Continuation/Research Progression Projects Form (7)

Required for projects that are a continuation/progression in the same field of study as a previous project.

This form must be accompanied by the previous year's abstract and Research Plan/Project Summary.

Student's Name(s) Kyra McCreery

To be completed by Student Researcher: List all components of the current project that make it new and different from previous research. The information must be on the form; use an additional form for previous year and earlier projects.

Components	Current Research Project	Previous Research Project: Year: 2018-2019
1. Title	Associations between the Slowdown in North Atlantic Tropical Cyclone Translation Speed and Intensifying Storm Precipitation	Assessing Regional North Atlantic Hurricane Landfall Frequencies Through Poisson Regression Modeling and Probability Simulations
2. Change in goal/ purpose/objective	The present study aims to quantify the slowdown in translation speed of North Atlantic hurricanes from 1851-2016 and link the observed decline in translation speed to storm precipitation intensity. This research does not examine the relationship between climatic variables and hurricane frequency, and the focus of the project is an investigation of hurricane speed and rainfall rather than an assessment of new methodologies.	The objective of this research is to develop a novel Poisson regression methodology for the modeling of North Atantic landfalling hurricane frequency as a function of five key atmospheric and oceanic variables. The project also incorporates a probability simulation that determines the regional likelihood of hurricane frequency based on variables identified as statistically significant in the Poisson analysis.
3. Changes in methodology	The current research project is a retrospective data analysis that investigates the translation speed of 1,857 North Atlantic hurricanes, examines the regional and temporal distribution of these speeds through a variety of statistical tests (k-means clustering, k-nearest neighbor outlier detection) and then assesses the rainfall intensity associated with a subset of these storms through a hierarchal clustering assessment.	This study involves three separate analyses that assess the success of Poisson regression hurricane models and relate certain climatic variables to the probability of hurricane landfall in distinct regions. The first assessment is a Kendall's tau correlation assessment, the second component is a Poisson regression analysis and Log-Likelihood quantification of model success, and the third is a probability simulation designed to predict regional hurricane frequency.
4. Variable studied	The shifts in North Atlantic hurricane translation speed over time are examined in this study. Rainfall precipitation intensity is also assessed over time, and the spatial similarities and differences between patterns in rainfall intensity of North Atlantic hurricanes are quantified.	This project assesses five atmospheric and oceanic variables, including El Nino Southern Oscillation and the Atlantic Meridional Mode. The relationship between these variables and the regional frequency of hurricane occurence is analyzed.
5. Additional changes		

Attached are:

☑ Abstract and Research Plan/Project Summary, Year 2018-19

I hereby certify that the above information is correct and that the current year Abstract & Certification and project display board properly reflect work done only in the current year.

Kyra McCreery — Student's Printed Name(s)

Kyra McCreery — Signature

11/19/19 — Date of Signature (mm/dd/yy)

Assessing Regional North Atlantic Hurricane Landfall Frequencies Through Poisson Regression Modeling and Probability Simulations

Kyra McCreery
Earth and Environmental Sciences

A. Rationale

Hurricanes in the Atlantic basin are devastating natural phenomena with immense impacts on the inhabitants, economies, and environments of coastal locations. The development of accurate statistical models that assess the complex relationships between hurricane activity in this region and various climatic factors has been an area of concern for researchers over the past several decades. Our ability to predict the frequency of future hurricanes in the Atlantic basin relies on our understanding of the factors that contribute to the variability in hurricane genesis and movement in this region (Villarini *et. al* 2011).

The use of models that adequately describe the relationship between North Atlantic landfalling hurricanes and global climate indices is central to an improved understanding of this variability. Several studies in recent years have contributed to this mission by investigating the impacts of different climate indices on Atlantic hurricane landfalls. The imminent thermodynamic conditions that hurricanes move through during their lifecycles certainly impact their movement and intensity, but more indirect and long-term climatic factors can modulate atmospheric and oceanic conditions to drastically influence patterns of hurricane activity. Frequently examined indices in recent literature include: El Niño-Southern Oscillation (ENSO; Klotzbach 2010), North Atlantic Oscillation (NAO; Elsner and Jagger 2004), sea surface temperature (SST) anomalies in the tropical regions (Dailey *et al.* 2006), Atlantic Meridional Mode (AMM; Vimont and Kossin 2007), and Atlantic Multidecadal Oscillation (AMO; Zhang and Delworth 2006).

