Key Points:

- Both ground-based radar and airborne radar observations are critical to analyzing and predicting the landfalling of hurricane Harvey
- The continuous ground-based radar observations benefit the analysis and prediction of Harvey more than the intermittent airborne datasets
- Combining both observations shows complementary effects and produces Harvey's overall best analysis and predictions

Correspondence to:

X. Lu,
luxu@ou.edu

Author Contributions:

Conceptualization: Xu Lu, Xuguang Wang
Data curation: Xu Lu, Xuguang Wang
Formal analysis: Xu Lu, Xuguang Wang
Funding acquisition: Xuguang Wang
Investigation: Xu Lu, Xuguang Wang
Methodology: Xu Lu, Xuguang Wang
Project Administration: Xuguang Wang
Resources: Xuguang Wang
Software: Xu Lu, Xuguang Wang
Supervision: Xuguang Wang
Validation: Xu Lu, Xuguang Wang
Visualization: Xu Lu, Xuguang Wang
Writing – original draft: Xu Lu
Writing – review & editing: Xu Lu, Xuguang Wang

A Study of Simultaneous Assimilation of Coastal Ground-Based and Airborne Radar Observations on the Prediction of Harvey (2017) With the Hourly 3DEnVar System for HWRF

Xu Lu[1] and **Xuguang Wang[1]**

[1]School of Meteorology, University of Oklahoma, Norman, OK, USA

Abstract This study investigates the relative impact of assimilating the ground-based WSR-88D radar (GBR) and the airborne Tail Doppler Radar (TDR) observations on the analysis and prediction of Hurricane Harvey (2017) during its landfalling stage. Results show that the assimilation of GBR (experiment "DAG") outperforms the assimilation of TDR (experiment "DAT") in multiple aspects. For example, "DAG" produces better predictions in V_{max}, radial velocity, brightness temperature, and precipitation than "DAT." The advantages of "DAG" over "DAT" are likely from the better analyzed thermodynamical structures in addition to the better data availability. The average Minimum Sea Level Pressure (MSLP) prediction error is the only aspect of "DAG" that is inferior to "DAT." Diagnostics show that such an inferior performance of MSLP for "DAG" is associated with the systematic bias from the Hurricane Weather Research and Forecasting model. The combined assimilation of both observations (experiment "DAB") shows complementary effects and performs the best overall among all experiments.

Plain Language Summary The Ground-based WSR-88D Radar (GBR) and the airborne tail Doppler Radar (TDR) observations have been separately assimilated and evaluated in the analysis and prediction of landfalling hurricanes in the past. This study explores their relative and combined impacts. Using a newly developed, GSI-based, hourly hybrid three-dimensional ensemble-Kalman-Filter-Variational data assimilation system for the Hurricane Weather Research and Forecasting, experiments are performed for the landfalling stage of Hurricane Harvey (2017). Results showed that the continuous availability allows GBR to produce better and more persistent improvements in the analysis and the subsequent forecasts than TDR. Investigations show that the improvements are not only in the wind fields but also in the precipitation and brightness temperature predictions. The combined assimilation shows complementary effects and further improves upon GBR alone.

1. Introduction

While landfalling hurricanes are notorious for their catastrophic impacts on coastal regions, accurate predictions remain a big challenge as always. One of the major difficulties for hurricane predictions, especially intensity predictions, comes from the need for more high-density and high-quality inner-core observations to provide accurate initial conditions for numerical models.

One common source of inner core observations is instruments used during airborne field campaigns. An example is the tail Doppler radar (TDR) observations onboard the National Oceanic and Atmospheric Administration (NOAA) WP-3D aircraft (Rogers et al., 2012). This sampling of TC inner-core regions using the X-band (3 cm wavelength) TDR can be traced back to the 1980s (Jorgensen, 1984; Marks & Houze, 1984). However, the assimilation of TDR observations only showed promising results in improving the hurricane predictions in the recent decade after high-resolution (e.g., convection-permitting) numerical models with ensemble-based data assimilation (DA) capabilities emerged (Aksoy et al., 2013; Feng & Wang, 2019; Lu, Wang, Li, et al. (2017); Lu, Wang, Tong, et al. (2017); Weng & Zhang, 2012; F. Zhang et al., 2011; S. Zhang & Pu, 2019). For example, using a newly developed, end-to-end Ensemble-Variational (EnVar) DA system with the Hurricane Weather Research and Forecasting (HWRF) model, Lu, Wang, Li, et al. (2017); Lu, Wang, Tong, et al. (2017) showed that the assimilation of TDR radial velocity observations could systematically improve hurricane intensity predictions. Consequently, the assimilation of TDR observations with a fully

cycled HWRF ensemble has been implemented into the operational HWRF since 2017 (Biswas et al., 2018). However, one major limitation of the airborne observations is the significant temporal and spatial discontinuity due to the nature of aircraft missions, for example, the restricted airtime. To be specific, the NOAA WP-3D aircraft are normally deployed twice a day, and each flight can spend only 3~4 hr or even less around the storm center.

For landfalling hurricanes, the existing coastal radar observations can be another good source for inner-core sampling. The Weather Surveillance Radar-1988 Doppler (WSR-88D) radar has been installed across the continental United States (CONUS), including coastal regions, as a part of the Next Generation Weather Radar (NEXRAD) system (Crum & Alberty, 1993) for decades. These high-resolution Doppler radar observations have been widely assimilated by numerical models for predicting convective-scale weather systems, such as thunderstorms and supercells (Snyder & Zhang, 2003; Stensrud et al., 2009; Sun, 2006). Nevertheless, these ground-based radars (GBR) observations have not been used by operational hurricane models like HWRF until recently (Sippel et al., 2021), and their applications on landfalling hurricanes are still limited (Zhang et al., 2009). There have been multiple studies assimilating those GBR observations to explore their impacts on landfalling hurricane predictions (Dong & Xue, 2013; Green et al., 2022; Li et al., 2012; Shen et al., 2016; Y. Wang & Pu, 2021; Xiao et al., 2007; Zhao & Jin, 2008; Zhao & Xue, 2009; Zhu et al., 2016). These early works showed encouraging results, especially with advanced DA techniques and high-resolution models. For example, Li et al. (2012) investigated the impact of assimilating radial velocity observations from two WSR-88D radar sites during hurricane Ike (2008) using an advanced three-dimensional ensemble-Kalman-Filter-Variational (3DEnVar DA) system at a 5-km grid spacing. They found improvements not only in the storm structure analyses, and track and intensity forecasts but also in the precipitation prediction skill. Such improvements were consistently found in the other early works, although the improvements varied depending on the DA system and the model resolution.

Even though the impacts of the assimilation of either airborne or ground-based Doppler radar observations have been studied individually, studies that examine their relative impacts on the analyses and predictions of landfalling hurricanes are limited. Green et al. (2022) investigated the relative performance of the assimilation of these two types of radar observations during the eyewall replacement cycle (ERC) of Hurricane Matthew (2016). They found distinct impacts from the individual assimilation of the two datasets. It is found that the continuous availability of GBR benefits the analysis and prediction of ERCs more than the discontinuous availability of TDR. GBR observations dominate the impact of the combined assimilation of both datasets. In their study, the superior performance of GBR was attributed to the rapid evolution of the ERC. Nevertheless, the impact of the assimilation of GBR observations versus TDR observations can depend on the types of processes that the TC experiences. Unlike Green et al. (2022), who's study focuses on the ERC, this study will investigate the relative and combined impacts of both types of inner-core radar observations for a different TC landfalling process.

Harvey was a category four hurricane, which formed a tropical depression on 17 August 2017 and transformed into an extratropical cyclone on 1 September 2017. It first landed on the north of San Jose Island around 0300 UTC on 26 August. Within 3 hr, Harvey hit the Texas mainland and moved northwestward until the late 26th. Then, the storm looped back to the ocean around 0300 UTC on 28 August to re-develop itself. The final landfall was on the southwestern coast of Louisiana at 0800 UTC on 30 August. This two-week-long tropical cyclone produced the most significant precipitation event for a landfalling TC in the history of the US and caused 68 direct casualties and about $125 billion in damages (Blake & Zelinsky, 2017). However, the then-operational models failed to provide a consensus on its track predictions before the first landfall (http://hurricanes.ral.ucar.edu/realtime/plots/northatlantic/2017/al092017/track_early/aal09_2017082518_track_early.png, accessed on 6 November 2022). The unprecedented precipitation and the large track uncertainties around landfall make Harvey a popular research topic (e.g., Galarneau & Zeng, 2020; Ko et al., 2020; P. Wang et al., 2019). Before Harvey's first landfall, four TDR missions were conducted to sample its inner-core evolution. The last flight was within the GBR range and captured the end of the intensification of Harvey before its first landfall. Therefore, this setting provides an opportunity to address further the impact of the assimilation of GBR and TDR for hurricane prediction.

This manuscript is organized as follows: The hourly 3DEnVar hybrid DA system for HWRF, the observations, and the experiment designs are described in Section 2. The results of the experiments are discussed in Section 3. Section 4 concludes and further discusses the paper.

2. Model, Data, and Experiment Design

2.1. System and Model Description

The hourly 3DEnVar DA system used in this study is based on a newly developed dual-resolution, hybrid 3DEnVar DA system for HWRF (Lu, Wang, Li, et al., 2017 Lu, Wang, Tong, et al., 2017) with adaptions to the hourly DA capabilities (Davis et al., 2021). Briefly, a 40-member double-nested, self-cycled HWRF EnKF ensemble runs in parallel with the triple-nested control, which runs at a higher model resolution than the ensemble. During GSI variational minimizations, the higher-resolution first guesses from the control will digest the coarser-resolution ensemble error covariance through an augmented control vector (GSI-ACV) method (X. Wang et al., 2013). No static error covariance is employed following Lu, Wang, Li, et al. (2017); Lu, Wang, Tong, et al. (2017). The horizontal and vertical localizations used in this study are configured to be 60 (~220 km cutoff radii) and −0.4 (~330 hPa) scale-height recursive filter localization length scale following Li et al. (2012). The ensemble square root Kalman Filter (EnSRF) adopted for the 40-member ensemble utilizes the observation pre-processing and forward operators from GSI and is tuned to be comparable with the GSI 3DEnVar in localization length scales. The mean of the EnKF analysis ensemble will be replaced by the 3DEnVar analysis through recentering. Then, both the EnKF analysis ensemble and the 3DEnVar analysis will be used to launch a 1 hr forecast to provide the background for the next DA cycle. The hourly lateral boundary conditions (LBC) and initial conditions (IC) are interpolated from the operational global forecast system (GFS) (X. Wang et al., 2013), which was stored at a three-hourly frequency. In the meantime, a 5-day free forecast is initialized from the control analysis at each hour. More details about this hybrid EnVar DA system can be found in Lu, Wang, Tong, et al. (2017), Lu and Wang (2019), and Davis et al. (2021).

