

Waves in Portugal vs. Hurricanes in the Atlantic

The title says everything. In the present report I try to find a correlation between waves height in the Portuguese coast with hurricanes happening in the Atlantic Ocean.

Context

An auditory stimulus makes millions of neuron cells in the brain to react and vibrate. This vibration can be measure by means of electroencephalography (EEG) at the surface (scalp), resulting in a wave as visible output.

Likewise, in the ocean waves arise due to storms (e.g. hurricanes), and can be measure with buoys floating at the surface.

With a neuroscience background, I ventured through the oceanography field with the help of my dad (an engineer in coastal management) and tried to find some correlations between ocean waves in Portugal and hurricane data from the Atlantic.

Sources of data

1) Waves measuring buoy data, 1995-2014

A buoy placed near Leixões (close to Porto, Portugal), around 20km from the coast. This data was kindly provided by the Hydrographic Institute, Portuguese Agency for Environment (IH - APDL/ Projecto MarRisk APA). These data can only be used for educational and investigational purposes and for the sake of its protection (and as agreed with me) I can't share the original files.

Among other measures, I used wave height (meters) and wave direction (degrees), as well as date of the recording (this was the point of connection with the below dataset).

2) Hurricanes and Typhoons in the Atlantic, 1851-2015

Location, wind, and pressure of ~50000 tropical hurricanes in the Atlantic. This set I downloaded the .csv file from Kaggle, which in turn was obtained from the National Hurricane Center.

Data analysis

1) Waves measuring buoy data, 1995-2014

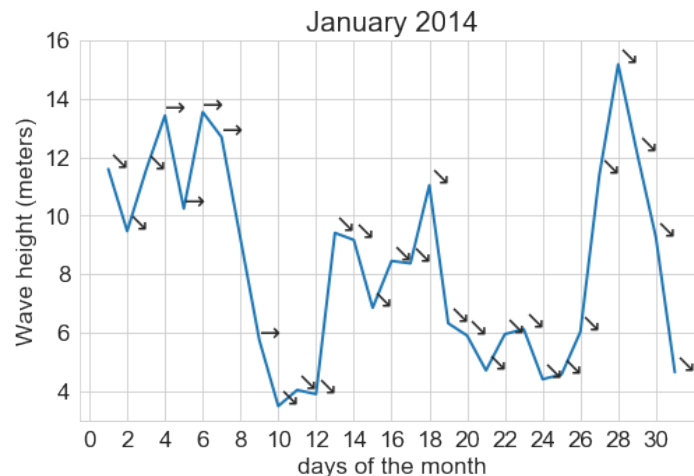
For the waves data, I dropped some columns that were not necessary for my analysis, and renamed the columns, as the original files were not column indexed. I converted the wave direction, in degrees, to actual compass directions (N, S, NW, SE, etc) and then to little arrows (for plotting purposes). There were several measures per day: wave height was sampled every 3h, but this was not consistent. For that, and for the sake of plotting, as well as presentation effect, I picked the highest measure per day. (Note: I only had measures for the winter months).

2) Hurricanes and Typhoons in the Atlantic, 1851-2015

For the hurricanes dataset, the data was pretty much organized. I just had to separate the 'Date' column into 3 columns ('day', 'month', 'year'), which was my 'point of contact' with the Portuguese waves. By means of pandas library in python, I filtered by date the hurricanes I wanted to see and compare with the Portuguese waves.

Plotting

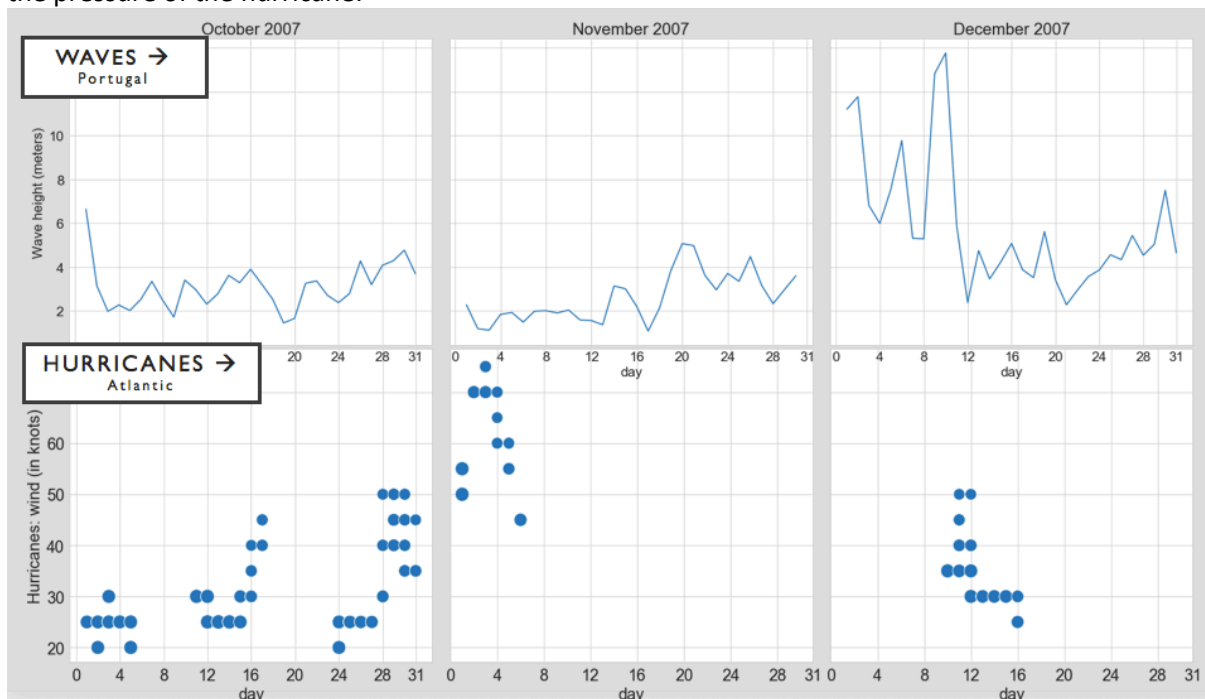
1) Waves measuring buoy data, 1995-2014



This was my favorite part. I had much fun in plotting the waves in Portugal as I could relate to some events that happened in the Portuguese coast (e.g. Hercules storm January 2014). I used a combination of matplotlib and seaborn libraries. Main challenge was the printing of the little arrows on the graph and sub-plotting (grid).

2) Hurricanes and Typhoons in the Atlantic, 1851-2015

For the hurricanes (bottom of below picture), I used the scatter plot from seaborn library. I plotted the wind (in knots) in function of the day. The size of the dots (not very significant) are in function of the pressure of the hurricane.



Correlation of data

Due to shortness of time as well as nature of the data, I didn't perform any type of correlation. There were too many variables (wind, direction, pressure, origin of hurricane) to correlate with the wave height. Some correlation can be inferred just by observing the plots (above) but it is very difficult to make any final conclusion, as too many variables were in game.

Final thoughts

I got a bit disappointed with not exploring the correlation and statistics part in this project. This was the new factor learned this week and was a shame I didn't put it into use in this project, due to lack of time.

Nevertheless, I really enjoyed this project. I learned a lot on the way, not only about ocean behavior but also how to handle huge amounts of data. The last one leaves me in the expectation for the next module – machine learning – I'm really curious about this new topic that will give me the tools to handle more data in a more effective way.

Filipa Lopes

Project 4

13th September 2019

#data-squad-21

Ironhack