JuliaEO - Global Workshop on Earth Observation with Julia 2024

Session:

POS2IDON: Pipeline for Ocean Features Detection with Sentinel-2

Andrea Giusti, Emanuel Castanho
AIR Centre

11.01.2024

AIRCENTRE

JuliaEO
Global Workshop on Earth
Observation with Julia 2024

What is Marine Debris

Marine debris is defined as any persistent solid material that is manufactured or processed and directly or indirectly, intentionally or unintentionally, disposed of or abandoned into the **marine environment**

Kikai et al., 2020 Photos: Careline Power

Plastics, metals, rubber, paper, textiles, and other lost or discarded items enter the ocean daily

One of today's greatest's pollution problems

Marine debris are a strongly related to the **problem of plastic pollution**, since most debris are made of plastic.

They are a threat to the Environment, Navigation safety, Economy and Human health.
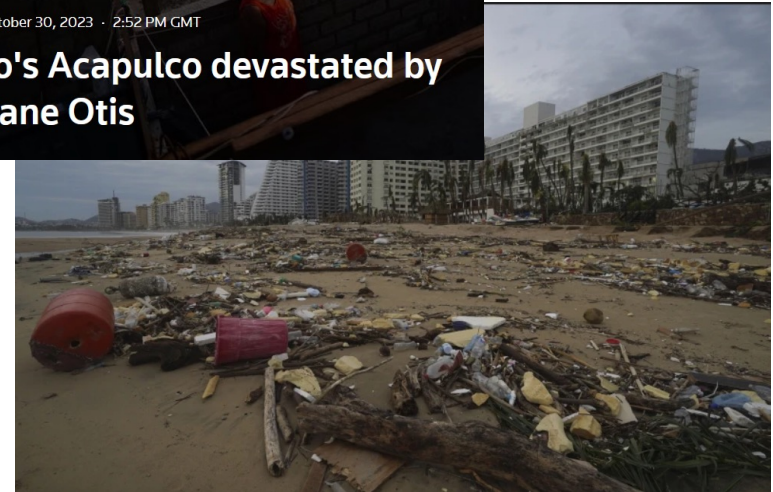
Flood events and other catastrophes can lead to great inputs into the ocean.

Pictures
33 Photos · October 30, 2023 · 2:52 PM GMT
Mexico's Acapulco devastated by Hurricane Otis

Video. Thousands of plastic bottles litter beach following catastrophic flood

AFP

Spain declares environmental emergency over plastic pellets spillage from ship

Satellite Technologies: Sentinel-2

Sentinel 2 provides a revolutionary way to study this...

- free-of-charge

- provides a trade between spatial resolution (10–20 m) and

- revisit frequency (5 days) for most coastal waters.

- has a global coverage.

- Multi-Spectral Instrument covers wavelengths of vis-NIR-SWIR, suitable for detecting and differentiating small floating matters.
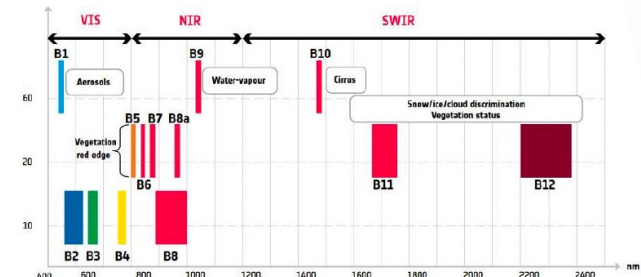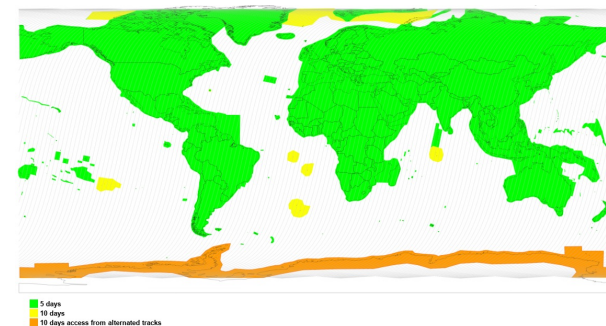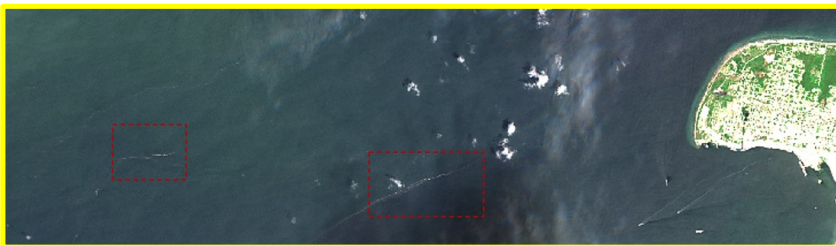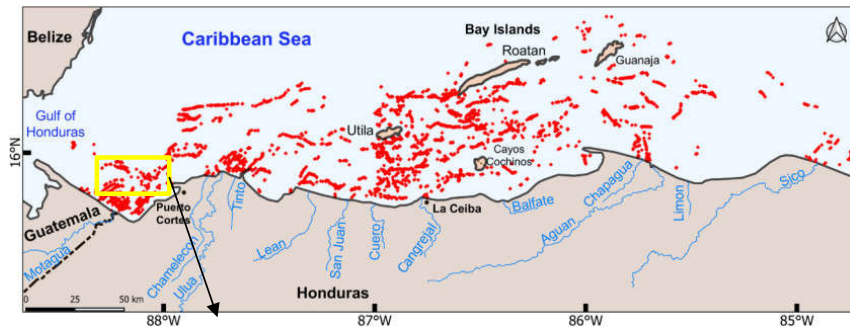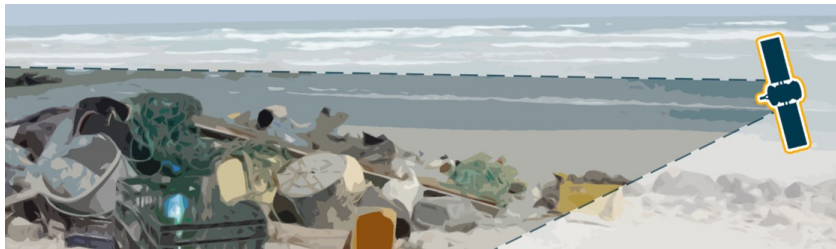