There is no definite consensus among the scientific community as to which combination of the aforementioned hurricane factors produces the most authentic model of North Atlantic hurricane landfalls (Villarini *et. al* 2011). As such, this study will take a holistic approach by initially incorporating several predictive factors. The statistically significant results of this preliminary investigation will then be more closely examined through a probability simulation that will assess the probability of a hurricane landfall in years during which certain climatic conditions are present. This approach will build upon prior research in this area, which has focused heavily on isolating the relationships between specific indices and Atlantic hurricane activity. Klotzbach (2010) assessed the influence of El Niño-Southern Oscillation on Atlantic basin hurricane activity and landfalls, refining the accepted conclusion that hurricane activity is typically lessened during El Niño years. Vimont and Kossin (2007) examined the relationship between AMM and seasonal Atlantic hurricane activity, noting the contributions of AMM to the variability of hurricane activity in the North Atlantic. Dailey *et al*. (2009) investigated the impact of Atlantic SST anomalies on U.S. landfalling hurricane counts. The current research will elaborate on the findings of these prior investigations by integrating a unique combination of indices and assessing individual indices based on initial conclusions.

In this investigation, multiple regression models will be developed for North Atlantic hurricane landfall counts, with a focus on three major landfall regions and specific hurricane “clusters” within these regions. The models developed in this study will be based on the Poisson regression; most hurricane researchers advocate the use of Poisson modeling for analyses of hurricane landfall frequencies, as hurricane landfalls can be represented as counts (non-negative integers), have known average rates, and occur independently of previous landfall events (Klotzbach 2010).

The hurricane landfall count data to be used in this study will be derived from Atlantic basin storm tracks separated into five distinct groups, or “clusters”. The purpose of assessing landfall counts within individualized clusters is to gain insight into the driving mechanisms that contribute to the disparities in the behavior of hurricanes that typically follow different tracks. Rather than group all North Atlantic landfalling hurricanes together, cluster analysis enables one to generate unique models that pertain to specific track patterns.

Nakamura *et. al* (2009) developed a unique clustering methodology using the HURDAT dataset available from the National Hurricane Center that will serve as the basis for the data used in this study. The clusters were developed through the use of a k-means clustering algorithm, an organic methodology for grouping data involving the identification of centroids that define the central points of each cluster and the subsequent allocation of specific data points to the nearest cluster (Nakamura *et. al*

2009). Hurricane clusters produced using this same methodology will be provided for use in this study, the only difference being that an improved dataset (HURDAT2) and a greater number of years (1851-2016) have been incorporated into this newly developed cluster dataset.

B. Research Questions

Over the course of this study, three major questions will be emphasized. These questions are summarized below:

1) How can the frequency of North Atlantic landfalling hurricanes in three distinct regions and five unique clusters be effectively modeled as Poisson processes?
2) Which climate indices indicate the greatest amount of statistical significance in their correlation to regional North Atlantic hurricane landfall frequency?
3) How can these statistically significant indices be further investigated through the use of a probability simulation?

C. Procedure

a. *Data*

1. Hurricane Cluster Data

The North Atlantic hurricane cluster landfall frequencies provided for use in this study were developed from the historical hurricane track dataset HURDAT2. This dataset is maintained by the National Hurricane Center (NHC) and contains updated information on storm position and live conditions such as wind speed and pressure. The HURDAT2 record is considered less accurate pre-1944, the year that reconnaissance aircraft flights began to reliably document hurricane activity in the North Atlantic (Villarini *et. al* 2011). Thus, the models that will be developed in this study will make use of hurricane landfall data from 1950-2016 to reflect the improved reliability of hurricane documentation following the implementation of flight-based data collection.

2. Climate Index Data

Five climate indices will be used as “predictors” in the regression models analyzing North Atlantic hurricane landfalls: Niño 3.4, Niño 1.2, SST anomaly in the Main Development Region (MDR), SST anomaly in the Gulf of Mexico (GOM), and the

Atlantic Meridional Mode (AMM). The indices will be averaged for the April-June season (the period considered pre-hurricane season). Raw yearly data for all five indices between 1950 and 2016 will be provided for use in this study (Nakamura, unpublished data).

b. Kendall's tau *Correlation Assessment*

To begin the assessment of North Atlantic hurricane landfall frequencies, a correlation table will be generated to visually assess initial relationships between the five climate indices (Niño 3.4, Niño 1.2, SST anomaly in the Main Development Region (MDR), SST anomaly in the Gulf of Mexico (GOM), and the Atlantic Meridional Mode (AMM)) and the five cluster tracks in each landfall region. The generated correlations will be Kendall's *tau* correlation coefficients (Kendall's *tau* is a non-parametric rank-based correlation typically used for data that does not fit a normal distribution). A two-tailed t-test will then be performed on each correlation coefficient to assess the statistical significance of the initial associations between each index and the frequency of hurricane landfalls in each cluster.

c*. Poisson Regression Modeling*

A standard Poisson generalized linear model approach will be undertaken to perform the regression analysis. Generalized linear models (GLMs) resolve one of the fundamental weaknesses of the standard linear model in that they account for errors, or residuals, that follow distributions other than normal (Gill 2001). A Poisson GLM will be fitted for each cluster within each landfall region using the five climate indices as dependent variables, or covariates. Each regression will be assessed with regard to AIC criterion; this method of model fitting maximizes the accuracy of a model while minimizing its complexity.