The HWRF model configuration used in this study is comparable with the 2018 operational HWRF (Biswas et al., 2018) with adaptions from Lu et al. (2022). The horizontal grid spacing of the control is approximately 1.5, 4.5, and 13.5-km for the inner (411×820 grid points), middle (343×682 grid points), and outer (390×780 grid points) domains, respectively. The ensemble members are only running with the middle and outer domains. The model is topped at 10 hPa with 75 vertical levels. The detailed model physics and parameterization schemes follow the 2018 operational HWRF and can be found in the study by Biswas et al. (2018).

2.2. Observations, Pre-Processing, and Verifications

In this study, the high-resolution inner-core observations to be discussed are TDR and GBR observations. Since 2013, TDR observations have been available in the operational HWRF data stream (Tallapragada et al., 2013). Therefore, TDR observations and other flight-level, conventional, and satellite observations are obtained directly from the operational HWRF data stream. The pre-processing and quality control of those data types thus follow the operational HWRF, and the details can be found in the study by Biswas et al. (2018).

GBR observations used in this study were obtained from the National Center for Environmental Information (NCEI), or the former National Climatic Data Center (NCDC). Since this study focuses on the storm evolution before Harvey turned eastward around 0000 UTC, 27 August 2017, we only address the radial velocity observations from the following three radar sites in Texas in this study: Corpus Christi (KCRP), Brownsville (KBRO), and Houston (KHGX) sites (Figure 1a). The obtained level II observations were then de-aliased, and quality controlled through the Warning Decision Support System - Integrated Information (WDSS-II) software (http://www.wdssii.org). An example sweep of the pre-processed GBR observations is shown in Figure 1b. The negative and positive values correspond to the winds toward and away from the KCRP radar site, respectively. Figure 1b shows southeast winds in the northeastern quadrant and northwest winds in the southwestern quadrant. Such a feature is consistent with the cyclonic circulation from Harvey, which centered around 27.1°N and 96.3°W at 2518 UTC, August 2017. Next, these pre-processed GBR observations were homogeneously thinned through 10-km $\times$ 10-km $\times$ 500-m boxes during the hourly 3DEnVar assimilation in GSI. The observation error is set to 2 ms^{-1} following Li et al. (2012).

In order to verify the analysis and forecast of each DA experiment, various independent observations are used in this study. Specifically, the best track data from the national hurricane center (NHC; obtained from https://www.nhc.noaa.gov/data/hurdat/hurdat2-1851-2019-052520.txt, accessed 6 November 2022) is used to verify the V_{max}, Minimum Sea Level Pressure (MSLP), and Track of the model forecasts every 6 hr (at the synoptic times) following the operational HWRF. Those Automated Tropical Cyclone Forecasting System (ATCF) -related values
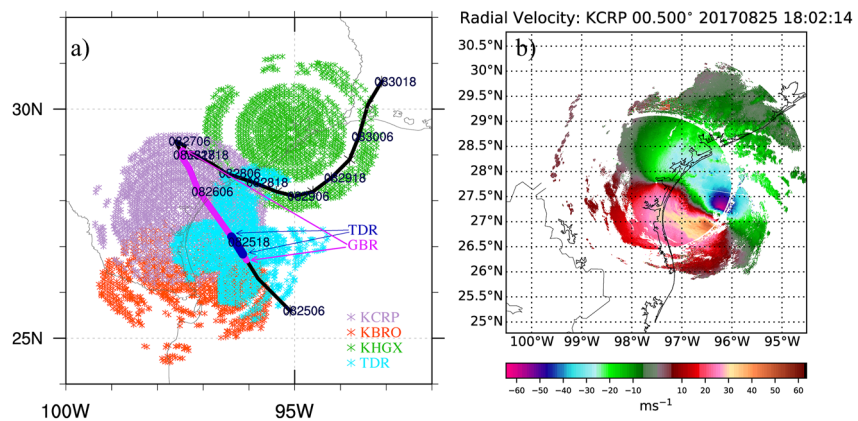
Figure 1. (a) Horizontal distribution sample of ground-based WSR-88D radar (GBR) and Tail Doppler Radar (TDR) radial velocity observations. The black line shows the best track of Harvey (2017), and the purple and blue lines show the period of GBR and TDR observations assimilated in this study, respectively. (b) An example sweep of the post-processed radial velocity observations from the KCRP radar site valid at 1802 UTC, 25 August 2017, at the 0.5° elevation angle.

(V_{max}, MSLP, and Track) from the models are calculated through the Geophysical Fluid Dynamics Laboratory tracker (Marchok, 2002, 2021). Since the best track data is only available at a 6-hourly interval, only the analyses and forecasts valid at the synoptic times are verified in this study. In other words, during the verifications, there will be 1–3 hr of offsets between the hourly DA cycles. For example, the 48-hr forecast initialized from 1800 UTC, 25 August 2017 will be compared with the 45-hr forecast initialized from 2100 UTC, 25 August 2017 when they are verified against the best track data valid at 1800 UTC, 27 August 2017.

To verify the analyzed inner-core dynamic structure of Harvey, a radar composite from the Hurricane Research Division (HRD; obtained from https://www.aoml.noaa.gov/hrd/Storm_pages/harvey2017/radar.html, accessed 6 November 2022) is used. This synthetic radar composite is post-processed at HRD and differs from the TDR observations assimilated. There will only be one HRD radar composite valid at 1800 UTC, 25 August 2017, available during this period of study.

Additionally, the mobile, C-band dual-polarimetric Shared Mobile Atmospheric Research and Teaching (SMART) radar observations are used as another independent verification metric for the dynamic fields. As a newly available dataset, the SMART radar observations supplement the GBR and are obtained from Alford et al. (2019). These high-frequency SMART radar observations are available between 2100 UTC to 0600 UTC, on 25 August 2017 at 5-min intervals.

To verify the precipitation forecast, a 4-km gridded Stage-IV precipitation product is used (obtained from https://data.eol.ucar.edu/dataset/21.093, accessed on 6 November 2022). These Stage-IV precipitation observations are produced by the 12 River Forecast Centers (RFCs) in the CONUS from the NEXRAD observations.

To verify the thermal structures of the predictions, a cloud-top brightness temperature (BT) product from the GOES-13 satellite band 4 (10.7 µm central wavelength) is used (obtained from https://www.avl.class.noaa.gov/saa/products/welcome, accessed 6 November 2022) following Lu and Wang (2020). The corresponding model-derived precipitation and BT outputs are generated through the Unified Post Processor (UPP; https://dtcenter.org/sites/default/files/community-code/upp-users-guide-v4.pdf, accessed 6 November 2022).

Note, due to the large track errors at long-term forecasts, the ETS score against the Stage IV observations and the correlation coefficients against the GOES-13 observations are all calculated in a storm-relative framework. Specifically, the predicted structures will be relocated to the observed locations before the ETS or correlation coefficient calculation. All the statistical significances between experiments are evaluated using the *t*-test with a null hypothesis that the sample means are from the same population.

2.3. Experiment Design

Three sets of experiments have been designed to address the scientific goals of this study. They are named "DAT", "DAG", and "DAB", respectively, as shown in Table 1. To be more specific, each experiment is detailed as follows.

Table 1
List of Experiments

Experiment name	DAT	DAG	DAB
High-resolution inner-core data assimilated	TDR + flight-level observations	GBR + flight-level observations	TDR + GBR + flight-level observations
Other operational data assimilated	Conventional Observations from prepbufr, TCVital, satellite radiances, and satellite-derived wind		

Experiment "DAT" assimilates all the operational observations, including TDR. The experiment is first initialized from 0600 UTC, 25 August 2017, using the GFS analysis and forecasts as IC and LBC following Lu, Wang, Li, et al. (2017); Lu, Wang, Tong, et al. (2017). After 6 hr of spin-up, a 6-hourly 3DEnVar is performed at 1200 UTC, 25 August 2017, following the operational HWRF and Lu, Wang, Li, et al. (2017); Lu, Wang, Tong, et al. (2017). No vortex relocation or modification was performed before any DA cycles in this study. Then a 3-hr forecast is launched to provide the background for the first hourly 3DEnVar DA cycle valid at 1500 UTC, 25 August 2017. Next, the hourly forecast and DA cycles follow Section 2.1, and the last cycle ends at 2100 UTC, 26 August 2017. In this "DAT" experiment, TDR observations are only available between 1625 UTC and 1838 UTC, 25 August 2017, as shown in Figure 1a.

Experiment "DAG" differs from "DAT" by replacing the TDR observations with the GBR observations during DA in the inner-core region. As shown in Figure 1a, the GBR observations are continuously available during all the 31 hourly DA cycles from 1500 UTC, August 25 to 2100 UTC, 26 August 2017. Hence, comparing "DAT" and "DAG" for the TDR available cycles is expected to provide a direct comparison of the relative impacts of the assimilation of each data type on the hurricane analysis and predictions. Furthermore, the comparisons between the two experiments for the TDR unavailable cycles after 1900 UTC, 25 August 2017, are expected to illustrate the necessity of continuous inner-core sampling and assimilation.

Experiment "DAB" assimilates all the observation types listed in Table 1, including both TDR and GBR. This experiment investigates whether the two types of radar wind observations can be complementarily assimilated.

3. Results

3.1. Statistics on All DA Cycles

As designed, all three experiments are continuously cycled at an hourly interval from 1500 UTC, 25 August to 2100 UTC, 26 August. The predictions initialized from these 31 DA cycles are first shown in Figure 2, verified against the best track.

Although the then-operational models did not have consistency on the track predictions during the period, all three experiments in this study do not show a significant diversity. All experiments at every cycle can capture the movement pattern of Harvey, which moved northwestward in the early stages, turned southeastward back to the ocean, and eventually made another landfall toward the northeast. Overall, "DAT" produces better track predictions at the early lead times. However, "DAT" tends to have an eastward bias at longer lead times (e.g., the location of the second landfall) as compared to the best track (Figure 2c). In contrast, "DAG" and "DAB" produce better track predictions after about 80 hr. Nevertheless, these two experiments assimilating GBR show eastward bias compared to the best track at early lead times in the first several cycles (Figures 2f and 2i). Such performances are quantitively shown by the averaged track errors shown in Figure 3a.

