Hurricane Analysis and Visualization Using R

Romane Goldmuntz, Vy Tran, and Jianqiong Zhan

2019-12-13

Contents

Chapter 1

Introduction

Hurricanes have caused tremendous losses on the country's economy (Weinkle et al. (2018)): they currently cost the government over $28 billion each year. Besides the government, several industries are heavily impacted by hurricanes, including the insurance industry. For example, according to Bloomberg, hurricane Dorian caused the insurance industry losses of up $25 billion, making it the most expensive natural disaster for the industry since 2017's Hurricane Maria (D'Souza (2019)). Therefore, the United States' Sustainable Development has identified economic losses from hurricane as a key indicator for the economic growth (SDG-Indicators (2018)). Disentangling the relative roles of variability and changes in Hurricane could contribute information useful to sustainable development goals.

Key frequently asked questions are: how climate affects hurricane, and what of the future hurricanes are more likely to be. To gain insight on these questions, we first have to look into historical records to discern the variability and changes of hurricanes, the analysis of which will help us to build predictive understanding of the influencing of climate on hurricane. Our working hypothesis is that: if greenhouse warming has caused a substantial increase in Atlantic hurricane activity, then the century scale increase in Atlantic hurricane since 1851 should have produced a long-term rise in measures of Atlantic hurricanes activity, similar to that seen for global temperature, for example. Therefore, we aim to dig into the original storm track file with more than 150 years of records of Atlantic tropical storm activities in order to examine variability and changes of Atlantic hurricane under global warming conditions. Specifically, three important questions we would like to address in this analysis are:

(1) What influences the change in category of hurricanes?

(2) Has hurricane activity shifted north-ward?

(3) Have humans already caused a detectable increase in Atlantic hurricane activity?

Chapter 2

Data Source

The storm track data presented in the analysis is downloaded from National Hurricane Center and Central Pacific Hurricane Center, which is also known as the Atlantic hurricane database [HURDAT2: 1851-2018] (https://www.nhc.noaa.gov/data/hurdat/hurdat2-1851-2018-120319.txt) and the HURDAT data is downloaded from the [NOAA's Atlantic Oceanographic and Meteorological Laboratory] (https://www.aoml.noaa.gov/).

We choose to work with these data because they provide us with updated, complete and accurate information about hurricanes over the past 150 years (1851 - 2018). Since the Atlantic hurricane activity has shown very strong year-to-year and decade-to-decade variability, longer records or hurricanes are much needed.

One important thing to note about the names of hurricanes is that there are six names lists that have been used in rotation and re-cycled every six years. This means that there can be more than one hurricanes having the same name but happened in different years. We have tackled this problem by giving each hurricane a unique `storm-id`.

Chapter 3

Data Transformation

The (HURDAT2) data has a comma-delimited, text format with six-hourly information on the location, maximum winds, central pressure, and (beginning in 2004) size of all known tropical cyclones and subtropical cyclones. The dataset is a combination of serveral subsets. Each subset is used for a storm track record which includes header information and values. please refer to this file for detail information.

Firstly, we extract storm `id`, `name`, and `subtext length` from each subtext header, then read in data according to each `subtext length`, and merge data subset by indexing it with storm `id`, `name`. In the original text, you will find that the `name` is non-unique. Currently, there are six lists that are used in rotation and re-cycled every six years, i.e., the 2013 list is used again in 2019. For more information, please see tropical cyclone names. To avoid the future confusion, we create `storm-id` variable by combining `name` and `year` together (e.g., *Sandy-2012*). In the original file, there are storms labeled with NAMEs but others labelled with `UNNAME`. Here, we use `name_id` variable to indicate whether a storm has a name or not.

Secondly, we estimate `category` variable from wind speed based on Saffir-Simpson storm category, calculate the diameter of the area experiencing hurricane strength winds (64 knots or above), `_ts_diameter_` from extent of 34 kt wind radii maximum extent in northeastern quadrant (in nautical miles, `extent_34_NE`), 34 kt wind radii maximum extent in southeastern quadrant (in nautical miles, `extent_34_SW`), 34 kt wind radii maximum extent in northeastern quadrant (in nautical miles, `extent_34_NW`), and 34 kt wind radii maximum extent in southeastern quadrant (in nautical miles, `extent_34_SE`), `_hu_diameter_` from extent of 64 kt wind radii maximum extent in northeastern quadrant - `extent_64_NE`, southeastern quadrant - `extent_34_SW`, northeastern quadrant, `extent_64_NW`), and southeastern quadrant - `extent_64_SE`.

Thirdly, we estimate the storm duration `tc_dur_track` to those with maximum sustained surface winds of at least 35 knot and defined storms and define `tc_dur_type` for type of the duration. Here, `S` indicates storms with duration of 2.0 days or less and will be mentioned in the following text as *short-lived* storms, and `L` represnts storms with duration of more than 2.0 days and will be referred as "medium-to-long lived" storms.

Finally, the ocean surface temperature is `.nc` formate, we use `ncdf4` to read in data. Note that there is there is an "ET" typo in *Status* of system in the HURDAT2, which has been corrected to `EX` in the output `data\clean\hurricanes.csv` file.

Meaning for each variables

`_id_`

Storm id, which is unique. An id is a combination of 8 characters,

for example, 'AL092011',

- AL (Spaces 1 and 2) – Basin – Atlantic

- 09 (Spaces 3 and 4) – ATCF cyclone number for that year

- 2011 (Spaces 5-8, before first comma) – Year

for detail information, please see dataformat

`_name_`

Storm Name, which is non-unique. There are six lists that are used in rotation and re-cycled every six years, i.e., the 2013 list is used again in 2019. For more information, please see tropical cyclone names.

