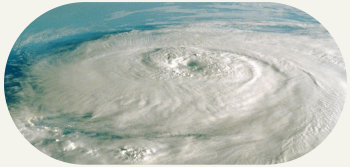
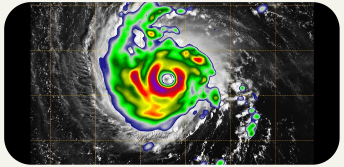
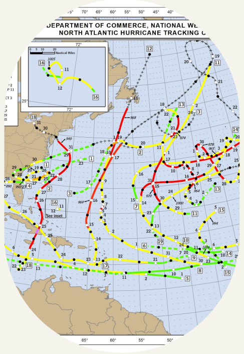
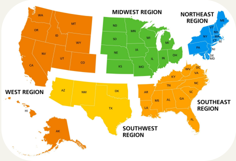
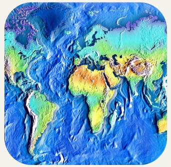
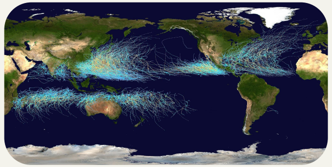
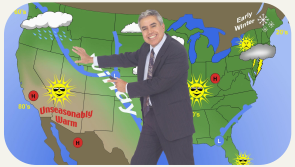
Predicting Hurricane Behavior:
Meteorological, Spatial, and Temporal Predictors

By: Lucy, Erika, Justin, & Jillian




Background: Why Study Hurricanes?

Basic Overview

- Hurricanes form over warm ocean water when heat and moisture cause air to rise and create lower surface pressure; as air continues to rush in & rise, the storm rotates and intensifies.
- Lower central pressure leads to stronger winds
- Hurricanes are ranked on a scale of 1-5 based on maximum sustained wind speed, with 5 being the most intense

Importance?

- Hurricanes are among the most destructive and common natural disasters
- Yearly, about 3 out of 14 hurricanes reach major hurricane status, which is any hurricane above a Category 3 level
- Being able to predict which storms will reach extreme intensities can save lives and guide emergency planning and preparedness

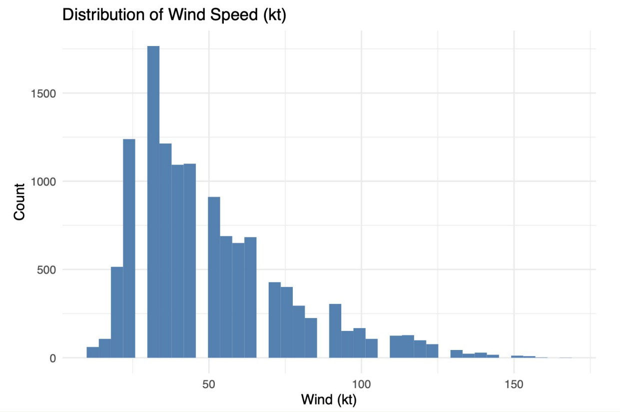
Research Question:

To what extent can early meteorological measurements, such as wind speed, central pressure, and storm location, predict whether a tropical hurricane will intensify into a major hurricane (Category 3-5)? Which variables are most significant in category prediction?

Data & Methodology:

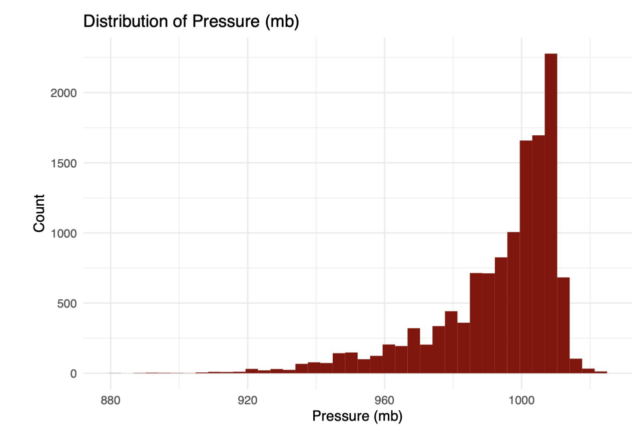
Important Variables:

- **Pressure:** Lower central pressure = stronger storm; drives wind speed.
- **Wind speed:** Direct measure of intensity; predictor for major classification.
- **Track Coordinates (latitude/longitude):** Storms behave differently by region (wind shear, water temps).
- **Date:** Seasonal cycles influence intensity (warmer waters mid-season).
- **Category:** Storm category reflects sustained wind speed; higher categories indicate more intense, damaging storms and help classify major (Category 3–5) hurricanes.
- **Naming Status:** Named storms typically reach higher intensity and are better monitored; early pressure and wind patterns can help predict whether a storm will become named.

