Developmental Testbed Center (DTC) Project Report

June 2015

Table of Contents

1. Introduction

Hurricane Irene was the first hurricane of the 2011 Atlantic hurricane season and ranked as the seventh-costliest hurricane in United States history with destructive damages on much of the Caribbean and East Coast of the United States during late August 2011. Figure 1 shows the Best Track of Hurricane Irene during its lifetime with categories marked by different colors (Avila and Cangialosi 2011).

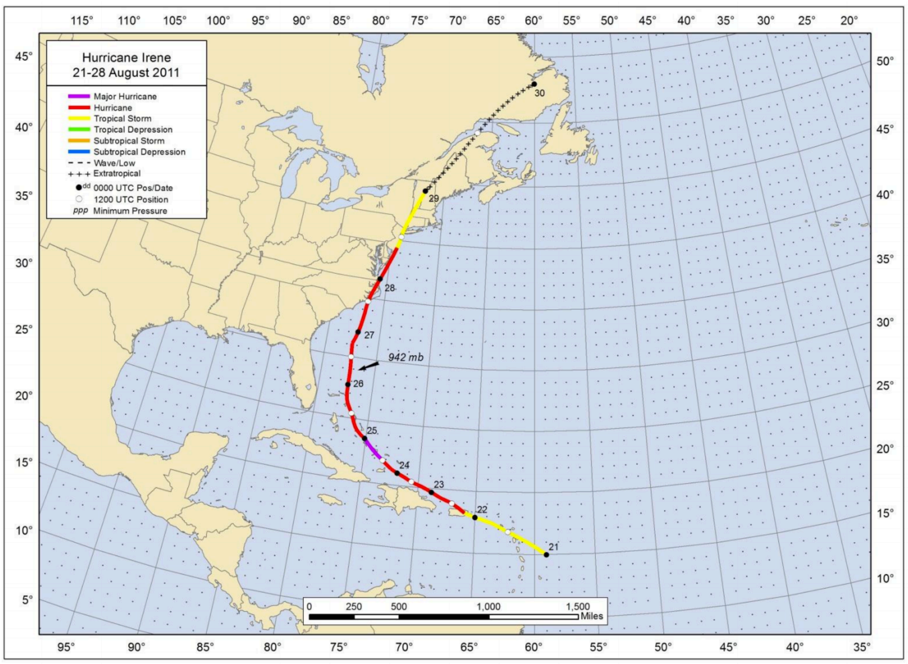

Figure 1. Hurricane Irene's Best Track.

Zhou et al. (2015) found that there is an intensity "spin-down" issue occurring in forecast of Irene's intensity using the Hurricane Weather Research and Forecasting model (HWRF) 2014 configuration. That is, the bias of Maximum surface wind speed is positive at the initial time and goes down to be negative shortly after the first a few hours. The intensity jump is called intensity "spin-down". The HWRF model uses the NCEP operational Gridpoint Statistical Interpolation (GSI) hybrid data assimilation and the background error covariances are calculated from global ensemble forecast instead of regional ensemble forecast that is expected to match the high-resolution HWRF model and therefore to provide better background error covariances in GSI Hybrid data assimilation. The intensity forecast and "spin-down" issue could be related to a couple of aspects. This study aims to use regional ensemble EnKF and forecast to look at the "spin-down" issue as well as the forecast performance.

2. Experimental setup

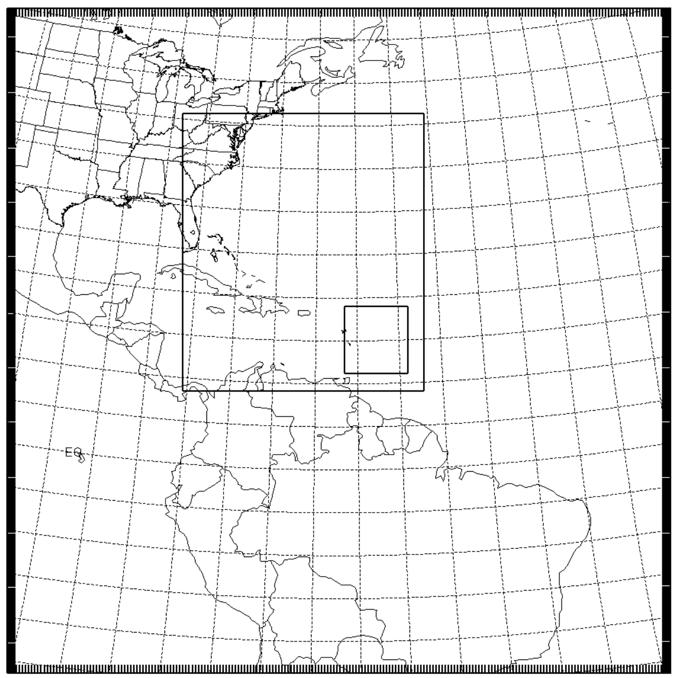

Figure 2. Domain setup. The outermost 27-km and the inner 9-km domains are fixed during cycles. The innermost 3-km domain moves with storm. Here shows the initial position of the 3-km domain at 1800 UTC 20 August 2011.

The Weather Research and Forecasting model version 3 with Nonhydrostatic Mesoscale Model (NMM) core is employed in advancing the hybrid and ensemble analyses as forecast model. Three nested domains are used at horizontal grid spacings of 0.18 (27-km), 0.06 (9-km), and 0.02 (3-km) degree with 216x432, 235x543, and 187x392 grid points respectively. As shown in Figure 2, the outermost 27-km domain and the inner 9-km domain are fixed throughout the simulation period. The 27-km domain covers an 80^ox80^o region, the same as that of the HWRF 2014 operational 27-km domain. The 9-km domain is also fixed to facilitate the ensemble run during data assimilation (DA) cycles, covering the entire storm track path of interest. The 3-km domain moves with storm, with a similar coverage as the HWRF 2014 analysis d03 domain, using the "vortex-following" method built in the WRF model. All domains have 61 vertical levels with model top of 2-hPa. Physical schemes are exactly the same as the HWRF 2014 operational configuration.

DA cycles were conducted between 1800 UTC 20 August, 6 hour prior to the time when Irene was identified as a topical storm, and 0000 UTC 27 August 2011, when Irene was moving close to the United States coast. The NCEP operational Gridpoint Statistical Interpolation (GSI) hybrid DA is used in this study, and the flow-dependent background error covariances come from an Ensemble Square Root Kalman filter (EnKF hereafter; Whitaker and Hamill 2002). The global ensemble forecast in spectral space was pre-processed to GRIB format and then an 80-member ensemble was generated at initial time 1800 UTC 20 August 2011.

