Ocean Meteorological Monitoring via National Oceanic and Atmospheric Administration Data

Masoud Masoumi

July 2019

Contents

1 Introduction

This section provides a brief description of the problem and the kind of data which will be used to answer that problem along with the possible limitations regarding the available data sets.

1.1 Background

The total available wave power resource is roughly estimated 2.7 terawatt. The available energy varies depending on water level, wave period, wave height, and the accessibility of the ocean for wave energy harvesting. This untapped resource creates an incentive to develop and commercialize harnessing technologies. Over the last few decades, there have been countless efforts in the US, Europe and Asia to come up with a viable technology with an efficient wave-to-wire energy transformation approach. "With more than 50 percent of the U.S. population living within 50 miles of coastlines, there is vast potential to provide clean, renewable electricity to communities and cities across the United States using wave energy" based on US Department of Energy (DOE) research. This was the main incentive for DOE to announce US Wave Energy Prize competition in 2016 and Wave Energy Water Desalination Prize Competition in 2019.

Among the challenges to come up with a mature technology, is the variations in ocean waves' characteristics in different areas and coastlines. Ocean waves vary in terms of height, directions, and time period which makes the development of an efficient harnessing mechanism a difficult task.

1.2 Problem Definition

1.2.1 Definition

This project tries to provide a region-dependent guideline for technology developers who are working to implement ocean wave energy for power production and water desalination. The aim is to create a platform with meteorological data from buoys' data reported by National Data Buoy Center (NDBC) and cluster the buoys based on:

- Wave Height
- Wave Period and Direction
- Wind Speed and Direction

over one year period or shorter. Some other parameters are also monitored and visually represented such as sea level pressure, sea surface temperature, and air temperature.

1.2.2 Limitations & Assumptions

Limitations are mostly caused by the available and accessible data for each category. For instance, the data provided by NDBC is reported for all buoys owned and operated by more that 130 centers, institutes, and universities. These buoys have different capabilities and not all of them can measure all the desired parameters.

Furthermore, there are 916 number of buoys and all of them are not operational all the time. There are some periods that no data is reported for some buoys. Also, some of these buoys were operational for few years only. So the data available for these buoys are limited to that period of time.

1.3 Stakeholders

The main beneficiary of this project is wave energy developers which are looking into technologies that can adapt itself into specific coastal regions for energy production and water desalination.

2 Data Acquisition

Here, a list of sources and the type of data acquired from those sources are provided.

2.1 Buoys Locations

The geographic coordinates of each buoy is generally reported along with the data provided for that buoy. Here, an API provided by Planet OS (`https://planetos.com/`) is used to access buys' latitudes and longitudes. First, an account has to be created on the website. Once the account is created, API Key can be found under Account Setting. The data set id for Standard Meteorological Data is `noaa_ndbc_stdmet_stations`. This data typically updates every hour.

2.2 List of Buoys

Since the data from the API mentioned in section 2.1 is not complete for each buoy and also have missing columns, we use the National Data Buoy Center website directly to get our information. The list of all the buoys with data provided on the website can be obtained by scraping the data on this `https://www.ndbc.noaa.gov/to_station.shtml` web page along with the institute which owns the buoy.

2.3 Availability of Historical Meteorological Data

As mentioned before, buoys have recorded various parameters in different time periods. First, there needs to be a list of available historical data for each buoy. To get such information, the web page at `https://www.ndbc.noaa.gov/historical_data.shtml` is scraped and the years for which each buoy was operational and measuring the standard meteorological data, were recorded. The range goes back to 1970 and is up to 2018.

2.4 Historical Meteorological Data

Finally, if the buoy has the meteorological data for a desired year, this data was downloaded directly from the website using a direct link knowing the buoy's name, and the year. The url formatting is as follows: https://www.ndbc.noaa.gov/view_text_file.php?filename=[**Buoy Name**]h[**Year**].txt.gz&dir=data/historical/stdmet/

The data description for these data sets along with the associated units are provided by NDBC at this web page: `https://www.ndbc.noaa.gov/measdes.shtml` under **Standard Meteorological Data**.

3 Data Cleaning

There are few important steps that need to be taken before exploring and visualizing the data. First, station names obtained from the web page scraping are in uppercase format while the ones from the API data collection are in lowercase format. So they should be fixed and matched.