`_storm_id_`

Storm name and id combined, i.e., Sandy-2012

`_unname_label_`

Storms have name or not ("yes", "no")

`_datetime, year, month, day, hour_`

Date of report (in Universal Time Coordinate)

`_record_identifier_`

C – Closest approach to a coast, not followed by a landfall

G – Genesis

I – An intensity peak in terms of both pressure and wind

L – Landfall (center of system crossing a coastline)

P – Minimum in central pressure

R – Provides additional detail on the intensity of the cyclone when rapid changes are underway

S – Change of status of the system

T – Provides additional detail on the track (position) of the cyclone

W – Maximum sustained wind speed

`_latitude,longitude_`

Location of storm center

`_status_`

Storm classification (Tropical Depression, Tropical Storm, or Hurricane)

TD – Tropical cyclone of tropical depression intensity (< 34 knots)

TS – Tropical cyclone of tropical storm intensity (34-63 knots)

HU – Tropical cyclone of hurricane intensity (> 64 knots)

EX – Extratropical cyclone (of any intensity)

SD – Subtropical cyclone of subtropical depression intensity (< 34 knots)

SS – Subtropical cyclone of subtropical storm intensity (> 34 knots)

LO – A low that is neither a tropical cyclone, a subtropical cyclone, nor an extratropical cyclone (of any intensity)

WV – Tropical Wave (of any intensity)

DB – Disturbance (of any intensity)

`_category_`

Saffir-Simpson storm category (estimated from wind speed. -1 = Tropical Depression, 0 = Tropical Storm)

`_max_wind_`

storm's maximum sustained wind speed (in knots)

`_min_pressure_`

Air pressure at the storm's center (in millibars)

`_ts_diameter_`

Diameter of the area experiencing tropical storm strength winds (34 knots or above)

`_hu_diameter_`

Diameter of the area experiencing hurricane strength winds (64 knots or above)

`_max_category_`

Maximum category of each storm track

`_max_status_label_`

Label ("TRUE", "FALSE") to indicate whether the measurement is for the maximum status of each track

`_max_max_wind_`

The maximum value of the `max_wind` for each track

`_min_min_pressure_`

The minimum value of the `min_pressure` for each track

`_max_ts_diameter_`

The maximum value of the `ts_diameter` for each track

`_max_hu_diamter_`

The maximum value of the `hu_diameter` for each track

!–test–>

Chapter 4

Missing Values

Here, we use `visna::extract` to look at the missing value patterns. *The bars* beneath the columns in Figure 4.1 show the proportions of missingness by variable, suggesting hurricane diameter *(hu_diameter)* and storm diameter *(ts_diameter)* both have the highest number of missing value and the columns suggest that they follow the same missing pattern, meaning when hurricane diameter is missing, storm diameter is also missing.

The third most missing variable is Pressure *(min_pressure)* and *the columns* show that Pressure is missing only when hurricane diameter and storm diameter are missing. The *bars* on the right show the relative frequencies of the missing patterns, which suggest the most frequent missing patterns are in the combination of hurricane diameter, storm diameter and pressure, followed by the combination of hurricane diameter and storm diameter. Non-missing data are in the third meaning most of rows are completeness.

Finally, hurricane diameter and storm diameter are the most missing values because they are calculated from *Wind Raddii* but *Wind Raddii* values were not used before 2004.

It is worth noting that almost all Pressure data are missing from 1850s to 1940s. The number of missing data is then decreasing from the 1940s, and there are no missing Pressure starting the 2000s (see Figure 4.2). The reasons for this are are the following: (1) in the early years, information about pressure was recorded by ships; those were few in numbers and thus a lot of pressure data could not be recorded (2) in more recent years, it became a common habit to replace missing pressure values by an analytical product such as sattelite data; (3) improvements in modern tools and technologies have provided us with a more powerful observation network than that in the early years.

Another noticeable pattern in missing value relates to the name of tropical cyclones (see Figure 4.3).
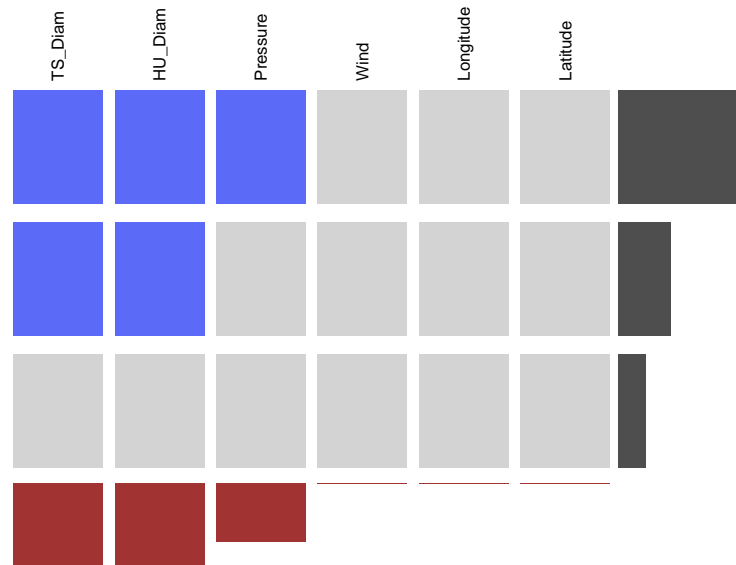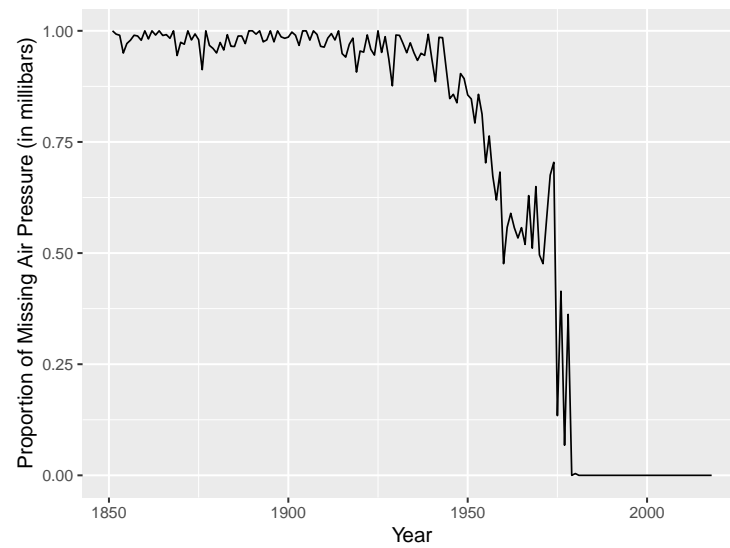
13