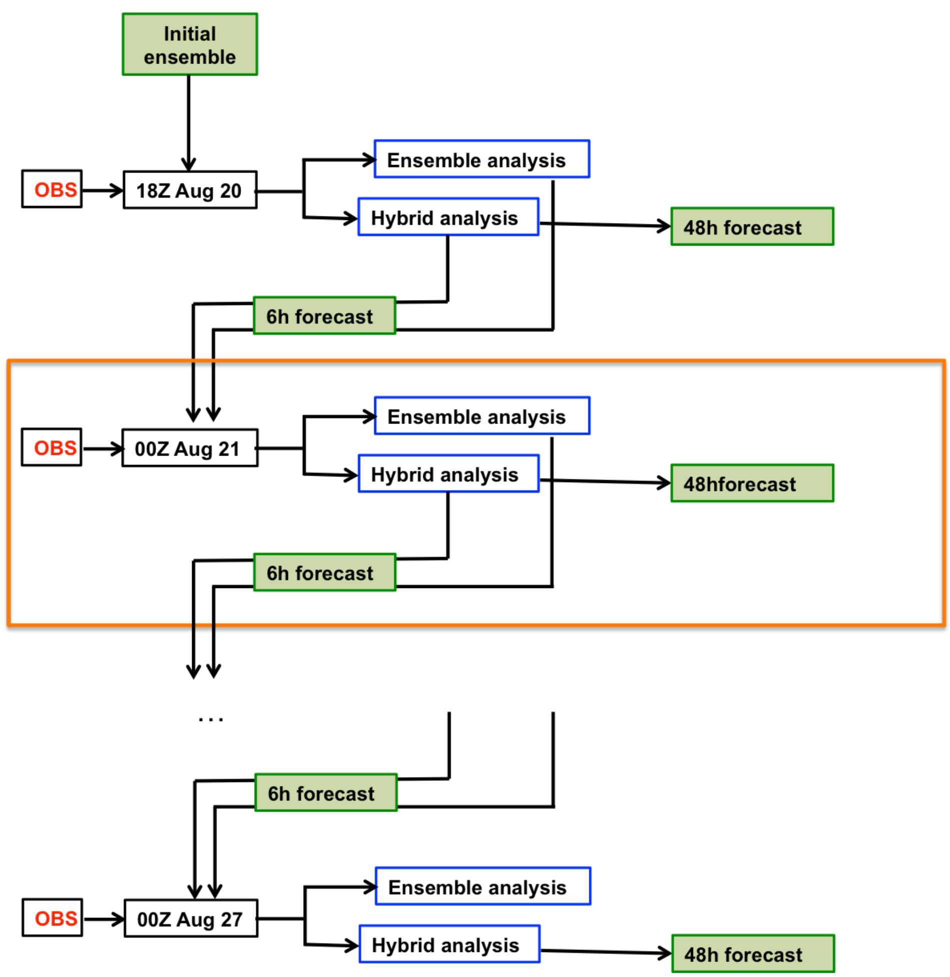

Figure 3. Schematic illustration of the data assimilation procedure.

Parallel DAs were conducted for both the GSI Hybrid and EnKF every 6 hour since 1800 UTC 20 August 2011. GSI Hybrid DA was conducted on all the 3 domains, while EnKF was conducted only on the 27-km and 9-km domains. Therefore the output includes the Hybrid analyses of the 27-km, 9-km and 3-km domains, and the analyses of ensemble members and mean of the 27-km and 9-km domains. The Hybrid 3-km DA uses the 9-km ensemble forecasts to calculate the flow-dependent background error covariances through the "dual resolution" option in GSI since 3-km ensemble forecast was not running limited by the computational resources. After getting the analyses on each DA cycle, both the GSI Hybrid and EnKF analyses were advanced 6h forward to next cycle time using the WRF model. Meantime, 48h deterministic forecast initialized from the GSI Hybrid analysis was also launched. This comprises a DA cycle as shown in Figure 3. 25 forecast samples were collected during the entire cycle time. Statistical analysis in Section 4 will base on these samples.

Table 1. Experimental configurations

EXP/configuration	Re-centering	Assi. TDR	Assi. TC vital
OnewaynTDR_hyb	NO	NO	NO
Oneway_hyb	NO	YES	NO
Twoway_hyb	YES	YES	NO
Twoway_hyb_tcv	YES	YES	YES

Table 1 lists all experiments conducted in this study. "Twoway" here refers to the re-centering procedure applied on EnKF analysis. In a "Twoway" experiment, all ensemble members are re-centered to the Hybrid analysis. To be more specifically, there are 3 steps for re-centering: 1) taking the mean of the ensemble members, 2) calculating the ensemble perturbations by subtracting the ensemble mean from each member, 3) adding the ensemble perturbation to the Hybrid analysis to generate new analysis for each member. As shown in Table 1, experiments "Twoway_hyb" and "Twoway_hyb_tcv" are applied the re-centering while experiments "Oneway_hyb" and "OnewaynTDR_hyb" are

not. All the experiments assimilate the Tailed Doppler Radar (TDR) data expect for “OnewanTDR_hyb”. In addition, “Twoway_hyb_tcv” is the only experiment assimilating TC vital data in this study.

Figure 4 gives an example of the observational coverage in the 3-km domain, showing the coverage of both the conventional data and the TDR data at 0000 UTC August 24 2011.

3. Results

Since the initial ensemble was generated from global ensemble forecasts, ensemble spreads during DA cycles are first examined to see if the ensembles spreads are reasonable for the regional run. Figure 5 shows the time series of both the total spread and the Root Mean Square Error (RMSE) of variables u, v and t for the 27-km and 9-km respectively. It can be seen that the ensemble spreads for all variables are stable and comparable to the RMSEs in both domains, indicating a well-behaving ensemble during the entire cycle period.

Track forecasts from all the experiments are verified against Best Track from the National Hurricane Center (NHC). It is found that there is no obvious difference for track forecasts among all the experiments. Figure 6 shows all the track forecasts of the experiment “Twoway_hyb”, which initialized 4 times daily between 1800 UTC 20 August and 0000 UTC 27 August. All the track forecasts are in good agreement with the Best Track and thus are not the major concern of this study. Figure 7 shows an example how the ensemble forecasts look like in a 6-h cycle by plotting the 6-h ensemble forecast together with the Best Track and the GSI Hybrid 48-h forecast.