Considering the use of climate indices as covariates in this generalized linear model, it is also possible to envision a scenario in which the Poisson rate is constant and does not change yearly depending on the value of these covariates. If the climate indices are not factored into the model, the probability of a certain number of hurricane landfalls could be expressed using a base Poisson rate calculated by dividing the sum of the landfalls in a specified timeframe by the number of years in this timeframe. This calculation will be performed for all cluster scenarios to provide a point of comparison in assessing the performance of the generalized linear model for each cluster.

d. Log Likelihood Plots

To assess the quality of the Poisson GLM for each cluster and landfall region, a likelihood-ratio test will be used to evaluate the goodness of fit of the generalized linear model against a model in which a constant Poisson rate is assumed across all 67 years of data (1950-2016). This test will determine how many times more likely the correct prediction of hurricane count data under the Poisson GLM is than the base-rate Poisson model. Essentially, the log-likelihood plot for each cluster will indicate whether the logarithm of the probability of correct prediction from the Poisson GLM exceeds the logarithm of the probability of correct prediction from the constant-rate Poisson model.

e. Probability Assessment

In addition to indicating the coefficient of each covariate, the output of each Poisson generalized linear model in R (R, Development Core Team, 2018) reveals individually statistically significant covariates. To more closely evaluate the nature of the relationship between hurricane count and these statistically significant covariates, several repetitions of the Poisson generalized linear model will be performed with these covariates isolated. These variables will be further investigated through a probability simulation designed to assess the probability of a hurricane landfall in a specific cluster at various thresholds for these indices.

Risk and Safety

Risk and safety precautions are not applicable to this research.

Data Analysis

Due to the retrospective nature of this research, the entirety of the project can be described as data analysis. This analysis is detailed above in the Procedure section.

D. Bibliography

Dailey, P. S., Zuba, G., Ljung, G., Dima, I. M., & Guin, J. (2009). On the Relationship between North Atlantic Sea Surface Temperatures and U.S. Hurricane Landfall Risk. *Journal of Applied Meteorology and Climatology*, *48*(1), 111–129.

Dake, C., Cane, M. A., Kaplan, A., Zebiak, S. E., & Huang, D. (2004). Predictability of El Nino over the past 148 years. *Nature; London*, *428*(6984), 733–736.

Davies, T. M. (2015). *The Book of R: A First Course in Programming and Statistics 1st.* San Francisco, CA: No Starch Press.

Elsner, J. B., & Jagger, T. H. (2004). A Hierarchical Bayesian Approach to Seasonal Hurricane Modeling. *Journal of Climate*, *17*(14), 2813–2827.

Fraza, E., & Elsner, J. B. (2015). A climatological study of the effect of sea-surface temperature on North Atlantic hurricane intensification. *Physical Geography*, *36*(5), 395–407.

Gill, J. (2001). *Generalized Linear Models: A Unified Approach.* Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-134. Thousand Oaks, CA: Sage.

Klotzbach, P. J. (2010). El Niño–Southern Oscillation's Impact on Atlantic Basin Hurricanes and U.S. Landfalls. *Journal of Climate*, *24*(4), 1252–1263.

Kossin, J. P., Camargo, S. J., & Sitkowski, M. (2010). Climate Modulation of North Atlantic Hurricane Tracks. *Journal of Climate*, *23*(11), 3057–3076.

Nakamura, J., Lall, U., Kushnir, Y., & Camargo, S. J. (2009). Classifying North Atlantic Tropical Cyclone Tracks by Mass Moments. *Journal of Climate*, *22*(20), 5481–5494.

Patricola, C. M., Saravanan, R., & Chang, P. (2014). The Impact of the El Niño–Southern Oscillation and Atlantic Meridional Mode on Seasonal Atlantic Tropical Cyclone Activity. *Journal of Climate*, *27*(14), 5311–5328.

Rayner, N. A. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, *108*(D14).

Smith, S. R., Brolley, J., O'Brien, J. J., & Tartaglione, C. A. (2007). ENSO's Impact on Regional U.S. Hurricane Activity. *Journal of Climate*, *20*(7), 1404–1414.

