

Diagnosing Tropical Cyclone Motion Forecast Errors in the 2015 HWRF Retrospective Test (H215)

Thomas J. Galarneau, Jr.¹
National Center for Atmospheric Research
Boulder, Colorado

Final Report for the 2014–15 DTC Visitor Program
31 July 2015

Contents of the report:

1. Overview and Summary of Key results
2. Data and methods
3. Track error statistics for 2012–2014
4. Analysis of forecast bust cases
5. Final comments
6. References
7. Figures

Corresponding author address:

Thomas J. Galarneau, Jr.
The University of Arizona
1118 E. 4th St, PAS 570
P.O. Box 210081
Tucson, AZ 85721
E-mail: tgalarneau@email.arizona.edu

¹ New affiliation: Department of Hydrology and Atmospheric Sciences, The University of Arizona, Tucson, Arizona

1. Overview and Summary of Key Results

The aim of this proposal was to examine tropical cyclone (TC) track forecasts from the 2015 Hurricane Weather Research and Forecasting model retrospective test (H215). As part of analyzing track error characteristics, the diagnostic approach developed by Galarneau and Davis (2013) to examine the sources of numerical model forecast error that contribute to degraded tropical cyclone (TC) motion forecasts was utilized. Emphasis was placed on 120-h (day 5) forecasts of TC track in order to assess whether H215 has added value over the global operational model in the medium range.

The results show that while medium-range H215 TC track forecasts improve on cases defined as “busts” for the GFS, the overall H215 absolute track error and bias is not significantly different from older versions of HWRF (H214) or the operational GFS. For the 2012–2014 North Atlantic seasons, the H215 slow track bias was dominated by forecasts from four TCs: Kirk, Leslie, Nadine, and Sandy. The forecasts for these cases, which all involved interaction with subtropical and midlatitude circulation features, were much worse than the pre-defined HFIP baseline error for 120-h forecasts. The steering flow analysis and motion error diagnosis showed that the H215 slow bias for these cases does not point toward an obvious problem with the model physics. The slow bias is driven primarily by errors in the environment wind, which points to errors in the synoptic-scale subtropical and midlatitude flow. It appears that these errors are likely inherited from the GFS forecast boundary conditions. The one exception is Sandy, where the environment wind errors appear to be linked to the treatment of convection on Sandy’s northwest side, as shown in previous work on other modeling systems in the literature (e.g., Bassill 2014, 2015; Torn et al. 2015). It may be useful to rerun these slow bias “bust” cases using the GFS analysis for the boundary conditions to see if there are expected improvements in the prediction of the synoptic-scale flow. This will provide a better sense of how much of the errors in the synoptic-scale flow are driven internally by H215 versus the GFS forecast boundary conditions.

2. Data and methods

Galarneau and Davis (2013) defined the steering flow as the spatially averaged environment wind that best matches the TC motion. They computed the steering flow by first computing the environment wind every 50 hPa in the 850–200 hPa layer using eight different radii ranging from 1–8° in which to remove the TC vortex. Then for each radius, the pressure-weighted vertically averaged environment wind was computed for layers of increasing depth ranging from 850–800 hPa to 850–200 hPa. For all possible depth and radius combinations, or candidate steering layer values, the magnitude of the vector difference between the actual TC motion and steering flow was computed. The radius and steering layer depth combination that produced the smallest vector residual magnitude was the optimal steering layer definition chosen for the TC at the given time. The optimal steering layer method showed a marked improvement over the deep-layer mean steering method (e.g., Velden and Leslie 1991), by sharpening the agreement between TC motion and steering flow. The close match between TC motion and steering flow allows for quantification of TC motion errors and partition of the difference in steering flow

between the forecast and analysis into errors in environment wind, vertical steering depth, and TC removal radius.

Using the steering flow definition defined above, Galarneau and Davis (2013) derived a diagnostic equation that quantitatively assesses the sources for forecast TC motion error. The TC motion error diagnostic equation is defined as

$$\underbrace{\vec{V}_m - \vec{V}_o}_{\text{storm motion error}} = \underbrace{\frac{1}{p_b - p_{to}} \int_{p_{to}}^{p_b} (\hat{v}_m - \vec{v}_o) dp}_{\text{environment wind error}} + \underbrace{\frac{1}{p_b - p_{tm}} \int_{p_{tm}}^{p_b} -(\hat{v}_m - \vec{v}_m) dp}_{\text{TC removal radius error}}$$

$$+ \underbrace{\frac{1}{p_b - p_{tm}} \left[\int_{p_{to}}^{p_b} \left(\frac{p_{tm} - p_{to}}{p_b - p_{to}} \right) \hat{v}_m dp + \int_{p_m}^{p_{to}} \hat{v}_m dp \right]}_{\text{TC steering depth error}} + \text{residual term}$$

where $p_b=850$ hPa and p_t is defined as the top of the optimal steering layer, \mathbf{V} is defined as the TC motion (using positions at time ± 12 h), \mathbf{v} is the area-averaged environment wind (TC removed) computed over the optimal radius, and $\hat{\mathbf{v}}$ is the area-average environment wind (TC removed) computed for the model TC over the radius determined for the observed TC. The subscripts ‘o’ and ‘m’ indicate the observed storm and steering flow and the model storm and steering flow, respectively. The term on the left hand side represents the vector error between the observed and model TC motion. The first term on the right hand side represents the contribution from model environment wind errors in the observed steering flow layer, the second term represents the contribution from differences in the TC removal radius between the model and observed TC in the model steering flow layer, and the third term represents the contribution from differences between the model and observed steering layer vertical depth. The fourth term represents the residual term, which includes analysis wind errors and storm motion errors introduced by uncertainties in TC position.

The environment wind error term is defined as the model environment wind minus the analysis environment wind integrated over the steering depth for the observed storm, both using the radius defined for the observed storm (r_o). Thus, the environment wind error term is a function of only the difference in wind between the analysis and forecast all due to vorticity and divergence differences outside of r_o since it is this vorticity and divergence that drives the wind inside of r_o . The next two terms – TC removal radius and vertical steering depth term – arise because we allow the radius and vertical steering depth to differ between the model and observed storm at a given time. In cases when $p_{tm}=p_{to}$ or when $\hat{v}_m(p)$ does not vary with pressure (no vertical shear) the steering depth term will go to zero. Likewise, when $r_m=r_o$ the radius term will go to zero. Because the TC motion error diagnostic equation uses a local coordinate system, Galarneau and Davis (2013) analyzed 24-h TC position forecasts so that the spatial variation of the environment wind between the observed and forecast TC locations could be ignored. It is likely, however, that notable 24-h forecast TC motion errors will

contribute to large position errors at longer lead times. However, there is no formal limitation to applying this diagnostic to any lead-time.

This study will apply the motion error diagnostic to medium-range forecasts of TC motion from the gridded H215 dataset. The motion diagnostic is performed on the forecasted and observed TC just prior to the development of large (> 200 km) position differences. The forecasted motion is computed using the “ATCF” output, and the steering flow using gridded winds on pressure levels from the outer domain. The forecasts are compared to the motion of the observed TC as determined from the Hurricane Best Track Database (HURDAT). The steering flow for the observed storms is computed using the NCEP Global Forecast System (GFS) analyses at a horizontal resolution of 0.5° . The H215 data is interpolated to a regular latitude-longitude grid at 0.5° grid spacing in order to match the resolution and grid structure of the GFS analyses. The details on how the TC vortex is removed for the steering layer flow calculation follows Galarneau and Davis (2013).

3. Track error statistics for 2012–2014

The mean absolute track error and bias for the GFS, H214, and H215 modeling systems during the 2012–2014 North Atlantic TC seasons are shown in Fig. 1. Note that all three modeling systems show significant improvement relative to the HFIP baseline track error in the 24–84-h forecast range (Fig. 1a). Despite these important improvements over the last seven years, the track forecasts are all worse than the HFIP baseline by 120-h. At all forecast lead times, the performance of each modeling system is not significantly different, with the H214 and H215 forecasts slightly worse than the GFS at all lead times. The H215 forecasts are slightly improved compared to the H214 forecasts at all lead times. The TC track bias, in motion-relative coordinates, is relatively consistent among the three modeling systems, with a distinct slow bias apparent at virtually all lead times (Fig. 1b). The H214 and H215 modeling systems also suggest a left-of-track bias, with the H215 system shifted to right-of-track by 120-h.

Closer inspection of the 120-h GFS TC track forecasts reveals that the absolute track error peaks below the HFIP baseline error of 411 km (Fig. 2a). It appears that the overall track forecast statistics are influenced by 18 “bust” cases (defined here as $+1.0$ sigma errors). These forecasts are comprised of 10 forecasts for Nadine, 3 for Sandy, 3 for Michael, and 2 for Leslie, highlighting that 2012 was a difficult year for TC track prediction over the North Atlantic. The H214 and H215 120-h track forecasts for these GFS bust cases shows significant improvement over the GFS (Fig. 3a). While both versions of HWRF add value to the GFS for the GFS bust cases, the H215 forecasts show more improvement over the GFS when compared to the H214 forecasts (Fig. 3b). This result suggests that H215 (and to a lesser extent H214) can help in situations where the GFS may be “going off the rails”, but the problem is that H215 has a similar distribution of absolute errors overall as compared to the GFS (Fig. 2b). Much like the GFS, the 120-h absolute track error for H215 peaks below the HFIP baseline, but the tail in the distribution negatively impacts the overall track error statistics. Some of the H215 bust

cases overlap with the GFS busts, but most do not, suggesting that while H215 can add value to GFS bust cases, it can also be significantly degraded relative to the GFS.

The 120-h TC track forecast bias for H215 is shown in Fig. 4. Note that while forecasts that are worse than the HFIP baseline can occur in any direction relative to the TC motion, the “bad” forecasts are most frequently a slow error that can produce the largest track errors overall. The remainder of this report will discuss a detailed analysis of the “bad” H215 forecasts characterized by a slow bias (the forecasts enclosed by the blue dashed circle in Fig. 4). These forecasts are comprised of one forecast for Kirk (initialized at 00Z/29 August 2012), one for Leslie (12Z/6 September 2012), four for Sandy (18Z/24, 00Z/25, 06Z/26, and 12Z/26 October 2012), and 20 for Nadine (00Z/13, 00Z/15, 12Z/15, 18Z/15, 00Z/16, 06Z/19, 12Z/19, 00Z/20, 06Z/20, 00Z/23, 06Z/23, 12Z/23, 00Z/24, 06Z/24, 12Z/24, 18Z/24, 00Z/25, 06Z/25, 12Z/25, and 00Z/29 August 2012).

4. Analysis of forecast bust cases

a. TC Sandy

The H215 and GFS track forecasts for the Sandy “bust” cases are shown in Fig. 5. The forecasts initialized at 1800 UTC 24 and 0000 UTC 25 October 2012 had the most difficulty with Sandy’s unusual left hook toward New Jersey (Figs. 5a,b). The forecast initialized at 0600 UTC 25 October started to capture the left hook, but was a bit slow compared to observations. The forecasts initialized at 0600 and 1200 UTC 26 October were much improved as most of the track error developed after landfall. In fact, these two forecasts could be considered “good” forecasts, unfortunately the large track errors that developed after landfall negatively impact the overall track error statistics. For all of the Sandy forecasts, the H215 model outperformed the GFS.

Inspection of the absolute and track-relative error for the forecasts initialized at 1800 UTC 24, 0000 UTC 25, and 0600 UTC 25 October 2012 reveals that the forecasts for Sandy were quite good through 1800 UTC 28 October (Fig. 6). Rapid error growth occurred after 0000 UTC 29 October, with most of the error in the along track direction. The forecast initialized at 0000 UTC 25 October will now be inspected in more detail, with an emphasis on the verifying time of 1200 UTC 28 October when the forecasted and observed TC were still within 100 km of each other (Fig. 6a).

Analysis of the steering flow for TC Sandy at 1200 UTC 28 October 2012 is shown in Fig. 7. The steering layer for Sandy is defined in the 850–400 hPa layer with $r_0=1^\circ$ for both the forecasted and observed system. The \mathbf{V}_{env} profile shows that the H215 forecast has a northeasterly error (Fig. 7a). The steering layer mean wind and vorticity shows that the observed system, as represented by the GFS analysis, is embedded in a southwesterly flow driven primarily by a region of cyclonic vorticity northwest of Sandy (Fig. 7b). Conversely, the 84-h H215 forecast has Sandy embedded in weak northwesterly flow (Fig. 7c). The key difference between the H215 forecast and the GFS analysis is the differences in near-storm vorticity. The H215 storm is more asymmetric, with much more

cyclonic vorticity located on the southeast flank of Sandy. The H215 forecast lacks cyclonic vorticity to the northwest of Sandy, which ultimately contributes to Sandy's motion out to sea in the forecast.

The 850 hPa height, temperature, and frontogenesis maps derived from the GFS analysis for 1200 UTC 27–0000 UTC 29 October 2012 shows that Sandy was moving northeastward into an increasingly baroclinic environment (Fig. 8). A persistent region of frontogenesis was concentrated on the northwest side of Sandy where its rotational flow interacted with a cooler air mass moving southeastward off of eastern North America. The 60–96-h H215 forecast indicates that the frontogenesis on the northwest side of Sandy was not apparent in the forecast, as the flow appears to be more parallel to the temperature contours (Fig. 9). During the same period, convection was organized along the band of frontogenesis in a very moist environment in mid-levels (Fig. 10). Consistent with the lack of frontogenesis in the forecast, convection was not organized on the northwest side of Sandy in the H215 forecast (Fig. 11). Additionally, dry air (with relative humidity values $< 50\%$) appeared to be wrapping westward on the northern side of Sandy on 28 October (Figs. 11b,c). The lack of convection on the northwest side of Sandy in the H215 forecast appears to be related to reduced forcing for ascent and decreased moisture.

Examination of the potential vorticity (PV) in the 700–500 hPa layer clearly shows a positive mid-level PV band on the northwest side of Sandy (Fig. 12). The positive PV band is consistent with the steering flow analysis (Fig. 7b), and is important in steering Sandy on a northward course leading to its left hook into New Jersey. In the 300–200 hPa layer, cyclonic roll-up of the midlatitude trough is well underway by 0000 UTC 28 October. The upper-level trough captured Sandy and steered it back toward New Jersey. The H215 forecast of PV shows a different evolution (Fig. 13). First, the mid-level band of positive PV is absent in the forecast, which is consistent with the earlier steering layer flow analysis (cf. Fig. 7c). The lack of positive PV northwest of Sandy is linked to the reduced frontogenesis and attendant convection. Also of interest is that the midlatitude trough over the Great Lakes and Appalachians is a bit more progressive in the H215 forecast (compare Figs. 12d and 13d). This difference in trough structure and position is also likely due to the differences in convective activity on the northwest side of Sandy. The upper-level outflow associated with this convection acted to “push back” on the upper-level trough, as marked by the increased ridging over the northeast U.S. and PV gradient over the Appalachians compared to the H215 forecast.

Vertical cross sections oriented northwest-to-southeast through the positive PV band are shown in Fig. 14. Although the Sandy vortex has similar structure in both the H215 forecast and GFS analysis at 1200 UTC 28 October, the surrounding environment is quite different. Note that the environment in the H215 forecast is much drier at mid-levels on the northwest side of Sandy with weaker and shallower convection (Figs. 14a,d). The axis of frontogenesis sloping toward the cold air in the GFS analysis is not apparent in the forecast (Figs. 14c,f). Also, the upstream trough is broader in scale in the forecast and has southwesterly flow that extends farther to the southeast (Figs. 14b,e).

In summary, the forecasts for TC Sandy initialized at 1800 UTC 24 through 0600 UTC 25 October 2012 were unable to capture Sandy's left hook toward New Jersey and carried Sandy out to sea. The steering flow analysis pointed toward a northeasterly error in the environment wind, which limited Sandy's progression northward compared to observations. It appears that a well-defined band of cyclonic vorticity on the northwest side of Sandy that was persistent during 27–29 October, and was important in steering Sandy northward, was absent in the H215 forecast. The absence of this vorticity band is linked to the model's inability to produce persistent convection on Sandy's northwest side, in response to weaker lifting/frontogenesis and reduced moisture.

b. TC Kirk

The H215 and GFS track forecast for Kirk initialized at 0000 UTC 29 August 2012 is shown in Fig. 15. Both the H215 and GFS forecasts recurve Kirk similar to observations. After recurvature, both model forecasts move Kirk northeastward too slowly (Fig. 15a). By 120-h, the H215 forecast is worse than the GFS forecast. The largest error growth for H215 occurs after 0000 UTC 1 September in the post-recurvature stage (Fig. 15b) and is primarily an along track slow error, with a smaller contribution from a left-of-track error (Fig. 15c).

Analysis of the 48-h H215 forecast steering flow for Kirk at 0000 UTC 1 September shows a northeasterly environment wind error present in the 850–500 hPa layer (Fig. 16a). The environment wind error appears to be linked to circulation errors in the midlatitude flow. Increased anticyclonic vorticity is present on the north side of Kirk in the H215 forecast compared to the GFS analysis, resulting in weaker flow over the vortex (Figs. 16b,c). The upstream midlatitude trough was more progressive in the forecast, which acted to pull the cyclonic vorticity streamer north of Kirk farther north in the H215 forecast, resulting in increased anticyclonic vorticity north of Kirk. Unlike for the Sandy case, this forecast “bust” for Kirk is related to errors in the midlatitude flow rather than the treatment of near-storm convection.

c. TC Leslie

The H215 and GFS track forecast for Leslie initialized at 1200 UTC 6 September 2012 is shown in Fig. 17. As with Kirk, both the H215 and GFS forecasts recurve Leslie similar to observations. After recurvature, however, both models fail to move Leslie on a northeasterly course toward Greenland (Fig. 17a). This error suggests that there was difficulty in capturing Leslie's interaction with the midlatitude flow in the H215 forecast. The H215 forecast absolute track error grows rapidly starting at 1200 UTC 10 September, primarily as a slow bias (Figs. 17b,c).

