Peer review

Houghton, Graham. 2025. "A Multivariable Regression Model Forecasting the Rapid Intensification of Hurricanes." *Journal of High School Science* 9 (1): 1–17.

A good start. However, I have some concerns that I list below and must be answered, explained or discussed in the manuscript.

1. you seem to arbitrarily change the definition of intensification from 30 to 34.5 knots "….After converting the 30 knot threshold for rapid intensification into 34.5 miles per hour…." which is probably because more predictive accuracy is achieved. However, this represents a heuristic classification thereby undermining your model and seems like you adjust the classifier cut-off to suit the model. I do not think this is justificable in a scientific context. Please explain and describe your rationale for this change in the manuscript.

2. If intensification is defined as >30 knots in 24 hours, does that translate into > 90 knots in 72 hours? Please explain and describe in the manuscript what defines intensification over a period of 72 hours?

3. For the variables that were eventually included in the final regression equation, please present their standard deviation over a period of 72 hours. For example, average OHC plus/minus SD, average HWS plus/minus SD and average calculated HIKE plus/minus SD.

4. From figure 1, your model underestimates HWS at 72 hours for actual speed increases of >20 and < 60 mph at -72 hours and overestimates HWS for actual speed increases of > 60 mph at -72 hours. This means that the HWS and HIKE terms influence the predicted HWS more when actual Deltaspeedincrease is between 20 and 60 miles per hour and the OHC term influences the predicted HWS to a larger extent when actual deltaspeedincrease > 60 miles per hour. This leads me back to my previous point. Should you consider using the value of [(OHC -72) – SD] instead of average OHC -72 when HWS -72 > 60 miles per hour. And, should you consider using [(HWS -72) – SD] instead of average HWS -72 and also using [(HIKE – 72) – SD] instead of average (HIKE -72) when HWS -72 < 60 miles per hour. In effect, use two equations to predict; one equation when HWS -72 is > 60 miles per hour and another equation when HWS – 72 is < 60 miles per hour. I don't know why the variables change explanatory mechanism at about 60 miles per hour hurricane speed. Is this a velocity below and above which the hurricane behaves differently ? Are there any literature references? (see: http://dx.doi.org/10.5772/14184, where the cut-off is 90 miles per hour for two different behaviors for vmax).

5. One of your explanatory variables is the speed at -72 hours. This makes HWS an autocorrelation temporal function. This is reflected in the regression coefficient of -0.6535 between delta HWS 72 and HWS -72 (table 2). This means the lower the speed at 72 hours out, the greater the speed is likely to be after 72 hours and vice versa (as you rightly point out "…..That is, storms are more likely to weaken from strength or intensify from weakness……") In Table 2, there is also a good correlation between HIKE -72 and HWS (- 0.6284). In fact, the correlation is much lesser for OHC -72 and HWS (0.39). Why then are the two explanatory variables HIKE -72 and HWS -72 not enough to predict HWS ? I would like to see correlation coefficients R2, with just OHC -72 against HWS, HWS -72 against HWS, HIKE -72 against HWS and finally a combination of HWS -72 and HIKE -72 against HWS calculated as regression equations.

6. From Table 3, HWS -72 and HIKE -72 are strongly correlated (0.767), If that is the case, they should both not be explanatory variables and part of the same equation. Explain and discuss in the manuscript. What was done to check for collinearity of these explanatory variables. I would like to see calculation of Variance Inflation Factor (VIF) values as evidence for non – collinearity. Please describe and explain in the manuscript.

7. You state "…..In order to calculate HIKE, wind- field data from NHC tropical cyclone forecast advisories [14] was entered into the Integrated Kinetic Energy Calculator available from NOAA [15]….." Please present a sample (example) calculation because it is not clear what the 'wind-field data from the NHC tropical cyclone forecast advisory' is.

8. Why is surface sea temp. not as good an explanatory variable as OHC. The two should be correlated (0.553 correlation from Table 3) if the mass and specific heat of the medium are constant (which they are). Please explain.

A good start. However, I have some concerns that I list below and must be answered, explained or discussed in the manuscript.

Thank you for reviewing this paper. I have answered your questions below and added to the manuscript in two areas as indicated. Please let me know if you have any further questions.

1. you seem to arbitrarily change the definition of intensification from 30 to 34.5 knots "….After converting the 30 knot threshold for rapid intensification into 34.5 miles per hour…." which is probably because more predictive accuracy is achieved. However, this represents a heuristic classification thereby undermining your model and seems like you adjust the classifier cut-off to suit the model. I do not think this is justifiable in a scientific context. Please explain and describe
your rationale for this change in the manuscript.

   The HWS variable is measured in miles per hour as reported by the National Hurricane Center. The dHWS predicted by the regression model is also in miles per hour. Since the rapid intensification threshold is defined in (Kaplan) as 30 knots, the threshold was converted to miles per hour so the comparison is done using the same units (1.15 mph = 1 knot).

2. If intensification is defined as >30 knots in 24 hours, does that translate into > 90 knots in 72 hours? Please explain and describe in the manuscript what defines intensification over a period of 72 hours?