Figure 4.1: Missing Values Patterns

Figure 4.2: Proportion of Missing Air Pressure at the Storm's Center By Year
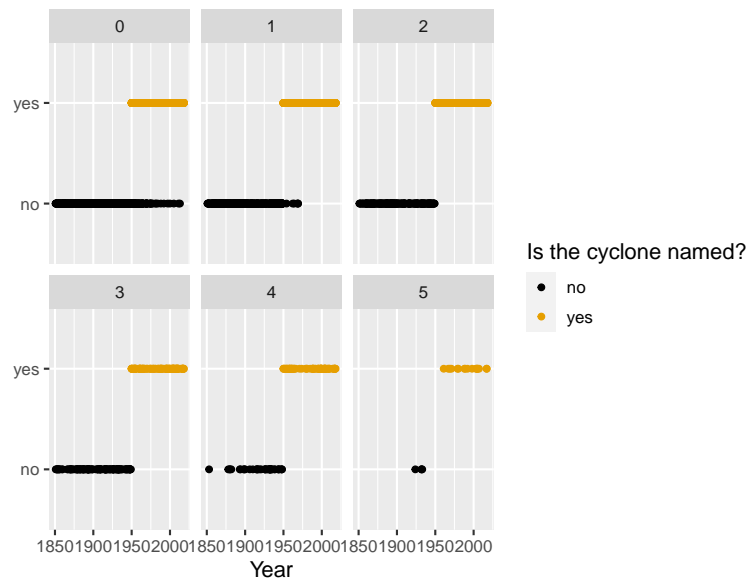
Figure 4.3: Cyclones received names starting 1950s

Before the 1950s, it was not common practice to name tropical cyclones. This explains why we cannot find any cyclone with a name before that time. A couple were still not named after that, and the reason can be linked to the category: those cyclones are of category 0 and 1 (more information about categories can be found in part 5), which means that they could have been overlooked and therefore didn't receive any name.

Chapter 5

Results

5.1 What influences the change in category of a hurricane?

Hurricanes are classified in different categories based on the Saffir-Simpson Hurricane Wind Scale, which classifies hurricanes with regards to their wind speeds and estimates potential property damages. The hurricanes categories come under five items:

(1) Category 1:

Very dangerous winds from 64 to 82 knot (kt) that will lead to some damage such as roof damage or snapped tree branches. A famous hurricane of category 1 is Danny, which hit Cuba and the Gulf of Mexico in 1985.

(2) Category 2:

Extremely dangerous winds from 83 to 95 kt that will produce extensive damage such as roof damage and siding, many toppled shallowly rooted trees and near-total power outage. A famous hurricane of category 2 is Erin which hit Florida in 1995 and caused 5 people to die and cost $700 million (1995 USD).

(3) Category 3:

Extremely dangerous winds from 96 to 112 kt will cause devastating damage such as removal of roof decking, power outage and power outage for days to weeks after the storm passes. Katrina is a famous hurricane of category 3 that lead to 1,836 deaths and $81 billion (2005 USD) in property damage in 2005. Major states impacted were Louisiana and Mississippi.

(4) Category 4:

Extremely dangerous winds from 113 to 136 kt that will cause catastrophic damage such as loss of most of the roof structure and exterior walls, fallen trees and power outage for weeks to months. A famous hurricane of category 4 is Harvey which hit Texas and Louisiana in 2017 with damage estimated to $125 billion (2017 USD) and 107 deaths in the United States.

(5) Category 5:

Extremely dangerous winds of 137 kt or higher that will cause catastrophic damage such as total roof and wall collapse, fallen trees, power outage for weeks to months and uninhabitable regions for weeks or months. A famous hurricane of category 5 is Andrew, which hit Louisiana and Florida in 1992 and whose damage accounts for $41.5 billion (2011 USD) and 54 deaths (approximately 1 million people were evacuated before the storm).

Besides, it is also worth distinguishing hurricanes from tropical storms, whose maximum sustained surface winds range from 34 to 63 kt. They are labeled category 0 in our dataset.

The strength of a hurricane can also be assessed by its barometric pressure, that is the pressure at its center: the lower the barometric pressure, the stronger the hurricane.

Finally, the size of a hurricane is defined by its diameter.

Let's now examine those features on a parallel coordinate plot and see how they relate to each other.

First of all, as showing in Figure 5.1, categories are clearly distinct in terms of windspeed and increase proportionally to this feature. This is consistent with the definition of the Saffir-Simpson Scale. The same conclusion can be drawn for the barometric pressure measured at the center of the hurricane: for low pressures, the hurricane will be overall of a higher category. The distinction is however not as clear as for windspeed and some overlap can be noticed.

Furthermore, no relationship can be established between the diameter of the area experiencing hurricane strength winds of 64 knots or above hurricane diameter *(hu_diameter)*, which would mean that the category, wind speed and min pressure will not determine the potential areas that will be affected by hurricane.