Analysis of the 72-h H215 forecast steering flow for Leslie at 1200 UTC 9 September shows a northeasterly environment wind error through a deep layer, but primarily in the 800–400 hPa layer (Fig. 18a). The GFS analysis of steering layer (850–200 hPa) flow and vorticity shows that Leslie was embedded in southerly flow associated with a subsynoptic-scale cyclonic vortex west of Leslie (Fig. 18b). This vortex helped to steer Leslie northward resulting in a more robust interaction with the midlatitude trough over the Great Lakes region. In the H215 forecast, this midlatitude trough is more progressive

and has already absorbed the subsynoptic-scale vortex west of Leslie (Fig. 19). As a result, Leslie is moving more slowly and eastward compared to observations and does not interact with the approaching midlatitude trough (Fig. 18c and 19d). The trough eventually passes to the north of Leslie in the H215 forecast (not shown). Much like the Kirk case, this forecast “bust” for Leslie is related to errors in the midlatitude flow.

d. TC Nadine

The complex life cycle of TC Nadine is summarized by a sequence of dynamic tropopause (DT; defined at the 2.0 PVU surface) maps for 17–22 September 2012 shown in Fig. 20. Nadine interacted with multiple midlatitude troughs that extended into the subtropics. On 17–18 September, Nadine was positioned on the southern flank of an upper-level trough (Figs. 20a–d). A shortwave trough approached from the west, which began to steer Nadine to the northeast through 1200 UTC 19 September (Figs. 20d–f). This shortwave trough was absorbed into Nadine’s circulation and weakened in the deformation flow ahead of the next upper-level trough approaching from the west. Ridge building over the western North Atlantic in response to cyclogenesis along the U.S. east coast acted to drive the downstream trough southward into lower latitudes (Figs. 20d–h). This second upper-level trough redirected Nadine to the southeast through 1200 UTC 21 September (Figs. 20h–j). At 1200 UTC 21 September, the upper-level trough fractured with the bulk of the disturbance moving eastward north of Nadine and the southern end that fractured into a cutoff low positioned to the west of Nadine (Fig. 20j). As the main trough moved to the east, Nadine interacted with the cutoff low and began to move on a westward course on 22 September (Figs. 20j–l).

Nadine proved to be a problematic case for both the GFS and H215. The first set of Nadine forecasts consist of forecasts that move too slow and recurve too early (Fig. 21). These forecasts were initialized during and just after the initial recurvature of Nadine. Generally, the GFS performed better than H215 for these forecasts of Nadine. The forecasts for Nadine initialized after 0000 UTC 15 September show rapid error growth beginning at 0000 UTC 18 September (Fig. 22a). The initial rapid error growth is primarily a left-of-track error, evolving into a slow error, which in this case is indicative of an early recurvature problem (Fig. 22b; cf. Fig. 21).

Analysis of the steering flow for the 24- (blue; 0000 UTC 16 September initialization) and 48-h (red; 0000 UTC 15 September) H215 forecasts verifying at 0000 UTC 17 September 2012 shows an easterly wind error throughout a deep layer (Fig. 23a). The weaker westerly steering flow is consistent with Nadine moving too slow and recurving early. For the 48- and 72-h H215 forecasts verifying at 0000 UTC 18 September, the easterly wind error is still an issue within the steering layer (Fig. 23b). Close inspection of the steering layer mean environment wind and vorticity for the GFS analysis and 48-h H215 forecast verifying at 0000 UTC 17 September shows that the trough north of Nadine is not apparent in the H215 forecast (Figs. 24a,b). As a result, the anticyclonic environment wind error north of Nadine, combined with a cyclonic wind error south of Nadine, conspire to produce an easterly environment wind error over Nadine (Fig. 25a). Similarly, the 72-h H215 forecast verifying at 0000 UTC 18

September also misplaced the trough north of Nadine (Figs. 24c,d). As a result, Nadine was embedded in an environment wind characterized by an easterly wind error (Fig. 25b).

The remaining forecasts of Nadine had difficulty with Nadine's unusual loop to the south and back to the west after 1200 UTC 21 September (Figs. 26 and 27). The forecasts initialized on 19–23 September all moved Nadine eastward to near northwest Africa by 120-h (Fig. 26). The time series of absolute track errors for the forecasts of Nadine are summarized in Fig. 28. For the forecasts shown in Fig. 26, the rapid error growth occurred after 48-hours as H215 moved Nadine toward Africa while in reality Nadine began to retreat to the west (Fig. 28a).

The forecast for Nadine initialized at 1200 UTC 19 September 2012 (Fig. 26b) is representative of all the forecasts initialized on 19–23 September, and will now be analyzed further. Analysis of the steering layer flow for the 48-h H215 forecast verifying at 1200 UTC 21 September shows a northwesterly environment wind error throughout a deep layer above 700 hPa (Fig. 29a). The observed storm followed a very shallow steering layer, defined here as the 850–800 hPa layer. The environment wind errors within this shallow layer are very small. The forecasted storm followed a much deeper steering layer, defined as the 850–500 hPa layer. The impact of differences in steering layer depth are shown by the TC motion error diagnostic equation terms in Fig. 29b. The actual motion error at 1200 UTC 21 September is toward the east-southeast, driven primarily by differences in steering layer depth. Therefore, much of the errors in TC position for Nadine are related to the much deeper steering layer flow in H215 versus the GFS analysis. Additionally, there is increased vertical wind shear in the 850–500 hPa layer in H215 compared to the GFS analysis (Fig. 29a).

The flow at 500 hPa in H215 is too strong from the northwest, and is related two to factors. First, the trough to the northeast of Nadine is located slightly closer to the Nadine in the H215 forecast (Figs. 30a,b). Additionally, the cutoff low that previously fractured from this trough and is now located on the west side of Nadine is displaced farther west away from Nadine in the H215 forecast. These differences in position of the trough and cutoff low resulted in a cyclonic error to the northeast and anticyclonic error to the west side of Nadine, which drives an enhanced northwesterly flow (and northwesterly 850–500 hPa shear) over Nadine (Fig. 30c).

So, why did the H215 TC follow a deeper steering layer flow than in the GFS analysis? Inspection of vertical cross sections of PV, potential temperature, and wind component normal to the cross section shows that there is little difference in the vertical extent of the PV tower in the GFS analysis and 48-h H215 forecast (Fig. 31). Therefore, the steering layer depth term in the motion error diagnostic equation is not really pointing to differences in vortex structure, rather it is telling us that the Nadine forecasts went off the rails because of big errors in the vertical shear in the environment.

5. Final comments

The aim of this proposal was to examine tropical cyclone (TC) track forecasts from the 2015 Hurricane Weather Research and Forecasting model retrospective test (H215). Emphasis was placed on 120-h (day 5) forecasts of TC track in order to assess whether H215 has added value over the global operational model in the medium range. The overall statistics for the 2012–2014 North Atlantic TC seasons showed that while H215 did improve 120-h forecasts for cases in which the GFS produced a “bust” forecast, the overall absolute track error and track bias statistics are not significantly different between the two modeling systems. Both the H215 and GFS have a robust slow bias. While the GFS and H215 120-h forecast track error peaks below the HFIP baseline, notable tails in the distribution (e.g., “busts”) negatively impact the seasonal statistics.

For 2012–2014, the H215 slow bias day 5 “busts” (defined here as forecasts worse than the HFIP baseline error) were dominated by cases that were interacting with the midlatitude flow. These cases were comprised of four TCs from the 2012 North Atlantic TC season: Sandy (4 forecasts), Kirk (1), Leslie (1), and Nadine (20). The large track errors for these four cases were driven by errors in the synoptic-scale flow. Errors for TC Sandy were linked to the generation of cyclonic vorticity on Sandy’s northwest side in conjunction with deep convection along a moist, frontogenetical baroclinic zone. Drier conditions and reduced forcing for ascent in the H215 forecast mitigated the generation of this cyclonic vorticity in the low- to mid-troposphere. This vorticity was important in steering Sandy northward prior to its left hook into New Jersey. The lack of cyclonic vorticity northwest of Sandy in the H215 forecast allowed Sandy to move northeastward out to sea. Forecasts for the other three TCs were driven by errors in midlatitude trough position and structure.

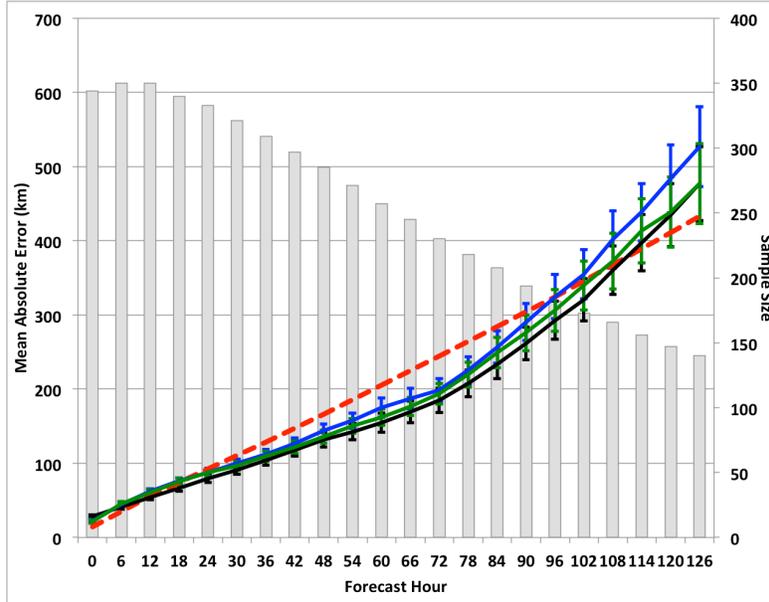
The steering flow analysis and TC motion error diagnostic for the H215 slow bias cases did not obviously point toward a problem with the model physics. It appears that the slow bias is an issue driven by errors in the synoptic-scale flow, which is likely inherited from the GFS forecast boundary conditions. This result is perhaps not surprising since we are looking at day 5 forecasts. To test the contribution of the GFS forecast boundary conditions to the big errors in the H215 forecast bust cases, these cases could be rerun using the GFS analysis as boundary conditions. From a broader perspective, since the H215 forecast did not add value to the GFS day 5 forecasts on average, then perhaps running H215 (or any region model) out to day 5 is not the best use of computer resources from the perspective of track prediction. One strategy may be to run H215 at increased resolution out to day 3 to address the intensity problem, and leave the medium range to the global models.

6. References

- Bassill, N. P., 2014: Accuracy of early GFS and ECMWF Sandy (2012) track forecasts: Evidence for a dependence on cumulus parameterization. *Geophys. Res. Lett.*, **41**, 3274–3281, doi:10.1002/2014GL059839.
- Bassill, N. P., 2015: An analysis of the operational GFS Simplified Arakawa Schubert parameterization within a WRF framework: A Hurricane Sandy (2012) long term track perspective. *J. Geophys. Res.*, **120**, 379–398, doi: 10.1002/2014JD022211.
- Galarneau, T. J., Jr., and C. A. Davis, 2013: Diagnosing forecast errors in tropical cyclone motion. *Mon. Wea. Rev.*, **141**, 405–430.
- Gall, R., F. Toepfer, F. Marks, and E. Rappaport, 2014: National Oceanic and Atmospheric Administration Hurricane Forecast Improvement Project Years Five to Ten Strategic Plan. Tech. Rep. HFIP2014-1.1a, NOAA/Hurricane Forecast Improvement Project, 44 pp. [Available online at http://www.hfip.org/documents/HFIP_StrategicPlan_Yrs5-10_Nov05_2014_Update.pdf.]
- Torn, R. D., J. S. Whitaker, P. Pegion, T. M. Hamill, and G. J. Hakim, 2015: Diagnosis of the Source of GFS Medium-Range Track Errors in Hurricane Sandy (2012). *Mon. Wea. Rev.*, **143**, 132–152.
- Velden, C. S., and L. M. Leslie, 1991: The Basic Relationship between Tropical Cyclone Intensity and the Depth of the Environmental Steering Layer in the Australian Region. *Wea. Forecasting*, **6**, 244–253.

7. Figures

a) Absolute track error (2012–2014)



b) Track bias (2012–2014)

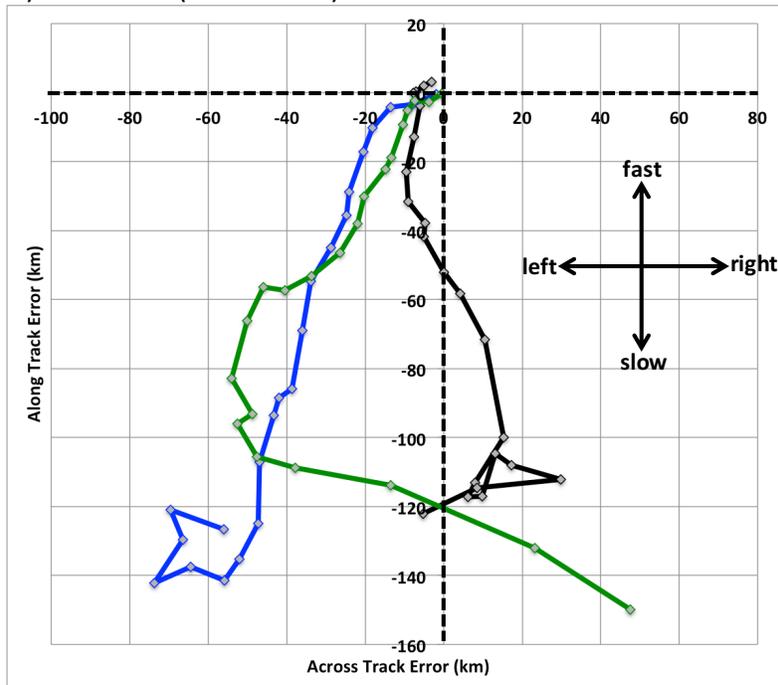


Figure 1: (a) Absolute track error (km) and (b) track bias (km) in TC motion-relative coordinates for the 2012–2014 North Atlantic TC seasons. The HFIP baseline track error (red dashed line; Gall et al. 2014) is shown in (a). The operational GFS, H214, and H215 track error is shown in black, blue, and green, respectively. The error bars in (a) represents the 90% confidence interval. The grey bars in (a) represents the sample size.

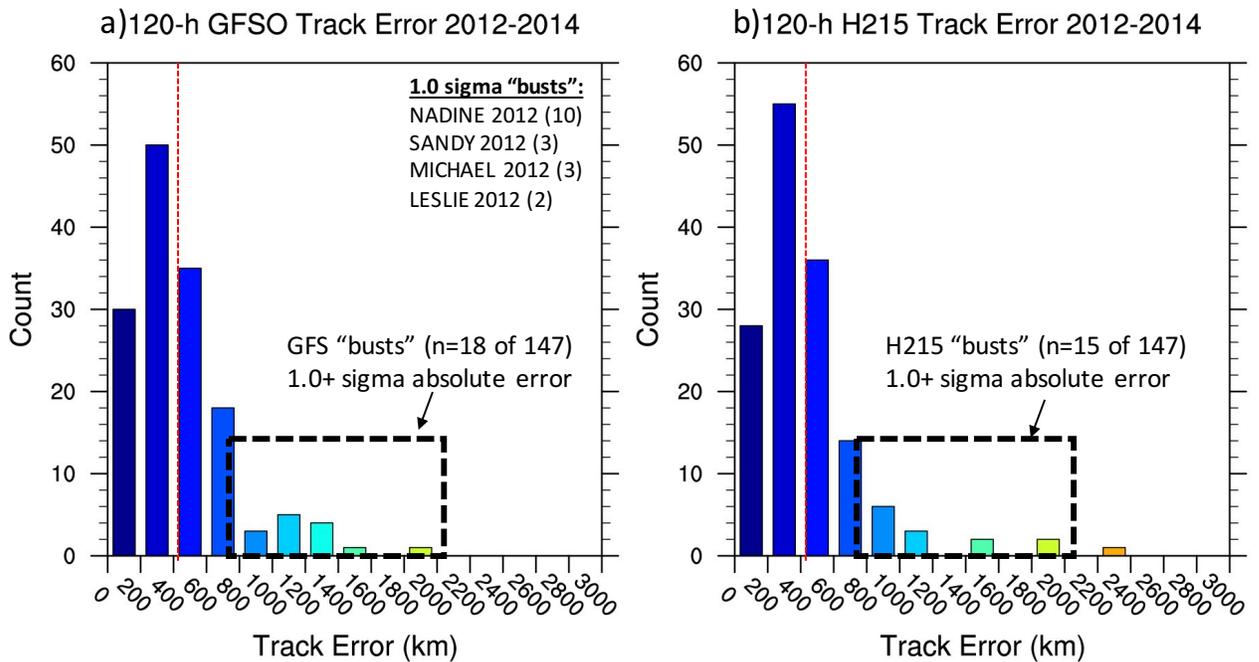
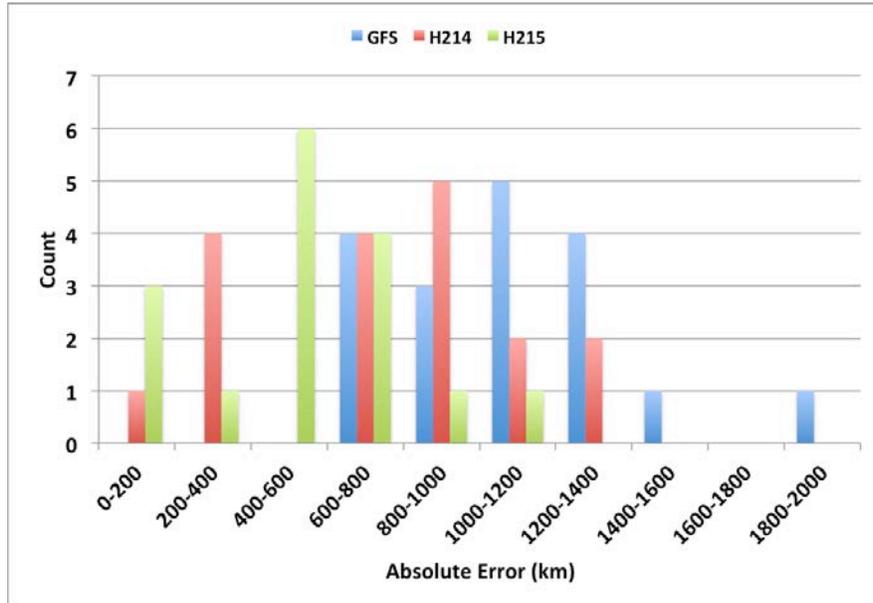


Figure 2: Histogram of absolute track error (km) for 120-h (a) GFS and (b) H215 120-h track forecasts for the 2012–2014 North Atlantic TC seasons. The forecast “busts” are defined as absolute track errors > 1.0 sigma over the mean error. The red dashed line marks the 120-h HFIP baseline track error of 411 km.

a) Distribution of day 5 absolute track error for GFS busts



b) HWRF improvement on GFS track error for day 5 bust cases

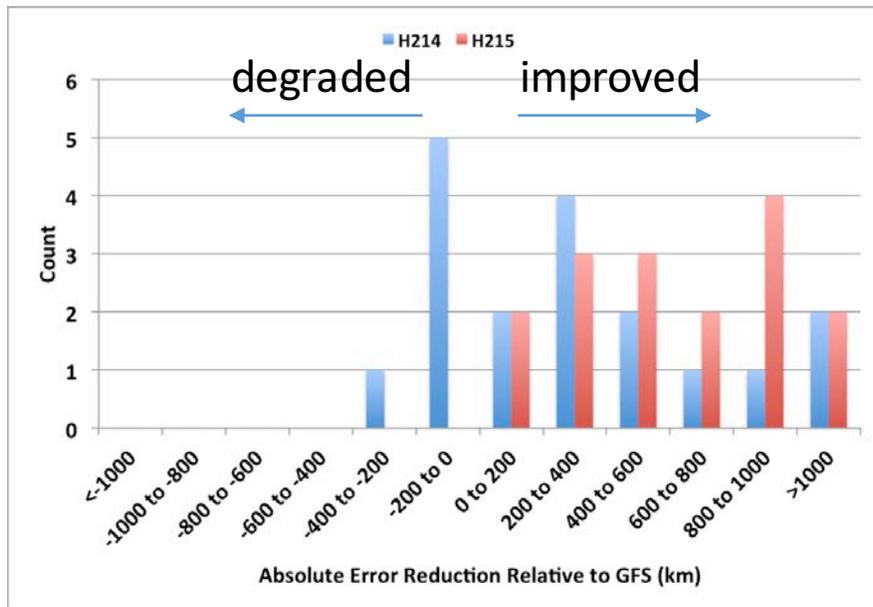


Figure 3: Histogram of (a) 120-h absolute track error (km) from the GFS (blue), H214 (red), and H215 (green) for GFS forecast bust cases, and (b) H214 (blue) and H215 (red) absolute track error improvement (km) over the operational GFS for GFS forecast bust cases.

