https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

A Statistical Model of Atlantic Hurricane Intensity Forecast Certainty

Zander Urquhart[1]

[1]AP Research student, Lovejoy High School, Lucas, 75002, USA

Correspondence to: Zander Urquhart (zander.t.urquhart@gmail.com)

Abstract. In recent decades, the skill of hurricane forecasts have improved dramatically, thanks to the continuing progress of NWP models. In the past 30 years, track forecast errors have decreased by about 67%, but intensity forecasts have only recently begun to improve. The difficulty of forecasting rapid intensification events remains an especially difficult factor for forecasters. Data from the GFS archives and HURDAT2, a database maintained by the National Hurricane Center of best-estimate storm center locations and maximum wind speeds were collected, and all tropical systems in the North Atlantic basin that were present in both the GFS archives and HURDAT2 were analyzed. Corrections for GFS initialization error according to a transfer function found by statistical regression between GFS-reported and HURDAT2-reported max wind speeds were applied. After these corrections, the average max wind speed forecast error and its correlations with certain atmospheric and ocean conditions were analyzed. Then, using these correlations, a statistical model was constructed to predict the error range of hurricane intensity forecasts. This model demonstrated far more skill in stronger storms, and positive skill was only observed in storms Category 2 and stronger. Hurricanes Harvey and Dorian were used as test cases, with which the model was generally successful, despite problems during rapid intensification events.

1 Introduction

Over the past few decades, Numerical Weather Prediction (NWP) has evolved from an impractical experimental technique to an essential for any forecaster. Events such as hurricane landfalls can now be forecasted 4 days in advance, according to Nick Bassill, University of Albany (2014). In particular, hurricane track forecasts have improved dramatically since 1990 due to NWP models as found in an analysis by Christopher Landsea, researcher in the Atlantic Oceanographic and Meteorological Laboratory at NOAA, and John Cangialosi, hurricane specialist at the National Hurricane Center. However, NWP techniques did not improve hurricane intensity forecasts until 2010 (Cangialosi, 2020). Particularly, Rapid Intensification (RI) events, where the maximum wind speed in a hurricane increases by 35 mph in 24 hours or less, complicate forecasting, with the timing and degree of the intensification both being difficult to predict, even after decades of research (Trabing and Bell, 2020).

https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

To address this, John Kaplan and Mark DeMaria, hurricane researchers at NOAA, developed the Statistical Hurricane Intensity Prediction Scheme (SHIPS) model in 1994. It predicts a hurricane's intensity based on four factors: the difference between the current intensity and the potential intensity based on Sea Surface Temperature as determined by Kaplan's previous research; 200 mb – 850 mb wind shear, the persistence of the shear, and eddy flux convergence. It used analysis data in the VICBAR model as a basis for development and adjustment, and remains in use by the National Hurricane Center 28 years later.

NWP models compute the future state of the weather by running current conditions through a system of equations that govern fluid flow, known as the Navier-Stokes equations. NWP models return a set of parameters that represent precise predictions. They are the most common type of computer model in meteorology. NWP was first attempted with digital computers by MIT professor Jule Gregory Charney, University of Copenhagen professor Ragnar Fjörtoft, and applied mathematician John Von Neumann. Some of the model's forecasts performed poorly, but others were much more accurate. (Charney et al., 1952). This showed that NWP had promise. Currently, models such as the ECMWF model are used to predict future movement and impacts of storms. The track of Hurricane Sandy was predicted to an accuracy within 1000 km by the ECMWF, a storm whose movement was notably poorly predicted by the GFS, as analyzed by Nick Bassell, Director of R&D at the University of Albany, in 2014. In fact, models have become quite skillful at predicting hurricane tracks.

Hurricane track errors have decreased by about 67% over the past 30 years, according to a 2018 analysis by Christopher Landsea and John Cangialosi. The average three-day track error in the Atlantic has been reduced from 300 nmi in 1990 to just 100 nmi in 2016 (Landsea and Cangialosi, 2018). Likewise, average errors have dropped from 225 nmi to 75 nmi in the North Pacific (Landsea and Cangialosi, 2018). However, intensity errors have not seen such an improvement. Cangialosi found in 2020 that progress only began around 2010, with no improvement over the previous thirty years. This did not stop tropical meteorologists from finding upper bounds. Kerry Emanuel (1988), professor of meteorology at MIT, performed a theoretical analysis of hurricane energy sources formulated a relation that governed the maximum possible rate of central pressure lowering in a tropical cyclone. However, Michael Montgomery and John Persing, professor and researcher respectively at the Naval Postgraduate School, ran simulations of hypothetical hurricanes and found that this maximum was violated after 15 simulated days, so some doubt has been cast on Emanuel's formulations.

The potential for Rapid Intensification/Weakening (RI/RW) events is a significant complicating factor for tropical forecasts. According to Benjamin Trabing and Michael Bell, student and professor at Colorado State University respectively, (2020), the overall forecast error for a hurricane season is positively correlated with the number of RI and RW events. These can explain about 20% of all errors in the Atlantic and 30% in the East Pacific (Trabing and Bell, 2020). Research on the factors that lead to a higher chance of RI is ongoing. John Kaplan, Mark DeMaria, and John Knaff (2010) developed a statistical model to predict the chance of rapid intensification. It used 850-200 mb wind shear, 850-700 mb relative humidity, Kaplan's Potential Intensity (1994b), 12-hr intensity change, 200 mb (11.7 km) divergence, percentage of storm area covered by cloud

https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

tops colder than -22F, the variability of cloud top temperatures over that region, and Ocean Heat Content (Kaplan, DeMaria, and Knaff, 2010).

Kaplan and DeMaria (1994a) also developed a statistical model currently in use by the National Hurricane Center to predict the future strength of active storm systems. To address the fact that skill in intensity forecasts was "sorely lacking" (p. 220), they used the difference between Potential Intensity, referencing Kaplan's previous work, and the actual intensity, 850-200 mb wind shear, persistence, and the flux convergence of eddy angular momentum at 200 mb. It can predict storm intensity to within 8 kts at a 24-hour lead time and 13 kts at a 48-hour lead time, better than any other model that existed at the time (Kaplan and DeMaria, 1994a). In 2005, they, along with colleagues, made further improvements to the SHIPS model. They included land effects, which alone reduced the error by 15% in the Atlantic and 3% in the East Pacific. They also included new data available from satellite imagery, including cloud top temperatures and an estimate of Ocean Heat Content for the Atlantic Ocean. This data reduced error by 3.5% in the Atlantic and 7% in the East Pacific (Kaplan, DeMaria, et. al., 2005).

