

14.4 Improvement in Track and Intensity Prediction of Hurricane Florence 2018 Using WRF and HWRF Models with Ensemble Techniques

Liping Liu^{1*}, Aniya Tyson¹, Yuh-Lang Lin¹, Rick Luettich²

¹North Carolina Agricultural and Technical State University, Greensboro, NC

²University of North Carolina at Chapel Hill, Chapel Hill, NC

1. Introduction

Hurricane Florence (2018) was a powerful and long-lived hurricane that caused extensive damage in the Carolinas in September 2018, primarily as a result of both freshwater and saltwater flooding. Florence originated from a strong tropical wave that emerged off the west coast of Africa on August 30th. The wave steadily organized and strengthened into a tropical depression on the next day, then acquired tropical storm strength on September 1st. As it moved north-westward, Florence intensified rapidly on September 4–5, reaching Category 4. Then it weakened rapidly to tropical storm on September 7. Florence re-intensified to hurricane strength on September 9 and major hurricane status by the following day. It reached its peak intensity on September 11, when the east coast received warnings from NHC (National Hurricane Center). Then it weakened to Category 1 on September 13, turned southward, and made the landfall the next day near Wilmington, North Carolina. Two days later it resumed northward then north-eastward direction, dissipated, and quickly went back to sea on September 18.

As shown in Fig. 1, over the sea, Florence track looks mostly straight in northwestward, with a slight northward deflection on September 6–10. Starting on September 13, Florence stalled for two and a half days while turning southward and making landfall in North Carolina as a weakened Category 1 hurricane. However, this Category 1 storm created storm damage equivalent to a Category 5 hurricane. At its peak just before landfall, the 500-mile-wide storm had an area of tropical storm force winds that was 300 miles wide (Armstrong, 2018). The combination of strong winds, size and slow

speed prompted a wide-spread storm surge of record high (9 to 13 feet) across eastern North Carolina (Stewart and Berg, 2019). Many places received record-breaking rainfall, with more than 30 inches (760 mm) measured in some places (Fig. 2). As the ninth most-destructive hurricane to hit United States, the storm caused a total of 54 deaths, property damage and economic losses of \$24.23 billion, with \$24 billion in damages in the Carolinas alone (Tanz et al., 2019).

Strong storms such as tropical cyclones often cause flooding hazards in coastal areas due to a combination of storm surge, gusty winds, and precipitation. Accurate predictions are critical for effective disaster mitigation. While advanced storm surge / flood modeling systems have been developed, they are heavily dependent on the accuracy of the tropical cyclone (TC) prediction (Li and Nie, 2017). Especially challenging is predicting instances when abrupt variations occur in TC track and intensity.

This study focuses on the accurate prediction of track and intensity near landfalling. For the forecasting errors, we refer to the data in NHC for the annual average error in Fig. 3 for the track and Fig. 4 for the intensity. In 2018, the track error is about 33 nautical miles for 24 hours, and the intensity error is about 7 knots for 24 hours. We are using these as references to evaluate our ensemble models' performance. Our goal is to improve these predictions by combining deterministic NWP (Numerical Weather Prediction) and statistical models.

2. NWP Experiments

2.1 WRF simulations

The WRF-ARW (Weather Research and Forecast, Advanced Research WRF, denoted as WRF) is a state-of-the-art atmospheric modeling system designed for both meteorological research and weather prediction. It offers a host of options for atmospheric processes and can

Corresponding author address: Liping Liu, Department of Mathematics and Statistics, NC A&T State University, Greensboro, NC. Email: lliu@ncat.edu

run on a variety of computing platforms (Skamarock et al., 2019). WRF excels in a broad range of applications across scales ranging from tens of meters to thousands of kilometers, including meteorological studies, real-time NWP, idealized simulations, data assimilation, earth system model coupling, and model training and educational support. In studying this hurricane, we utilized WRF v4.4 and conducted numerous simulations (Liu, 2020). The varied configuration options include the domain size, resolution, microphysics scheme, boundary layer model, cumulus parameterization, sea surface temperature updating, and pressure top. The starting time varied from September 10 to 11. The model initialization and boundary data are ERA-Interim and ERA5 (<https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-interim>, or <https://rda.ucar.edu/datasets/ds627.0>). Both are from the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis.

The WRF simulated results are compared with the observation data (best track data) from HURDAT2 (<https://www.aoml.noaa.gov/hrd/hurdat/hurdat2.html>) for the track and intensity. Our study (Liu, 2020) reveals that the WRF tracks are sensitive to the simulation domain with slight improvement from the frequent SST updating and deep high pressure-top. Overall, the WRF results match reasonably well with the observation data after two days' simulation, although some error persists. Figure 5 shows the track results of some simulation cases. In Fig. 5, we can see two sets of tracks: one with a huge southward turning over the sea and made landfall in South Carolina; the other matches better with the observation, but still with big errors at the southward turning near the landfall.