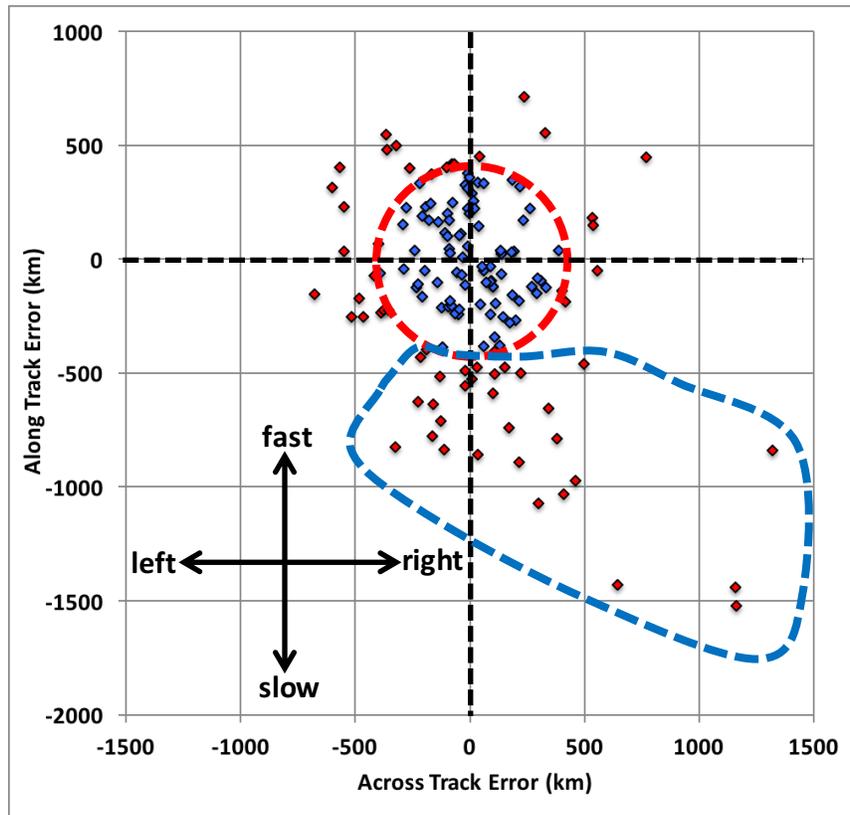


Figure 4: Track bias (km) in TC motion-relative coordinates for 120-h H215 track forecasts from the 2012–2014 North Atlantic TC seasons. The blue (red) diamonds mark forecasts that are better (worse) than the HFIP baseline track error. The forecasts enclosed by the blue dashed circle mark the slow bias to be analyzed further.

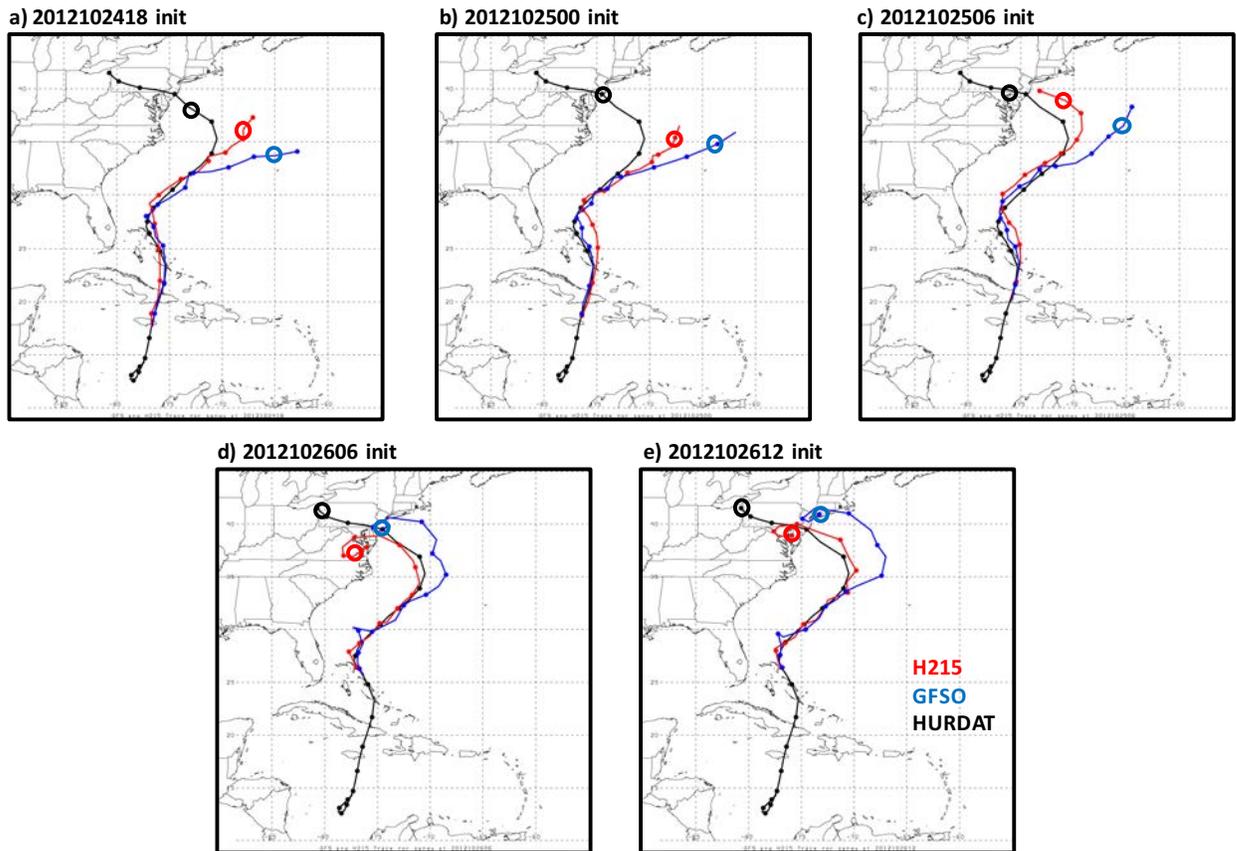
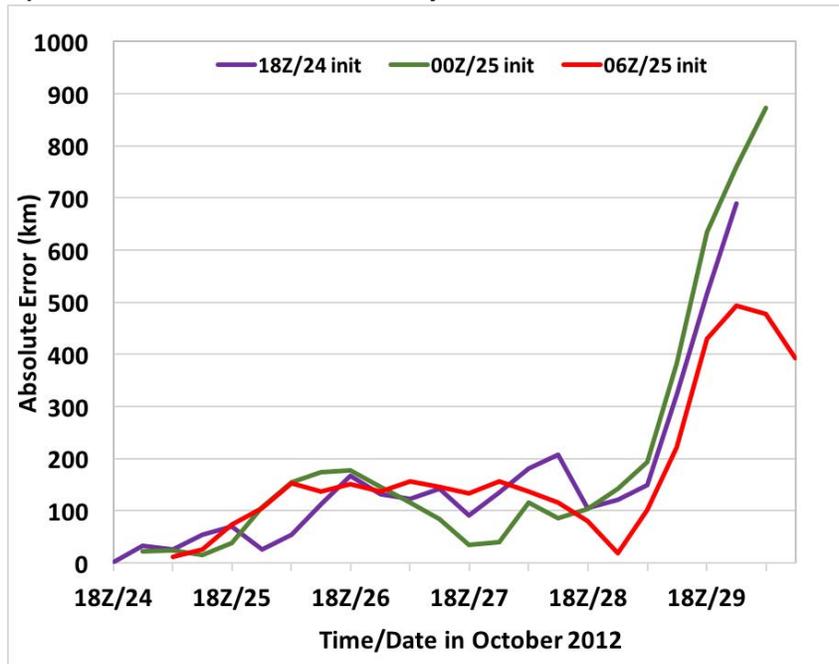


Figure 5: HURDAT (black line) and GFS (blue line) and H215 (red line) track forecasts for TC Sandy initialized at (a) 1800 UTC 24, (b) 0000 UTC 25, (c) 0600 UTC 25, (d) 0600 UTC 26, and (e) 1200 UTC 26 Oct 2012. The filled circles mark the positions at 0000 and 1200 UTC. The open circles mark the 120-h forecast position.

a) Absolute Track Error for Sandy



b) Track-Relative Error for Sandy

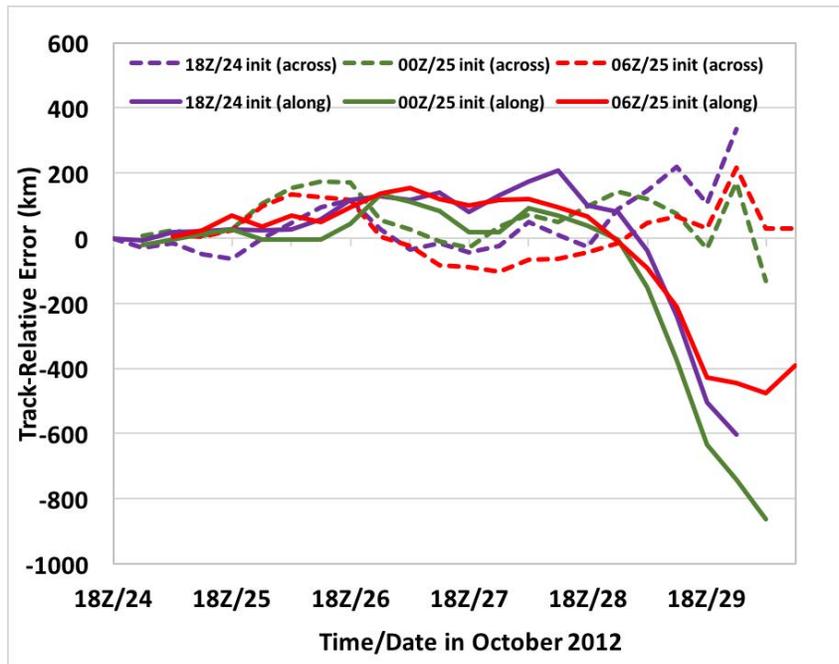


Figure 6: Time series of (a) absolute track error (km) and (b) track-relative error (km) for TC Sandy forecasts initialized at 1800 UTC 24 (purple), 0000 UTC 25 (green), and 0600 UTC 25 Oct 2012 (red). In (b), the across (along) track error is shown with solid (dashed) lines.

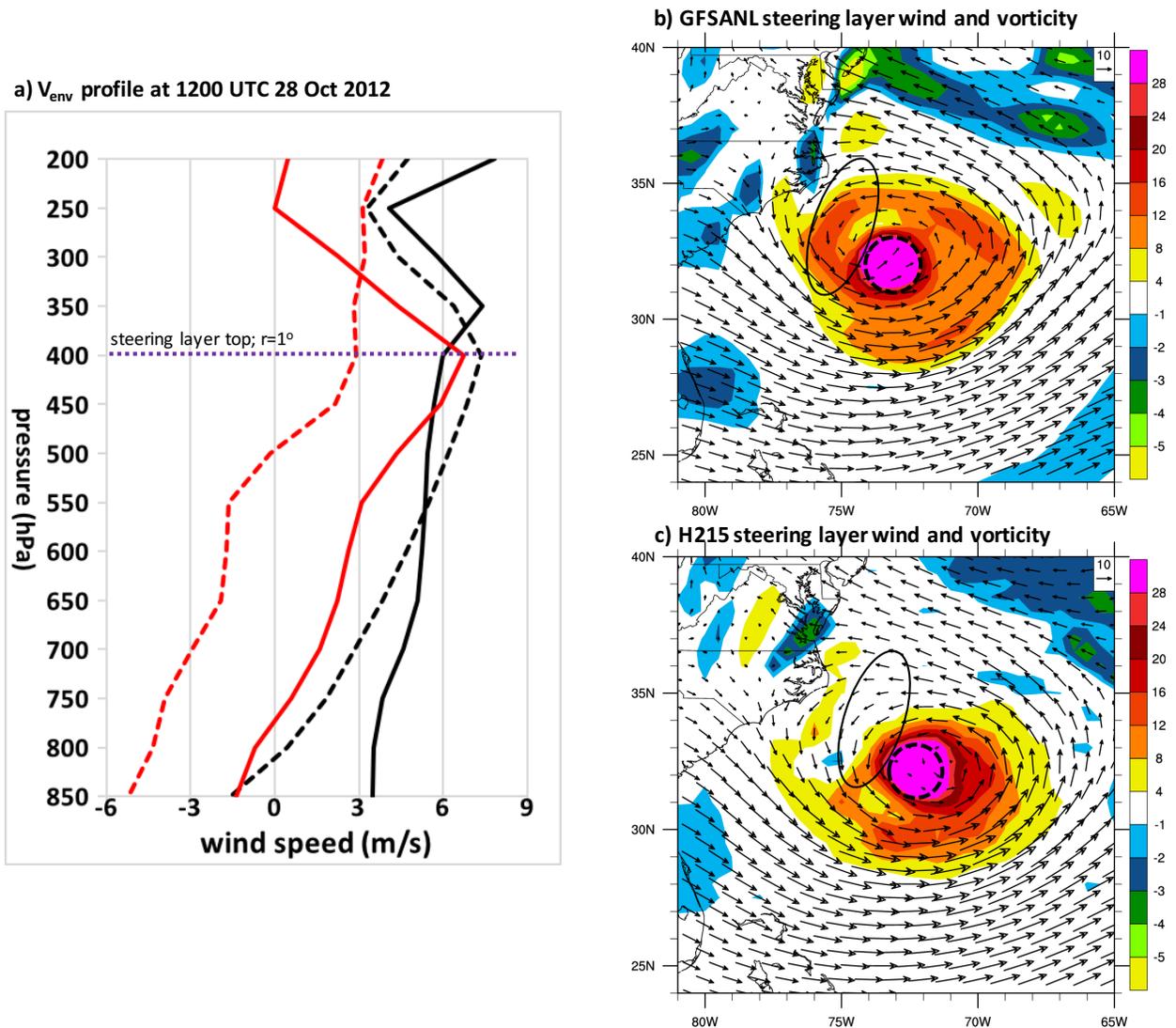


Figure 7: (a) TC Sandy V_{env} (m/s; defined using $r_0=1^\circ$) profile from the GFS analysis (black lines) and 84-h H215 forecast (red lines) verifying at 1200 UTC 28 Oct 2012. The u component is plotted using solid lines and the v component is plotted using dashed lines. (b) GFS analysis and (c) 84-h H215 forecast of V_{env} (arrows in m/s) and relative vorticity (shaded according to the color bar in 10^{-5} s^{-1}) vertically averaged in the 850–400 hPa layer verifying at 1200 UTC 28 Oct 2012. The r_0 is illustrated by the black dashed circle in (b) and (c). The cyclonic vorticity lobe to the northwest of Sandy is marked by the solid black oval.