However, the SHIPS model does not include a certainty indicator, which would be almost as important as the forecast itself. Preparations for a hurricane landfall forecasted at 75-85 mph could be much different than between 60-100 mph. A method similar to that used for the SHIPS was applied in the pursuit of measuring the average error and bias of intensity forecasts using analyses from the GFS 0.25 Degree model as a basis for my statistical analysis.

2 Data and Methods

2.1 Hurricane Location and Strength Data

The National Hurricane Center's HURDAT2 database, which contains the storm center location, maximum sustained winds, and minimum central pressure of every Atlantic tropical cyclone from 1851 to 2020 at 6-hour intervals, was used to collect ground truth data on the location and intensity of storms.

2.2 Hurricane Forecast and Surrounding Atmosphere Data

For the forecasted storm intensity and the of the surrounding atmosphere, 0-hour and 6-hour forecast files from the GFS 0.5 Degree model in GRIB2 format from the UCAR-RDA database (NCEP, 2015) were downloaded for all 6-hour intervals in a time period from 15 Jan 2015 to 31 Dec 2020 for which a tropical cyclone was present in the Atlantic according to the HURDAT2 database. All data that fell between 65 N and 5 S in latitude and 110 W and 10 E in longitude, as depicted in Figure 1, was trimmed to minimize file size. The list of fields from the GRIB file is provided in Table 1.

https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

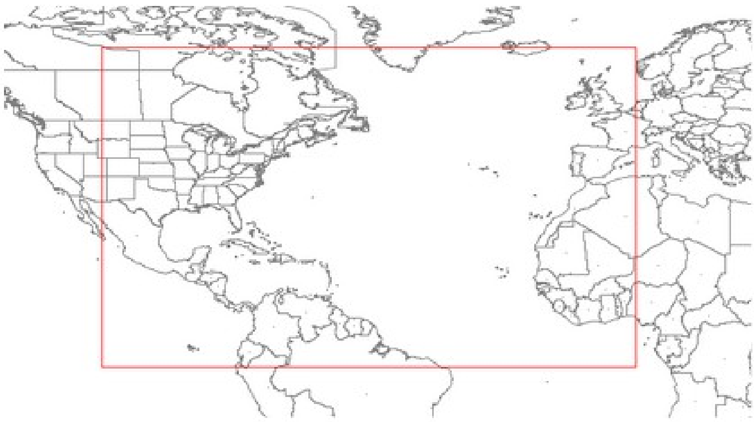


Figure 1. Study Area Outlined in Red

Field name in GRIB2 file	**Field name in plain text**
PRMSL	Mean sea level air pressure
TMP:surface	Temperature of the land or ocean surface
TMP:500 mb	Air temperature at the 500 mb level
RH:850 mb	Relative humidity at the 850 mb level
WIND:10 meter	Wind speed 10 m above the land or ocean surface
WIND:850 mb	Wind speed at the 850 mb level
WIND:250 mb	Wind speed at the 250-millibar

Table 1. GFS Fields used

2.3 GFS Model Error Correction

To correct for any initialization errors in the GFS, historical GFS files from the aforementioned UCAR-RDA database were compared with HURDAT2 data to find an approximate conversion between GFS hurricane strength and actual hurricane strength, to use with +006h forecasts. The GFS hurricane strength is defined here as the GFS data point with the highest

https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

surface wind velocity within 423 km of the storm center, the average radius of tropical cyclones (Chavas and Emanuel 2010). This value was calculated for every hurricane at every 6-hour interval in the study time period. A curve of best fit was determined by statistical regression and used as a conversion function using the GFS hurricane strength as an input and outputting the estimated real-life wind speed. All samples present in both the UCAR-RDA and HURDAT databases were included. The strength correction function is found by means of a nonlinear regression.

2.4 Error Interval Model Design

A model to predict the error interval for hurricanes was constructed, called the Hurricane Forecast Certainty Model (HFCM). In the first step of its calculations, the average of a value over a storm's area is defined as the weighted average value of a quantity over all GFS data points within 423 km of the storm center. As the GFS uses an equirectangular grid, points representing smaller areas will be given smaller weights. The exact calculation for this average is shown below

$$m = \frac{s}{w}$$

$$s = \sum_{p}\left[\left(r\cos^{-1}\left(\sin p_{\varphi}\sin\varphi_0 + \cos p_{\varphi}\cos\varphi_0\cos|p_{\lambda} - \lambda_0|\right) < k\right) \rightarrow p_q \cos p_{\lambda}\right]$$

$$w = \sum_{p}\left[\left(r\cos^{-1}\left(\sin p_{\varphi}\sin\varphi_0 + \cos p_{\varphi}\cos\varphi_0\cos|p_{\lambda} - \lambda_0|\right) < k\right) \rightarrow \cos p_{\lambda}\right]$$

with m the Mean Storm-Area quantity, p a GFS data record, p_{φ} the longitude of that record, p_{λ} the latitude of that record, p_q the value of the data stored at that record, φ_0 the longitude of the storm's center, and λ_0 the latitude of the storm's center, and the coefficients k = 423 km for the storm radius as found by Chavas and Emanuel (2010) and r = 6378.137 km for the Earth's radius.

To explain the formula more intuitively, the Mean Storm-Area quantity will be calculated by first selecting only points within the tropical cyclone, using the 423 km from the 2010 Chavas and Emanuel study. The distance is calculated using the Spherical Law of Cosines. Then, a weighted average of all selected values is calculated. Since the area represented by a GFS datapoint is proportional to the cosine of its latitude, weights are assigned as such.

