Journal of High School Science

A multivariable regression model forecasting the rapid intensification of hurricanes

Houghton G

Submitted: July 23, 2024, Revised: version 1, December 22, 2024
Accepted: January 6, 2025

Abstract
Forecasting hurricane strength timely and accurately is of critical importance to ensure that preparations and evacuations can be carried out as necessary to save lives and property. In recent years, hurricanes have become more likely to experience rapid intensification, making them more dangerous and difficult to forecast with sufficient confidence, precision, and lead time. In general, rapid intensification is defined as a sudden increase in storm intensity within a short period of time, and often this intensification occurs just prior to landfall. Recent studies have linked the rapid intensification of storms to changes in the atmosphere and oceans resulting from climate change. The proportion of storms experiencing rapid intensification tripled from 1982 to 2009, suggesting that existing forecasting models based on historical storm data may be insufficient to account for today's climate environment. Significant advances in satellite observation technology and an increase of *in-situ* storm data collected by the National Oceanic and Atmospheric Administration (NOAA) aircraft have greatly increased the information available for more accurate predictions of these storms. This paper will investigate several factors that may contribute to hurricane rapid intensification and assess their predictive value for Atlantic storms. A multivariable regression model was fit to the data to quantify predictive relationships and to provide a tool that can be used to help forecast rapid intensification of future storms. The analysis indicated a strong relationship between observed meteorological and physical storm data and hurricane rapid intensification, and the model produced an acceptable forecast success when applied to recent Atlantic storm history.

Keywords
Hurricane, Hurricane strength, Forecasting hurricane strength, Storm, Intensification, Landfall, Hurricane forecast models, Predicting hurricane intensification

Graham Houghton, The Pingry School, 131 Martinsville Rd, Basking Ridge, NJ 07920, USA.
grahamhoughton06@gmail.com

Introduction

Hurricanes constitute a major threat to coastal communities along the Atlantic, Caribbean, and Gulf of Mexico every summer from May through November. The proportion of storms experiencing rapid intensification tripled from 1982 to 2009, suggesting that existing forecasting models based on historical storm data may be insufficient to account for today's climate environment (1). On average, 14 named storms and 7 hurricanes form each year in the Atlantic with 3 of these becoming major storms rated Category 3 or above on the Saffir-Simpson scale with winds exceeding 110 *mph* (2). Storm activity and associated costs have risen in recent years with > $660 billion in property damage from 23 major hurricanes during the 2017-2023 period. Hurricanes have resulted in over 3,600 fatalities in the United States alone during this period with most of these deaths occurring during major hurricanes classified as Category 3 or greater by the National Hurricane Center (3).

Global warming is thought to contribute to this increasing frequency and severity of storms as rising atmospheric, oceanic, and sea surface temperatures create more favorable conditions for hurricane development. Recent storms have demonstrated an increased likelihood of undergoing rapid intensification, defined by a 30*kt* (knots) increase in windspeed within a 24 hour time period (4). Sudden unexpected surges in hurricane strength are extremely difficult for local governments to manage and expose coastal communities to greater risk. Likewise, unpredicted rapid weakening can lead to repeated false alarms, thereby contributing to a lack of public response to evacuation orders.

Hurricane forecasting has evolved in parallel with technological advances over the last 100 years. Before the 1940s, data was collected only for storms that made landfall or by chance ship encounters at sea. Aircraft surveillance of hurricanes began in the 1940s leading to the development of the first forecast models using early computers for calculations (5). The first weather satellite was launched in April 1960 and provided aerial views of global cloud formations (6). Today, information is collected from a global network of geostationary and polar orbiting satellites as well as from the center of storms by Hurricane Hunter aircraft, a joint effort of the United States Air Force and NOAA (7). The vast amount of data currently available has led to the development of forecasting models that are capable of producing predictions for the path and intensity of storms. Forecasts for a given storm are generally updated every 6 hours to incorporate the latest observed conditions.

During the 1986-1990 period, rapid intensification was observed on average 9 times per season, and in 2016-2020 this frequency spiked to 24.8 times per season (8). While forecast models have significantly improved over time, forecast errors for rapidly intensifying hurricanes are 2 to 3 times larger than those for other storms as current models lack training on new climate patterns to be able to confidently predict the behavior of today's storms (9). According to recent research, in the 2016-2020 period the best performing model for the Atlantic hurricane basin correctly forecast rapid intensification in only 21% of cases where it occurred with a 50% false alarm ratio (8).

This study aims to develop a tool that helps forecast rapid intensification of hurricanes by fitting a regression model to those evolving environmental conditions that have been linked to the increase in this phenomenon. It is limited in scope to Atlantic hurricanes that recently made landfall to address the urgent need to establish more confidence for emergency management purposes and to control for land interaction effects. Utilizing the extensive network of weather data now available for recent storms, the study verifies the predictive value of these atmospheric and storm variables and establishes a basis for future investigation of their relationships.

