STEM Fellowship Big Data Challenge 2019-2020

Predicting Hurricane Frequency and Intensity Using Ocean Data and Boosting Algorithms

Alex Zhuang*, Jason Xiong, Baker Jackson, Nathan Kim

University of Toronto Schools

November 10, 2020

*corresponding author

Abstract

Recent years mark some of the most destructive hurricanes ever recorded in the Atlantic, causing record-breaking levels of damage to coastal communities and the environment. The increase in hurricane intensity has been hypothesized to be linked to anthropogenic ocean warming, highlighting the dangers of accelerating climate change. In this paper, we investigate the proposed relationship using a two-part analysis of over 40 years of oceanographic data and hurricane records in the North Atlantic ocean. Using recent advances in machine learning and boosting algorithms, we aim to make sense of the large amounts of data used in climate stewardship. First, we compare the overall trend in surface temperatures (SST), sea-level pressures and wind shear in the North Atlantic since 1980 with corresponding trends in hurricane intensity as measured by the Saffir-Simpson Hurricane Wind Scale (SSHWS). Second, we partition the measurements by geographic location and create features based on the monthly average measurement of each variable (SST, wind shear, pressure). Using these features as input, we trained two gradient boosting models to predict the monthly average frequency and intensity respectively of a storm in the North Atlantic.

1 Keywords

XGBoost, Hurricane, Saffir-Simpson, Sea Surface Temperature.

2 Introduction

In 2019, the “ultra-intense” Hurricane Dorian caused 4.6 billion USD in damages across Canada, the US and the Bahamas, becoming the second-most powerful Atlantic storm recorded. Over 400,000 cases of power loss occurred in Nova Scotia during this event, affecting up to 80% of the province’s population. Recordbreaking hurricanes in 2017 such as Irma and Harvey regularly exceeded forecasts, while policymakers continue to make cuts to the funding of disaster relief agencies. In 2018, President Donald Trump proposed $61m in cuts to the Federal Emergency Management Agency’s Pre-Disaster Mitigation grant, despite the National Institute of Building Sciences estimating that "every $1 spent can save $6 in future disaster costs.” The potential rise in intensity and frequency of hurricanes raises concerns regarding the impact of tropical anthropogenic warming on extreme weather, and recent trends may act as precursors to future climate-related catastrophes.

In this study, we investigate how different factors linked to climate change intensify the impact of hurricanes, through the analysis of over 40 years of oceanographic data and hurricane records. This data can be used by relief groups in the mitigation of risk for susceptible regions and to spur strong climate action and effective governance of the global commons.

Several conditions that precipitate cyclogenesis were chosen as potential indicators in our analysis. Firstly, tropical winds should accelerate with rising sea surface temperatures (SST), as higher temperatures stimulate the evaporation and heat transfer which fuels the formation of hurricanes (3) (4). According to the literature, a minimum threshold of 26.5 $^{\circ}$C was determined to be necessary for hurricane formation (4). Wind shear is furthermore investigated due to the role that trans-Atlantic wind drafts play in the formation of hurricanes as well as their disruption of the formation of a lowpressure “eye” necessary for the rotation of thunderstorms. The maximum wind speed, duration, and frequency of hurricanes are used to measure future impacts. The Saffir-Simpson Hurricane Wind Scale is used to qualitatively describe the expected damage of a hurricane.

3 Materials & Methods

3.1 Variables

The independent variables determined were first theoretically suggested in literature (3).

- SST

- Wind shear
- Pressure

The three characteristics of hurricanes measured are as follows:

- Saffir-Simpson Hurricane Wind Scale (SSHWS)
- Hurricane duration
- Maximum wind speed

These variables best represent the potential impacts of future hurricanes. The SaffirSimpson wind scale, ranging from 1 to 5, is a measure of the damage a hurricane is able to cause based on its wind speed.

3.2 Datasets

We collected data from four, large scale, internationally compiled datasets obtained from the *National Oceanic and Atmospheric Administration* (NOAA) and the *National Center for Atmospheric Research* (NCAR):

- Extended Reconstructed Sea Surface Temperature v5 (ERSST) (5)
- The International Best Track Archive for Climate Stewardship (IBTracs) (6)
- Blended Sea Winds (7)
- NCAR Gridded Pressure (8)

Sea surface temperature (SST) data is provided by the ERRST dataset, derived from the *International Comprehensive OceanAtmosphere Dataset*. SST data is measured for every month from 1854 to 2019 and is stored on a 2 $^\circ$ by 2 $^\circ$ grid (roughly 200km by 200km), obtained from Argo floats with spatial completeness enhanced using statistical methods. Spatial coverage ranges from 88.0N to 88.0S and to 0.0E to 358.0E (5).