The error interval was computed by first categorizing all hurricanes into quantized interval categories, determined individually for each component. The quantized interval categories are as follows:

•Storm Center Latitude: 60 N – 0 N, 1 degree intervals

•Storm Center Longitude, 105 W – 0 W, 1 degree intervals

•Mean Storm-Area Sea Surface Temperature: 20 C to 35 C, 1 C intervals

•Mean Storm-Area 850-250 mb Wind Shear: 0 mph – 220 mph, 5 mph intervals

https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

•Mean Storm-Area 850 mb Relative Humidity: 0 % - 100 %, 5 % intervals

•Maximum 10 m Wind Speed: 30 mph – 190 mph, 5 mph intervals

•Actual Intensity divided by Potential Intensity (Kaplan and DeMaria 1994b): 0 % - 100 %, 5 % intervals

For each component, the root mean square error (RMSE), a standard technique to measure deviation, of each GFS hurricane forecast was used to calculate the average forecast error, which was then stored in the program. For example, if there were three hurricanes that were in the 10 W - 11 W interval, and their forecasts differed from ground truth by 10 mph above, 5 mph below, and 5 mph below, the value put in the longitude table at 10 W - 11 W cell would be 7.07 mph.

To find the error multipliers, a hurricane's parameters, such as its Average Storm-Area Sea Surface Temperature, would be determined, and the error multipliers of each component would be determined by taking the value from each lookup table and dividing it by the standard deviation of the whole dataset. For example, if the 7.07 mph value from earlier is used, and the standard deviation of the whole dataset is set to 10 mph, the longitude error multiplier would be 0.707.

Finally, to determine the error range, every multiplier was multiplied together, and then multiplied with the standard deviation of the dataset. Then, the lower bound of the error interval was set as the GFS forecast minus the error range, and the upper bound was set by the GFS forecast plus the error range. If the GFS forecast is 80 mph and the error range is 7.1, the error interval is 72.9 - 87.1 mph.

2.5 Error Interval Model Skill Evaluation

The accuracy of the HFCM was measured as the percentage of forecasts that fell within the confidence interval. For example, if a hurricane was forecasted to be 72.9 - 87.1 mph, and the hurricane actually became 84 mph, the forecast would be counted as a success. However, if it were to strengthen to 97 mph, the forecast would be counted as a failure.

Another model was constructed for use as a baseline to test the skill of the HFCM, following the same method except for the error range. This is set to the overall average forecast error for the whole dataset, 12.53 mph. Skill was measured as the percentage point difference between the accuracy of the HFCM and the baseline.

3 Results

3.1 GFS Error Correction

A scatter plot of the results shows the GFS has significant error in hurricane maximum wind speed, sometimes significantly overestimating its strength while at times severely underestimating it. This results in the data taking on the appearance of a

https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

loose correlation, as seen in Figure 2. Further to the right of the graph, the cluster begins to deviate from the line defined by y = x, shown by the diagonal dashed line, showing a tendency for the GFS to consistently underestimate the strength of strong hurricanes, especially those above 89.5 mph (40 m/s, 77.8 kt).

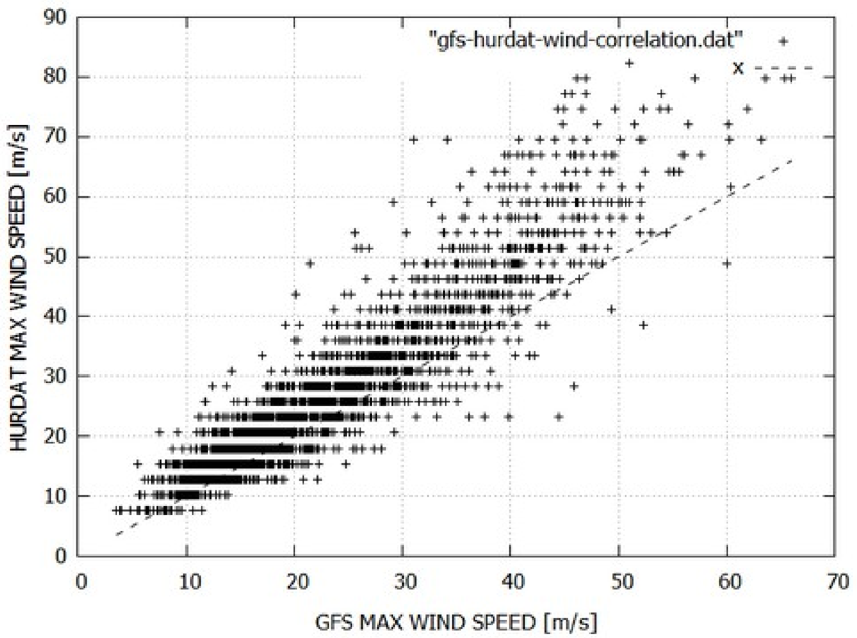


Figure 2. Hurricane Strength: HURDAT vs. GFS

The imprecision inherent to the current initialization procedure of the GFS does limit the quality of any intensity forecast derived from its forecasts, but the actual strength of the hurricane can be estimated. A quadratic form ($y = Ax^2 + Bx + C$) with the coefficients A = 0.00910682, B = 0.753648, and C = 4.08778 accounts for 86.32% of the variation. This curve is depicted in Figure 3. An exponential model ($y = AB^x + C$) with the coefficients A = 63.9202, B = 1.01352, and C = -60.7188 is almost as accurate, accounting for 86.27% of the variation. Due to its slightly higher accuracy over the exponential model, the quadratic model has been used for correcting the wind values from GFS forecasts.