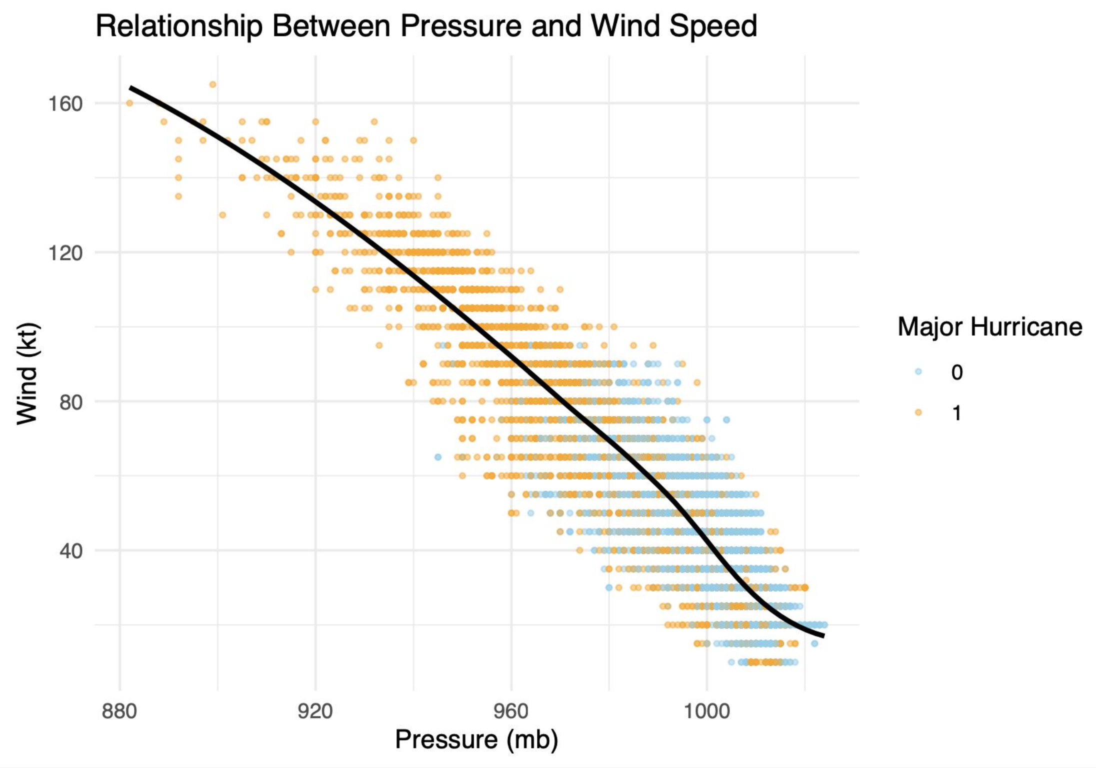


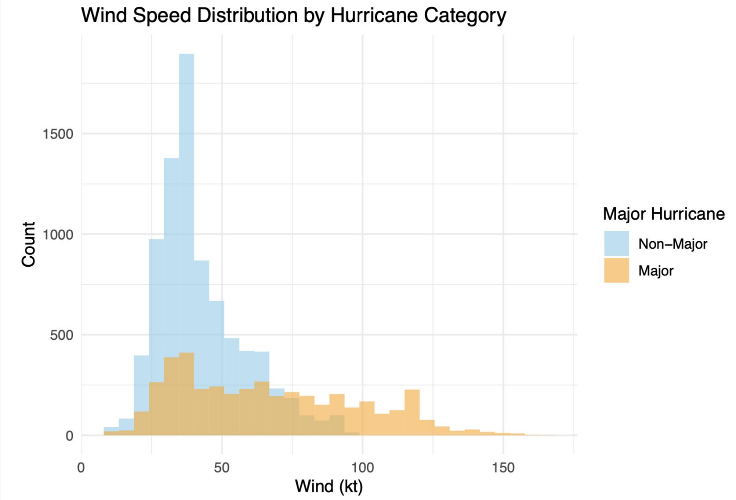
EDA and Cleaning:

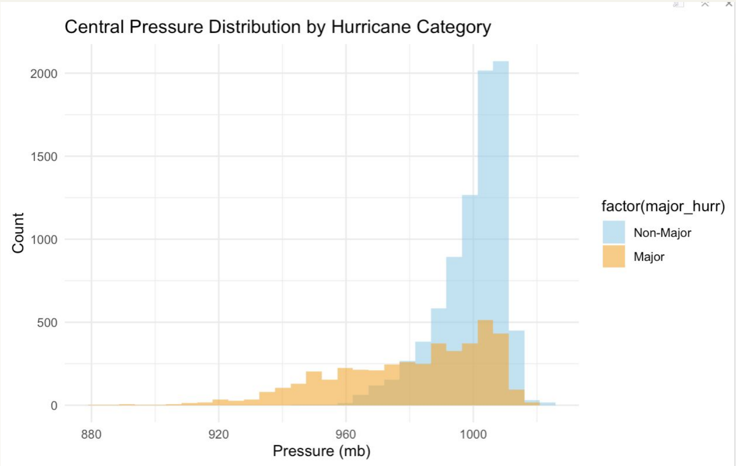
- Began with ~13,665 storm observations before cleaning.
- Removed missing values for pressure, wind speed, and tracking gaps
- Removed entries with impossible/incomplete values (e.g., pressure = 0; tropical storms with wind speed = 0)
- Feature engineered variables **category** (from max wind), **region** (from coordinates), and **hemisphere** (from serial number provided)
- After cleaning we had approximately 12,674 torm observations



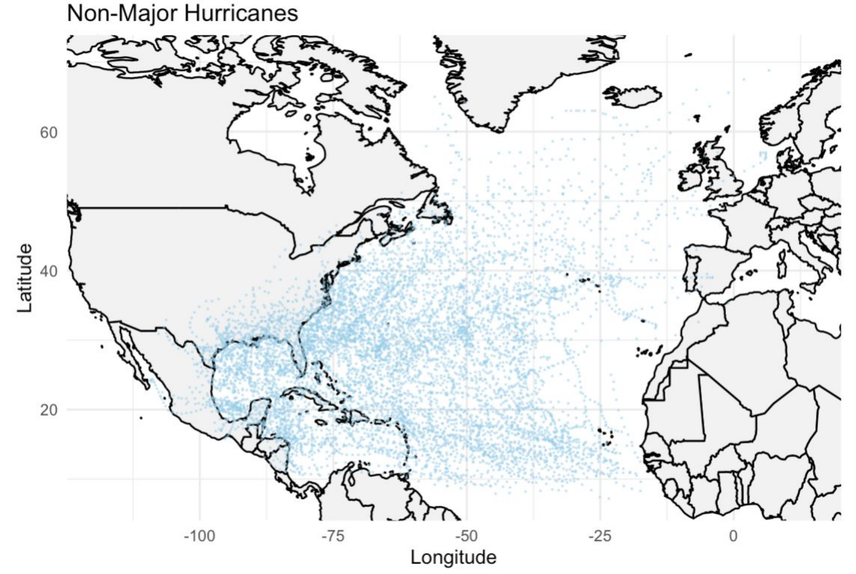
Initial Visuals:

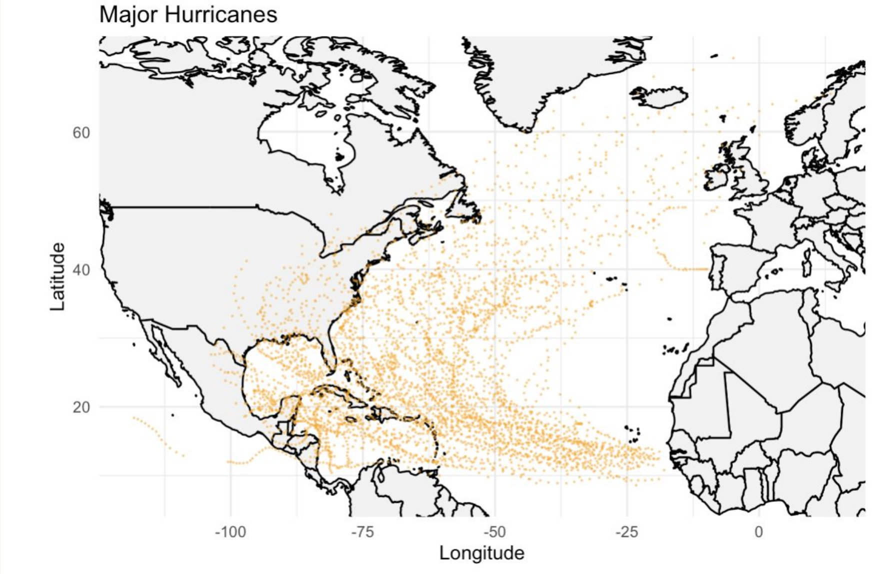






Initial Visuals:





Models: Logistic Regression

Model 1: Main Effects Only

Predictors: wind speed, pressure, latitude, longitude, region.

- **AIC: 12,435**
- **AUC: 0.797**

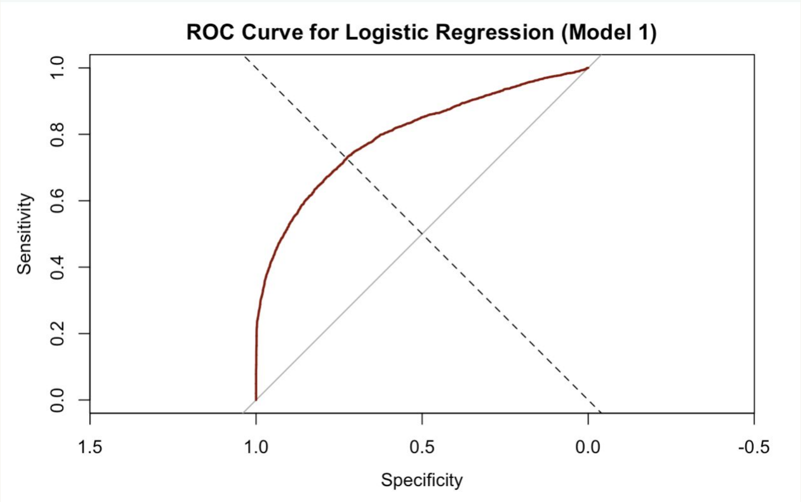
Threshold: 0.5
Model's ability to classify major hurricanes is around 48%
Models ability to classify non-major hurricanes was about 92.7%
Balanced accuracy: 0.7035

Optimal threshold from ROC curve of 0.336
Model's ability to classify major hurricanes is around 64.3%
Models ability to classify non-major hurricanes was about 78.1%
Balanced accuracy: 0.712

- Lower pressure → higher odds of major hurricane
- HIgher winds early in storm → higher odds of major hurricane

Limitation: Missed many major hurricanes → suggests nonlinear/conditional effects not captured.

- Under-predicts major hurricanes (high false-negative rate)
- Suggests important interactions or nonlinearities are missing



Models: Logistic Regression

Model 2: Added Interaction Terms

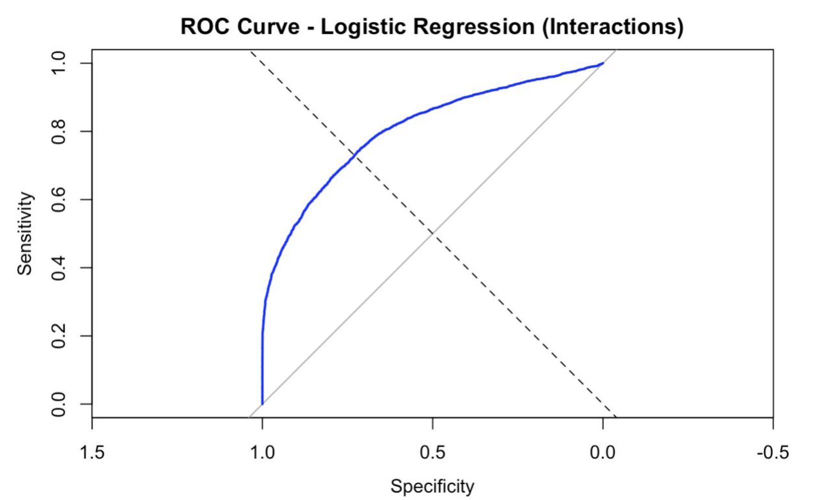
Added interactions: wind × pressure and wind × latitude.

