

## ABSTRACT

FRIEDMAN, LUKE ALEXANDER. A Comparative Study of Atlantic Seasonal Accumulated Cyclone Energy and Tropical Cyclone Counts. (Under the direction of Dr. Lian Xie).

Tropical cyclones (TCs) pose a significant threat to life and property around the globe. Two common metrics to measure seasonal TC activity in the Atlantic are accumulated cyclone energy (ACE) and overall number of TCs, or counts. These two metrics are not always consistent with each other when describing Atlantic TC activity on a seasonal basis. In this study, the ACE to Counts Ratio was defined to quantify these inconsistencies. Then, spatial and temporal analysis was performed on the ratio to explore its properties and connect the ratio to the various large scale phenomena that influence overall Atlantic TC activity. The analysis found significant differences in Atlantic sea surface temperature and vertical wind shear between different ACE to Counts Ratio environments, along with dominant patterns of variability of the ratio on interannual and multidecadal timescales. These signals can be connected to the patterns that affect Atlantic tropical activity, such as El Niño-Southern Oscillation and the Atlantic Multidecadal Oscillation.

Also in this thesis, statistical models were constructed to predict and classify Atlantic seasonal ACE. 34 climatic indices were used as possible predictors for the models, and statistical and machine learning methods were used for predictor selection and numerical ACE prediction. Finally, two different methods were used to classify the ACE predictions. The goal is for model performance to exceed the skill of climatology and perform at least comparatively to prediction models of other organizations. In the final section, this methodology will be used to predict the ACE to Counts Ratio and compare to predictions using climatology.

© Copyright 2022 by Luke Friedman

All Rights Reserved

A Comparative Study of Atlantic Seasonal Accumulated Cyclone Energy and Tropical Cyclone Counts

by  
Luke Friedman

A thesis submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of the  
requirements for the degree of  
Master of Science

Marine, Earth, and Atmospheric Science

Raleigh, North Carolina  
2022

APPROVED BY:

---

Dr. Lian Xie  
Committee Chair

---

Dr. Anantha Aiyyer

---

Dr. Sarah Larson

---

Dr. Xipeng Shen

## **BIOGRAPHY**

Luke Friedman was raised in the NYC suburbs in New Jersey and New York. He moved to the Carolinas for college, where he has since spent his adult life. Though his undergraduate degree is in engineering, he wanted to explore a different field and decided to pursue a Master's degree in Meteorology at NC State. In his spare time, he enjoys playing piano and disc golf, and he and his wife have an extensive board game collection.

## ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Lian Xie, for taking a chance on a student with no prior academic meteorology experience. I also want to thank him for supporting me in my research and giving me the opportunity to contribute to official seasonal hurricane prediction releases. I want to thank Dr. Viney Aneja and Matt Borgia for providing me valuable advice and direction during my transition from an engineering career to my meteorology studies. I also want to acknowledge my committee members, Dr. Anantha Aiyyer, Dr. Sarah Larson, and Dr. Xipeng Shen, for providing different perspectives to make this thesis as thorough as possible. Graduate students in my advisor's research group also supported me early on in my research and helped with many of the datasets I used in the analyses. Finally, I want to thank my wife, Rachel, for imparting the experience she recently had as a graduate student and supporting me emotionally through this journey.

## TABLE OF CONTENTS

LIST OF TABLES .....	v
LIST OF FIGURES .....	vi
<b>Chapter 1: Introduction</b> .....	1
<b>Chapter 2: Data</b> .....	6
2.1 Hurricane Data.....	6
2.2 Environmental Data .....	6
2.3 Climatic Indices.....	7
<b>Chapter 3: Methods</b> .....	11
3.1 Seasonal Hurricane Activity Classification Predictions .....	11
3.2 ACE to Counts Ratio .....	14
3.3 EOF Analysis and Track Density Analysis of Vertical Wind Shear.....	18
3.4 Ratio Predictability .....	19
<b>Chapter 4: Results</b> .....	24
4.1 Seasonal Hurricane Activity Classification Predictions .....	24
4.2 ACE to Counts Ratio .....	25
4.2.1 Frequency Analysis of ACE to Counts Ratio .....	25
4.2.2 Connections to Environmental Variables of TC Activity .....	28
4.2.2a Sea Surface Temperature.....	28
4.2.2b Vertical Wind Shear.....	31
4.2.2c Mid-Tropospheric Moisture .....	32
4.3 Spatial Variability of Atlantic Vertical Wind Shear and Analysis.....	33
4.3.1 EOF Spatial Analysis .....	33
4.3.2 Correlations of Vertical Wind Shear EOFs with ENSO and Other Indices.....	34
4.3.3 Track Density Analysis .....	36
4.4 Ratio Predictability .....	37
<b>Chapter 5: Conclusions</b> .....	73
5.1 Prediction Models.....	73
5.2 ACE to Counts Ratio .....	73
<b>Chapter 6: Future Work</b> .....	76

## LIST OF TABLES

Table 2.1	The first 16 of the 31 predictors (listed alphabetically) that were used in the ACE classification prediction models .....	9
Table 2.2	The other 15 of the 31 predictors (listed alphabetically) that were used in the ACE classification prediction models .....	10
Table 3.1	Number of years between 1948 and 2020 that fall into each class of each ACE classification system .....	21
Table 4.1	Results of the ACE classification prediction models. Comparison models from Colorado State University, Tropical Storm Risk, and simple climatologies are also shown. Each percentage is a different model run and shows the percent of test years classified correctly. Darker (lighter) highlights indicate higher (lower) accuracy. Models left blank are duplicate models .....	39
Table 4.2	Correlation coefficients of the AMO to the ratio from 1857-2020. Each month of the AMO was correlated to the ratio individually; a yearly average of the AMO index was also correlated. All correlation coefficients are statistically significant at the 95% confidence level. ....	40
Table 4.3	Correlation coefficients between the monthly hurricane season (MHS) VWS ECs and climate indices Arctic Oscillation (AO), Atlantic Multidecadal Oscillation (AMO), North Atlantic Oscillation (NAO), and ENSO indices (NINO12, NINO3, NINO34, and NINO4). The top panel shows the EOFs from the global domain (30°S - 30°N, 90-5°W), and the bottom panel shows the EOFs from the hemispheric domain (0-45°N, 90°W-0). Statistically significant correlation coefficients at the 95% confidence level are highlighted .....	41
Table 4.4	Correlation coefficients between the hurricane season average (HSA) VWS ECs and the unfiltered ACE to Counts Ratio, the ratio with a high pass filter applied, and yearly ACE. The top panel shows the EOFs from the global domain (30°S - 30°N, 90-5°W), and the bottom panel shows the EOFs from the hemispheric domain (0-45°N, 90°W-0). Statistically significant correlation coefficients at the 95% confidence level are highlighted .....	42
Table 4.5	Results of the ACE to Counts Ratio prediction models. Each row is a different model run. The average MSE from the ten sliding windows of each model run and the MSSS when compared to three different climatologies are shown. Positive (green highlighted) MSSS values show positive skill compared to climatology, and negative (red highlighted) values show negative skill. MSE values for the climatologies are also shown.....	43

## LIST OF FIGURES

Figure 1.1	Atlantic TC Counts (solid line, left axis) and Atlantic ACE (dashed line, right axis) from 1951 to 2020. Both datasets were calculated using the HURDAT2 Best Track Data.....	5
Figure 3.1	ACE to Counts Ratio (solid line). The ratio is centered around 1, marked by the dashed line. The 18 high ratio years and 18 low ratio years are marked by red and green circles, respectively .....	21
Figure 3.2	Comparison of the ACE to Counts Ratio (solid line, left axis) with ACE per storm (dashed line, right axis). ACE per storm is normalized by its standard deviation. Correlation coefficient between the two indices is also shown.....	22
Figure 3.3	Track Density Analysis zones, with latitude ( $^{\circ}$ N) and longitude ( $^{\circ}$ W) extents included.....	23
Figure 4.1	The detrended ACE to Counts Ratio from 1951-2020 (left), and Fourier analysis of the ratio (right). The five most dominant frequency bands occur at 2.67, 3.12, 8, 18 and 42 years .....	44
Figure 4.2	Wavelet analysis of the ACE to Counts Ratio from 1951-2020. (a) shows the ratio, normalized by standard deviation. (b) shows the wavelet power spectrum, with power scale to the right. The black contours and black cone show the statistically significant power contours and the cone of influence, respectively. (c) shows the global wavelet spectrum; global power (solid line) and 95% significance contour (dashed line). (d) shows the variance of the 2-8 year scale-average time series (solid line), along with the 95% significance value (dashed line) .....	45
Figure 4.3	The detrended ACE to Counts Ratio from 1851-2020 (left), and Fourier analysis of the ratio (right). The eight most dominant frequency bands occur at 2.22, 2.59, 3.2, 5.57, 8, 17.07, 36.57, and 128 years .....	46
Figure 4.4	Wavelet analysis of the ACE to Counts Ratio from 1851-2020. (a) shows the ratio, normalized by standard deviation. (b) shows the wavelet power spectrum, with power scale to the right. The black contours and black cone show the statistically significant power contours and the cone of influence, respectively. (c) shows the global wavelet spectrum; global power (solid line) and 95% significance contour (dashed line). (d) shows the variance of the 2-8 year scale-average time series (solid line), along with the 95% significance value (dashed line) .....	47

Figure 4.5	Atlantic Ocean SST Composite Analysis of ACE to Counts Ratio for JJA (left) and ASO (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.....	48
Figure 4.6	Atlantic Ocean SST Composite Analysis of ACE to Counts Ratio for September (left) and October (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant .....	49
Figure 4.7	EMD Analysis of the ACE to Counts Ratio. Top panel shows the original ratio timeseries (signal), followed by the top 5 IMFs in descending order from highest frequency to lowest frequency. The bottom panel shows the residual trend of the ratio. The x-axis denotes the years of the timeseries, from 1951-2020.....	50
Figure 4.8	The ACE to Counts Ratio with a high pass filter applied. The lowest frequencies corresponding to IMFs 3, 4, and 5 from Figure 4.7 are removed. The 18 high ratio years and 18 low ratio years of the high pass ratio are marked by red and green circles, respectively .....	51
Figure 4.9	Atlantic Ocean SST Composite Analysis of the high pass ACE to Counts Ratio for JJA (left) and ASO (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.....	52
Figure 4.10	Atlantic Ocean SST Composite Analysis of the high pass ACE to Counts Ratio for September (left) and October (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant .....	53
Figure 4.11	Pacific Ocean SST Composite Analysis of ACE to Counts Ratio for JJA (left) and ASO (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.....	54
Figure 4.12	Pacific Ocean SST Composite Analysis of ACE to Counts Ratio for September (left) and October (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.....	54
Figure 4.13	Pacific Ocean SST Composite Analysis of the high pass ACE to Counts Ratio for DJF (left) and MAM (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.....	55

Figure 4.14 Pacific Ocean SST Composite Analysis of the high pass ACE to Counts Ratio for JJA (left) and ASO (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.....	56
Figure 4.15 Pacific Ocean SST Composite Analysis of the high pass ACE to Counts Ratio for September (left) and October (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.....	57
Figure 4.16 Atlantic Ocean VWS Composite Analysis of ACE to Counts Ratio for JJA (left) and ASO (right). VWS differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.....	58
Figure 4.17 Atlantic Ocean VWS Composite Analysis of ACE to Counts Ratio for September (left) and October (right). VWS differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.....	59
Figure 4.18 Atlantic Ocean VWS Composite Analysis of the high pass ACE to Counts Ratio for JJA (left) and ASO (right). VWS differences between high and low ratio years above, statistical significance (90% confidence level) below, with maroon areas being statistically significant .....	60
Figure 4.19 Atlantic Ocean VWS Composite Analysis of the high pass ACE to Counts Ratio for September (left) and October (right). VWS differences between high and low ratio years above, statistical significance (90% confidence level) below, with maroon areas being statistically significant .....	61
Figure 4.20 Atlantic Ocean RH700 Composite Analysis of ACE to Counts Ratio for JJA (left) and ASO (right). RH700 differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.....	62
Figure 4.21 Atlantic Ocean RH700 Composite Analysis of ACE to Counts Ratio for September (left) and October (right). RH700 differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant .....	63
Figure 4.22 Atlantic Ocean RH850-500 Composite Analysis of ACE to Counts Ratio for JJA (left) and ASO (right). RH850-500 differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.....	64

Figure 4.23 Atlantic Ocean RH850-500 Composite Analysis of ACE to Counts Ratio for September (left) and October (right). RH850-500 differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant .....	65
Figure 4.24 Percent variance explained by the top 10 monthly hurricane season VWS EOFs in the global domain. The truncation point is marked on the graph.....	66
Figure 4.25 Percent variance explained by the top 10 monthly hurricane season VWS EOFs in the hemispheric domain. The truncation point is marked on the graph.....	66
Figure 4.26 Percent variance explained by the top 10 yearly hurricane season average VWS EOFs in the global domain. The truncation point is marked on the graph.....	67
Figure 4.27 Percent variance explained by the top 10 yearly hurricane season average VWS EOFs in the hemispheric domain. The truncation point is marked on the graph .....	67
Figure 4.28 Spatial patterns of the top 4 monthly hurricane season VWS EOFs in the global domain. EOF1 (top left), EOF2 (top right), EOF3 (bottom left), and EOF4 (bottom right).....	68
Figure 4.29 Spatial patterns of the top 4 monthly hurricane season VWS EOFs in the hemispheric domain. EOF1 (top left), EOF2 (top right), EOF3 (bottom left), and EOF4 (bottom right) .....	69
Figure 4.30 Spatial patterns of the top 4 yearly hurricane season average VWS EOFs in the global domain. EOF1 (top left), EOF2 (top right), EOF3 (bottom left), and EOF4 (bottom right) .....	70
Figure 4.31 Spatial patterns of the top 4 yearly hurricane season average VWS EOFs in the hemispheric domain. EOF1 (top left), EOF2 (top right), EOF3 (bottom left), and EOF4 (bottom right) .....	71
Figure 4.32 Percent difference of storm starts by zone between high/low VWS EOF1 years and all years from 1951-2018. High (low) VWS EOF1 years represented by dashed (solid) lines. Analysis was done for all TCs (top left), tropical depressions (top right), tropical storms (bottom left), and hurricanes (bottom right). Latitude/longitude bounds of each zone can be found in Figure 2.3.....	72
Figure 5.1 Flowchart of conclusions connecting remote effects from the Pacific and local effects from the Atlantic to the variability of the ACE to Counts Ratio.....	75

# CHAPTER 1

## Introduction

Tropical cyclones (TCs) pose a significant threat to life and property around the globe. In the Atlantic Ocean basin, hurricanes have the potential to affect a large area of population centers encompassing nations in the Caribbean, eastern Mexico and the eastern and Gulf coasts of the United States. From 1949 to 2006, Atlantic hurricanes accounted for the highest storm-related financial losses in the United States, averaging \$2.6 billion in annual insured losses (Changnon, 2008). In a study of U.S. billion dollar weather disasters from 1980-2020, Smith (2021) found that TCs caused the most damage and had the highest average cost per weather event. According to the Insurance Information Institute, there were 2,003 deaths in the United States from hurricanes between 2005 and 2020 (Facts + Statistics: Hurricanes, 2021). It is clear that due to this threat, preparedness for hurricanes is key to mitigating property loss and saving lives.

To combat the threat of hurricanes, extensive resources are devoted toward the seasonal prediction of TCs globally and hurricanes in the Atlantic basin. In 2019, Klotzbach et al. performed a review of twelve different organizations that perform seasonal TC predictions using statistical, dynamical, and/or mixed statistical-dynamical models. Some of the organizations that predict hurricane activity in the Atlantic basin include the National Oceanic and Atmospheric Administration (NOAA), the UK Met office, the European Centre for Medium-Range Weather Forecasts, Colorado State University (CSU), and Tropical Storm Risk (TSR). Most of these organizations have been performing predictions for as long as decades. They generally issue/update predictions several times per year, as early as the December prior to each hurricane season (Klotzbach et al. 2019). Though this paper will be highlighting statistical efforts to predict seasonal TC activity, it is important to acknowledge the vast area of research dedicated to

dynamical TC predictions. Some of the ways this can be done are tracking TC-like vortices in a general circulation model (Vitart et al. 1997) and replicating TC activity by forecasting basin SSTs (Vitart and Stockdale 2001). This is not a comprehensive accounting of dynamical TC prediction methods, and more can be found in Klotzbach et al. 2019.

Two metrics of seasonal hurricane activity commonly predicted are number of TCs (hereafter referred to as “counts”) and accumulated cyclone energy (ACE). First developed by William Gray as the Hurricane Destruction Potential (Gray 1988), ACE was coined by NOAA and is calculated by summing the squared maximum sustained wind speed (knots<sup>2</sup>) of each system every six hours while the storm is at least tropical storm strength (>35 knots) (Waple et al. 2001).

Predictions of numerical values of ACE and counts have been done extensively, by organizations such as the ones listed above and in various publications (Gray, 1984; Vecchi et al., 2011; Chen and Lin, 2013; Sun et al., 2021). For the public, it is more straightforward to understand that a certain number of named storms, hurricanes, or major hurricanes will occur in a given season than describing hurricane activity using ACE. To make it easier for public understanding, ACE can be predicted by classifying it, such as in Chand and Walsh (2012). Chand and Walsh set up the ACE prediction problem as a binary classification problem. However, using a larger number of classes for more clarity to the public is also possible. In addition, accurately classifying the activity of hurricane seasons can have other benefits, such as improving accuracy in landfall count predictions (Yan et al. 2014). In their review of the 2004-2005 Atlantic hurricane seasons, Lea and Saunders mention that TSR issues probabilistic forecasts in tercile levels of TC activity with categories of “high,” “medium,” and “low activity” (Lea and Saunders 2006). NOAA has also created classifications for North Atlantic seasonal

hurricane activity based on tercile partitioning of ACE values. They classify hurricane activity into four classes: extremely active, above-normal, below-normal, and near-normal (“Background Information,” 2021). Though classification predictions are performed by some of the well-known prediction organizations, less is written in the literature on this topic. The published NOAA activity classes are a starting point for creating classification prediction models of ACE, which will be shown later in this paper.

Since ACE is a sum of squared velocities, it is a proxy for the kinetic energy of hurricanes. Measuring the strength of a hurricane season by this metric is supposed to be consistent with measuring a season using counts. However, this is not always the case, as shown in Figure 1.1 which plots the counts (solid line) and ACE values (dashed line) of each year from 1951 to 2020. For about a decade from 1968 to 1979, most years had very low ACE values compared to the rest of the dataset, but rather high counts values. The opposite is true in years such as 1961, 1995 and 2017, where the years’ ACE values are high relative to the corresponding counts values. These inconsistencies can be explored by creating a new metric, the ACE to Counts ratio (referred throughout as the “ratio”), to explain the disparity between a given year’s ACE and counts. Then, this metric can be explored to see how it changes over time, and the skill to which it can be predicted.

The objectives of this thesis fall into four main categories. The first is to design a statistical prediction model that classifies hurricane activity into five classes. The goal is for model performance to exceed the skill of climatology and perform at least comparatively to prediction models of other organizations. The second is to define and analyze the ACE to Counts ratio, which will measure how seasonal hurricane activity differs as measured by ACE and by number of TCs. The dominant frequency patterns of the ratio will be determined and the

ratio will be connected to the environmental variables conducive to TC formation. The third is to study the overall variability in Atlantic vertical wind shear (VWS) by using EOF analysis to determine the dominant spatial modes of variability. In addition, these modes will be correlated to the ACE to Counts ratio and other large scale climate patterns to determine possible links between these phenomena. A link will also attempt to be drawn between the VWS spatial modes and the location of TC genesis in the Atlantic. The fourth is to predict the ratio using similar prediction methods as the hurricane activity classification models, and exceed the skill of climatology with the prediction results.

In accordance with accomplishing these objectives, the rest of the thesis will be laid out as follows. Section 2 describes the datasets used in this study. Section 3 describes the methods used to create the prediction models. It also defines the ratio and explains methods used to analyze it. Section 4 presents the results in four parts, where section 4.1 focuses on the classification prediction models, section 4.2 the ACE to Counts Ratio, section 4.3 the VWS analysis, and section 4.4 the ratio prediction models. The thesis is concluded in Section 5, with possible directions for future work outlined in Section 6.

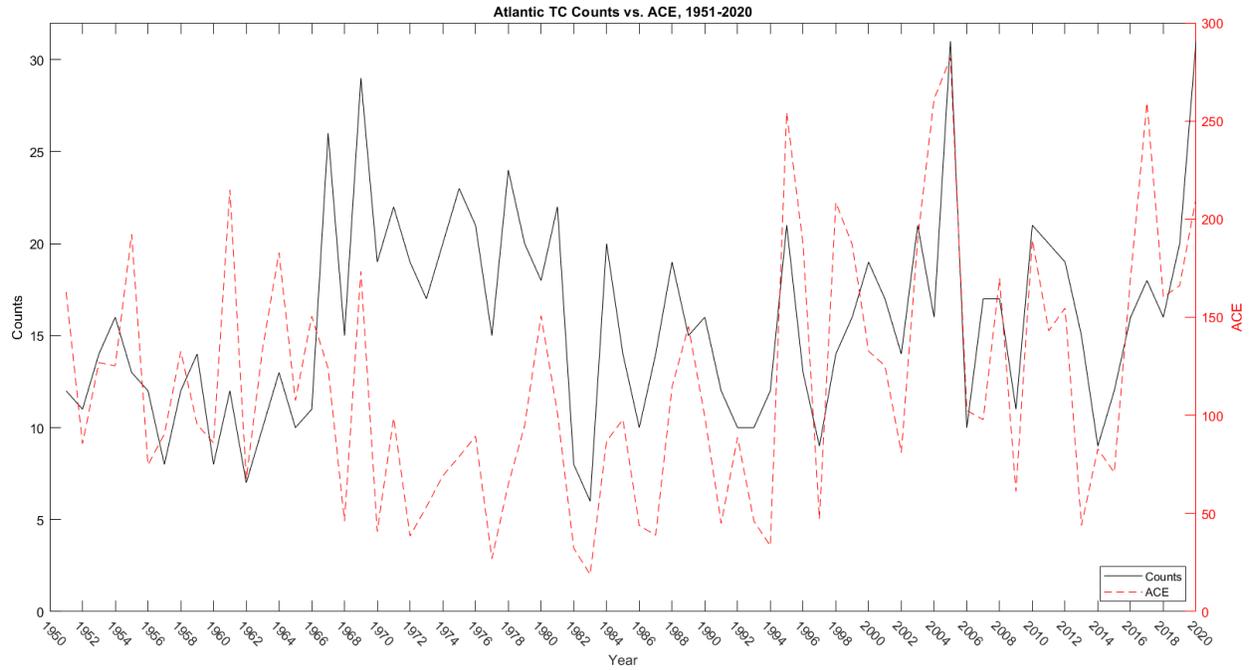


Figure 1.1. Atlantic TC Counts (solid line, left axis) and Atlantic ACE (dashed line, right axis) from 1951 to 2020. Both datasets were calculated using the HURDAT2 Best Track Data.

## CHAPTER 2

### Data

#### 2.1 Hurricane Data

Two metrics of TC activity that are being studied are ACE and the ACE to Counts Ratio (abbreviated as “ratio”). To obtain historical TC data for these two metrics, the HURDAT2 best track dataset from the National Hurricane Center’s hurricane database was used (Landsea and Franklin 2013). HURDAT2 includes six hourly information on location, maximum wind speed and minimum central pressure for TCs in the Atlantic basin since 1851. Since hurricane cyclone frequency and intensity is known to be underestimated prior to the aircraft reconnaissance and satellite eras (Vecchi and Knutson 2008), for the majority of the analysis the data has been truncated to the period beginning in either 1948 or 1951. ACE was calculated as follows:

$$ACE = 10^{-4} \sum v_{max}^2 \quad 1$$

where  $v_{max}$  is the 6-hourly maximum 1-minute average wind speeds for each storm in the database.

ACE has been calculated per month and per year, though primarily yearly ACE values will be used in the analysis later in the paper. To calculate the ratio, a yearly count of TCs starting in 1948 was also obtained from the HURDAT2 data. Finally, the TC starting points (lat-lon) for the track density analysis were also obtained from HURDAT2.

#### 2.2 Environmental Data

The physical mechanisms being used to study the ratio are SST, mid-tropospheric relative humidity at the 700mb and 850-500mb average levels (RH700 and RH850-500, respectively), and 850-200mb VWS. All of these datasets were originally obtained from the NCEP/NCAR Reanalysis project (Kalnay et. al, 1996). They are at a spatial resolution of  $1^{\circ} \times 1^{\circ}$

latitude/longitude and a monthly temporal resolution from 1891-present. The SST and RH850-500 datasets were used in their original form from Kalnay et. al. For RH850-500, data from the 850mb, 700mb, 600mb and 500mb levels were averaged together at each grid point to obtain a three-dimensional latitude X longitude X time dataset.