Methods

Multiple linear regression

This paper utilized statistical regression techniques to model the relationship between explanatory and response variables. Regression produces a mathematical equation that best fits a given set of input data. The optimization of a multiple linear regression formula is achieved by minimizing the sum of the squared error terms of the regression equation. Consider a dataset that includes n observations of the response variable y_i for i from 1 to n and k explanatory variables $x_{i,1}$, $x_{i,2}$, ... $x_{i,k}$ also for i from 1 to n. A model regression equation can then be defined by equation 1, where the coefficients β_0, β_1, ... β_k are selected such that the sum of the squared error terms (SSE) is minimized, as shown in equation 2.

$$y_i = \beta_0 + \beta_1 x_{i,1} + \ldots + \beta_k x_{i,k} + \epsilon_i \tag{1}$$

$$SSE = \sum_{i=1}^{n} \epsilon_i^2 \tag{2}$$

The coefficient of determination (R^2) of a regression model measures the proportion of variation in a response variable that can be explained by the coefficients of the regression equation. R^2 varies between 0 and 1 with higher values indicating that the equation provides a strong predictive relationship.

$$R^2 = 1 - \frac{SSE}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3}$$

An F-test of significance for a regression model produces a statistic indicating the likelihood that the model tested is not better than simply using the response variable average. The F statistic is distributed according to an F distribution with k and $n-k-1$ degrees of freedom. The associated value on the cumulative density function for F represents $1-p$ or the likelihood that the regression model provides a better prediction for the response variables in place of $\bar{y}$. Given k explanatory variables and n observations, the F-test statistic is calculated from equation 4, where an analysis of residuals identifies bias or types of patterns that could indicate a non-linear relationship between the variables.

$$F = \frac{\frac{\sum_{i=1}^{n}(y_i - \epsilon_i - \bar{y})^2}{k}}{\frac{\sum_{i=1}^{n}\epsilon_i^2}{n-k-1}} \tag{4}$$

Hurricane selection

To capture the changing climate conditions that are driving the increase in rapid intensification, data was collected for Atlantic storms that reached hurricane strength in recent years. Since key satellite data used for the study was not collected prior to 2012, the study included storms from 2012 to 2023. In addition, to control for factors that may be different as a storm approaches land and to address the immediate need for greater confidence around emergency management plans, only those storms that made landfall were considered. During this period there were 40 Atlantic storms that made landfall. A list of these storms used to construct the dataset and calibrate the model can be found in Table 1.

Table 1. Atlantic hurricanes making landfall 2012 - 2023

Hurricane	Date	Location	Hurricane	Date	Location
Ernesto	Aug 7, 2012	Mexico	Isaias	Aug 3, 2020	N. Carolina
Isaac	Aug 29, 2012	Louisiana	Laura	Aug 27, 2020	Louisiana
Sandy	Oct 25, 2012	Cuba	Nana	Sep 3, 2020	Belize
Ingrid	Sep 16, 2013	Mexico	Sally	Sep 16, 2020	Louisiana
Arthur	Jul 4, 2014	N. Carolina	Gamma	Oct 3, 2020	Mexico
Earl	Aug 5, 2016	Mexico	Delta	Oct 9, 2020	Louisiana
Hermine	Sep 2, 2016	Florida	Zeta	Oct 28, 2020	Louisiana
Matthew	Oct 4, 2016	Haiti	Eta	Nov 3, 2020	Nicaragua
Otto	Nov 24, 2016	Nicaragua	Iota	Nov 16, 2020	Nicaragua
Franklin	Aug 7, 2017	Mexico	Elsa	Jul 7, 2021	Florida
Harvey	Aug 25, 2017	Texas	Grace	Aug 19, 2021	Mexico
Irma	Sep 10, 2017	Florida	Henri	Aug 22, 2021	Rhode Island
Katia	Sep 8, 2017	Mexico	Ida	Aug 29, 2021	Louisiana
Maria	Sep 20, 2017	Puerto Rico	Nicholas	Sep 14, 2021	Texas
Nate	Oct 8, 2017	Mississippi	Ian	Sep 28, 2022	Florida
Florence	Sep 14, 2018	N. Carolina	Julia	Oct 9, 2022	Nicaragua
Michael	Oct 10, 2018	Florida	Lisa	Nov 2, 2022	Belize
Barry	Jul 13, 2019	Louisiana	Nicole	Nov 10, 2022	Florida
Dorian	Sep 1, 2019	Bahamas	Franklin	Aug 23, 2023	Dom. Republic
Hanna	Jul 25, 2020	Texas	Idalia	Aug 30, 2023	Florida

Explanatory variables and data collection

Recent research has linked rapid intensification to a number of meteorological factors such as sea surface temperature, relative humidity, and wind shear (9). Additionally, oceanic heat content provides further information about the heat energy contained beneath the immediate surface of the ocean. Separately from these environmental factors, a hurricane's physical characteristics such as wind speed, central barometric pressure, Saffir-Simpson strength, and overall kinetic energy potentially provide predictive value for rapid intensification.

Hurricane data for these variables was collected and evaluated at 72 hours prior to landfall ($^{-72}$). 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 forecast path with the need to provide sufficient time for emergency management. Hurricane data was then collected at landfall to determine actual intensification over the 3-day (72 hours) period.