A more distinct difference between the experiments is from the intensity predictions in the early DA cycles. There were only limited TDR observations available during the 1600 UTC DA cycle. Specifically, the flight was still about 200 km northeast of the storm and was not sampling the storm center (to be shown later in Figure 5i). Therefore, there are hardly any inner-core corrections for "DAT" until 1700 UTC. Figure 2a shows that the early lead-time V_{max} and MSLP predictions in "DAT" are much weaker than the best track in the first two DA cycles. After the inner-core samplings from TDR are available beginning the third cycle, "DAT" quickly modifies its V_{max} and MSLP analyses and the corresponding forecasts. However, issues like the spuriously strong V_{max} to be discussed in Section 3.3 can be found in the later cycles when the short-lived TDR observations are no longer available. Also, the V_{max} predictions during the rapid weakening (re-intensification) stage are weaker (stronger) than the best track.
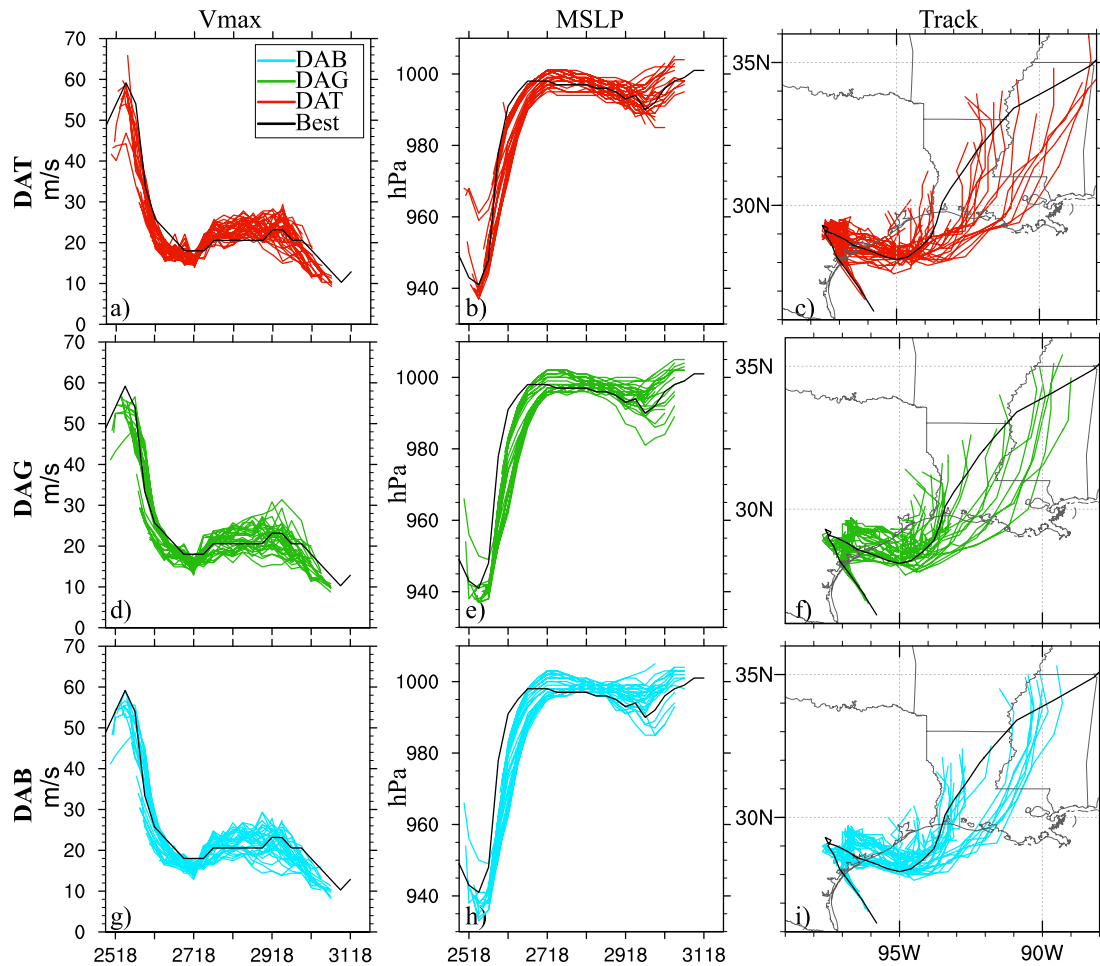
Figure 2. (a), (d), (g) V_{max}, (b), (e), (h) minimum sea level pressure, and (c), (f), (i) track predictions for all 31 data assimilation cycles for (a)-b) "DAT", (d)-f) "DAG", and (g)-i) "DAB" in verification with the best track (black). The average of forecasts valid at the same time for each experiment is shown in the purple line for reference.

Compared to the TDR observations, the GBR observations are always available during the hourly cycles (Figure 1a). The major issue of GBR is the ~200-km range limit (Figure 1b). Also, due to the minimum elevation angle, when the storm is around the edge of the detecting range, GBR can only cover the very upper portions of the inner core. Therefore, the first DA cycle in "DAG" and "DAB" is only slightly better than the "DAT" experiment as the storm is about 200 km away. However, as Harvey moves closer to the mainland, the V_{max} and MSLP analysis and predictions from "DAG" and "DAB" quickly catch up with the best track starting from the second DA cycle. Additionally, due to the continuity of the dataset, both GBR experiments do not possess the sudden V_{max} jump issue found in "DAT." As a result, the average V_{max} predictions are produced by "DAG" and "DAB" in Figure 3b, which are statistically significantly better than "DAT" most of the time. Due to the lack of lower-level inner-core corrections, the early DA cycles from "DAG" tend to produce a weaker peak V_{max} prediction than the best track (Figure 2d). In comparison with "DAG", "DAB" can better depict the peak V_{max} evolution due to the additional assimilation of the complementary TDR observations (Figure 2g). Such a complementary effect will be discussed in more detail with an example case in Section 3.2.

Although "DAG" and "DAB" show apparent improvements over "DAT" in the V_{max} predictions, the improvements are not consistently found in the MSLP predictions. Instead, "DAG" and "DAB" produces worse MSLP predictions than "DAT" (Figure 3c). Such a disagreement between the MSLP and V_{max} predictions is found to be related to the systematic bias in the HWRF model. The inconsistent V_{max} and MSLP relationship has been a known issue for the HWRF and is likely due to the physics representations (Bao et al., 2012). With improvements in the HWRF physics, we do see improvements in the V_{max} and MSLP relationship in some cases (e.g., "6H-3DEnVar" in figure 3c of Davis et al., 2021). But as shown in this study, additional investigations into the
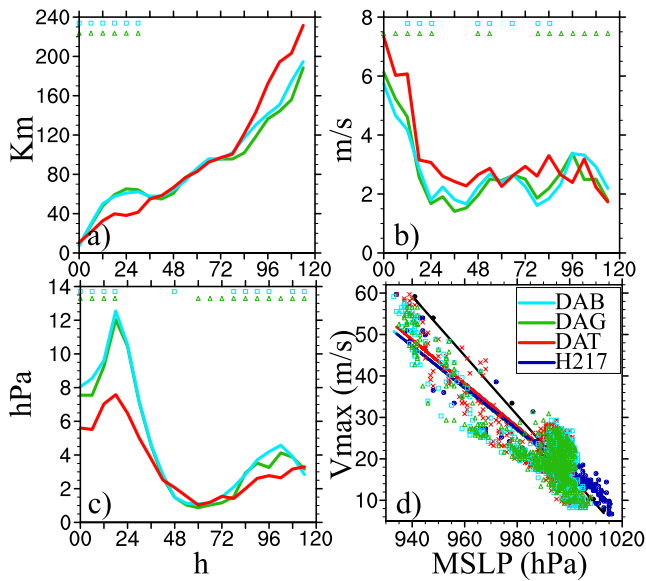
Figure 3. Mean absolute error of all 31 data assimilation cycles for (a) track, (b) V_{max}, and (c) minimum sea level pressure (MSLP) predictions for "DAT" (red), "DAG" (green), and "DAB" (cyan). The cyan squares indicate a statistically significant difference between "DAB" and "DAT" at a 90% level. The green triangles indicate a statistically significant difference between "DAG" and "DAT" at a 90% level. No statistically significant difference can be found between "DAB" and "DAG" in all metrics and therefore not shown. The V_{max} and MSLP relationship for each experiment is shown in (d) in verification with the best track (black) and the 2017 operational Hurricane Weather Research and Forecasting (blue).

issue are still necessary for future work. The MSLP and Vmax relationship for all analyses and predictions for each experiment is shown in Figure 3d. The corresponding relationship from the 2017 operational HWRF is also plotted for reference (data obtained from ATCF A-Deck: https://ftp.nhc.noaa.gov/atcf/archive/2017/, accessed on 6 November 2022). Figure 3d shows that the observations have a steeper slope in the wind and pressure relationship than all experiments, including the operational HWRF. Such a relationship means that given the same V_{max} value, the MSLP produced by the model tends to be lower than the best track. Therefore, when the V_{max} predictions are corrected in "DAG" and "DAB", their MSLP predictions become lower than "DAT", inconsistent with the reality.

Following Section 2.2, additional verifications are performed against various independent observations to quantify the impacts on the analysis and forecasts of each experiment. First, diagnostics are performed using the SMART radar observations to verify the DA analyses. Due to the limitation of data availability, only the hourly analyses between 2100 UTC, 25 August, and 0600 UTC, 26 August are verified against the corresponding SMART radar observations. The level-dependent correlation coefficients of each analysis against the observations are shown in Figure 4a. The higher the correlation coefficients, the better the similarities between the analysis and the observations. Among the three experiments, "DAT" shows the smallest coefficients at all levels, which indicates the worst analysis among all experiments. Between the two GBR assimilated experiments, "DAG" matches the SMART radar observations slightly better than "DAB." However, only the difference between "DAG" and "DAT" at 7,500-m is statistically significant at a 90% level.