https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

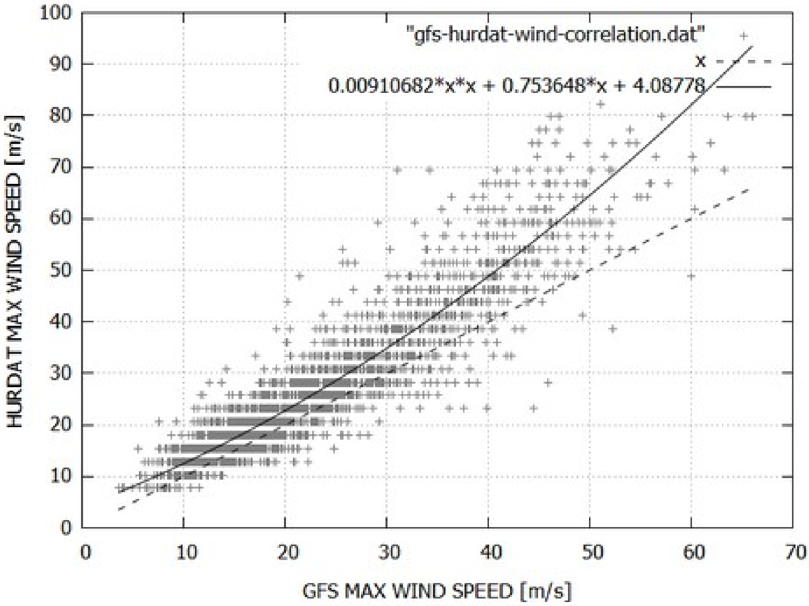


Figure 3. GFS Max Wind Correction Function

l3.2 HFCM Lookup Tables

As the lookup tables were generated, it was clear that the values within were noisy, meaning that they were highly variable in a random way. In Figure 4, the noise creates significant variation on the smaller scale, possibly due to the limited size of the dataset. Some lookup tables showed strong trends despite this noise, such as in Figure 5. There, cyclones with lower fractions of their Maximum Potential Intensity (MPI) (Kaplan and DeMaria, 1994b) are found to have significantly lower forecast errors than those with higher MPIs. The average forecast error in the ≥100% band was nearly 27 mph, while only 6 mph in the 15-20% band. The higher MPI bands were noisier, possibly due to the lower sample amount. Further research needs to be conducted to determine the effect of more samples.

https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

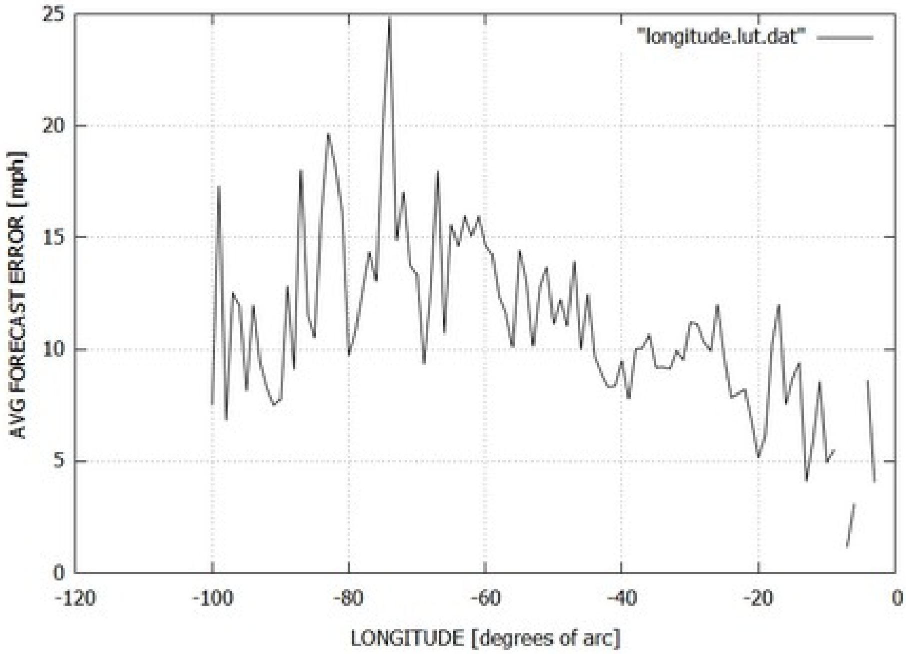


Figure 4. Longitude Lookup Table Plot

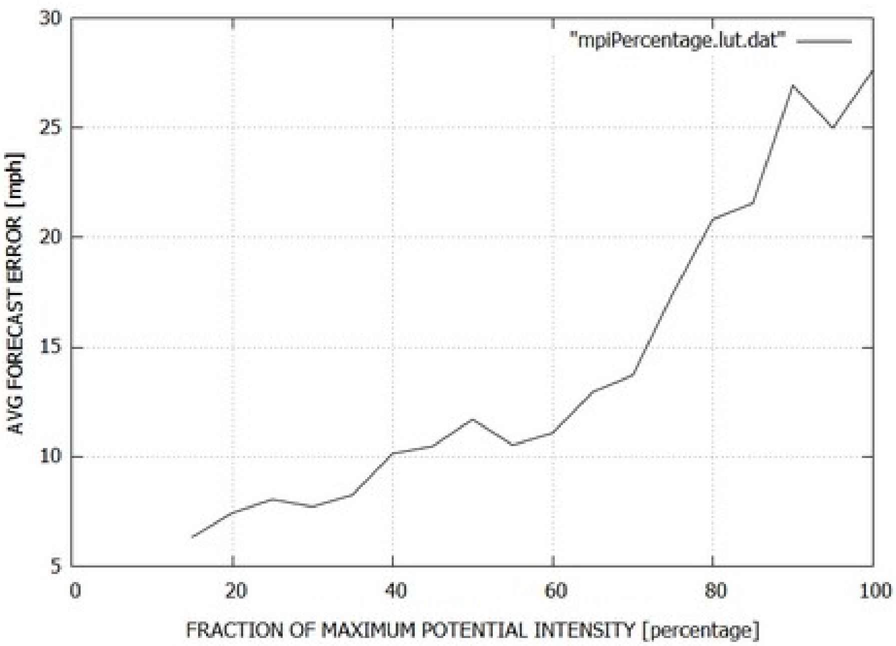


Figure 5. MPI Percentage Lookup Table Plot

https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

3.3 HFCM Skill

Table 2 shows that there is a significant difference in skill across the spectrum of tropical cyclone strengths. The HFCM performs 52.8 points worse than the baseline for tropical depressions, but 42.8 points better for Category 3 hurricanes. Overall, the model performs worse than the baseline; however, most of this negative skill is with tropical depressions and storms. The model performs positively in all hurricane categories except for Category 1. The best performance is seen in Category 4 and Category 5 hurricanes, both with over 80 points of skill.