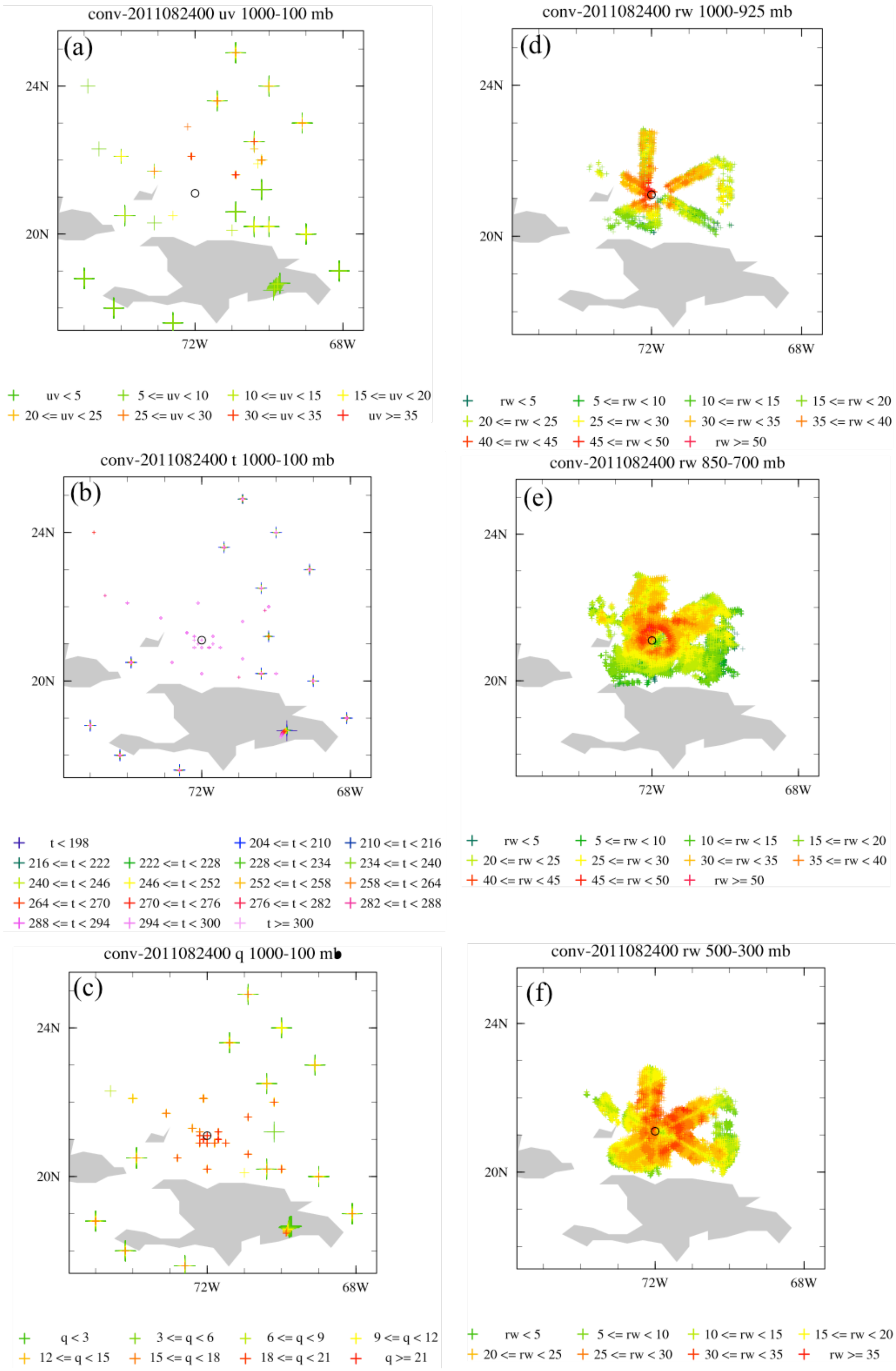
conv-2011082400 uv 1000-100 mb
(a)
24N
20N
72W
68W
uv < 5
5 <= uv < 10
10 <= uv < 15
15 <= uv < 20
20 <= uv < 25
25 <= uv < 30
30 <= uv < 35
uv >= 35
conv-2011082400 rw 1000-925 mb
(d)
24N
20N
72W
68W
rw < 5
5 <= rw < 10
10 <= rw < 15
15 <= rw < 20
20 <= rw < 25
25 <= rw < 30
30 <= rw < 35
35 <= rw < 40
40 <= rw < 45
45 <= rw < 50
rw >= 50
conv-2011082400 t 1000-100 mb
(b)
24N
20N
72W
68W
t < 198
204 <= t < 210
210 <= t < 216
216 <= t < 222
222 <= t < 228
228 <= t < 234
234 <= t < 240
240 <= t < 246
246 <= t < 252
252 <= t < 258
258 <= t < 264
264 <= t < 270
270 <= t < 276
276 <= t < 282
282 <= t < 288
288 <= t < 294
294 <= t < 300
t >= 300
conv-2011082400 rw 850-700 mb
(e)
24N
20N
72W
68W
rw < 5
5 <= rw < 10
10 <= rw < 15
15 <= rw < 20
20 <= rw < 25
25 <= rw < 30
30 <= rw < 35
35 <= rw < 40
40 <= rw < 45
45 <= rw < 50
rw >= 50
conv-2011082400 q 1000-100 mb
(c)
24N
20N
72W
68W
q < 3
3 <= q < 6
6 <= q < 9
9 <= q < 12
12 <= q < 15
15 <= q < 18
18 <= q < 21
q >= 21
conv-2011082400 rw 500-300 mb
(f)
24N
20N
72W
68W
rw < 5
5 <= rw < 10
10 <= rw < 15
15 <= rw < 20
20 <= rw < 25
25 <= rw < 30
30 <= rw < 35
rw >= 35

Figure 4. Conventional observation coverage for (a) u and v, (b) t, (c) q, and TDR data coverage between (d) 1000-925 hPa, (e) 850-700 hPa, and (f) 500-300 hPa, at 0000 UTC 24 August 2011 in the 3-km domain.

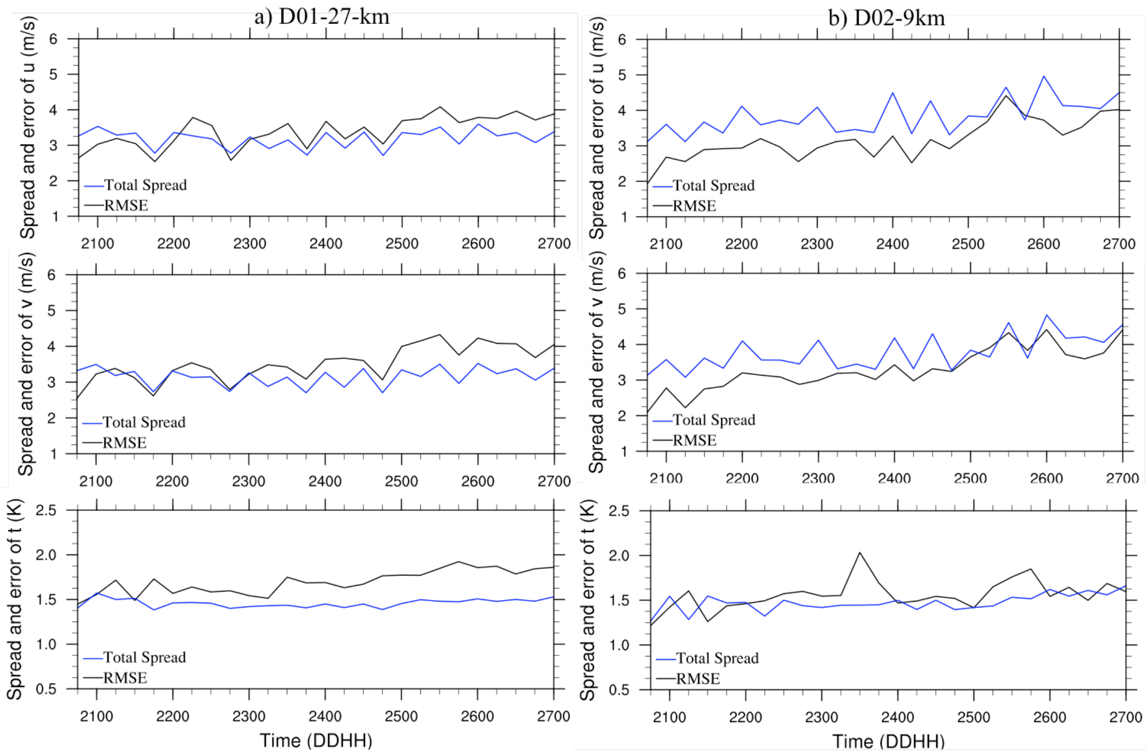

Figure 5. Time series of total spread and RMSE for variables u, v and temperature of a) the 27-km domain, and b) the 9-km domain.