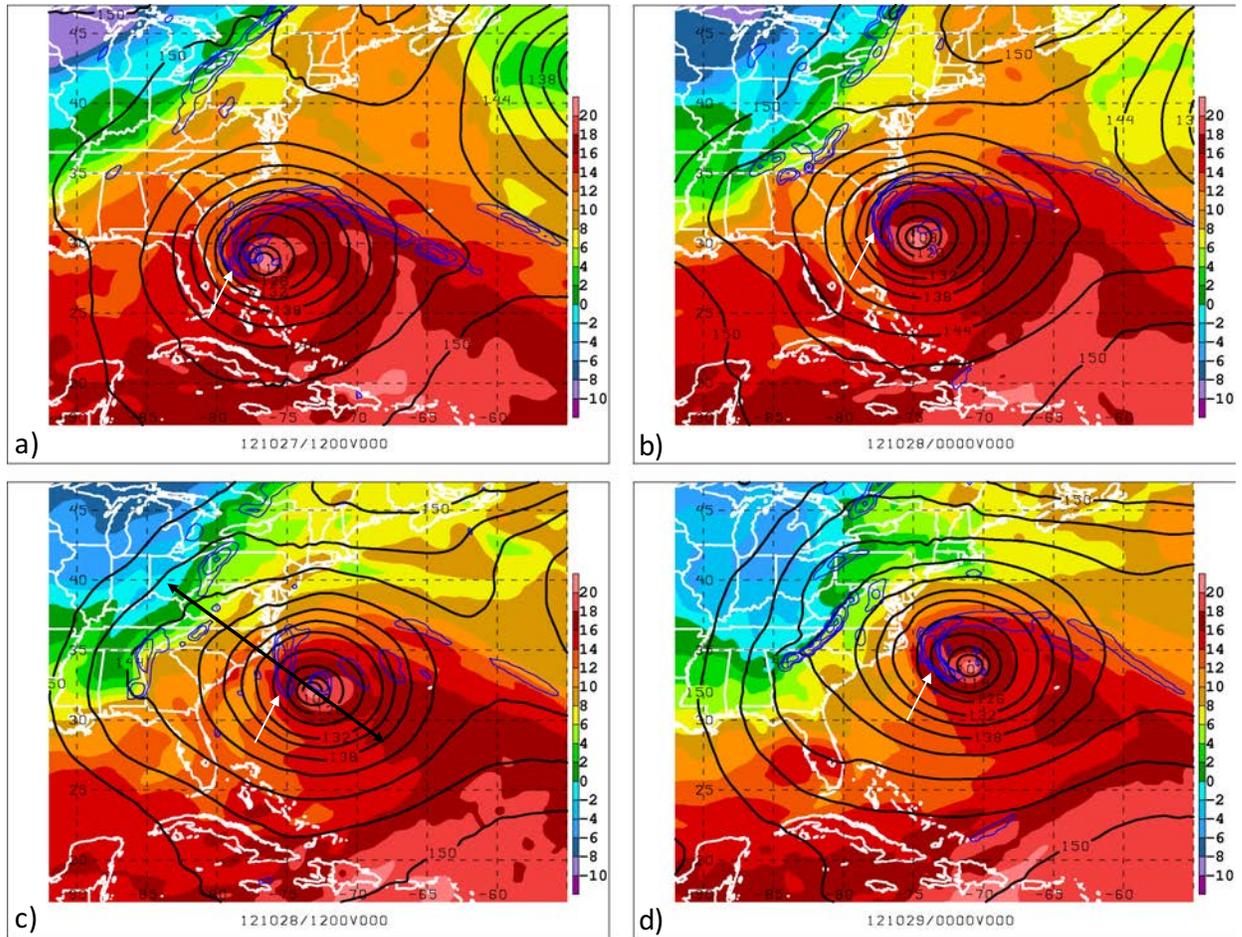


Figure 8: GFS analysis of 850 hPa temperature (shaded in $^{\circ}\text{C}$), height (solid black contours every 3 dam ≥ 132 dam and 6 dam ≤ 126 dam), and frontogenesis [solid blue contours at 1.0, 2.0, 4.0, 8.0, 16.0, and 32.0 $^{\circ}\text{C} (100 \text{ km})^{-1} (3 \text{ h})^{-1}$] at (a) 1200 UTC 27, (b) 0000 UTC 28, (c) 1200 UTC 28, and (d) 0000 UTC 29 Oct 2012. The black double sided arrow in (c) marks the cross section orientation shown in Figs. 14a–c. The white arrow marks the axis of frontogenesis northwest of Sandy.

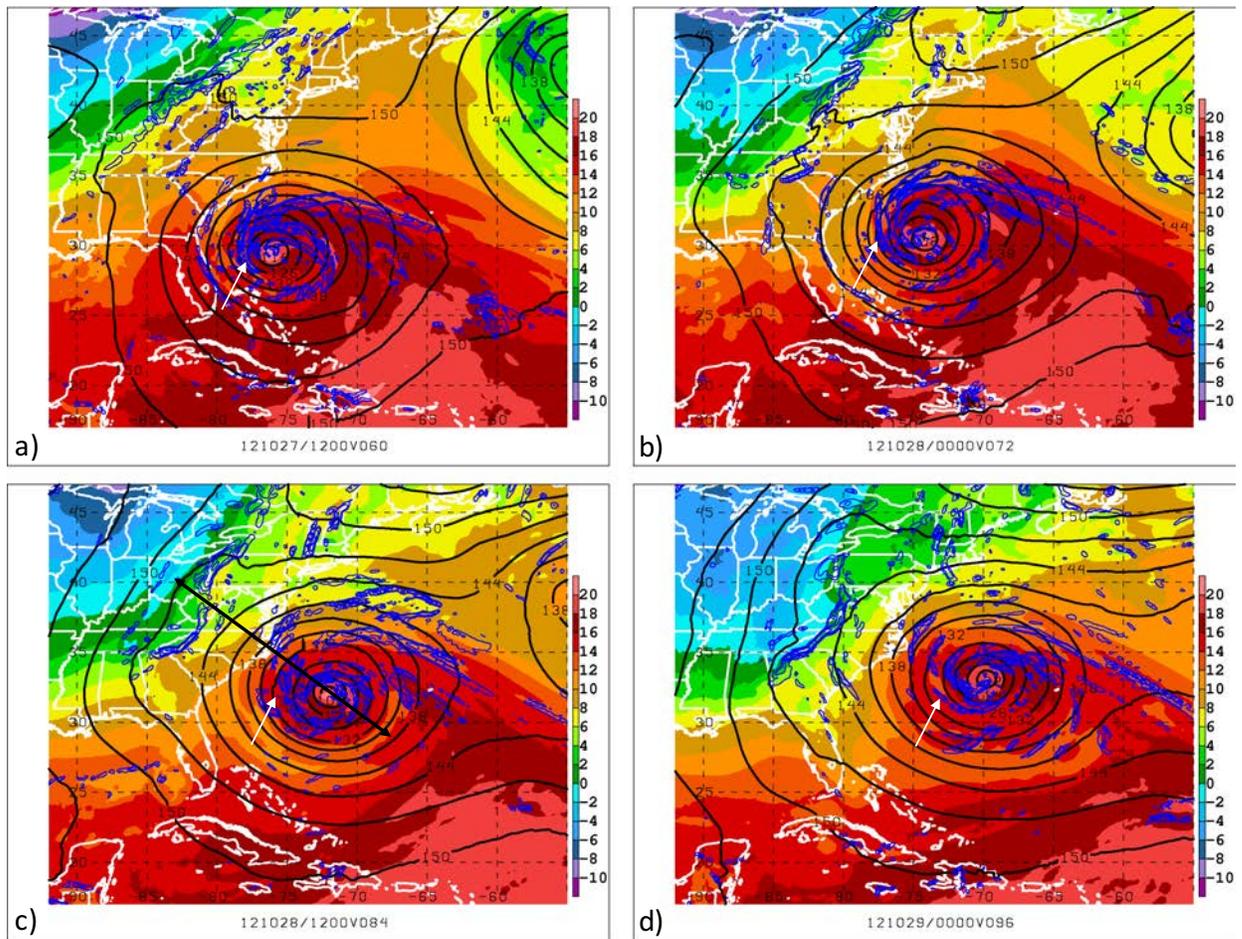


Figure 9: As in Fig. 8, except for the (a) 60, (b) 72, (c) 84, and (d) 96-h H215 forecast initialized at 0000 UTC 25 Oct 2012. The black double sided arrow in (c) marks the cross section orientation shown in Figs. 14d–f. The white arrow marks the axis of frontogenesis northwest of Sandy.

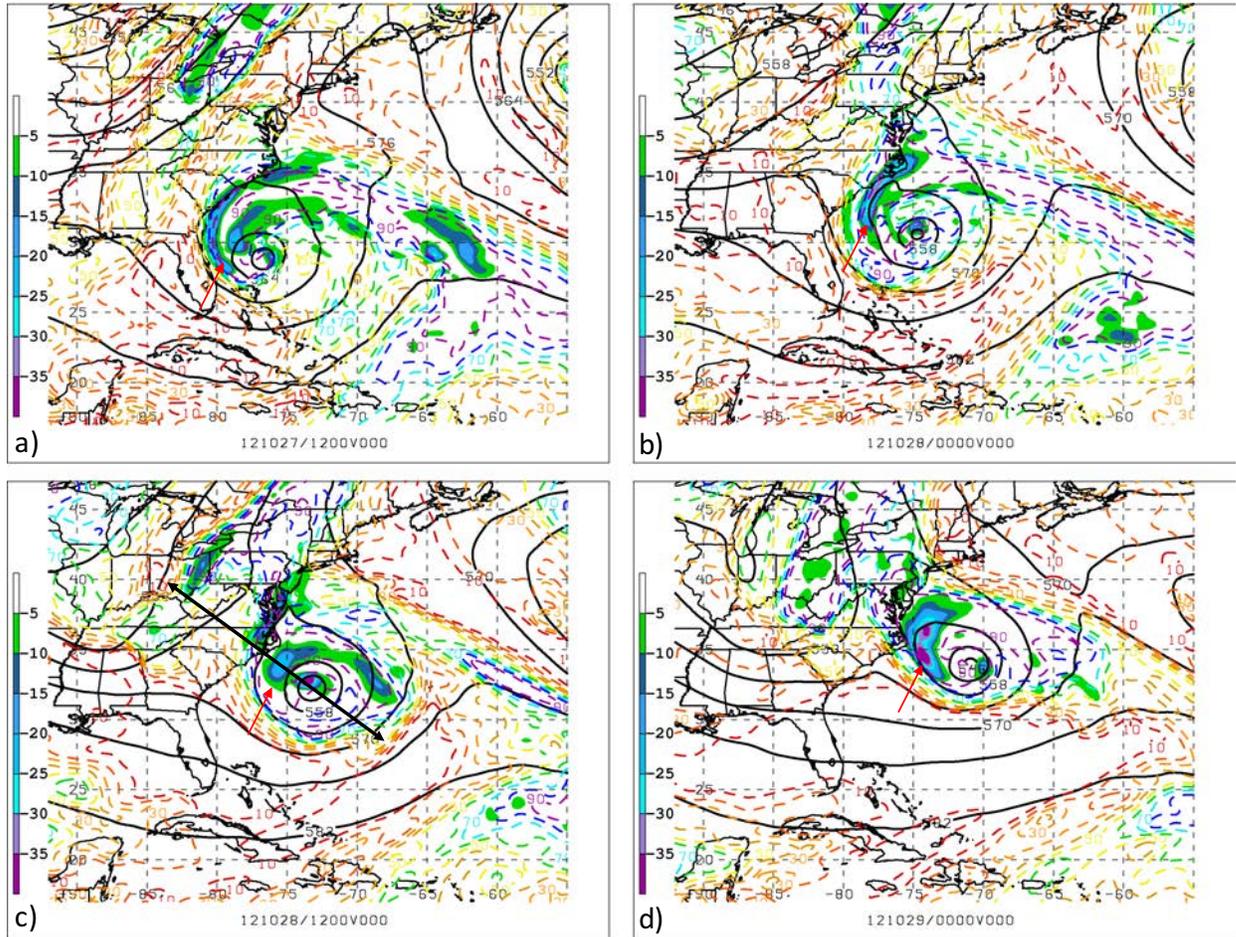


Figure 10: GFS analysis of 600–400 hPa layer mean vertical velocity (shaded in $\times 10^{-3}$ hPa s^{-1}), 500 hPa height (solid black contours every 6 dam), and 700–500 hPa layer mean relative humidity (dashed contours every 10%) at (a) 1200 UTC 27, (b) 0000 UTC 28, (c) 1200 UTC 28, and (d) 0000 UTC 29 Oct 2012. The black double sided arrow in (c) marks the cross section orientation shown in Figs. 14a–c. The red arrow marks the axis of frontogenesis northwest of Sandy.

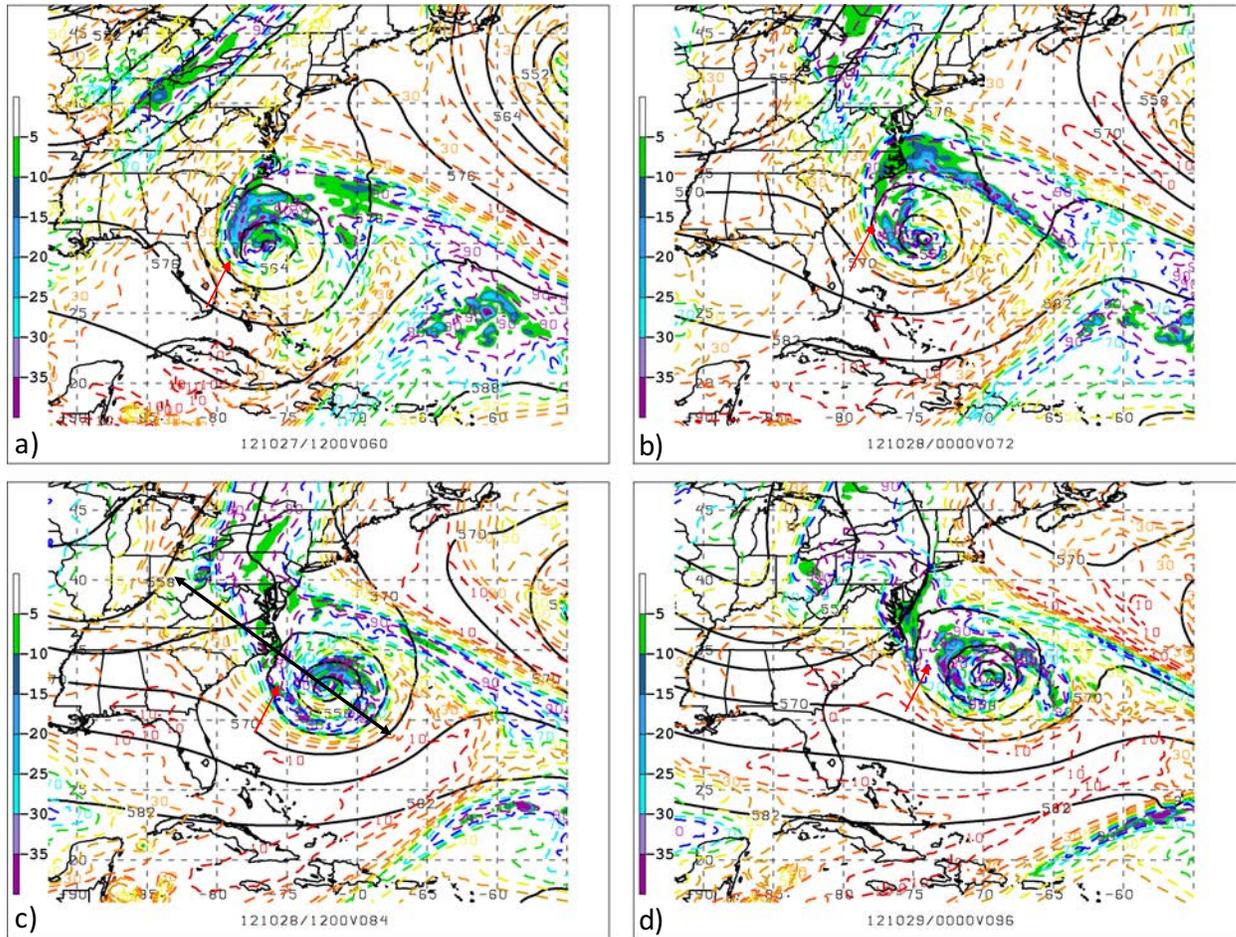


Figure 11: As in Fig. 10, except for the (a) 60, (b) 72, (c) 84, and (d) 96-h H215 forecast initialized at 0000 UTC 25 Oct 2012. The black double sided arrow in (c) marks the cross section orientation shown in Figs. 14d–f. The red arrow marks the axis of frontogenesis northwest of Sandy.

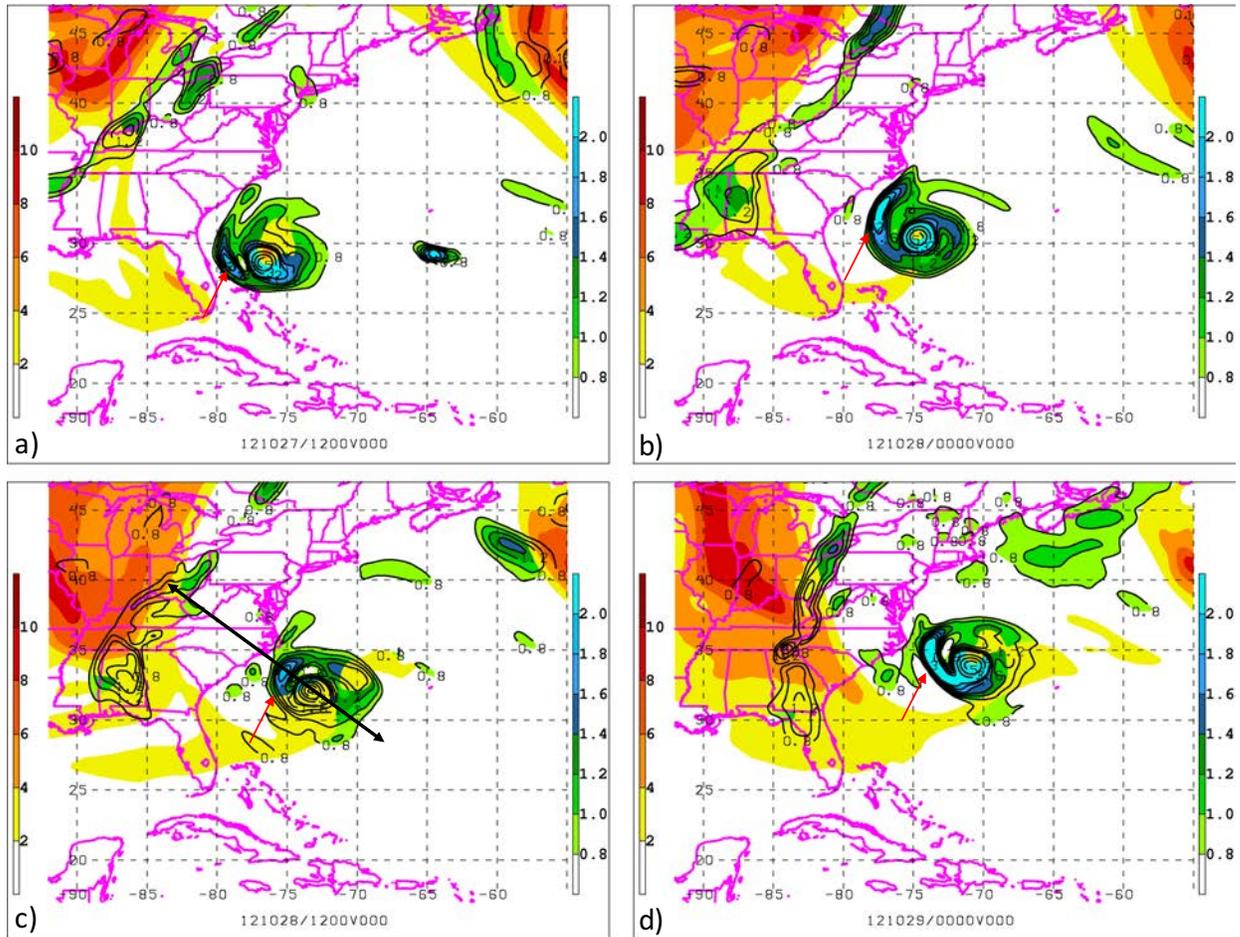


Figure 12: GFS analysis of potential vorticity (PVU) in the 300–200 hPa (shaded in warm colors) and 700–500 hPa layers (shaded with cool colors and contours) at (a) 1200 UTC 27, (b) 0000 UTC 28, (c) 1200 UTC 28, and (d) 0000 UTC 29 Oct 2012. The black double sided arrow in (c) marks the cross section orientation shown in Figs. 14a–c. The red arrow marks the axis of frontogenesis northwest of Sandy.