Note that all tropical storm have a diameter equal to 0: this is because the variable used for the graph is hurricane diameter, which accounts for the diameter of hurricanes. The diameter of the area experiencing tropical storm strength winds (34 knots or above) is reported in a separated variable called storm diameter *(ts_diameter)*. As this project focuses on hurricanes, and these two parameters are highly correlated to each other, the decision was made to not include it in the graph.
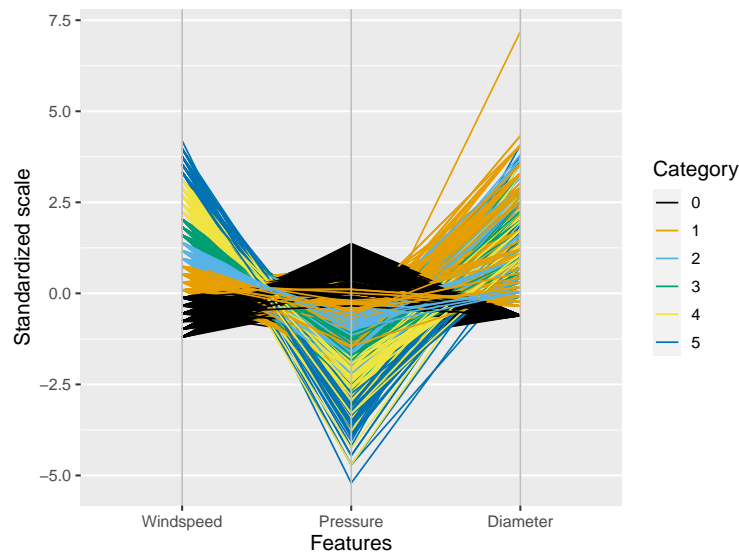
Figure 5.1: Windspeed and Pressure are indicators of cyclone category

5.1.1 A first look into the evolution of the category of hurricanes

Let's now check whether a hurricane sticks to one category in its lifetime or if it changes overtime. To do so, we select three hurricanes of the same year: Irma (Florida), Maria (North Carolina) and Gert (North of Atlantic Ocean), which all occurred in 2017.

From Figure 5.2, it can be noted that both Irma and Maria grew in intensity to reach category 5, while Gert **only** reached category 2 as a maximum. A difference in lifetime can also be highlighted (almost 2 weeks vs a couple days). As both Irma and Maria hit Southern states and Gert occured at a Northern location, the following hypothesis is stated: is the maximum category reached by a hurricane dependent on its location? That is, *is there a difference between the maximum category reached by Southern and Northern hurricanes?*

Before testing that hypothesis, it is worthy to note that the vertical jumping lines in the time series of hurricanes (Figure 5.3). These suggest that hurricanes are highly dynamic and can lose/gain energy in only a couple hours. This concept is illustrated by the following time series (Figure 5.4), showing a close up of hurricane Irma evolution on September 9th 2017; in a couple hours, the hurricane when from category 5 to category 2.
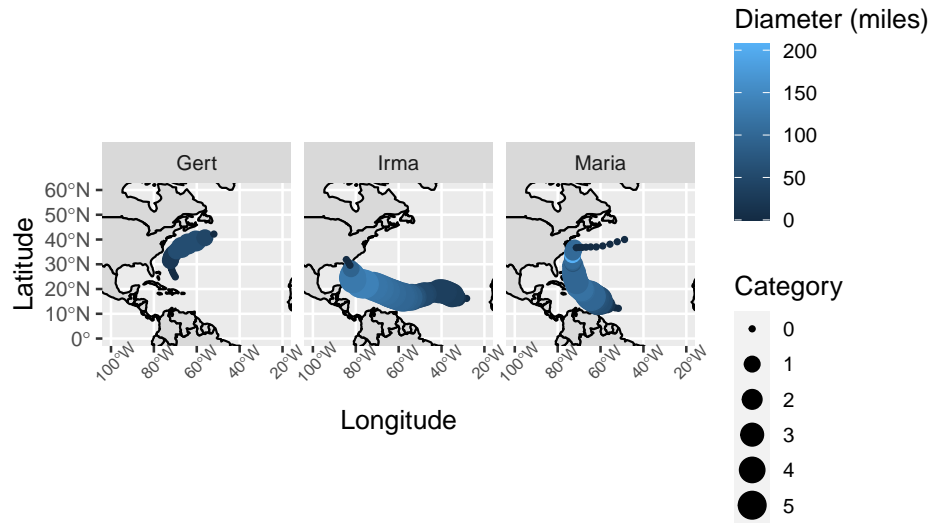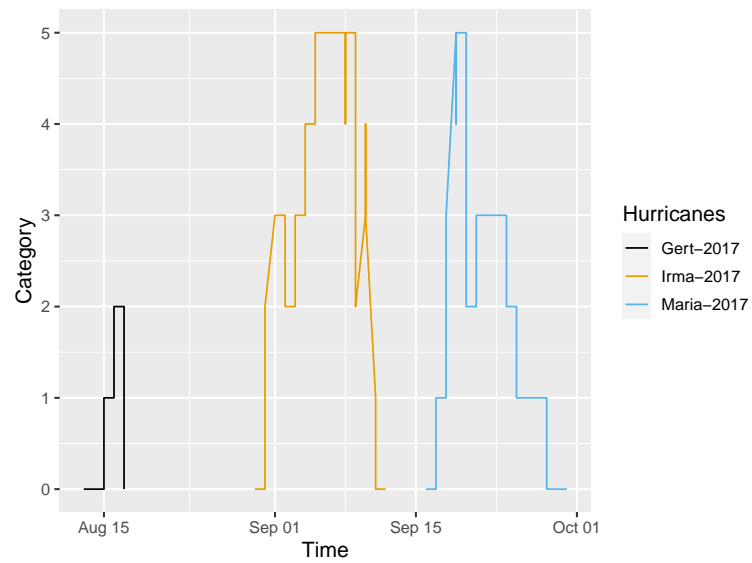
Figure 5.2: Three Example Hurricanes in 2017

