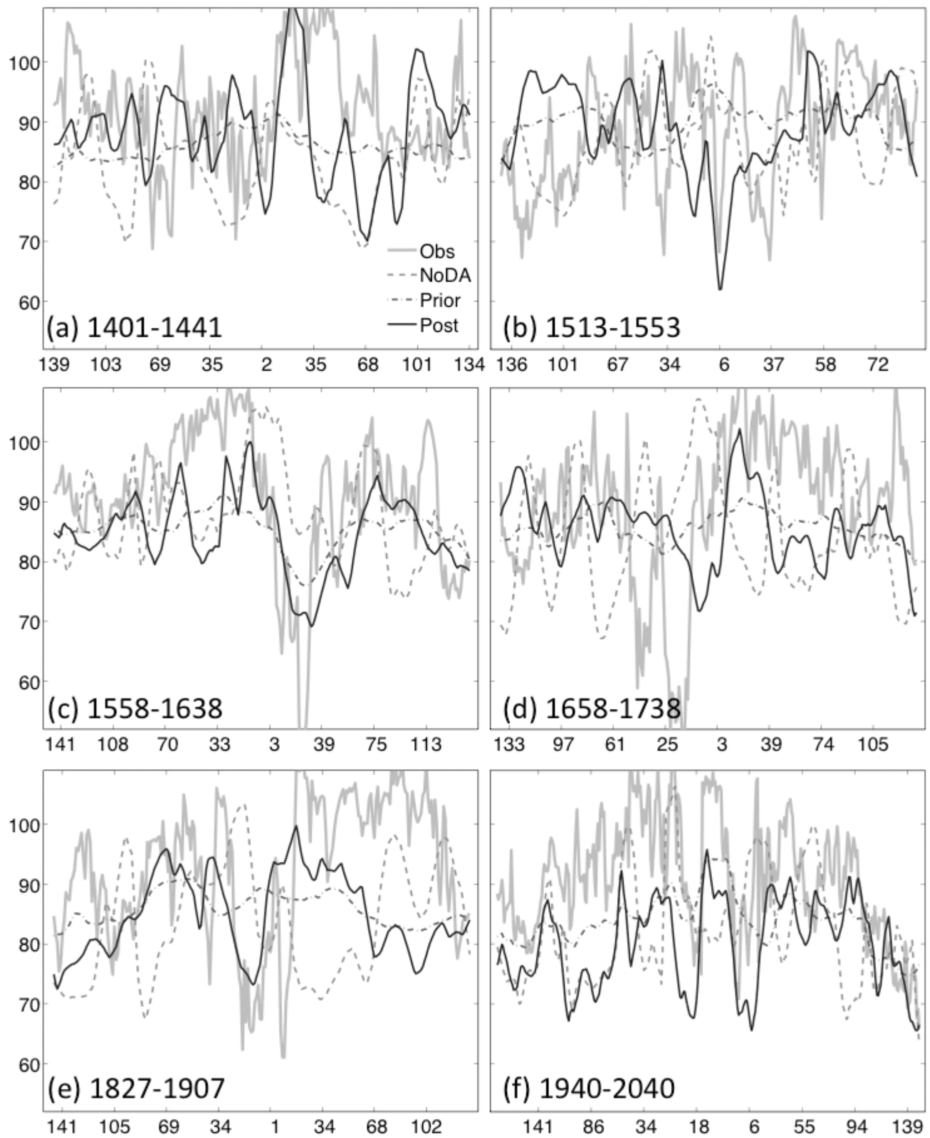

Figure 8. Same as that of Figure 5, but for flight-level relative humidity (%).