The other datasets include the RH700 and VWS data. VWS was calculated using 850mb and 200mb zonal and meridional wind data as follows:

$$VWS = \sqrt{(U200 - U850)^2 + (V200 - V850)^2} \quad 2$$

where

U200 = 200mb zonal wind

U850 = 850mb zonal wind

V200 = 200mb meridional wind

and V850 = 850mb meridional wind.

These four wind component datasets, along with RH700, were originally obtained from Kalnay et. al (1996) but were later smoothed to a 5°X5° latitude/longitude resolution. The latitude range is limited to 75°S to 45°N, and the longitude is spanned globally. Like the original NCEP/NCAR data, the temporal resolution of this data is also monthly, but spans the time period from 1951-2018.

### 2.3 Climatic Indices

For the forecasts performed in this paper, a set of 31 predictors were used. These predictors and their names can be found in Tables 2.1 and 2.2. Most of these predictors are either indices of atmospheric circulation patterns (ex. North Atlantic Oscillation – NAO), ocean patterns (ex. Atlantic Multidecadal Oscillation – AMO), or area-averaged SSTs (ex. NINO34) that are known to affect Atlantic hurricane activity. They are consistent with model predictors

used in past publications such as Keith and Xie, 2009 and Xie and Liu, 2014. A majority of the predictors were obtained from the NOAA Physical Sciences Laboratory (PSL), while some of the datasets originated in other locations, such as NOAA's Climate Prediction Center ("Climate Indices," 2021). Three additional predictors, mdrushear, mdrvshare, and mdrshare (see Table 2.1) were calculated using the MDR wind predictors obtained from NOAA PSL. Mdrushear was obtained by subtracting the mdru850 values from the mdru200 values for each time step. Mdrvshare was similarly calculated by subtracting mdrv850 from mdrv200. Finally, mdrshare was found by taking the resultant vector of the mdrushear and mdrvshare values at each time step. With these three predictors added, a total of 34 predictors were available for the forecasts. All of these datasets are in the form of monthly time series from 1951-2020. The 34 predictors were then standardized by subtracting each value from that predictor's mean and dividing by its standard deviation.

The response variable for these forecasts is yearly ACE values. ACE was obtained and calculated as described in Section 2.1.

Table 2.1. The first 16 of the 31 predictors (listed alphabetically) that were used in the ACE classification prediction models.

Predictor Name	Predictor Long Name	Description	Source
ammsst	Atlantic Meridional Mode	Meridional variability in the tropical Atlantic Ocean	NOAA Physical Sciences Laboratory -From U Wisconsin using the NCEP SST
amon	Atlantic Multidecadal Oscillation	An index of N Atlantic temperatures, calculated from the Kaplan SST Dataset	NOAA Physical Sciences Laboratory
ao	Arctic Oscillation	Constructed by projecting 1000mb height anomalies poleward of 20°N onto the loading pattern of the AO.	NOAA Climate Prediction Center (CPC) -Height Anomalies from NCEP/NCAR reanalysis
censo	Bivariate ENSO Time Series	Combination of standardized SOI and standardized Niño3.4 SST timeseries	NOAA Physical Sciences Laboratory -SOI from Climatic Research Unit of East Anglia, England -SST from HadISST 1.1
epo	East Pacific/North Pacific Oscillation	Spring-Summer-Fall height anomaly pattern centered in Alaska, N Pacific and E. N America	NOAA Climate Prediction Center (CPC) -Height Anomalies from NCEP/NCAR reanalysis
ggst	Global Land-Ocean Temperature Index	Combined Land-Surface Air and Sea-Surface Water Temperature Anomalies, global mean	NASA Goddard Institute for Space Studies -Land Data: GHCN v4 / Ocean Data: ERSST v5
mdroir	MDR Top of Atmosphere Outgoing Longwave Radiation	Top of Atmosphere Outgoing Longwave Radiation over the Main Development Region; 10.5-20N, 80.6-20.6W	NOAA Physical Sciences Laboratory -NCEP/NCAR Reanalysis monthly means
mdrslp	MDR Surface Sea Level Pressure	Surface Sea Level Pressure over the Main Development Region; 10-20N, 80-20W	NOAA Physical Sciences Laboratory -NCEP/NCAR Reanalysis monthly means
mdrsst	MDR Sea Surface Temperature	Surface Gaussian SST over the Main Development Region; 10.5-20N, 80.6-20.6W	NOAA Physical Sciences Laboratory -NCEP/NCAR Reanalysis monthly means
mdru200	200mb Pressure Level u wind	200mb Pressure Level u wind	NOAA Physical Sciences Laboratory -NCEP/NCAR Reanalysis monthly means
mdru850	850mb Pressure Level u wind	850mb Pressure Level u wind	NOAA Physical Sciences Laboratory -NCEP/NCAR Reanalysis monthly means
mdrv200	200mb Pressure Level v wind	200mb Pressure Level v wind	NOAA Physical Sciences Laboratory -NCEP/NCAR Reanalysis monthly means
mdrv850	850mb Pressure Level v wind	850mb Pressure Level v wind	NOAA Physical Sciences Laboratory -NCEP/NCAR Reanalysis monthly means
mdrushear	200-850mb zonal wind shear	mdru200-mdru850	Self-calculated
mdrvshear	200-850mb meridional wind shear	mdrv200-mdrv850	Self-calculated
mdrshear	200-850mb zonal-meridional wind shear	$\sqrt{(\text{mdrushear}^2 + \text{mdrvshear}^2)}$	Self-calculated

Table 2.2. The other 15 of the 31 predictors (listed alphabetically) that were used in the ACE classification prediction models.

Predictor Name	Predictor Long Name	Description	Source
meiv2	Multivariate ENSO Index Version 2	Time series of the leading combined EOF of SLP, SST, zonal wind, meridional wind, and OLR over tropical Pacific; 30S-30N, 100E-70W	NOAA Physical Sciences Laboratory -SLP, SST and zonal/meridional winds: JRA-55 global reanalysis -OLR: NOAA Climate Data Record
nao	North Atlantic Oscillation	N-S dipole of height/pressure anomalies centered in Greenland and C. N Atlantic	NOAA Climate Prediction Center (CPC) -Height Anomalies from NCEP/NCAR reanalysis
ngst	Northern Hemisphere Land-Ocean Temperature Index	Combined Land-Surface Air and Sea-Surface Water Temperature Anomalies, N Hemisphere mean	NASA Goddard Institute for Space Studies -Land Data: GHCN v4 / Ocean Data: ERSST v5
nino12	Niño1+2 Anomaly Index	Average SST anomalies across E Pacific; 0-10S, 90W-80W	NOAA Physical Sciences Laboratory (NOAA ERSST V5)
nino3	Niño3 Anomaly Index	Average SST anomalies across E. C Pacific; 5N-5S, 150W-90W	NOAA Physical Sciences Laboratory (NOAA ERSST V5)
nino34	Niño3.4 Anomaly Index	Average SST anomalies across C Pacific; 5N-5S, 170W-120W	NOAA Physical Sciences Laboratory (NOAA ERSST V5)
nino4	Niño4 Anomaly Index	Average SST anomalies across C Pacific; 5N-5S, 160E-150W	NOAA Physical Sciences Laboratory (NOAA ERSST V5)
pdo	Pacific Decadal Oscillation	Leading principle component of monthly SST anomalies in the North Pacific Ocean	NOAA Physical Sciences Laboratory
pna	Pacific North American Index	Low-frequency variability in the Northern Hemisphere extratropics, centered around Hawaii and the intermountain region of N America	NOAA Climate Prediction Center (CPC) -Height Anomalies from NCEP/NCAR reanalysis
qbo	Quasi-Biennial Oscillation	Calculated from the zonal average of the 30mb zonal wind at the equator	NOAA Physical Sciences Laboratory -Wind Data: NCEP/NCAR Reanalysis
sgst	Southern Hemisphere Land-Ocean Temperature Index	Combined Land-Surface Air and Sea-Surface Water Temperature Anomalies, S Hemisphere mean	NASA Goddard Institute for Space Studies -Land Data: GHCN v4 / Ocean Data: ERSST v5
soi	Southern Oscillation Index	Observed sea level pressure difference between Tahiti and Darwin, Australia	NOAA Climate Prediction Center (CPC)
solar	10.7cm Solar Flux Data	Solar radio flux density values	NOAA Physical Sciences Laboratory -Solar Radio Monitoring Program, National Research Council Canada
tna	Tropical Northern Atlantic Index	Average SST anomalies across southern N Atlantic; 5.5-23.5N, 15-57.5W	NOAA Physical Sciences Laboratory -HadISST and NOAA OI 1x1 SST datasets
tni	Trans-Niño Index	Difference in normalized SST anomalies between Niño1+2 and Niño4 regions (C - E Pacific SST anomaly gradient)	NOAA Physical Sciences Laboratory -HadISST and NCEP OI SST Datasets
tsa	Tropical Southern Atlantic Index	Average SST anomalies across north-central S Atlantic; 0-20S, 10E-30W	NOAA Physical Sciences Laboratory -HadISST and NOAA OI 1x1 SST datasets
whwp	Western Hemisphere warm pool	Monthly anomaly of the ocean surface area warmer than 28.5°C in the Atlantic and eastern North Pacific	NOAA Physical Sciences Laboratory -HadISST and NOAA OI 1x1 SST datasets
wp	Western Pacific Index	Low-frequency variability over N Pacific in all months, affects location and intensity of Pacific Jet entrance region	NOAA Climate Prediction Center (CPC) -Height Anomalies from NCEP/NCAR reanalysis

## CHAPTER 3

### Methods

#### 3.1 Seasonal Hurricane Activity Classification Predictions

A 5-class classification system to classify and predict Atlantic hurricane seasons by ACE was developed. ACE was calculated for each year from 1951-2020 using Equation 1 in Section 2 and is displayed in Figure 1.1 (dashed line). The five classes used to classify ACE are extremely below average (class 1), below average (class 2), near average (class 3), above average (class 4), and extremely above average (class 5). For these classifications, numerical ACE values were predicted first and then classified into the classes above.

A total of 34 predictors will be used for the prediction models. The predictors were first correlated by month with the ACE time series. For each predictor, the month with the largest magnitude correlation coefficient was used for further analysis. In this way, each predictor has a single month of values from 1951-2020 to predict the 1951-2020 ACE values. To further constrain the model and allow for a more accurate comparison with other prediction organizations, two different model runs were performed: a “preseason” and “midseason” model. The preseason (midseason) model only includes predictor months of January-April (January-August). Once the 34 predictor months were obtained, backward stepwise regression was performed, with the criterion to remove predictors varied from  $p > 0.01$  to  $p > 0.25$ . This resulted in multiple models with different combinations of predictors to use for the next two prediction methods.

Two separate prediction methods were used to predict ACE for the test set years: least absolute shrinkage and selection operator (LASSO) regression and k-nearest neighbors (KNN) classification. LASSO regression attempts to minimize:

$$\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \sum_{j=1}^p |\beta_j| = RSS + \lambda \sum_{j=1}^p |\beta_j| \quad 3$$

Where  $y_i$  is the ACE value of year  $i$

$X_{ij}$  are the predictors  $j$  for each year  $i$

$\lambda$  is the shrinkage parameter of the coefficients  $\beta_j$

and RSS is the Residual Sum of Squares:

$$RSS = \sum_{i=1}^n (Y_i - f(X_i))^2 \quad 4$$

LASSO attempts to shrink the coefficients of the predictors toward zero using  $\lambda$  in order to reduce variance. As  $\lambda$  increases, more coefficients are shrunk to zero, and thus the corresponding predictors are not used. In these ACE prediction models, the largest value of  $\lambda$  is used that does not return a null model (i.e., the largest possible number of coefficients are shrunk to zero without eliminating all predictors) (Xie 2021d). The distribution of the response variable ACE is assumed to resemble a Poisson distribution. The LASSO model returns the predicted ACE value of each of the test set years, which will then be classified to compare to the known classifications.

KNN classification uses test data to classify data points based on similarity to neighbors in the training set (Subramanian 2019). It uses a certain number of  $k$ -neighbors in the training data to perform this classification, and the similarity is calculated by Euclidean distance, which is represented by

$$\sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad 5$$

Where  $x_i$  is a specific predictor value in each training year

And  $y_i$  is the same predictor value in the test year that is to be classified (Band 2021).

In each KNN model, the value of  $k$  is varied from 1 to 10, and the value of  $k$  from the model that gives the lowest RSS is used. Thus, the exact value of  $k$  may vary from model to model. Once the  $k$ -nearest training years are chosen using the minimized Euclidean distance of their predictors to those of the test year, the ACE of those training years is averaged using a simple average, to represent the predicted ACE of the test year.

Due to the relatively small sample size of 70 years of predictor and ACE values, a sliding-window cross validation method was used, as defined by Alfaro Córdoba (2017). The 70 years of predictor and ACE values were split into windows of training and test sets of 40 and 20 years, respectively. Though the sample size for each window is smaller than just using the entire dataset, having multiple windows to make predictions increases the total number of predictions that can be made. The training/test set window advances forward by 1 year each time, for a total of 10 sliding windows (the first window was 1951-1990 predicting 1991-2010, and the last window was 1960-1999 predicting 2000-2019).

Once the ACE values were predicted for the test years, they were classified into 4 or 5 classes using two different classification methods. The first method uses categories borrowed from the Climate Prediction Center (CPC) and is based on values above or below a certain percentile of the data (CPC 2021). It is a 4-class system, where the extremely below average and below average classes are combined together. The below average class (class 2) contains years with ACE values below the 33<sup>rd</sup> percentile of the 1951-2020 ACE data. The near average class (class 3) represents ACE values between the 33<sup>rd</sup> and 67<sup>th</sup> percentile of the data. The above average class (class 4) consists of ACE values above the 67<sup>th</sup> percentile, but less than 165% of the median value of the 1951-2020 ACE data. The extremely above average class (class 5) are ACE values larger than 165% of the median 1951-2020 ACE value. The CPC system does not

have classification parameters for extremely below average ACE, so four classes will be used for this system. The second classification system is based on how far from the 1951-2020 mean each ACE value lies. The extremely below (above) average classes (class 1 and 5) are ACE values larger in magnitude than 1.2 standard deviations below (above) the mean, respectively. The below (above) average classes (class 2 and 4) consist of values between negative (positive) 0.4 and 1.2 standard deviations from the mean. Finally, the near average class (class 3) has ACE values that lie between 0.4 standard deviations below the mean and 0.4 standard deviations above the mean. These two classification systems will later be referred to as the “CPC Class” and the “Z-score Class,” respectively, and the number of years between 1948-2020 that fall into each class of these systems can be found in Table 3.1.

### 3.2 ACE to Counts Ratio

Two common metrics for measuring hurricane activity are ACE and number of TCs (counts). In the process of performing hurricane predictions using these two metrics, an interest in comparing the two methods arose. The interest was to determine whether there were hurricane seasons that had at the same time a relatively high value of ACE but a low value for counts. This would indicate a hurricane season with a smaller number of relatively stronger storms. The opposite could also be possible: seasons with a low ACE value, but a high number of TCs (a season that has a larger number of relatively weaker storms). From this, the ACE to Counts Ratio (abbreviated as “ratio”) was developed. It is described as follows:

$$ACE\ to\ Counts\ Ratio = \frac{\frac{\% \ Above\ Mean_{ACE}}{Range(\% \ Above\ Mean_{ACE})} + 1}{\frac{\% \ Above\ Mean_{Counts}}{Range(\% \ Above\ Mean_{Counts})} + 1} \quad 6$$

where

$$\% \ Above\ Mean_{ACE} = \frac{ACE\ Value}{\bar{x}_{ACE\ (1948-2020)}} - 1 \quad 7$$

and

$$\% \text{ Above Mean}_{Counts} = \frac{\text{Counts Value}}{\bar{x}_{Counts (1948-2020)}} - 1 \quad 8$$

First, the percent above the mean (PAM) for each parameter was calculated by dividing each year's ACE or counts value by the 1948-2020 mean of that parameter, and then subtracting 1 (Equations 7 and 8). This centers each PAM around zero, instead of one. Each of these quantities is then normalized by the range of the PAM values. The 1 is added back to the result, to avoid division by zero errors for the final step, which is dividing the normalized PAM of ACE by the normalized PAM of counts to get the ACE to Counts Ratio (Equation 6).

Figure 3.1 shows the ACE to Counts Ratio plotted for the years 1951-2020. A ratio of 1 means that the year's ACE relative to the 1951-2020 ACE values was the same as the year's counts relative to the 1951-2020 counts values. A ratio greater than 1 signifies a year with a relatively high value of ACE and a relatively low value for counts. A ratio less than 1 signifies the opposite (low ACE but high counts). Thus, it can be generalized that "high ratio" years (greater than 1) may tend to have a fewer number of relatively stronger storms, and "low ratio" years (less than 1) may tend to have a larger number of relatively weaker storms.

Figure 3.1 also contains 18 years each denoted by red and green circles. These are the 18 years with the highest ratio and lowest ratio, respectively (later referred to as "high ratio" and "low ratio" years). Composite analysis was performed on these years to determine if the high end and low end of the ratio exhibit any particular patterns among parameters conducive to hurricane development. The composite analysis was performed by averaging each parameter in the 18 high ratio years, and subtracting the parameter averaged among the 18 low ratio years. Those studied were sea surface temperature (SST), vertical wind shear (VWS, calculated using equation 2 in the data section), and mid-tropospheric moisture. These conditions were averaged

at every gridpoint over a domain for the Atlantic Ocean (30°S-30°N, 90°W-0°, or the Atlantic domain). SST was also averaged at every gridpoint over a domain for the Pacific Ocean (30°S-30°N, 150°E-75°W, or the Pacific domain). To determine whether the composite differences were large enough to be significant, two sample t-tests were performed on the two averages (high ratio and low ratio), with a 95% confidence level for the differences.

An idea similar to the ACE to Counts Ratio is ACE per storm, which is discussed in papers such as Camargo et al. 2008 and Maue 2009. ACE per storm, measured on a seasonal scale, at first appears to be a similar metric to the ratio as seen in Figure 3.2. ACE per storm (dashed line) was normalized by its standard deviation and plotted alongside the ratio (solid line). The two metrics correlate positively with a 0.96 correlation coefficient, and the trend for the ratio appears to follow the trend for ACE per storm. However, there are some years where the two metrics differ, such as 1969, 1982-1983, 1993-1994, 1997 and 2009. On a conceptual level, the ACE to Counts Ratio measures how seasonal activity differs (or is consistent) between the two metrics of ACE and number of TCs, whereas ACE per storm measures the average value of accumulated cyclone energy of all the TCs in a given season. An ACE per storm value would be the same if calculated for one storm with a certain ACE value or 15 storms that average that same particular ACE value. However, the ratio would be different in these two cases since it also takes into account the counts value relative to the counts of all seasons from 1948-2020. Since the ratio is considering both the variation in counts and in ACE each year, it adds an extra layer of understanding to the ACE per storm metric.

Various forms of frequency analysis were also performed on the ACE to Counts Ratio, in order to analyze its temporal and spatial characteristics. One of these is Fourier Analysis, which

represents signals as a sum of sine and cosine waves. It transforms a signal from the time domain to the frequency domain (Bohnenstiehl 2021). It can be represented by the equation:

$$f(\sigma) = \int_{t=-\infty}^{\infty} F(t)e^{-i2\pi\sigma t} dt \quad 9$$

Where  $t$  is time (years)

and  $\sigma$  is frequency (years<sup>-1</sup>) (Xie 2021b).

Fourier analysis was used to determine the dominant temporal patterns of the ratio, which will later be connected to various large scale phenomena such as El Niño-Southern Oscillation (ENSO) and the AMO.

Another analysis method used is Wavelet Analysis, which can be used to determine the statistical significance of the dominant frequencies found using Fourier Analysis. Wavelet Analysis uses wavelets that have varying frequency and limited duration, as opposed to the infinite, fixed frequency sine and cosine waves in Fourier Analysis. It is represented by the equation:

$$\gamma(s, \tau) = \int f(t)\Psi_{s,\tau}(t)dt \quad 10$$

Where  $\gamma(s, \tau)$  is the coefficient of the wavelet with scale  $s$  and time  $\tau$

$f(t)$  is the time series

and  $\Psi_{s,\tau}(t)$  is the complex conjugate of the wavelet with scale  $s$  and time  $\tau$  (Xie 2021c).

The ACE to Counts ratio is also analyzed using Empirical Mode Decomposition (EMD). EMD is similar to Fourier Analysis with a couple differences. In EMD, the output remains in the time domain, instead of being transformed to the frequency domain. Also, EMD does not assume the original time series is periodic; unlike Fourier and Wavelet Analysis, it can be used to analyze non-linear, non-stationary time series (Ryan 2019). In this analysis, EMD

decomposes the ratio into Intrinsic Mode Functions (IMFs), which each represent a different temporal signal of the ratio. These range from interannual signals to multidecadal signals to a simple trend of the time series. Each of these signals has a time series of coefficients, which can be removed from the ratio to create a filtered version of it. This will be discussed further in the composite analysis results in Section 4.2.2.

### 3.3 EOF Analysis and Track Density Analysis of Vertical Wind Shear

Empirical Orthogonal Function (EOF) analysis was performed on Atlantic VWS to analyze the dominant spatial patterns of VWS and to determine if these patterns are significantly correlated to phenomena such as (ENSO) and the ACE to Counts ratio. There were two domains selected for this analysis: 30°S-30°N, 90°W-0° (northern and southern Atlantic Ocean, or the “global” domain) and 0-45°N, 90°W-0° (northern Atlantic only, or the northern “hemispheric” domain). In the EOF analysis, singular value decomposition (SVD) was used to obtain the leading EOF modes of the monthly VWS data. The percent variance explained by each of the top 10 EOF modes was plotted to determine a truncation point, where the EOFs before this point would sum to explain a significant portion of the total variance, and the EOFs after this point would not be considered. Then, EOFs are plotted and analyzed to understand the dominant spatial patterns of the VWS. In addition, the varimax algorithm in MATLAB is used to rotate the EOFs, and those rotated EOFs (REOFs) that explain a significant portion of the total variance are also plotted and analyzed. The corresponding expansion coefficients (ECs), or weights of the EOFs (and the REOFs), are then correlated to the phenomena (ENSO, ACE to Counts ratio, etc.) to determine which EOFs significantly correlate with which phenomena (Xie 2021e).

A variation on the analysis of hurricane track density, proposed by Anderson and Gyakum (1989) in the Pacific and performed by Xie et al. (2005) in the Atlantic, was performed

on the starting points of tropical cyclones in the Atlantic. To do this, the Atlantic was divided into five zones: three along the Atlantic Main Development Region (referred to as the Eastern MDR, Central MDR, and Caribbean) and two north of the MDR (the NW North Atlantic and NE North Atlantic). These zones are visually depicted in Figure 3.3. Then, for the time period of 1951-2018 (referred to as the base period), the latitude-longitude starting point of every TC was determined. This was further broken down into the starting points of only tropical depressions, only tropical storms, and only hurricanes. If a TC weakened and then restrengthened into either a tropical storm or hurricane, only the first time it became a tropical storm/hurricane was considered as a start point. This process was repeated for two subsets of years, and the number of start points per zone for those subsets of years was compared to the number of start points per zone for the base period. Percent difference from the base period was then calculated. These results will be shown in Section 4.4.3.

### 3.4 Ratio Predictability

The ACE to Counts Ratio was also predicted, using similar methods as in Section 3.1. For this prediction, only a midseason model was ran and just the numerical values of the ratio were predicted (as opposed to classifying them). In addition, since the ratio is a newly developed metric, the only comparison used to measure model skill is comparison to climatology. The metrics used to measure model performance are mean squared error (MSE) and mean squared skill score (MSSS). MSE values for 10, 30, and 50 year climatology were calculated, and the MSSS, used to evaluate the prediction performance relative to climatology, was determined. The formulas for MSE and MSSS are:

$$MSE = \left(\frac{1}{n}\right) \sum (y_{actual} - y_{predicted})^2 \quad 11$$

$$MSSS = 1 - \frac{MSE_{prediction}}{MSE_{climatology}} \quad 12$$

where a positive MSSS value indicates skill over climatology (Xie 2021a).

Table 3.1. Number of years between 1948 and 2020 that fall into each class of each ACE classification system.