2.2 HWRF simulations

The HWRF (Hurricane WRF) model is a specialized version of the WRF and is the operational backbone for hurricane track and intensity forecasts by NHC. The HWRF system includes the WRF model software infrastructure, the NMM-E (Non-Hydrostatic Mesoscale Model on the E Grid) dynamic core, the MIPOM-TC (Message Passing Interface Princeton Ocean Model-Tropical Cyclone), and the NCEP (National Centers for Environmental Prediction) coupler (Tallapragada et al. 2014). Studies show accurate predictions from HWRF for the track

and intensity of hurricanes. The HWRF package v4.0 is well wrapped in Python scripts and the components are tightly streamlined with optimal schemes (Biswas et al., 2018). The initialization and boundary data are the GFS (Global Forecast System) forecasting data (<https://rda.ucar.edu/datasets/ds084.1/>).

The HWRF simulations were set on a parent domain covering $80^\circ \times 80^\circ$ with two moving nest domains ($24^\circ \times 24^\circ$ and $7^\circ \times 7^\circ$) on a rotated latitude/longitude E-staggered grid. The center of the stationary parent domain was at the location of the initial storm, and the nest domains moved along with the storm using a two-way interactive nesting. The resolutions were set at 0.099° (about 11km), 0.033° (about 3.67km) and 0.011° (about 1.22km). We used 10 hPa for model top and 75 for vertical levels. The other schemes included SASAS (Scale-Aware Simplified Arakawa-Schubert) for cumulus parameterization, Ferrier-Aligo package for microphysics, GFS eddy-diffusivity mass flux scheme for planetary boundary layer, Monin-Obukhov scheme for surface flux, and RRTMG (Rapid Radiative Transfer Model for General Circulation Models) for radiation effects.

In HWRF simulations, we mainly varied the starting times (Liu, 2022). The individual tracks in comparison with the observation data were plotted in Fig. 6 for starting times on 09/10 and 09/11, and in Fig. 7 for starting times on 09/12 and 09/13. It can be seen from these two figures that in general the closer the starting time is to the landfall time the better the simulation result is. The tracks from the cases with starting times on 09/10 are all straight in northwest direction, while one day later (09/11) the tracks start to deflect toward north before the landfall, which can be seen clearly in case 9/11_18Z in Fig. 6. The cases starting on 09/12 showed clear southward then northward turning after the landfall. In Fig. 7, the cases starting late 09/12 start to curve/turn north before the landfall, while the cases starting on 09/13 show clear curving and turning northward before the landfall and southward after the landfall, matching very well with the observation track.

The intensity results (not shown here) are consistent with the track results. The maximum winds from the cases with early starting times on 09/10 and 09/11 show large discrepancies, while the later starting times on 09/12 and 09/13 show

small errors with the simulation results varying around the observation values up and down. Overall, the case with starting time 09/13_00Z, which is 36 hours before the landfall time, provides the best results in track and intensity near the landfall.

2.3 Error analysis

To better evaluate the performance of the simulation cases, we calculated the storm center errors (in distance) for the tracks (Fig. 8) and the errors of the maximum winds (Fig. 9) for intensity. Around landfall at 9/14_12Z, both cases with starting times at 9/13_00Z and 9/14_00Z have the error of 30km. A simple average of these two may result in a better case with a distance error 17km at the landfall. In general, on average the location distance error within 24 hours is 44km. The intensity error in Fig. 9 also shows the best two cases of 9/13 and 9/14 with small errors of 2 and 3 knots. On average the intensity error is 8 knots within 24 hours.

3. Methodology

3.1 Data collection

In this study, we retrieved and obtained data for Florence 2018 from various NWP models: WRF (one case), HWRF (one case), GFS (forecast data), ECMWF (ERA-Interim, ERA5). The GFS forecast, ERA-Interim and ERA5 are downloaded from the RDA website. The initial time for these data is 09/11_00Z. These data from different NWP models are used in our statistical model-based ensemble method to improve the accuracy of prediction, especially for track and intensity. A simple analysis of these data on track and intensity is shown in Fig. 10.

Overall, the track data from NWP models match well with the observation to some extent. Some discrepancies occur at the beginning in the WRF and ERA-Interim data. The big discrepancies occurred at the turnings before and after the landfall. As shown in Fig. 10a, the GFS forecast track stalled at the coast and never made landfall, the HWRF track went straight northwest and made the landfall 12 hours early, the WRF track followed well with the observation until half day after the landfall turning northward early. Relatively, the ERA-

Interim and ERA5 tracks matched well with the observation in Fig. 10a.