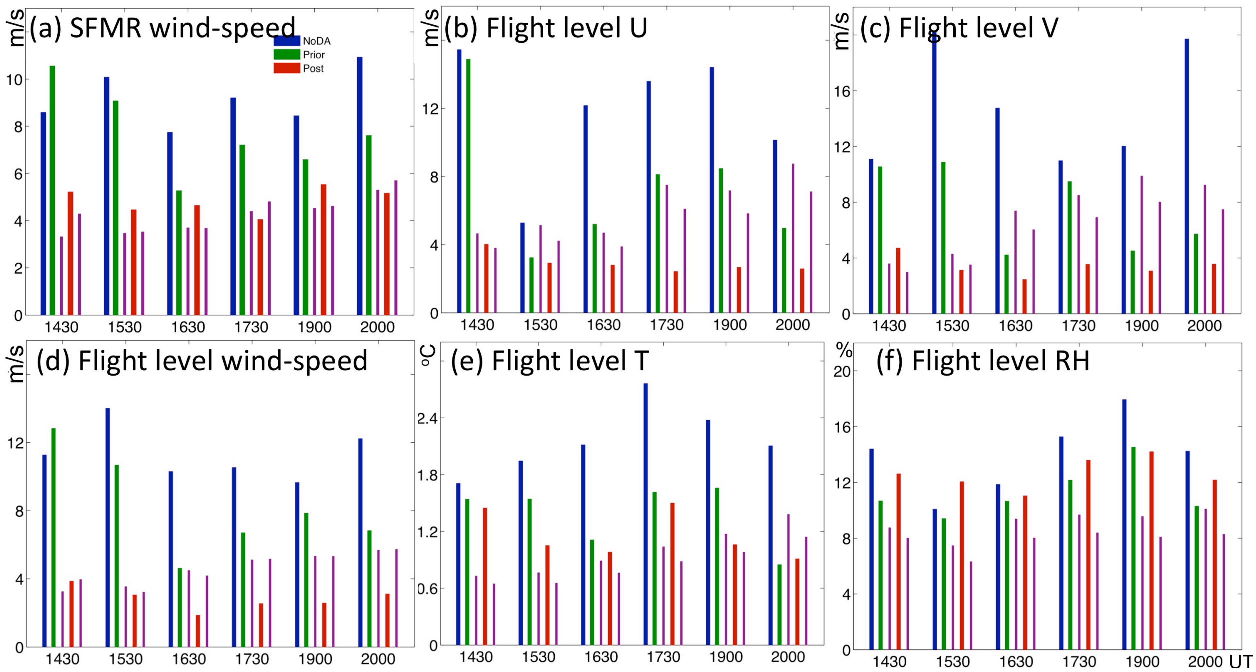

Figure 9. RMSEs of NoDA (blue), prior (green) and posterior (red) and ensemble STDs of prior (magenta next to green) and Posterior (magenta next to red) for (a) SFMR surface wind speed, (b) flight level u, (c) flight level v, (d) flight level wind speed (m·s^{-1}), (e) flight level temperature (°C) and (f) flight level relative humidity (%), respectively.