Figure 5.3: Time Series of Three Example Hurricanes in 2017
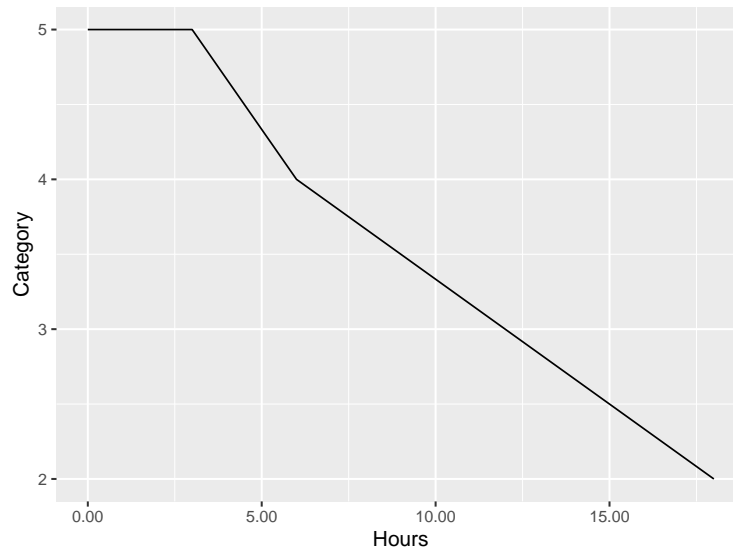
Figure 5.4: Close Up of Time Series of Hurricane Irma in 2017

5.1.2 Is there a difference between the maximum category reached by Southern and Northern hurricanes?

Let's now verify whether the hypothesis holds. The Figure 5.2 depicts the path of each hurricane across time and also colors the diameter the area experiencing hurricane strength winds of 64 knots or above and sized according their categories.

Each hurricane formed in the low latitude east with a low category, then intensified (large circle and bluer) as it extracted energy from the warm/moist air over the oceans. As a result, its category evolved with time until it reached a certain point then dissipated. The hurricanes then broke apart because they move through cold water (e.g, Maria) or move over land (e.g., Irma) which made them losing touch with the hot water that powered them.

When hurricanes start at different location point, they absorb different amount of energy throughout their progression towards the west or northwest. This will lead to different energy formation that will determine the highest wind level measured, and in turn decide what that category of that hurricane will be.

The comparison of the hurricane track and intensity indicates that the longer the hurricane stay over the southern warm water, the stronger it gets to be (see Irma), while the northern formed storms, like Gert, could not reach high category and they die out a few hours after they start because they do not have enough energy to become a hurricane.

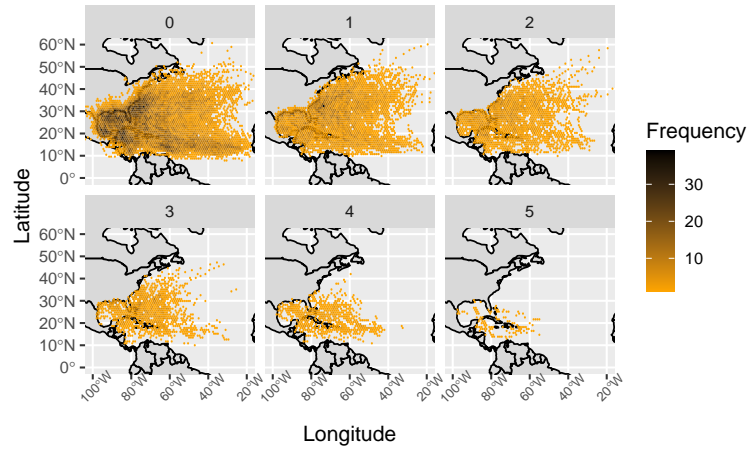This seems to confirm the hypothesis stated above: there seems to be an inter-

Figure 5.5: Frequency of Hurricanes across the United States

esting relationship between location and category of a hurricane. In addition, hurricanes seem to reach higher categories when located in the South of the Atlantic Ocean than in the North.

Let's now check whether the hypothesis can be extended in terms of frequency. The frequency of hurricane by location for each category. Figure 5.5 shows that storms and hurricanes disappear when they reach the land or the northern water.

The low category storms (e.g., category 0 and 1) with the highest density along the coastline can reach further North, while the high category hurricanes (4 and 5) only appear in the South. The highest density along the coastline is related to the progress of the small tropical depressions. When these small depressions travel over the ocean to the US eastern coast, they gain energy and update to the category 0 storm. It is worth to note that as category increases, their north boundary seems to shift toward the south, suggesting the feature of storms and hurricanes are location-related.

We build a Shiny Web Application for the Interactive Component in order to assist with visualizing this result.

5.2   Does hurricane activity shift north-ward?

#{r echo=FALSE} #df <- read.csv("../data/clean/hurricanes.csv") #

Next, we want to assess whether hurricane activity is shifting North. A recent study showed that the locations of tropical cyclones is shifting poleward (i.e. north-ward) globally (Thompson). If true, this implies heavy consequences for the Northern part of the United States. Southern states such as Florida
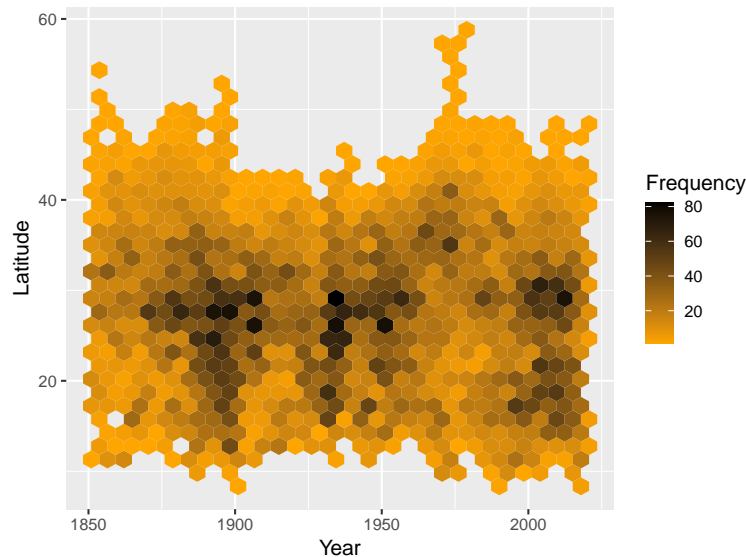
Figure 5.6: Frequency of Hurricanes by Latitude over Years

have already taken steps to improve its approach to hurricanes, such as property structure improvements, rescue boats and vehicles investments, and emergency workers trainings; if the frequency of tropical cyclones is indeed shifting North, Northern states should consider taking similar actions.