The intensity in Figs. 10b & 10c, however, showed large discrepancies. In Fig. 10b for the maximum winds, the worst case was ERA-Interim which remained 20-30m/s the whole time, the winds in ERA5 were about 10m/s higher than the ERA-Interim. Both ECMWF data sets never picked up the hurricane strength. The GFS were about 60m/s till 6 hours after the landfall when the GFS wind started to match with the observation and decreased following the observation. The WRF simulation data started with 60m/s and picked up the hurricane strength in about two days, over predicted with higher values before the landfall, then decreased along with the observation but remained above after the landfall. The HWRF simulation provided the best matching values for the hurricane intensity. With the largest error as 30m/s at the landfall, the HWRF data followed the observation in the entire simulation time. In Fig. 10c for the minimum MSLP, the worst case is again ERA-Interim with the highest values above the observation. The values from GFS forecast and WRF were comparable to each other and did not match with the observation until almost one day after the landfall. Unlike the wind result, the ERA5 pressure values were lower/better than the GFS/WRF/ERA-Interim data. The best pressure data was from HWRF, starting with almost the same value 945mb, and remaining within 10mb error in the whole simulation time.

From the above analysis, the WRF and HWRF simulations with starting time 9/11 have large errors for the prediction of landfall location but their intensity match with the observation data with little errors. The ERA-Interim and ERA5 match well on track with small error but have large errors on intensity. With large variation and discrepancies, this is a good set of data to test the regression-based ensemble techniques.

3.2 Statistical modeling

The main steps of the regression-based ensemble technique are: first interpolate the data, then separate the data into training set and testing or forecasting set, use the training set to build the regression model, test or forecast using the testing set, and lastly revise the regression

model(s) accordingly. In regression models, the usual sampling size is about 1000 data points (James et al. 2021). However, the 6-hourly observation data has only 4 data points for one day. The other data sets could be 3-hourly and hourly, which are still not enough. Interpolating the data into smaller time intervals (i.e., 3 mins) helps with the consistent time stamps for the data points and provides sufficient sampling size for regression model training. The goal is to predict landfalling, the location and the intensity. The earlier times data are used to train the regression model. Thus, we separate the data into training set covering 9/11_00Z to 9/14_00Z and testing set for the rest for 9/14_00Z to 9/15_00Z. We use the training set to train the model, then use the model to test. The observation data (interpolated) is used to supervise the data in the training stage and to verify (evaluate) in the test (forecasting) stage. Lastly, we revise the model based on the verification and evaluation results. For a new model, the last two steps (training and testing) may be repeated.

Figure 11 shows an example of interpolating the observed track data in latitude and longitude. The left is the original data, and the right is the interpolated data. The thick short blue line shows the cut off place for the training set and testing/forecasting set.

The statistical models employed in this study are multiple linear regression and regression based random forest. The data from NWP models serve as predictors, and the observed data serves as the real value of the predictant. In this study, different statistical models are built/trained separately for different variable fields. All the following numerical experiments and statistical modeling were conducted in R.

4. Results and Discussions

4.1 Latitude

When considering all 5 NWP models data as predictors, the following summary was obtained:

	Estimate	Std. Error	t-value	Pr(> t)
HWRF	-0.0154	0.009296	-1.659	0.09746 .
WRF	-0.0087	0.003232	-2.709	0.00685 **
GFS	0.1445	0.010434	13.852	<2e-16 ***
ERA-I	0.3184	0.018298	17.402	<2e-16 ***
ERA5	0.5614	0.014705	38.180	<2e-16 ***

The trained regress model is

$$Y = -0.0154X_1 - 0.0087X_2 + 0.1445X_3 + 0.3184X_4 + 0.5614X_5$$

where X_1 is for the latitude from HWRF, X_2 is for WRF, X_3 for GFS, X_4 for ERA-I, and X_5 for ERA5. The above model is called full model. In the summary table, the 3 stars besides the Pr-value for GFS, ERA-I and ERA5, indicating that these three are the statistically significant predictors.

Next, we considered only these three predictors in the regression model. We obtained the following summary:

	Estimate	Std. Error	t-value	Pr(> t)
GFS	0.1254	0.009229	13.59	<2e-16 ***
ERA-I	0.3369	0.017682	19.05	<2e-16 ***
ERA5	0.5380	0.012701	42.36	<2e-16 ***

We call this model the short model. In this short model, all the coefficients are positive, and they add up to 1. In a sense, these coefficients can be considered as the weights. It tells us that the ERA5 and ERA-I resemble the observation track very well.

We also did the simple ensemble by averaging all the ensemble members. To compare with different regression models, we employed the random forest regression model which picks up significant contributors automatically. The results are shown in Table 1. For latitude, Random Forest model gives the best result, with error 15 nautical miles. The NHC error for the track is about 33 nautical miles. The simple ensemble on the short model members also provides a good result.

4.2 Longitude

For some reason, the performance of the models on the longitude is quite different. Similar to the models for latitude, we obtained the full model with all 5 NWP models as predictors and a short model. The results of the root mean square errors are displayed in Table 2. Surprisingly, the results from random forest did not improve the MLR. The best model (the smallest error) is the simple ensemble on the full model, with an error of 22.8 n miles, comparable with the NHC error.