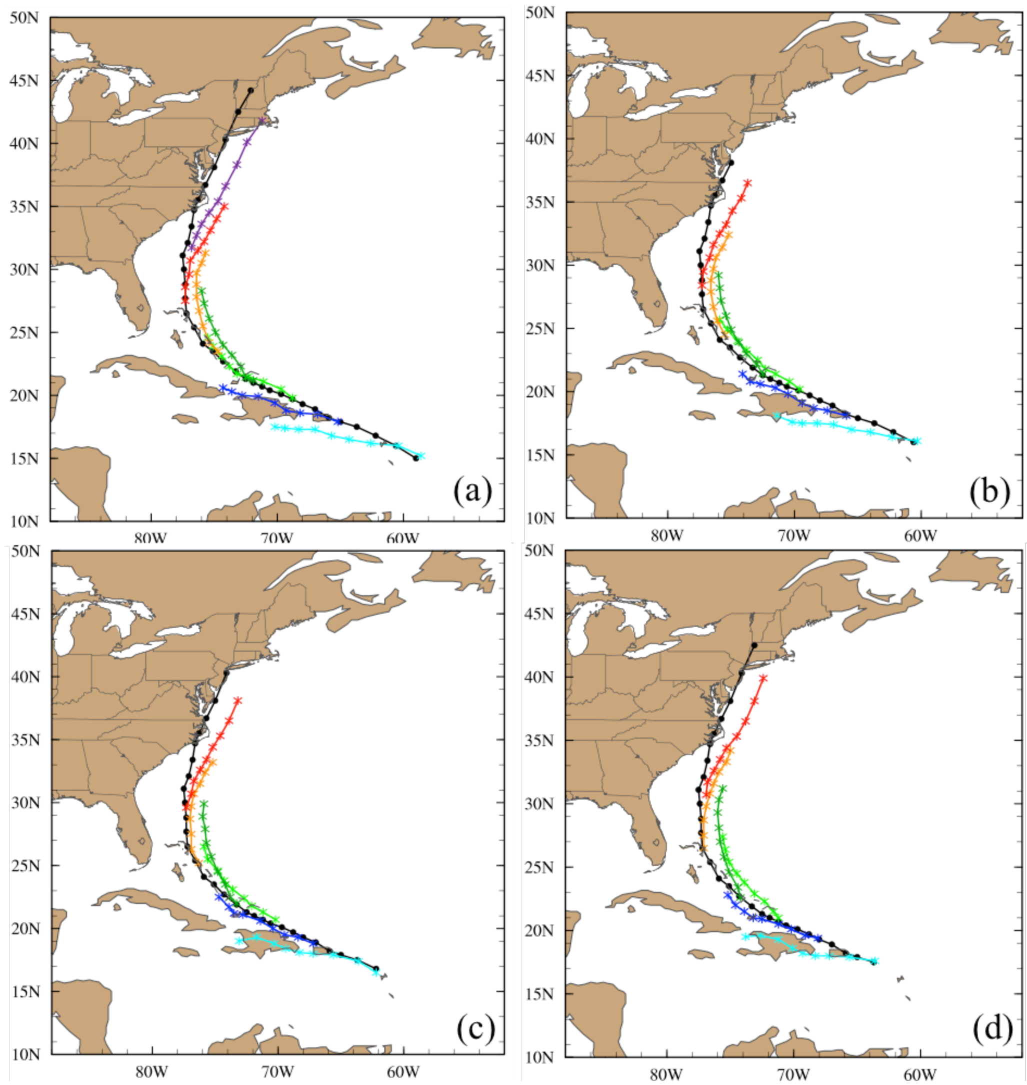

Figure 6. Track forecasts initialized from "Twoway_hyb" analyses at (a) 0000 UTC, (b) 0600 UTC, (c) 1200 UTC, and (d) 1800 UTC between 1800 UTC 20 August and 0000 UTC 27 August 2011. The Best Track is plotted in black and Hybrid forecasts initialized from different date are presented in color curves.

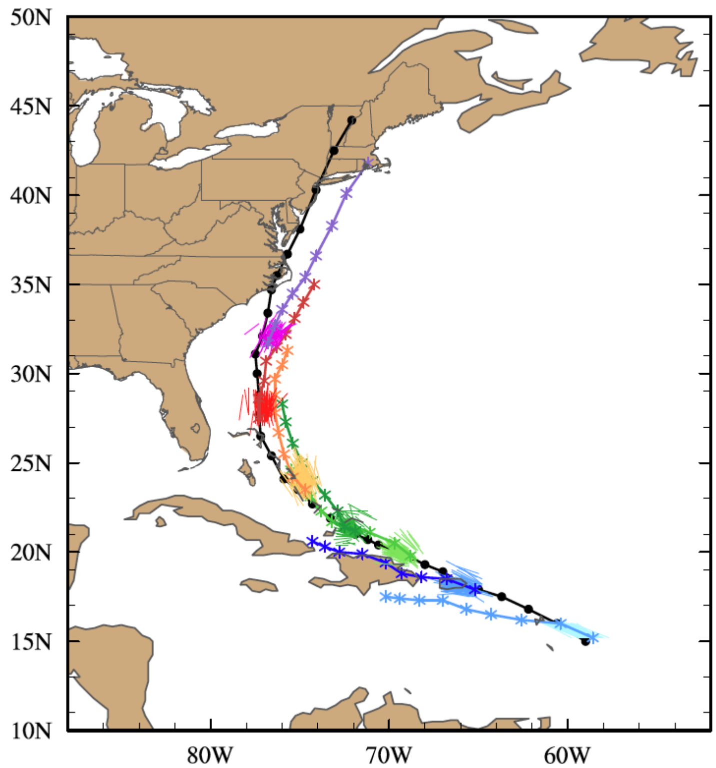

Figure 7. An example showing how the ensemble forecast looks like in a 6-h cycle. Best Track is plotted in black and Hybrid forecasts are plotted in thicker dark colors. The grouped light color curves are ensemble track forecasts.

Results are analyzed statistically based on all forecast samples from all the experiments. In addition, results from two more experiments, "HDTC" and "HNVI", are also displayed in this section for comparison. "HDTC" is the experiment using the HWRF 2014 trunk code and "HNVI" is run similarly to "HDTC" but skipping the vortex initialization step. Both "HDTC" and "HNVI" use global ensemble forecast for background error covariance calculation in GSI Hybrid DA, and are cold-started every 6-h expect "HDTC" uses vortex initialization by which the inner storm could be cycled from previous 6-h forecast.

Figure 8 shows the aggregated errors of hurricane intensity of the experiments "Oneway_hyb", "Twoway_hyb", "HDTC", and "HNVI" averaging over all the forecast samples as a function of forecast lead time. As mentioned earlier, the HWRF 2014 operational configuration run "HDTC" shows the intensity "spin-down" issue with positive Maximum Surface wind speed (MSW) bias at initial time and negative bias a few hours later (Figure 8a). The positive (stronger) MSW bias is mainly caused by the vortex initialization procedure in the HWRF system that makes a "strong" storm. "HNVI", which is run similarly as the "HDTC" expect for the vortex initialization step, shows a -15 knots bias at initial time. Both "Oneway_hyb" and "Twoway_hyb" obtain lower MWS bias than "HNVI" does throughout the 48-h lead time and than "HDTC" does after the first 12-h lead time. The aggregated biases of the Minimum Sea Level Pressure (MSLP) among all the four experiments are consistent as that of the MSW (Figure 8b). "HNVI" has the largest MSLP bias throughout the 48-h. The MSLP bias of "HDTC", again, is nearly zero at initial time due to the vortex initialization, and exceeds "Oneway_hyb" after 12-h. Clearly, "Oneway_hyb" performs better than "Twoway_hyb" in terms of MSLP forecast. It should be remembered that the only difference between these two experiments is that "Twoway_hyb" uses Hybrid analysis to re-center ensemble analysis but "Oneway_hyb" does not. To answer the question why "Oneway_hyb" produces smaller MSLP bias, a series of diagnoses are made for investigation in the following paragraphs.