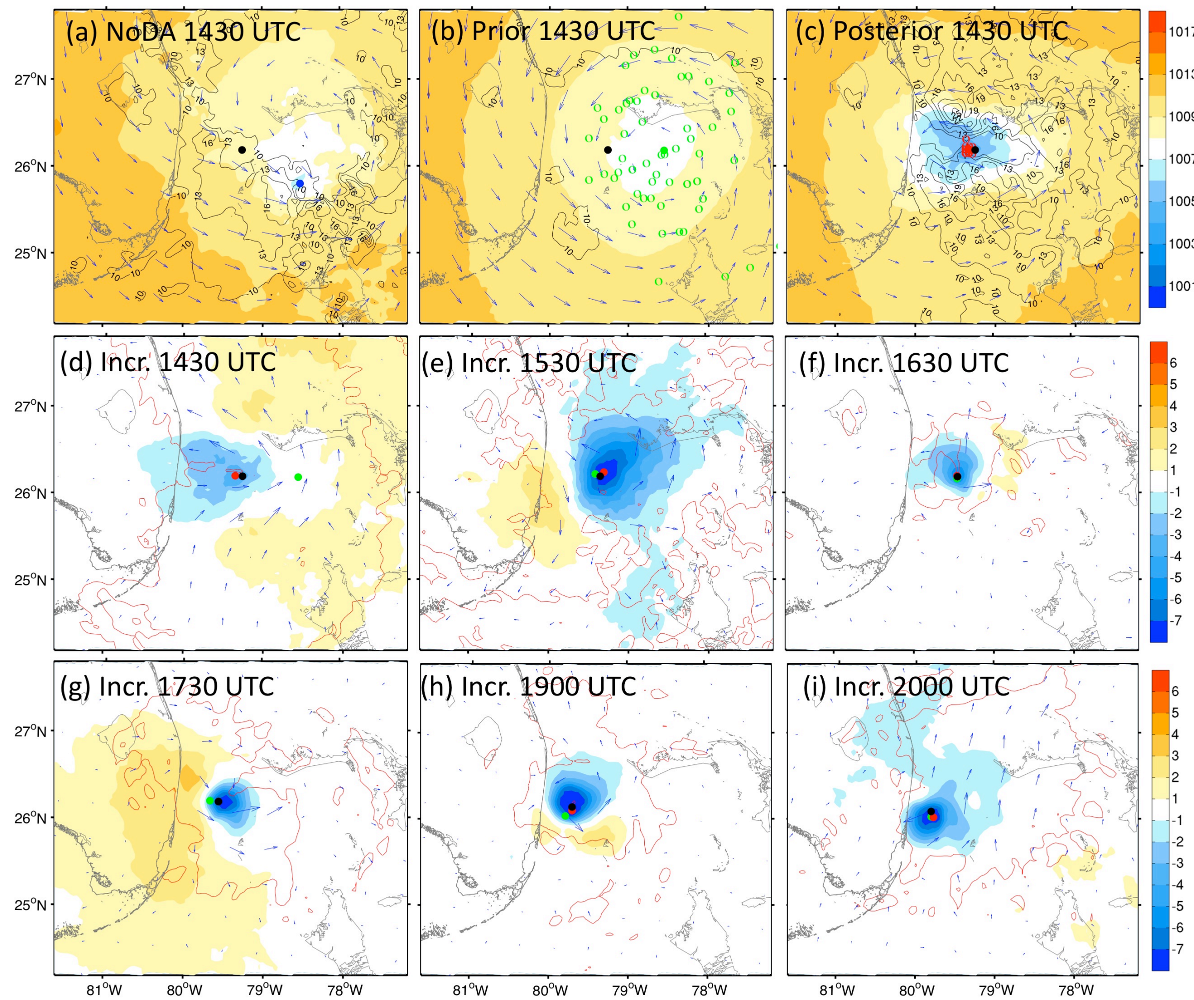

Figure 10. Sea level pressure (SLP) (shaded), the lowest level winds (arrows) and wind speed (contours started at 10 m s^{-1} every 3 m s^{-1}) for (a) NoDA, (b) prior and (c) posterior at 1430UTC 25 Aug 2005; and increments of SLP (shaded), lowest model level winds (arrows) and 3 m s^{-1} wind speed (red contour) between posterior and prior after assimilating the airborne radar SOs of 1-6 legs at (d) 1430, (e) 1530, (f) 1630, (g) 1730, (h) 1900 and (i) 2000 UTC, respectively. Black dot is the storm center position linearly interpolated from the NHC Best Track; blue dot is the storm center of NoDA; green dot is the storm center position calculated from the ensemble mean prior; red dot is that of posterior that assimilated the first P3 leg radar radial velocity; green circles are the centers of the ensemble members of prior; and red circles (overlapped by red dots) are the centers of the ensemble members of posterior.

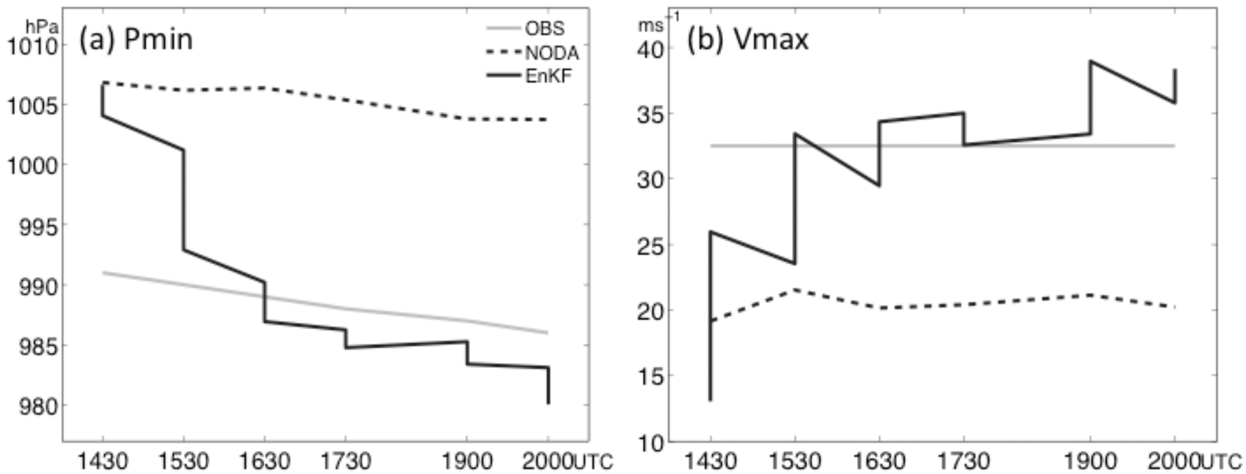

Figure 11. Evolution of (a) minimum sea level pressure (hPa) and (b) maximum surface wind speed (m s^{-1}) for OBS (gray solid) that is linearly interpolated from the NHC Best Track, NoDA (dash) and EnKF (solid).