   This is explained on page 4 of the manuscript: "For purposes of this study, the rapid intensification detection threshold was expanded from 24 to 72 hours to balance the range of uncertainty of the forecasted path with the need to provide sufficient time for emergency management."

   The time window was modified because 24 hours does not provide enough time to mobilize emergency management resources and evacuate residents. The severity threshold was not adjusted because an increase of 34.5mph in storm windspeed is material to the amount of potential damage regardless of the timeframe for intensification.

3. For the variables that were eventually included in the final regression equation, please present their standard deviation over a period of 72 hours. For example, average OHC plus/minus SD, average HWS plus/minus SD and average calculated HIKE plus/minus SD.

Addressed in #4 below.

4. From figure 1, your model underestimates HWS at 72 hours for actual speed increases of >20 and < 60 mph at -72 hours and overestimates HWS for actual speed increases of > 60 mph at -72 hours. This means that the HWS and HIKE terms influence the predicted HWS more when actual Deltaspeedincrease is between 20 and 60 miles per hour and the OHC term influences the predicted HWS to a larger extent when actual deltaspeedincrease > 60 miles per hour. This leads me back to my previous point. Should you consider using the value of
[(OHC -72) – SD] instead of average OHC -72 when HWS -72 > 60 miles per hour. And, should you consider using [(HWS -72) – SD] instead of average HWS -72 and also using [(HIKE – 72) – SD] instead of average (HIKE -72) when HWS -72 < 60 miles per hour. In effect, use two equations to predict; one equation when HWS -72 is > 60 miles per hour and another equation when HWS – 72 is < 60 miles per hour. I don't know why the variables change explanatory mechanism at about 60 miles per hour hurricane speed. Is this a velocity below and above which the hurricane behaves differently ? Are there any literature references? (see: http://dx.doi.org/10.5772/14184, where the cut-off is 90 miles per hour for two different behaviors for vmax).

There is a paragraph on page 11 of the manuscript which discusses the outlier storms you refer to above. Instead of looking into a HWS cutoff of the data, I added an indicator variable to the regression for deep water landfall locations (n=7). This improved the regression significantly (R2 increased to 0.75 from 0.56 and the outliers were significantly reduced). Given that n=7 is quite low, I included additional information about this result by expanding the "Other Storms" paragraph on page 12.

The literature reference you provide indicates a linear relationship between hurricane intensity and frequency up to 90mph, and then hurricane frequency decays more rapidly above 90mph, as those extreme storms are more rarely observed. This does not seem to address individual storm windspeed behavior, it is an analysis rather of the overall incidence rates of maximum observed windspeeds for each storm.

5. One of your explanatory variables is the speed at -72 hours. This makes HWS an autocorrelation temporal function. This is reflected in the regression coefficient of -0.6535 between delta HWS 72 and HWS -72 (table 2). This means the lower the speed at 72 hours out, the greater the speed is likely to be after 72 hours and vice versa (as you rightly point out "…..That is, storms are more likely to weaken from strength or intensify from weakness……") In Table 2, there is also a good correlation between HIKE -72 and HWS (- 0.6284). In fact, the correlation is much lesser for OHC -72 and HWS (0.39). Why then are the two explanatory variables HIKE -72 and HWS -72 not enough to predict HWS ? I would like to see correlation coefficients R2, with just OHC -72 against HWS, HWS -72 against HWS, HIKE -72 against HWS and finally a combination of HWS -72 and HIKE
-72 against HWS calculated as regression equations.

Generally, I was looking to achieve an R2 in the 0.6 range and as I sequentially removed explanatory variables with poor p-values, the regression with OHC, HWS and HIKE met the criteria of having a good R2 and also significance for each individual coefficient. While HWS provides good information about the maximum windspeed of the storm at its center, HIKE provides information about how much energy is present throughout the storm, so while these are certainly correlated, there is a different kind of information contained in each variable.

A summary of the R2 for the regressions you requested is here:

Variable(s)	R2
OHC-72	0.153
HWS-72	0.427
HIKE-72	0.395
HWS-72 and HIKE-72	0.466

6. From Table 3, HWS -72 and HIKE -72 are strongly correlated (0.767), If that is the case, they should both not be explanatory variables and part of the same equation. Explain and discuss in the manuscript. What was done to check for collinearity of these explanatory variables. I would like to see calculation of Variance Inflation Factor (VIF) values as evidence for non – collinearity. Please describe and explain in the manuscript.

Here are the VIF values:

Variable	VIF
OHC	1.022
HWS	2.424
HIKE	2.444

These are below the VIF threshold for collinearity.

I added a short paragraph about this collinearity test in the manuscript on page 8.

Given the relatively high correlation ($\rho = 0.767$) between HWS and $HIKE$, a test for collinearity was performed on the final regression model. Variance inflation factors for each explanatory variable appear in Table 5 below and are all under 2.5, well within an acceptable range for non-collinearity.