4.3 Intensity

Prediction of intensity is always much harder than that of track. Here is a quick show for some of the regression models on the maximum wind. Again, we have the full model and short model. Only the ERA-Interim is excluded in the short model. The summary of the full model for maximum wind of MLR:

	Estimate	Std. Error	t-value	Pr(> t)	
(Int.)	86.2848	7.034765	12.265	<2e-16	***
HWRf	0.26694	0.026322	10.141	<2e-16	***
WRF	-0.5412	0.007779	-69.567	<2e-16	***
GFS	2.57476	0.118608	21.708	<2e-16	***
ERA-I	0.02239	0.097575	0.229	0.819	
ERA5	-0.48353	0.051040	-9.474	<2e-16	***

It is clearly indicated that ERA-I is not a significant contributor in representing the maximum winds. The summary of the short model is:

	Estimate	Std. Error	t-value	Pr(> t)	
(Int.)	87.11558	6.029308	14.449	<2e-16	***
HWRf	0.26502	0.024951	10.622	<2e-16	***
WRF	-0.5405	0.007376	73.296	<2e-16	***
GFS	2.58118	0.115280	22.390	<2e-16	***
ERA5	-0.48596	0.049906	-9.738	<2e-16	***

Similar work has been conducted to the minimum SLP (Sea Level Pressure). All the results for maximum wind and minimum SLP for intensity are shown in Table 3. For both, the averaging on the short model provides the best result, comparable with the NHC error. The random forest regression improved the MLR on the maximum wind, but not for the minimum SLP. The result from the MLR on the maximum wind is comparable with the best result and NHC error, not too far off.

5. Conclusion

In conclusion, we applied simple ensemble (SE), multiple linear regression (MLR), and random forest regression (RFR) techniques. For latitude, the RFR provides the best result, improving the NHC's forecast significantly. For longitude, the SE on the full model provides the best result, improving the NHC's forecast. For intensity, the SE on the short model provides the best result, comparable with the NHC's forecast.

Overall, the hybrid of SE with MLR and RFR is recommended.

6. Acknowledgements

This research was funded by NC A&T SU and UNC CH Looking Forward Pilot Program 136069, DHS SRT for MSI program by ORISE under DOE contract number DE-SC0014664, NCAR/CISL for computing resources under Project UNCT0006, and RENCI at UNC CH for computing support.

This research was developed by DHS Summer Research Team Program for Minority Serving Institutions, administered for the U.S. Department of Homeland Security (DHS) by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between DHS and the U.S. Department of Energy (DOE). ORISE is managed by Oak Ridge Associated Universities (ORAU) under DOE contract number DE-SC0014664. This document has not been formally reviewed by DHS. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of DHS, DOE, or ORAU/ORISE. DHS, DOE and ORAU/ORISE do not endorse any products or commercial services mentioned in this publication.

7. References

- Armstrong, 2018: Post Tropical Cyclone Report...Tropical Storm Florence (Report). National Weather Service Columbia, South Carolina. Retrieved Oct. 1, 2018.
- Biswas, M. K., Carson, L., Newman, K., Stark, D., Kalina, E., Grell, E., and Frimel, J., 2018: Community HWRf Users' Guide v4.0a. NCAR: Boulder, CO, USA.
- El Gharamti, M., McCreight, J. L., Noh, S. J., Hoar, T. J., RafieeiNasab, A., & Johnson, B. K., 2021: Ensemble streamflow data assimilation using WRF-hydro and DART: Hurricane Florence flooding. *Hydrol. Earth Syst. Sci. Discuss.*, 2021, 1-31.
- James, G., Witten, D., Hastie, T., & Tibshirani, R., 2021: *An introduction to statistical learning: With applications in R*. Boston: Springer.

Johnson, T., 2021: The WRF Modeling and Ensemble Prediction of Hurricane Florence 2018. Graduate Project Report, *Department of Mathematics and Statistics, North Carolina Agricultural and Technical State University*, Greensboro, NC. May 2021.

Krishnamurti, T. N., Kumar, V., Simon, A., Bhardwaj, A., Ghosh, T., and Ross, R., 2016: A review of multimodel superensemble forecasting for weather, seasonal climate, and hurricanes. *Reviews of Geophysics*, 54(2), 336-377.

Li, J. and Nie, B., 2017: Storm surge prediction: present status and future challenges, *IUTAM Symposium on Storm Surge Modelling and Forecasting*, 2017.

Liu, L. (2020), Combined atmospheric and storm surge modeling of Hurricane Florence 2018, *DHS Summer Research Training Team for Minority Serving Institutions report*, August 2020.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., ... & Huang, X. Y., 2019: A description of the advanced

research WRF version 4. *NCAR tech. note ncar/tn-556+ str*, 145.

Stewart, S. R., and Berg, R., 2019: *Tropical cyclone report: Hurricane Florence (AL062018)*. NHC Tech. Rep., 98 pp., www.nhc.noaa.gov/data/tcr/AL062018_Florence.pdf.

Tanz, L. J., Hoffman, M. N., Dandeneau, D., Faigen, Z., Moore, Z., Proescholdbell, S., and Kansagra, S. M., 2019: Notes from the field: hurricane Florence—related emergency department visits—North Carolina, 2018. *Morbidity and Mortality Weekly Report*, 68(28), 631.

