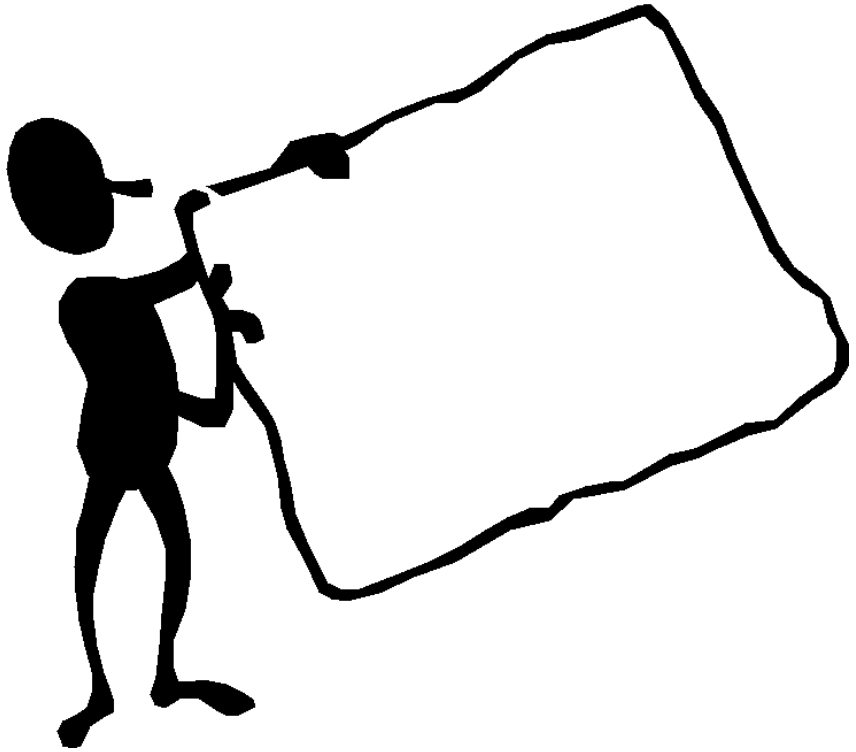
Study of AI & Machine Learning in Ocean Coast Applications Science

Cesar A. Rojas
Computer Science PhD Candidate
Florida International University
Summer 2021 – Fall 2022

Outline

- Personal Background
- Topic Introduction
- Research Motivation
- Related Work
- Early Approach
- Proposed Methods
- General Pipeline
- Results
- Publications
- References
- Acknowledgements