Forecast Category	Baseline Accuracy	HFCM Accuracy	HFCM Skill	Amt. Samples in Category
Overall	78.60%	53.20%	-25.4	2863
Tropical Depressions	94.70%	41.90%	-52.8	472
Tropical Storms	86.40%	47.90%	-38.5	1640
Hurricanes	51.30%	72.00%	20.7	751
Category 1 Hurricanes	64.20%	58.50%	-5.7	369
Category 2 Hurricanes	50.00%	70.50%	20.5	156
Category 3 Hurricanes	48.60%	91.40%	42.8	105
Category 4 Hurricanes	17.70%	97.90%	80.2	96
Category 5 Hurricanes	8.00%	100.00%	92	25

Table 2. HFCM and Baseline Model Accuracy by Category

3.4 Storm Intensity-Forecast Error Correlation in Longitude Lookup Table

The lookup table for Longitude contained notably higher values for average forecast error in between 80 W and 40 W. As seen in Figure 6, longitude bands outside of that range frequently had an average forecast error of 10 to 18 mph, while within the range, errors of 16 - 24 mph were common. This data was analyzed for a correlation between storm intensity and forecast error using Pearson's *r* for the correlation coefficient and a least-squares regression for the line of best fit.

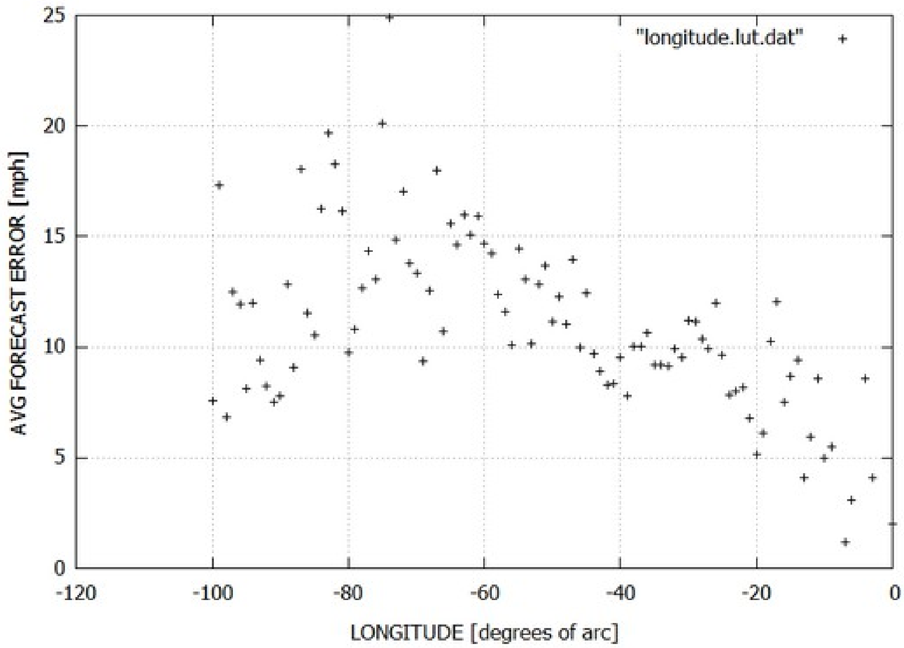


Figure 6. Longitude Lookup Table Point Plot

The data had a correlation coefficient of 0.7535, a moderately high positive correlation. The line of best fit is shown in Figure 7 and described by $E = 0.236521s - 3.93024$, s meaning average storm strength and E meaning average forecast error. That line of best fit explains 56.78% percent of the variance within the data. This shows that there is some relation between storm strength and forecast error in the lookup table, but other influencing factors likely exist.

https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

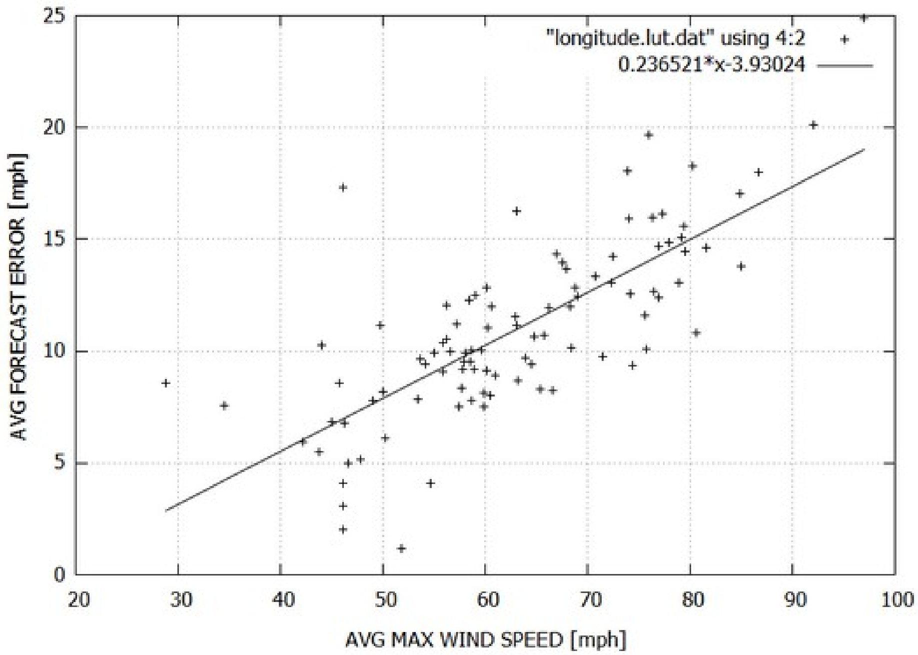


Figure 7. Longitude and Strength Correlation

3.5 Lookup Table Noise-Sample Size Correlations

Variability of the amount of noise was apparent in the data. To analyze the exact behavior of this noise, n-width noisiness was defined as the average of the absolute values of the differences between the band being evaluated and its n neighboring bands in each direction. For example, the 3-width noisiness of the latitude band 25 N would take the absolute values of the differences between the datum in that band and the data in the bands 22 N, 23 N, 24 N, 26 N, 27 N, and 28 N. Data from bands with zero samples are excluded. A variable n-width has been chosen to account for the differences in size and bandwidth between different lookup tables. An n-width of 3 was used for Latitude, Longitude, Wind Speed, and Wind Shear. An n-width of 1 was used instead for the others on the basis that those lookup tables were smaller.