Impact:

- Wind effect depends on pressure (combined low pressure + high wind → higher intensification odds).
- Wind influence varies by latitude (high winds at higher latitudes less predictive).
- Region remained important; certain regions (South, Northeast, sparse areas) show lower odds

AIC dropped to 12138
AUC value of .8055

Optimal threshold from ROC curve of 0.3091
Model's ability to classify major hurricanes improved from about 64% to 67%
Models ability to classify non-major hurricanes was about 80% (which was similar to model 1)
Balanced Accuracy = 0.731



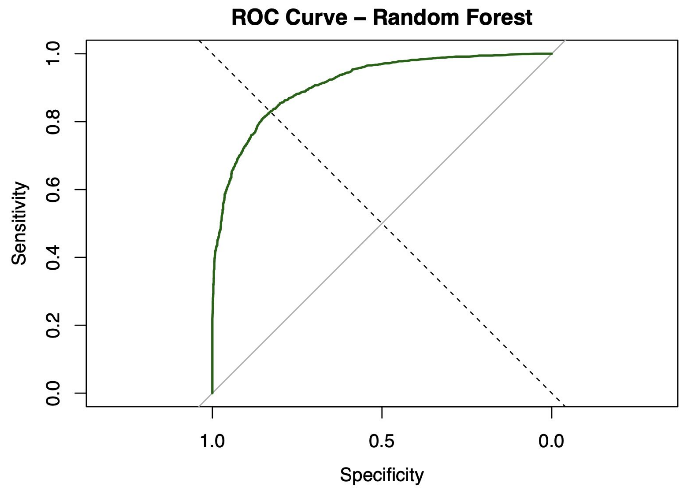
Key Takeaways:

- Interaction terms capture conditional meteorological relationships important for intensification.
- Model 2 is more accurate and contains nuance that model 1 does not

Models: Random Forest

Model 3: Random Forest

- Predictors: wind speed, pressure, latitude, longitude, and cleaned regional classification.
- Motivation: Logistic regression could not fully capture nonlinear patterns and complex interactions.
- Model’s ability to classify major hurricanes was 81.8% (better detection of major hurricanes)
- Models ability to classify non-major hurricanes was about 85% (strong identification of non-major storms))
- AUC: 0.911, outperforming logistic regression (AUC ~0.805).



Model 4: Enhanced Random Forest with Interaction Features

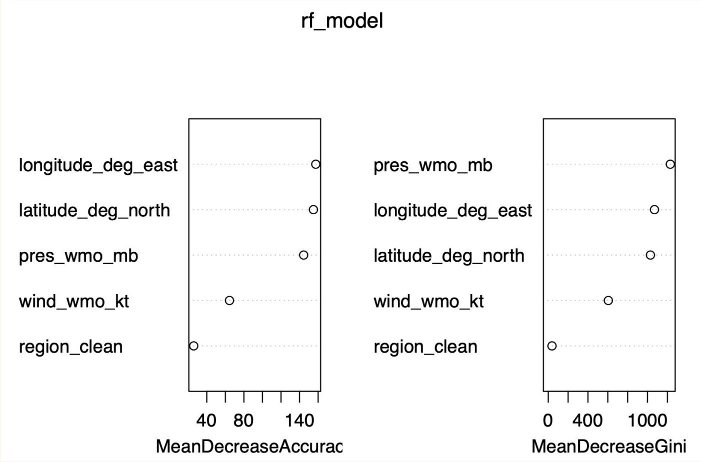
- Added the same engineered interaction terms from our 2nd logistic regression model
- Further captured nonlinear and conditional effects where:
  - Wind’s impact depends on pressure.
  - Wind's predictive value varies with latitude.
- Adding the terms allows the model to detect more complex meteorological relationships that single predictors cannot reveal.
- AUC: 0.8994 (slightly lower than simpler random forest)
- Model’s ability to classify major hurricanes was 0.766 (lower than model 3))
- Model’s ability to classify non-major hurricanes was 85.8% (better than model 3)
- AUC and accuracy drop slightly – RF already takes interaction into account so this model was slightly redundant