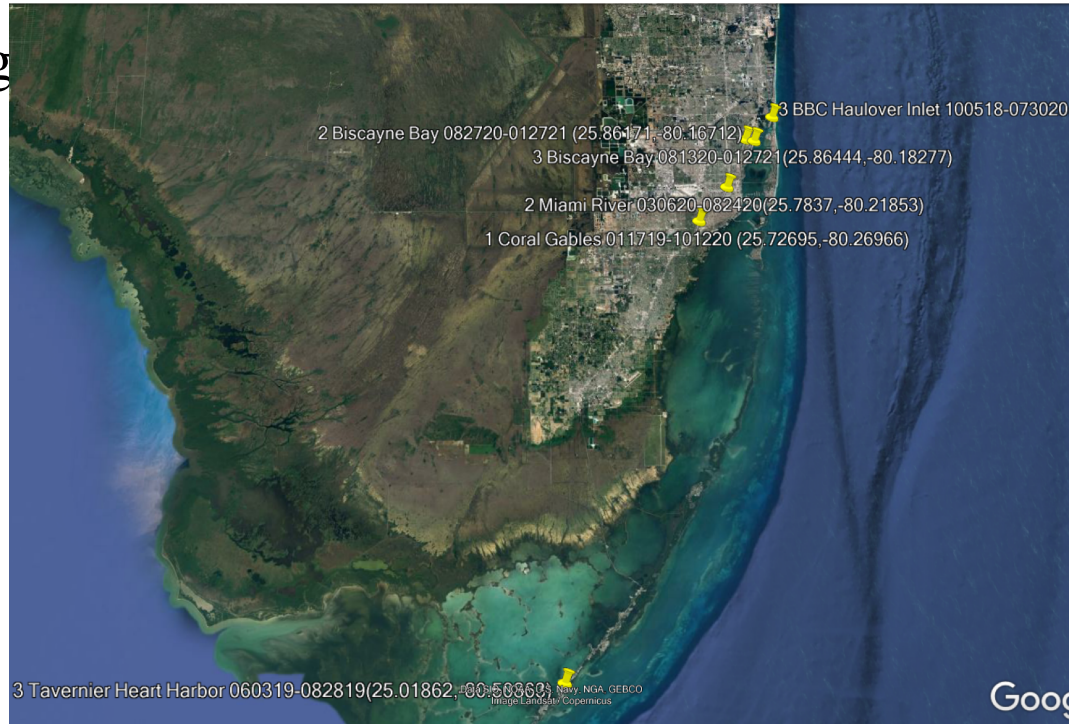
Background

- Third Year Computer Science PhD candidate.
- Former Microsoft Application Development consultant.
- GEM Fellow sponsored by NASA.
- Began researching applications of Secure Multi-party Computation in Robotics in 2020
- Began researching Remote Sensing Machine Learning applications in 2021.

https://www.linkedin.com/in/cesarrojas001

Introduction

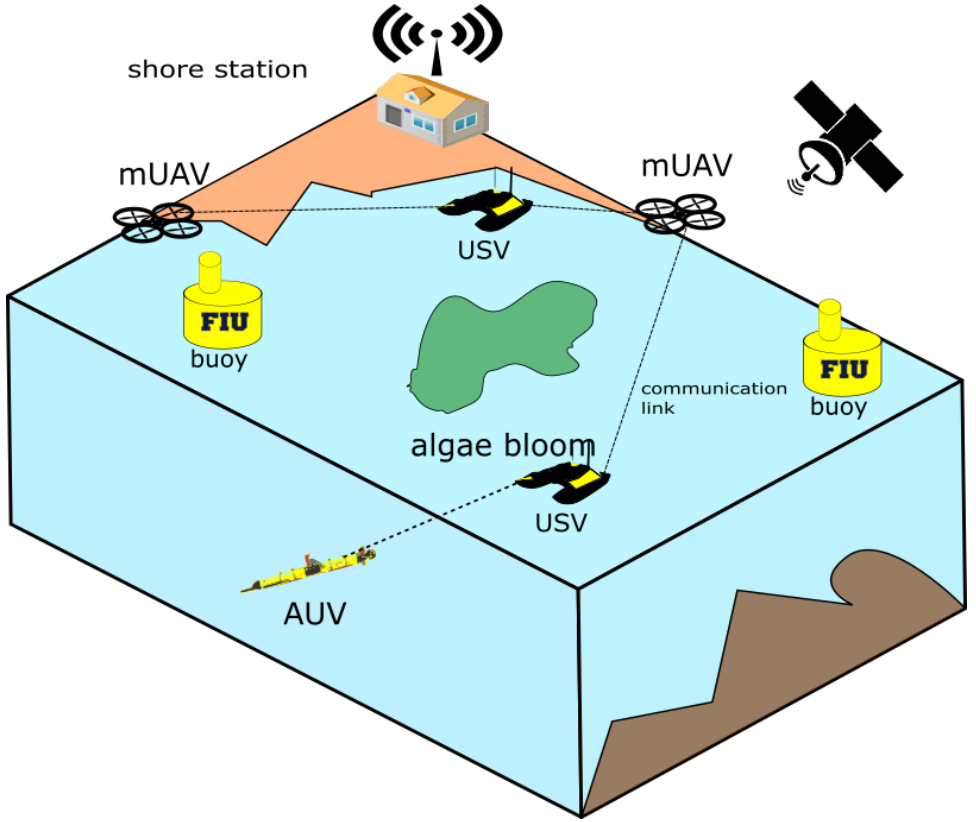
- Water quality measurement Machine Learning applications is a popular research area of interest.
- FIU Water sensing Buoys collecting data in coastal waters along Biscayne Bay.[1]
- Remote sensing data is publicly available and abundant.
- General purpose machine learning algorithms have been developed and researched.



Google Earth image of past and present FIU water buoy data collection missions.