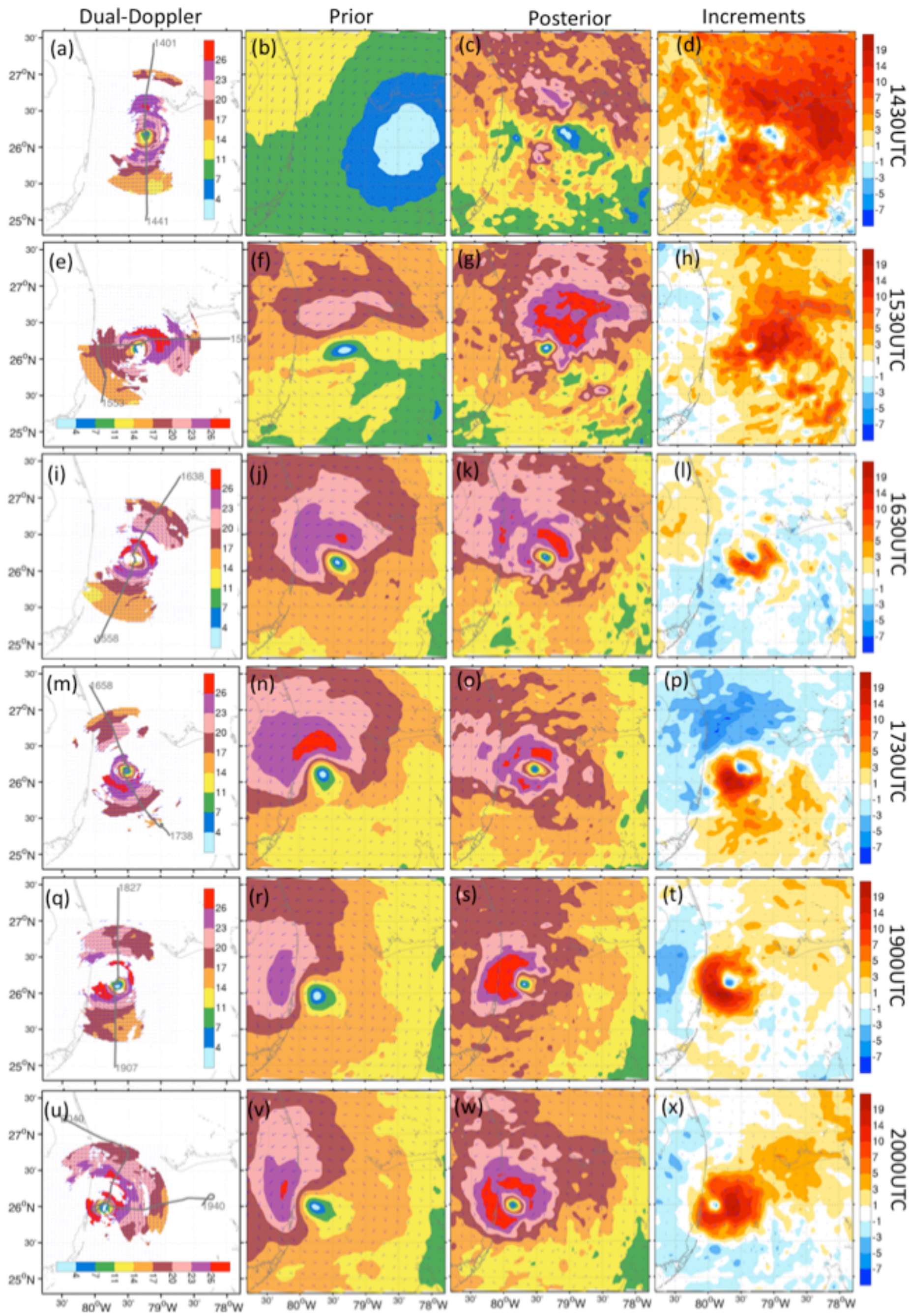

Figure 12. Winds and wind speed at 3 km height. The columns from left to right are the dual-Doppler radar analyses, prior, posterior and EnKF increments (posterior minus prior), respectively. The rows from the top to the bottom are for 1430, 1530 1630, 1730, 1900 and 2000 UTC, respectively. The gray lines in the 1st column are tracks of the NOAA P-3 flight. The circled shaded areas in the last column indicate the increments are less than 1 m s^{-1}.