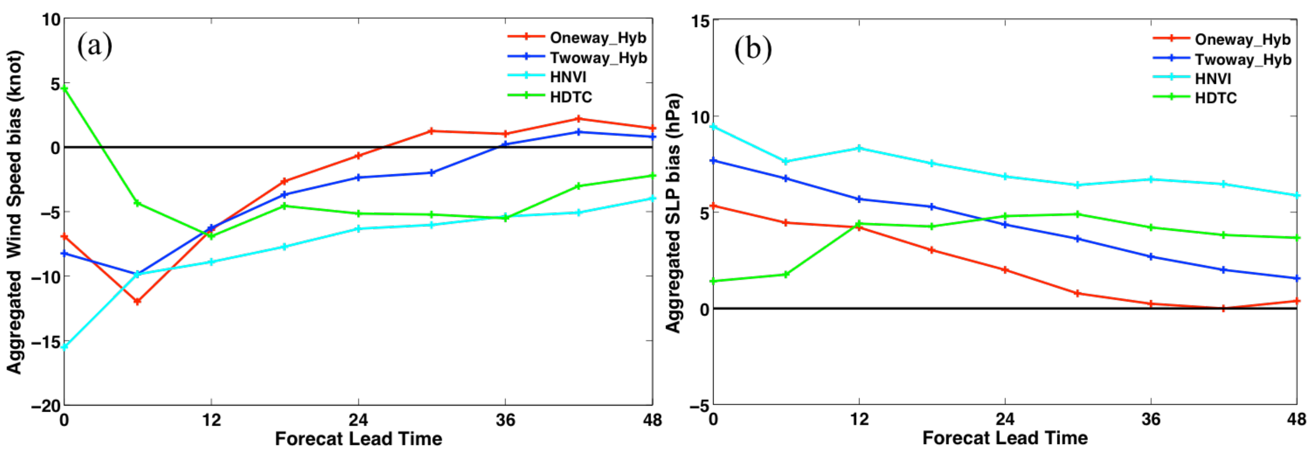

Figure 8. Aggregated bias of (a) Maximum surface wind speed, and (b) Minimum sea level pressure averaging over all the forecasts as a function of forecast lead time.

Figures 9 and 10 compare the fit-to-obs RMSEs and biases of analysis (O-A) and background (O-B) averaging over the entire cycle period as a function of pressure level and over the vertical column as a function of time. It can be seen that the RMSE reduction in analysis relative to background is evident due to the impact of assimilated data. The results, however, are mixed, and it is hard to tell which experiment is better in terms of the RMSEs and biases. Table 2 lists the amounts of assimilated "uv" observations at each analysis time in both experiments. It does show that "Oneway_hyb" assimilates more observations than "Twoway_hyb" does. The difference, however, is not significantly large to cause the MSLP bias difference between those two experiments. The assimilated observation amount and the corresponding fit-to-obs difference are not the major reason.

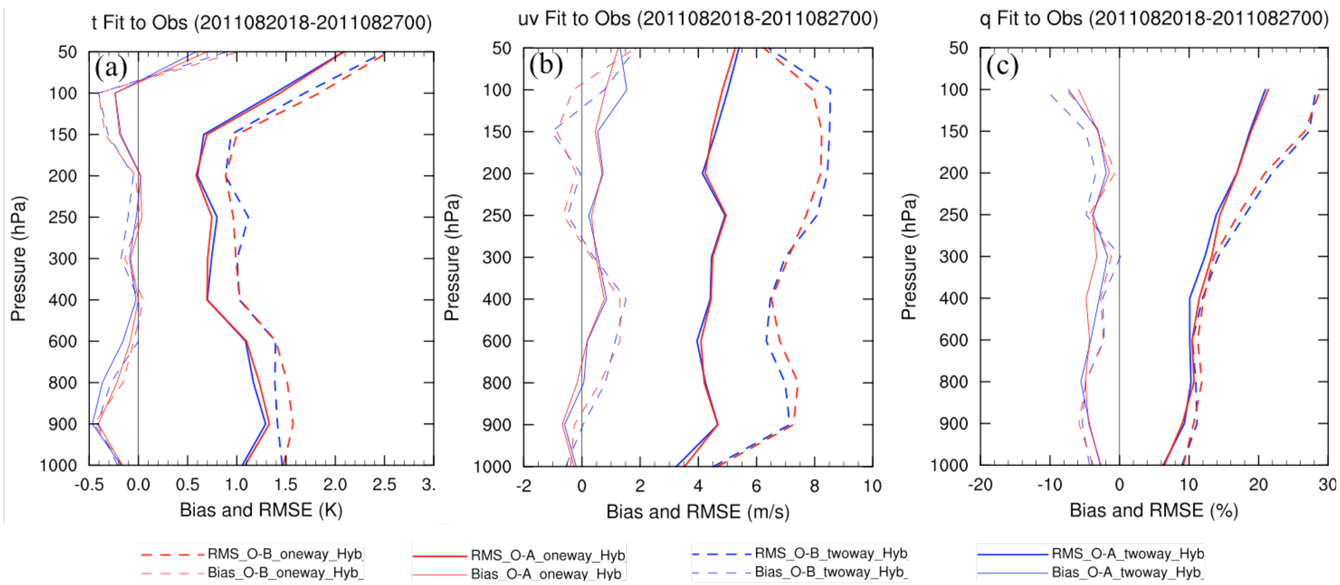

Figure 9. Fit to observation (fit-to-obs) RMSE and bias of (a) t, (b) uv, and (c) q, averaging over the entire period as a function of pressure level.

Aggregated MSW biases are calculated separately by averaging all the forecasts with TDR data assimilated and over forecasts with no TDR available at analysis time (Figure 11). All the experiments share similar MSW bias feature and trend for each type of forecast, showing increasing storm intensity (negative to positive MSW bias) for TDR forecasts and stable negative bias for nTDR forecasts with forecast lead time. The bias of "HDTC_TDR" starts from positive at forecast hour zero and goes down to be negative a few hours later and then becomes positive again. Meanwhile, The biases of "HNVI", "Oneway_hyb_TDR", and "Twoway_hyb_TDR" are all negative at the initial time due to lack of vortex initialization that is applied only in "HDTC". The biases, however, become positive quickly with time, and grow fastest in “Twoway_hyb_TDR”. For the nTDR forecasts, both “Oneway_hyb_nTDR” and “Twoway_hyb_nTDR” have smaller biases than “HNVI_nTDR” all the time, and “Oneway_hyb_nTDR” obtains smaller bias than “HDTC_nTDR” after the first 12h.

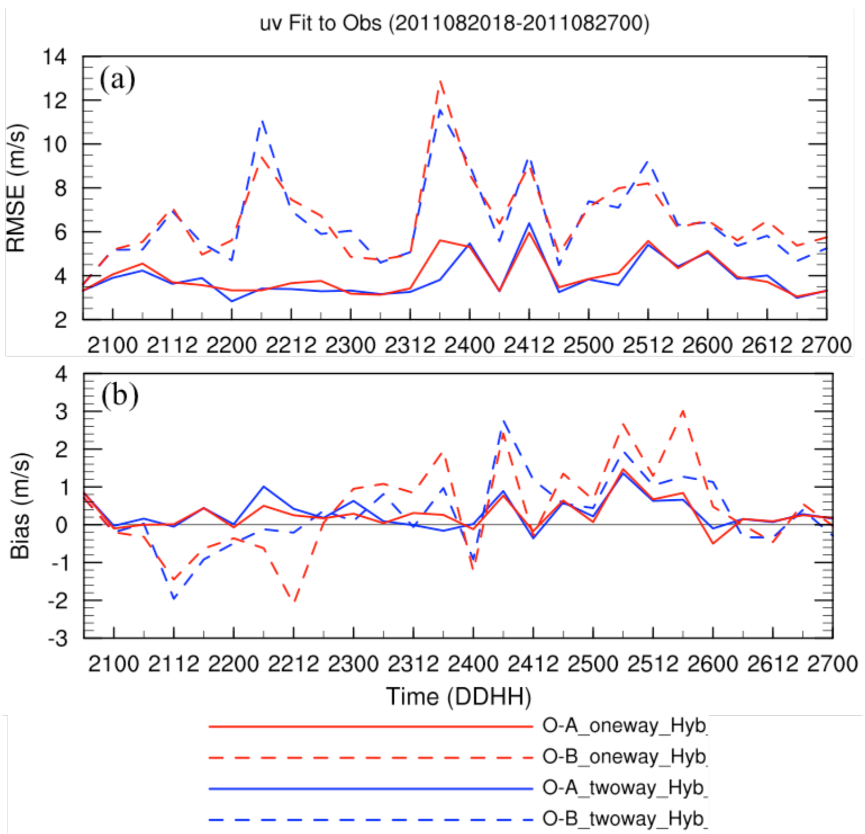

Figure 10. Fit to observation (fit-to-obs) RMSE and bias of (a) t, (b) uv, and (c) q,

Table 2. Assimilated observation amount during cycles.