Number of years in each class		
Class	Z-score Class	CPC Class
1	7	0
2	22	23
3	20	26
4	16	6
5	8	18

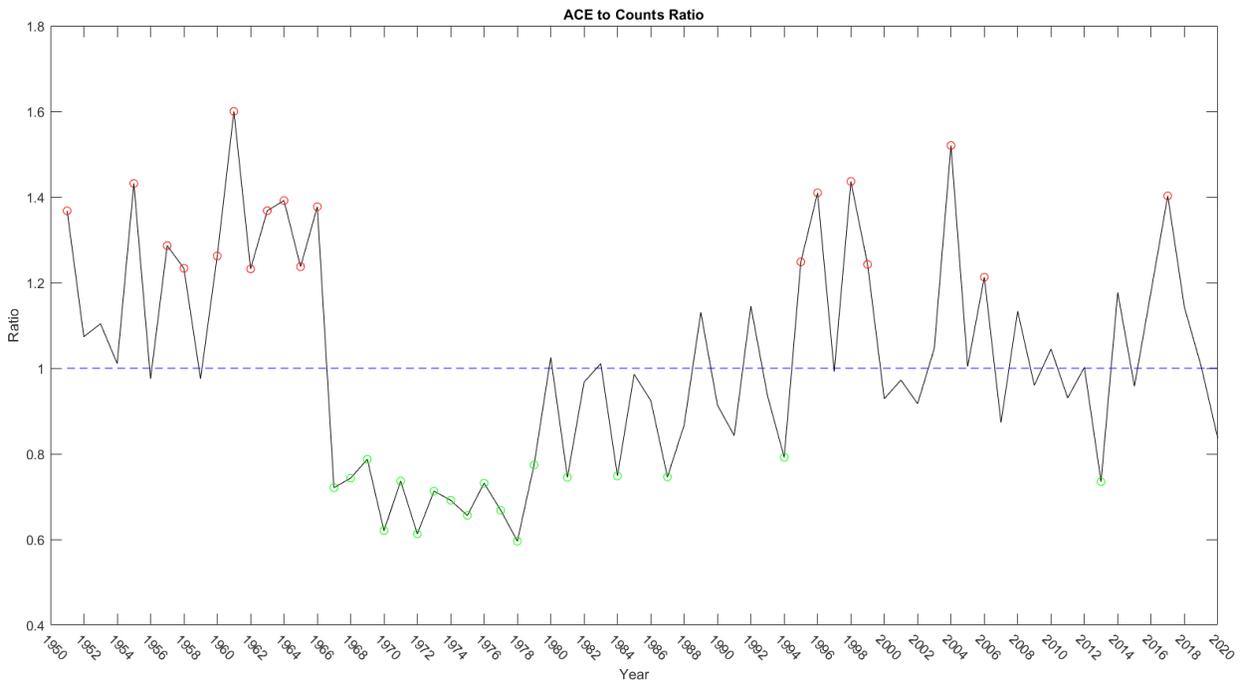


Figure 3.1. ACE to Counts Ratio (solid line). The ratio is centered around 1, marked by the dashed line. The 18 high ratio years and 18 low ratio years are marked by red and green circles, respectively.

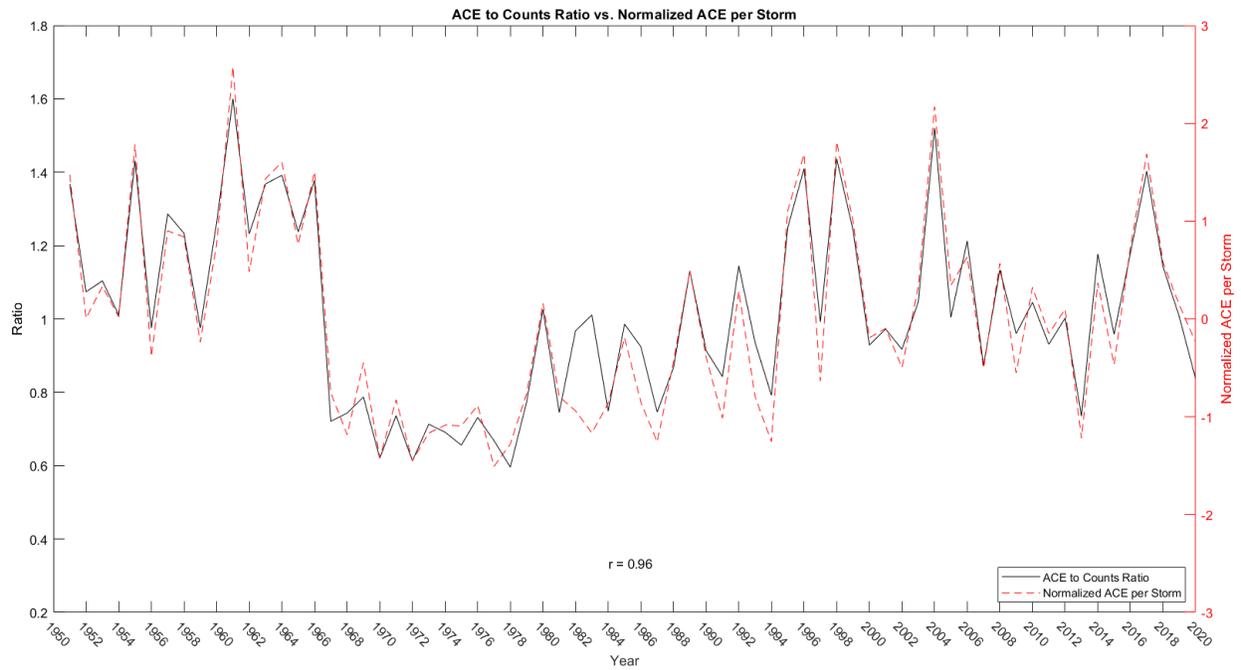


Figure 3.2. Comparison of the ACE to Counts Ratio (solid line, left axis) with ACE per storm (dashed line, right axis). ACE per storm is normalized by its standard deviation. Correlation coefficient between the two indices is also shown.

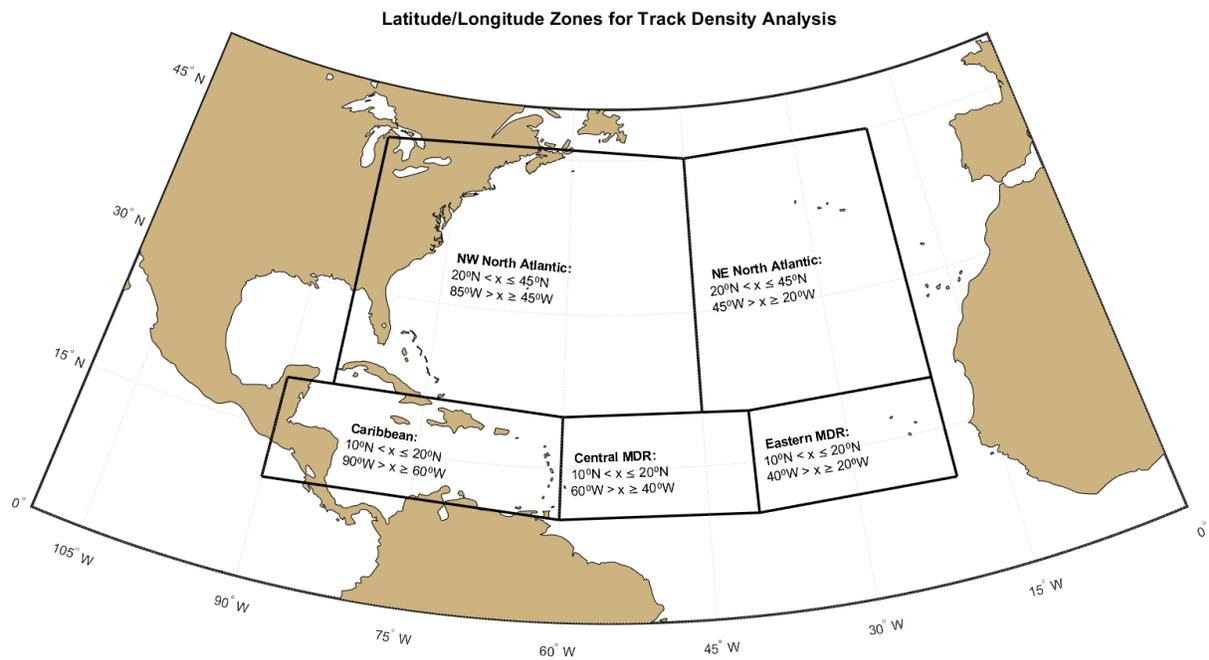


Figure 3.3. Track Density Analysis zones, with latitude ( $^{\circ}\text{N}$ ) and longitude ( $^{\circ}\text{W}$ ) extents included.

## CHAPTER 4

### Results

#### 4.1 Seasonal Hurricane Activity Classification Predictions

Seasonal ACE was predicted for each year for each sliding window test set. The ACE values were then classified into one of the five classes that was described in the methods section and compared to the actual classification for that year. The ACE predictions of two comparison organizations, Colorado State University (CSU) and Tropical Storm Risk (TSR), and multiple climatology predictions were also classified and compared to the actual classifications. This can be seen in Table 4.1, where the percentages correspond to the percent of years classified correctly. Each pair of left-right adjacent percentages is the same ACE prediction classified by the CPC Class and Z-score Class methods. The climatologies, in the bottom right panel of Table 4.1, have classification accuracies ranging from 20-35%. The classification accuracy of the comparison organizations is shown in the top right panel, listed by the month of the year that each organization released its predictions. The accuracy tends to increase when predictions are made later in the year.

The other sets of top-bottom panels show the results for the preseason (left) and midseason (center) models. Each row represents a different p-value used in backward stepwise regression, resulting in models with different numbers of predictors. For clarity, a model with a larger number of predictors in it includes all of the same predictors as a model with fewer predictors, plus an additional number. Models left blank are duplicates. The top (bottom) section of each panel are the LASSO (KNN) models.

When looking at the CPC Class results, almost all model runs outperform climatology. This is generally true for the Z-score Class as well, with the exception of some of the preseason

and three of the midseason predictions not outperforming the 30-year climatology. Overall, 83% of the model runs either match or exceed all predictions using climatology. For a fair comparison to the predictions of CSU and TSR, those organizations' April and May releases are compared to the preseason model, and their August releases to the midseason model. When doing this, the preseason models improve upon the accuracy of the preseason forecasts issued by the comparison organizations. The midseason models generally do as well, especially in the case of the LASSO predictions. However, for the CPC Class, half of the KNN predictions are outperformed by CSU and two of them by TSR.

In terms of the performance of the different methods studied in this research, LASSO generally tends to slightly outperform KNN, and the CPC Class method classifies predictions better than the Z-score Class method in majority of the cases (81%). There was no pattern found as to models with more or fewer predictors having better prediction accuracy. The most common predictors to appear in the models were AMO and TSA, the Tropical Southern Atlantic index (refer to Tables 2.1 and 2.2 for descriptions). Other predictors that appeared frequently were NINO34, MDRSST (Main Development Region Sea Surface Temperature), AO (Arctic Oscillation), MDROLR (Main Development Region Outgoing Longwave Radiation), and AMM (Atlantic Meridional Mode). Several of these predictors have previously been shown to affect Atlantic hurricane activity: AMO (e.g. Goldenberg et al., 2001), NINO34/ENSO (e.g. Goldenberg and Shapiro, 1996), and AMM (e.g. Patricola et al., 2014).

## 4.2 ACE to Counts Ratio

### 4.2.1 Frequency Analysis of ACE to Counts Ratio

A Discrete Fourier Transform was performed on the ACE to Counts Ratio (abbreviated as "ratio") to determine the most prominent frequency bands in the time series. The ACE to

Counts Ratio was first detrended using a first order polynomial, to remove any linear influence or trend on the ratio. This is shown in the left panel of Figure 4.1. The results of the Fourier Transform can be found in the right panel of Figure 4.1. The five most dominant frequency bands occur at 2.67 and 3.12 years (interannual signals), 8, 18 and 42 years, and are marked in the figure. The interannual signals occur on a timescale consistent with the frequency of ENSO events. The 8 year band is very close to what could be considered a decadal oscillation of the ACE to Counts Ratio. Finally, the 18 and 42 year periods suggest multidecadal variability.

To verify the significance of these frequency bands, Wavelet Analysis was also performed on the ACE to Counts Ratio using wavelet software provided by Torrence and Compo (1998). In this instance the ratio was first normalized by its standard deviation instead of being detrended. Figure 4.2a shows the normalized ratio. Figure 4.2b shows the wavelet power spectrum, with the contours showing spikes in power. Those that have statistical significance at the 95% confidence level are outlined in black. The “cone of influence,” outlined in black, shows which frequencies are considered as statistically significant at each time. Figure 4.2c is the global wavelet spectrum, which shows the overall wavelet power averaged over the length of the time series at each frequency. The dashed line represents a line of statistical significance, where global wavelet powers larger (to the right) of the line are significant with 95% confidence. The interannual signals identified in the Fourier Analysis are inside the cone of influence and are significant at the 95% confidence level. They also have large enough power on the global scale to be statistically significant in the global wavelet spectrum. The 8 year period is not statistically significant at the 95% level, but is mostly within the cone of influence, and is likely short enough for the dataset. The 18 and 42 year periods are both not statistically significant, and are either barely within or not at all within the cone.

Since the decadal and multi-decadal signals were not able to be evaluated in the wavelet analysis, the time series for the ACE to counts ratio was expanded to the years 1851-2020. The ratio was calculated in the same way as for the shorter time series, with the knowledge that the ACE and TC count data from earlier years in the time period may not be as reliable than more recent data. It is reasonable to believe, however, that the uncertainties in the earlier data would not be large enough to alter the frequency characteristics of the ACE to counts ratio time series. Fourier and Wavelet analysis was again performed on the expanded time series, and is shown in Figures 4.3 and 4.4. The dominant peaks in frequency in the Fourier analysis were determined by setting a threshold for the amplitude difference between a given frequency and the two neighboring frequencies. There were eight peaks in frequency, corresponding to periods of 2.22, 2.59, 3.2, 5.57, 8, 17.07, 36.57, and 128 years. Again, the 2-5 year frequencies can be considered a part of the same interannual frequency band. The interannual signals, along with the 8, 17.07, and 36.57 year periods are similar to the results for the shorter time series above.

Figure 4.4 shows the Wavelet analysis for the 169 year ratio timeseries. Similar to the shorter time series, the interannual signals show statistically significant levels of power in areas of the local wavelet spectrum (Figure 4.4b) and the global wavelet spectrum (Figure 4.4c) at 95% significance. The approximately decadal (8 year) and multidecadal (17.07 and 36.57 year) periods are all now within the cone of influence, for either a majority of the timeseries or close to half of the time series (71 out of 169 years) in the case of the 36.57 year period. The 128 year period is far outside the cone of influence because the time series is not long enough to resolve such low frequency signals, but the other frequency bands can be considered short enough for the timeseries when considering the Fourier analysis.

#### 4.2.2 Connections to Environmental Variables of TC Activity

Composite analysis was performed to determine the extent of a difference in environments between high ratio years and low ratio years. Figure 2.1 shows which years between 1951 and 2020 are considered high ratio and low ratio. Three different parameters conducive to tropical cyclone development were considered: SST, VWS, and mid-tropospheric moisture. The two sets of years that were used in the composite analysis were years with a ratio value of 1.2 or greater and years with a ratio value of 0.8 or less. There are 18 years total in each sample.

#### 4.2.2a Sea Surface Temperature

Composite analysis of SST was performed in the Atlantic domain (30°S-30°N, 90°W-0°, Figures 4.5-4.6, 4.9-4.10) and the Pacific domain (30°S-30°N, 150°E-75°W, Figures 4.11-4.15). The figures show the magnitude of the SST difference between the high and low ratio years, and directly below the results of the statistical significance test of those differences (1=significant at the 95% level, 0=not significant). Figures 4.5 and 4.6 show Atlantic domain SST composites during the North Atlantic hurricane season. Figure 4.5 shows composites for the 3-month JJA average (left) and 3-month ASO average (right). Figure 4.6 shows similar composites for September (left) and October (right). All four sets of composite analysis show a swath of positive SST differences across much of the North Atlantic MDR and Caribbean Sea, corresponding to higher SSTs for the high ratio years. These differences are generally statistically significant, with a slightly smaller area of significance corresponding to slightly lower SST differences in October. Overall, the ratio appears to be controlled partly by the hurricane season North Atlantic SST.

Further analysis was performed to determine the dominant frequencies of the observed SST differences between the high and low ACE to counts ratio years. It is well known that

North Atlantic SSTs fluctuate on a multidecadal scale in the form of the AMO (Enfield et al., 2001; Knight et al., 2005) or the Atlantic SST dipole mode (Servain, 1991), which are also shown to affect overall Atlantic hurricane activity (Goldenberg et al., 2001; Knight et al., 2006) and the variability of Atlantic hurricane track patterns (Xie et al., 2005). Because of this, the relationship between the AMO and the ratio was explored further. The ratio was correlated with both the yearly average of the monthly AMO index and each month of the AMO index from 1857-2020. The ratio that was expanded to 1851-2020 was used for these correlations. Correlation coefficients can be found in Table 4.2; all correlations are significantly positive at the 95% confidence level. This lends evidence to the idea that the ratio is modulated by the AMO on multidecadal timescales.

To see whether other significant SST patterns exist in the ratio unrelated to multidecadal influences, the portion of the ratio signal that could be influenced by the AMO was removed. This was done by performing EMD analysis on the original 70-year ratio (1951-2020), shown in Figure 4.7. IMF 3, 4, and 5 correspond roughly to the 18 year, 42 year, and longer multidecadal frequencies found in the Fourier Analysis. Next, a high pass filter was applied to the ratio by removing the amplitude of the IMF 3, 4, and 5 patterns. The new high pass ratio is plotted in Figure 4.8. Composite Analysis of Atlantic SST was then performed on the high pass ratio, using the 18 years with the highest and lowest ratio values. The SST composites are shown in Figures 4.9 (JJA and ASO averages) and 4.10 (September and October). The warmer pattern of SSTs for high ratio years still exists in the early part of the hurricane season (JJA) in the eastern Atlantic from about 15°N-30°N. However, the pattern is diminished and is not statistically significant during the peak and late season months (ASO average, September, and October). In these

months, the multidecadal signals appear to have more control over the SST differences, whereas in the early part of the season, interannual and decadal signals may still play a role.

This analysis was repeated for the Pacific domain, first for the original, unfiltered ratio [Figures 4.11 (JJA, left; ASO, right) and 4.12 (September, left; October, right)]. Though there are some weak positive differences centered around 155°W, 18°S and 160°E, 18°N, the areas are relatively small compared to the significant areas in the Atlantic. Across the equatorial Pacific, including the NINO34 region, there are no statistically significant differences in SST in any of the months. SST does not exhibit the same control over the unfiltered ratio in the Pacific as it does in the Atlantic, suggesting that remote effects from the Pacific do not directly affect the ACE to Counts ratio on all timescales.

To further understand any possible relationship between Pacific influences and the ratio, Pacific SST composite analysis was then performed on the high pass ratio, where the multidecadal oscillations were removed. The results are shown in Figures 4.13 (DJF and MAM; Atlantic hurricane preseason), 4.14 (JJA and ASO) and 4.15 (September and October). The results during the preseason (DJF, Figure 4.13) and the early hurricane season (JJA, Figure 4.14) show a swath of positive SST differences across the equatorial eastern Pacific, corresponding to higher SSTs for the high ratio years. A signal of positive SST differences in that area appears during the other months, though it is not as strong. When performing the significance test (bottom panels of Figures 4.13-4.15), these differences are not found to be statistically significant at either the 95% or 90% confidence levels. However, the magnitude of these SST differences is similar to the magnitude of the statistically significant differences of the unfiltered ratio in the Atlantic (Figures 4.5 and 4.6). Pacific SSTs do appear to control the interannual and

decadal signals of the ratio, but more work will have to be done to determine the significance of this connection.

#### 4.2.2b Vertical Wind Shear

Composite analysis of VWS was performed on the ratio in the same Atlantic domain as for the SST composites and is shown in Figures 4.16 (JJA and ASO averages) and 4.17 (September and October). The results show a zonal gradient in shear in the central N. Atlantic at about 40°W, with higher shear values to the west and lower shear values to the east for high ratio years. For low ratio years these differences are swapped and higher shear exists in the eastern N. Atlantic (relative to the high ratio years). Both the positive and negative differences in shear are statistically significant through most of the hurricane season, though the gradient is slightly less pronounced in October. A theory that arises from this is that in the lower ratio years (a relatively larger number of less powerful storms), TCs forming off the W. African coast immediately encounter higher shear, so they may not develop as far eastward in the MDR as is the case for other years. Decreased shear in the central MDR (between 40°-60°W) could lead to more storms developing in low ratio years, but they would have less time to strengthen because they are already further westward in the Atlantic. This theory will be expanded on in section 4.4.3 using track density analysis to determine the frequency of TC starting points in different zones in the Atlantic.

The composite analysis of VWS was repeated for the high pass ratio, like it was done for SST (Figure 4.18, JJA/ASO; Figure 4.19, September/October). The statistical significance of VWS differences was tested at both the 95% and 90% levels; the significant areas at 95% confidence were limited to just a few gridpoints, so the significant areas at 90% confidence are shown instead. The zonal gradient is still present in September, though the shear magnitude

differences are lower ( $\pm 2$  m/s) than in Figure 4.17 ( $\pm 3-4$  m/s) and only the positive differences are significant. In the other cases, specifically JJA and October, the zonal gradient is highly diminished. In general, the significance in the composites for the higher frequencies of the ratio is concentrated more locally, as opposed to covering a larger, more general area in the Atlantic when the high pass filter is not applied, suggesting that the VWS difference between high and low ACE to counts ratio years appears mostly at multidecadal frequencies like the Atlantic SST composite signals.

#### 4.2.2c Mid-Tropospheric Moisture

Composite analysis of RH700 was performed on the ratio and is shown in Figures 4.20 (JJA and ASO averages) and 4.21 (September and October). Unlike the SST and VWS results, the area of significant RH700 differences between high and low ratio years is not very large, nor does it encompass the majority of the MDR. There is one dry region (for high ratio years) that is statistically significant across the ASO, September and October composites, but it is concentrated at only one grid point longitudinally (45-50°W).

To further explore a possible relationship, RH850-500 was also studied and is found in Figures 4.22 (JJA and ASO averages) and 4.23 (September and October). The signal of the composite magnitude and significance is not exactly the same as in RH700. While the central Atlantic still shows drier conditions for high ratio years, the magnitude of the differences is larger around 0-15°N latitude, as opposed to being largest around 15-30°N in the RH700 composites. In addition, the differences in October are less pronounced, and the overall patterns of statistical significance look quite different. The lack of a coherent signal differentiating the high ratio years from the low ratio years suggests that mid-tropospheric moisture, in the form of RH, may not be a distinguishing factor for the ratio.

### 4.3 Spatial Variability of Atlantic Vertical Wind Shear and Analysis

#### 4.3.1 EOF Spatial Analysis

The composite analysis results for SST were interesting to note the Atlantic exhibited control over the ratio more conclusively than the Pacific. However, it has been shown that ENSO modulates the VWS in the Atlantic MDR on interannual timescales (Goldenberg and Shapiro 1996). A question then arose of whether the effects on Atlantic VWS from ENSO would influence the ratio. To answer, Atlantic VWS was spatially decomposed using EOF analysis, to see which modes (if any) would correlate with ENSO indices, the ratio, overall Atlantic TC activity, and other indices. Two domains were chosen for the analysis: 30°S - 30°N, 90-5°W (the “global” domain) and 0-45°N, 90°W-0 (the northern “hemispheric” domain). In both domains, EOF analysis was performed on monthly VWS using only the Atlantic hurricane season months (JJASON). Separate EOF analysis was also performed on an average of the JJASON VWS values. The monthly hurricane season values and hurricane season average values will be referred to as MHS and HSA, respectively. For validation, EOF analysis using the full time series (all months) was also performed, and the spatial pattern results and percent variance explained were very similar to that of the MHS analysis.

As mentioned in the methods section, the percent variance explained by the top 10 principal components (PCs) for each EOF analysis was plotted to determine, through a truncation point, which PCs are significant. These plots are shown in Figures 4.24 (global domain, MHS VWS), 4.25 (hemispheric domain, MHS VWS), 4.26 (global, HSA VWS), and 4.27 (hemispheric, HSA). From these plots, the top 7 global and top 5 hemispheric MHS EOFs, and the top 8 global and top 6 hemispheric HSA EOFs, were chosen. Similar truncation analysis was

performed on the REOFs. The top 5 global and top 3 hemispheric MHS REOFs, and the top 4 global and top 4 hemispheric HSA REOFs, were chosen.

Figures 4.28 and 4.29 show the top EOF spatial patterns for MHS VWS in the global and hemispheric domains, respectively. The percent variance explained by each EOF is also noted. For brevity, only the top 4 EOFs are shown. In both domains, EOF1 shows a north-south gradient. EOF3 is also similar in both domains, showing a north-south dipole mode. EOF2 is slightly different between the two domains; it shows an equatorial gradient in the global domain and a pattern centered at 20°N in the hemispheric domain. EOF4 shows a northwest-southeast dipole mode in both domains, though the negative pole is not as well defined/shared between the two domains. The REOFs (not shown) for these domains exhibit similar spatial patterns to each of their corresponding EOFs.