Table 5: *Explanatory Variable Variance Inflation Factors*

Variable	VIF-value
OHC^{-72}	1.022
HWS^{-72}	2.424
$HIKE^{-72}$	2.444

7. You state "…..In order to calculate HIKE, wind- field data from NHC tropical cyclone forecast advisories [14] was entered into the Integrated Kinetic Energy Calculator available from NOAA [15]….." Please present a sample (example) calculation because it is not clear what the 'wind-field data from the NHC tropical cyclone forecast advisory' is.

Here is the windfield data from the NHC forecast advisory for Hurricane Michael (https://www.nhc.noaa.gov/archive/2018/al14/al142018.fstadv.016.shtml)

```
ESTIMATED MINIMUM CENTRAL PRESSURE  928 MB
EYE DIAMETER  15 NM
MAX SUSTAINED WINDS 125 KT WITH GUSTS TO 150 KT.
64 KT....... 40NE  35SE  25SW  30NW.
50 KT....... 80NE  80SE  40SW  50NW.
34 KT.......150NE 140SE  80SW 120NW.
12 FT SEAS..120NE 240SE 210SW 120NW.
WINDS AND SEAS VARY GREATLY IN EACH QUADRANT.  RADII IN NAUTICAL
MILES ARE THE LARGEST RADII EXPECTED ANYWHERE IN THAT QUADRANT.
```

Here is the NOAA IKE Calculator webpage after populating the data fields and calculating the result (HIKE = 45.276 TJ in this instance):

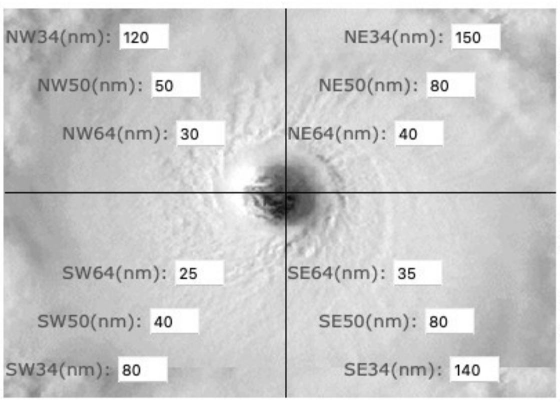
Integrated Kinetic Energy

I.K.E. Calculator

This calculator will let you estimate the IKE and Surge Destructive Potential based on estimates of the outermost extent of tropical storm (TS), 50 kt, and Hurricane Force (H) winds in each storm quadrant. The results will give you ball park estimates if you use the operational wind radii. These estimates are based on some very coarse wind field assumptions. An improved estimate would come from a gridded wind analysis or model input data using the methods described in Powell and Reinhold 2007.

Update: Please note that IKE calculations prior to October 23, 2012 used a surface air density 0f 1.0 Kg per cubic meter. Current calculations (after October 23, 2012) use a more representative value of 1.15 kg/cubic meter based on GPS sonde measurements in hurricanes.

Rmax(nm): 15 Vmax(kt): 125



Results

Entire Storm

Storm Total IKE $_{TS-50}$ (TJ)	24.783
Storm Total IKE $_{50-H}$ (TJ)	13.017
Storm Total IKE $_{H}$ (TJ)	7.476
IKE$_{TS}$(TJ)	45.276
SDP	3.688

Sample Storm

Calculate reset

NW Quadrant

IKE $_{TS-50}$ (TJ)	6.411
IKE $_{50-H}$ (TJ)	1.660
IKE $_{H}$ (TJ)	1.360
IKE $_{TS}$ (TJ)	9.431

NE Quadrant

IKE $_{TS-50}$ (TJ)	8.674
IKE $_{50-H}$ (TJ)	4.979
IKE $_{H}$ (TJ)	3.308
IKE $_{TS}$ (TJ)	16.961

SW Quadrant

IKE $_{TS-50}$ (TJ)	2.586
IKE $_{50-H}$ (TJ)	1.011
IKE $_{H}$ (TJ)	0.876
IKE $_{TS}$ (TJ)	4.474

SE Quadrant

IKE $_{TS-50}$ (TJ)	7.112
IKE $_{50-H}$ (TJ)	5.368
IKE $_{H}$ (TJ)	1.931
IKE $_{TS}$ (TJ)	14.411

8. Why is surface sea temp. not as good an explanatory variable as OHC. The two should be correlated (0.553 correlation from Table 3) if the mass and specific heat of the medium are constant (which they are). Please explain.

SST provides information about the temperature of the surface of the ocean, and thus provides data about the amount of heat immediately available to be transferred to an overpassing hurricane. The regression model favors OHC over SST as a predictive indicator, likely because it measures heat in the layers of water underneath as well as at the surface. As hurricanes pass over, water is churned up by the winds and if water underneath is also hot (and not cold), this provides more energy to the storm.

Thank you for addressing my comments. What I would like for you to do is to incorporate all the responses into the manuscript at appropriate locations with a contextual discussion. Please resubmit when done with two documents - one where the responses included in the manuscript are tracked and another clean version. Both should be word docs.
Please resubmit when done.

Thank you for reviewing this paper. I have resubmitted the manuscript with a tracked version included as well as the clean file. My original responses to your comments appear below for your convenience. Thank you

---

Thank you for addressing my comments. Accept.