Before exploring the above hypothesis, it is worth redefining latitude and longitude:

Latitude is the measurement of a place on Earth with regards to its distance north or south-wise from the Earth's equator. Latitude is expressed in degrees followed by the letter N or S whether the location measured is North or South from the equator. Interesting examples for this project are Miami's latitude (~26°N) and New York City's latitude (~41°N).

Longitude is the measurement of a place on Earth with regards to its distance west or east-wise from the meridian at Greenwich, England. It is also measured in degrees and followed by the letter W or E depending on whether the location of interest is West or East from the meridian.

We will focus on latitude to verify whether hurricane activity seems to be changing directions.

Figure 5.6 shows a heatmap of hurricane activity based on time and location.

While tropical cyclones frequency seems to be consistently higher at lower latitudes, one can observe an increase in cyclones activity at higher latitudes in the 1970s. However, this shift does not seem to be permanent as the frequency of tropical cyclones seems to decrease again in the following decades.
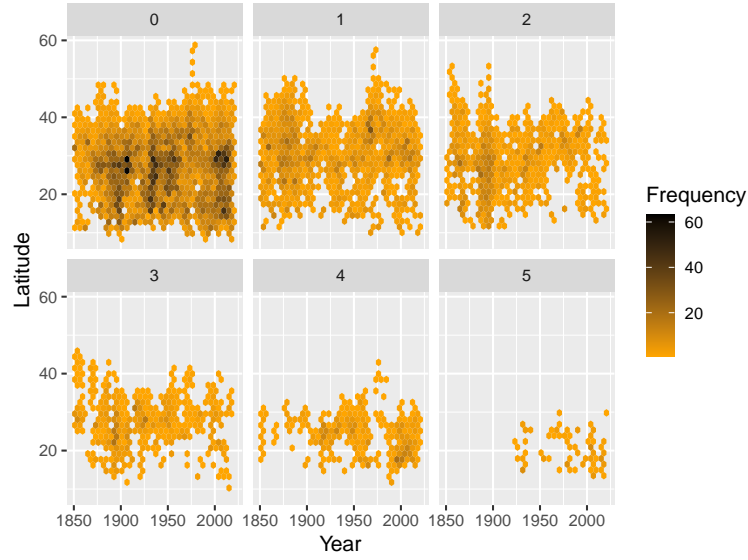
Figure 5.7: Frequency of each Hurricane Category over Years

Besides, if we look at the tropical cyclones activity faceted on category (Figure 5.7), the shift in location seems to be focused on tropical storms (category 0) and hurricanes of category 1. Hurricanes of category 4 are also more frequent on higher latitudes in the 1970s but the frequency remains at a low level and the activity returns to lower latitudes in the following decades. These observed trends would therefore not justify investments in hurricane-proofed facilities for Northern States.

5.3 Have humans already caused a detectable long-term increase in Atlantic hurricane activity?

As we can see from Figure 5.8, the global CO2 concentration has increased, and the tropical Atlantic sea surface temperature show significant warming trends (Figure 5.9).

Besides, one can also observed an increase in the number of storms and hurricanes in the Atlantic from 1851 to 2018 (Figure 5.10). We used this finding to ask the following question: *is there a link between greenhouse-induced warming and changes in the Atlantic storms frequency (Figure 5.11)?*.

Based on past Atlantic hurricanes records from 1878-2006, scientists concluded that the increasing trend in the number of Atlantic storms is almost completely
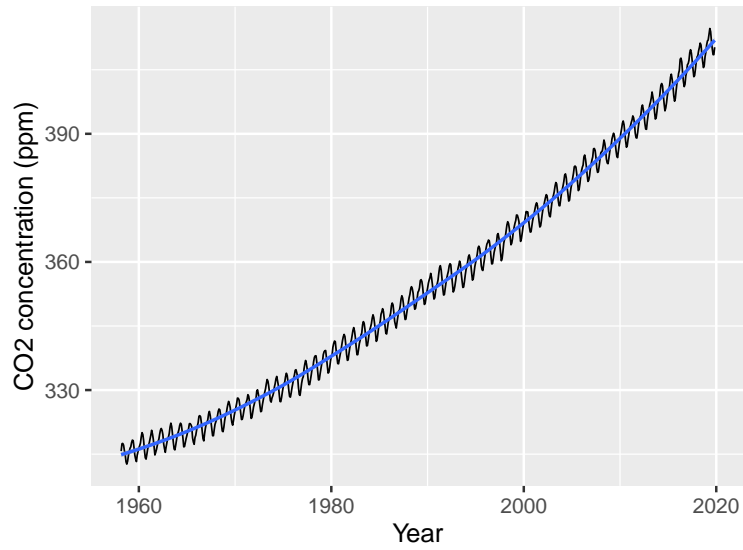
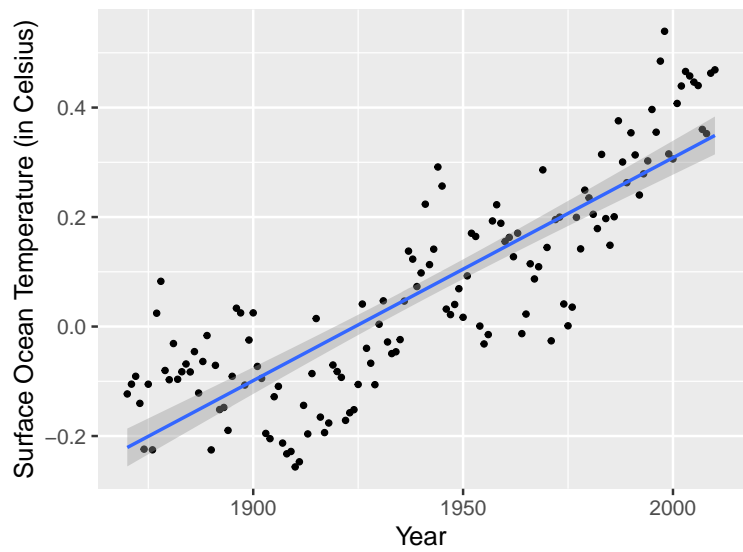Figure 5.8: Time Series of Atmospheric CO2 (ppm)