For each of the storms in Table 1, the following information was collected in order to analyze possible factors influencing storm intensity changes during the 72 hours prior to landfall. 1] Average sea surface temperature (*SST*) along the hurricane's eventual path as of 72 hours prior to landfall in $^{\circ}C$ from the NOAA Physical Sciences Laboratory High Resolution Dataset (10). 2] Average oceanic heat content (*OHC*) along the hurricane's eventual path as of 72 hours prior to landfall in kJ/cm^2 from the University of Miami Rosenstiel School (11). 3] Average atmospheric moisture content measured through total precipitable water (*ATPW*) along the hurricane's eventual path as of 72 hours prior to landfall in *mm* from the NOAA Center for Satellite Applications and Research (12). 4] Layer averaged wind shear (*WS*) along the hurricane's eventual path as of 72 hours prior to landfall in *kts* from the CIMSS Tropical Cyclones Data Archive (13). 5] Saffir-Simpson strength category (*SS*) as of 72 hours prior to landfall from the NHC and Central Pacific Hurricane Center Tropical Cyclone Advisory Archive (14). Tropical Storms, Tropical Depressions and Disturbances were assigned values of 0, −1, and −2 respectively for *SS*. 6] Storm wind speed (*VMAX*) in *mph* and central barometric pressure (*MSLP*) in *mb* as of 72 hours prior to landfall and at the time of landfall from the NHC and Central Pacific Hurricane Center Tropical Cyclone Advisory Archive (14). Hurricane integrated kinetic energy (*HIKE*) in *TJ* was calculated as of 72 hours prior to landfall and at the time of landfall. In order to calculate *HIKE*, windfield data from NHC tropical cyclone forecast advisories (14) was entered into the Integrated Kinetic Energy Calculator available from NOAA (15).

Explanatory and response variable assessment

An analysis was conducted on the collected data to determine which of the explanatory variables would likely provide the best regression model inputs for prediction of rapid intensification. This analysis was performed for three potential response variables in order to identify the most useful response variable for the model. The response variable candidates analyzed were the 72-hour change in storm windspeed ($\Delta VMAX^{72}$), which is a direct measurement of rapid intensification, as well as the 72-hour changes in *MSLP* and *HIKE* ($-\Delta MSLP^{72}$ and $\Delta HIKE^{72}$). The correlations of each explanatory variable against the three response variable candidates are shown in Table 2.

Table 2. Input variable averages (μ) and correlation between input variables and response variable candidates

	μ	$\Delta VMAX^{72}$	$-\Delta MSLP^{72}$	$\Delta HIKE^{72}$
SST^{-72}	29.3	0.3260	0.2986	0.0524
OHC^{-72}	82.0	0.3912	0.4482	0.1304
$ATPW^{-72}$	48.7	0.0075	−0.0002	−0.2942
WS^{-72}	22.3	−0.3439	−0.3666	0.1021
$VMAX^{-72}$	56.5	−0.6535	−0.4049	−0.0515
$MSLP^{-72}$	997	0.6241	0.4037	0.0217
SS^{-72}	−0.1	−0.5949	−0.3605	−0.1537
$HIKE^{-72}$	9.2	−0.6284	−0.4109	0.0952
	$Avg(\lvert\rho\rvert)$	0.4462	0.3367	0.1126

The strongest overall correlations were found between the explanatory variables and $\Delta VMAX^{72}$, with an absolute correlation of 0.4462, making $\Delta VMAX^{72}$ the most suitable response variable for purposes of the regression model. Although significant correlation was also found between the explanatory variables and $-\Delta MSLP^{72}$, with an absolute correlation of 0.3367, it was smaller than the average for $\Delta VMAX^{72}$. The explanatory variables did not correlate well to $\Delta HIKE^{72}$ with an absolute correlation of 0.1126 which suggested that $\Delta HIKE^{72}$ was not well explained linearly by the variables chosen.

With respect to the meteorological explanatory variables, the initial correlation assessment found a strong positive correlation between the sea surface temperature and oceanic heat content with $\Delta VMAX^{72}$ while wind shear was strongly negatively correlated. Atmospheric moisture showed near zero correlation. This could suggest that a larger structural weather feature associated with high moisture was limiting *ATPW's* impact on intensification. The storm's physical characteristics showed very strong negative (positive in the case of *MSLP* which was inversely related to storm strength) correlation to $\Delta VMAX^{72}$. This finding indicated that storms were unlikely to maintain their current condition for 72 hours. That is, storms were more likely to weaken from strength or to intensify from weakness.

Results

Regression model

Incorporating the identified inputs and the selected response variable, the initial regression relationship is shown in Equation 5.

$$\Delta \widehat{VMAX}_i^{72} = \beta_0 + \beta_1 SST_i^{-72} + \beta_2 OHC_i^{-72} + \beta_3 ATPW_i^{-72} + \beta_4 WS_i^{-72} + \beta_5 VMAX_i^{-72} + \beta_6 MSLP_i^{-72} + \beta_7 SS_i^{-72} + \beta_8 HIKE_i^{-72} + \epsilon_i \quad (5)$$

The constants β_0, β_1, β_2, ..., and β_8 were determined by minimizing the sum of the squared error terms. The resulting model, including the coefficients that minimize *SSE*, for prediction of rapid intensification is shown in Equation 6.

$$\begin{aligned}\Delta \widehat{VMAX}_i^{72} &= -42.49 + 3.27 \times SST_i^{-72} + 0.32 \times OHC_i^{-72} \\ &-0.08 \times ATPW_i^{-72} - 0.28 \times WS_i^{-72} - 0.64 \times VMAX_i^{-72} \\ &+0.02 \times MSLP_i^{-72} + 3.82 \times SS_i^{-72} - 0.73 \times HIKE_i^{-72}\end{aligned} \quad (6)$$

A test of significance was performed on the explanatory variables to determine which inputs had the greatest influence on the results. Only *OHC* was significant near the 10% level indicating potential correlation among the explanatory variables and a need to restrict the inputs in the final regression model. The *p*-values and correlations for the explanatory variables are shown in Table 3 below.