Wind shear data is provided by the Blended Sea Winds dataset, which stores high-resolution ocean surface vector winds on a global 0.25 $^\circ$ grid from July 9, 1987, to the present day with a maximum resolution of 6-hours.

Sea-level surface pressure data is recorded in the NCAR Gridded dataset, with a 2.5 $^\circ$ by 2.5 $^\circ$ spatial coverage between 90N - 90S and 0E - 357.5E.

Finally, all hurricane data is provided by the International Best Track Archive for Climate Stewardship (IBTracs) dataset, a compilation of best-track datasets from ten meteorological organizations across the globe. Measurements for sustained wind speed (knots), central pressure (mb) and storm center of circulation (latitude/longitude) going back to 1841 are stored within a spatial resolution of 1°.

3.3 Analysis

In our preliminary exploration of the increase in surface temperatures and hurricane intensities over time, we used gridded monthly sea surface temperature (SST) data and cyclone event data since 1980 from ERSST and IBTracs respectively.

To obtain a monthly average for SST in the North Atlantic region, we averaged all data points within each month over 2° by 2° latitude and longitude grids between 28°S and 42°N, and 98°W and 14°E.

To obtain yearly measures of hurricane intensity we averaged the maximum Saffir-Simpson wind scale and storm duration in the respectively in the North Atlantic region over each year and plotted them over time.

Although our analysis is focused on the North Atlantic region, the methods in this study are equally applicable to the other regions measured in this dataset.

In the first phase of analysis, we visualize general trends in the above data. For this purpose, the measurements for each variable are preprocessed into monthly mean time series using the Python library Pandas. For each month for which data is recorded, all measurements of a particular variable in the given month are selected and averaged; the resulting monthly average is entered into a Pandas DataFrame with two columns: the month, and the monthly average measurement for the given month. In order to observe the trend in each variable over time, we fit a linear model to each monthly mean time series using the linear regression tools in scikitlearn, a Python library for scientific computing. Each linear model takes the form

$$y = \alpha + \beta t,$$

where t is the time in months since January 1980, y is the monthly average measurement for the given variable, and α and β are coefficients determined by the linear regression and specific to each model. Then, the coefficient of determination R^2 for each linear model is calculated in order to quantify the degree of correlation between each variable and changing time. For simple linear estimators, the R^2 value is calculated as the square of the Pearson correlation coefficient R, where

$$R = \frac{n(\sum ty) - \sum t - \sum y}{\sqrt{[n\sum t^2 - (\sum t)^2][n\sum y^2 - (\sum y)^2]}}$$

In addition, the relationship between oceanographic data and hurricane characteristics is explored in a similar fashion.

For the second phase of analysis, we investigate whether oceanographic data could be used to predict the frequency and intensity of North Atlantic hurricanes using machine learning techniques. We divide each dataset into a grid of partitions based on latitude and longitude, and for each of 288 months from October 1987 to September 2011, average measurements during the previous month of each oceanographic variable in each partition are aggregated as distinct features in a feature vector. For SST data, partition monthly averages from three months prior were aggregated into features. Two models, using these feature vectors as inputs, are then trained to predict, given a particular month's feature vector, the average SSHWS and hurricane frequency of the next month across the entire North Atlantic ocean. These models used (10)XGBoost, a gradient boosting framework that forms an ensemble of weak decision trees to classify the data. In total, the models were trained on 9343 features. Models are trained using an 80:20 train-test split and are evaluated and optimized on the squared error of their predictions from the true value. In each epoch of training, the model evaluates its predictions on all the training data, then makes predictions for all examples in the test data. In order to prevent overfitting, the training process is stopped if the accuracy rate on the test data does not improve for 10 consecutive epochs.

To optimize the two models, Grid Search Cross Validation was performed, which tested 300 different sets of parameters for each of the models. The parameters that were tested were the maximum depth of a tree, the sub-sampling value, the η value, and the minimum loss reduction value required to make a partition on a leaf node of the tree.

4 Results

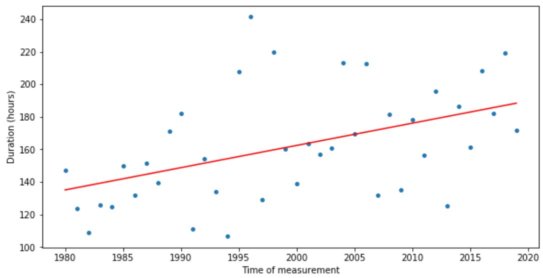

Figure 1: Change in Average Cyclone Lifespan Over Time. This plot shows the growth of cyclone lifespan from 1980, with the linear regression modeling the data with an R^2 value of 0.217. For n = 40, this value indicates a statistically significant relationship at p = 0.05.