Shown in Figure 8, some noise analyses showed strong inverse correlations between sample size and noise amount, most notably for the Wind Shear analysis. Using a nonlinear regression model, the curve of best fit is described by $y = 1/sqrt(Ax)$, $A = 0.00214806$. Such a model has an $R2$ of 0.6917. Similar patterns were found in the MPI Percentage, 850 mb Relative Humidity, SST, and Wind Speed analyses.

https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

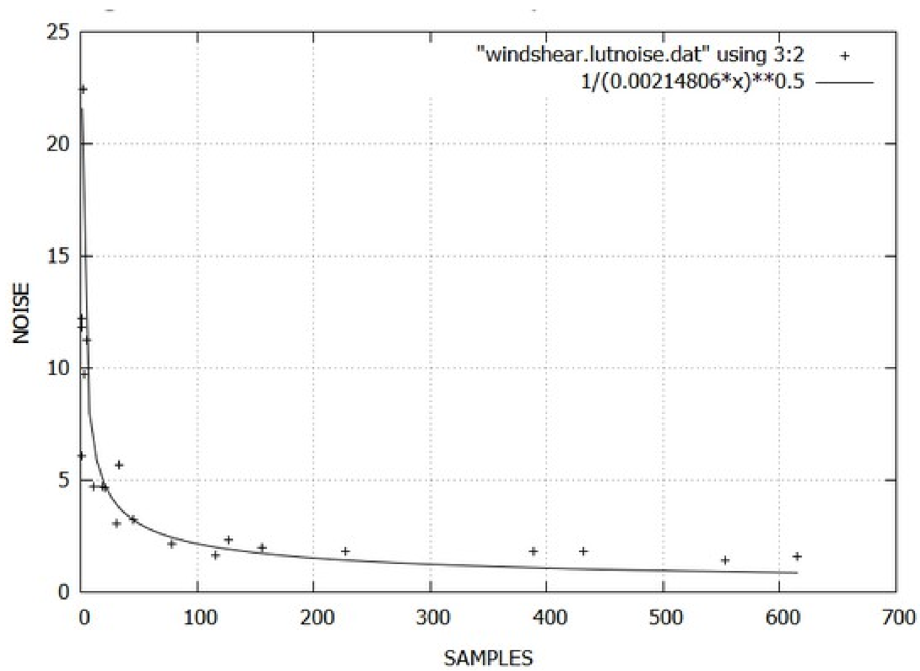
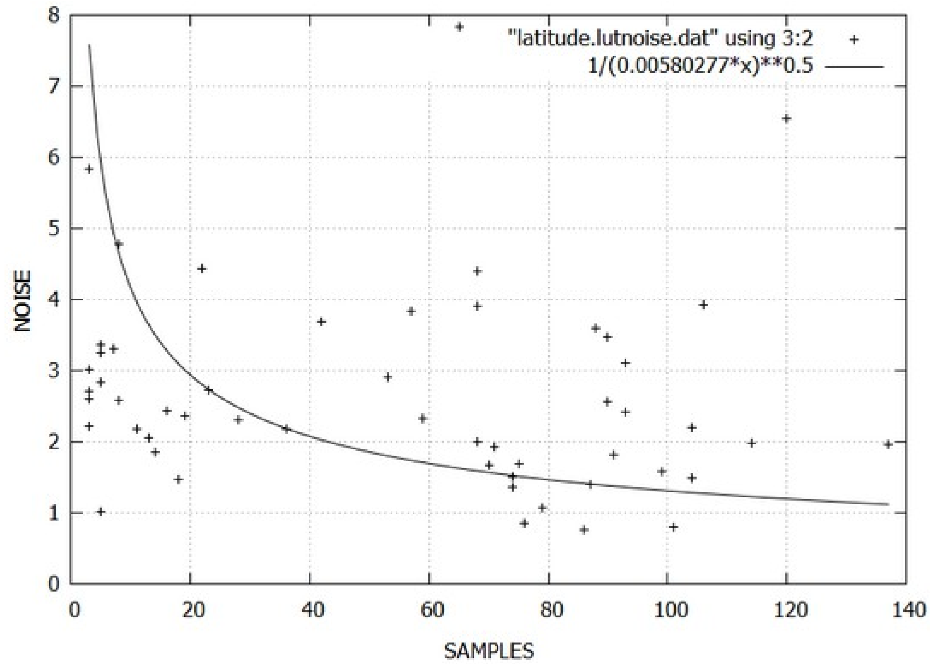


Figure 8. Wind Shear Noise Analysis

However, other analyses showed little to no correlation. The Latitude analysis showed no significant correlation at all between the sample size and the noise amount (Figure 9). The same regression method as earlier finds a curve of best fit using *A = 0.00580277*, but this curve has an *R2* of -1.114, meaning that this model performs worse than a constant function that always predicts the mean noise value of the dataset. A similar pattern was found in the Longitude analysis.