Tallapragada, V., Bernardet, L., Biswas, M. K., Gopalakrishnan, S., Kwon, Y., Liu, Q., ... & Zhang, X. (2014). Hurricane Weather Research and Forecasting (HWRF) model: 2013 scientific documentation. *HWRF Development Testbed Center Technical Report*, 99.

8. Appendix

8.1 Appendix A: Figures

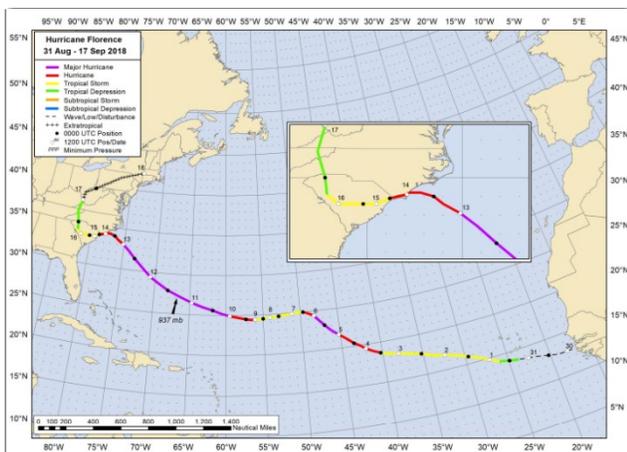


Figure 1. Hurricane Florence 2018 track from NHC

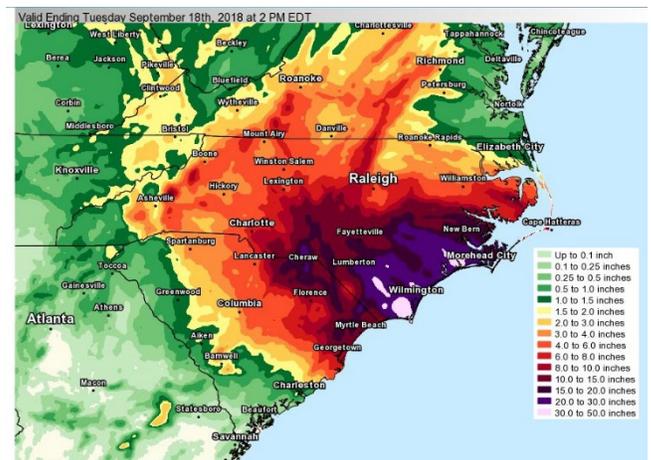


Figure 2. Observed precipitation of Hurricane Florence 2018

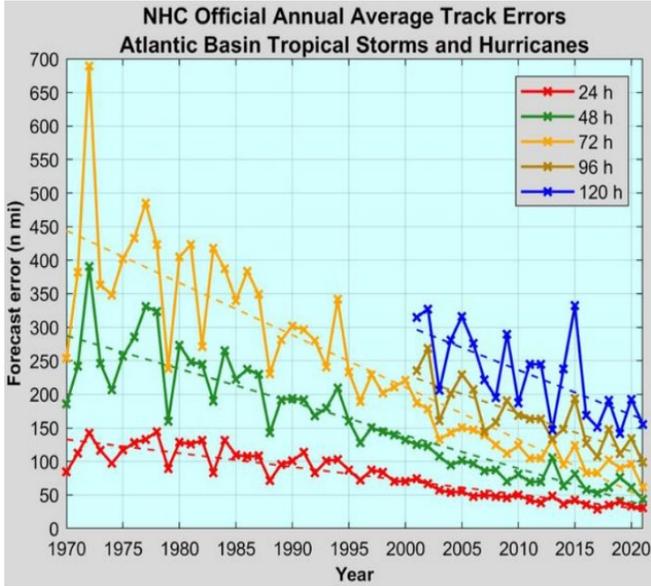


Figure 3. Track error from NHC