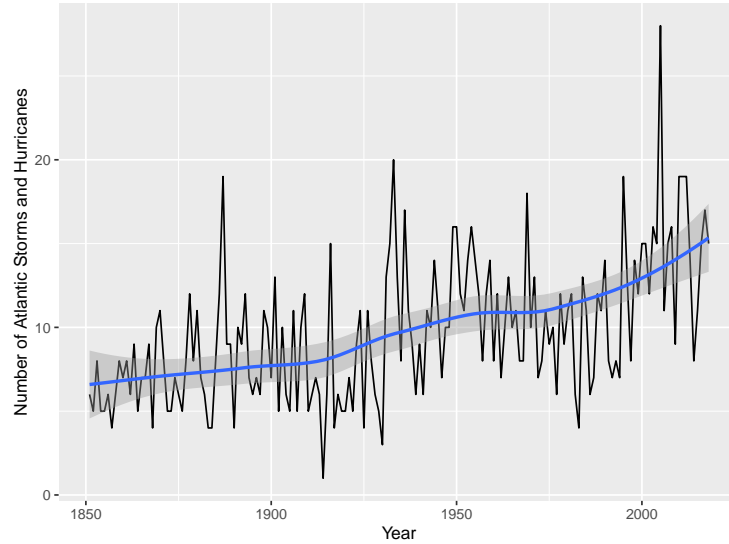
Figure 5.9: Time Series of Ocean Surface Temperature

Figure 5.10: Number of Atlantic Storms and Hurricanes Over Time

due to the increase in short-duration storms (less than 2 days) (Landsea et al. (2010)), suggesting that only the count of short-duration storms follow a similar trend as the CO2 concentration and surface ocean temperature.

However, at the time, the focus was on storms that could have had an impact on ship traffic and, therefore, it is very likely that short-lived storms were overlooked in the earlier parts of the record. Besides, once our group explored the dataset with extra ten years' record, we found that there is an increase trend on both short-lived and long-lived storms.

This means that the evolution of the number of cyclones over time follows a similar trend as the CO2 concentration and surface ocean temperature. This could mean that humans have had an impact on the increase of hurricane activity, but to prove that this is a causation and not only a correlation, deeper research and statistical tests should be performed.
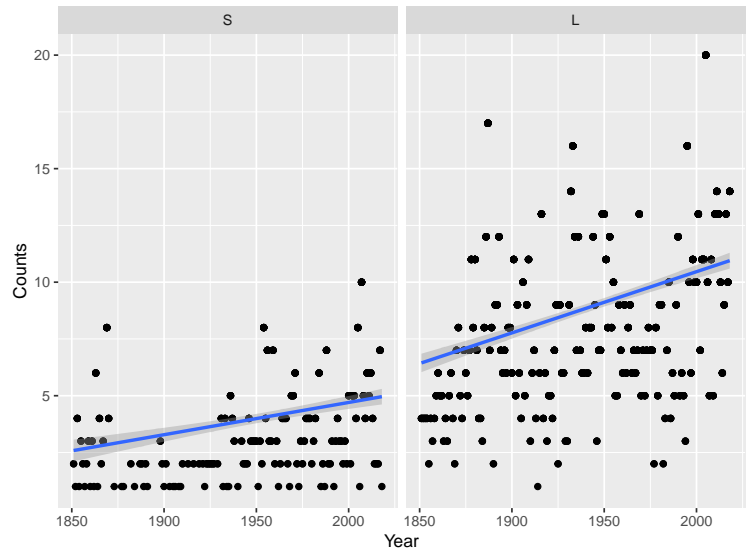
Figure 5.11: Time Series of Atlantic Short-lived ('S') and Long-lived ('L') Storms and Hurricanes Number

Chapter 6

Interactive Component

The following Shiny App that helps the user to assess hurricanes' features evolution across time and location as done in part 5.1 of this analysis.

The graph on the left displays the hurricane path in the Atlantic Ocean and its change in category across locations. The category can be assessed by hovering on the path; the date at which the hurricane was at each location can be assessed by clicking on the path.

The graph on the right displays the evolution of windspeed, barometric pressure and hurricane diameter across time of a hurricane. If values for the selected variable are missing, the graph will be blank.

If the server is too slow and the app doesn't appear on your screen, follow this Github link to download the necessary files and run the app locally.