Second, there are buoys which are available on API linked data set but are not listed in NDBC web page. So those buoys are discarded since there is no data available for them on NBDC database.

Third, if a buoy is not measuring a specific variable, the whole column corresponding to that variable is filled with 99, 999, or 9999 depending on the format of the data in that column. So these columns should be discarded for that buy.

Finally, the data for each buoy might have missing points which are replaced by 99, 999, and 9999. These data cells are removed and replaced by the mean value for that column.

4 Exploring the Data For One Buoy

Before we dive deep into tackling the problem, data for one buoy is obtained over a period of time to explore the outputs. First, let's take a quick look at the world map showing the locations of all the buoys along with coloring the spots based on different owners in Figure 1. A better quality and interactive version of this image can be found at `https://nbviewer.jupyter.org/github/MasoudMiM/Python-Codes/blob/master/WorldMapBuoysOwners.html`, where you can click on each buoy to see its name. You can also filter the buoys based on the owner from the top right of the map.

Data from the buoy, 44025 between February 25th, 2018 and December 1st, 2018 was selected. The significant wave height and average wave periods are shown in Figure 2.

Wind and wave rose plots which represent their intensity and direction are shown in Figures 3 and 4. It is worth mentioning that wave height is considered as the wave intensity. Finally, the frequency of wave frequencies and wave height plotted for that buoy during the time period are shown in Figure 5.

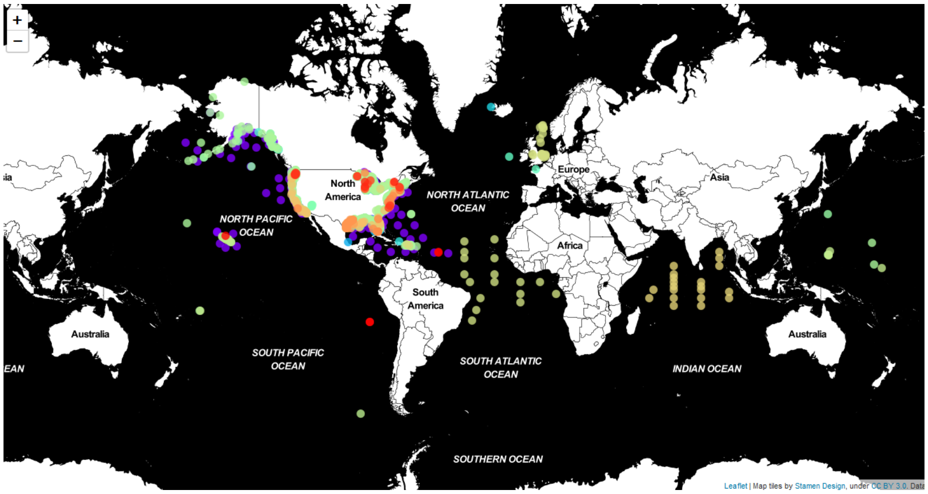
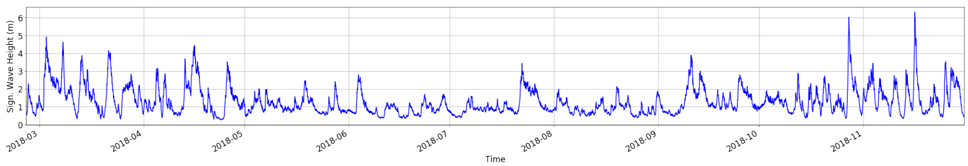


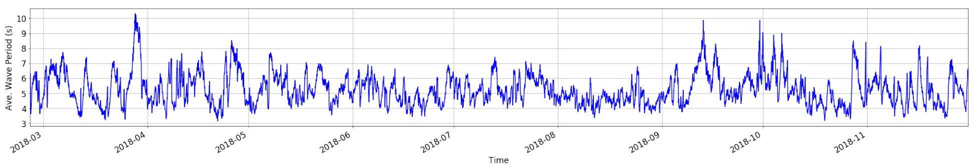
Figure 1: Locations of all the buoys on the map colored based on their owners

5 Clustering for a Time Period

For each year or any period of time, available data recorded for each buoy is downloaded, cleaned, and analyzed. Using this data, buoys are clustered based on the average wave height, average or dominant wave period and wave direction, and average wind speed and wind direction in the area. Therefore, areas suitable for power production/ water desalination using the energy of ocean waves with specific characteristics can be identified and used for the design process.