Introduction

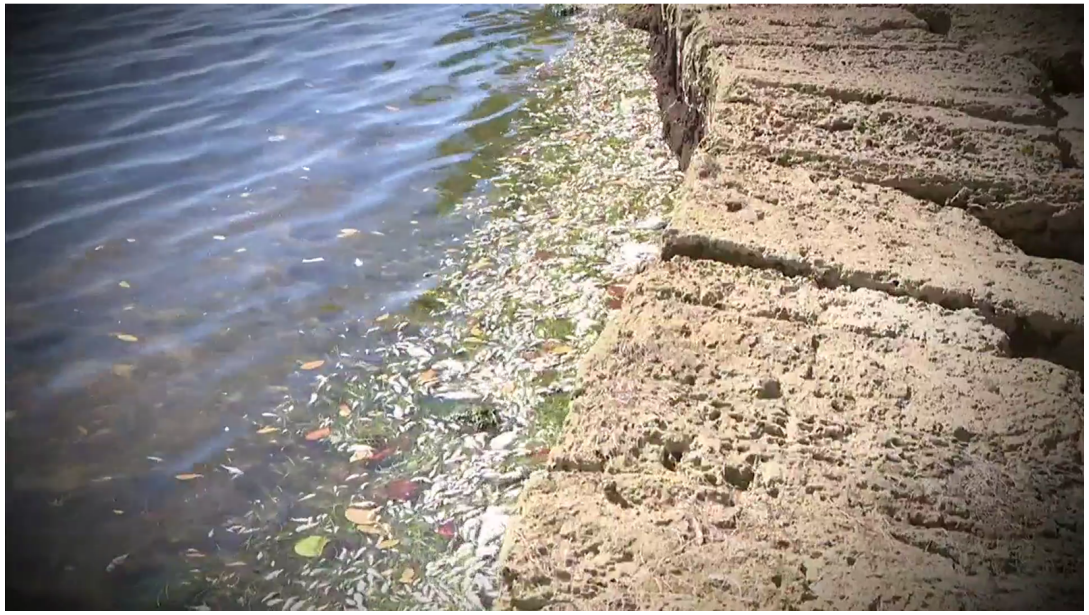
- Sensor data collection is carried out by specialized robotics system designed to collect data from land, air, and sea.
- Time series sensor data collection missions are expensive.
- Available data may only cover small area.
- It is precise but sparse.
- Enhance available data with remote sensing data.
- Remote sensing can provide guidance to which locations we should focus on.
- Part of a bigger agenda.



Data collection missions carried out by robotic systems.

Motivation

- Tampa Fish kill on August 2018 to November 2018 (4 months) resulted in 3+ million pounds and June 2021 to July 2021 (1.5 months) has same amount.[2]
- Biscayne Bay fish kill events in 2020, 2021, and 2022.[3]
- Harmful algal blooms (HABs) depends on wind direction and strength, temperature, nutrients, and salinity.[4]
- Water quality measurements and HAB forecasting allow for better disaster response and recovery.[5]



Dead fish and eels collecting in the Northern Biscayne Bay area of Miami, Florida.[3]

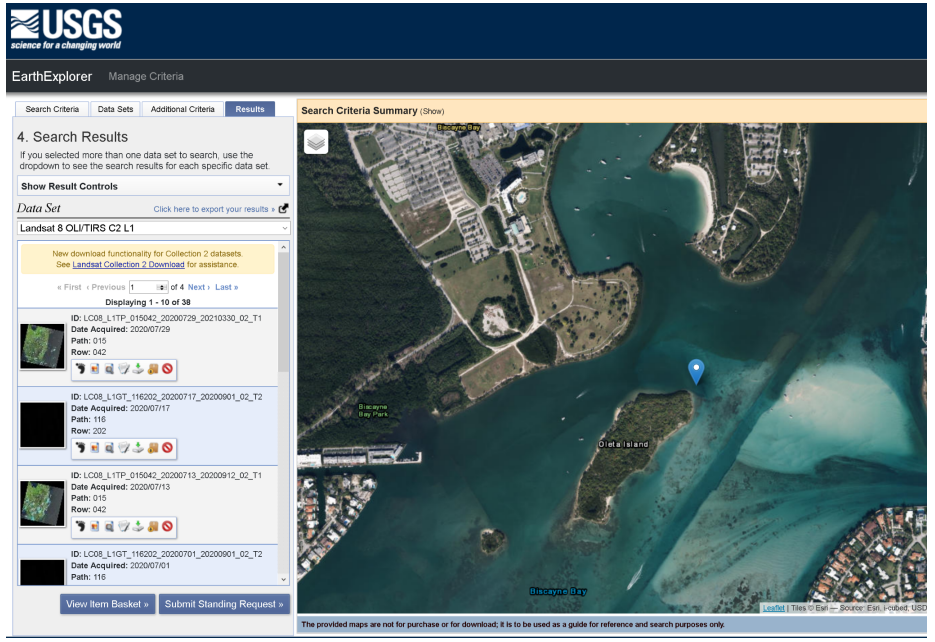
Related Work

- Landsat 8/9, Sentinel 2, and Sentinel 3 satellites are collecting valuable data and imagery used in agriculture, education, business, science, and government. .
- Remote sensing data for HAB detection and mitigation for Chesapeake Bay preservation has been studied [6]
- NASA launching Plankton, Aerosol, Cloud, & ocean Ecosystem (PACE) mission in 2023
- Remote sensing and in-situ data being combined for robot assisted space exploration has been studied.[7]

Artist concept of Landsat 8 satellite.[8]

Early Approach

- USGS Earth Explorer provides download access of Landsat and Sentinel images.
- Has bulk download capability but...
- Individual scenes are very large (850+MB)
- Downloaded as zip archives containing multiple GEOTIFF files.
- Each GEOTIFF file is processed by a python script using georasters library.
- Hard to create **time series** data.
- Too much data, only data from the location of buoy is needed.