Table 3. Explanatory variable initial significance test results and correlations

p-value	Variable	SST^{-72}	OHC^{-72}	$ATPW^{-72}$	WS^{-72}	$VMAX^{-72}$	$MSLP^{-72}$	SS^{-72}	$HIKE^{-72}$
0.536	SST^{-72}	1.000	0.552	0.083	−0.31	−0.14	0.082	−0.14	−0.09
0.103	OHC^{-72}		1.000	0.144	−0.22	−0.11	0.083	−0.09	−0.15
0.874	$ATPW^{-72}$			1.000	−0.12	0.101	−0.14	0.195	−0.01
0.604	WS^{-72}				1.000	0.304	−0.36	0.309	0.283
0.442	$VMAX^{-72}$					1.000	−0.97	0.953	0.767
0.987	$MSLP^{-72}$						1.000	−0.93	−0.74
0.733	SS^{-72}							1.000	0.675
0.197	$HIKE^{-72}$								1.000

In order to address the low significance and potential redundancy in the initial regression inputs, explanatory variables were removed sequentially, starting with *MSLP* (*p*-value = 0.987), until all remaining inputs had satisfactory *p*-values. This process resulted in a final regression model with *OHC*, *VMAX*, and *HIKE* as the explanatory variables. While all of the variables except *ATPW* showed strong correlation to $\Delta VMAX^{72}$, the three variables ultimately used provided the strongest value as independent predictors. The final regression model is presented in Equation 7 and the updated significance test results are shown in Table 4.

$$\Delta\widehat{VMAX}_i^{72} = 46.26 + 0.41 \times OHC_i^{-72} - 0.50 \times VMAX_i^{-72} - 0.76 \times HIKE_i^{-72} \quad (7)$$

Table 4. Explanatory variable final significance test results

Variable	*p*-value
OHC^{-72}	0.00998
$VMAX^{-72}$	0.02167
$HIKE^{-72}$	0.13325

Given the relatively high correlation (ρ = 0.767) between *VMAX* 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 were all < 2.5, well within an acceptable range for non-collinearity.

Table 5. Explanatory variable variance inflation factors

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

Model fit

A graph of actual increase in storm windspeed over 72 hours ($\Delta VMAX_i^{72}$) along the *x*-axis versus the model's predicted storm windspeed increase ($\Delta\widehat{VMAX}_i^{72}$) along the *y*-axis for each of the 40 storms is shown in Figure 1. The plot shows a linear trend with reasonable correlation throughout the range.

R^2 and significance test

The R^2 of the regression model was 0.557, meaning that the model is able to explain 55.7% of the variation in $\Delta VMAX^{72}$. In addition, an *F* test of significance for the model parameters using Equation 4 with $k = 3$ and $n = 40$ resulted in an *F* statistic of 15.113, which has a corresponding *p* value of 1.5×10^{-6}. The relatively high R^2 value and low *p* value together indicate a significant relationship between the explanatory and response variables which provided the basis for the model.

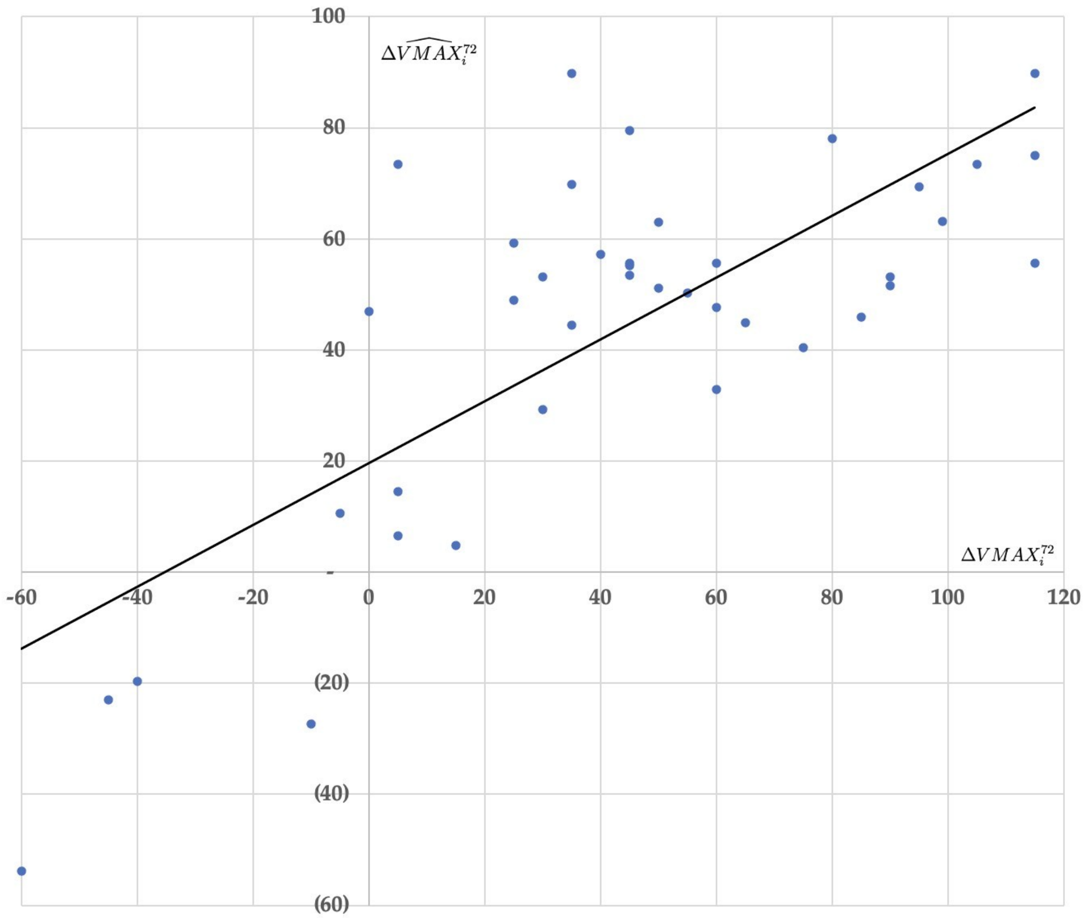

Figure 1. Model Prediction Plot $\Delta VMAX_i^{72}$ vs. $\Delta\widehat{VMAX}_i^{72}$