	2018	2100	2106	2112	2118	2200	2206	2212	2218	2300	2306	2312	2318	2400	2406	2412	2418	2500	2506	2512	2518	2600	2606	2612	2618	2700
Twoway_hyb O-B	5	32	6	100	192	188	58	162	234	311	189	143	20	510	96	358	15	412	86	483	535	1061	390	793	578	61
Twoway_hyb O-A	5	32	6	100	192	189	86	162	234	311	189	143	20	517	96	363	15	415	86	484	535	1066	391	793	578	61
Oneway_hyb O-B	5	32	6	100	192	215	60	162	232	309	188	192	20	517	96	388	14	384	86	477	534	1092	396	796	582	61
Oneway_hyb O-A	5	32	6	100	192	220	86	162	234	309	188	193	20	517	96	389	14	384	86	486	535	1096	397	798	584	62

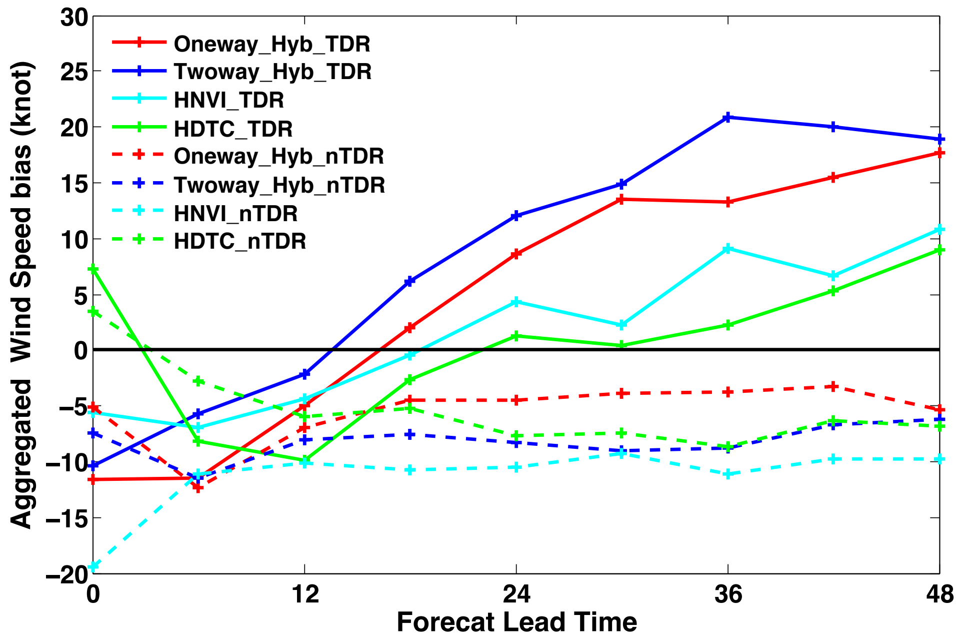

Figure 11. Aggregated bias of Maximum surface wind speed averaging over forecasts with TDR data assimilated (TDR, solid curves) and forecasts without TDR data assimilated (nTDR, dashed curves) at the analysis time as a function of forecast lead time. TDR data are avaibable at 0000 UTC 24, 1200 UTC 24, 1200 UTC 25, 1800 UTC 25, 0000 UTC 26, 1200 UTC 26, and 0000 UTC 27 August 2011. There are 18 nTDR/7 TDR forecasts used in calculation.

It should be mentioned that the results of nTDR cases are more meaningful because there are much more samples for nTDR forecast, and "Oneway_hyb_nTDR" shows the smallest biases at most of times. It is, however, still worthy investigating the impact of TDR data to see if TDR data plays a positive or negative role. The experiment "OnewaynTDR_hyb", which does not assimilate TDR data, was therefore conducted. Aggregated MSW biases of "OnewaynTDR_hyb" are compared with "Oneway_hyb" as well as "HDTC" and "HNVI" in Figure 12. Overall, the exclusion of TDR data in "OnewaynTDR_hyb" produces weaker storms indicated by bigger negative biases throughout the period (Figure 12a). Although no TDR data is assimilated in "OnewaynTDR_hyb", aggregated biases are still separately examined over forecasts with TDR assimilated and over forecasts without TDR data assimilated in "Oneway_hyb" (Figure 12b). For the forecasts initialized from analyses without TDR data assimilated in "Oneway_hyb", "OnewaynTDR_hyb" produces larger biases than "Oneway_hyb", similarly to that of the overall bias. For the forecasts initialized from analyses with TDR data assimilated in "Oneway_hyb", the negative bias in "OnewaynTDR_hyb" is larger in the first 24-h and grows with time to be positive. Although the bias of "Oneway_hyb" is positively lager than that of "OnewaynTDR_hyb" at 48-h, the increase rates are almost the same in both experiments. In a word, the assimilation of TDR data makes the storm intensification faster, but does not significantly influence the forecasts initialized from analyses with no TDR data assimilated that account for the majority of the forecast samples in statistical calculation.

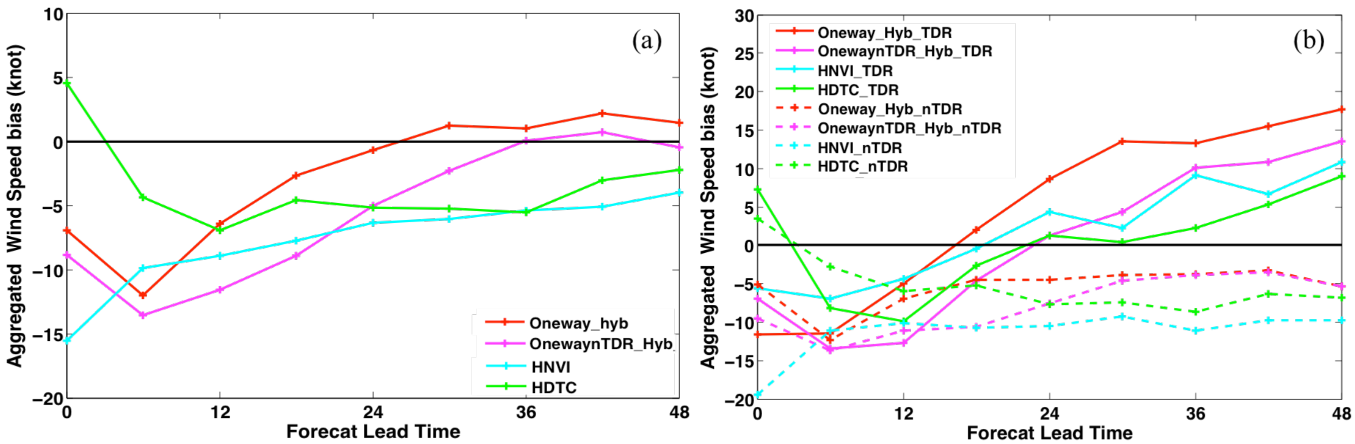

Figure 12. Aggregated bias of Maximum surface wind speed (a) averaging over all the forecasts, and (b) averaging over forecasts with TDR data assimilated in "Oneway_hyb" (TDR, solid curves) and forecasts without TDR data assimilated in "Oneway_hyb" (nTDR, dashed curves) at the analysis time as a function of forecast lead time. TDR data are available at 0000 UTC 24, 1200 UTC 24, 1200 UTC 25, 1800 UTC 25, 0000 UTC 26, 1200 UTC 26, and 0000 UTC 27 August 2011. "OnewaynTDR_hyb" does not assimilate any TDR. Forecast cases are separately analyzed based on if TDR data is assimilated in "Oneway_hyb" in this figure.