Image Credits: ESA

- monitor **larger areas**, as well as **remotely, hard-to-reach** areas

- provide **uniform coverage, regular and longer temporal coverage** and **harmonized data collection**

- support **effective monitoring** for management and remediation policies.

- be a **cost-effective and replicable solution**

Image Credits: European Space Agency, Katerina Kikaki et.al,2020

Ocean Features Detection

Marine Debris are typically seen as "bright" features...

Date: 18/04/2019
Target configuration:
Single target 5x20 m (5x10 m PET and 5x10 m LDPE)
Plastic Coverage: 100%

Date: 03/05/2019
Large targets configuration:
Four 5x5 m targets; two with PET and two with LDPE
Plastic Coverage: 100%

Topouzelis et al., 2020

Kikai et al., 2020 Photos: Careline Power

Marine Debris

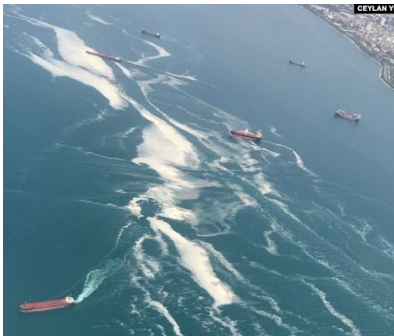Macroalgae Sargassum

Macroalgae Ulva

Microalgae Noctiluca

Microalgae Cyanobacteria

Sea Snot

Jellyfish

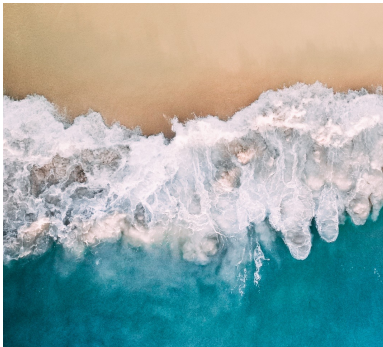Sea Pumice

Shrimp Eggs

Tree Pollen

But there are various other "floating features"...

JuliaEO
Global Workshop on Earth Observation with Julia 2024

AIRCENTRE

and more other features in Sentinel-2...

MARIDA - Marine Debris Archive

dataset based on the multispectral Sentinel-2 (S2) satellite data, which distinguishes Marine Debris from various marine features that co-exist, including Sargassum macroalgae, Ships, Natural Organic Material, Waves, Wakes, Foam, dissimilar water types (i.e., Clear, Turbid Water, Sediment-Laden Water, Shallow Water), and Clouds.