Vecchi, G. A., Msadek, R., Anderson, W., Chang, Y.-S., Delworth, T., Dixon, K., … Zhang, S. (2013). Multiyear Predictions of North Atlantic Hurricane Frequency: Promise and Limitations. *Journal of Climate*, *26*(15), 5337–5357.

Villarini, G., Vecchi, G. A., & Smith, J. A. (2011). U.S. Landfalling and North Atlantic Hurricanes: Statistical Modeling of Their Frequencies and Ratios. *Monthly Weather Review*, *140*(1), 44–65.

Vimont, D. J., & Kossin, J. P. (2007). The Atlantic Meridional Mode and hurricane activity. *Geophysical Research Letters*, *34*(7).

Wolter, K., & Timlin, M. S. (2011). El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI.ext). *International Journal of Climatology*, *31*(7), 1074–1087.

1. **Humans Participants Research**

Human participants will not be involved in this research.

2. **Vertebrate Animals Research**

Vertebrate animals will not be involved in this research.

3. **Potentially Hazardous Biological Agents Research**

Potentially hazardous biological agents will not be involved in this research.

4. **Hazardous Chemicals, Activities, & Devices**

Hazardous chemicals, activities, and devices will not be involved in this research.

NO ADDENDUMS EXIST

OFFICIAL ABSTRACT and CERTIFICATION

Assessing Regional North Atlantic Hurricane Landfall Frequencies Through Poisson Regression Modeling and Probability Simulations

Kyra McCreery
North Shore High School, Glen Head, NY, USA

The development of models that accurately assess the complex relationships between major climate indices and North Atlantic hurricane activity has become of pressing importance as the frequency and intensity of Atlantic landfalling hurricanes continues to increase. This study elaborates on prior research by modeling time series of North Atlantic landfalling hurricanes over the period 1950-2016 using covariates in the form of indices for Main Development Region sea surface temperature (SST) anomalies, Gulf of Mexico SST anomalies, Niño 3.4 SST anomalies, Niño 1.2 SST anomalies, and the Atlantic Meridional Mode (AMM). Hurricane landfall counts grouped into five unique clustered datasets are utilized to examine distinct trends amongst storms that tend to follow disparate tracks. An initial Kendall's tau correlation assessment reveals early relationships between the clustered datasets and indices that are further examined through the use of Poisson generalized linear models individually evaluated for quality based on Akaike information criterion (AIC). The results of the Poisson regression analysis do not point to a single all-encompassing model as superior, but rather reveal the differences in the way various climate indices modulate multiple landfall regions and distinctly grouped tracks within these regions. However, two indices, Niño 3.4 and MDR, are routinely observed to indicate a high degree of association with the frequency of hurricane landfalls cross-cluster; these repeatedly statistically significant variables are further investigated through a probability simulation designed to assess the probability of a hurricane landfall at various thresholds for each index.

Category
Pick one only—mark an "X" in box at right

- ☐ Animal Sciences
- ☐ Behavioral & Social Sciences
- ☐ Biochemistry
- ☐ Biomedical & Health Sciences
- ☐ Biomedical Engineering
- ☐ Cellular & Molecular Biology
- ☐ Chemistry
- ☐ Computational Biology & Bioinformatics
- ■ Earth & Environmental Sciences
- ☐ Embedded Systems
- ☐ Energy: Chemical
- ☐ Energy: Physical
- ☐ Engineering Mechanics
- ☐ Environmental Engineering
- ☐ Materials Science
- ☐ Mathematics
- ☐ Microbiology
- ☐ Physics & Astronomy
- ☐ Plant Sciences
- ☐ Robotics & Intelligent Machines
- ☐ Systems Software
- ☐ Translational Medical Sciences

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check ALL that apply):

   ☐ human participants ☐ potentially hazardous biological agents

   ☐ vertebrate animals ☐ microorganisms ☐ rDNA ☐ tissue

2. I/we worked or used equipment in a regulated research institution or industrial setting: ■ Yes ☐ No

3. This project is a continuation of previous research. ☐ Yes ■ No

4. My display board includes non-published photographs/visual depictions of humans (other than myself): ☐ Yes ■ No

5. This abstract describes only procedures performed by me/us, reflects my/our own independent research, and represents one year's work only ■ Yes ☐ No

6. I/we hereby certify that the abstract and responses to the above statements are correct and properly reflect my/our own work. ■ Yes ☐ No

This stamp or embossed seal attests that this project is in compliance with all federal and state laws and regulations and that all appropriate reviews and approvals have been obtained including the final clearance by the Scientific Review Committee.