Residual assessment

In order to assess the linearity assumption of the model, the residual terms were graphed against each explanatory variable in the final model presented in Equation 7. Each residual graph displayed a satisfactory scatterplot without signs of bias, indicating that the linear model was appropriate for the data supplied. The graphs of the error terms versus each explanatory variable are shown in Figures A1 through A3 in Appendix A.

Discussion

Rapid intensification forecast accuracy

The regression model presented in Equation 7 can be employed to generate rapid intensification forecasts for hurricanes 72 hours prior to anticipated landfall. When the model is applied to the 40 storms listed in Table 1, it provides a reasonable prediction for the change in hurricane windspeed. After unit converting the 30 (knot) threshold for rapid intensification into 34.5 (miles per hour), an inspection of the actual windspeed change ($\Delta VMAX_i^{72}$) for each of the 40 historical storms showed that 26 of them met this intensification criteria, while 14 did not. The model's predicted change in windspeed ($\Delta\widehat{VMAX}_i^{72}$) provided an accurate prediction of rapid intensification for 25 of the 26 storms that rapidly intensified. Among the 14 storms that did not rapidly intensify, the model correctly identified 9. The results of the model forecast accuracy are shown in Table 6.

Table 6. Model forecast accuracy

Actual Rapid Intensification	Model Forecast Rapid Intensification	
	Yes	No
Yes	25	1
No	5	9

The overall forecast success of 85.0% (34/40) indicates the model's strong predictive potential. The associated precision of the model, defined by the percentage of correct rapid intensification forecasts, is 83.3% (25/30). Low precision could lead to unnecessary evacuation orders and wasted resources. The recall, or detection rate of the model, defined by the percentage of actual rapid intensification events forecast by the model, was 96.2% (25/26). Low detection could result in a dangerous storm making landfall without adequate warning.

Training and testing datasets

In order to assess the methodology's performance against storms not included in the model development, an alternate regression equation was calibrated to a training set of 30 storms randomly selected from the 40 storms in Table 1. The input values for the 10 remaining storms were then used as a testing set for the alternate regression model to produce $\Delta\widehat{VMAX}_i^{72}$ forecasts for each storm. The associated rapid intensification predictions demonstrated high accuracy, with an overall forecast success rate of 90% (9/10), a precision rate of 87.5% (7/8), and a detection rate of 100% (7/7).

Comparison with NHC forecasts

To compare model forecasts to those of the NHC, the forecast landfall windspeed of each storm at 72 hours prior to landfall was collected from the NHC forecast advisories available in the archive [14]. These figures were used to generate an NHC-forecast-based $\Delta\widehat{VMAX}_i^{72}$ for each storm, which was then evaluated against the 34.5*mph* rapid intensification threshold.

The NHC forecasts for the 40-storm dataset have a calculated R^2 of 0.475 and thus explain 47.5% of the variation in $\Delta VMAX^{72}$ compared to 55.7% explained by the regression model. The NHC forecasts accurately predicted the rapid intensification outcome for 26 of the 40 storms versus the 34 outcomes correctly predicted by the model. The NHC forecasts were less likely to predict strong increases in intensity reflected by the high 92.3% forecast precision rate combined with a low 50% detection rate. The results of the NHC forecast accuracy are shown in Table 7 and a comparison of forecast accuracy metrics for NHC and the regression model are presented in Table 8. A detailed listing of the actual and forecast windspeed changes by storm appears in Table B1 in Appendix B.

Table 7. NHC Forecast accuracy

Actual Rapid Intensification	NHC Forecast Rapid Intensification	
	Yes	No
Yes	13	13
No	1	13

Table 8. Comparison of NHC and model forecast accuracy

	NHC	Regression Model
Forecast Success	65.0%	85.0%
Forecast Precision	92.3%	83.3%
Detection	50.0%	96.2%

It is important to note that the model was calibrated specifically using the storms in the analysis. To provide an unbiased comparison of the NHC and the regression model's predictive abilities, the analysis was repeated using the alternative testing model. For the 10-storm testing set, the NHC forecasts have a calculated R^2 of 0.555 and thus explain 55.5% of the variation in $\Delta VMAX^{72}$ compared to 57.3% explained by the alternative regression model. The NHC forecasts accurately predicted the rapid intensification outcome for 8 of the 10 storms in the testing set versus 9 of 10 outcomes correctly predicted by the alternative regression model. A listing of the actual and forecast windspeed changes by storm for the testing set analysis appears in Table 9.