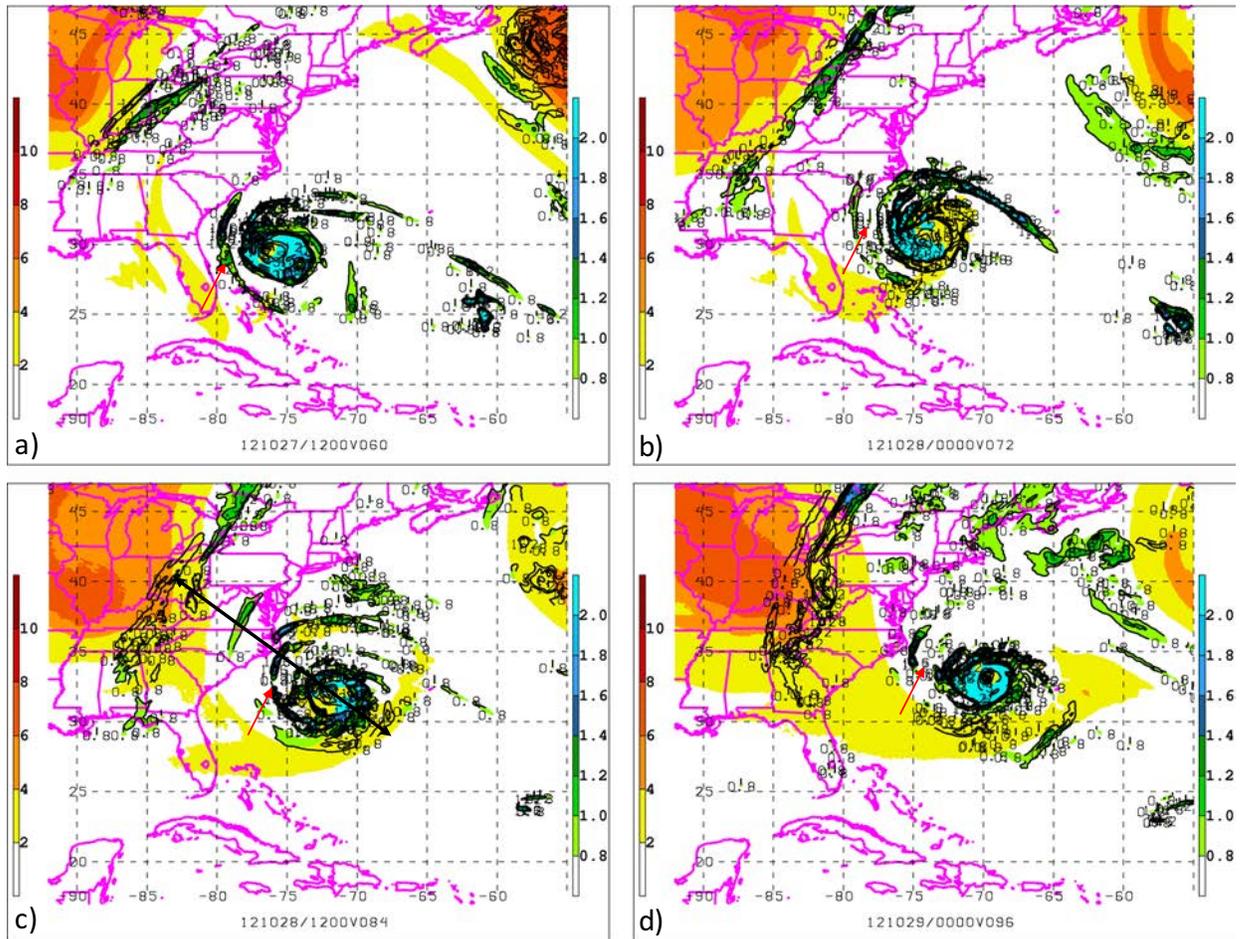


Figure 13: As in Fig. 12, except for the (a) 60, (b) 72, (c) 84, and (d) 96-h H215 forecast initialized at 0000 UTC 25 Oct 2012. The black double sided arrow in (c) marks the cross section orientation shown in Figs. 14d–f. The red arrow marks the axis of frontogenesis northwest of Sandy.

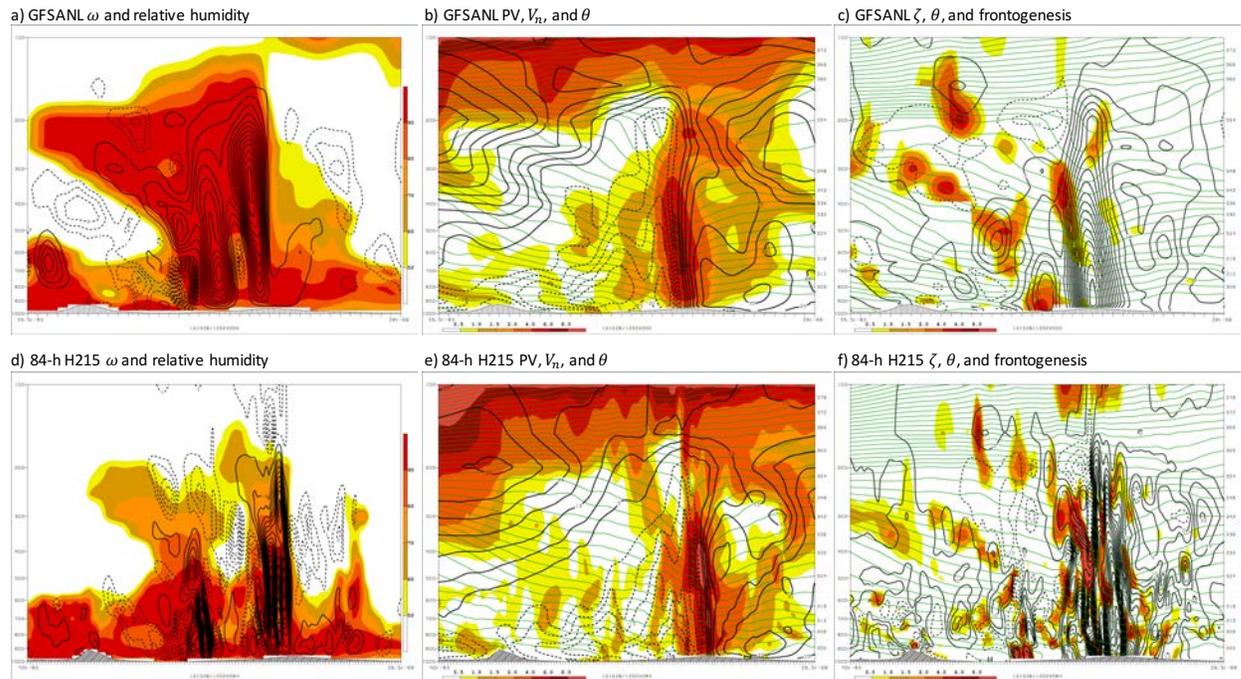
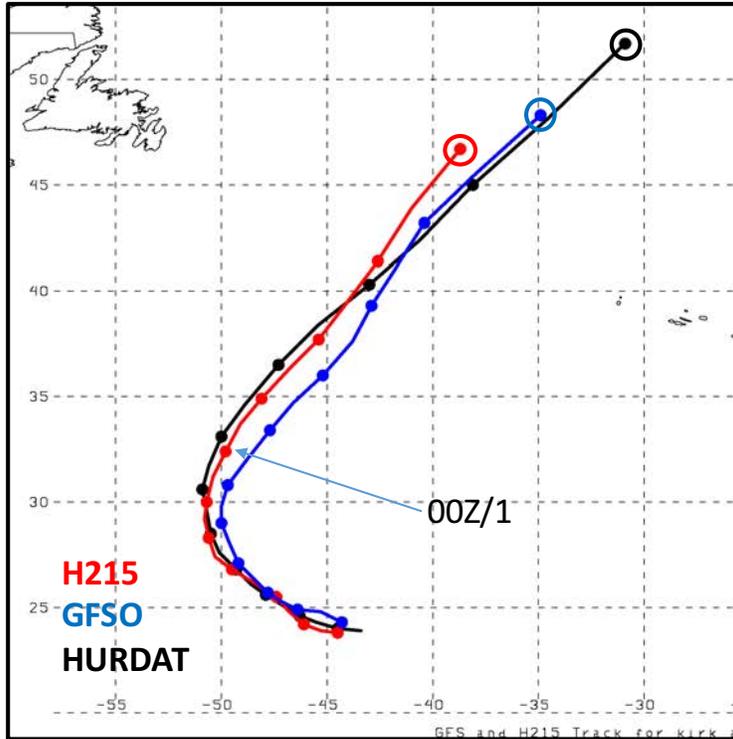
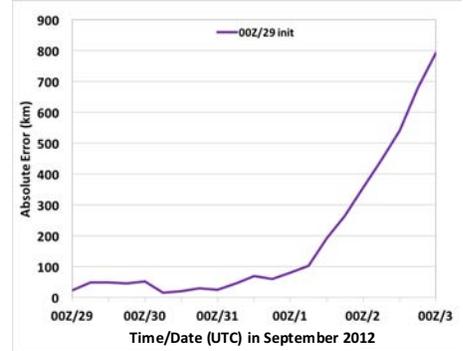


Figure 14: Vertical cross section of GFS analysis (a) vertical velocity (ascent in solid contours every $3 \times 10^{-3} \text{ hPa s}^{-1}$; descent in dashed contours every $1.0 \times 10^{-3} \text{ hPa s}^{-1}$) and relative humidity (shaded in %), (b) potential vorticity (shaded in PVU), potential temperature (green contours every 3 K), and wind component normal to the cross section (black contours every 5 m/s; zero contour omitted; negative dashed; positive solid), and (c) relative vorticity (contours every $5 \times 10^{-5} \text{ s}^{-1}$; positive and zero solid; negative dashed), potential temperature (solid green contours every 3 K), and frontogenesis [shaded in $^{\circ}\text{C} (\text{100 km})^{-1} (\text{3 h})^{-1}$] at 1200 UTC 28 Oct 2012. (d)–(f) As in (a)–(c), except for the 84-h H215 forecast verifying at 1200 UTC 28 Oct 2012. The cross section orientations for the GFS analysis and H215 forecast is shown in Figs. 12c and 13c, respectively.

a) Track forecast for Kirk initialized at 0000 UTC 29 Aug 2012



b) absolute track error



c) track-relative error

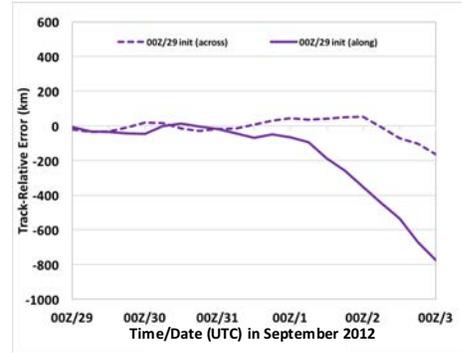


Figure 15: (a) HURDAT (black line) and GFS (blue line) and H215 (red line) track forecast for TC Kirk initialized at 0000 UTC 29 Aug 2012. Time series of (b) absolute track error (km) and (c) track-relative error (km) for TC Kirk forecast initialized at 0000 UTC 29 Aug 2012. In (c), the across (along) track error is shown with solid (dashed) lines. The open circles mark the 120-h forecast position.

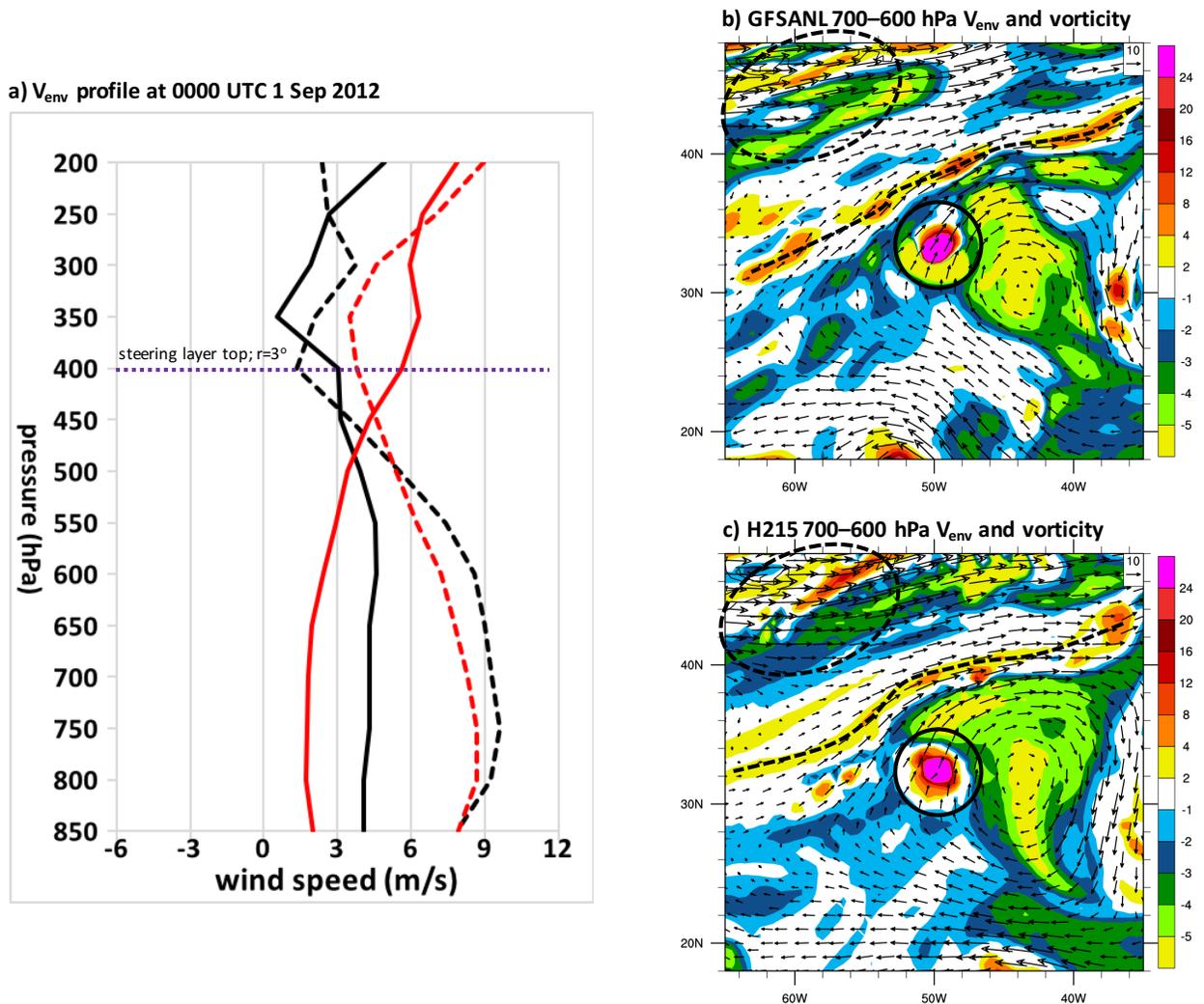
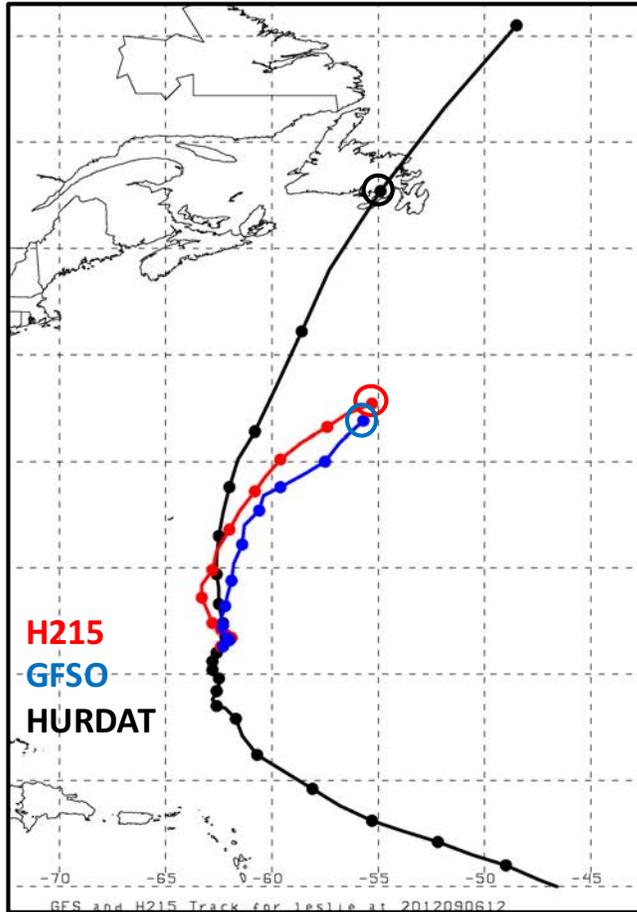
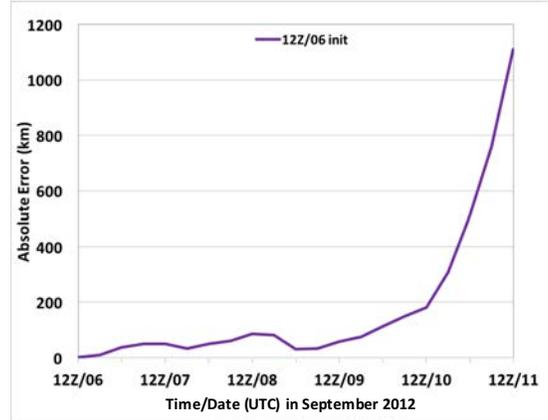


Figure 16: (a) TC Kirk V_{env} (m/s; defined using $r_0=3^\circ$) profile from the GFS analysis (black lines) and 72-h H215 forecast (red lines) verifying at 0000 UTC 1 Sep 2012. The u component is plotted using solid lines and the v component is plotted using dashed lines. (b) GFS analysis and (c) 72-h H215 forecast of V_{env} (arrows in m/s) and relative vorticity (shaded according to the color bar in $10^{-5} s^{-1}$) vertically averaged in the 700–600 hPa layer verifying at 0000 UTC 1 Sep 2012. The r_0 is illustrated by the black solid circle in (b) and (c). The cyclonic vorticity strip and midlatitude trough to the northwest of Kirk is marked by the dashed line and circle, respectively.

a) Track forecast for Leslie initialized at 1200 UTC 6 Sep 2012



b) absolute track error



c) track-relative error

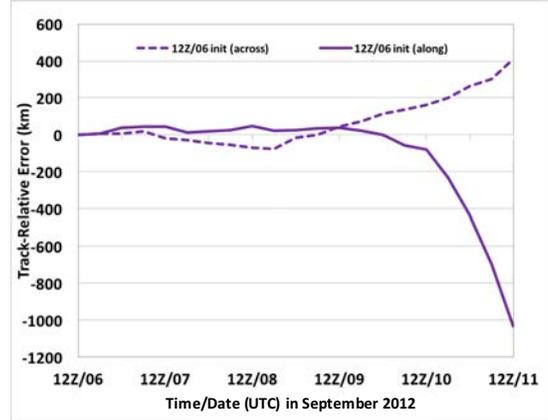


Figure 17: As in Fig. 15, except for the TC Leslie forecast initialized at 1200 UTC 6 Sep 2012.