Then, the GBR, Stage IV, and GOES-13 observations are used to verify the dynamic and precipitation predictions from all cycles. Figure 4b shows the root mean square error (RMSE) of the model predictions verified against the GBR observations up to forecast hour 108. Overall, "DAT" produces the largest RMSE among all experiments at almost every forecast lead time. "DAB" is comparable to "DAG" in the first 72 hr and improves afterward. These results suggest that the continuous assimilation of the GBR observations benefits the model predictions more than the TDR observations. Also, they show that the assimilation of both observations can further improve the predictions of the dynamic structures of hurricanes at longer lead times. The averaged storm-relative Equitable Threat Score (ETS) of the accumulated precipitation predictions at the 150 mm threshold from all three experiments is calculated against the corresponding Stage IV observations. Figure 4c shows that the assimilation of GBR observations produces better rainfall predictions than "DAT" during the heavy precipitation period of Harvey at most of the forecast lead times (seven out of nine). "DAB" is overall comparable with "DAG" in the precipitation predictions. The averaged storm-relative correlation coefficients against the GOES-13 BT product are shown in Figure 4d. The results show that the performance of the BT predictions from the three experiments is overall mixed. The major disadvantages of "DAB" and "DAG" over "DAT" are between hours 84–96. Additional diagnostics show that the issue can be attributed to the timing of the final landfall (not shown). To better understand the differences in the statistics above, example cycles are selected and discussed in more detail in the following sections.

3.2. Analysis and Prediction Performance During the TDR Available Cycles

As shown in Figure 5i, although the TDR data is available between 1625 UTC and 1836 UTC, 25 August 2017 (Figure 5i), 1800 UTC is the final TDR cycle with a complete sampling of the TC inner core. Therefore, this section will focus on this particular DA cycle to compare the relative impact of the DA experiments in both analysis and predictions.

Figure 5 shows the horizontal wind analyses (valid at 1800 UTC, 25 August 2017) produced by each experiment in verification against the HRD radar composite in both vertical and horizontal directions. At this time, Harvey (2017) was already a category three hurricane with a surface wind maximum (V_{max}) of around 54 ms^{-1} and was
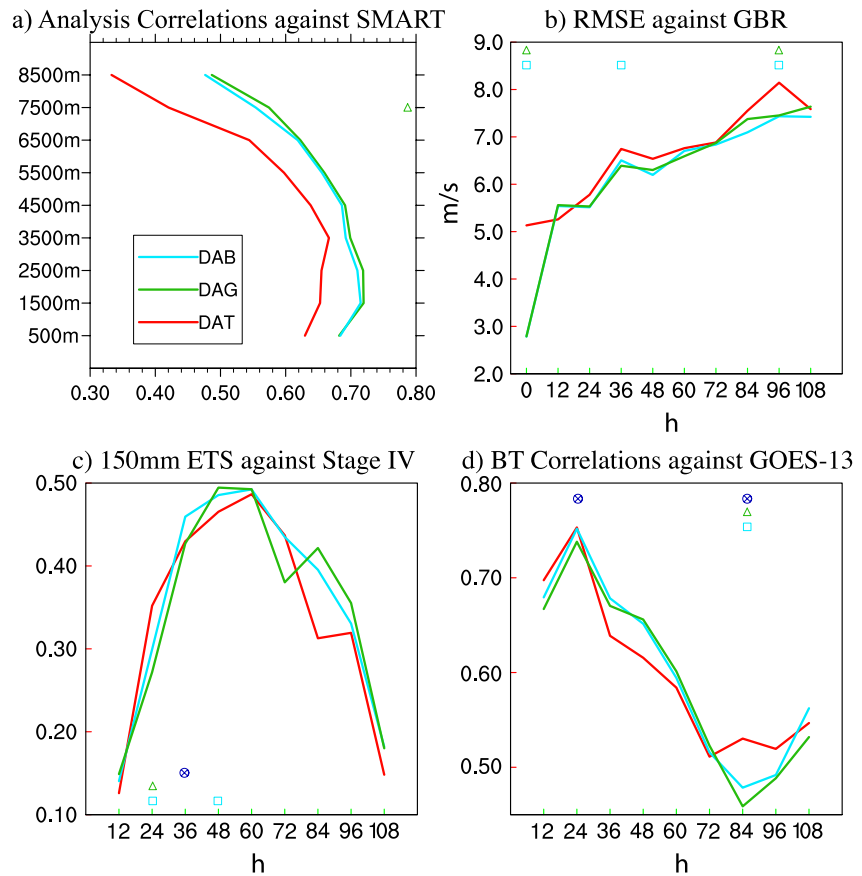
Figure 4. (a) Level dependent correlation coefficient of each experiment analysis against the Shared Mobile Atmospheric Research and Teaching radar observations between 2100 UTC, August 25, and 0600 UTC, 26 August 2017. (b) Mean root mean square error of the predictions against the ground-based WSR-88D radar observations, (c) mean 150 mm accumulated storm-relative Equitable Threat Score against the stage IV product, and (d) mean correlation coefficient against the GOES-13 Band 4 brightness temperature observations from each experiment over 31 data assimilation cycles. The cyan squares indicate a statistically significant difference between "DAB" and "DAT" at a 90% level. The green triangles indicate a statistically significant difference between "DAG" and "DAT" at a 90% level. The blue circles indicate a statistically significant difference between "DAG" and "DAB" at a 90% level.

rapidly intensifying into category four. Figures 5a, 5e and 5i suggest that the wind maxima of Harvey at this time were primarily in the northeast quadrant of the storm below 4 km height. In the meantime, the NOAA WP-3D aircraft was penetrating the storm from east to west (Figure 5i). Such a flight pattern is a typical sampling strategy for TDR and nicely captured the inner-core region of Harvey along the flight (Figures 5b, 5f and 5j). With these abundant inner-core observations assimilated, the corresponding analysis from "DAT" reproduces major features in the HRD radar composite. For instance, the cross-sections in Figures 5b and 5f show comparable storm sizes as those in Figures 5a and 5e. The wind maxima centered around 3 km height in the east and around 1 km height in the north are also consistent with the observations. Figure 5j shows a comparable storm in eye size and wind patterns as Figure 5i as well. Such a rational analysis is within expectation when "DAT" properly assimilates the TDR observations, which were also used to composite the verifications, following early TDR assimilation studies like Lu and Wang (2019), Lu, Wang, Li, et al. (2017), and Lu, Wang, Tong, et al. (2017). However, there are still differences between the analysis and the observations. In particular, the wind maxima region (the gray area greater than 55 ms^{-1}) in the eastern cross-section extends to 6 km height while it is only up to 4.5-km in the observations (Figure 5b vs. Figure 5a). Also, the wind maxima at the 3-km height in Figure 5j are in the due east of the storm. In contrast, those in the observations are more northeastward (Figure 5i).

Compared with TDR, GBR partially covered the storm center as Harvey was still far from the coast. For example, Figure 5 shows that the GBR scans only captured the middle to the upper part of the inner-core regions of Harvey. The lower portion of the storm, especially the eastern portion, was missed due to the limitation in the minimum
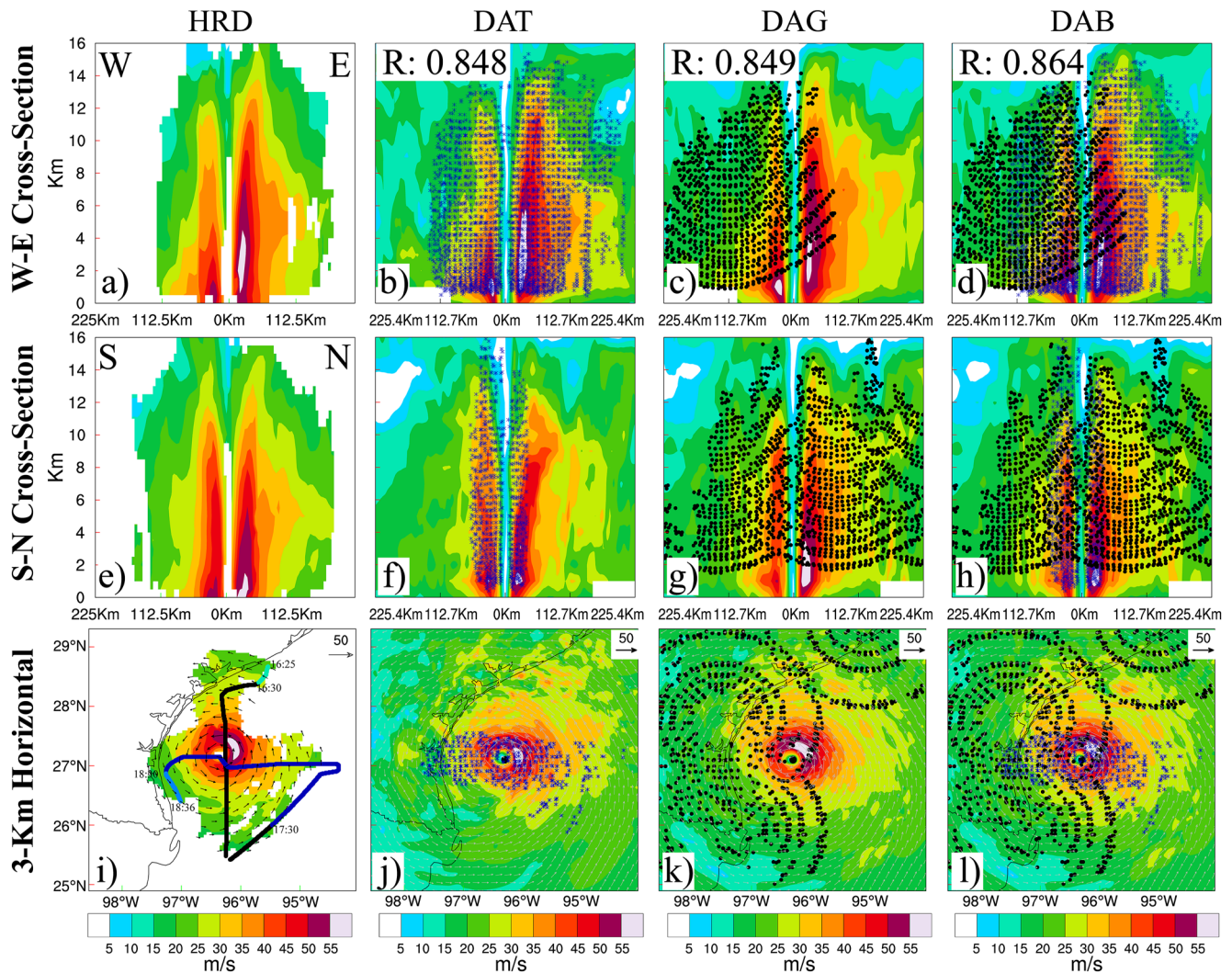
Figure 5. Horizontal wind (vector) and wind speed (shading) in (a)-(d) west-to-east cross-section, (e)-(h) south-to-north cross-section and (i)-l) 3,000-m height for (a), (e), (i) the Hurricane Research Division radar composite centered at 1817 UTC, 25 August 2017, (b, f, j) analysis from "DAT", (c, g, k) analysis from "DAG", and (d, h, l) analysis from "DAB" valid at 1800 UTC, 25 August 2017. The Tail Doppler Radar and ground-based WSR-88D observations at each cross-section and height plot are shown in blue and black dots, respectively. The black dot is the best track location. The co-located 3D correlation coefficient R of each analysis against the Hurricane Research Division radar composite is shown in (b)-d). The flight track of WP-3D is shown in (i) with colors indicating the assimilation period for each hourly data assimilation cycle.