Outlier storms

Figure 1 shows that windspeed increases were over-predicted for storms with $\Delta VMAX_i^{72}$ in the 0 - 40*kt* range. An inspection of the underlying data for the 6 storms with the highest error values ($\Delta \widehat{VMAX}_i^{72}$- $\Delta VMAX_i^{72}$) found the landfall locations of these storms to be Mexico (4), Belize, and the Dominican Republic. This group includes Hurricane Earl, which can be seen in Table 9 as an outlier in the test model results. A number of factors such as land interaction, coastal currents, upwelling, and the continental shelf could be inhibiting rapid intensification for these landfall areas. The continental shelf does not extend far from shore at these locations, potentially leading to greater upwelling of cold water as hurricanes approach land.

Table 9. Actual and Model Forecast Windspeed Change and Rapid Intensification

Hurricane	Actual $\Delta VMAX^{72}$	NHC $\Delta \widehat{VMAX}^{72}$	Test model $\Delta \widehat{VMAX}^{72}$	Rapid intensification Actual	NHC	Test model
Idalia	90	57	52	Yes	Yes	Yes
Sandy	80	33	68	Yes	No	Yes
Laura	85	39	41	Yes	Yes	Yes
Eta	105	51	65	Yes	Yes	Yes
Earl	0	9	48	No	No	Yes
Florence	-40	8	-6	No	No	No
Katia	40	34	55	Yes	No	Yes
Ida	115	74	76	Yes	Yes	Yes
Delta	-45	-19	5	No	No	No
Arthur	65	40	47	Yes	Yes	Yes

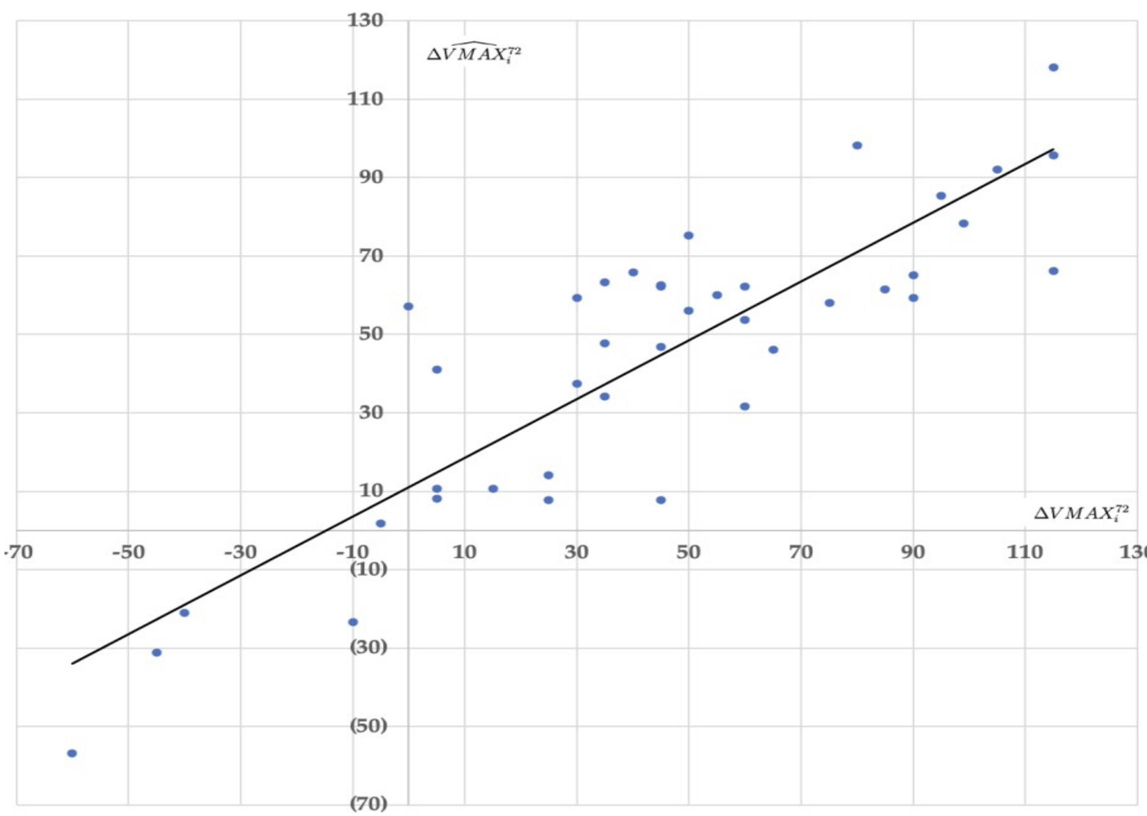

Figure 2. Model with LWD Prediction Plot $\Delta VMAX_i^{72}$ vs. $\Delta \widehat{VMAX}_i^{72}$

In order to test the performance of the regression with the addition of continental shelf type at landfall, an indicator variable, *LWD*, was added to the model. Locations with gradual shelf dropoff (n = 33) were assigned *LWD* = 0, while locations with steep shelf dropoff (n = 7) were assigned *LWD* = 1. The results were encouraging with the refined model producing an R^2 of 0.75. Further research across a large number of storms is needed to investigate *LWD* or potentially an alternate variable as the source of moderated intensification in the outlier storms. The regression model including *LWD* is presented in Equation 8 and a prediction plot is shown in Figure 2.

$$\Delta \widehat{VMAX}_i^{72} = 40.36 + 0.65 \times OHC_i^{-72} - 0.58 \times VMAX_i^{-72} - 0.79 \times HIKE_i^{-72} - 54.95 \times LWD \quad (8)$$

Limitations

The overall lack of extensive storm data was a limiting factor of the regression model. Given that satellite data for key variables became available in 2012, and that the study was limited to landfall hurricanes, the dataset consisted of only 40 storms from the last 12 years.