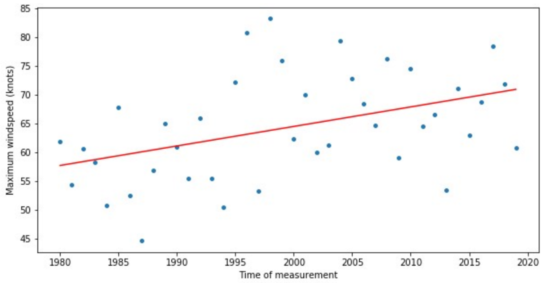

Figure 2: Change in Average Cyclone Maximum Windspeed Over Time. This plot shows the increasing relationship between wind speed and time, with the linear regression having an R^2 value of 0.1822. For n = 40, this value indicates a statistically significant relationship at p = 0.05.

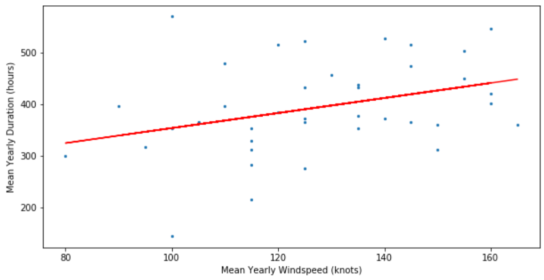

Figure 3: Correlation Between Mean Windspeed and Storm Duration. The plot shows an increasing relationship between storm duration and windspeed. The linear approximation of the data has an R^2 value of 0.119. For n = 40, this value indicates a statistically significant relationship at p = 0.05.

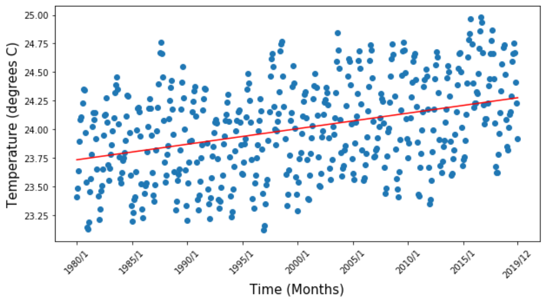

Figure 4: Monthly Sea Surface Temperature Over Time. This relationship shows the effects of climate change on increasing sea surface temperature. The R^2 value of the linear approximation is 0.144. For n = 40, this value indicates a statistically significant relationship at p = 0.05.

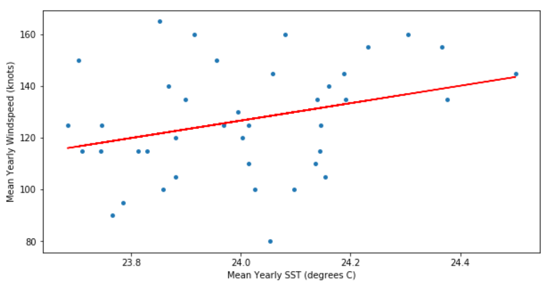

Figure 5: Correlation Between Windspeed and Sea Surface Temperature. The data shows a weak positive correlation between the two variables, where the linear approximation has an R^2 value of 0.098. For n = 40, this value indicates a statistically significant relationship at p = 0.05.

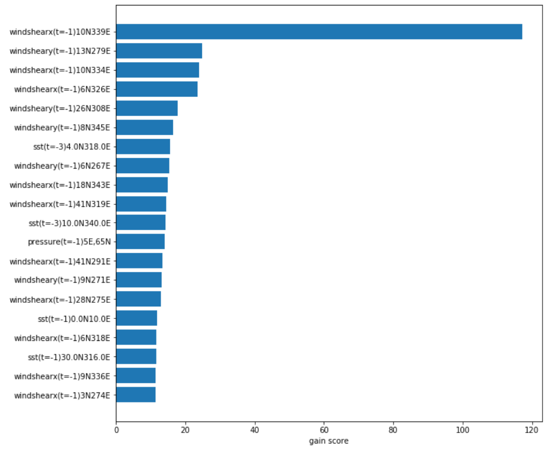

Figure 6: Feature Importance Values of Frequency Model. Gain is a measure of the relative importance of a feature in predicting the outcome of the model. While this model performed quite poorly in comparison with the SSHWS model, these features remained differentiated predictors of cyclone event frequency within the next month.