Spatial patterns for the top 4 HSA EOFs are also shown in Figures 4.30 (global domain) and 4.31 (hemispheric). They are fairly similar to the MHS spatial patterns, especially in the hemispheric case. The minor differences are highlighted as follows. The global EOF2 shows less of an equatorial gradient, and more of a polar pattern with two negative poles centered near 5°N, 30°W and 10°S, 60°W. The positive patterns are concentrated across all longitudes in the domain around 30°N and 30°S. The global EOF4 shows a similar dipole pattern to the hemispheric case, but the positive pole is shifted much further southward. Once again, the REOFs exhibit similar spatial patterns to the corresponding EOFs in the HSA case.

#### 4.3.2 Correlations of Vertical Wind Shear EOFs with ENSO and other indices

Table 4.3 shows the correlation of the expansion coefficients (ECs) for the MHS VWS EOFs with multiple parameters in both the global and hemispheric domains. The Arctic Oscillation (AO), Atlantic Multidecadal Oscillation (AMO), North Atlantic Oscillation (NAO),

and multiple ENSO parameters (NINO12, NINO3, NINO34, NINO4) were studied. The correlation coefficients that are statistically significant for the length of the timeseries at the 95% level are highlighted. For the AO and the NAO, correlations with certain ECs and rotated ECs are statistically significant, but they are all domain dependent (i.e. not significant across both domains). There is one EC that correlates significantly with the AMO across both domains, EC5, but the correlation doesn't hold when EC5 is rotated. ENSO, however, does have a significant negative correlation with the VWS EC2 through the NINO12, NINO3 and NINO34 indices. The significance holds for all three indices in both the rotation of EC2 and across domains. EC5 is the only other expansion coefficient to correlate significantly with all the ENSO indices, but it is domain dependent, only correlating with VWS in the global domain.

It was also desired to correlate each of the ECs with the ratio and a measure of hurricane activity (ACE was chosen). Since the ratio and ACE have both been used as yearly quantities in this study, the HSA EOFs were used for these correlations. This is shown in Table 4.4, where they are correlated with the top 4 ECs for HSA VWS in both the global and hemispheric domains. Correlations with the high pass ratio are also performed. The significance testing of the correlation coefficients requires higher values than that of the MHS time series since there are fewer values in the HSA case. The ECs with statistically significant correlations to any of these parameters are EC1 and EC2. EC1 and rotated EC1 correlate significantly with the ratio in the hemispheric domain. However, there is close to zero correlation of EC1 with ACE. EC2 correlates significantly with both the ratio and ACE in each domain, and the significance holds in the global domain when EC2 is rotated. When the high pass filter is applied to the ratio, the correlations with the VWS ECs generally become weaker. However, the correlation with EC2 is

still statistically significant in both domains, as well as the correlation with rotated EC1 in the global domain.

The above findings suggest that EC1 which is significantly correlated with Atlantic SST signals (such as AMO) is likely a factor affecting the ACE to counts ratio, whereas EC2, which is significantly correlated with ENSO, is an important factor for the interannual variability of the overall Atlantic hurricane activity (ACE). After finding significant correlations with each of the top two ECs of VWS, correlations of those ECs with other climate indices were also performed. These results will be discussed in the conclusion section.

#### 4.3.3 Track Density Analysis

The VWS composite results for the ratio showed a zonal gradient in wind shear across the Atlantic, with lower shear values in the eastern Atlantic and higher shear values in the central Atlantic for high ratio years (refer to Figures 4.16 and 4.17). Analysis of the starting points of tropical cyclones in the Atlantic was performed to see if a link exists between the VWS gradient in the ratio and the location of TC genesis. As mentioned earlier, the five zones the Atlantic was divided into were the Eastern MDR, Central MDR, Caribbean, NW North Atlantic and NE North Atlantic. These zones are visually depicted in Figure 3.3. The latitude-longitude starting point of every TC for two subsets of years was determined: the 17 years with the highest and lowest values of EC1 in the yearly VWS EOF. The number of start points per zone for those subsets of years was compared to the number of start points per zone for the base period. Percent difference from the base period was calculated and is shown in Figure 4.32 for all TC starts, only tropical depressions, only tropical storms, and only hurricanes.

A trend can be seen in the Eastern MDR and the Caribbean that the low VWS EOF1 years have a higher number of storm starts than the base period (with the exception of tropical

depressions), and the high VWS EOF1 years have fewer starts than the base period. These two areas correspond to areas with higher wind shear for low ratio years (Figures 4.16 and 4.17). From Table 4.4, EOF1 has a significant positive correlation with the ratio in the hemispheric domain. From this connection between VWS EOF1 and the ratio, the increased number of storm starts for low EOF1 years could be connected to the VWS composite signals in the ratio. However, the positive correlation indicates that low EOF1 years would correspond to low ratio years. For low ratio years, there are higher shear values in the Eastern MDR and Caribbean (Figures 4.16 and 4.17), which should reduce the number of storms that would develop. The track density analysis for VWS EOF1 shows the opposite result. So, the theory of TCs starting further westward in the MDR during low ratio years, thus giving them less time to strengthen across the Atlantic, cannot be confirmed at this time. Plotting tracks, instead of just starting points, may provide more information. See Xie et al. 2005 for more details.

#### 4.4 Ratio Predictability

Since the ratio is a yearly parameter, attempts to predict it were made, similar to predicting other metrics of seasonal Atlantic hurricane activity. The same methodology used in the 5-class classification predictions of ACE was used to predict the ratio, except classifications were not made. Also, the expansion coefficients from the HSA VWS EOF analysis for both the global and hemispheric domains were included as predictors, in addition to the predictors listed in Tables 2.1 and 2.2.

A midseason model was ran using the same predictor month restrictions as the midseason classification model, and the results can be found in Table 4.5. The numerical value of the ratio was predicted for each sliding window, and the MSE was calculated for each sliding window using the actual ratio values. The top panel shows the results for the LASSO models, and the

middle panel the KNN models. MSE values for the three climatologies are shown in the bottom panel. The “MSE Average” for each model refers to the average MSE of the 10 sliding windows used in that model run, and the MSSS is calculated using this average value. The green MSSS values correspond to the model runs that outperform each climatology, and red values are predictions with less skill than climatology. Backward stepwise regression was performed with p-values of 0.15, 0.2, and 0.25, similar to the classification models, but those predictions failed to outperform climatology in both the LASSO and KNN models and are not included in the paper.

Table 4.5 shows that the LASSO method was more successful in predicting the ratio than the KNN method. Almost all LASSO model runs outperform all three climatologies, with the model with least error being the 17-predictor model. In the case of KNN, only the 30 year climatology is outperformed (by all but one model run). Overall, there is still room for improvement in predicting the ratio.

Table 4.1. Results of the ACE classification prediction models. Comparison models from Colorado State University, Tropical Storm Risk, and simple climatologies are also shown. Each percentage is a different model run and shows the percent of test years classified correctly. Darker (lighter) highlights indicate higher (lower) accuracy. Models left blank are duplicate models.

	Preseason (Jan-Apr) Model			Midseason (Jan-Jul) Model			Comparison			
	LASSO			LASSO						
p value	# Predictors	CPC %	ZSC %	# Predictors	CPC %	ZSC %	Agency	Month	CPC %	ZSC %
0.01	1			2	50%	48%	CSU	April	14%	14%
0.02	4	35%	43%	8	53%	51%	CSU	June	21%	21%
0.03	4			9	56%	46%	CSU	July	29%	14%
0.04	4			11	43%	36%	CSU	August	43%	21%
0.05	4			11			TSR	April	22%	17%
0.1	4			13	52%	39%	TSR	May	20%	20%
0.15	10	42%	33%	15	50%	38%	TSR	June	42%	25%
0.2	15	41%	26%	17	49%	29%	TSR	July	41%	24%
0.25	17	42%	29%	18	55%	39%	TSR	August	39%	6%
	KNN			KNN						
p value	# Predictors	CPC %	ZSC %	# Predictors	CPC %	ZSC %	Climatology		CPC %	ZSC %
0.01	1			2	33%	38%	10-year Climatology		30%	25%
0.02	4	28%	34%	8	58%	33%	30-Year Climatology		28%	35%
0.03	4			9	62%	34%	50-Year Climatology		28%	20%
0.04	4			11	41%	38%				
0.05	4			11						
0.1	4			13	37%	38%				
0.15	10	31%	29%	15	39%	43%				
0.2	15	42%	45%	17	44%	38%				
0.25	17	41%	38%	18	45%	43%				

Table 4.2. Correlation coefficients of the AMO to the ratio from 1857-2020. Each month of the AMO was correlated to the ratio individually; a yearly average of the AMO index was also correlated. All correlation coefficients are statistically significant at the 95% confidence level.

Month of AMO Index	Correlation to Ratio
January	0.296
February	0.297
March	0.260
April	0.278
May	0.331
June	0.368
July	0.324
August	0.308
September	0.266
October	0.274
November	0.308
December	0.317
Yearly AMO	0.345

Table 4.3. Correlation coefficients between the monthly hurricane season (MHS) VWS ECs and climate indices Arctic Oscillation (AO), Atlantic Multidecadal Oscillation (AMO), North Atlantic Oscillation (NAO), and ENSO indices (NINO12, NINO3, NINO34, and NINO4). The top panel shows the EOFs from the global domain (30°S - 30°N, 90-5°W), and the bottom panel shows the EOFs from the hemispheric domain (0-45°N, 90°W-0). Statistically significant correlation coefficients at the 95% confidence level are highlighted.

**June - November Monthly EOFs, Global Domain**

	ao	amo	nao	nino12	nino3	nino34	nino4
EC1	0.01	-0.10	0.02	-0.07	-0.04	-0.01	-0.04
EC2	0.07	-0.08	-0.04	-0.13	-0.13	-0.11	-0.08
EC3	-0.14	0.08	-0.15	0.03	0.07	0.06	0.09
EC4	0.10	-0.01	0.07	0.00	0.07	0.08	0.11
EC5	0.02	0.12	-0.03	0.10	0.12	0.13	0.18
EC6	0.09	0.09	0.04	0.11	0.07	0.05	0.11
EC7	-0.06	0.04	-0.04	-0.10	-0.06	-0.04	-0.05
Rot. EC1	0.06	-0.10	-0.03	-0.13	-0.12	-0.10	-0.08
Rot. EC2	0.08	-0.07	-0.04	-0.12	-0.13	-0.12	-0.08
Rot. EC3	-0.11	0.06	-0.13	0.02	0.06	0.05	0.08
Rot. EC4	0.07	-0.09	0.08	-0.06	-0.03	-0.02	-0.04
Rot. EC5	0.13	0.06	0.07	0.08	0.12	0.13	0.20

**June - November Monthly EOFs, Hemispheric Domain**

	ao	amo	nao	nino12	nino3	nino34	nino4
EC1	0.05	-0.06	-0.02	-0.07	-0.07	-0.04	-0.08
EC2	0.07	-0.12	0.00	-0.14	-0.14	-0.13	-0.09
EC3	-0.08	0.00	-0.04	0.00	-0.07	-0.09	-0.12
EC4	-0.09	0.11	-0.12	0.05	0.097	0.095	0.17
EC5	0.04	-0.10	0.01	0.08	0.04	0.03	0.03
Rot. EC1	0.07	-0.10	-0.01	-0.11	-0.11	-0.09	-0.10
Rot. EC2	0.05	-0.11	0.01	-0.13	-0.13	-0.12	-0.06
Rot. EC3	-0.09	0.00	-0.05	0.00	-0.06	-0.08	-0.11

Table 4.4. Correlation coefficients between the hurricane season average (HSA) VWS ECs and the unfiltered ACE to Counts Ratio, the ratio with a high pass filter applied, and yearly ACE.

The top panel shows the EOFs from the global domain (30°S - 30°N, 90-5°W), and the bottom panel shows the EOFs from the hemispheric domain (0-45°N, 90°W-0). Statistically significant correlation coefficients at the 95% confidence level are highlighted.

**June - November Yearly EOFs,  
Global Domain**

	Ratio	High Pass Ratio	ACE
EC1	0.17	0.19	-0.04
EC2	0.38	0.25	0.46
EC3	-0.20	0.00	-0.02
EC4	-0.16	0.10	-0.22
Rot. EC1	0.36	0.26	0.28
Rot. EC2	0.29	0.18	0.50
Rot. EC3	0.05	0.04	-0.19
Rot. EC4	-0.19	0.01	-0.07

**June - November Yearly EOFs,  
Hemispheric Domain**

	Ratio	High Pass Ratio	ACE
EC1	0.31	-0.01	0.02
EC2	0.29	0.25	0.32
EC3	-0.05	-0.10	0.04
EC4	0.20	0.00	0.09
Rot. EC1	0.37	0.20	0.28
Rot. EC2	-0.19	0.00	-0.10
Rot. EC3	-0.10	-0.23	-0.17
Rot. EC4	-0.06	-0.01	0.06

Table 4.5. Results of the ACE to Counts Ratio prediction models. Each row is a different model run. The average MSE from the ten sliding windows of each model run and the MSSS when compared to three different climatologies are shown. Positive (green highlighted) MSSS values show positive skill compared to climatology, and negative (red highlighted) values show negative skill. MSE values for the climatologies are also shown.

August Model - LASSO Predictions						
Backward Stepwise		MSE Average	MSSS - 10 Year Clim	MSSS - 30 Year Clim	MSSS - 50 Year Clim	
p value	# Predictors					
0.01	5	4.751	-0.18	0.07	-0.13	
0.02	11	4.021	0.00	0.21	0.05	
0.03	14	3.820	0.05	0.25	0.09	
0.04	14	3.820	0.05	0.25	0.09	
0.05	17	3.541	0.12	0.31	0.16	
0.1	21	3.895	0.03	0.24	0.08	

August Model - KNN Predictions						
Backward Stepwise		MSE Average	MSSS - 10 Year Clim	MSSS - 30 Year Clim	MSSS - 50 Year Clim	
p value	# Predictors					
0.01	5	5.940	-0.48	-0.16	-0.41	
0.02	11	4.526	-0.13	0.11	-0.07	
0.03	14	4.245	-0.06	0.17	-0.01	
0.04	14	4.245	-0.06	0.17	-0.01	
0.05	17	4.475	-0.11	0.12	-0.06	
0.1	21	5.099	-0.27	0.00	-0.21	

Climatology Validation			
Climatology	10 Year	30 Year	50 Year
MSE	4.021	5.112	4.214

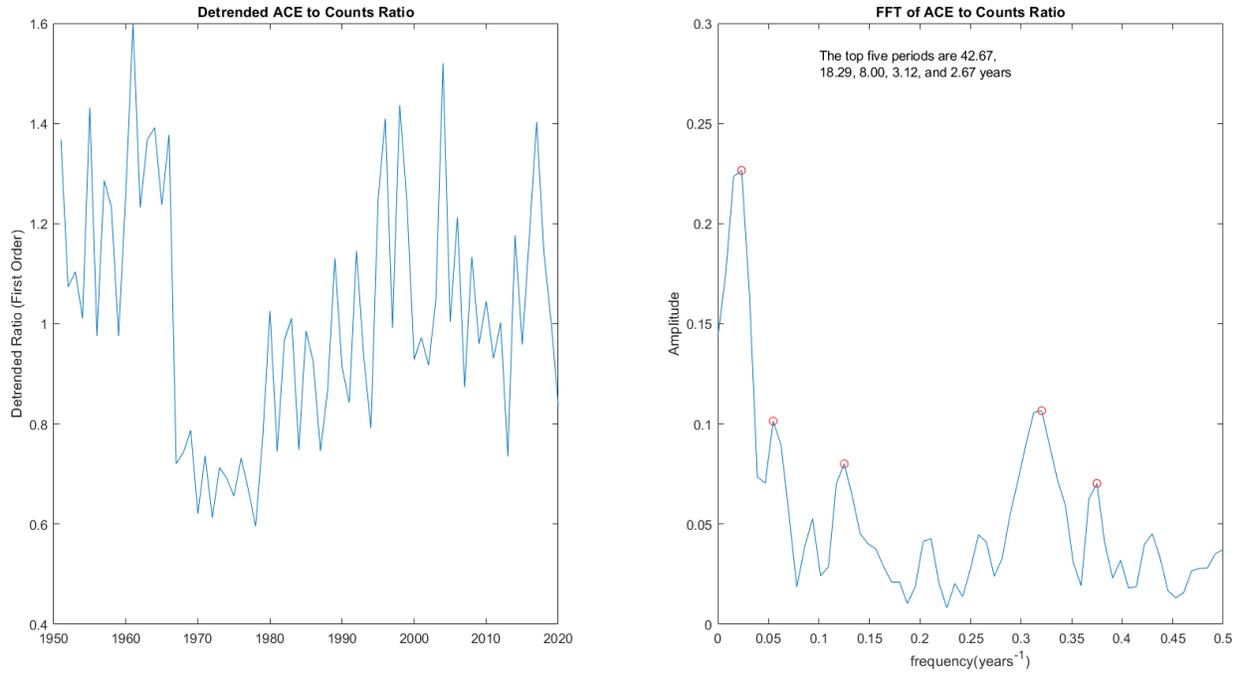


Figure 4.1. The detrended ACE to Counts Ratio from 1951-2020 (left), and Fourier analysis of the ratio (right). The five most dominant frequency bands occur at 2.67, 3.12, 8, 18 and 42 years.

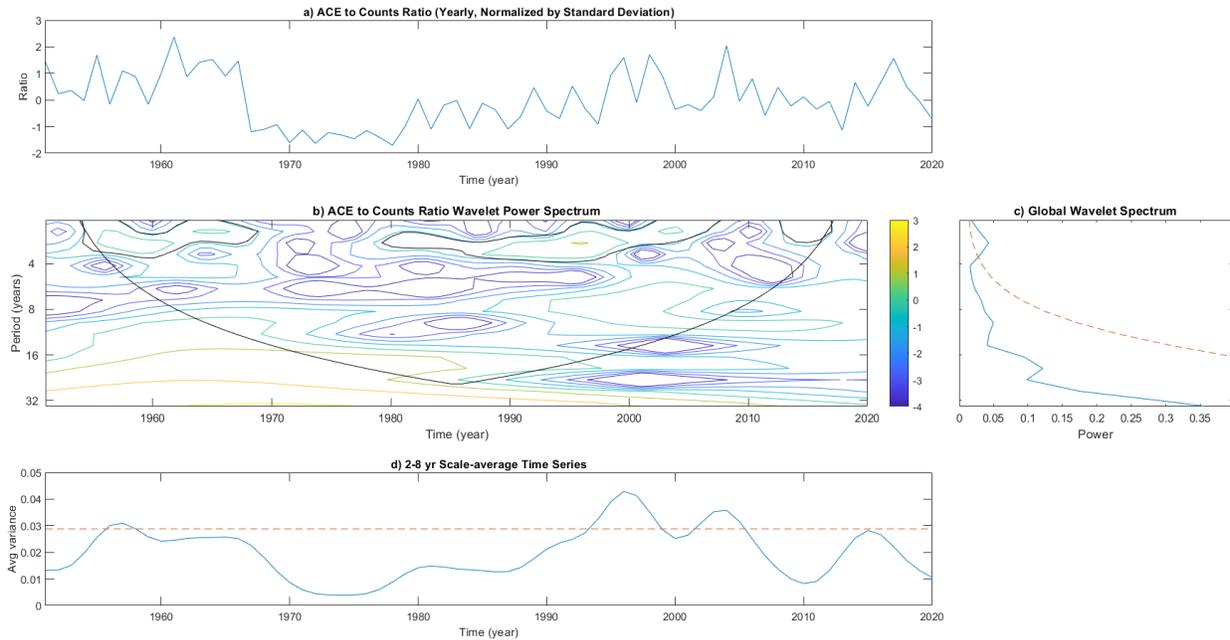


Figure 4.2. Wavelet analysis of the ACE to Counts Ratio from 1951-2020. (a) shows the ratio, normalized by standard deviation. (b) shows the wavelet power spectrum, with power scale to the right. The black contours and black cone show the statistically significant power contours and the cone of influence, respectively. (c) shows the global wavelet spectrum; global power (solid line) and 95% significance contour (dashed line). (d) shows the variance of the 2-8 year scale-average time series (solid line), along with the 95% significance value (dashed line).

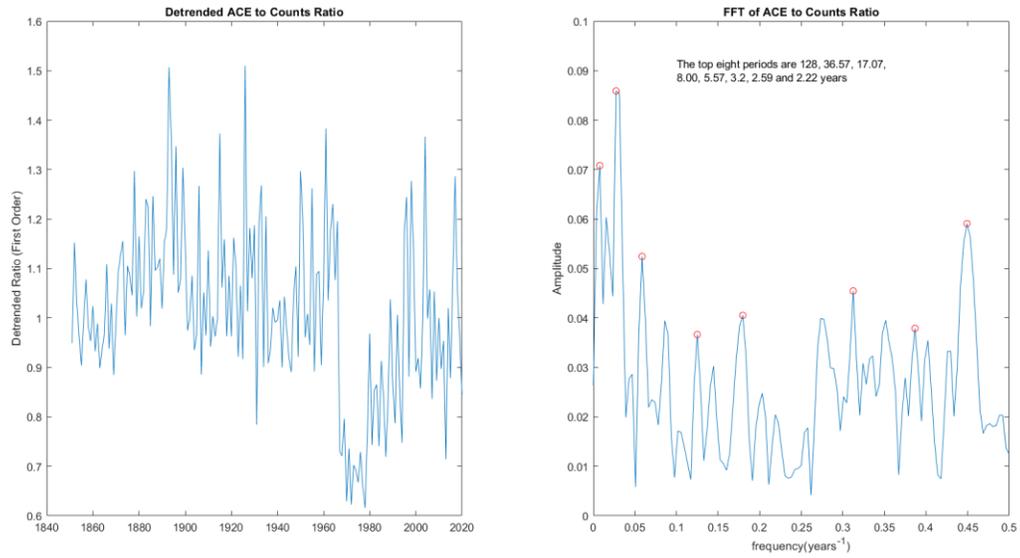


Figure 4.3. The detrended ACE to Counts Ratio from 1851-2020 (left), and Fourier analysis of the ratio (right). The eight most dominant frequency bands occur at 2.22, 2.59, 3.2, 5.57, 8, 17.07, 36.57, and 128 years.

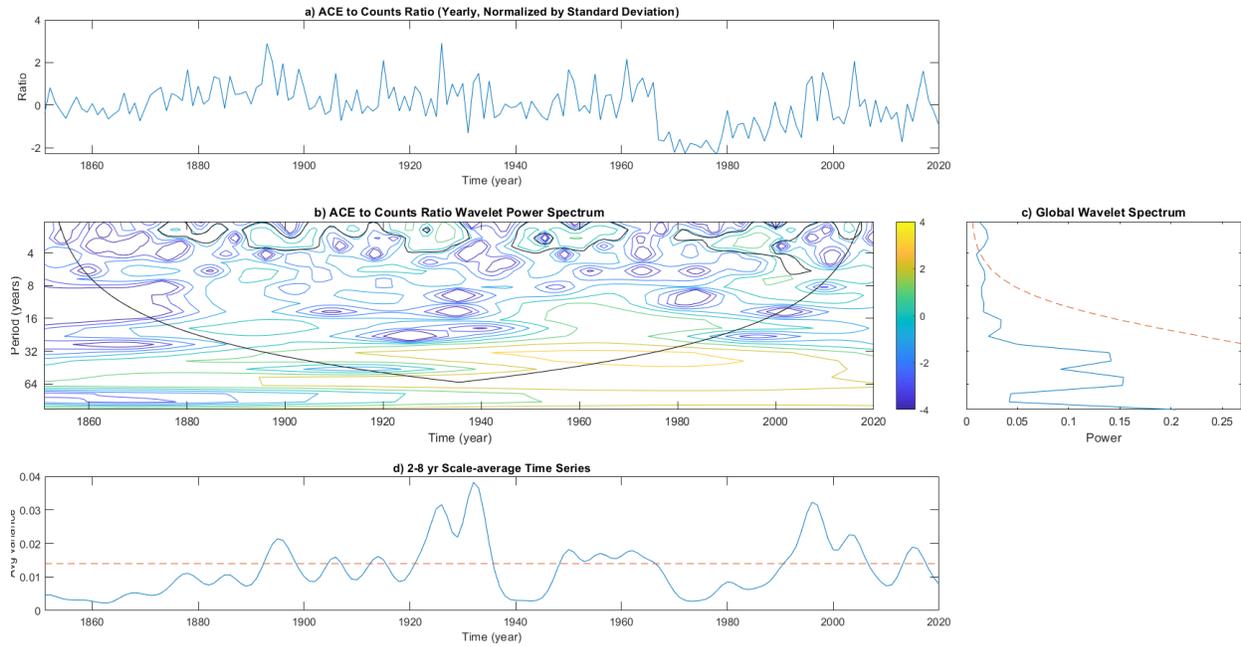


Figure 4.4. Wavelet analysis of the ACE to Counts Ratio from 1851-2020. (a) shows the ratio, normalized by standard deviation. (b) shows the wavelet power spectrum, with power scale to the right. The black contours and black cone show the statistically significant power contours and the cone of influence, respectively. (c) shows the global wavelet spectrum; global power (solid line) and 95% significance contour (dashed line). (d) shows the variance of the 2-8 year scale-average time series (solid line), along with the 95% significance value (dashed line).