elevation angle of GBR. Still, the analysis produced by "DAG" is comparable with both the observations and "DAT." Note, although there are flight-level observations assimilated in addition to GBR in the experiment "DAG", sensitivity experiments show that the improvements in wind patterns are primarily from the assimilation of the GBR observations. The co-located three-dimensional (3D) correlation coefficient from "DAG" calculated against the HRD radar composite is even slightly higher than that in "DAT" (Figure 5). "DAG" outperforms "DAT" by constraining the vertical extent of the wind maxima region (the gray area greater than 55 ms^{-1}) in the eastern cross-section (Figure 5c).

When assimilating both types of Doppler radar wind observations, the analysis from experiment "DAB" fits the HRD composite the most. For instance, the vertical extents of the wind maxima in both the South–North and the West–East cross-sections are more consistent with the observations in Figures 5d and 5h than those in the other two experiments. So is the position of the wind maxima at 3 km height in Figure 5l. Such a better analysis is consistent with the complementary features in individual experiments "DAT" and "DAG." Consequently, "DAB" quantitively produces the highest correlation coefficient against the observations among all experiments (Figure 5d).
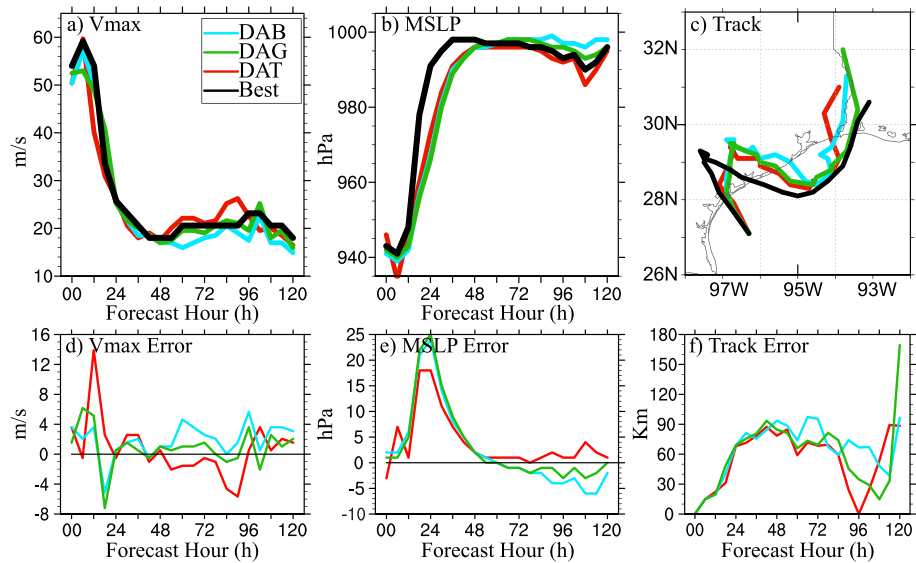
Figure 6. (a) V_{max}, (b) minimum sea level pressure, and (c) track predictions from "DAT" (red), "DAG" (green), and "DAB" (cyan) initialized from 1800 UTC, 25 August 2017, in verification with the best track (black). (d)-f) are the corresponding errors.

Figure 6 compares the V_{max}, Minimum Sea Level Pressure (MSLP), and track predictions of the three experiments against the best track for this particular cycle. In general, all three experiments did a reasonable job. In the track prediction, all experiments move eastward in agreement with the best track in the long term (Figure 6c). The end of the intensification and the subsequent rapid weakening in the short term is also well captured by all experiments, too (Figures 6a and 6b). Among all experiments, "DAT" produces the highest V_{max} peak and the lowest MSLP minimum. Its peak V_{max} matches the best track the most, but its peak MSLP is much lower than the best track. Further diagnostics into the thermodynamic fields suggest that the overly low MSLP predictions in "DAT" are likely due to a relatively stronger upper-level warm-core in the analysis (not shown). Although no other upper-level thermal observations were available to verify if such a stronger warm-core is unrealistic, it is well-known from the early studies that the warmer upper-level warm-core is efficient in reducing the MSLP (e.g., Zhang et al., 2012). Compared to "DAT", the Vmax prediction produced by "DAG" is generally weaker than the best track during the intensification period, although its MSLP predictions match the best track better. Even though "DAG" produces larger track errors than "DAT" in the long-term predictions due to faster movement (Figures 6c and 6f), it predicts the second landfall location more consistently with the best track. Combining both observations in "DAB" produces better peak V_{max} predictions, but slightly degrades the long-term predictions in almost all metrics compared to "DAG" for this particular cycle.

To further understand the performance of the forecast, the rest of this subsection compares the structural forecasts from three experiments in verification against diverse types of independent observations. The examples are selected to best represent the differences between experiments, as shown in Figure 4. Figure 7 shows the 9-hr horizontal wind prediction compared to the SMART radar observations. Harvey was making landfall then, and the observations showed that maximum winds were centered in the northeast quadrant at the 3,500-m altitude in Figure 7d. "DAT" can predict the northeast wind maxima though it is weaker in magnitude. Also, its storm moves more northwestward than observed (Figure 7a). In the west-to-east cross-section, "DAT" is consistently weaker in both east and west quadrants than the observation (Figure 7e vs. Figure 7h), likely due to its early landfall from the track error. In comparison with "DAT", "DAG" produces a more accurate track prediction, though its wind speed in the east-northeast quadrant is even weaker than "DAT" in Figures 7b and 7f. "DAB" produces the best 9-hr forecast of the dynamical field among all experiments. For example, the magnitude and position of the 3,500-m wind maxima and the storm position are in Figure 7c is more consistent with the observations in Figure 7d compared to the other experiments. The magnitude of the wind maxima in the east of the storm (Figure 7g) also matches the observations (Figure 7h) better than the others.

Further investigations are performed with the 12-hr accumulated precipitation predictions in verification with the Stage-IV observations. During this landfalling period, Harvey produced maximum accumulated precipitation of over 500 kg m^{-2} between 1800 UTC 25 August and 0600 UTC 26 August near the Port Bay area (~28ºN,
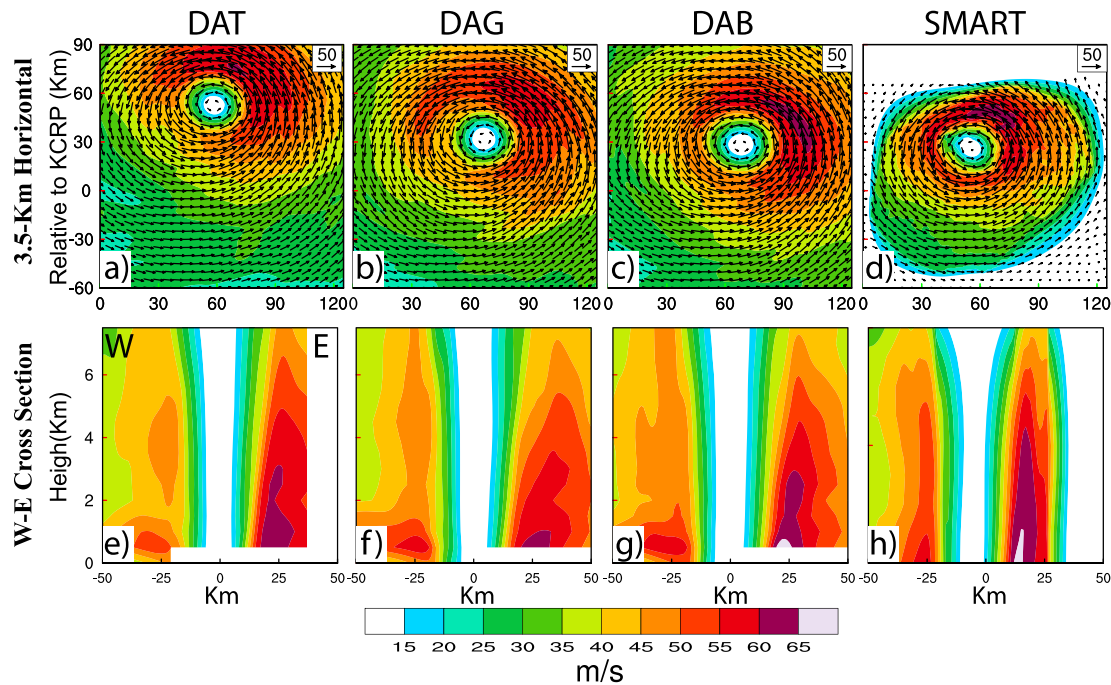
Figure 7. 9-hr horizontal wind (vector) and wind speed (shading) predictions at (a)-d) 3500-m height and (e)-h) west-to-east cross-section for (a), (e) "DAT", (b), (f) "DAG", and (c), (g) "DAB" in verification with (d), (h) the SMART radar observations valid at 0103 UTC, 26 August 2017,. The x-axis and y-axis in (a)-d) are the distance (Km) relative to the KCRP radar site.