Hurricane Shiny App

```
knitr::include_app('https://hurricane.shinyapps.io/01_01/', height = '600px')
```
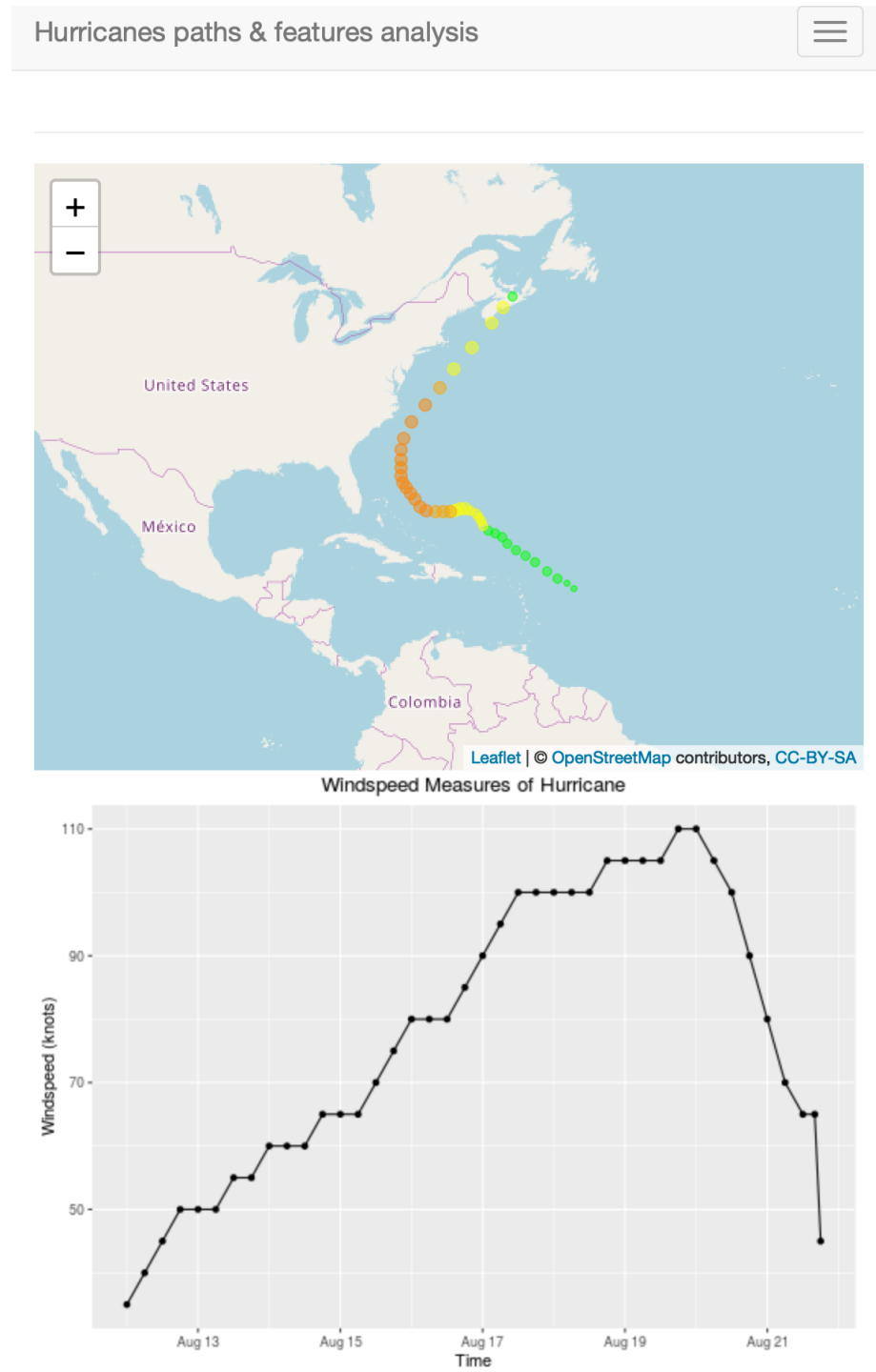
Figure 6.1: app

Chapter 7

Conclusion

In this analysis, we explore the century long storm records [HURDAT2: 1851-2018] (https://www.nhc.noaa.gov/data/hurdat/hurdat2-1851-2018-120319.txt) to investigate the characteristics of Atlantic storms and hurricanes and examine their changes under global warming conditions. The analysis results indicate that:

(1) The Saffir-Simpson scale that is currently widely used to define category of hurricane reflects well the strength of wind and pressure. However, no relationship could be established between the category and size of the cyclones, indicating that a high category hurricane can be of small size and vice versa. Finally, Southern cyclones seem to reach a higher category than Northern cyclones in their lifetime, indicating a relationship between the category and the location. The century long Atlantic storms and hurricanes tracks and features can be easily visualized from Hurricane Shiny App.

(2) While recent studies indicate a global North-ward shift in cyclones activity, we did not find strong evidence of this trend in our data. An increase of activity in the North was observed in the 1970s but seemed to decrease in the following decades. Besides, that shift seemed to be due to mostly an increase in frequency of cyclones of category 0 and 1, which are less worrying than cyclones of higher categories.

(3) Human activities may have already caused the changes in the Atlantic storm and hurricane activity, reflected by the increasing number of the Atlantic cyclones. However, exploratory analysis can only help us identify similarity in trends and not confirm that one is the cause of the other. In order to do the latter, further research and statistical testing should be performed.

Bibliography

D'Souza, D. (2019). *Hurricane Dorian: Measuring the Economic Impact.* Investopedia.

Landsea, C. W., Vecchi, G. A., Bengtsson, L., and Knutson, T. R. (2010). Impact of duration thresholds on atlantic tropical cyclone counts. *Journal of Climate*, 23(10):2508–2519.

SDG-Indicators (2018). Global indicator framework for the sustainable development goals and targets of the 2030 agenda for sustainable development.

Weinkle, J., Landsea, C., Collins, D., Musulin, R., Crompton, R. P., Klotzbach, P. J., and Pielke, R. (2018). Normalized hurricane damage in the continental united states 1900–2017. *Nature Sustainability*.