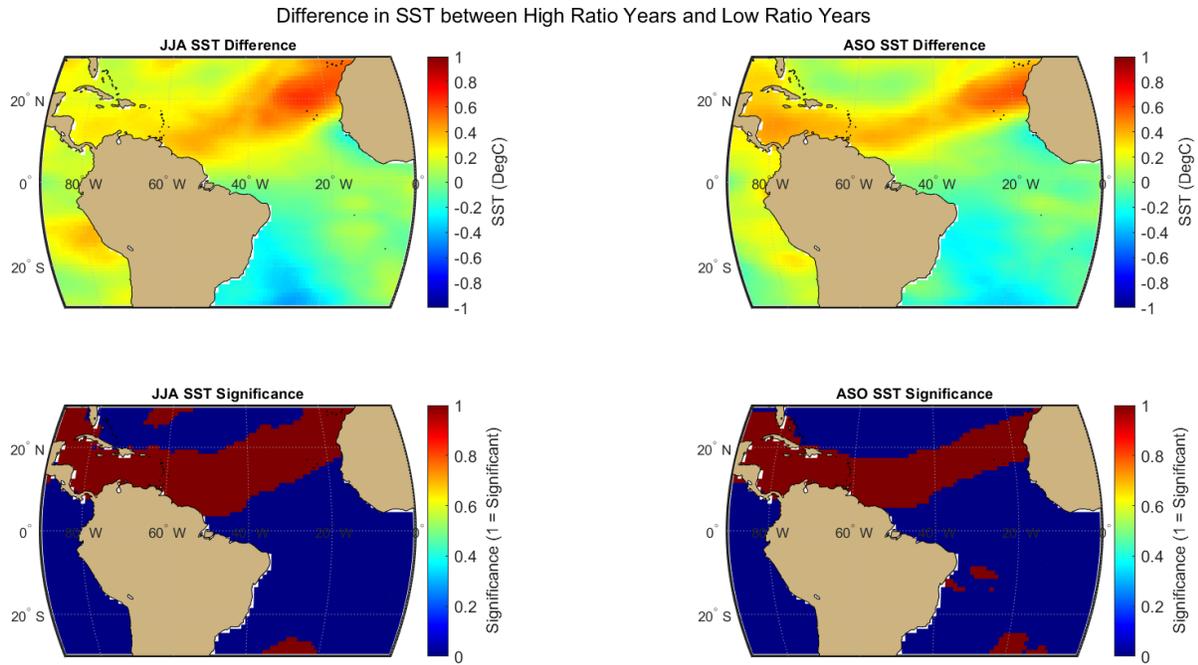


Figure 4.5. Atlantic Ocean SST Composite Analysis of ACE to Counts Ratio for JJA (left) and ASO (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

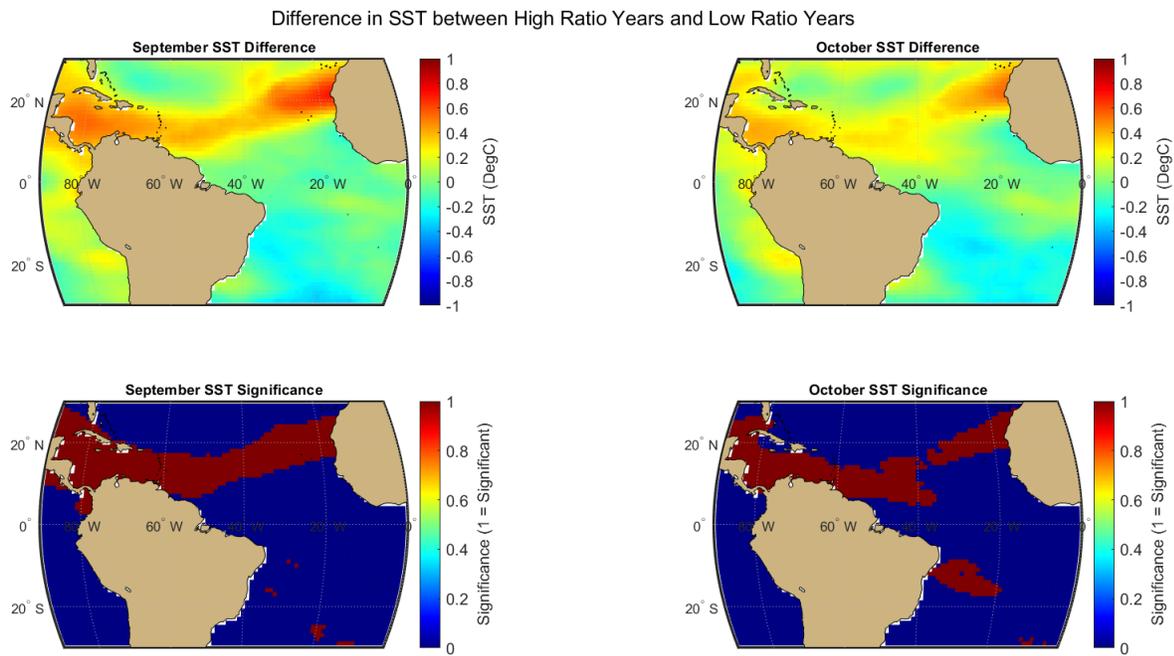


Figure 4.6. Atlantic Ocean SST Composite Analysis of ACE to Counts Ratio for September (left) and October (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

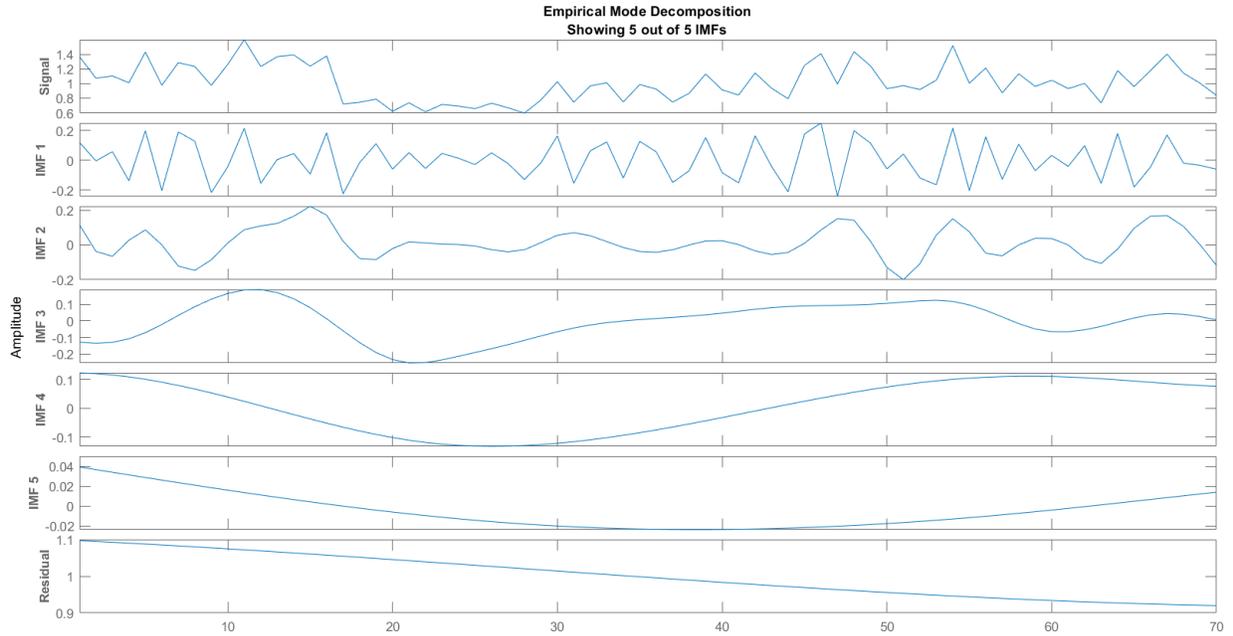


Figure 4.7. EMD Analysis of the ACE to Counts Ratio. Top panel shows the original ratio timeseries (signal), followed by the top 5 IMFs in descending order from highest frequency to lowest frequency. The bottom panel shows the residual trend of the ratio. The x-axis denotes the years of the timeseries, from 1951-2020.

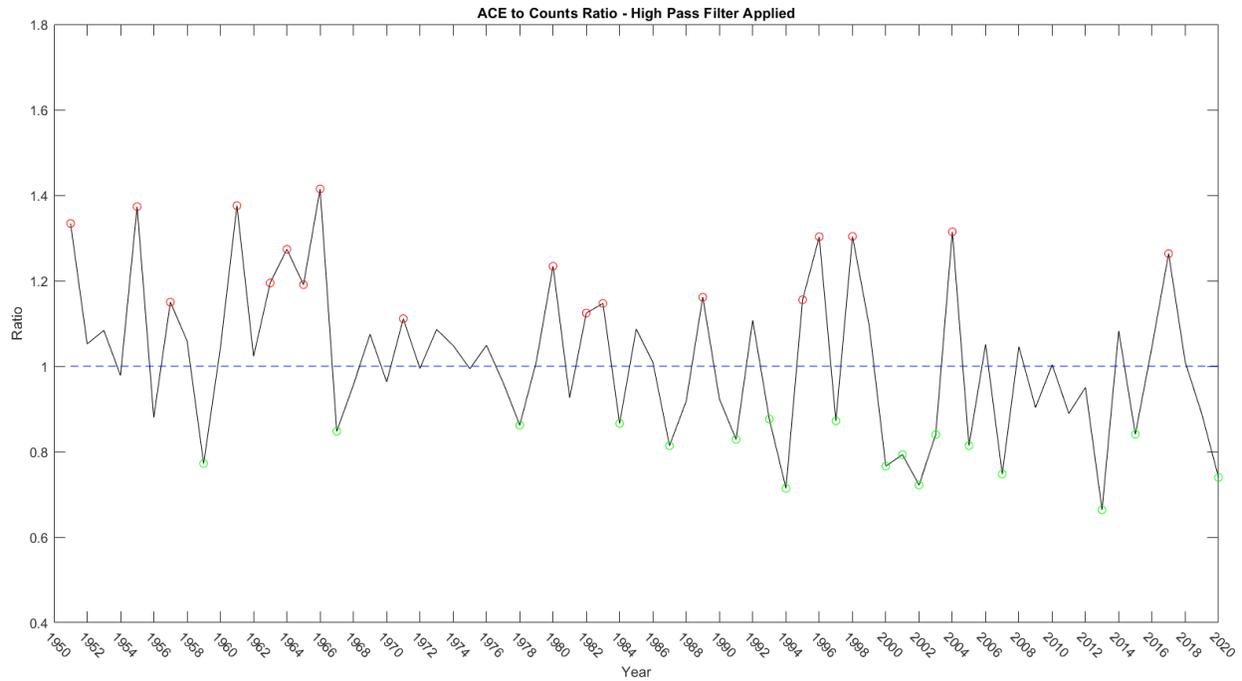


Figure 4.8. The ACE to Counts Ratio with a high pass filter applied. The lowest frequencies corresponding to IMFs 3, 4, and 5 from Figure 4.7 are removed. The 18 high ratio years and 18 low ratio years of the high pass ratio are marked by red and green circles, respectively.

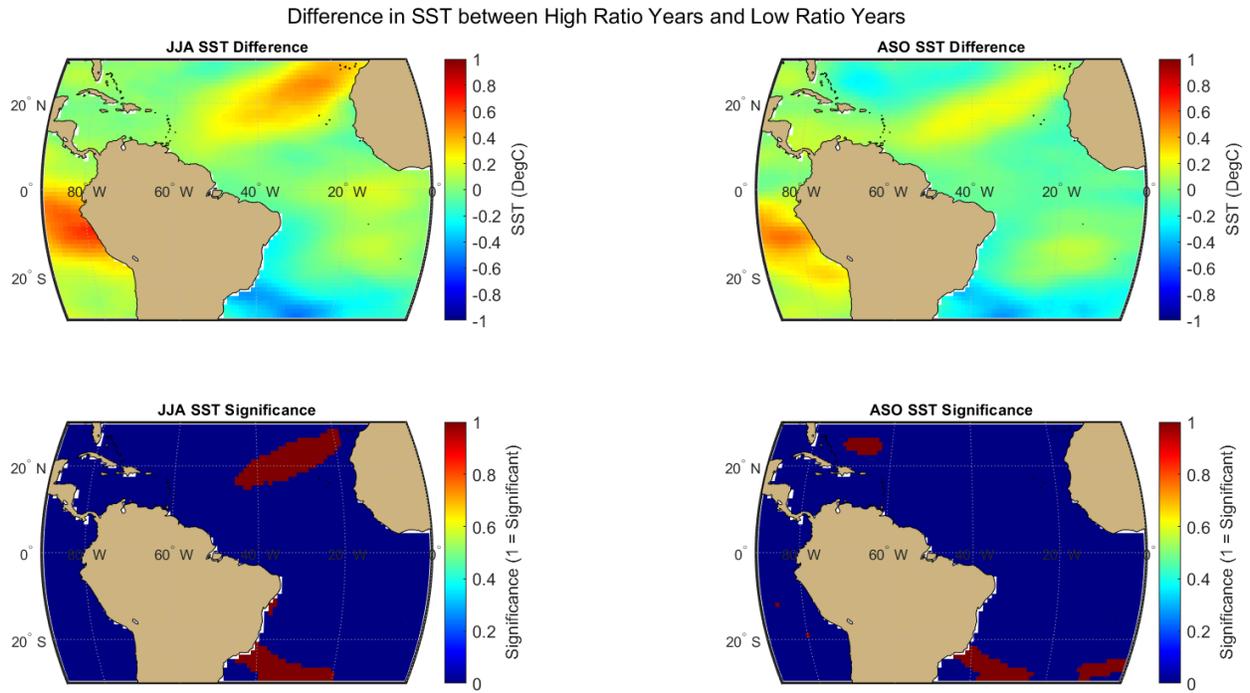


Figure 4.9. Atlantic Ocean SST Composite Analysis of the high pass ACE to Counts Ratio for JJA (left) and ASO (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

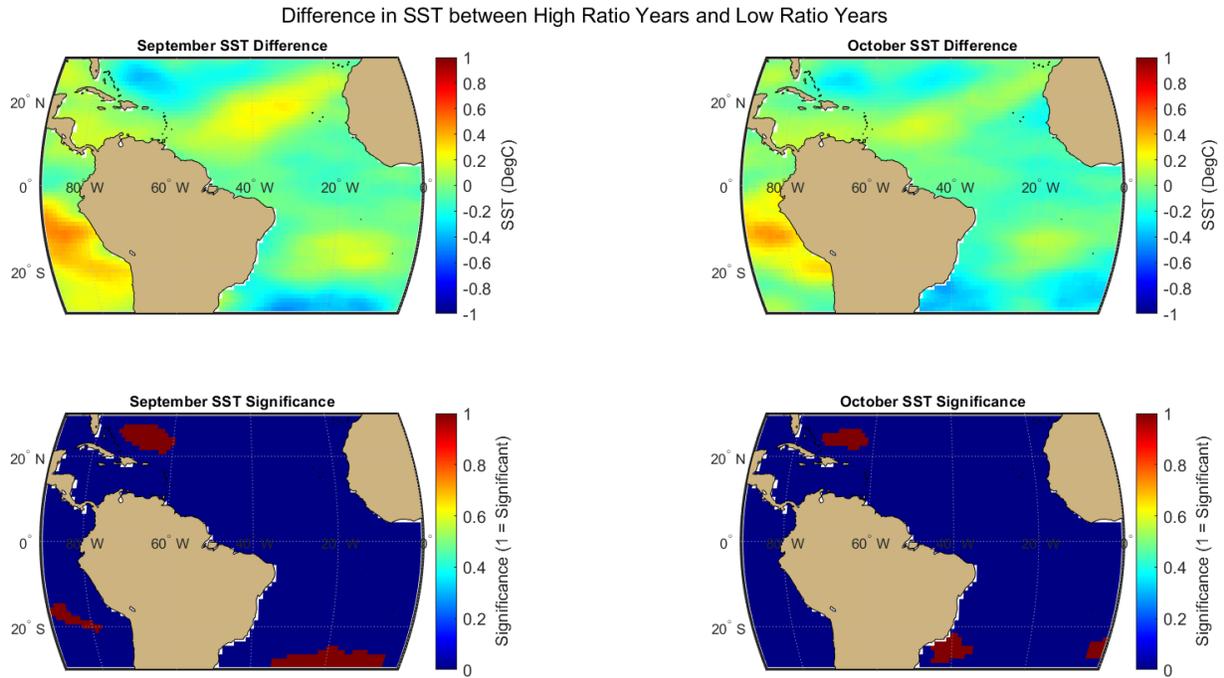


Figure 4.10. Atlantic Ocean SST Composite Analysis of the high pass ACE to Counts Ratio for September (left) and October (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

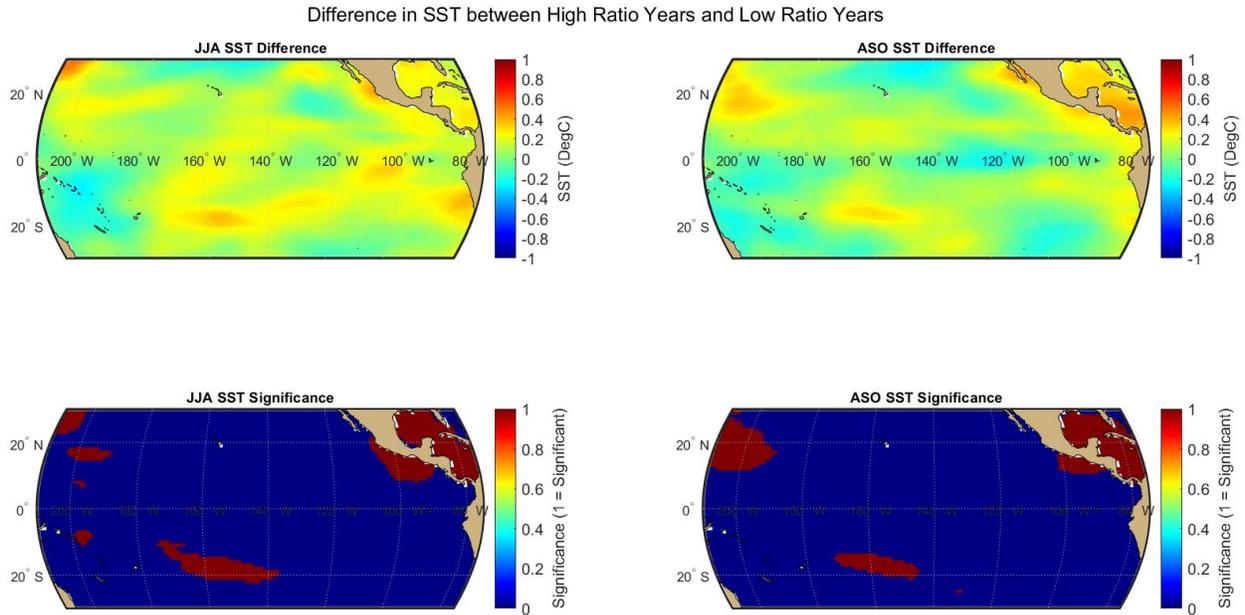


Figure 4.11. Pacific Ocean SST Composite Analysis of ACE to Counts Ratio for JJA (left) and ASO (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

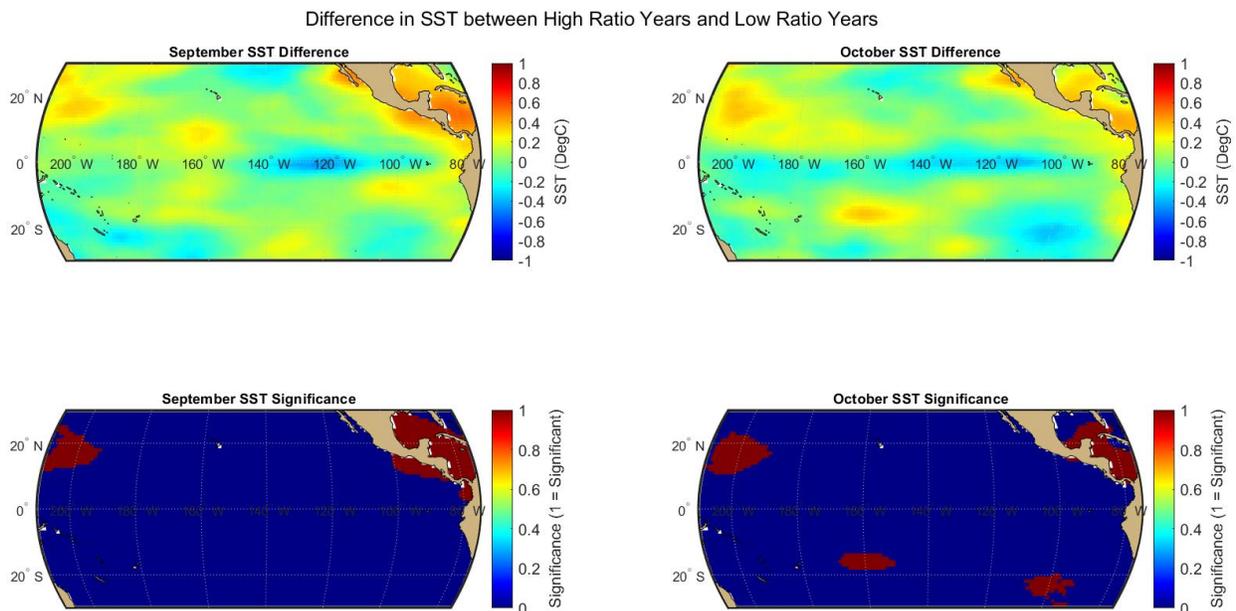


Figure 4.12. Pacific Ocean SST Composite Analysis of ACE to Counts Ratio for September (left) and October (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

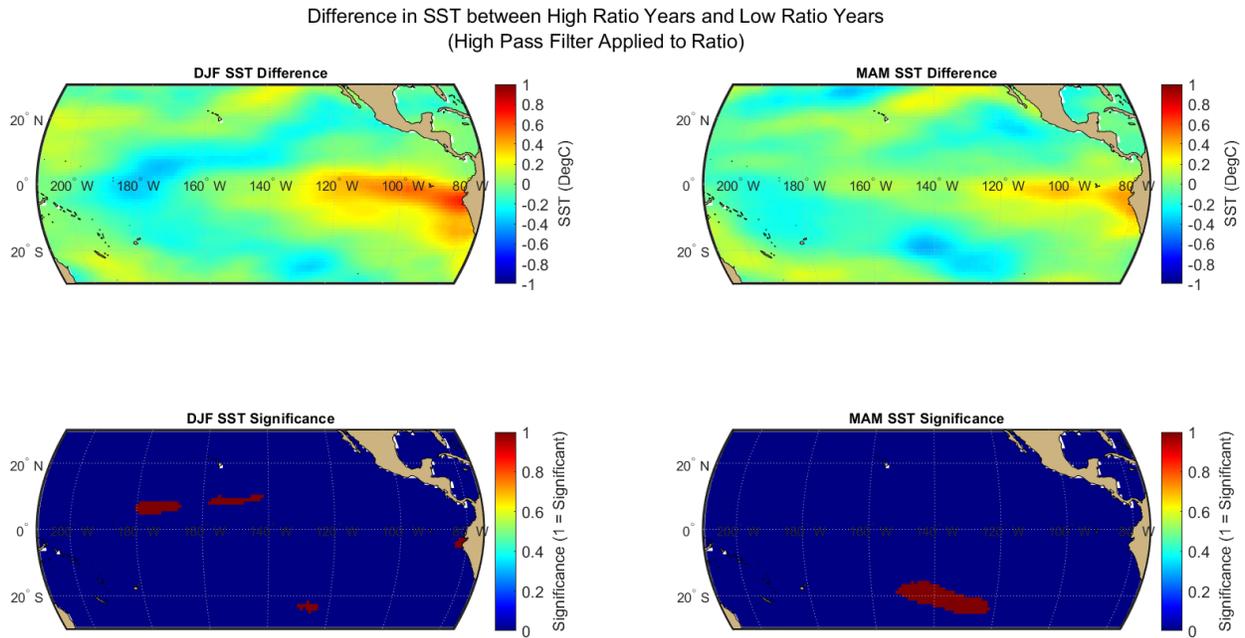


Figure 4.13. Pacific Ocean SST Composite Analysis of the high pass ACE to Counts Ratio for DJF (left) and MAM (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