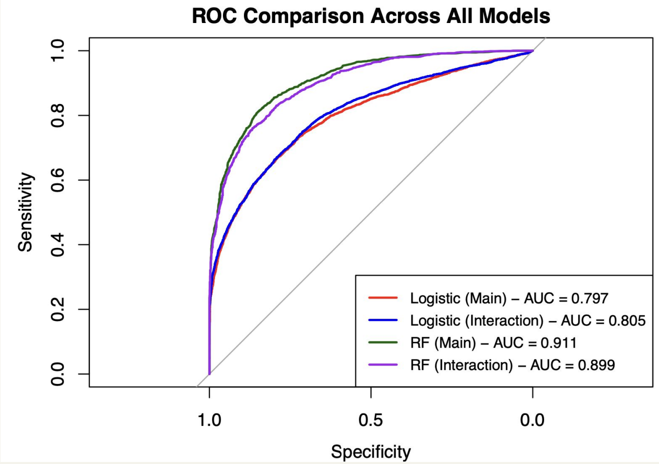
Confusion Matrix and Statistics

	Reference	
Prediction	0	1
0	2114	249
1	372	1067

Results & Analysis: Predicting Major Hurricanes

- Best performing model: Model 3 – Random Forest using wind speed, pressure, latitude, longitude, and cleaned regional classification.
- The variable importance in our Random Forest Models helps identify which meteorological factors matter most, even though individual trees are complex.
- Final models ability to classify major hurricanes was around 82%, which is a relatively high accuracy
- The AUC of 0.91 indicates the model reliably separates major vs non-major hurricanes
- The most influential predictors across the models where wind speed and pressure,
    - High wind speed & low pressure highly increased the likelihood of a major hurricane classification
- With our final model, we can predict a major or non major hurricane with relatively high confidence, improving early warning potential for major storms





Limitations:

Feature Limitations:

Some important meteorological factors (e.g., sea-surface temperature, humidity) were not available, limiting model completeness.

Accuracy of earlier observations in the data could be lower due to less technological advancements at the time

Major hurricanes are less common than non-major storms, resulting in a class imbalance that can skew models

Data Cleaning Limitations:

Removing missing or impossible values (e.g., pressure = 0) may introduce selection bias by excluding atypical storms.

Filtering incomplete tracks could remove storms that behaved differently, reducing diversity in training data.

Model Limitations:

Logistic regression struggled with nonlinear relationships, potentially underestimating key interactions.

Random Forests, while powerful, can be less interpretable, making it harder to understand

No model will be able to fully capture the complex dynamics of hurricane intensification.

Models trained on historical patterns may not fully account for changing climate conditions, which influence storm behavior over time

Conclusion & Importance

Key Findings

01 Central pressure and wind speed are the strongest indicators of future intensification

02 Storm location (latitude + region) meaningfully influences intensification potential

03 Predictive accuracy for early intensification is reasonably high

Thus: Early meteorological measurements can predict if a tropical cyclone will intensify into a major hurricane with strong accuracy. These predictions can help emergency responders, governments, and at risk communities to make earlier and more informed decisions to reduce the damage potential major category hurricanes could cause

References:

- Emanuel, K. (2021). *Tropical cyclones and climate change*. Annual Review of Earth and Planetary Sciences, 49, 1-32.

- Holland, G., & Bruyère, C. (2014). *Recent intense hurricane trends in the context of climate change*. Nature, 509, 349-352.

- Knutson, T., et al. (2019). *Tropical cyclones and climate change assessment: Part I: Detection and attribution*. Bulletin of the American Meteorological Society, 100(10), 1987-2007.

- Vecchi, G., & Knutson, T. (2008). *On estimates of historical North Atlantic tropical cyclone activity*. Journal of Climate, 21(14), 3580-3600.

Thank You!