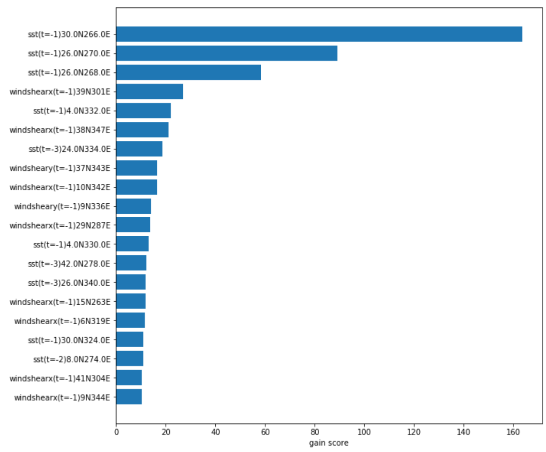

Figure 7: Feature Importance Values of SSHWS Model. While the top 20 features are dominated by a variety of different data categories, roughly 74% of all non-zero importance features were wind shear data.

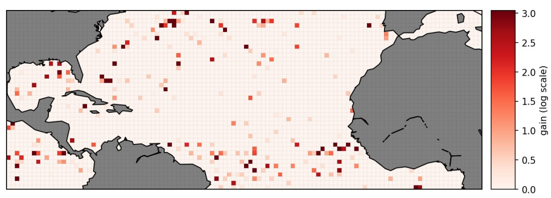

Figure 8: Figure 9: Feature importance of EastWest windshear by 2°partitions in the SSHWS model. Below a latitude of 12°N a qualitative increase in feature importance can be observed in the figure.

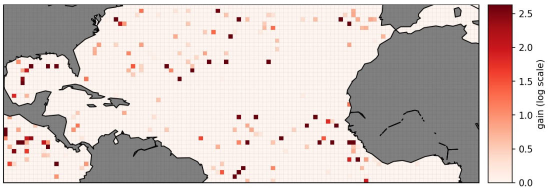

Figure 9: Feature importance of North-South windshear by 2°partitions in the SSHWS model. Relatively higher importance values are shown in the Gulf of Mexico, throughout the central Atlantic, and along the African coast.

5 Discussion

5.1 Data Analysis

As demonstrated by Figure 1 and Figure 2, there is a strong positive correlation between time and the duration and intensity of hurricanes. With the increasing duration, frequency, and intensity of hurricanes as well as the increase in sea temperatures, we then analyze the specific correlation between these variables. The analysis of changing sea surface temperatures in Figure 4 indicates the specific rate of ocean warming, with approximately a 0.5° increase in the past 40 years. Ocean warming is increasing at an alarming rate, and its relationship with hurricane windspeed in Figure 5 indicates it has a moderate impact on rising hurricane intensity. The Pearson coefficient is used as a measure of the statistical significance because of its capacity to measure bi-variate association.

5.2 Frequency Boosting Model

Based on the Grid Search Cross-Validation, the final parameter values for the Frequency XGBoost Model were a max tree depth of 5, $\eta = 0.1$, $\gamma = 0$, and a sub-sample of 0.75.

The final model had an accuracy of 55.17% when predicting from testing data and a mean score across 3-fold cross-validation of 0.6285. Furthermore, the model predicted 89.29% of all testing data within one storm of the actual data point.

In Figure 6, it can be seen that the most important feature for predicting the frequency of storms is the wind shear in the North-South Direction at 10° North 21° West, followed by the wind shear in the East-West Direction at 13° North, 81° West.

5.3 Intensity Boosting Model

Based on the Grid Search Cross-Validation, the final parameters for the Intensity Boosting Model were as follows: maximum tree depth of 8, a sub-sample of 0.75, $\gamma = 10$, and $\eta = 0.5$.

The final model had an accuracy of 91.38% when classifying the intensity of future storms on the SSH scale. Furthermore, the model predicted 94.64% of all testing data within one category of the actual testing data. This model shows the very strong connections between pressure, wind shear, and sea surface temperature when predicting hurricane intensity.

From Figure 7, it can be seen that the most influential features when modelling intensity of storms are the sea surface temperature at 30° North, 94° West, followed by sea surface temperature at 26° North, 90° West. The data suggests that this section of the ocean will be crucial when predicting the intensity of future hurricanes. The importance of each feature is indicated in Figure 8 and Figure 9 to show the relative importance of each region when modelling hurricane intensity.

5.4 Limitations of Methods

While XGBoost is a powerful model, it comes with certain limitations. Our dataset containing 288 data samples, 231 samples of which were part of the train set, was likely a relatively small size for the training process. This quantity of data may have limited the complexity of the model and its ability to generalize to more recent data.