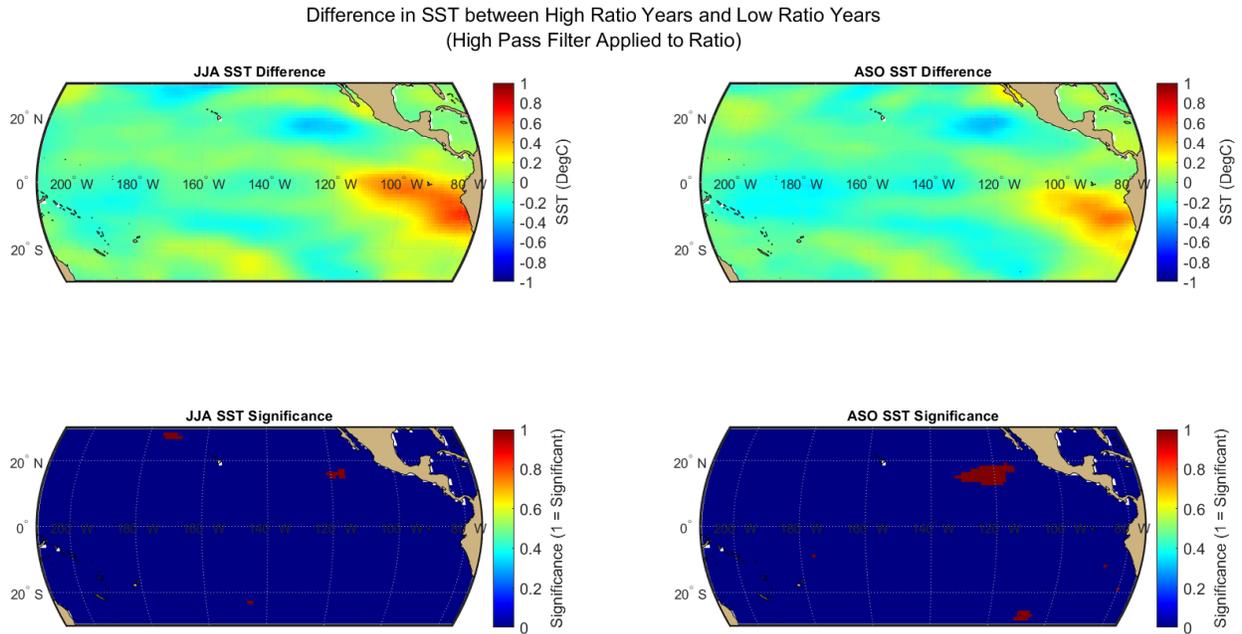


Figure 4.14. Pacific Ocean SST Composite Analysis of the high pass ACE to Counts Ratio for JJA (left) and ASO (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

Difference in SST between High Ratio Years and Low Ratio Years  
(High Pass Filter Applied to Ratio)

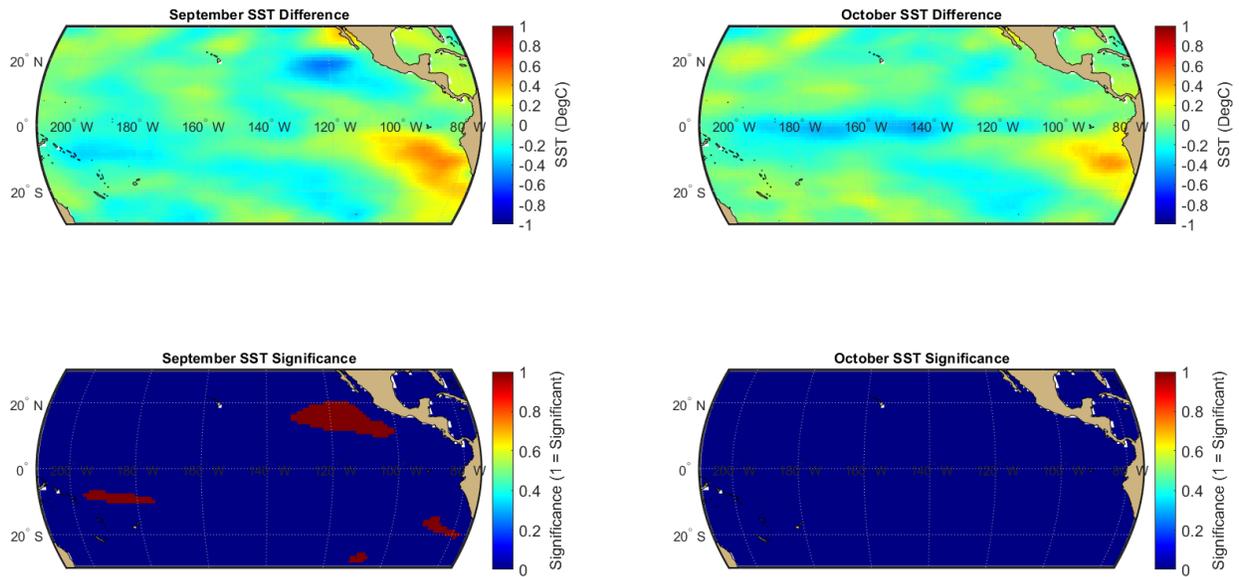


Figure 4.15. Pacific Ocean SST Composite Analysis of the high pass ACE to Counts Ratio for September (left) and October (right). SST differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

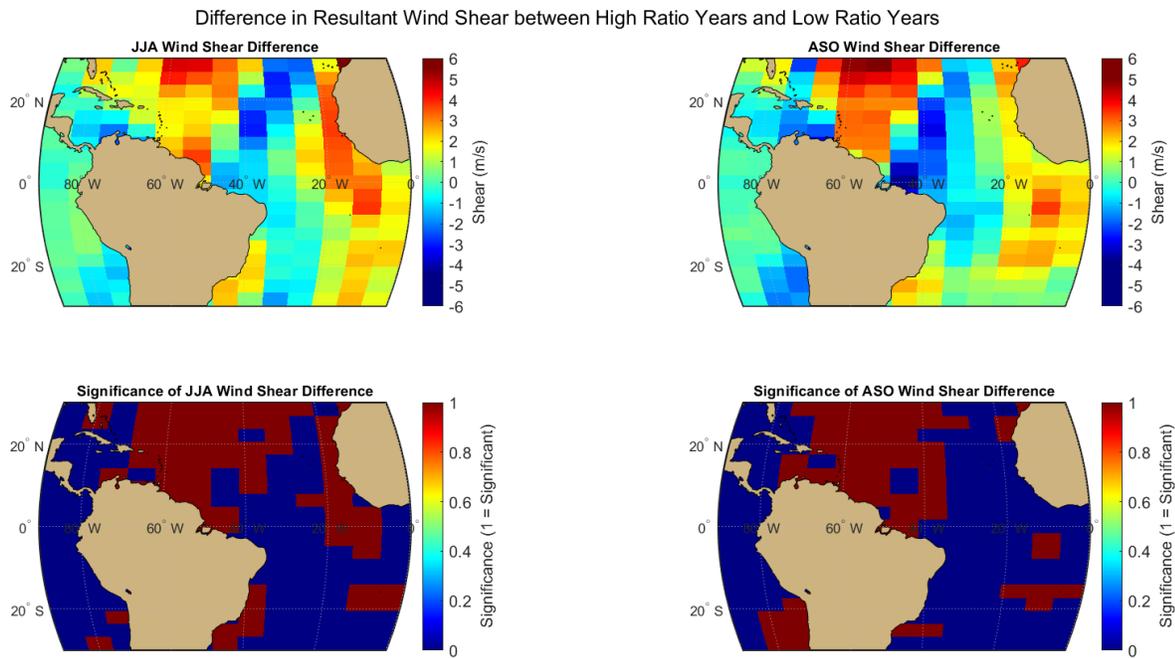


Figure 4.16. Atlantic Ocean VWS Composite Analysis of ACE to Counts Ratio for JJA (left) and ASO (right). VWS differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

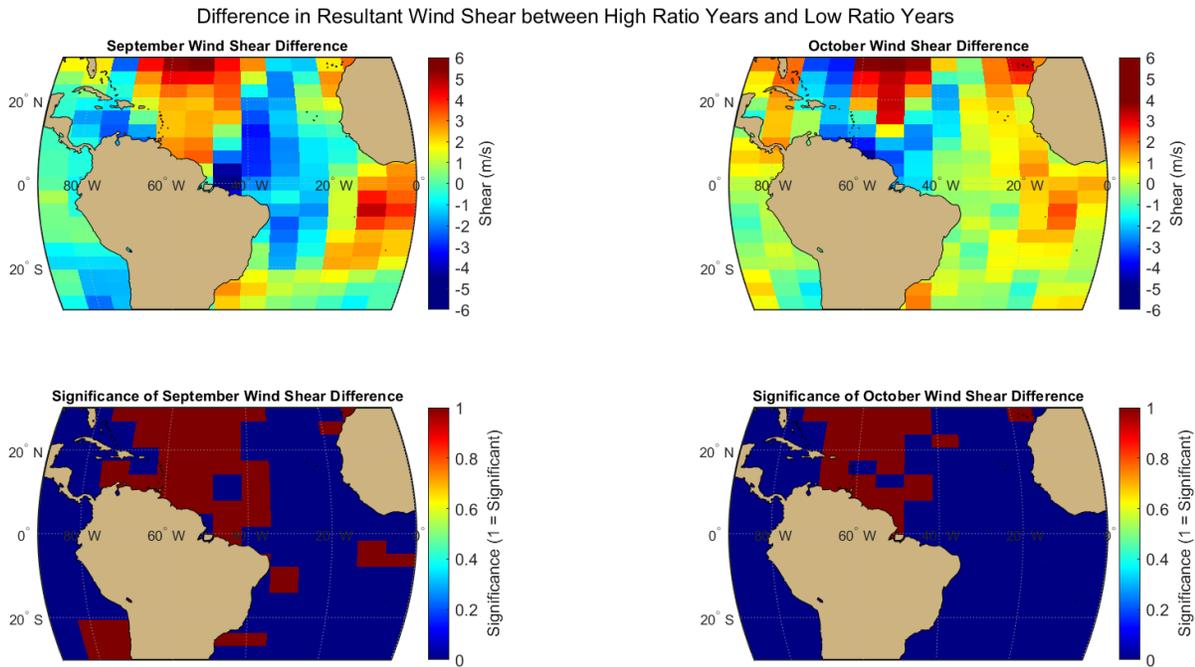


Figure 4.17. Atlantic Ocean VWS Composite Analysis of ACE to Counts Ratio for September (left) and October (right). VWS differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

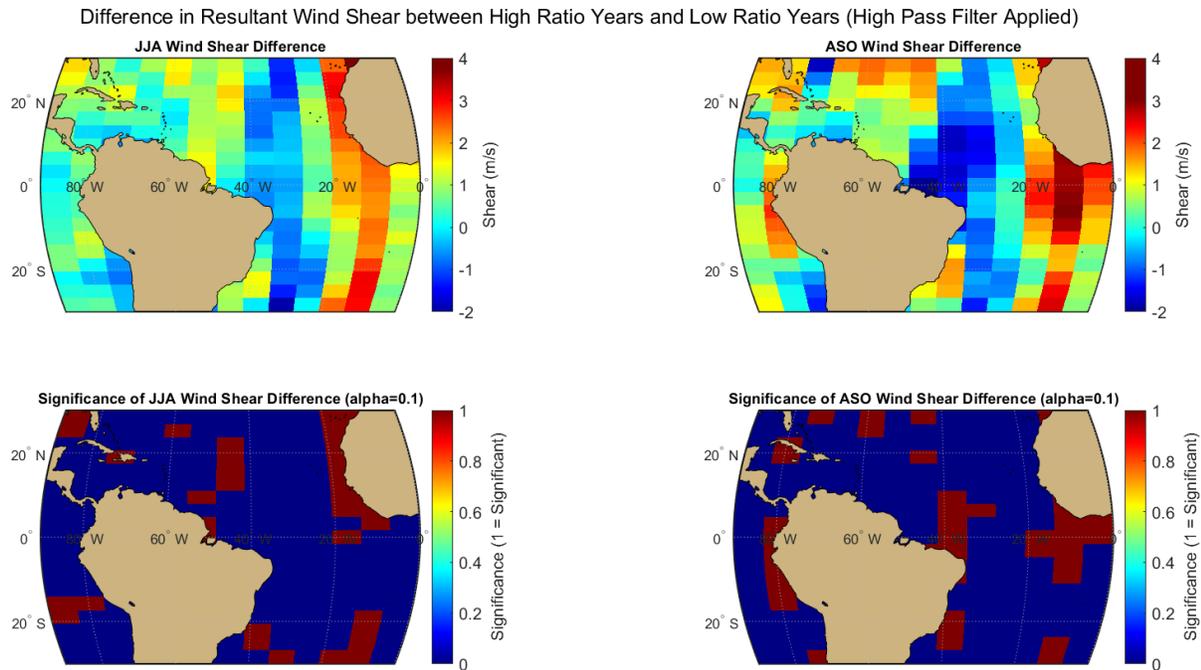


Figure 4.18. Atlantic Ocean VWS Composite Analysis of the high pass ACE to Counts Ratio for JJA (left) and ASO (right). VWS differences between high and low ratio years above, statistical significance (90% confidence level) below, with maroon areas being statistically significant.

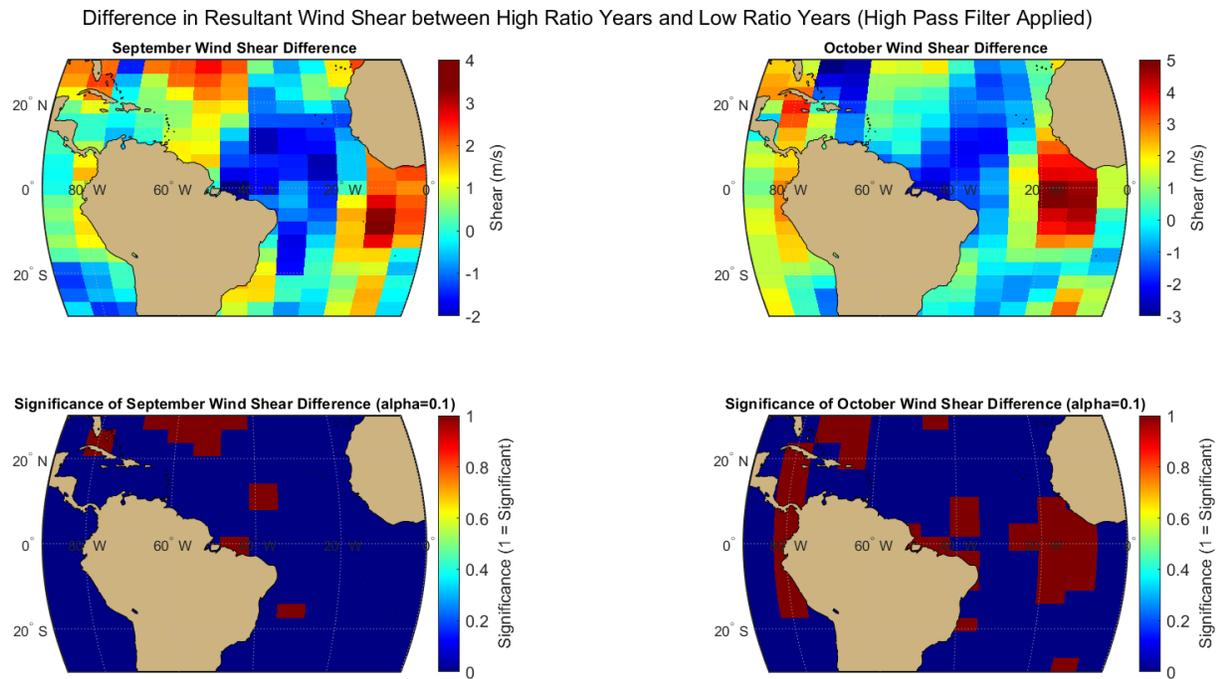


Figure 4.19. Atlantic Ocean VWS Composite Analysis of the high pass ACE to Counts Ratio for September (left) and October (right). VWS differences between high and low ratio years above, statistical significance (90% confidence level) below, with maroon areas being statistically significant.

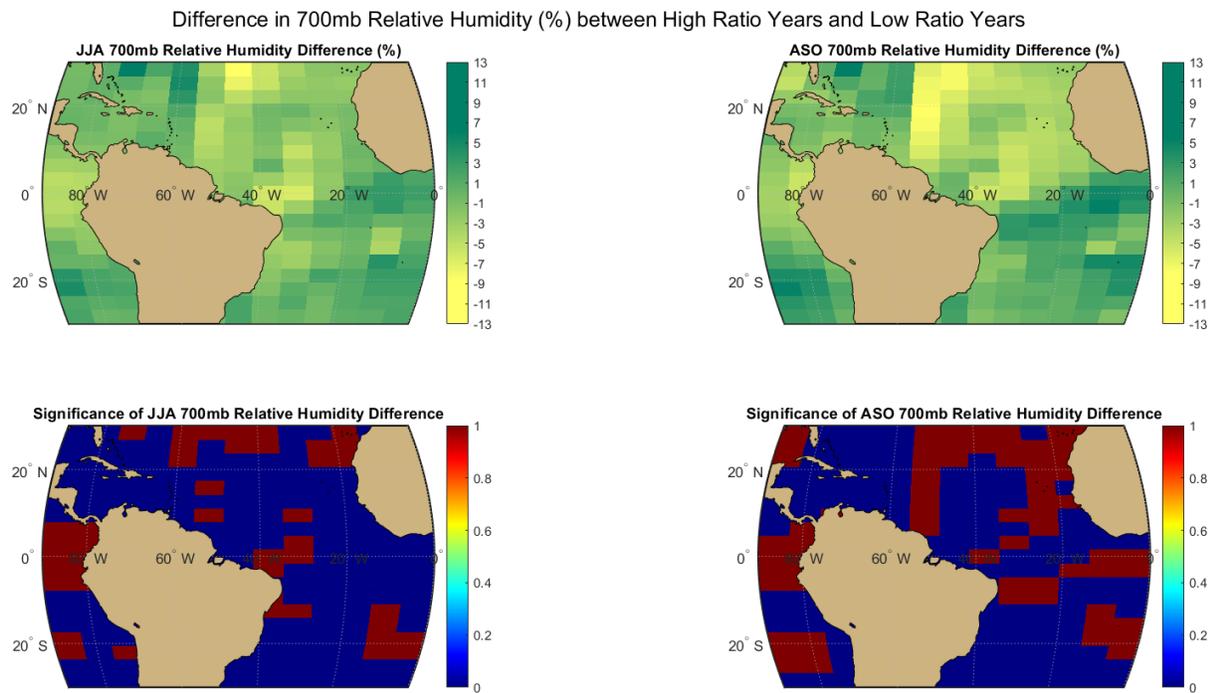


Figure 4.20. Atlantic Ocean RH700 Composite Analysis of ACE to Counts Ratio for JJA (left) and ASO (right). RH700 differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

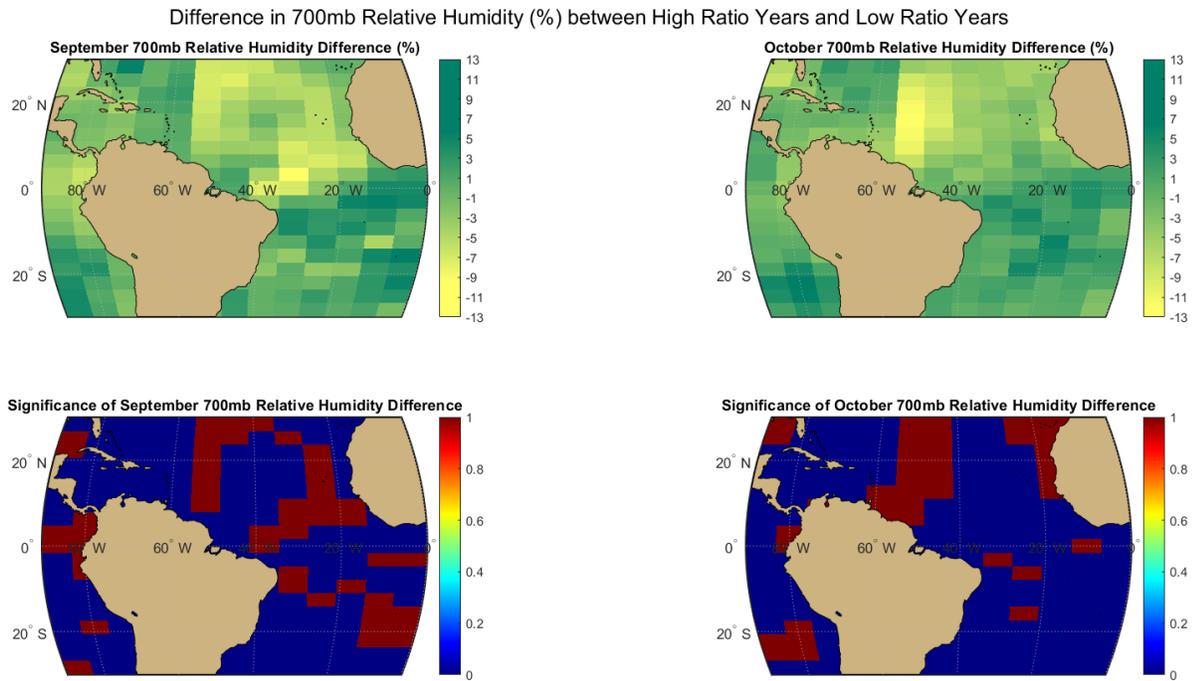


Figure 4.21. Atlantic Ocean RH700 Composite Analysis of ACE to Counts Ratio for September (left) and October (right). RH700 differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

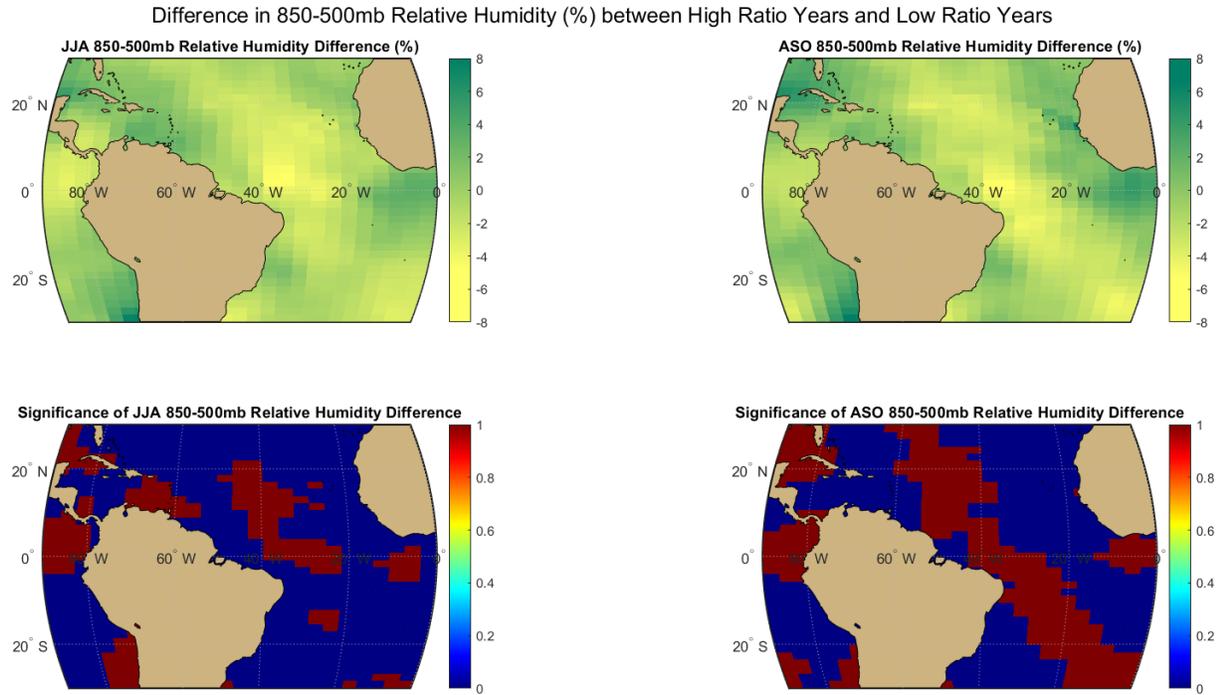


Figure 4.22. Atlantic Ocean RH850-500 Composite Analysis of ACE to Counts Ratio for JJA (left) and ASO (right). RH850-500 differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

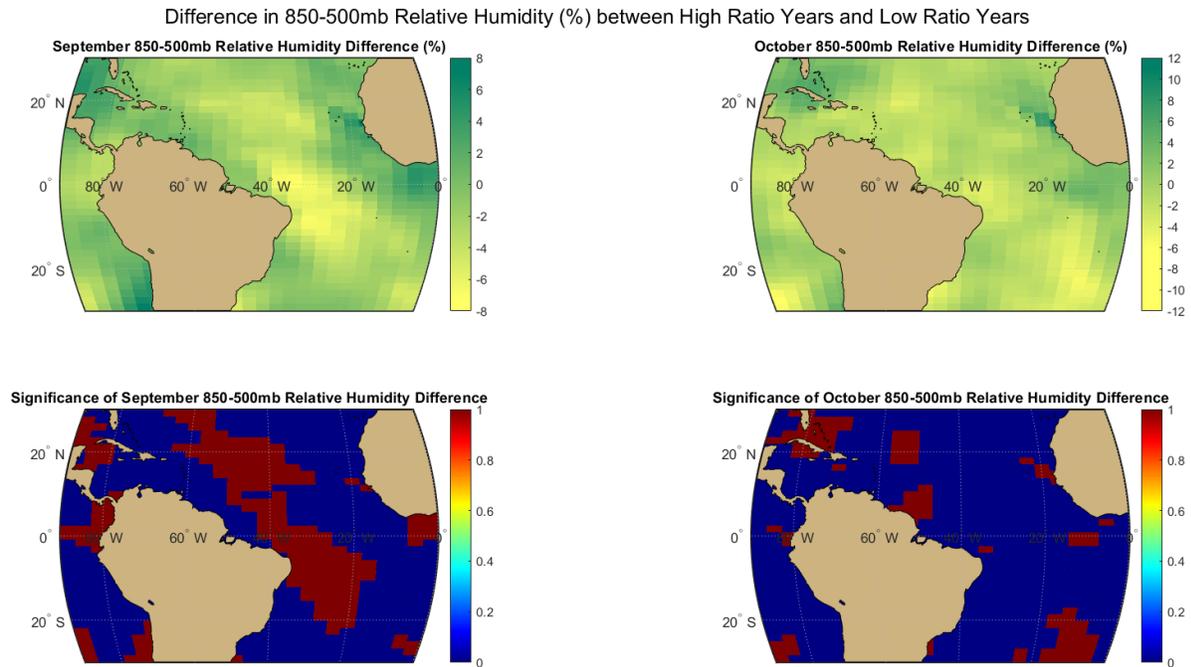


Figure 4.23. Atlantic Ocean RH850-500 Composite Analysis of ACE to Counts Ratio for September (left) and October (right). RH850-500 differences between high and low ratio years above, statistical significance (95% confidence level) below, with maroon areas being statistically significant.