97.15ºW; Figure 8d). It appears that "DAT" can produce a precipitation maximum above 450 kg m^{-2}, which is close to the observations in terms of magnitude (Figure 8a). However, "DAT" produces the most precipitation over the ocean instead of over land, although it moves much faster and more inland than the observations or the other two experiments. Additional investigations suggest that the early precipitation in the ocean is likely a result of the inner-core thermodynamical structure inconsistencies found in the DA analysis. The lower MSLP brought by the upper-level warm-core induces stronger secondary circulation in "DAT." The stronger updraft triggers more precipitation in the first several hours of model integration with a wetter middle-to-low-level inner-core moisture field in the analysis (not shown). "DAG" is not able to produce such an extraordinary amount of precipitation during the period (about 100 kg m^{-2} less than "DAT"; Figure 8b), but it can produce more rainfall over land to be more consistent with the observations than "DAT." The magnitude of the accumulated precipitation maximum in "DAB" is comparable with "DAG", but "DAB" follows the observations better by producing more rainfall around the Port Bay area and has fewer high precipitation tails over the ocean.

Figure 9 compares the 24-hr BT predictions between the model simulations and the GOES-13 observations. This period is about 15 hr after Harvey's landfall in Texas. Harvey was at the end of its rapid weakening and slowly moving northwestward. The satellite image shows a partially covered central dense overcast (CDO) cloud pattern that curves about halfway around the storm center. Unlike the observations, all model simulations have no transparent cloud void regions around the eye. "DAT" differs from the observations the most as its CDO region is much smaller than the observations and is positioned in the east-southeast of the storm center (Figure 9a vs. Figure 9d). As a result, the gap between the CDO region and the primary cloud band is much larger than observed. Both "DAG" and "DAB" predictions perform better than "DAT" by positioning the CDO region from the south to east part of the storm (Figures 9b and 9c). The larger size of the CDO region in those GBR assimilated experiments ("DAG" and "DAB") is also more consistent with the observations.

3.3. Issues of Temporal Discontinuity for TDR

As stated in the introduction, the temporal discontinuity of the TDR is one of the major limitations of this type of airborne-based observation. As a result, in a fully cycled hourly 3DEnVar DA system, experiment "DAT" can experience issues when the TDR observations are absent. For example, the abnormally high V_{max} analysis in Figure 2a. Hence, this subsection mainly discusses the issue of the particular DA cycle.
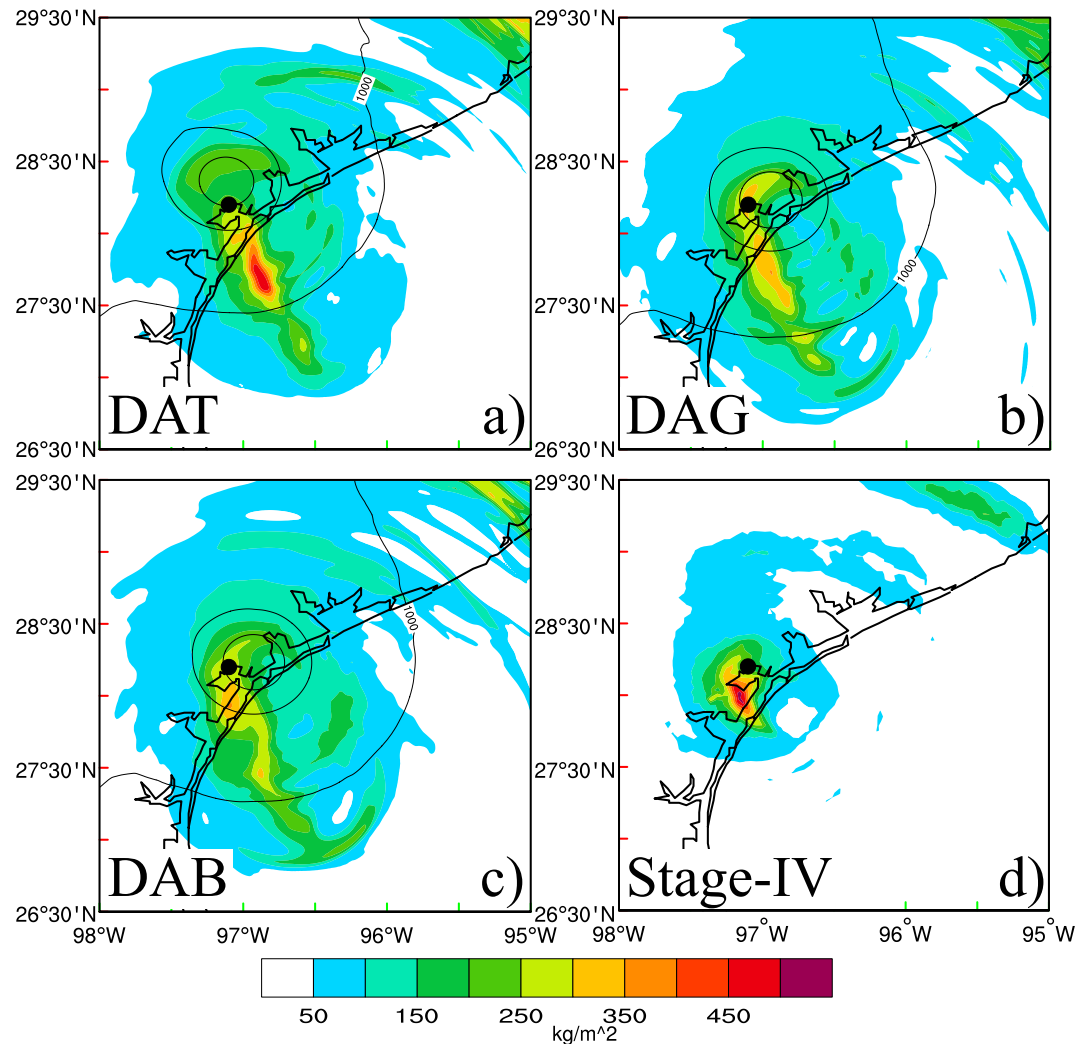
Figure 8. 12-hr accumulated precipitation prediction (shading) from (a) "DAT", (b) "DAG", and (c) "DAB" in verification with (d) the stage IV product valid at 0600 UTC, 26 August 2017. The background sea level pressure of each experiment is shown in black contour. The black dot in each figure is the best track location.

Figure 10 shows the horizontal wind analyses from all experiments at 500-m height in verification against the SMART radar observations at 0100 UTC on 26 August. This was about 4.4 hr after the final TDR observations were sampled, and Harvey was only about 20 km away from landfall. A few station observations over the coast (not shown) provided limited partial inner-core information to correct the storm location. At this time, the analysis produced by "DAT" presents an abnormally strong wind maximum of almost 10 ms^{-1} greater than the observations. Besides, the large area of the wind maxima (Figures 10d and 10a) also shows that "DAT" positions the wind maxima to the east of the storm instead of the north. The increments in Figure 10e show that such a large analysis of wind maximum anomaly is due to the large location error between the background and the observations. The background forecast in "DAT" at this time is about 19 km away from the interpolated best track, while the radius of maximum wind in the background storm is only about 20 km. Note, there is no vortex relocation performed to correct the location of the storms for each DA cycle. Therefore, when no inner-core observations are available to correct the storm location, such a large background location error is an accumulated result of several DA cycles lacking efficient inner-core corrections. Early studies like Chen and Snyder (2007) showed that when the location error is comparable with the vortex radius, the Gaussian assumptions during DA can be violated. Therefore, the produced analysis can be suboptimal. As a result, Figure 10e shows a dramatic adjustment to the background, which can produce a maximum increment greater than 60 ms^{-1}. In comparison with "DAT", with the continuous availability of the GBR observations, both "DAG" and "DAB" have less track
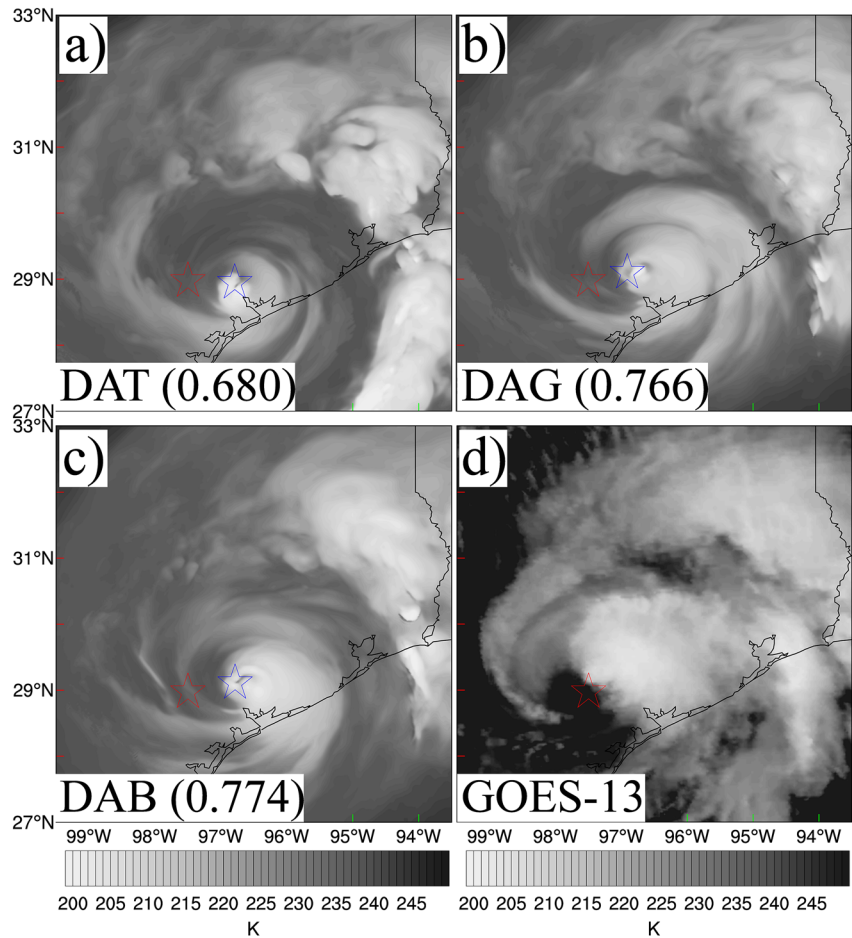
Figure 9. 24-hr brightness temperature prediction from (a) "DAT", (b) "DAG", and (c) "DAB" in verification with (d) the GOES-13 band four observations valid at 1800 UTC, 26 August 2017. The red and blue stars in each figure show the location of the storm from the best track and each experiment, respectively. The values in the bracket indicate the correlation coefficient of each forecast against the observations.

error in the background forecast initially. Therefore, the increment patterns in the corresponding Figures 10f and 10g are more reasonable than that from "DAT." Moreover, the analyses shown in Figures 10b and 10c are more consistent with the observations.