Furthermore, the data used for both frequency and intensity was imbalanced. Because most months do not have hurricanes, a large portion of the data consisted of imputed values of a frequency of 0 and an SSHWS of -3. Without sufficient representations of each predicted class, the model would be susceptible to noise and overfitting. K-Fold Cross-Validation and oversampling of rare data points were applied to ensure that the model was balanced; however, this issue could not be fully mitigated.

6 Conclusions

Through our research, we have established both the relationship between rising ocean temperatures and its implications on hurricane formation and intensity. Our predictive model is furthermore able to predict the intensity of future hurricanes with 91.38% accuracy, as well as mapping many of the key features. This allows for more precise measurements in the most important regions of the ocean for pressure, sea surface temperature, and wind shear, which will support the development of better hurricane models in the future.

The devastation caused by tropical storms in developing nations highlights the importance of investment into hurricane-response mechanisms to reduce long-lasting impacts on growing communities. Analysis of key features will be instrumental in the prediction of future risks, many months in advance, to best aid in mitigating hurricane damage before they strike.

We strongly encourage governments and the technology sector to increase the use of different types of machine learning techniques to model the climate; the development of these methods has seen near exponential growth, representing a wealth of opportunity for breakthrough discoveries.

For further study, we highlight the use of more sophisticated time-series forecasting, including the consideration of more granular data to identify more consistent spatial and temporal trends, as relevant areas. Furthermore, although our analysis was limited to the North Atlantic ocean, much of our data supports can support similar analyses on other regions of the world's oceans and we encourage others to do so. Finally, including monetary damages as an observed variable would be invaluable in quantifying the monetary impact of extreme weather on cities, and in determining investment into relief funds or infrastructure to mitigate risk. With the onset of additional extreme weather events such as forest fires and extreme temperatures, the study of the wide-ranging impacts of climate change and utilization of big data sources will be critical in disaster relief. Continual monitoring and the development of improved predictive models will be key in responding to and preventing future climate challenges in a changing world.

Acknowledgements

We would like to extend our gratitude to everyone who helped make this project possible. We would like to thank our teacher Mr. Wilson for his invaluable guidance and support, as well as Jacob Hazen for mentoring us throughout this process. Finally, we would like to extend special recognition to the STEM Fellowship for this amazing opportunity to challenge ourselves and create an impact on the world.

References

[1] Masters, Jeff. "A Review of the Atlantic Hurricane Season of 2019." Scientific American Blog Network, Scientific American, 25 Nov. 2019, blogs.scientificamerican.com/eye-of-the-storm/ a-review-of-the-atlantic-hurricane-season-of-2019/ .

[2] "America Is Good at Dealing with Hurricanes on the Mainland-after They Strike." The Economist, The Economist Newspaper, 1 June 2018, www.economist. com/graphic-detail/2018/06/01/ america-is-good-at-dealing-with-hurricanes -on-the-mainland-after-they-strike.

[3] Emanuel, K. A. The theory of hurricanes. Annu. Rev. Fluid Mech. 23, 179–196 (1991).

[4] Stull, Roland. Meteorology for Scientists and Engineers. Dept. of Earth, Ocean & Atmospheric Sciences, 2011. https:

//www.eoas.ubc.ca/books/Practical_Meteorology/mse3/Ch16-Hurricane.pdf

[5] Boyin Huang, Peter W. Thorne, Viva F. Banzon, Tim Boyer, Gennady Chepurin, Jay H. Lawrimore, Matthew J. Menne, Thomas M. Smith, Russell S. Vose, and Huai-Min Zhang (2017): NOAA Extended Reconstructed Sea Surface Temperature (ERSST), Version 5. [indicate subset used]. NOAA National Centers for Environmental Information. doi:10.7289/V5T72FNM

[6] Schreck, Carl & National Center for Atmospheric Research Staff (Eds). Last modified 06 Nov 2013. The Climate Data Guide: IBTrACS: Tropical cyclone best track data. Retrieved from https://climatedataguide.ucar.edu/climate-data/ ibtracs-tropical-cyclone-best-track-data.

[7] Zhang, Huai-Min. "Blended Sea Winds." National Climatic Data Center, NOAA, www.ncdc.noaa.gov/data-access/marineocean-data/blended-global/blended-sea-winds.

[8] Kalnay et al.,The NCEP/NCAR 40-year reanalysis project, Bull. Amer. Meteor. Soc., 77, 437-470, 1996.

[9] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

[10] Chen, Tianqi, and Carlos Guestrin. "XGBoost." Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD '16 (2016): n. pag. Crossref. Web.