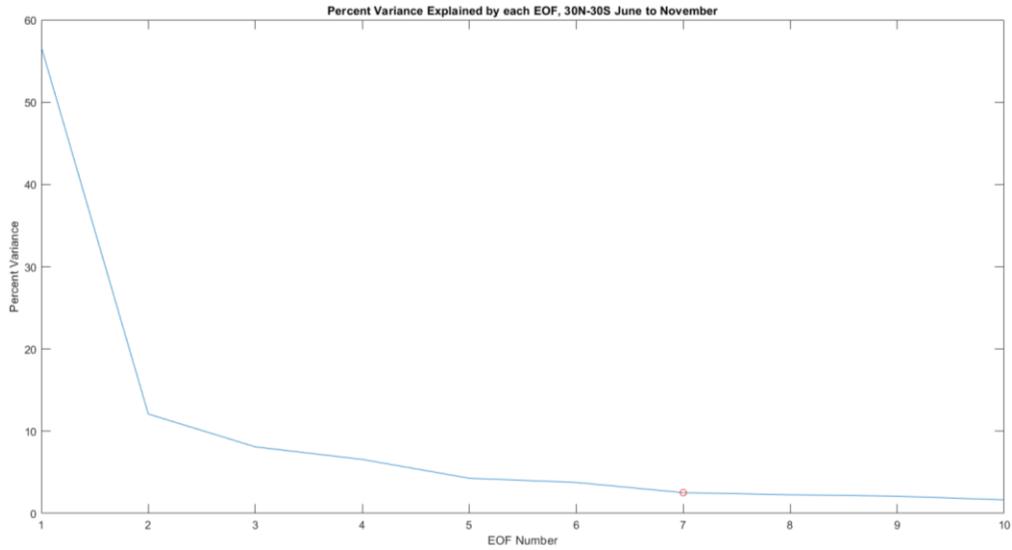


Figure 4.24. Percent variance explained by the top 10 monthly hurricane season VWS EOFs in the global domain. The truncation point is marked on the graph.

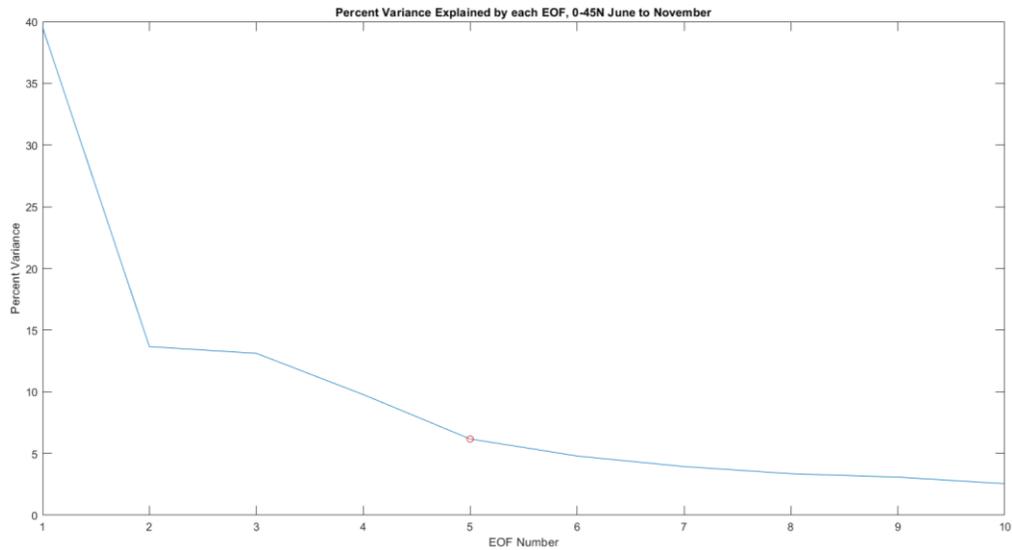


Figure 4.25. Percent variance explained by the top 10 monthly hurricane season VWS EOFs in the hemispheric domain. The truncation point is marked on the graph.

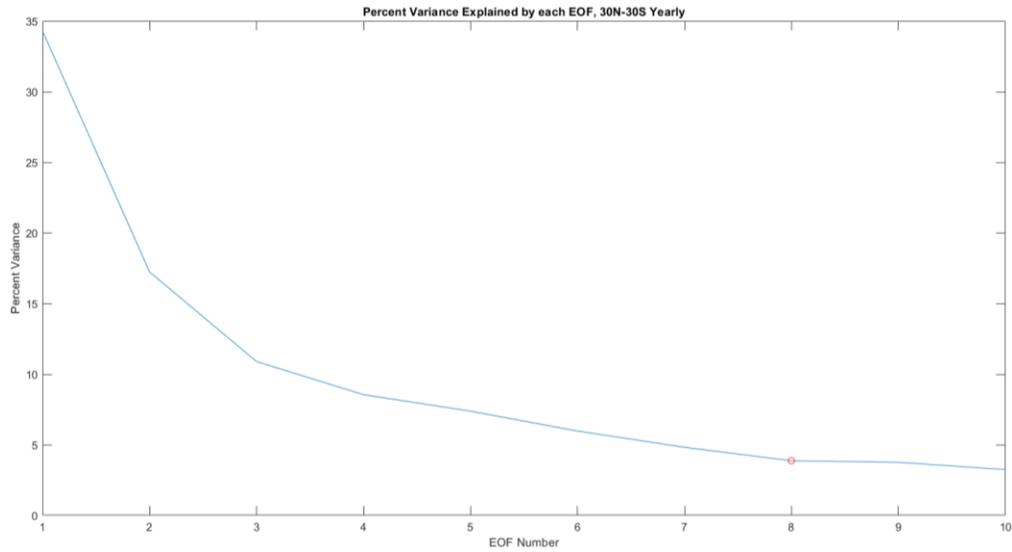


Figure 4.26. Percent variance explained by the top 10 yearly hurricane season average VWS EOFs in the global domain. The truncation point is marked on the graph.

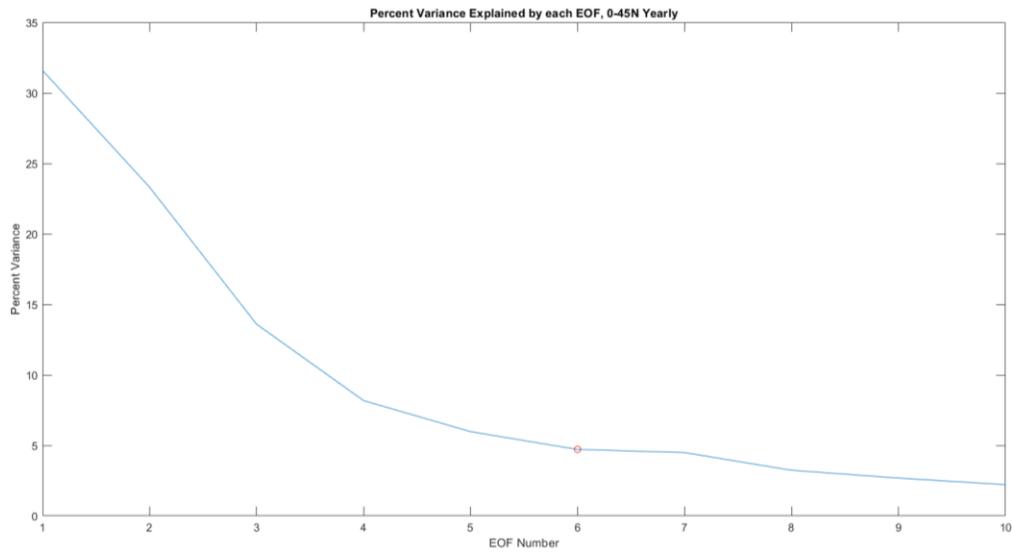


Figure 4.27. Percent variance explained by the top 10 yearly hurricane season average VWS EOFs in the hemispheric domain. The truncation point is marked on the graph.

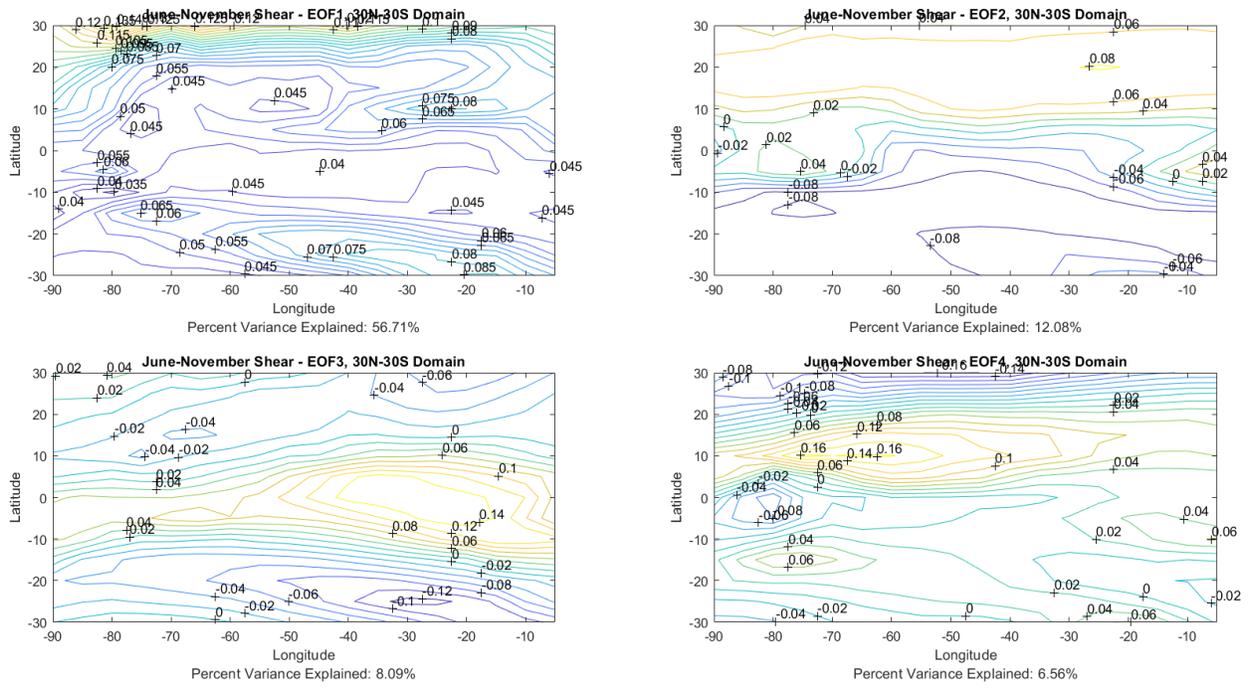


Figure 4.28. Spatial patterns of the top 4 monthly hurricane season VWS EOFs in the global domain. EOF1 (top left), EOF2 (top right), EOF3 (bottom left), and EOF4 (bottom right).

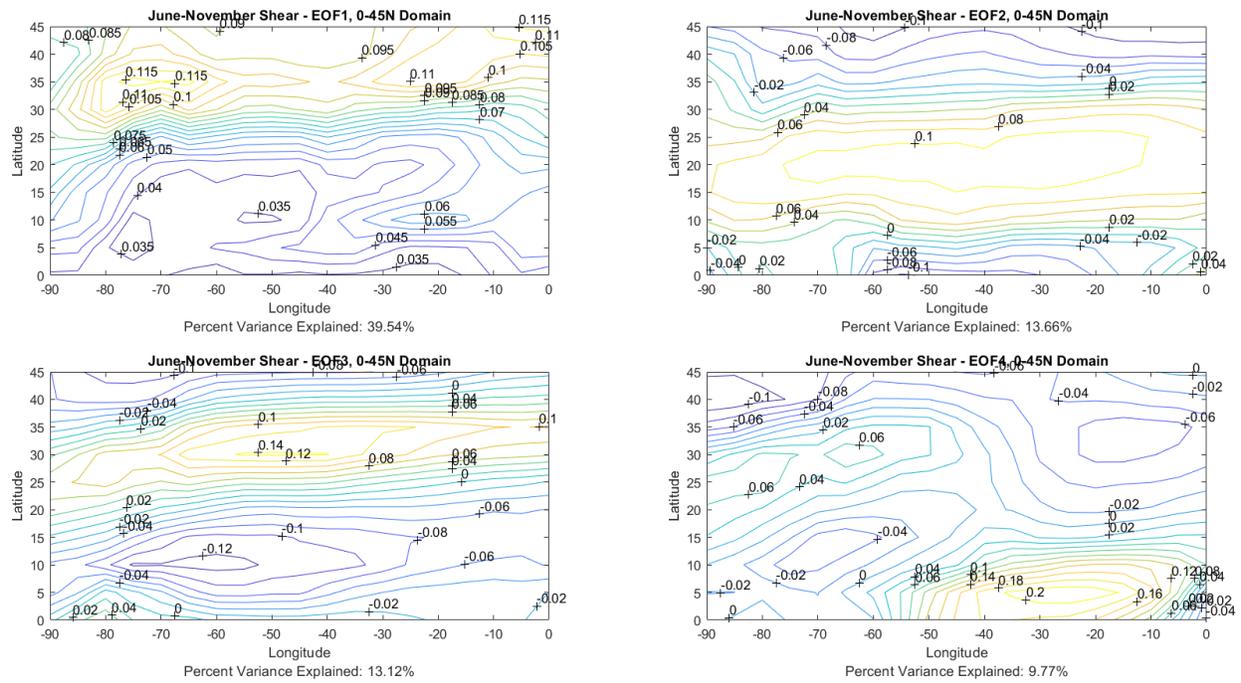


Figure 4.29. Spatial patterns of the top 4 monthly hurricane season VWS EOFs in the hemispheric domain. EOF1 (top left), EOF2 (top right), EOF3 (bottom left), and EOF4 (bottom right).

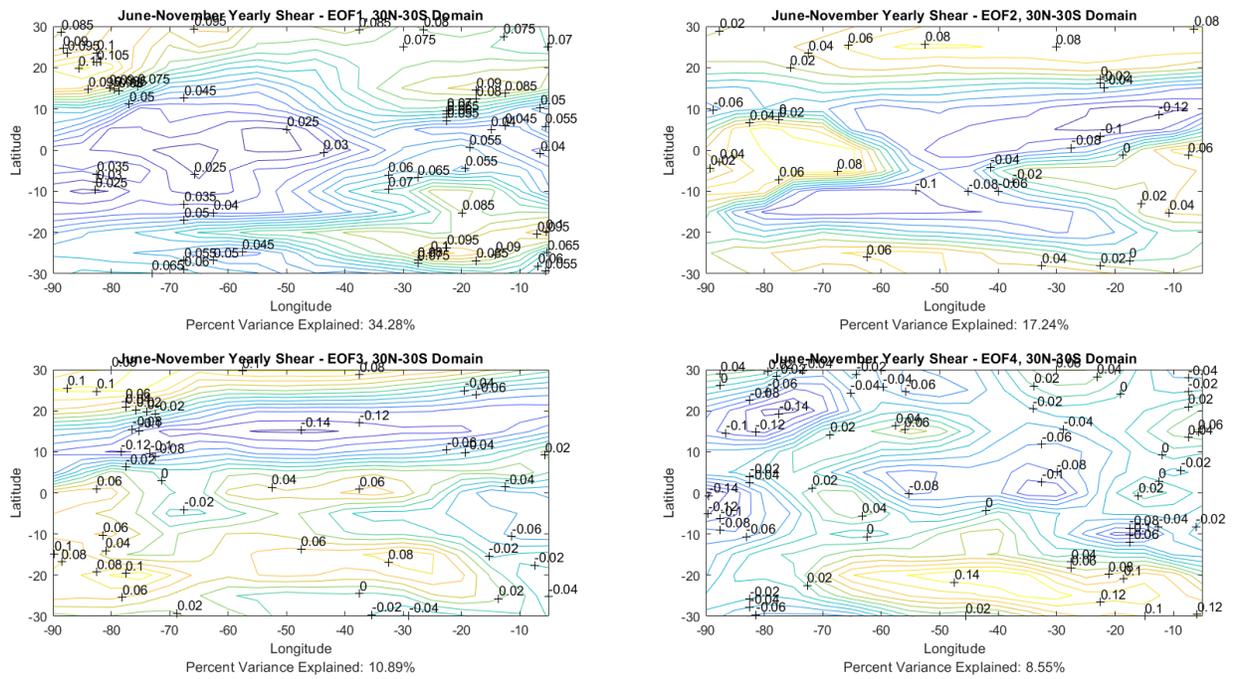


Figure 4.30. Spatial patterns of the top 4 yearly hurricane season average VWS EOFs in the global domain. EOF1 (top left), EOF2 (top right), EOF3 (bottom left), and EOF4 (bottom right).

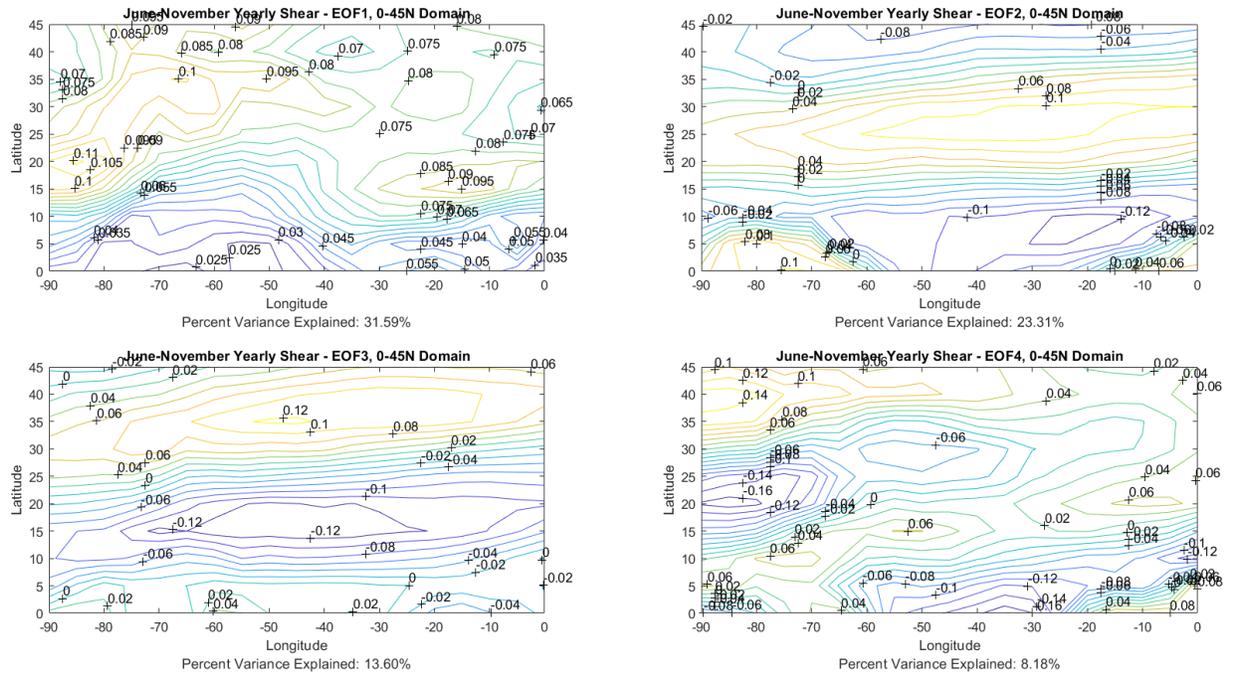


Figure 4.31. Spatial patterns of the top 4 yearly hurricane season average VWS EOFs in the hemispheric domain. EOF1 (top left), EOF2 (top right), EOF3 (bottom left), and EOF4 (bottom right).

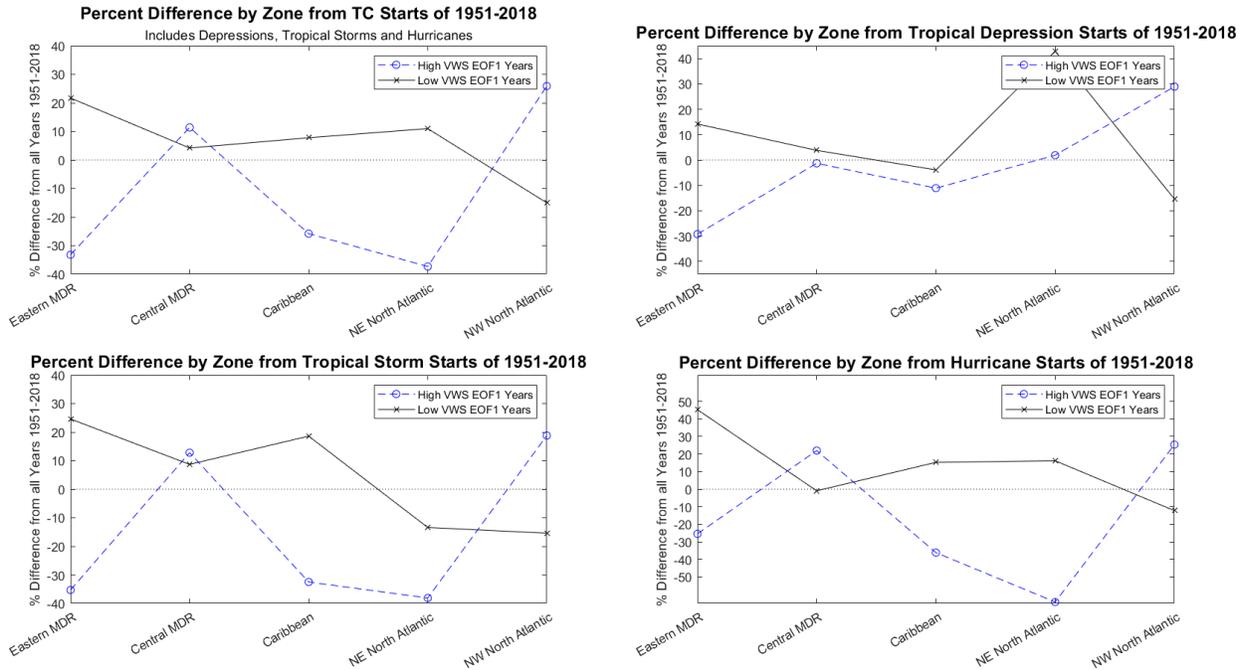


Figure 4.32. Percent difference of storm starts by zone between high/low VWS EOF1 years and all years from 1951-2018. High (low) VWS EOF1 years represented by dashed (solid) lines. Analysis was done for all TCs (top left), tropical depressions (top right), tropical storms (bottom left), and hurricanes (bottom right). Latitude/longitude bounds of each zone can be found in Figure 2.3.