4. Summary and Discussion

This study investigates the impact of the individual assimilation of two types of inner-core Doppler radar radial velocity observations and their combined assimilation impacts on the analysis and predictions of Harvey (2017) during its landfalling period using a self-cycled GSI-based hourly 3DEnVar DA system for HWRF. A series of experiments were conducted to address the objective. Comprehensive verifications are first performed on the statistical results of all 31 hourly DA cycles during Harvey (2017). Then detailed diagnostics are performed on two particular cycles to understand the results found in the statistics.

Results from this study show that the assimilation of GBR observations has significant advantages over the assimilation of TDR in the all-cycle accumulated statistical results, including V_{max}, radial velocity structure, and precipitation prediction verifications, except for the MSLP prediction statistics. The inconsistency between V_{max} and MSLP prediction performance is found to be attributed to a systematic bias in the HWRF model during Hurricane Harvey. The assimilation of both observations shows a complementary effect and presents comparable or slightly better results than the assimilation of GBR alone.

A detailed diagnostic into one particular DA cycle shows that the full assimilation of either TDR or GBR observations can both produce a reasonable inner-core DA analysis and do well in the subsequent intensity and structural
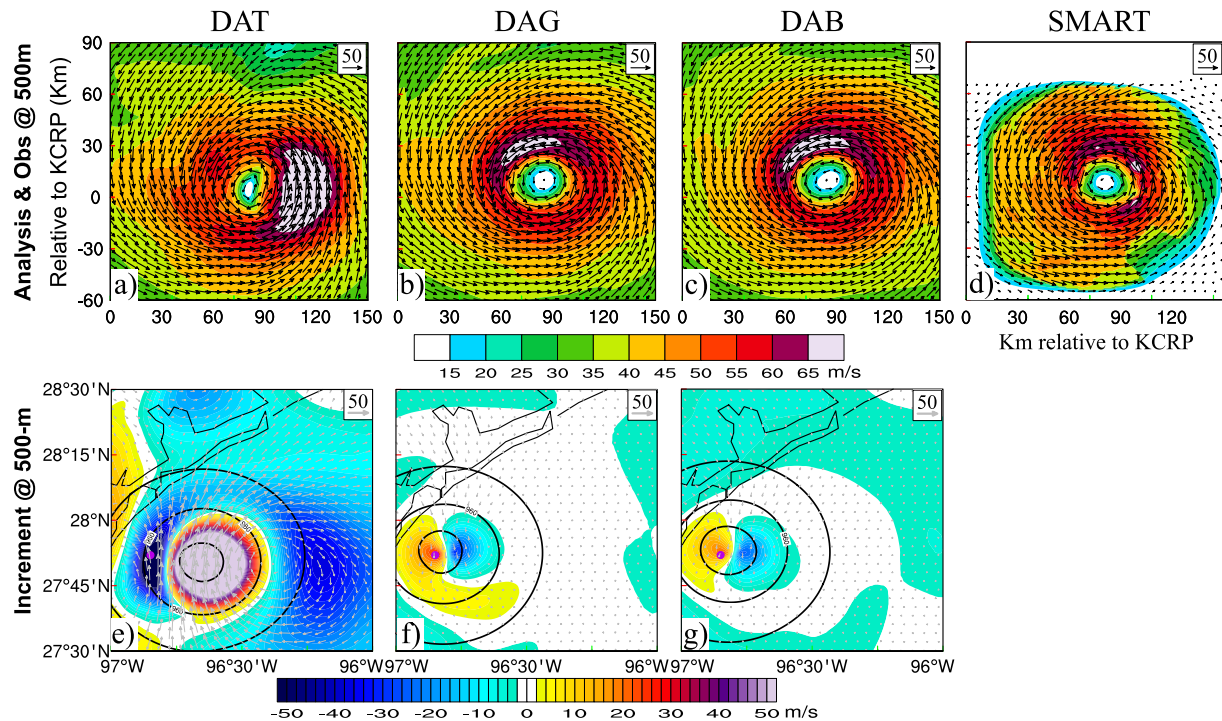
Figure 10. (a)-d) 500-m height horizontal wind (vector) and wind speed (shading) analyses and (e)-g) the corresponding wind increment (shading) and background sea level pressure (contour) for (a), (e) "DAT", (b), (f) "DAG", and (c), (g) "DAB" valid at 0100 UTC, 26 August 2017, in verification with (d) the SMART radar observations valid at 0102 UTC, 26 August 2017. The purple dot in (e)-g) indicates the best track storm location.

predictions. However, differences between the assimilation of the two data types also exist. Specifically, "DAT" tends to produce a stronger upper-level warm-core and therefore produces the much-lower-than-observed MSLP peak predictions during the intensification period, although it also produces better V_{max} predictions due to its better coverage in the lower portion of the inner-core region. Such results are slightly different from Green et al. (2022), who found limited dynamic improvements from TDR over GBR during the ERC evolutions. The differences could be attributed to the case difference between Matthew's ERC and Harvey's intensification. "DAG" outperforms "DAT" in the MSLP prediction but underpredicts the Vmax peak. The inconsistent Vmax and MSLP update suggests that the pure dynamic observations may cause inconsistent thermodynamical corrections from the cross-variable corrections, given the biased wind and pressure relationship in HWRF. Early studies like Lu and Wang (2020) showed that the issue is likely to be resolved if more corresponding inner-core thermodynamics observations are available during the DA. Additional prediction verifications indicate "DAG" produces more accurate dynamic, rainfall, and BT predictions than "DAT." Combining both observations shows the best performance in all aspects, suggesting the complementary effect of both observation types.

Additional investigations after TDR available cycles show that the temporal discontinuity of inner-core observations potentially harms a continuously cycled DA system for landfalling hurricanes, especially when no vortex relocation is performed. This result suggests that unless we have continuous inner-core sampling like the GBR observations to help regularly correct the storm locations and inner-core structures, a proper vortex relocation method is still necessary for a continuously cycled DA system. Such results are consistent with early findings from Lu, Wang, Tong, et al. (2017).

Overall, this study suggests that although the assimilation of the high-resolution inner-core observations from either TDR or GBR can improve our numerical analysis and prediction of hurricanes, which has been demonstrated in a lot of early studies, the two types of radial velocity observations can be complementary to each other. We do need both observations to achieve better predictions for landfalling hurricanes. Furthermore, additional inner-core thermodynamical observations are preferred add-ons for even better predictions. Also, consistent with findings from Green et al. (2022), this study indicates that the continuity of inner-core sampling can be necessary for a continuously cycled DA system that analyzes the rapid development phase of the hurricane. Given that this study is limited to hurricane Harvey, more studies with larger samples should be conducted in future studies.

Data Availability Statement

The operational datasets, including the control and ensemble analysis and forecast from the Global Forecast System (GFS, 2003; X. Wang et al., 2013) used in this study can be found at https://dtcenter.org/sites/default/files/community-code/hwrf/HWRF_input_data_sources.pdf. The level-II GBR observations (Crum et al., 1993) can be ordered from National Centers for Environmental Information (NCEI, 1995). The WDSS-II software (Lakshmanan et al., 2007) can be obtained from http://www.wdssii.org/download.shtml. The HWRF & GSI software (Biswas et al., 2018) can be downloaded from https://dtcenter.org/community-code/hurricane-wrf-hwrf/download.

References

Aksoy, A., Aberson, S. D., Vukicevic, T., Sellwood, K. J., Lorsolo, S., & Zhang, X. J. (2013). Assimilation of high-resolution tropical cyclone observations with an ensemble Kalman filter using NOAA/AOML/HRD's HEDAS), evaluation of the 2008-11 vortex-scale analyses. *Monthly Weather Review*, *141*(6), 1842–1865. https://doi.org/10.1175/Mwr-D-12-00194.1

Alford, A. A., Biggerstaff, M. I., & Carrie, G. D. (2019). Mobile ground-based SMART radar observations and wind retrievals during the landfall of Hurricane Harvey (2017). *Geoscience Data Journal*, *6*(2), 205–213. https://doi.org/10.1002/gdj3.82

Bao, J. W., Gopalakrishnan, S. G., Michelson, S. A., Marks, F. D., & Montgomery, M. T. (2012). Impact of physics representations in the HWRFX on simulated hurricane structure and pressure–wind relationships. *Monthly Weather Review*, *140*(10), 3278–3299. https://doi.org/10.1175/mwr-d-11-00332.1

Biswas, M. K., Bernardet, L., Abarca, S., Ginis, I., Grell, E., Kalina, E., & Zhang, Z. (2018). Hurricane weather research and forecasting (HWRF) model), 2018 scientific Documentation DEVELOPMENTAL TESTBED CENTER. Retrieved from https://dtcenter.org/HurrWRF/users/docs/scientific_documents/HWRFv4.0a_ScientificDoc.pdf

Blake, E. S., & Zelinsky, D. A. (2017). Hurricane Harvey. Retrieved from https://www.nhc.noaa.gov/data/tcr/AL092017_Harvey.pdf

Chen, Y., & Snyder, C. (2007). Assimilating vortex position with an ensemble Kalman filter. *Monthly Weather Review*, *135*(5), 1828–1845. https://doi.org/10.1175/MWR3351.1

Crum, T. D., & Alberty, R. L. (1993). The WSR-88D and the WSR-88D operational support facility. *Bulletin of the American Meteorological Society*, *74*(9), 1669–1687. https://doi.org/10.1175/1520-0477(1993)074<1669:TWATWO>2.0.CO;2

Davis, B., Wang, X., & Lu, X. (2021). A comparison of HWRF 6-hourly 4DEnVar and hourly 3DEnVar assimilation of inner core tail dopper radar observations for the prediction of hurricane Edouard (2014). *Atmosphere*, *12*(8), 942. https://doi.org/10.3390/atmos12080942

Dong, J., & Xue, M. (2013). Assimilation of radial velocity and reflectivity data from coastal WSR-88D radars using an ensemble Kalman filter for the analysis and forecast of landfalling hurricane Ike (2008). *Quarterly Journal of the Royal Meteorological Society*, *139*(671), 467–487. https://doi.org/10.1002/qj.1970