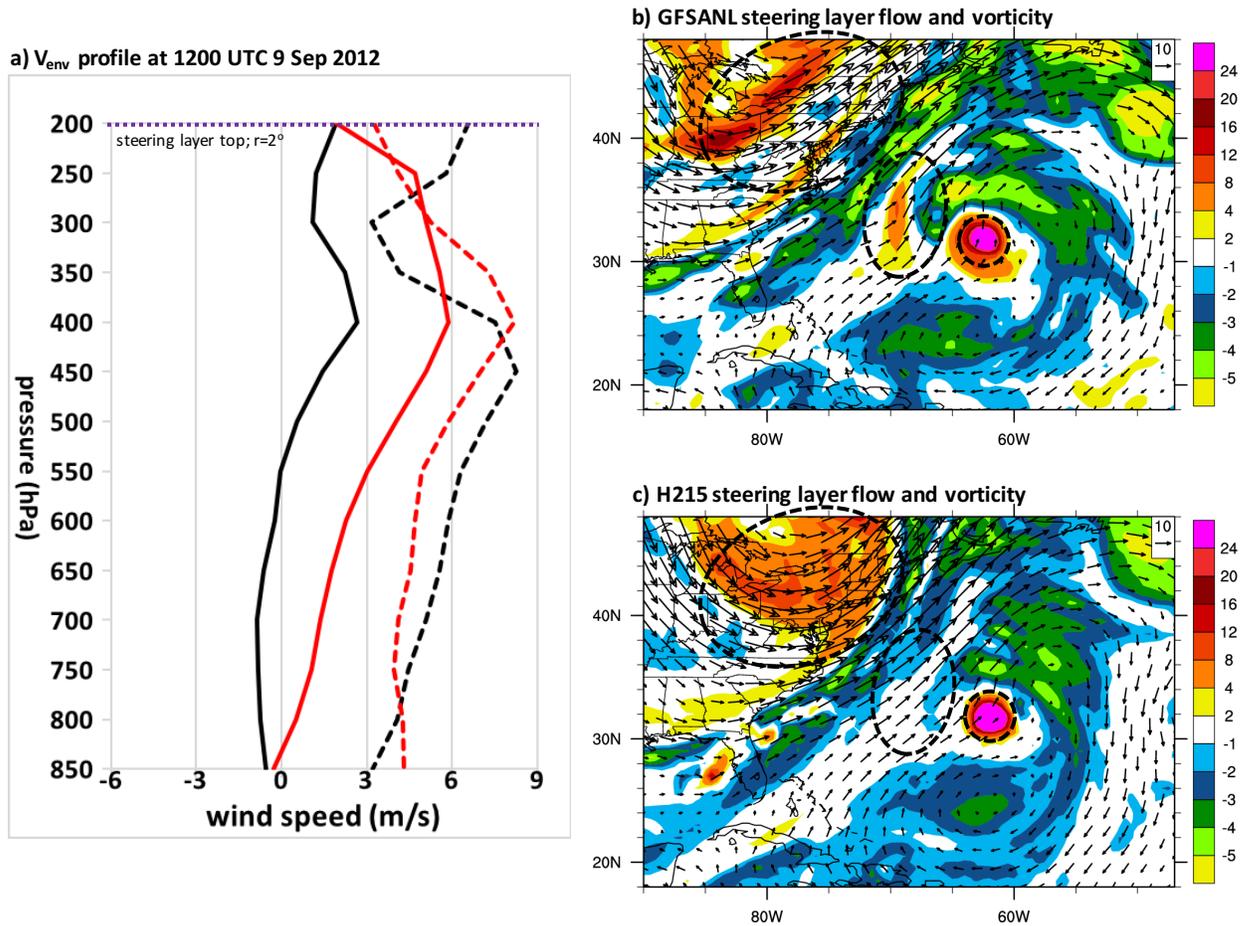


Figure 18: (a) TC Leslie V_{env} (m/s; defined using $r_0=2^\circ$) profile from the GFS analysis (black lines) and 72-h H215 forecast (red lines) verifying at 1200 UTC 9 Sep 2012. The u component is plotted using solid lines and the v component is plotted using dashed lines. (b) GFS analysis and (c) 72-h H215 forecast of V_{env} (arrows in m/s) and relative vorticity (shaded according to the color bar in 10^{-5} s^{-1}) vertically averaged in the 850–200 hPa layer verifying at 1200 UTC 9 Sep 2012. The r_0 is illustrated by the black solid circle in (b) and (c). The subsynoptic-scale vortex and midlatitude trough to the northwest of Leslie is marked by the dashed ovals.

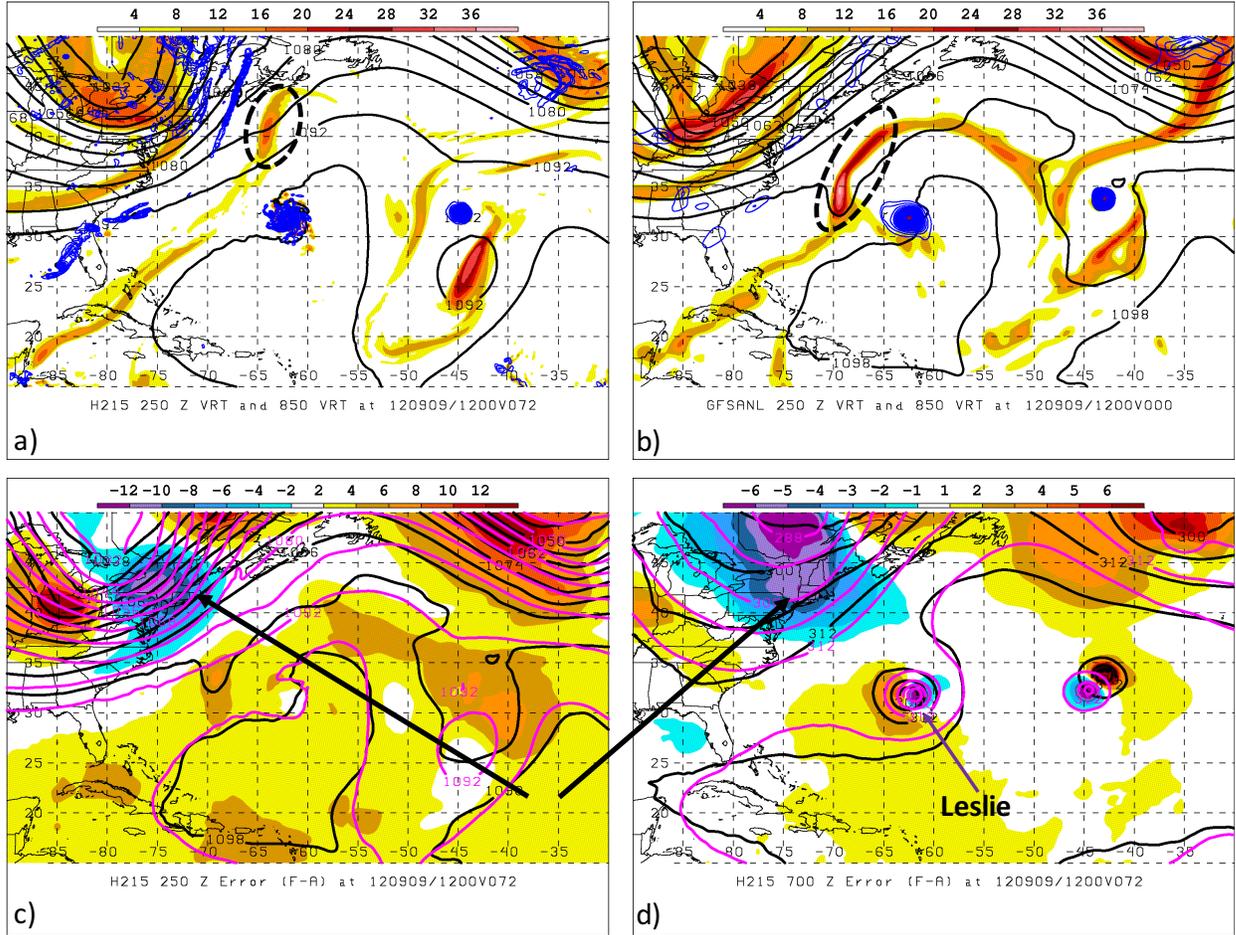


Figure 19: (a) 72-h H215 forecast and (b) GFS analysis 250 hPa height (solid black contours every 6 dam), relative vorticity (shaded in 10^{-5} s^{-1}), and 850 hPa relative vorticity (blue contours every $4 \times 10^{-5} \text{ s}^{-1}$ starting at $8 \times 10^{-5} \text{ s}^{-1}$) verifying at 1200 UTC 9 Sep 2012. (c) 250 hPa height 72-h H215 forecast (magenta contours every 6 dam) and GFS analysis (black contours every 6 dam) and height error (H215 minus GFS analysis; shaded in dam) and (d) 700 hPa height 72-h H215 forecast (magenta contours every 6 dam) and GFS analysis (black contours every 6 dam) and height error (H215 minus GFS analysis; shaded in dam) verifying at 1200 UTC 9 Sep 2012. The subsynoptic-scale vortex northwest of Leslie is marked by the black dashed oval in (a) and (b). The large midlatitude trough error is marked by a thick black arrow in (c) and (d). The position of Leslie is indicated by a thick purple arrow in (d).

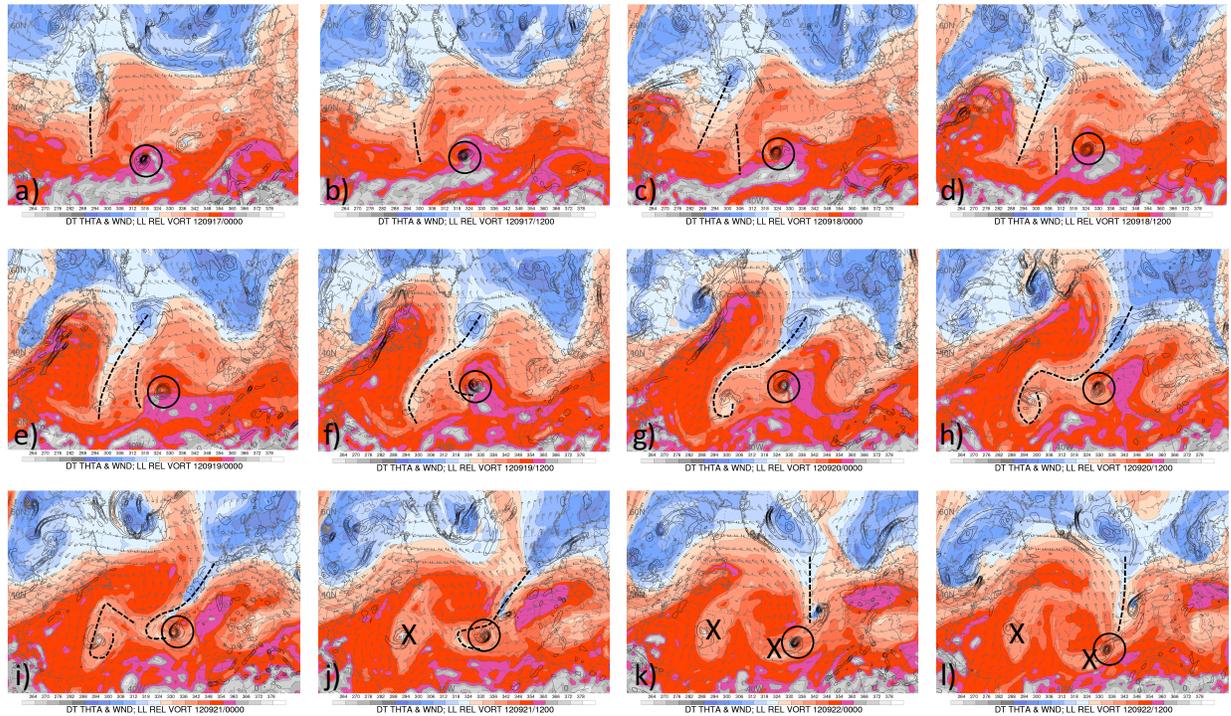


Figure 20: GFS analysis dynamic tropopause potential temperature (shaded in K), wind (standard barbs in knots), and 925–850 hPa layer mean relative vorticity (solid black contours every $5 \times 10^{-5} \text{ s}^{-1}$ starting at $5 \times 10^{-5} \text{ s}^{-1}$) at (a) 0000 UTC 17, (b) 1200 UTC 17, (c) 0000 UTC 18, (d) 1200 UTC 18, (e) 0000 UTC 19, (f) 1200 UTC 19, (g) 0000 UTC 20, (h) 1200 UTC 20, (i) 0000 UTC 21, (j) 1200 UTC 21, (k) 0000 UTC 22, and (l) 1200 UTC 22 September 2012. Nadine is marked by the solid black circle. Key upper-level troughs are marked with a dashed line, and cutoff circulations with an 'X'. Images were obtained from <http://www.met.nps.edu/~hmarcham/2012.html>.

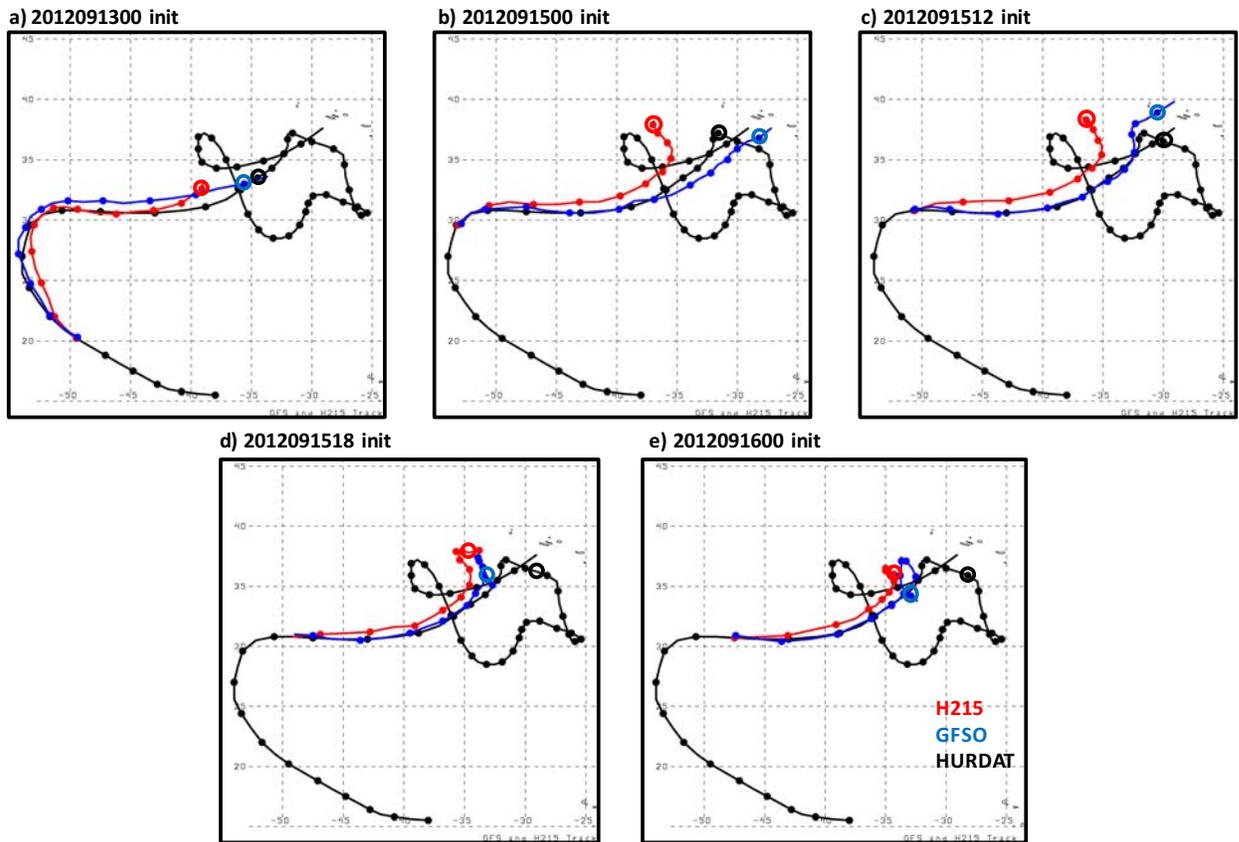
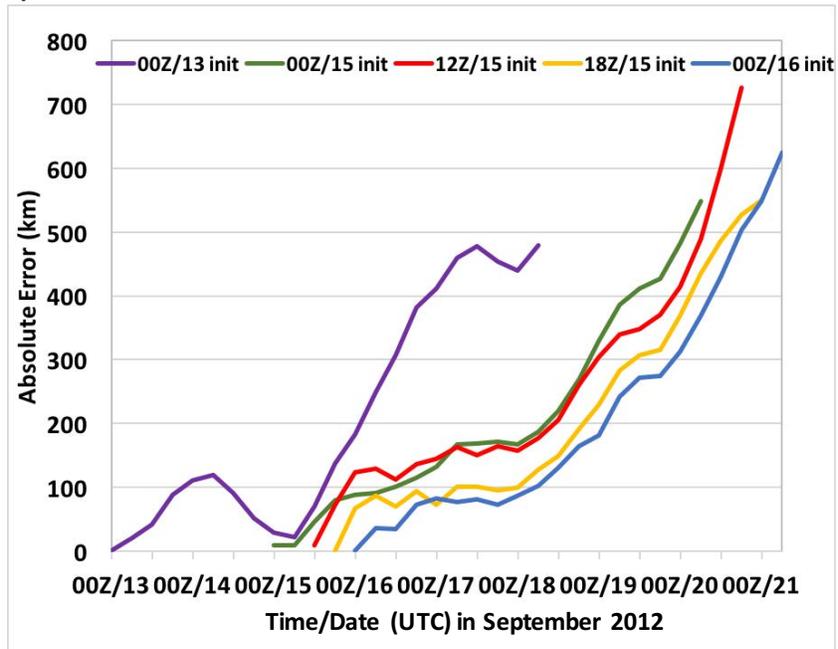