## CHAPTER 5

### Conclusions

#### 5.1 Prediction Models

The results of the ACE classification prediction models show that using a combination of statistical and machine learning algorithms, classification accuracy of seasonal ACE can be improved. Almost all model runs show improvement over climatology, and most show improvement over the comparison organizations. Predictions of the ACE to Counts Ratio still have room to improve. Though most of the LASSO model runs outperform climatology, the KNN method only outperformed one of the three climatologies calculated.

#### 5.2 ACE to Counts Ratio

Based on the correlations performed on the VWS ECs, two primary conclusions can be drawn. They are described below and are also visually depicted in Figure 5.1. The first is that VWS EC2 can be considered the ENSO induced VWS, or the Atlantic VWS that's influenced by remote effects from the Pacific. EC2 correlates well with NINO12, NINO3 and NINO34. In addition, correlations performed with other indices indicate statistically significant correlations with the bivariate ENSO time series (CENSO) and the Western Hemisphere Warm Pool (WHWP) index in both domains. The first of the two indices is a function of Pacific SST, and the second is partially a function of Pacific SST. This component of VWS strongly modulates overall tropical activity in the Atlantic, as shown by its strong correlation with ACE. It also modulates the ratio, most likely by virtue of its modulation of ACE (since ACE is in the numerator of the ratio).

The second is that VWS EC1 is the locally induced wind shear, driven by effects in the Atlantic. For all of the indices studied so far, either the correlation with EC1 is not significant,

or the significance is domain dependent. Other indices that correlate strongly with EC1 in both domains include the Tropical Northern Atlantic (TNA) index and the Tropical Southern Atlantic (TSA) index, both functions of Atlantic SST. Finally, the yearly hurricane season average (HSA) EC1 correlates significantly with the ratio, but not ACE, indicating a direct influence on the ratio by EC1. The ratio is likely modulated on decadal and multidecadal timescales by the Atlantic effects, based on the results from the composite analysis. Both SST and VWS significantly affect the ratio, but the significance is diminished when the high pass filter is applied, suggesting the lower frequency influences.

Therefore, EC1 is related to the ratio through local effects of N. Atlantic wind shear and Atlantic SST on decadal and multidecadal timescales, whereas EC2 is related to the ratio through its modulation of ACE from remote, Pacific influences on interannual timescales. Since the Pacific influences related to ENSO are on a higher than decadal frequency, the ratio still correlates significantly with EC2 in both domains even after the high pass filter is applied. However, the ratio no longer correlates significantly with EC1 once the high pass filter is applied.

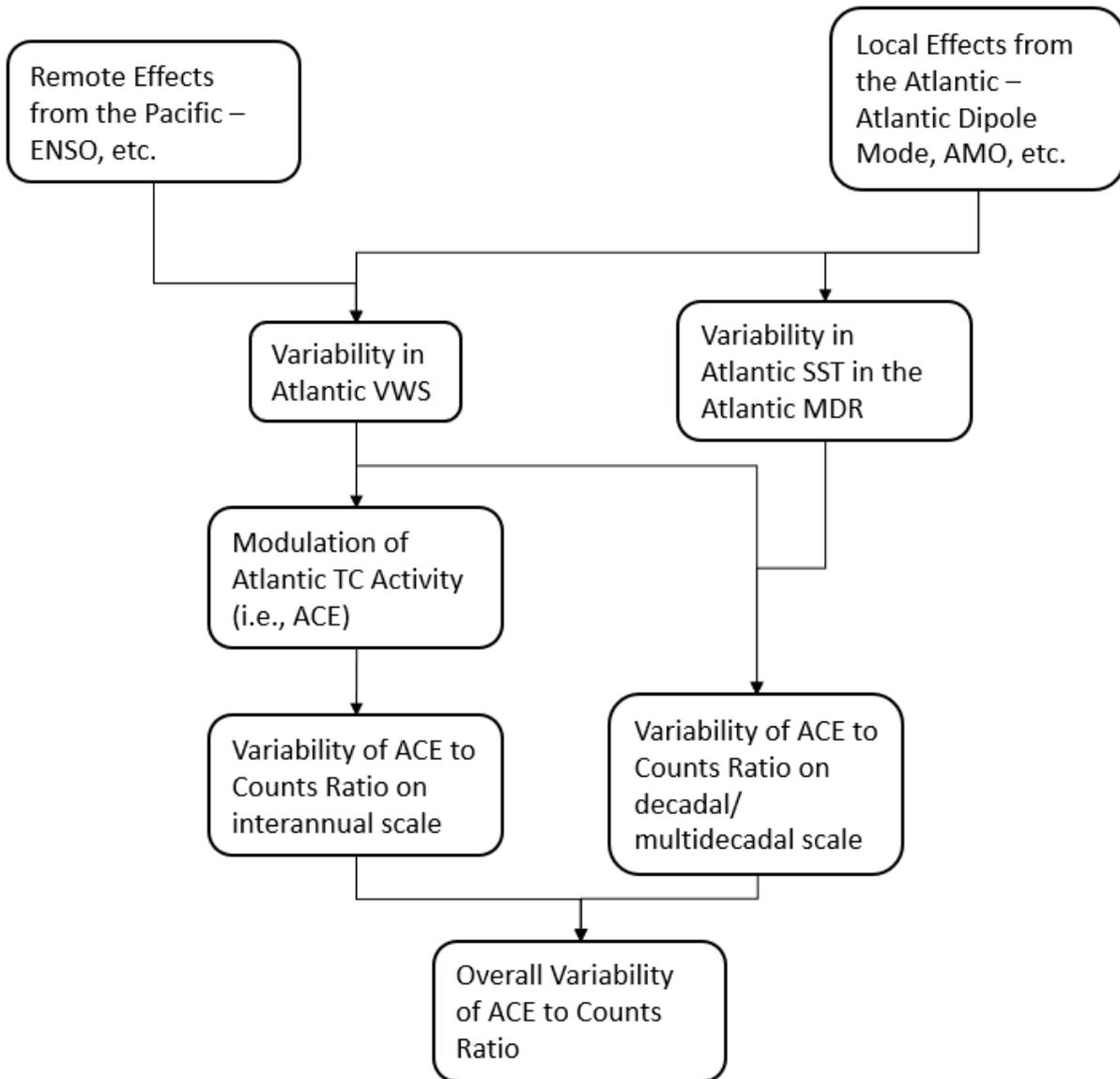


Figure 5.1. Flowchart of conclusions connecting remote effects from the Pacific and local effects from the Atlantic to the variability of the ACE to Counts Ratio.

## CHAPTER 6

### Future Work

In their review of organizations that perform seasonal hurricane predictions, Klotzbach et al. mentioned several areas as sources of improvement for TC modeling. New historical datasets, such as ERA5 and JRA55, are becoming available with more reliable historical upper-level data. North Atlantic hurricane seasons are continuously undergoing reanalysis, to provide more accurate historical TC data. Dynamical models are improving their ability to predict large scale oscillations such as ENSO. Finally, the spring predictability barrier is still an obstacle yet to be overcome (Klotzbach et al. 2019).

Another area that would benefit from more research that is not mentioned by Klotzbach et al. is the prediction of extreme hurricane seasons. Lea and Saunders (2006) discuss the accuracy of four prediction agencies for the extreme 2004 and 2005 seasons. The 2005 season was correctly predicted (ACE and counts) to be above average by TSR, CSU, NOAA, and the Meteorological Institute, Cuba. However, their count predictions were still 18-33% lower than the actual total of 27, and their ACE predictions 5-37% lower. The 2020 Atlantic Hurricane Season was a similar story, where by August, NOAA, CSU, and TSR predicted 25, 24 and 24 named storms, a 16-20% underprediction of the actual total of 30 (Temming 2020). The impact to the public from extreme hurricane seasons, especially those with landfalling major hurricanes, is large. The 2004 and 2005 hurricane seasons had a combined seven “intense” hurricane landfalls (maximum 1-minute sustained winds greater than 111mph) in the United States. 2005 saw the highest insured damage from a hurricane season at the time of US\$67 billion, breaking the record of US\$28 billion set in the previous year. The overall damage bills of the 2004 and 2005 seasons were US\$56 billion and US\$180 billion, respectively (Lea and Saunders 2006).

The ability to accurately predict extreme hurricane seasons will be increasingly important as time goes on.

Though the mid-tropospheric RH composites didn't show significant results, a link between mid-tropospheric moisture and the ratio could still be investigated in the future. This could be done using other proxies for moisture, such as entropy deficit. For the prediction models, the next iteration of the classification prediction models could involve direct classification predictions of ACE. This could be done using a tool such as decision tree analysis, used in Gao et al. 2016 and other works. In addition, prediction performance of the ratio could be improved. Though the midseason ratio models mostly outperformed climatology, a model using all available predictor months from January to December was not able to do so.

Additionally, model performance in predicting the ratio and other metrics of TC activity such as ACE or TC counts could be compared. The track density analysis could be expanded on as well by plotting the entire tracks of the TCs instead of just the starting points. Xie et al., 2005 indicate that the Atlantic SST dipole mode affects the track pattern, and that phenomenon may have some interaction with the ACE to Counts Ratio in the context of hurricane tracks.

## REFERENCES

- 2021: BACKGROUND INFORMATION: NORTH ATLANTIC HURRICANE SEASON. *National Weather Service Climate Prediction Center*,. <https://www.cpc.ncep.noaa.gov/products/outlooks/Background.html> (Accessed July 26, 2021).
- 2021: Facts + Statistics: Hurricanes. *iii.org*,. <https://www.iii.org/fact-statistic/facts-statistics-hurricanes> (Accessed May 9, 2022).
- Alfaro Córdoba, M., 2017: Variable Selection Methods with Applications to Atmospheric Sciences. North Carolina State University.
- Anderson, J., and J. Gyakum, 1989: A Diagnostic Study of Pacific Basin Circulation Regimes as Determined from Extratropical Cyclone Tracks. *Monthly Weather Review*, 117, 2672-2686, doi:10.1175/1520-0493(1989)117<2672:adsopb>2.0.co;2. [https://journals.ametsoc.org/view/journals/mwre/117/12/1520-0493\\_1989\\_117\\_2672\\_adsopb\\_2\\_0\\_co\\_2.xml?tab\\_body=pdf](https://journals.ametsoc.org/view/journals/mwre/117/12/1520-0493_1989_117_2672_adsopb_2_0_co_2.xml?tab_body=pdf) (Accessed July 1, 2022).
- Band, A., 2021: How to find the optimal value of K in KNN?. *Towards Data Science*,. <https://towardsdatascience.com/how-to-find-the-optimal-value-of-k-in-knn-35d936e554eb> (Accessed November 3, 2021).
- Bohnenstiehl, D., 2021: Fourier Transforms.
- Camargo, S., A. Robertson, A. Barnston, and M. Ghil, 2008: Clustering of eastern North Pacific tropical cyclone tracks: ENSO and MJO effects. *Geochemistry, Geophysics, Geosystems*, 9, n/a-n/a, doi:10.1029/2007gc001861. <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2007GC001861> (Accessed May 24, 2022).
- Chand, S., and K. Walsh, 2012: Modeling Seasonal Tropical Cyclone Activity in the Fiji Region as a Binary Classification Problem. *Journal of Climate*, 25, 5057-5071, doi:10.1175/jcli-d-11-00507.1. <https://journals.ametsoc.org/view/journals/clim/25/14/jcli-d-11-00507.1.xml> (Accessed May 24, 2022).
- Changnon, S., 2008: Characteristics of severe Atlantic hurricanes in the United States: 1949–2006. *Natural Hazards*, 48, 329-337, doi:10.1007/s11069-008-9265-z. <https://link.springer.com/article/10.1007/s11069-008-9265-z> (Accessed May 9, 2022).
- Chen, J., and S. Lin, 2013: Seasonal Predictions of Tropical Cyclones Using a 25-km-Resolution General Circulation Model. *Journal of Climate*, 26, 380-398, doi:10.1175/jcli-d-12-00061.1. [https://journals.ametsoc.org/view/journals/clim/26/2/jcli-d-12-00061.1.xml?tab\\_body=abstract-display](https://journals.ametsoc.org/view/journals/clim/26/2/jcli-d-12-00061.1.xml?tab_body=abstract-display) (Accessed May 24, 2022).
- Enfield D. B., Mestas-Nunez A. M., Trimble P. J., The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US. *Geophys. Res. Lett.* **28**, 2077–2080 (2001).

Gao, S., W. Zhang, J. Liu, I. Lin, L. Chiu, and K. Cao, 2016: Improvements in Typhoon Intensity Change Classification by Incorporating an Ocean Coupling Potential Intensity Index into Decision Trees. *Weather and Forecasting*, 31, 95-106, doi:10.1175/waf-d-15-0062.1. [https://journals.ametsoc.org/view/journals/wefo/31/1/waf-d-15-0062\\_1.xml?tab\\_body=pdf](https://journals.ametsoc.org/view/journals/wefo/31/1/waf-d-15-0062_1.xml?tab_body=pdf) (Accessed November 17, 2021).

Goldenberg, S. B., C. W. Landsea, A. M. Mestas-Nuñez, and W. M. Gray (2001), The recent increase in Atlantic hurricane activity: Causes and implications, *Science*, **293**, 474–479.

Goldenberg, S., C. Landsea, A. Mestas-Nuñez, and W. Gray, 2001: The Recent Increase in Atlantic Hurricane Activity: Causes and Implications. *Science*, 293, 474-479, doi:10.1126/science.1060040. <https://www.science.org/doi/10.1126/science.1060040> (Accessed July 11, 2022).

Goldenberg, S., and L. Shapiro, 1996: Physical Mechanisms for the Association of El Niño and West African Rainfall with Atlantic Major Hurricane Activity. *Journal of Climate*, 9, 1169-1187, doi:10.1175/1520-0442(1996)009<1169:pmftao>2.0.co;2. [https://journals.ametsoc.org/view/journals/clim/9/6/1520-0442\\_1996\\_009\\_1169\\_pmftao\\_2\\_0\\_co\\_2.xml?tab\\_body=pdf](https://journals.ametsoc.org/view/journals/clim/9/6/1520-0442_1996_009_1169_pmftao_2_0_co_2.xml?tab_body=pdf) (Accessed May 24, 2022).

Gray, W., 1988: Forecast of Atlantic Seasonal Hurricane Activity For 1988. Colorado State University, Fort Collins, CO,.

Kalnay, E. et al. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* 77, 437–471 (1996).

Keith, E., and L. Xie, 2009: Predicting Atlantic Tropical Cyclone Seasonal Activity in April. *Weather and Forecasting*, 24, 436-455, doi:10.1175/2008waf2222139.1. [https://journals.ametsoc.org/view/journals/wefo/24/2/2008waf2222139\\_1.xml](https://journals.ametsoc.org/view/journals/wefo/24/2/2008waf2222139_1.xml) (Accessed July 6, 2022).

Klotzbach, P. et al., 2019: Seasonal Tropical Cyclone Forecasting. *Tropical Cyclone Research and Review*, 8, 134-149, doi:10.1016/j.tcr.2019.10.003. <https://www.sciencedirect.com/science/article/pii/S2225603219300700>.

Knight, J., 2005: A signature of persistent natural thermohaline circulation cycles in observed climate. *Geophysical Research Letters*, 32, doi:10.1029/2005gl024233. <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2005GL024233> (Accessed May 24, 2022).

Knight, J., C. Folland, and A. Scaife, 2006: Climate impacts of the Atlantic Multidecadal Oscillation. *Geophysical Research Letters*, 33, doi:10.1029/2006gl026242. <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2006GL026242> (Accessed May 24, 2022).

- Landsea, C. W. and J. L. Franklin, 2013: Atlantic Hurricane Database Uncertainty and Presentation of a New Database Format. *Monthly Weather Review*, 141, 3576-3592.
- Lea, A., and M. Saunders, 2006: How well forecast were the 2004 and 2005 Atlantic and US hurricane seasons?. *Weather*, 61, 245-249, doi:10.1256/wea.95.06. <https://www.tropicalstormrisk.com/docs/LeaSaunders2006.pdf>.
- Maue, R., 2009: Northern Hemisphere tropical cyclone activity. *Geophysical Research Letters*, 36, doi:10.1029/2008gl035946. <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2008GL035946> (Accessed May 24, 2022).
- NOAA Physical Sciences Laboratory. 2021. *Climate Indices: Monthly Atmospheric and Ocean Time Series*. [online] Available at: <<https://psl.noaa.gov/data/climateindices/list/>> [Accessed 10 June 2021].
- Patricola, C., R. Saravanan, and P. Chang, 2014: The Impact of the El Niño–Southern Oscillation and Atlantic Meridional Mode on Seasonal Atlantic Tropical Cyclone Activity. *Journal of Climate*, 27, 5311-5328, doi:10.1175/jcli-d-13-00687.1. [https://journals.ametsoc.org/view/journals/clim/27/14/jcli-d-13-00687.1.xml?tab\\_body=pdf](https://journals.ametsoc.org/view/journals/clim/27/14/jcli-d-13-00687.1.xml?tab_body=pdf) (Accessed July 8, 2022).
- Ryan, M., 2022: Decomposing Signal Using Empirical Mode Decomposition. Towards Data Science,. <https://towardsdatascience.com/decomposing-signal-using-empirical-mode-decomposition-algorithm-explanation-for-dummy-93a93304c541> (Accessed May 24, 2022).
- Servain, 1991: Simple climate indices for the tropical Atlantic Ocean and some applications, *J. Geophys. Res.*, 96, 15,137– 15,146.
- Smith, A., 2021: 2020 U.S. billion-dollar weather and climate disasters in historical context. NOAA Climate.gov., <https://www.climate.gov/disasters2020> (Accessed May 9, 2022).
- Subramanian, D., 2019: A Simple Introduction to K-Nearest Neighbors Algorithm. *Towards Data Science*,. <https://towardsdatascience.com/a-simple-introduction-to-k-nearest-neighbors-algorithm-b3519ed98e> (Accessed May 6, 2021).
- Sun, X., L. Xie, S. Shah, and X. Shen, 2021: A Machine Learning Based Ensemble Forecasting Optimization Algorithm for Preseason Prediction of Atlantic Hurricane Activity. *Atmosphere*, 12, 522, doi:10.3390/atmos12040522. <https://www.mdpi.com/2073-4433/12/4/522>.
- Temming, M., 2020: Predictions for the 2020 Atlantic hurricane season just got worse. *Science News*,. <https://www.sciencenews.org/article/predictions-2020-atlantic-hurricane-storm-season-worse> (Accessed June 16, 2021).
- Torrence, C. and G. P. Compo, 1998: A Practical Guide to Wavelet Analysis. *Bull. Amer. Meteor. Soc.*, 79, 61-78.

Vecchi, G., and T. Knutson, 2008: On Estimates of Historical North Atlantic Tropical Cyclone Activity. *Journal of Climate*, 21, 3580–3600, doi:10.1175/2008jcli2178.1. <https://journals.ametsoc.org/view/journals/clim/21/14/2008jcli2178.1.xml>.

Vecchi, G., M. Zhao, H. Wang, G. Villarini, A. Rosati, A. Kumar, I. Held, and R. Gudgel, 2011: Statistical–Dynamical Predictions of Seasonal North Atlantic Hurricane Activity. *Monthly Weather Review*, 139, 1070–1082, doi:10.1175/2010mwr3499.1. <https://journals.ametsoc.org/view/journals/mwre/139/4/2010mwr3499.1.xml> (Accessed May 24, 2022).

Vitart, F., J. Anderson, and W. Stern, 1997: Simulation of Interannual Variability of Tropical Storm Frequency in an Ensemble of GCM Integrations. *Journal of Climate*, 10, 745–760, doi:10.1175/1520-0442(1997)010<0745:soivot>2.0.co;2. [https://journals.ametsoc.org/view/journals/clim/10/4/1520-0442\\_1997\\_010\\_0745\\_soivot\\_2.0.co\\_2.xml?tab\\_body=pdf](https://journals.ametsoc.org/view/journals/clim/10/4/1520-0442_1997_010_0745_soivot_2.0.co_2.xml?tab_body=pdf) (Accessed July 11, 2022).

Vitart, F., and T. Stockdale, 2001: Seasonal Forecasting of Tropical Storms Using Coupled GCM Integrations. *Monthly Weather Review*, 129, 2521–2537, doi:10.1175/1520-0493(2001)129<2521:sfotsu>2.0.co;2. [https://journals.ametsoc.org/view/journals/mwre/129/10/1520-0493\\_2001\\_129\\_2521\\_sfotsu\\_2.0.co\\_2.xml](https://journals.ametsoc.org/view/journals/mwre/129/10/1520-0493_2001_129_2521_sfotsu_2.0.co_2.xml) (Accessed July 11, 2022).

Waple, A. M., Lawrimore, J. H., Halpert, M. S., Bell, G. D., Higgins, W., Lyon, B., Menne, M. J., Gleason, K. L., Schnell, R. C., Christy, J. R., Thiaw, W., Wright, W. J., Salinger, M. J., Alexander, L., Stone, R. S., and Camargo, S. J., 2002: CLIMATE ASSESSMENT FOR 2001. *Bulletin of the American Meteorological Society*, 83(6), S1–S62. <http://www.jstor.org/stable/26215369>

Xie, K., and B. Liu, 2014: An ENSO-Forecast Independent Statistical Model for the Prediction of Annual Atlantic Tropical Cyclone Frequency in April. *Advances in Meteorology*, 2014, 1–11, doi:10.1155/2014/248148. [https://www.researchgate.net/publication/279163540\\_An\\_ENSO-Forecast\\_Independent\\_Statistical\\_Model\\_for\\_the\\_Prediction\\_of\\_Annual\\_Atlantic\\_Tropical\\_Cyclone\\_Frequency\\_in\\_April](https://www.researchgate.net/publication/279163540_An_ENSO-Forecast_Independent_Statistical_Model_for_the_Prediction_of_Annual_Atlantic_Tropical_Cyclone_Frequency_in_April) (Accessed July 7, 2022).

Xie, L., T. Yan, L. Pietrafesa, J. Morrison, and T. Karl, 2005: Climatology and Interannual Variability of North Atlantic Hurricane Tracks. *Journal of Climate*, 18, 5370–5381, doi:10.1175/jcli3560.1. <https://journals.ametsoc.org/view/journals/clim/18/24/jcli3560.1.xml> (Accessed July 9, 2022).

Xie, L., T. Yan, and L. Pietrafesa, 2005: The effect of Atlantic sea surface temperature dipole mode on hurricanes: Implications for the 2004 Atlantic hurricane season. *Geophysical Research Letters*, 32, doi:10.1029/2004gl021702. [https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2004GL021702?casa\\_token=wDi9ydVdWZsAAAAA%3AOCzVWEdYaNsCGkvPqieYqsRnENWaEXu-GfLQZo186K-ZbQRG9BBhEvS8AoWkSvEEIk7b2l6bjwG2fQ](https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2004GL021702?casa_token=wDi9ydVdWZsAAAAA%3AOCzVWEdYaNsCGkvPqieYqsRnENWaEXu-GfLQZo186K-ZbQRG9BBhEvS8AoWkSvEEIk7b2l6bjwG2fQ) (Accessed May 26, 2022).

Xie, L., 2021: Forecasting Atlantic hurricane activities from a season to a year in advance: the magical power and the limitations of Machine Learning.

Xie, L., 2021: Lecture 3: Fourier Transform.

Xie, L., 2021: Lecture 4: Wavelet Analysis.

Xie, L., 2021: Lecture 7: Regression Analysis (II).

Xie, L., 2021: Lecture 9: EOF/PC Analysis.

Yan, T., L. Pietrafesa, D. Dickey, P. Gayes, and S. Bao, 2014: Seasonal prediction of landfalling hurricanes along Eastern Seaboard of the United States. *International Journal of Climatology*, 35, 2647-2653, doi:10.1002/joc.4163. [https://rmets.onlinelibrary.wiley.com/doi/full/10.1002/joc.4163?casa\\_token=56Egxqro9xoAAAAA%3A3xteB-hC5aPjRo1xOyDoXdR2-ESNPkS7gXmNasWXz5CqAUoXrmjeWeQmyEl3SadS2FpCZJ-a7fbLzw](https://rmets.onlinelibrary.wiley.com/doi/full/10.1002/joc.4163?casa_token=56Egxqro9xoAAAAA%3A3xteB-hC5aPjRo1xOyDoXdR2-ESNPkS7gXmNasWXz5CqAUoXrmjeWeQmyEl3SadS2FpCZJ-a7fbLzw) (Accessed May 24, 2022).