Figures 8 and 11 have showed that "Oneway_hyb" produces smaller MSLP and MSW biases than "Twoway_hyb" does, and the difference between these two experiments is that the latter applies re-centering. That means EnKF and Hybrid analyses affect each other in DA cycle. Position spreads of the ensemble forecast, EnKF analysis and re-centered analysis during cycling period are compared in Figure 13. It can be seen that EnKF analysis does not change the position spread much relative to that of the ensemble forecast because there is not too much information constraining the storm location in EnKF. The re-centering procedure, however, relatively greatly reduces the position spread. This reduction can be seen more clearly in Figure 14 that shows the horizontal distribution of ensemble positions at various times. Figure 14 shows not only that the position spreads are reduced after re-centering but also that the positions are re-placed to be closer to the hybrid storm position. In a hurricane environment, if the positions of the EnKF forecast and analysis are two far away from the hybrid storm position, the re-centering could modify (destroy) the dynamic balance in EnKF analysis and therefore make the "Twoway" result worse than the "Oneway" results.

To make sure the positions of EnKF analysis do not depart two far away from the hybrid analysis that is used to re-center the EnKF analysis and causes problem in the experiment "Twoway_hyb", another experiment "Twoway_hyb_tcv" are conducted. In "Twoway_hyb_tcv", both the Hybrid and EnKF DA assimilate TC vital data that include the storm location and MSLP information. By assimilating these TC vital data, the intensity biases are greatly reduced and are nearly zero at initial time (Figure 15). However, the MSW bias becomes negative shortly after initial time and grows quickly with time. By the end of the 48-h forecast, the bias is much larger than that of "Twoway_hyb" that does not assimilate TC vital data. Simply assimilating TC vital could help correct storm's location and intensity at analysis time, but the one-point observation could destroy the balance in other fields and cause other problems.

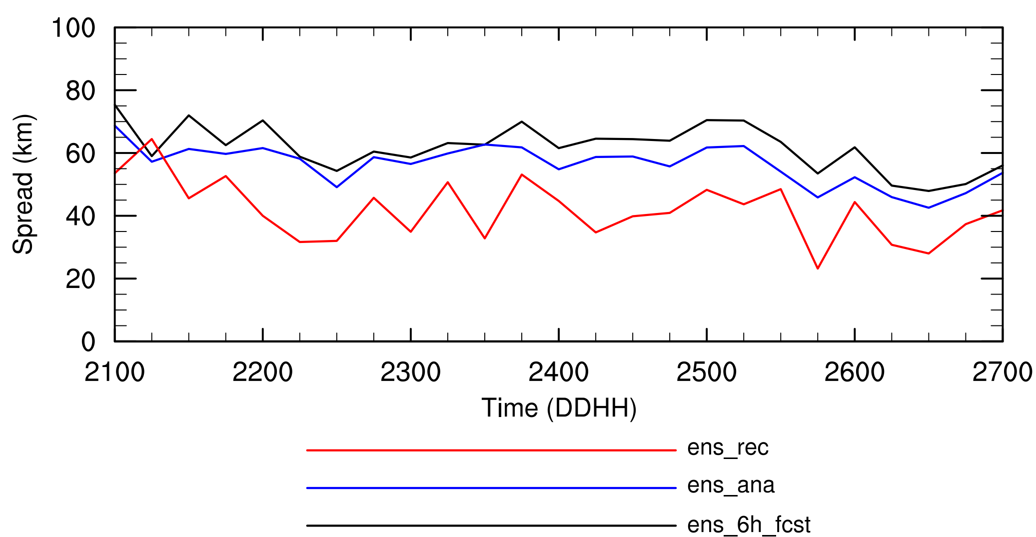

Figure 13. Ensemble position spreads of ensemble forecast, EnKF analysis and re-centered EnKF analysis as a function of cycle time. The mean position is calculated by

averaging all the ensemble members' positions. Then position spread is calculated using the mean position and ensemble member positions.