Figure 21: As in Fig. 5, except for TC Nadine forecasts initialized at (a) 0000 UTC 13, (b) 0000 UTC 15, (c) 1200 UTC 15, (d) 1800 UTC 15, and (e) 0000 UTC 16 Sep 2012.

a) Absolute Track Error for Nadine



b) Track-Relative Error for Nadine

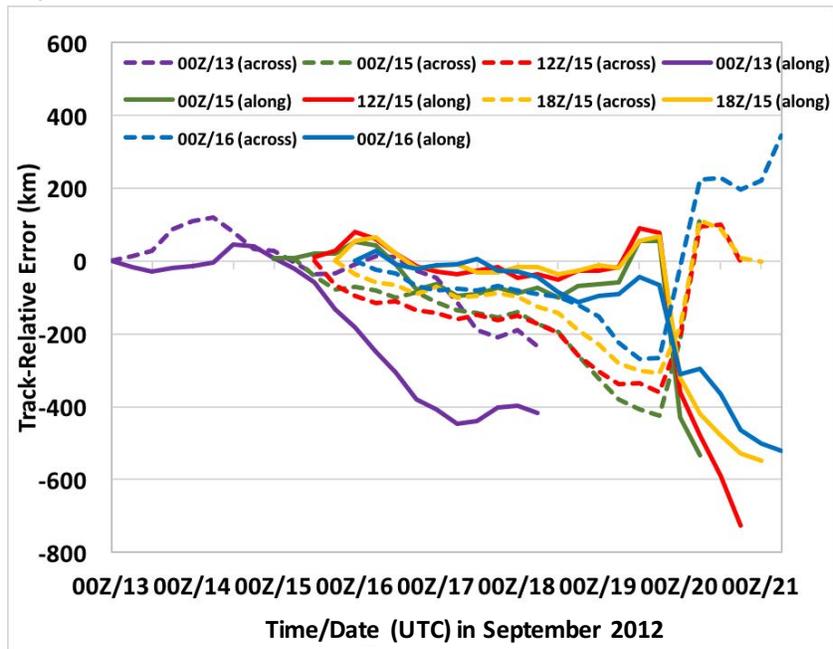
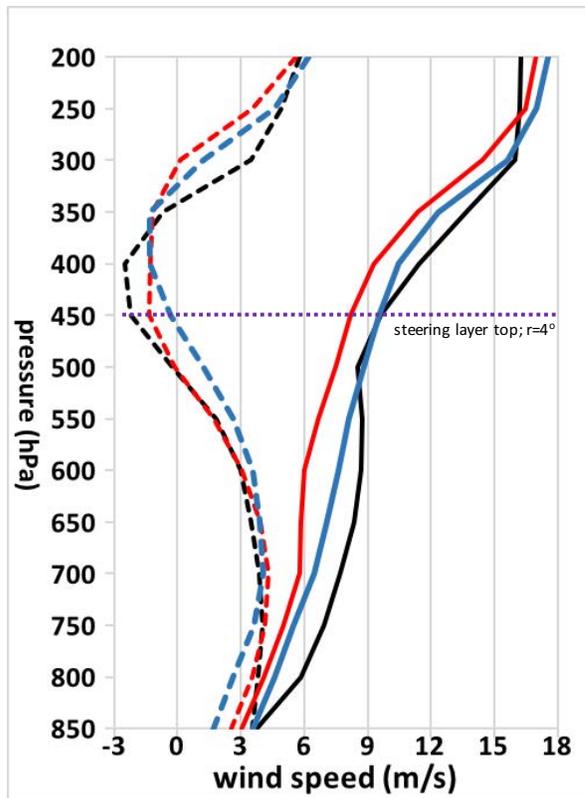


Figure 22: As in Fig. 6, except for TC Nadine forecasts initialized at (purple) 0000 UTC 13, (green) 0000 UTC 15, (red) 1200 UTC 15, (yellow) 1800 UTC 15, and (blue) 0000 UTC 16 Sep 2012.

a) V_{env} profile at 0000 UTC 17 Sep 2012



b) V_{env} profile at 0000 UTC 18 Sep 2012

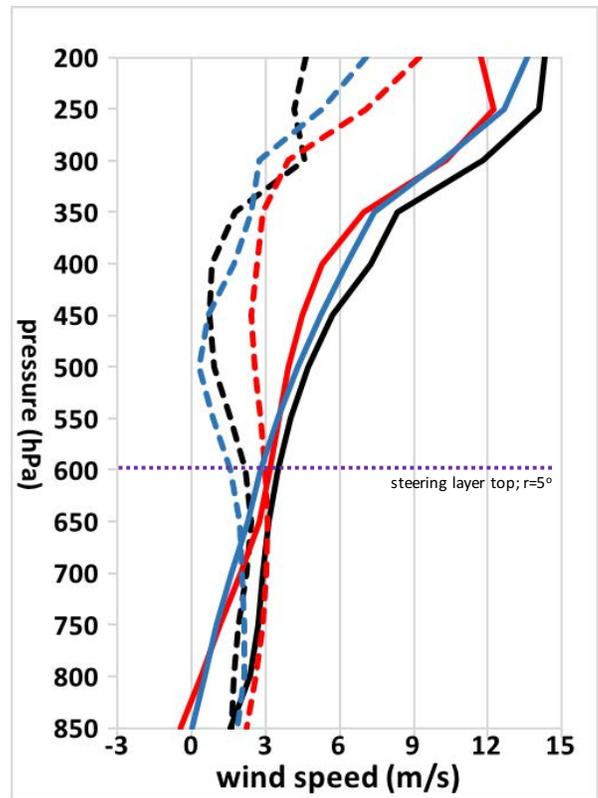


Figure 23: (a) TC Nadine V_{env} (m/s; defined using $r_0=4^\circ$) profile from the GFS analysis (black lines), 24-h (blue lines), and 48-h H215 forecast (red lines) verifying at 0000 UTC 17 Sep 2012. (b) As in (a), except for the 48- and 72-h H215 forecasts verifying at 0000 UTC 18 Sep 2012 with V_{env} defined using $r_0=5^\circ$. The u component is plotted using solid lines and the v component is plotted using dashed lines. The steering layer top is marked by a dashed purple line.

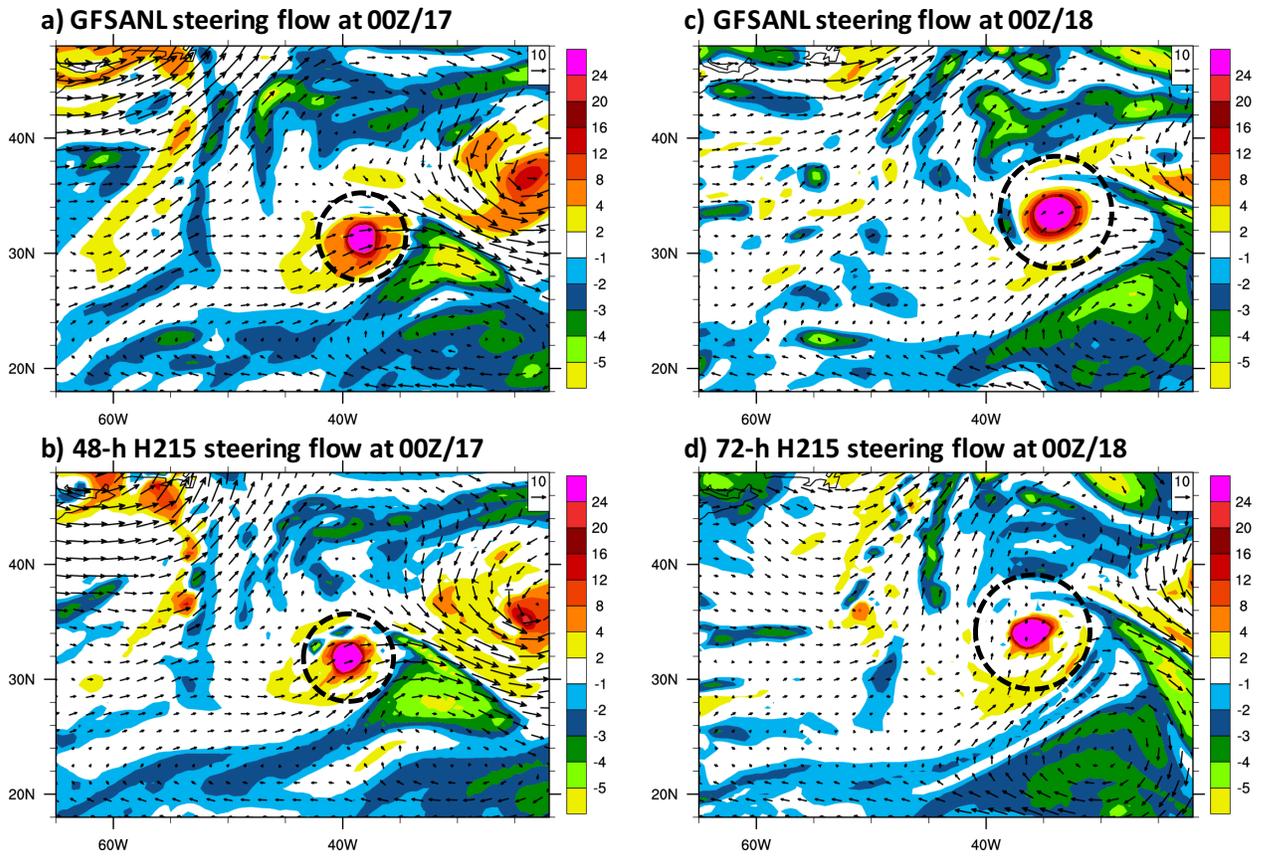


Figure 24: Steering layer (as defined on Fig. 22) mean \mathbf{V}_{env} (arrows in m/s) and relative vorticity (shaded in 10^{-5} s^{-1}) for the (a) GFS analysis and (b) 48-h H215 forecast verifying at 0000 UTC 17 and (c) GFS analysis and (d) 72-h H215 forecast verifying at 0000 UTC 18 Sep 2012. The radius r_0 is marked by the black dashed circle.

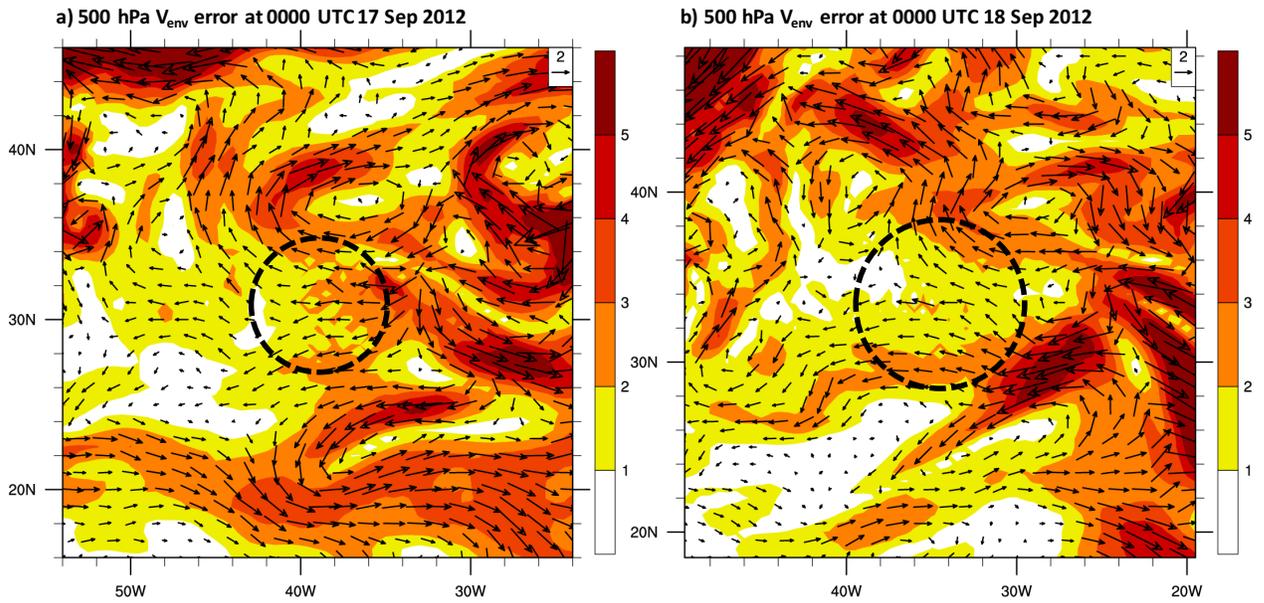


Figure 25: The 500 hPa V_{env} error (H215 minus GFS analysis; magnitude with vectors shaded in m/s) for the (a) 48-h H215 forecast verifying at 0000 UTC 17 and (b) 72-h H215 forecast verifying at 0000 UTC 18 Sep 2012. The radius r_0 is marked by the black dashed circle. The H215 forecast TC is shifted to the observed position prior to computing V_{env} error.

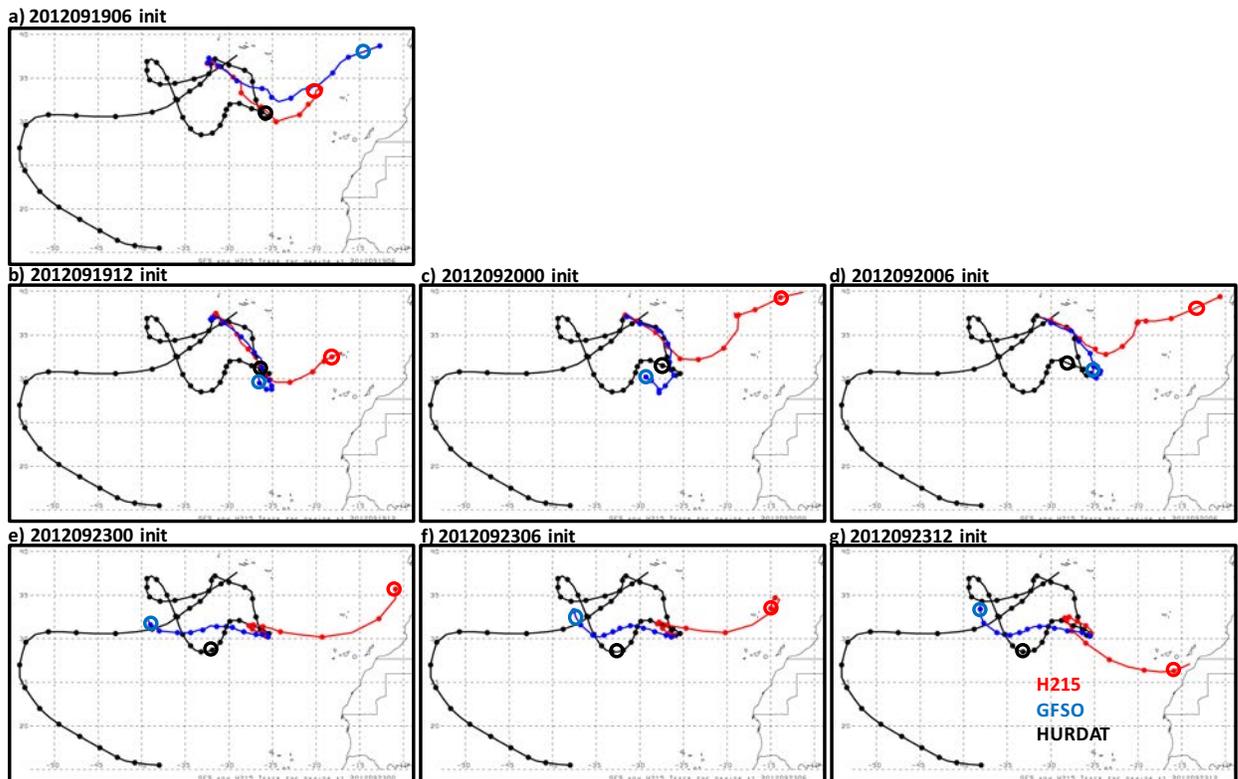


Figure 26: As in Fig. 5, except for TC Nadine forecasts initialized at (a) 0600 UTC 19, (b) 1200 UTC 19, (c) 0000 UTC 20, (d) 0600 UTC 20, (e) 0000 UTC 23, (f) 0600 UTC 23, and (g) 1200 UTC 23 Sep 2012.

