

CSC110 Project Report:

The Effect of Global Warming on Hurricane and Typhoon Occurrence

Mu "Samm" Du

Friday, December 14, 2020

Introduction

The NASA Earth Observatory defines Global Warming as “the unusually rapid increase in Earth’s average surface temperature over the past century” (Riebeek). As the global temperature has risen over the past decades, we are seeing more frequent occurrences of hurricanes and typhoons across the globe; these natural disasters have damaged properties, ruined communities, and destroyed people’s lives (Miller). **Is the increase in hurricanes and typhoons related to global warming? If so, how are they related?**

I will analyze global temperature data procured by Berkeley Earth across land and ocean, as well as hurricane and typhoon occurrence data collected by the US National Oceanic and Atmospheric Administration limited to the Atlantic basin and the North Pacific Ocean regions.

Datasets

- Berkeley Earth — **Monthly Global Average Temperature data** (Land + Ocean) 1850-Recent
Data format: txt
Columns used: Year, Month, Monthly Anomaly
Regression range: 1851-2009 (inclusive)
Test range: 2010-2014 (inclusive)
Source: <http://berkeleyearth.org/data-new/>
- NOAA — **Hurricanes and Typhoons** (Atlantic basin and the North Pacific Ocean) 1851-2014
Data format: csv
Columns used: ID, Date
Regression range: 1851-2009 (inclusive)
Test range: 2010-2014 (inclusive)
Source: <https://www.kaggle.com/noaa/hurricane-database/>

Computational Overview

- I will complete the project in a Jupyter Notebook, so that I can present my code, descriptions, output, and interactive elements in a consistent format
- I will filter the datasets to only include relevant information to my analysis, such as the temperature values and timestamps in the global temperature data
- I will limit the time range such that the two datasets completely overlap in time
- I will aggregate the hurricane and typhoon data by counting the number of unique events that occur for every month. I will identify unique events by the ID column of each row, counting each unique ID as one event

- I will try to correlate global temperature with the number of hurricane and typhoon occurrence through a few regression functions, compare them, and pick out the one with the best fit without over-fitting;

It is also important to acknowledge that while the temperature data is on a global scale, the hurricane and typhoon occurrence data is only limited to the Atlantic and Pacific regions, therefore a degree of error is involved with the correlation;

I will use the **scikit-learn** library for the purpose of construction a regression model, since **scikit-learn** has a trove of regression functions built-in, and I will use **plotly** to present the data graphically, since it is a versatile plotting library that I have become familiar with in class

- I will do a projection for the number of hurricanes and typhoons that are likely to occur in the next 5 and next 10 years, if the temperature keep rising based on the regression model constructed above;

It is also important to note that since the hurricane and typhoon data is only as recent as 2014, the projection into the future may not be as accurate

- To test the accuracy of my model, I will designate data between years 2010 and 2014 (5-year interval) as the testing range, and compare the projection from my model with the data during that interval
- I will create an interactive element (using the **ipywidget** library) where the user can adjust the slider to increase and decrease global temperature and see how that affects the number of hurricanes and typhoons in the future, based on the regression model, acknowledging the fact that the correlation derived from the model doesn't necessarily imply causation

Obtaining Datasets and Executing Code

Obtaining dataset files:

Method 1: You may clone the entire project and file structure from the following GitLab URL:

```
git clone https://gitlab.com/sammdu/csc110-final-project
```

Method 2: You may download the 3 dataset files (from the two original sources) from the following GitLab URL:

```
https://gitlab.com/sammdu/csc110-final-project/-/tree/master/data
```

After downloading the 3 dataset files, you must place them under a folder named **data**, that is directly under the project root. You may reference the GitLab project from **Method 1**.

Executing code locally:

The project code is written in a **Jupyter Notebook** file, **main.ipynb**. After installing all dependencies in the **requirements.txt** file, you can open a terminal in the project root and execute **jupyter notebook**. In the opened browser window, select **main.ipynb**. In the notebook file, simply execute every code block from top to bottom in order; alternatively, select “Kernel” ⇒ “Restart & Run All” to automatically run all code blocks in order. Make sure the correct Python 3 kernel is selected to avoid environment errors.

Further instructions are available within the **main.ipynb** file.

Discussion

TODO: include discussion

References (MLA)

“Changes in Earth’s Global Average Percentage Surface Temperature.” Berkeley Earth, Accessed 11 Dec. 2020, berkeleyearth.org/data-new/.

“Hurricanes and Typhoons, 1851-2014.” Kaggle, National Oceanic and Atmospheric Administration, 20 Jan. 2017, www.kaggle.com/noaa/hurricane-database.

Miller, Peter. “Weather Gone Wild.” National Geographic, Sept. 2012, www.nationalgeographic.com/magazine/2012/09/extreme-weather-global-climate-change-effects/.

Riebeek, Holli. "Global Warming." Earth Observatory, NASA, 3 June 2010, earthobservatory.nasa.gov/features/GlobalWarming.