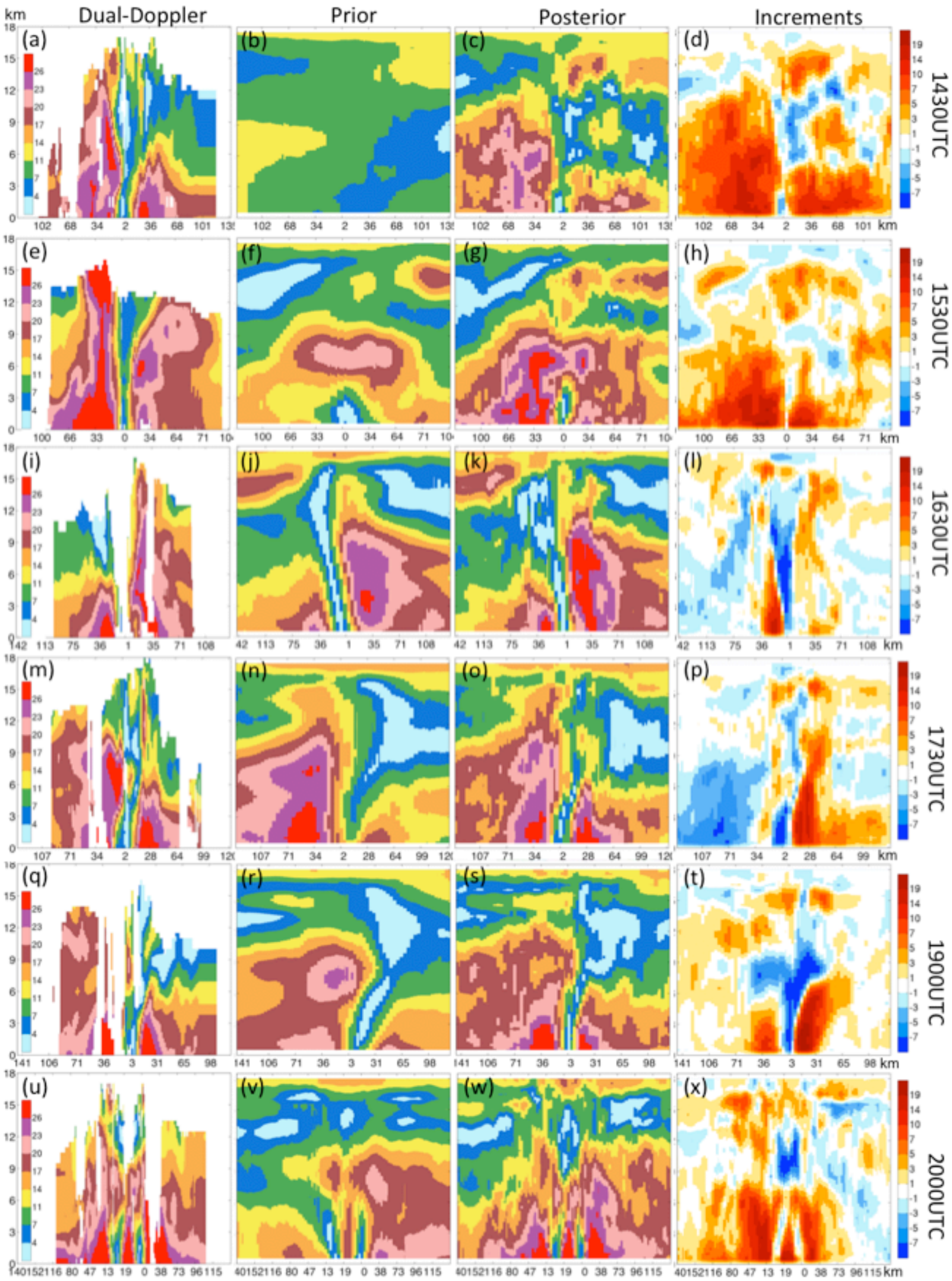

Figure 13. Track-height wind speed sections along the leg tracks shown in the first column of Fig. 12. Other information is the same as that in Fig. 12.