Feng, J., & Wang, X. (2019). Impact of assimilating upper-level Dropsonde observations collected during the TCI field campaign on the prediction of intensity and structure of hurricane Patricia (2015). *Monthly Weather Review*, *147*(8), 3069–3089. https://doi.org/10.1175/MWR-D-18-0305.1

Galarneau, T. J., & Zeng, X. (2020). The hurricane Harvey (2017) Texas rainstorm: Synoptic analysis and sensitivity to soil moisture. *Monthly Weather Review*, *148*(6), 2479–2502. https://doi.org/10.1175/MWR-D-19

GFS. (2003). Legacy NCEP NOAA Port Analysis and Forecasts [Dataset]. National Centers for Environmental Information. Retrieved from https://www.ncei.noaa.gov/products/weather-climate-models/global-forecast

Green, T., Wang, X., & Lu, X. (2022). Impact of assimilating ground-based and airborne radar observations for the analysis and prediction of the eyewall replacement cycle of hurricane Matthew (2016) using the HWRF hybrid 3DEnVar system. *Monthly Weather Review*, *150*(5), 1–40. https://doi.org/10.1175/mwr-d-21-0234.1

Jorgensen, D. P. (1984). Mesoscale and convective-scale characteristics of mature hurricanes. Part I), general observations by research aircraft. *Journal of the Atmospheric Sciences*, *41*(8), 1268–1285. https://doi.org/10.1175/1520-0469(1984)041<1268:macsco>2.0.co;2

Ko, M. C., Marks, F. D., Alaka, G. J., & Gopalakrishnan, S. G. (2020). Evaluation of hurricane Harvey (2017) rainfall in deterministic and probabilistic HWRF forecasts. *Atmosphere*, *11*(6), 666. https://doi.org/10.3390/atmos11060666

Lakshmanan, V., Smith, T., Stumpf, G., & Hondl, K. (2007). The warning decision support system–integrated information. *Weather and Forecasting*, *22*(3), 596–612. https://doi.org/10.1175/waf1009.1

Li, Y., Wang, X., & Xue, M. (2012). Assimilation of radar radial velocity data with the WRF hybrid ensemble-3DVAR system for the prediction of hurricane Ike (2008). *Monthly Weather Review*, *140*(11), 3507–3524. https://doi.org/10.1175/MWR-D-12-00043.1

Lu, X., Davis, B., & Wang, X. (2022). Improving the assimilation of enhanced atmospheric motion vectors for hurricane intensity predictions with HWRF. *Remote Sensing*, *14*(9), 2040. https://doi.org/10.3390/rs14092040

Lu, X., & Wang, X. (2019). Improving hurricane analyses and predictions with TCI, IFEX field campaign observations, and CIMSS AMVs using the advanced hybrid data assimilation system for HWRF. Part I), what is missing to capture the rapid intensification of hurricane Patricia (2015). *Monthly Weather Review*, *147*(4), 1351–1373. https://doi.org/10.1175/MWR-D-18-0202.1

Lu, X., & Wang, X. (2020). Improving hurricane analyses and predictions with TCI, IFEX field campaign observations, and CIMSS AMVs using the advanced hybrid data assimilation system for HWRF. Part II), observation impacts on the analysis and prediction of Patricia (2015). *Monthly Weather Review.*, *148*(4), 1407–1430. https://doi.org/10.1175/MWR-D-19-0075.1

Lu, X., Wang, X., Li, Y., Tong, M., & Ma, X. (2017). GSI-based ensemble-variational hybrid data assimilation for HWRF for hurricane initialization and prediction), the impact of various error covariances for airborne radar observation assimilation. *Quarterly Journal of the Royal Meteorological Society*, *143*(702), 223–239. https://doi.org/10.1002/qj.2914

Lu, X., Wang, X., Tong, M., & Tallapragada, V. (2017). GSI-based, continuously cycled, dual-resolution hybrid ensemble-variational data assimilation system for HWRF), System description and experiments with Edouard (2014). *Monthly Weather Review*, *145*(12), 4877–4898. https://doi.org/10.1175/MWR-D-17-0068.1

Marchok, T. (2002). How the NCEP tropical cyclone tracker works. In *Preprints, 25th Conf. On hurricanes and tropical Meteorology* (Vol. 13). Amer. Meteor. Soc. P1.

Marchok, T. (2021). Important factors in the tracking of tropical cyclones in operational models. *Journal of Applied Meteorology and Climatology*, *60*(9), 1265–1284. https://doi.org/10.1175/jamc-d-20-0175.1

Marks, F. G., & Houze, R. A. (1984). Airborne Doppler radar observations in hurricane Debby. *Bulletin of the American Meteorological Society*, *65*(6), 569–582. https://doi.org/10.1175/1520-0477(1984)065<0569:adroih>2.0.co;2

NCEI. (1995). NEXRAD Data Inventory [Dataset]. National Centers for Environmental Information. Retrieved from https://www.ncei.noaa.gov/has/HAS.FileAppRouter?datasetname=6500&subqueryby=STATION&applname=&outdest=FILE

Rogers, R., Aberson, S., Aksoy, A., Annane, B., Black, M., Cione, J., et al. (2012). NOAA's hurricane intensity forecasting experiment (IFEX)), A progress report. *Bulletin of the American Meteorological Society*, 130109100058001. https://doi.org/10.1175/BAMS-D-12-00089

Shen, F., Min, J., & Xu, D. (2016). Assimilation of radar radial velocity data with the WRF Hybrid ETKF-3DVAR system for the prediction of Hurricane Ike (2008). *Atmospheric Research*, *169*, 127–138. https://doi.org/10.1016/j.atmosres.2015.09.019

Sippel, J., Zhang, Z., Bi, L., & Mehra, A. (2021). Recent Advances in operational HWRF data assimilation. 34th conference on hurricanes and tropical Meteorology. May 10 – May 14, virtual online.

Snyder, C., & Zhang, F. (2003). Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Monthly Weather Review*, *131*(8), 1663–1677. https://doi.org/10.1175//2555.1

Stensrud, D. J., Xue, M., Wicker, L. J., Kelleher, K. E., Foster, M. P., Schaefer, J. T., et al. (2009). Convective-scale warn-on-forecast system: A vision for 2020. *Bulletin of the American Meteorological Society*, *90*(10), 1487–1499. https://doi.org/10.1175/2009BAMS2795.1

Sun, J. (2006). Convective-scale assimilation of radar data: Progress and challenges. *Quarterly Journal of the Royal Meteorological Society*, *131*(613), 3439–3463. https://doi.org/10.1256/qj.05.149

Tallapragada, V., Bernardet, L., Biswas, M. K., Gopalkrishnan, S., Kyon, Y., Liu, Q., et al. (2013). Hurricane weather research and forecasting (HWRF) model), 2013 scientific Documentation authors. (in alphabetical order by last name).

Wang, P., Li, J., Li, Z., Lim, A. H. N., Li, J., & Goldberg, M. D. (2019). Impacts of observation errors on hurricane forecasts when assimilating hyperspectral infrared sounder radiances in partially cloudy skies. *Journal of Geophysical Research: Atmospheres*, *124*(20), 10802–10813. https://doi.org/10.1029/2019JD031029

Wang, X., Parrish, D., Kleist, D., & Whitaker, J. (2013). GSI 3DVar-based ensemble–variational hybrid data assimilation for NCEP global forecast system: Single-resolution experiments. *Monthly Weather Review*, *141*(11), 4098–4117. https://doi.org/10.1175/MWR-D-12-00141.1

Wang, Y., & Pu, Z. (2021). Assimilation of radial velocity from coastal NEXRAD into HWRF for improved forecasts of landfalling hurricanes. *Weather and Forecasting*, *36*(2), 587–599. https://doi.org/10.1175/waf-d-20-0163.1

Weng, Y., & Zhang, F. (2012). Assimilating airborne Doppler radar observations with an ensemble Kalman filter for convection-permitting hurricane initialization and prediction), Katrina (2005). *Monthly Weather Review*, *140*(3), 841–859. https://doi.org/10.1175/2011MWR3602.1

Xiao, Q., Kuo, Y. H., Sun, J., Lee, W. C., Barker, D. M., & Lim, E. (2007). An approach of radar reflectivity data assimilation and its assessment with the Inland QPE of Typhoon Rusa (2002) at landfall. *Journal of Applied Meteorology and Climatology*, *46*(1), 14–22. https://doi.org/10.1175/JAM2439.1

Zhang, D. L., & Chen, H. (2012). Importance of the upper-level warm core in the rapid intensification of a tropical cyclone. *Geophysical Research Letters*, *39*(2). https://doi.org/10.1029/2011gl050578

Zhang, S., & Pu, Z. (2019). Numerical simulation of rapid weakening of hurricane Joaquin with assimilation of high-definition sounding system dropsondes during the tropical cyclone intensity experiment), comparison of three-and four-dimensional ensemble-variational data assimilation. *Weather and Forecasting*, *34*(3), 521–538. https://doi.org/10.1175/WAF-D-18-0151.1

Zhang, F., Weng, Y., Gamache, J. F., & Marks, F. D. (2011). Performance of convection-permitting hurricane initialization and prediction during 2008-2010 with ensemble data assimilation of inner-core airborne Doppler radar observations. *Geophysical Research Letters*, *38*(15), 2–7. https://doi.org/10.1029/2011GL048469

Zhang, F., Weng, Y., Sippel, J., Meng, Z., & Bishop, C. H. (2009). Cloud-resolving hurricane initialization and prediction through assimilation of Doppler radar observations with an ensemble Kalman filter. *Monthly Weather Review*, *137*(7), 2105–2125. https://doi.org/10.1175/2009mwr2645.1

Zhao, Q., & Jin, Y. (2008). High-resolution radar data assimilation for hurricane Isabel (2003) at landfall. *Bulletin of the American Meteorological Society*, *89*(9), 1355–1372. https://doi.org/10.1175/2008BAMS2562.I

Zhao, K., & Xue, M. (2009). Assimilation of coastal Doppler radar data with the ARPS 3DVAR and cloud analysis for the prediction of Hurricane Ike (2008). https://doi.org/10.1029/2009GL038658

Zhu, L., Wan, Q., Shen, X., Meng, Z., Zhang, F., Weng, Y., et al. (2016). Prediction and predictability of high-impact Western Pacific landfalling Tropical Cyclone Vicente (2012) through convection-permitting ensemble assimilation of Doppler radar velocity. *Monthly Weather Review*, *144*(1), 21–43. https://doi.org/10.1175/MWR-D-14-00403.1