Screenshot of browsing scenes of the Haulover inlet in Miami, Florida using USGS Earth Explorer.

Proposed Methods

- Develop a python script that:
  - Queries sentinel hub for remote sensing data in bulk.
  - Covers the timeline of all water buoy collection missions
  - Only for pixels associated with gathered in-situ data.
  - Stores downloaded data in a reusable format that can be analyzed, processed, and co-located with in-situ data.
  - For Landsat 8/9, Sentinel 2, and Sentinel 3
  - Resulting data-set can be used for machine learning applications.
  - Sentinel Hub Cloud API makes this possible.

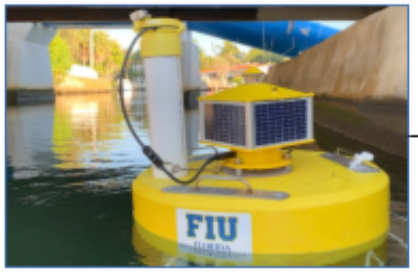
Sentinel Hub Cloud API can be accessed with python package sentinelhub-py

General Pipeline

Remote Sensing Data

sentinelhub

- Sentinel 2
- Sentinel 3
- Landsat 8



In-situ Data

CSV Data Files

Co-location module

Data Preparation

pandas

- Sampling
- Normalization
- Data Frames

Data Visualization

seaborn

- Scatter Plots
- Heat Maps

Regression Learning

scikit learn

- Nearest Neighbors
- Linear Regression
- Logistics Regression

Results

- Combining time series remote sensing data with in-situ data to create logistic regression model that can identify areas of interest for future data collection missions instead of waiting for fish kills to happen. (Published)
- Using remote sensing data as an additional sensor in an extended kalman filter for autonomous underwater vehicle localization. (Published)
- Generalized data management plan for projects that develop applications that combine remote sensing data and in-situ data. (Publication in development)

Publications

- Cesar A. Rojas, Gregory M. Reis, Arif R. Albayrak, Batuhan Osmanoglu, Leonardo Bobadilla, and Ryan N. Smith. ”*Combining Remote and In-situ sensing for Persistent Monitoring of Water Quality.*” In IEEE Oceans 2022 Chennai Conference Online Proceedings
- Cesar A. Rojas, Paulo V. Padrao, Jose E. Fuentes, Arif R. Albayrak, Batuhan Osmanoglu, Leonardo Bobadilla, and Ryan N. Smith. ”*Combining Remote and In-situ Sensing for Autonomous Underwater Vehicle Localization and Navigation*” In IEEE Oceans 2022 Hampton Roads Conference Online Proceedings

References

1. Florida International University, *Center for Aquatic Chemistry and Environment* [Source]
2. W. Lane, *Pinellas County already matches 2018 fish kill by cleaning up over 3 million pounds of dead fish*, 2021 [Source]
3. Miami Waterkeeper, *Fish Kill Press Releases*, 2022.[Source]
4. Brand L, Campbell L, and Bresnan E. *Karenia: The biology and ecology of a toxic genus*, 2012 [Source]
5. Hill P, Kumar A, Temimi M, and Bull D. *HABNet: Machine Learning, Remote Sensing-Based Detection of Harmful Algal Blooms*, 2020 [Source]
6. J. Wolny et al. *Current and Future Remote Sensing of Harmful Algal Blooms in the Chesapeake Bay to Support the Shellfish Industry*, 2020 [Source]
7. D. Thompson et al. *Spatio-spectral exploration combining in situ and remote measurements*, 2015 [Source]
8. NASA, *NASA-USGS Landsat 8 Satellite Celebrates First Year of Success*, 2014 [Source]
9. Cruz R, Costa P, Vinga S, Krippahl L, and Lopes M. *A review of recent machine learning advances for forecasting harmful algal blooms and shellfish contamination*, 2021 [Source]
10. USGS, *Reprocessing of Landsat 8 Data Acquired from November 2020 to February 2021 is Now Complete*, 2021 [Source]

Acknowledgments

- Project is supported in part by the ESA Network of Resources Initiative