The environmental variables used to calibrate the regression model were collected with knowledge of the hurricanes' actual paths, although at 72 hours prior to landfall the NHC track forecasts have a relatively narrow range of uncertainty and the variables would likely have similar readings at points within that range. Utilization of the model for predictions of future storms in real time with unknown future paths will require reliance on the accuracy of the NHC track forecast.

It should also be noted that rapid intensification can occur at any time during a hurricane's development, but since this model is specifically designed to predict rapid intensification within 3 days of landfall, its use is limited to this timeframe.

Conclusion

The mathematical model presented in this report is intended to help validate rapid intensification predictions of more robust forecasting models for Atlantic hurricanes actively threatening coastal communities and to enhance the information available to make critical decisions surrounding disaster preparedness. Overall, the model accurately detected rapid intensification in the historical dataset with a detection accuracy of 96.2% with an overall forecast accuracy of 85%. The required inputs consisted of readily available environmental data from the storm's projected path at 72 hours prior to forecast landfall. The study found the most significant explanatory variables for the final regression model to be oceanic heat content, storm wind speed, and hurricane integrated kinetic energy.

The climate environment is continuously evolving, and the data available for the study of rapid intensification is expanded greatly each year. As such, the model should be tested against future actual experience to refine, recalibrate, and potentially improve its forecasting ability. Outliers observed in the more southern regions of the study, especially in the Caribbean Sea, suggest that additional investigation is needed surrounding potential factors that may inhibit or interact with a storm's intensity as it approaches land. Future research may also expand the geographic region of study to the Eastern and Western Pacific areas threatened by tropical cyclones and include rapid intensification analysis for the entirety of a storm's path.

References

1. Bhatia KT, Vecchi GA, Knutson TR, Murakami H, Kossin J, Dixon K, et al. Recent increases in tropical cyclone intensification rates. *Nature Communications.* 10, 635 (2019). https://doi.org/10.1038/s41467-019-08471-z

2. National Hurricane Center and Central Pacific Hurricane Center, Tropical Cyclone Climatology (2024). https://www.nhc.noaa.gov/climo/

3. National Centers for Environmental Information, Billion-Dollar Weather and Climate Disasters (2024). https://www.ncei.noaa.gov/access/billions/summary-stats

4. Kaplan J, DeMaria M. Large-Scale Characteristics of Rapidly Intensifying Tropical Cyclones in the North Atlantic Basin. Weather and Forecasting. 18, 1093-1108, (2003). https://doi.org/10.1175/1520-0434(2003)018%3C1093:LCORIT%3E2.0.CO;2

5. University of Rhode Island Graduate School of Oceanography, Brief History of Hurricane Forecast Models (2020). https://hurricanescience.org/science/forecast/models/modelshistory/

6. National Environmental Satellite, Data, and Information Services, Celebrating 60 Years of the World's First Weather Satellite (2020). https://www.nesdis.noaa.gov/news/celebrating-60-years-of-the-worlds-first-weather-satellite

7. United States Air Force, 53rd Weather Reconnaissance Squadron "Hurricane Hunters" (2021). https://www.403wg.afrc.af.mil/About/Fact-Sheets/Display/Article/192529/53rd-weather-reconnaissance-squadron-hurricane-hunters/

8. DeMaria M, Franklin JL, Onderlinde MJ, Kaplin J. Operational Forecasting of Tropical Cyclone Rapid Intensification at the National Hurricane Center. Atmosphere. 12(6), 683 (2021). https://doi.org/10.3390/atmos12060683

9. Bhatia K, Baker A, Yang W, Vecci G, Knutson T, Murakami H, et al. A potential explanation for the global increase in tropical cyclone rapid intensification. Nature Communications. 13, 6626 (2022). https://doi.org/10.1038/s41467-022-34321-6

10. NOAA Physical Sciences Laboratory, NOAA OI SST V2 High Resolution Dataset (2024). https://www.psl.noaa.gov/mddb2/makePlot.html?variableID=4115

11. University of Miami Rosenstiel School Upper Ocean Dynamics Laboratory, North Atlantic Ocean Heat Content (2024). https://isotherm.rsmas.miami.edu/heat/weba/atlantic.php

12. NOAA Center for Satellite Applications and Research, MIRS High Resolution Products (2024). https://www.star.nesdis.noaa.gov/mirs/highresolutionv.php

13. Cooperative Institute for Meteorological Satellite Studies / University of Wisconsin-Madison, CIMSS Tropical Cyclones Data Archive Request (2024). https://tropic.ssec.wisc.edu/archive/

14. National Hurricane Center and Central Pacific Hurricane Center, Tropical Cyclone Advisory Archive (2024). https://www.nhc.noaa.gov/archive/

15. National Oceanic & Atmospheric Administration Hurricane Research Division, Integrated Kinetic Energy (2024). https://www.aoml.noaa.gov/hrd/ike/Calculator_AllQuad.php