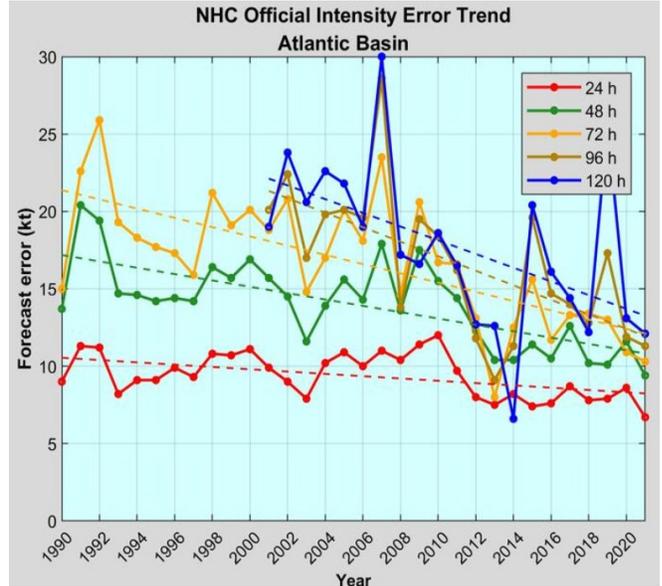


Figure 4. Intensity error from NHC

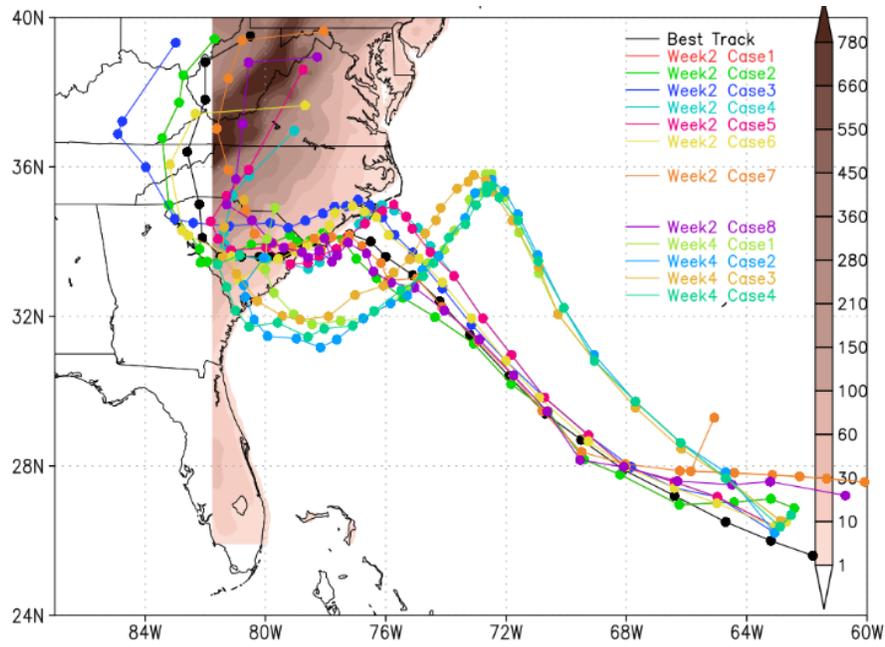


Figure 5. WRF-ARW simulations

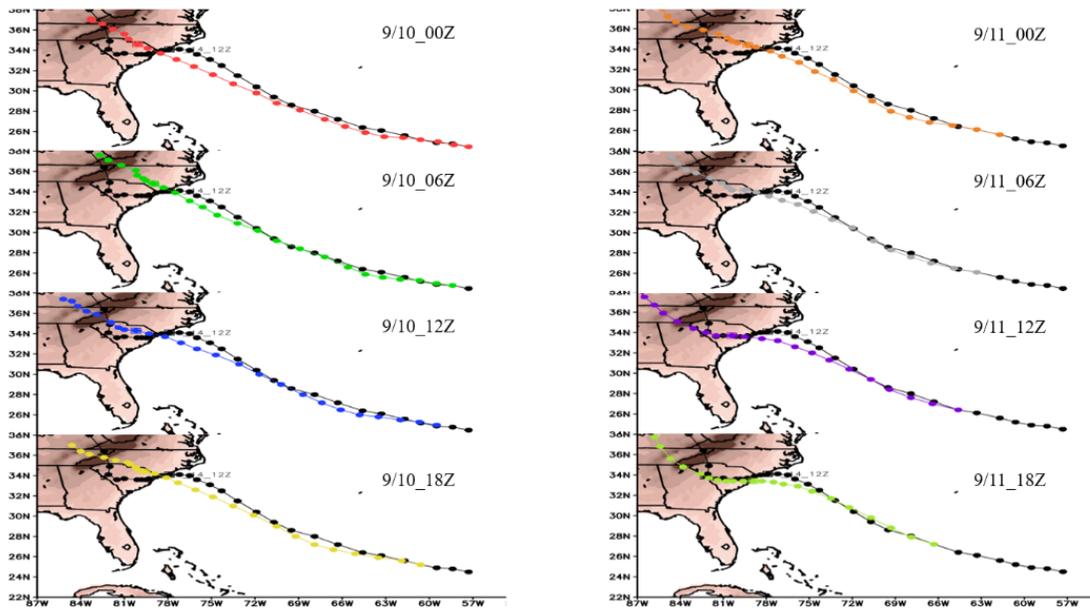


Figure 6. HWRf simulations

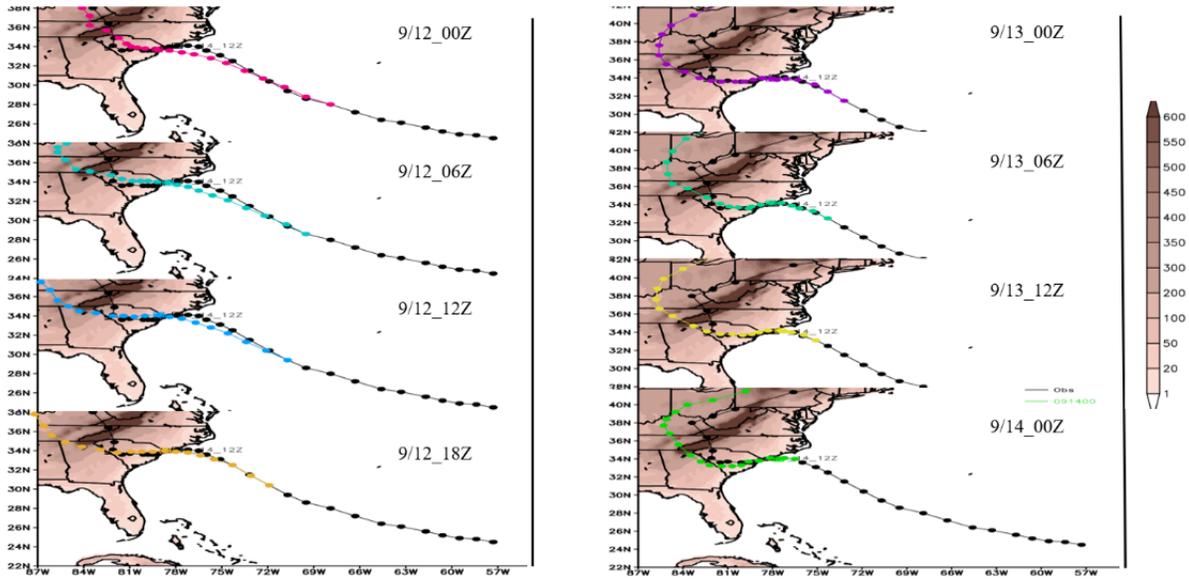


Figure 7. HWRf simulations

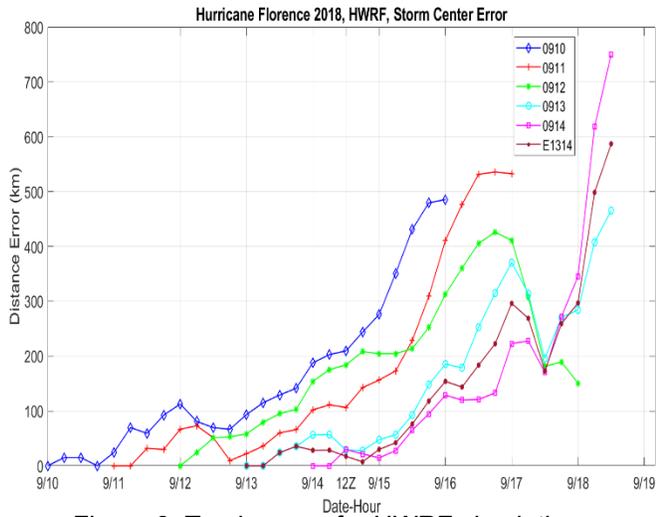


Figure 8. Track errors for HWRf simulations

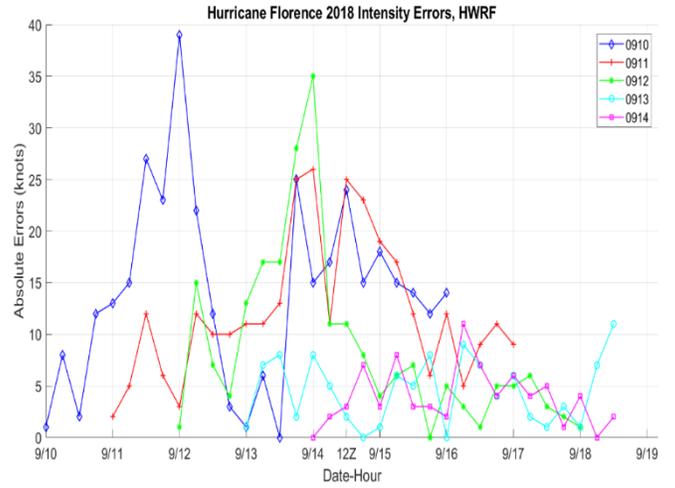


Figure 9. Intensity errors for HWRf simulations

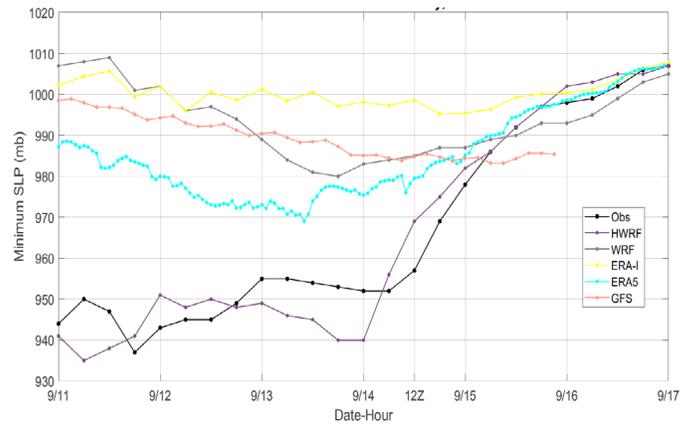
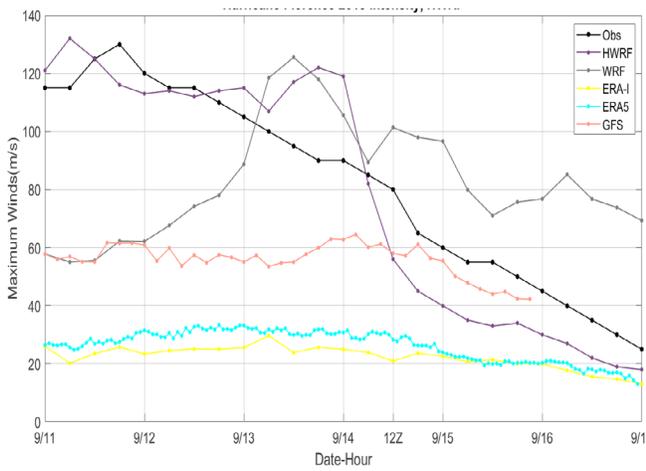