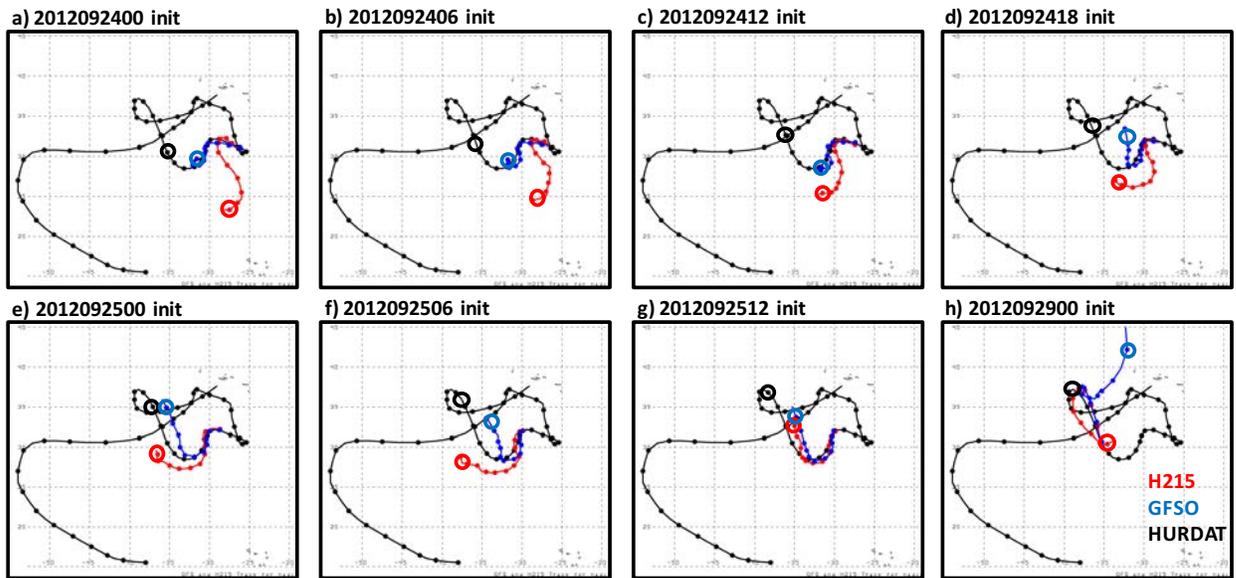
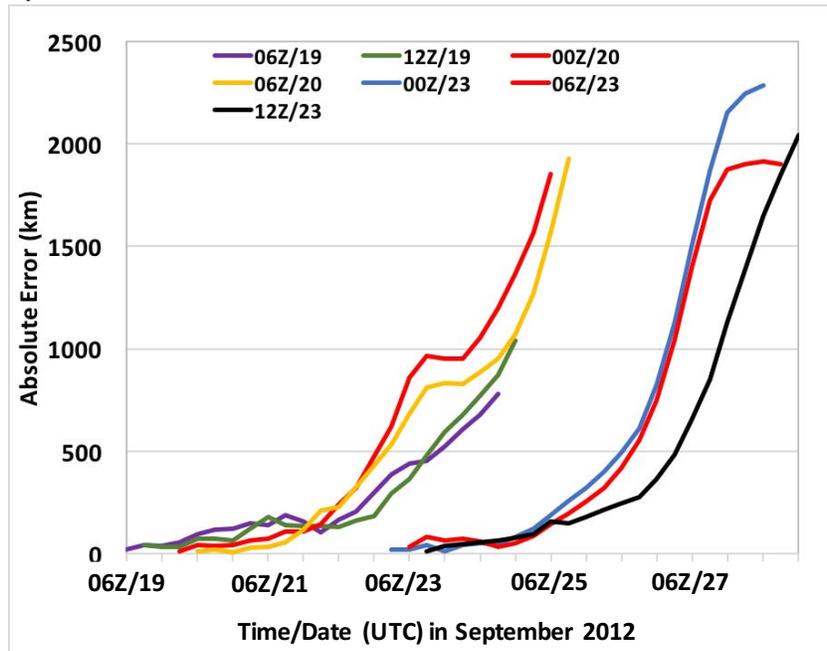


Figure 27: As in Fig. 5, except for TC Nadine forecasts initialized at (a) 0000 UTC 24, (b) 0600 UTC 24, (c) 1200 UTC 24, (d) 1800 UTC 24, (e) 0000 UTC 25, (f) 0600 UTC 25, (g) 1200 UTC 25, and (h) 0000 UTC 29 Sep 2012.

a) Absolute Track Error for Nadine



b) Absolute Track Error for Nadine

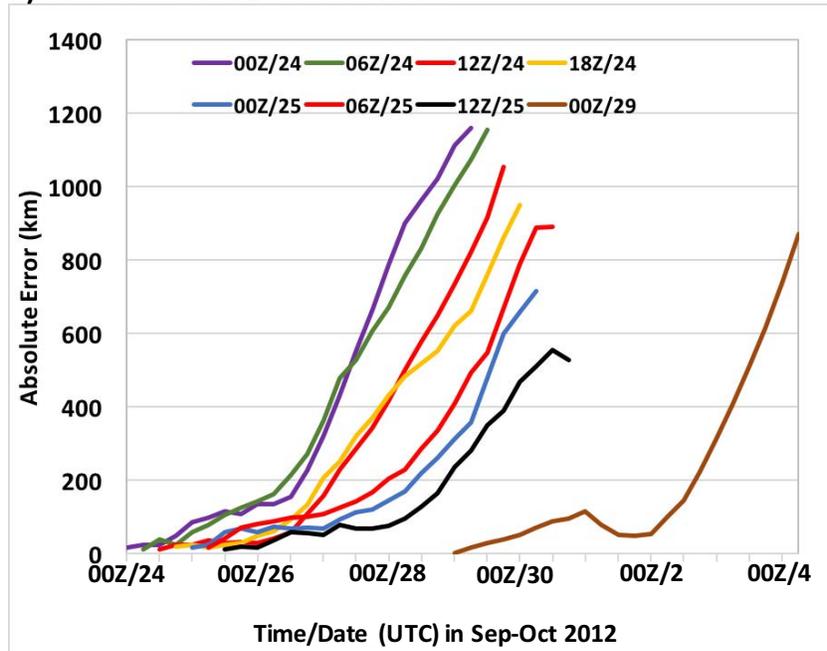
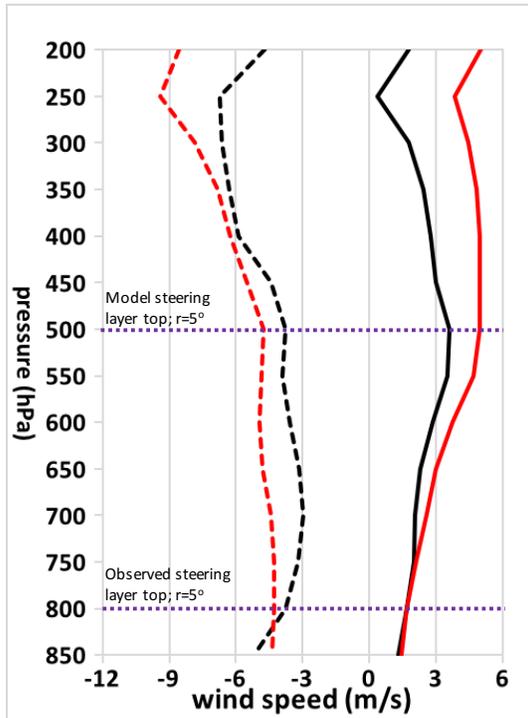


Figure 28: As in Fig. 6, except for TC Nadine forecasts initialized at (a) (purple) 0600 UTC 19, (green) 1200 UTC 19, (red) 0000 UTC 20, (yellow) 0600 UTC 20, (blue) 0000 UTC 23, (red) 0600 UTC 23, and (black) 1200 UTC 23, and (b) (purple) 0000 UTC 24, (green) 0600 UTC 24, (red) 1200 UTC 24, (yellow) 1800 UTC 24, (blue) 0000 UTC 25, (red) 0600 UTC 25, (black) 1200 UTC 25, and (brown) 0000 UTC 29 Sep 2012.

a) V_{env} profile at 1200 UTC 21 Sep 2012



b) Motion error diagnostic at 1200 UTC 21 Sep 2012 (48-h H215)

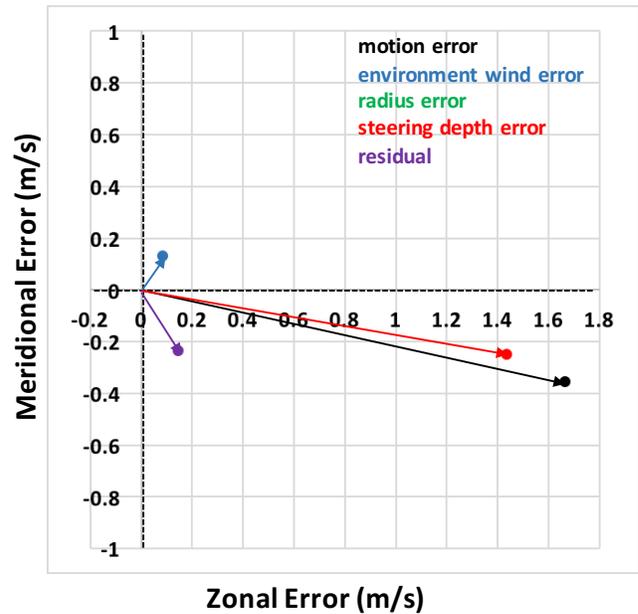


Figure 29: (a) TC Nadine V_{env} (m/s; defined using $r_0=5^\circ$) profile from the GFS analysis (black lines) and 48-h H215 forecast (red lines) verifying at 1200 UTC 21 Sep 2012. The u component is plotted using solid lines and the v component is plotted using dashed lines. The steering layer top is marked by a dashed purple lines. (b) TC motion error diagnostic equation terms (arrows in m/s) for the 48-h H215 forecast verifying at 1200 UTC 21 Sep 2012. The radius term is zero. The arrows are colored according to the key on the upper-right inset.

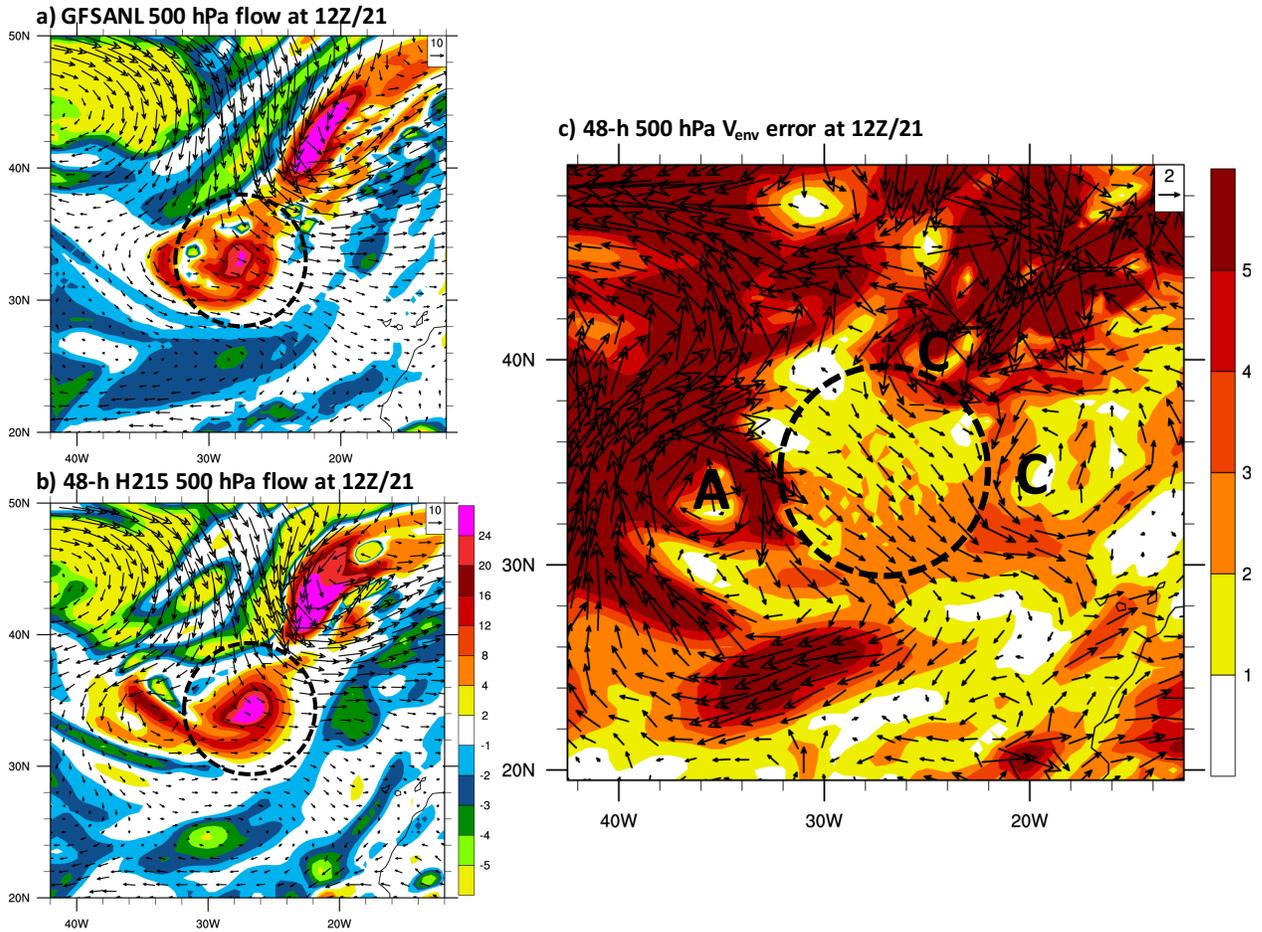
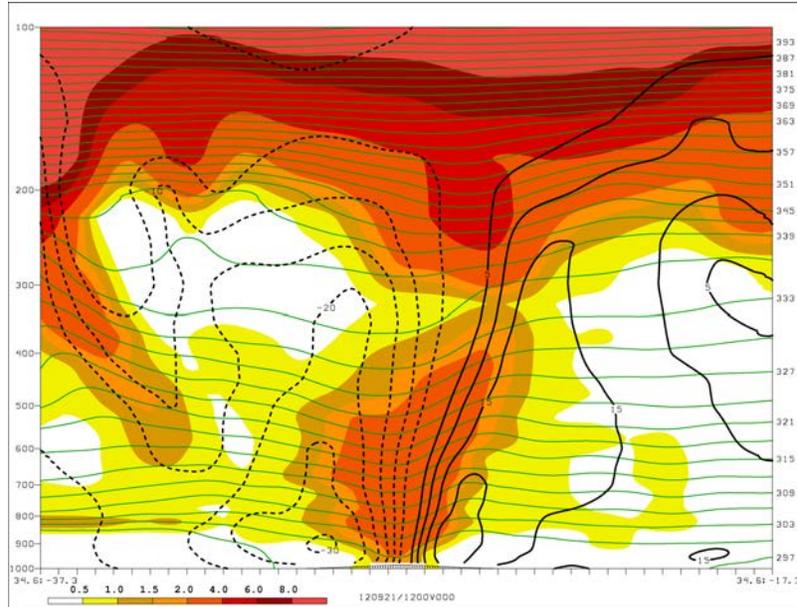


Figure 30: (a) GFS analysis and (b) 48-h H215 forecast 500 hPa V_{env} (arrows in m/s) and relative vorticity (shaded in 10^{-5} s^{-1}) and (c) 500 hPa V_{env} error (H215 minus GFS analysis; magnitude shaded with arrows in m/s) verifying at 1200 UTC 21 Sep 2012. The radius r_0 is marked by the black dashed circle. The H215 forecast TC is shifted to the observed position prior to computing V_{env} error in (c). (Anti)Cyclonic circulation errors are marked C (A) in (c).

a) GFSANL PV, V_n , and θ



b) 48-h H215 PV, V_n , and θ

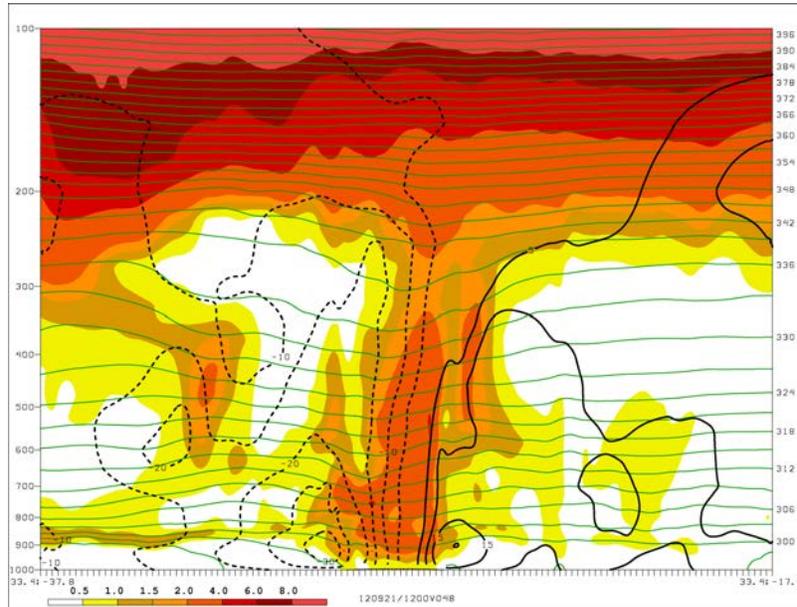


Figure 31: (a) GFS analysis and (b) 48-h H215 forecast vertical cross section of potential vorticity (shaded in PVU), potential temperature (green contours in K), and wind component normal to the cross section (black contours in m/s; positive solid; negative dashed) verifying at 1200 UTC 21 Sep 2012. The cross section orientation is zonal covering 20° of longitude centered on the TC.