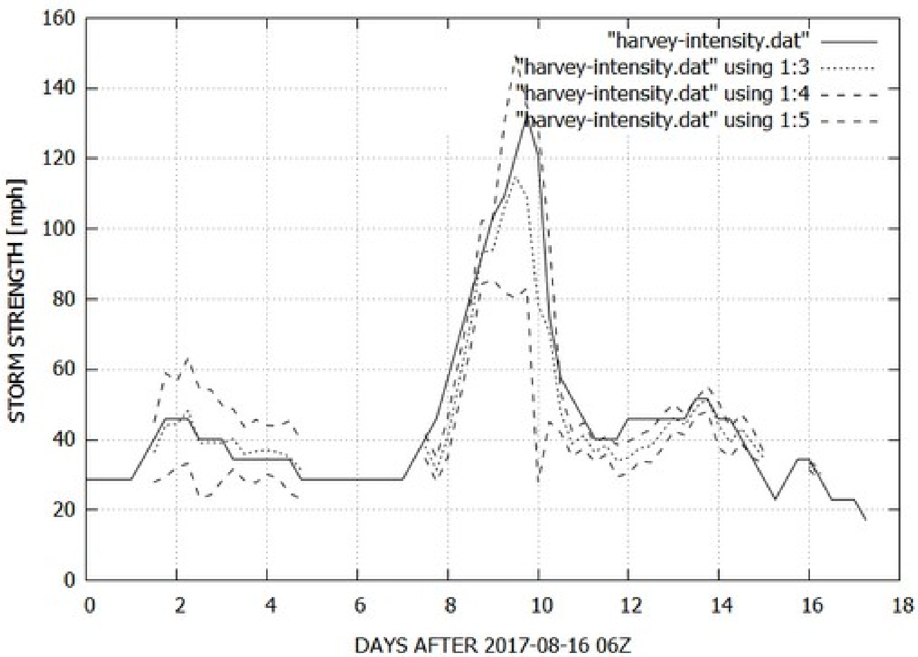


https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

Figure 9. Latitude Noise Analysis

3.6 HFCM Test Runs

To examine its behavior in a real-life storm forecasting situation, the HCFM model was tested on Hurricane Harvey, a Category 4 storm that caused severe flooding in Houston, Texas in 2017 (Amadeo 2019). The storm caused $125 billion dollars worth of damage, a figure matched only by Hurricane Katrina in 2005 (Amadeo 2019). Figure 10 shows how the cyclone rose and fell in strength, and took until August 24, 2017 to begin intensifying in earnest. The dashed lines show the error intervals forecasted by the HCFM, and the dotted line represents the storm strength forecasted by the GFS. There were two periods during which Harvey was too weak for use with the HFCM model, from August 21 at 00 UTC to August 23 at 18 UTC and from August 30 at 06 UTC to August 31 at 06 UTC. Winds during this period of rapid intensification were drastically underestimated by the GFS. The storm strength continued rising for 6 hours after its downturn was forecasted, possibly due to a later-than-forecast landfall. The storm weakened very rapidly, but then stabilized as it flooded Houston. It finally dissipated over a week after landfall. The HCFM performed very well between August 17 and 21, but forecasted too narrow of intervals at the beginning of the rapid intensification. The peak of the storm was forecasted very well, as demonstrated by the successful HFCM forecasts despite the large difference between the forecasted and actual strengths on August 25th and during the rapid weakening on August 26th. The model failed on August 27th and 28th, during the stable period, but succeeded for all of Harvey's remaining life.



https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

Figure 10. Hurricane Harvey HFCM Run

The HFCM model was also run on Hurricane Dorian, a Category 5 storm that caused $3.4 billion in economic damages in the Bahamas (ECLAC, 2019). The storm gradually intensified as it crossed the Atlantic before making landfall on Great Abaco Island in the Bahamas, and then drifted northwards along the United States coast. Figure 11 shows this intensification between August 28, 2019 and September 1, 2019, along with the upper and lower bounds forecasted by the HCFM.

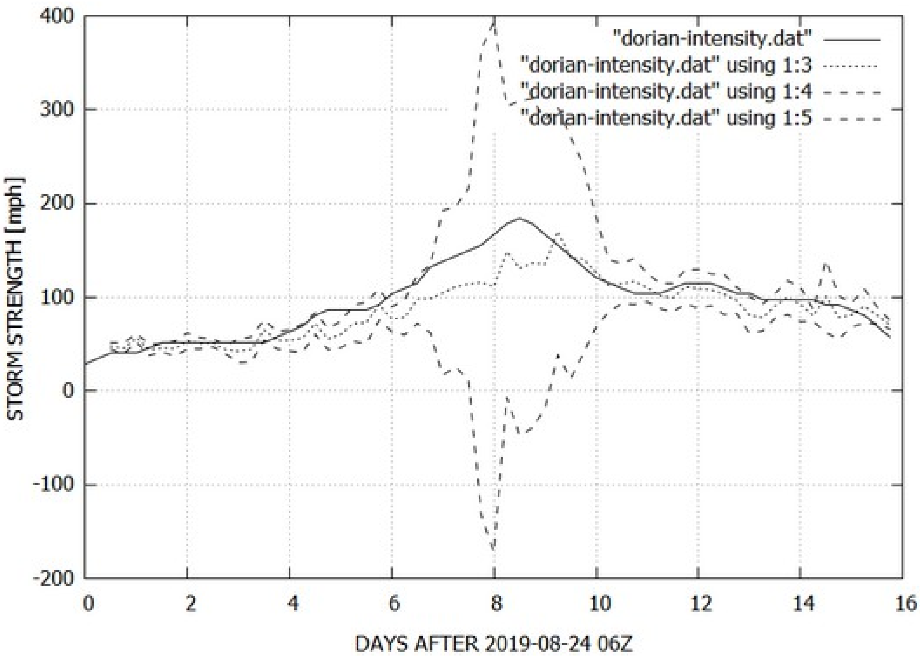


Figure 11. Hurricane Dorian HFCM Run

The forecast shows low error ranges, depicted with the two dashed lines in the Figure, near the beginning of the storm's lifetime, but higher error ranges during the period of intensification. However, these higher ranges were not enough to correct for the GFS's underestimation of Dorian's strength in most instances, shown in the dotted line. Starting on August 31, the error range had grown absurdly large, often returning lower bounds well below zero. The widest error interval was seen on September 1 at 06 UTC, with a lower bound of -172.0 mph and an upper bound of 394.2 mph. These forecasts were counted as successes as per the method, but the use of such results in a real-life forecasting situation would be limited.

4 Limitations

This model has three major limitations: firstly, it is only designed for use within the Atlantic basin. The longitude and latitude fields only extend from 105 W to 5 E and 0 N to 60 N respectively. Furthermore, the existence of two separate SHIPS models, one for the Atlantic basin and the eastern North Pacific basins (Kaplan and DeMaria, 1999), shows that there may exist differences in the statistics of the hurricanes that form within the two basins. Secondly, the GFS's initialization inaccuracies inherently limit the accuracy of any HFCM forecasts made with GFS forecasts. The GFS-HURDAT2 wind speed correction does help, but it cannot completely correct these inaccuracies due to the lack of a 1:1 correlation between the storm strength as reported by the GFS and the storm strength in reality. Thirdly, the small sample sizes in some bands and the noise that brings to the lookup tables may cause hurricanes with similar characteristics to be given different error ranges by the HFCM. Further research is needed to determine the effect of a larger sample size as well as the application of techniques to reduce or smooth out the noise.