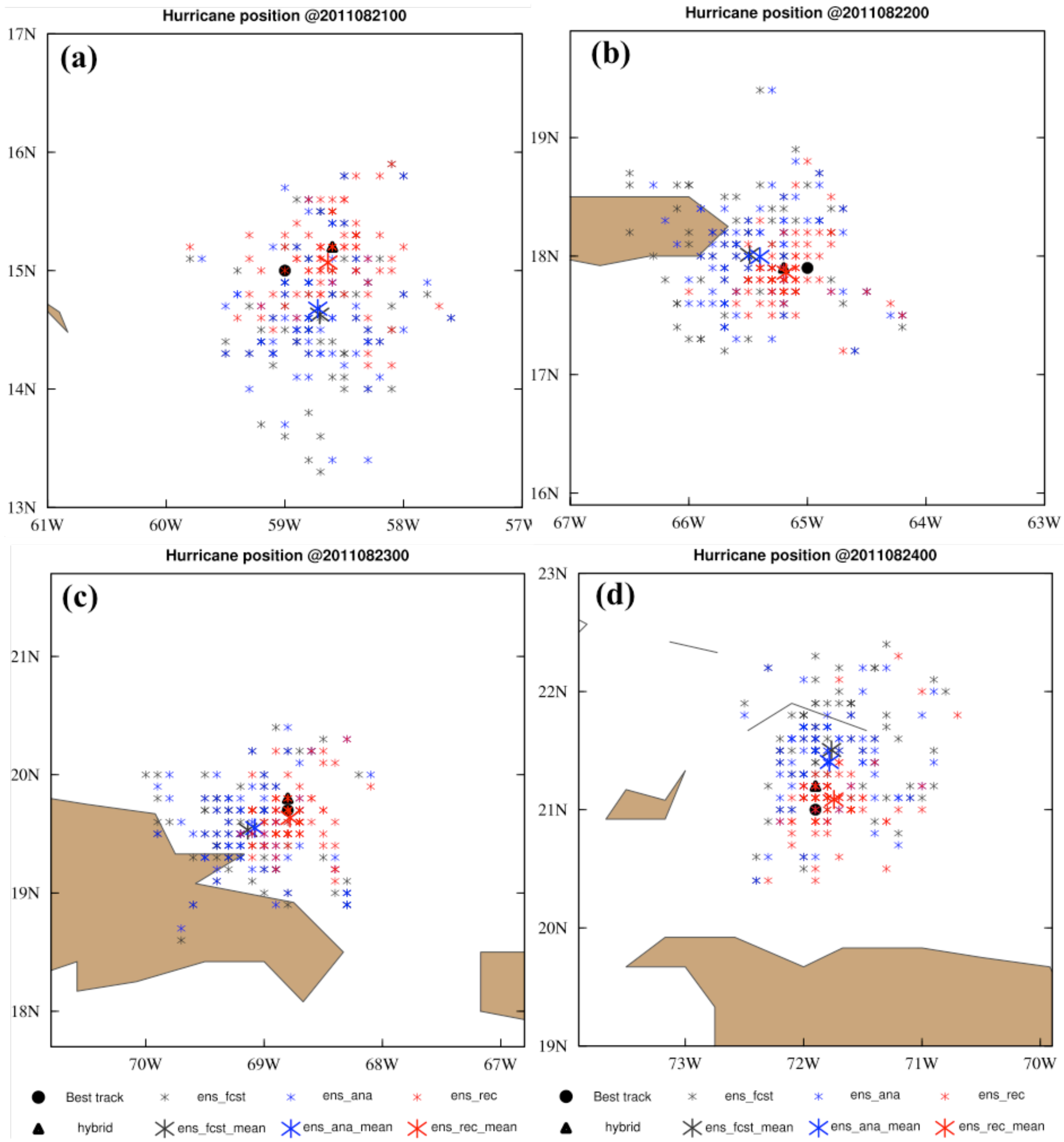

Figure 14. Ensemble positions of all members of ensemble forecast, EnKF analysis and re-centered analysis at (a) 0000 UTC 21, (b) 0000 UCT 22, (c) 0000 UTC 23, and (d) 0000 UTC 24 August 2011.

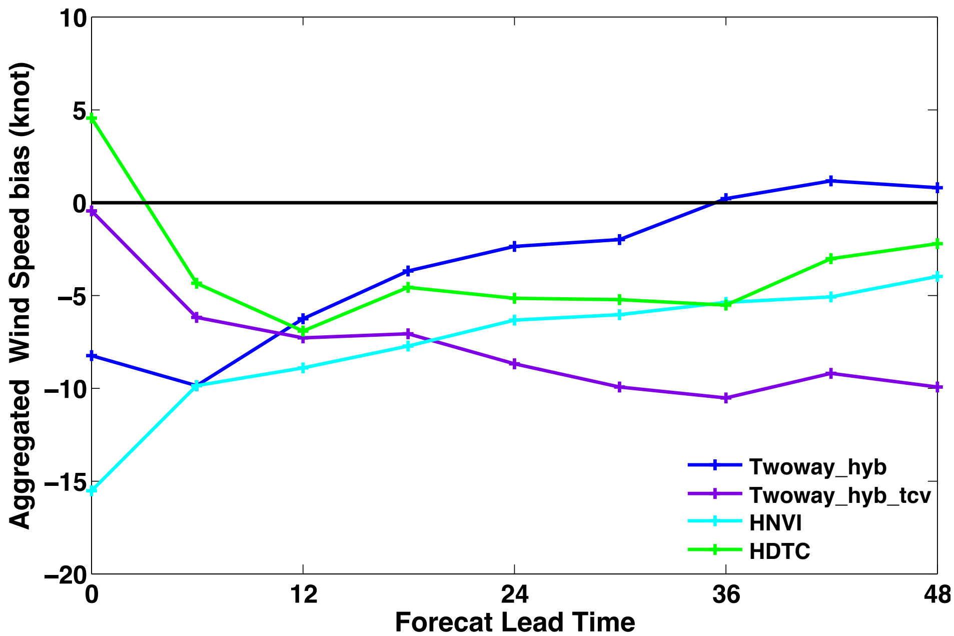

Figure 15. Aggregated bias of Maximum surface wind speed averaging over all the forecasts as a function of forecast lead time.

4. Discussion

Regional EnKF and ensemble forecast with 27-km and 9-km grid spacings are run together with the GSI Hybrid data assimilation cycles. The regional EnKF provides background error covariances for GSI Hybrid DA, which is different from the HWRF model that uses global forecast. The GSI hybrid can, in turn, re-center EnKF analysis. A series of experiments with various configurations are conducted to investigate the performance of the regional system. Specifically, experiments are designed to look at the impacts of the re-centering, TDR data and TC vital data. 6 hourly cycling DAs for the GSI Hybrid analysi are carried out between 1800 UTC 20 August and 0000 UTC 27 August 2011, and 25 forecast samples are collected. Statistical results are presented based on these samples as well as forecasts samples from the other two experiments with HWRF operational configurations.

Both the "Oneway_hyb" and "Twoway_hyb" obtain overall lower MWS biases than the "HNVI" throughout the 48-h lead time. This is benefited from the regional ensemble because all these three experiments have no vortex initialization applied. Both

"Oneway_hyb" and "Twoway_hyb" also have lower MWS biases than "HDTC" after the first 12-h lead time.

It is also noted that "Oneway_hyb" performs better than "Twoway_hyb" in terms of MSLP forecast. By examining the ensemble position spread and distribution of the ensemble forecast, analysis and re-centered analysis, it is found that the large position spread could be the reason that explains why "Twoway_hyb" does not gain better performance in intensity forecast. The re-centering applied in "Twoway_hyb" reduces position spread and re-places hurricane locations, but also changes the dynamic balance in the EnKF analysis in a bad way. To make sure the hurricane locations are not too different in the Hybrid analysis and the EnKF analysis in a "Twoway" experiment, experiment "Twoway_hyb_tcv", which assimilates TC vital data, is conducted. The MSW bias is nearly zero at initial time in "Twoway_hyb_tcv", but becomes negative shortly after initial time and grows quickly with time. By the end of the 48-h forecast, the bias is much larger than that of "Twoway_hyb" that does not assimilate TC vital data. Simply assimilating TC vital could help correct storm's location and intensity at analysis time, but cause other problems. Overall, the experiment assimilating TC vital data does not outperform the experiment without TC vital assimilation.

While simply assimilating one-point TC vital data does not help improve the intensity forecast, the assimilation of TDR data is proven helpful. The TDR data provides valuable inner core information that usually is missing in a hurricane environment. However, results show that the assimilation of TDR could also cause problems. For instance, the storm intensifies too fast and positive biases could be caused in longer forecast lead times for the forecasts with TDR assimilated. This at least indicates the TDR data does impact the intensity forecast, although the use of these data in a more positive way is still a question.

It was pointed out that the re-centering procedure does not help improve the intensity forecast as expected. There are potential solutions in future study. Now only 9-km ensemble EnKF and forecast are running, and the 3-km Hybrid uses the 9-km ensemble forecast for background error covariances with "dual resolution". Future work can run 3-km ensemble given the computational resources permitted. By doing so, more small-scale information can be gained in the 3-km EnKF, and more detailed background error structures can be passed to the 3-km Hybrid DA. With well behaving EnKF analysis, "Twoway" run could produce better forecast. Vortex initialization technique is also worthy trying in future study. A proper vortex initialization method not only corrects the storm location, but also maintains the balance in other fields, and therefore can keep the positive impact of data assimilation in the following forecast. Combining vortex initialization with regional EnKF in GSI Hybrid DA is expected to beat the forecast in current study.

Reference

Avila, L. A., and J. Cangialosi, 2011: Hurricane Irene (AL092011), tropical cyclone report, National Hurricane Center, Miami, FL.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.,* **130**, 1913–1924.

Zhou, C., H. Shao, L. R. Bernardet, M. Tong, and V. Tallapragada: Regional Applications of the GSI-Hybrid Data Assimilation for High-Resolution Tropical Storm Forecasts. 95th AMS annual meeting, Phoenix, AZ, January 2015.