(a)



(b)

Figure 2: (a) Significant wave height, and (b) average Wave period

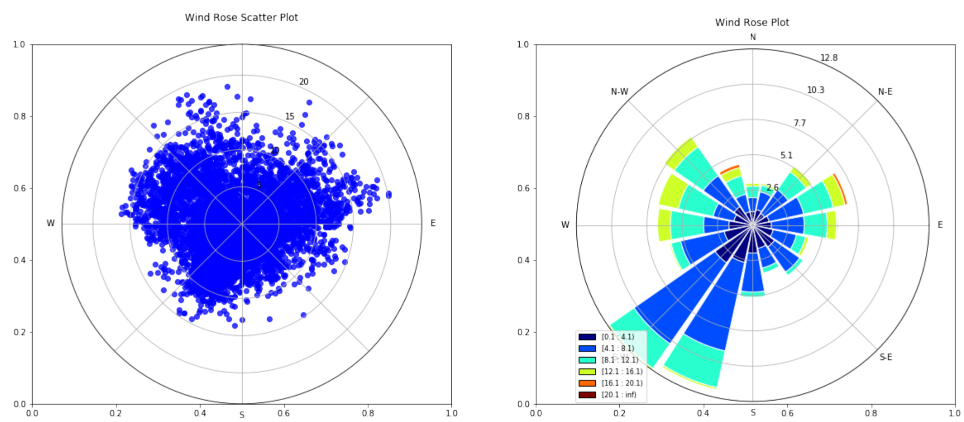


Figure 3: Wind Rose plot

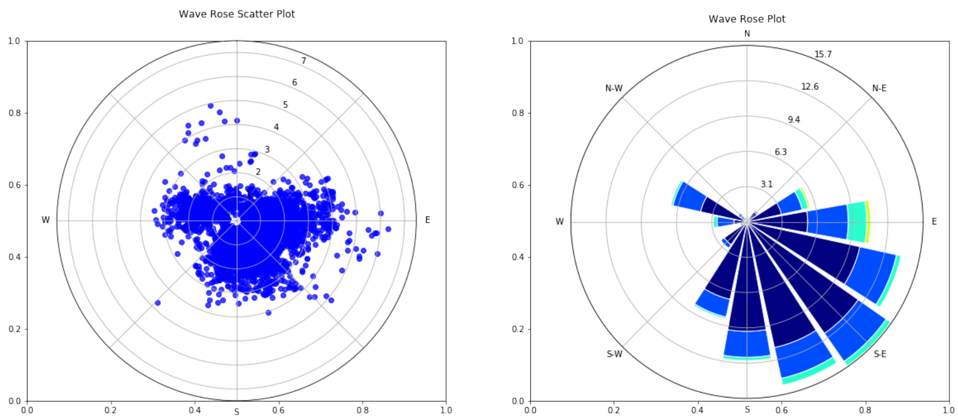


Figure 4: Wave Rose plot

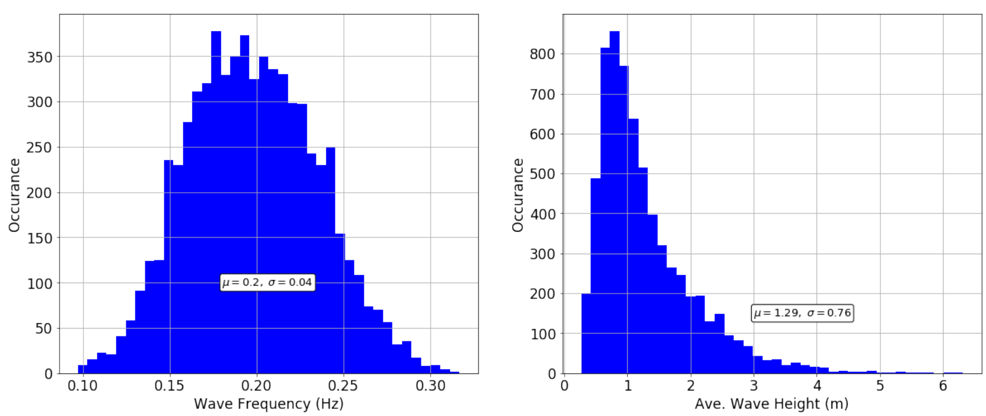


Figure 5: Frequency if occurrence for