5 Discussion and Conclusions

While the HFCM demonstrated negative skill with tropical depressions and tropical storms, -52.8 and -38.5 percentage points respectively, it demonstrated +20.7 percentage points of skill with hurricanes on average. It performed best with Category 5 hurricanes, showing +92.0 points of skill and an overall accuracy of 100.0%, given that the samples of Category 5 hurricanes were limited. In fact, the HFCM tended to forecast extremely wide intervals for significantly strong storms. The difference in performance between the HFCM and the accuracy baseline were astounding, showing that the HFCM should only be used with stronger storms. The HFCM could be used by a citizen living in a hurricane-prone area who needs more information about a cyclone's forecasted strength in order to decide whether to evacuate or to shelter in place through the storm. It could also be used by forecasting agencies to evaluate the possible range of outcomes and by emergency management agencies to make more informed decisions on whether to issue evacuation orders for a given region.

This study leaves four clear next steps to take with future research. Constructing a similar model for the North Pacific Basin could be of similar utility for regions affected by North Pacific hurricanes, such as the Mexican States of Sinaloa and Baja California Sur, Hawaii, as well as the US States of California and Arizona. Furthermore, a model constructed with a larger dataset in place of the UCAR-RDA GFS archive would likely have less statistical noise within its lookup tables, as was shown in the analysis. Thirdly, as the GFS was not specifically intended for forecasting hurricanes, the use of a model such as the HWRF for data about the surrounding environment as well as the forecasted wind speeds may yield more precise results. Fourthly, a more sophisticated model that tempers the extremely wide error ranges as seen with hurricane Dorian would drastically improve the usefulness of this model in such scenarios. Those four steps can help construct the tools to help citizens and emergency management organizations make more informed decisions regarding hazardous tropical cyclones.

https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

Data availability. The GFS archives are available at https://doi.org/10.5065/D65D8PWK. Accessed 27 Dec 2021. The HURDAT2 dataset is available at https://www.nhc.noaa.gov/data/. All data computed as part of this study are accessible upon request to the author.

Code availability. Code used to perform the analyses included in this study is accessible upon request to the author

Author contributions. ZU performed all work associated with this study.

Competing interests. The author has declared that they do not have any competing interests.

Acknowledgements. Zander Urquhart thanks Mrs. Anderson for teaching the AP Capstone program at their high school, and especially for the skills taught in AP Research, in which this study was conducted.

References

Amadeo, K.: Hurricane Harvey Facts, Damage and Costs, The Balance, 2019.

Bassill, N.: Accuracy of early GFS and ECMWF Sandy (2012) track forecasts: Evidence for a dependence on cumulus parameterization. Geophysical Research Letters, 41, 3274-3281, https://doi.org/10.1002/2014GL059839, 2014.

Cangialosi, J.: National Hurricane Center Forecast Verification Report: 2019 Hurricane Season. National Hurricane Center. 2020.

Charney, J. G., Fjörtoft, R., and Von Neumann, J.: Numerical Integration of the Barotropic Vorticity Equation. Tellus, 2, 237-254, https://doi.org/10.1111/j.2153-3490.1950.tb00336.x, 1954.

Chavas, D. R. and Emanuel, K.: A QuikSCAT climatology of tropical cyclone size. Geophysical Research Letters, 37, doi:10.1029/2010GL044558, 2010.

ECLAC (UN Economic Commission for Latin America and the Caribbean): Assessment of the Effects and Impacts of Hurricane Dorian in the Bahamas. Inter-American Development Bank. 2019.

Fraedrich, K.: Estimating the Dimensions of Weather and Climate Attractors. Journal of the Atmospheric Sciences, 43, 419-432. https://doi.org/10.1175/1520-0469(1986)043<0419:ETDOWA>2.0.CO;2, 1986.

Harr, P., Elsberry, R., and Hogan, T.: Extratropical Transition of Tropical Cyclones over the Western North Pacific. Part II: The Impact of Midlatitude Circulation Characteristics. Monthly Weather Review, 128, 2634-2653. https://doi.org/10.1175/1520-0493(2000)128<2634:ETOTCO>2.0.CO;2, 2000.

Kaplan, J. and DeMaria, M.: A Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic Basin. Weather and Forecasting, 9, 209–220. https://doi.org/10.1175/1520-0434(1994)009<0209:ASHIPS>2.0.CO;2, 1994a

https://doi.org/10.5194/egusphere-2022-343
Preprint. Discussion started: 23 May 2022

Kaplan, J. and DeMaria, M.: Sea Surface Temperature and the Maximum Intensity of Atlantic Tropical Cyclones. Journal of Climate, 7, 1324-1334, https://doi.org/10.1175/1520-0442(1994)007<1324:SSTATM>2.0.CO;2, 1994b.

Kaplan, J. and DeMaria, M.: An updated statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic and eastern North Pacific basins. Weather and Forecasting, 14, 326-337, https://doi.org/10.1175/1520-0434(1999)014<0326:AUSHIP>2.0.CO;2, 1999.

Kaplan, J., DeMaria, M. and Knaff, J.: A Revised Tropical Cyclone Rapid Intensification Index for the Atlantic and Eastern North Pacific Basins. Weather and Forecasting, 25, 220-241, https://doi.org/10.1175/2009WAF2222280.1, 2010.

Kaplan, J., DeMaria, M., et. al.: Further Improvements to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). Weather and Forecasting, 20, 531-543, https://doi.org/10.1175/WAF862.1, 2005.

Emanuel, K.: The Maximum Intensity of Hurricanes. Weather and Forecasting, 45, 1143-1155, https://doi.org/10.1175/1520-0469(1988)045<1143:TMIOH>2.0.CO;2, 1988.

Landsea, C. and Cangialosi, J.: Recent Progress in Tropical Cyclone Intensity Forecasting at the National Hurricane Center. Weather and Forecasting, 35, 1913-1922, https://doi.org/10.1175/WAF-D-20-0059.1, 2020.

National Centers for Environmental Prediction/NCEP. NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, CO, , 2015.

Persing, J. and Montgomery, M.: Hurricane Superintensity. Journal of the Atmospheric Sciences, 60, 2349-2371, https://doi.org/10.1175/1520-0469(2003)060<2349:HS>2.0.CO;2, 2003.

Trabing, B. and Bell, M.: Understanding Error Distributions of Hurricane Intensity Forecasts during Rapid Intensity Changes. Weather and Forecasting, 35, 2219-2234, https://doi.org/10.1175/WAF-D-19-0253.1, 2020.