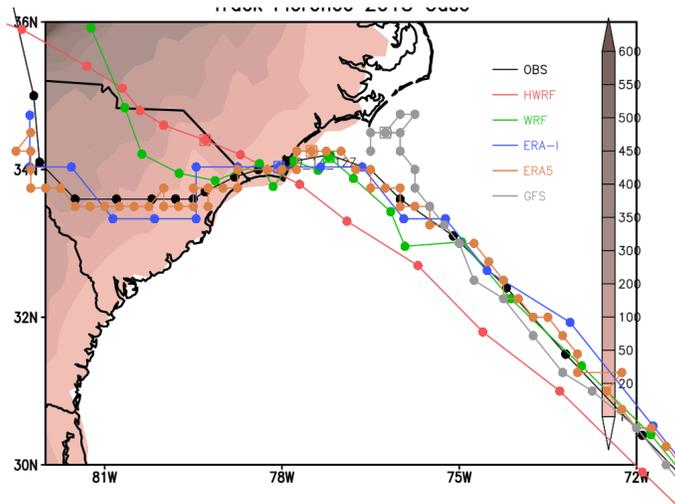


Figure 10. Track and intensity for the data collected.

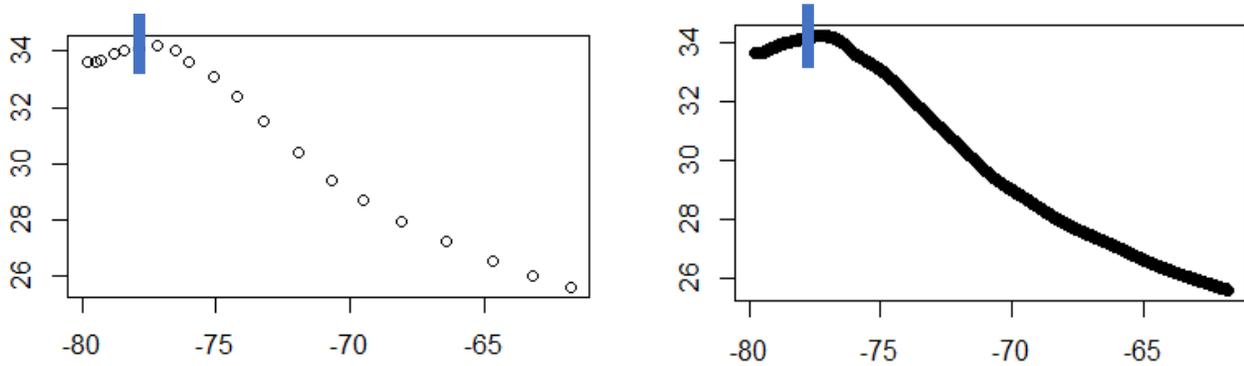


Figure 11. An example of interpolated data for the observation track. (a) The original track data; (b) The interpolated data.

8.2 Appendix B: Tables

Table 1. Errors from Various Models for Latitude

Model	MLR Full Model	Simple Ensemble Full model	MLR Short Model	SE Short Model	Random Forest Model	NHC
RMSE (deg)	0.4653696	0.4343811	0.4690583	0.3124698	0.2503998	
Distance (nautical miles)	27.94	26.08	28.16	18.76	15.05	33

Table 2. Errors from Various Models for Longitude

Model	MLR Full Model	Simple Ensemble Full model	MLR Short Model	SE Short Model	Random Forest Model	NHC
RMSE (deg)	0.6375	0.38	0.674112	0.49	1.4577	
Distance (nautical miles)	38.25	22.8	40.45	29.4	87.62	33

Table 3. Errors from Various Models for intensity

Model	MLR Full Model	Simple Ensemble Full model	MLR Short Model	SE Short Model	Random Forest Model	NHC
Max Wind RMSE (kt)	15.2449	19.52458	14.4	0.3124698	34.40646	13
Min SLP RMSE (mb)	21.98858	10.6756	21.34215	9.19942	12.36754	