Appendix A

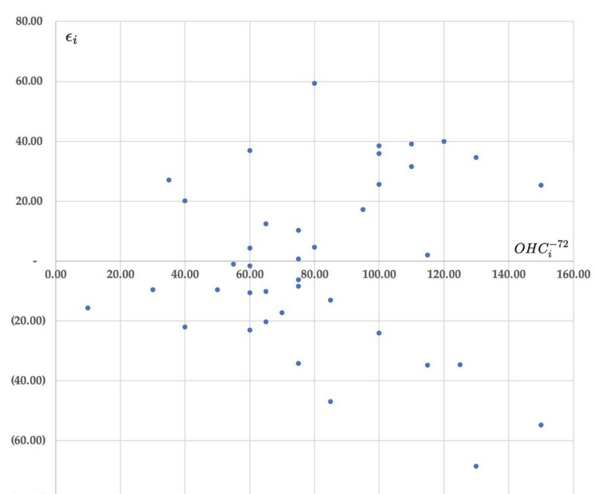

Figure A1. Model Error Plot OHC_i^{-72} vs. ϵ_i

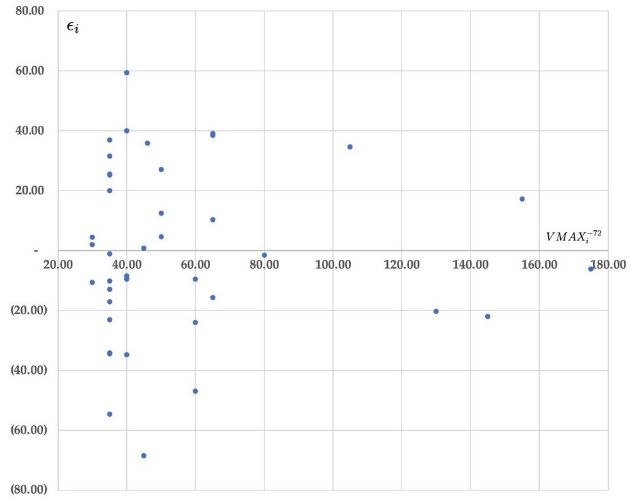

Figure A2. Model Error Plot $VMAX_i^{-72}$ vs. ϵ_i

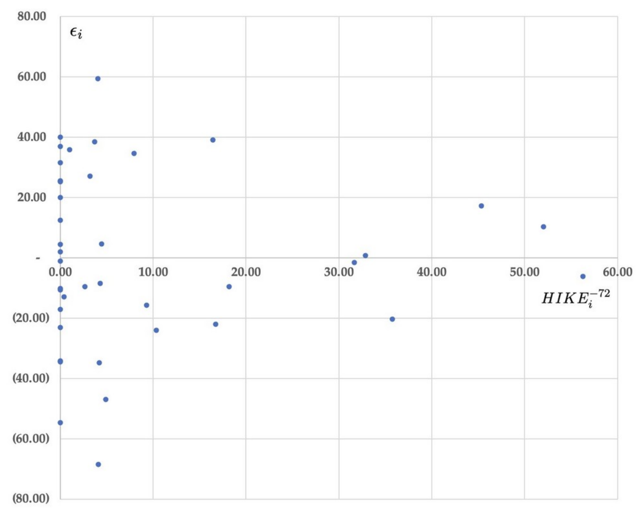

Figure A3. Model Error Plot $HIKE_i^{-72}$ vs. ϵ_i

Appendix B

Table B1. Actual, NHC, and Model Windspeed Change and Rapid Intensification

	Actual	NHC	Model	Rapid Intensification		
Hurricane	$\Delta VMAX^{72}$	$\Delta \widehat{VMAX}^{72}$	$\Delta \widehat{VMAX}^{72}$	**Actual**	**NHC**	**Model**
Ernesto	25	15	49	No	No	Yes
Isaac	15	39	5	No	Yes	No
Sandy	80	33	78	Yes	No	Yes
Ingrid	30	28	53	No	No	Yes
Arthur	65	40	45	Yes	Yes	Yes
Earl	0	9	47	No	No	Yes
Hermine	45	23	55	Yes	No	Yes
Matthew	-10	-29	-27	No	No	No
Otto	60	25	48	Yes	No	Yes
Franklin ('17)	25	17	59	No	No	Yes
Harvey	95	34	69	Yes	No	Yes
Irma	-60	-43	-54	No	No	No
Katia	40	34	57	Yes	No	Yes
Maria	90	62	51	Yes	Yes	Yes
Nate	50	46	51	Yes	Yes	Yes
Florence	-40	8	-20	No	No	No
Michael	115	41	75	Yes	Yes	Yes
Barry	45	56	56	Yes	Yes	Yes
Dorian	75	27	40	Yes	No	Yes
Hanna	60	16	56	Yes	No	Yes
Isaias	5	-11	7	No	No	No
Laura	85	39	46	Yes	Yes	Yes
Nana	35	18	70	Yes	No	Yes
Sally	55	48	50	Yes	Yes	Yes
Gamma	35	5	90	Yes	No	Yes
Delta	-45	-19	-23	No	No	No
Zeta	60	25	33	Yes	No	No
Eta	105	51	73	Yes	Yes	Yes
Iota	115	81	56	Yes	Yes	Yes
Elsa	5	-3	15	No	No	No
Grace	45	-1	80	Yes	No	Yes
Henri	-5	21	11	No	No	No
Ida	115	74	90	Yes	Yes	Yes
Nicholas	35	23	44	Yes	No	Yes
Ian	99	92	63	Yes	Yes	Yes
Julia	50	34	63	Yes	No	Yes
Lisa	45	35	53	Yes	Yes	Yes
Nicole	30	24	29	No	No	No
Franklin ('23)	5	18	73	No	No	Yes
Idalia	90	57	53	Yes	Yes	Yes