MARIDA: Marine Debris reports across different countries and continents for the period 2015–2021.

https://marine-debris.github.io/

MARIDA: A benchmark for Marine Debris detection from Sentinel-2 remote sensing data

Katerina Kikaki[1,2]*, **Ioannis Kakogeorgiou**[1], **Paraskevi Mikeli**[1], **Dionysios E. Raitsos**[3], **Konstantinos Karantzalos**[1,4]

1 Remote Sensing Laboratory, National Technical University of Athens, Athens, Zografou, Greece,
2 Institute of Oceanography, Hellenic Centre for Marine Research, Athens, Anavyssos, Greece,
3 Department of Biology, National and Kapodistrian University of Athens, Athens, Zografou, Greece,
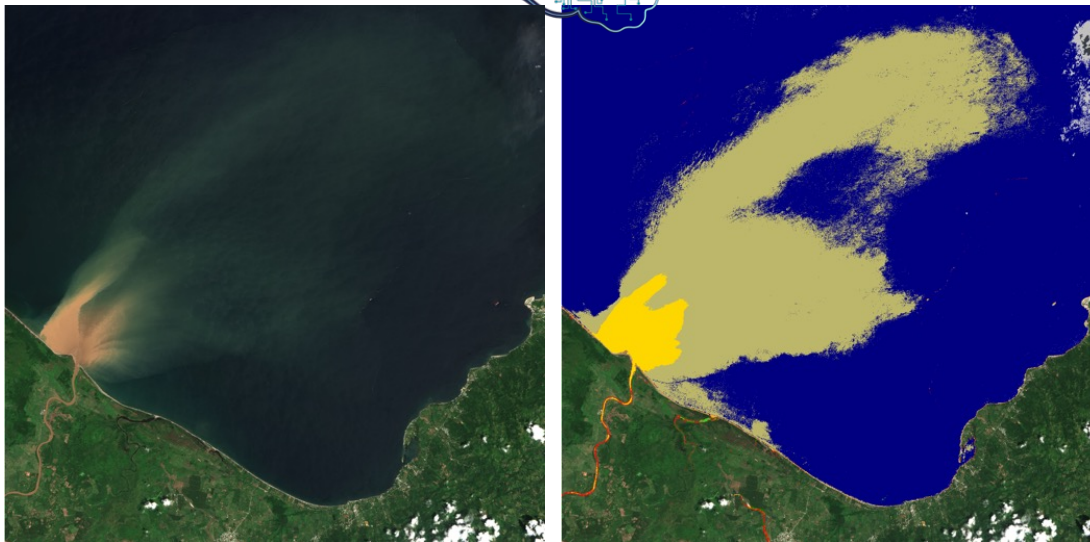**4** Athena Research Center, Athens, Greece

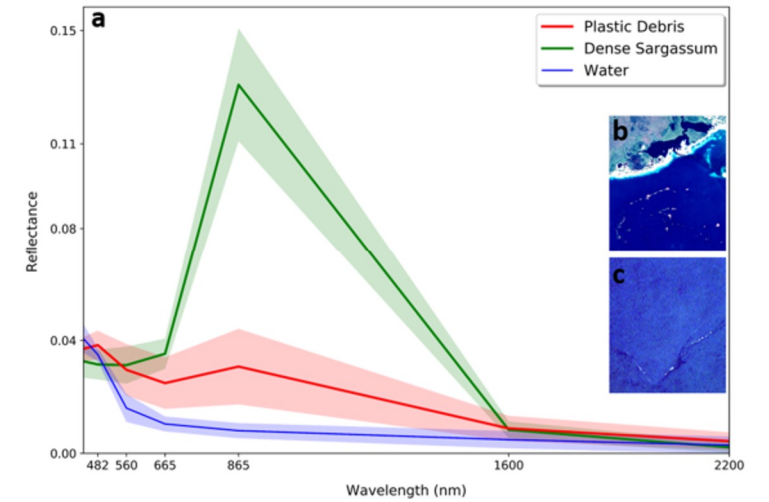* akikakh@central.ntua.gr
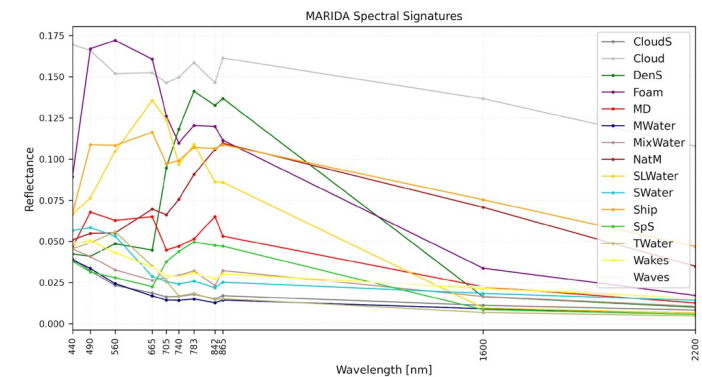
Different spectral signatures for ocean water, floating vegetation (e.g., sargassum, organic matter), floating plastics.

Machine learning techniques can help the discrimination of floating plastic debris from other features.

Ocean water – low reflectance
Floating Plastics – high reflectance w/ flat spectrum
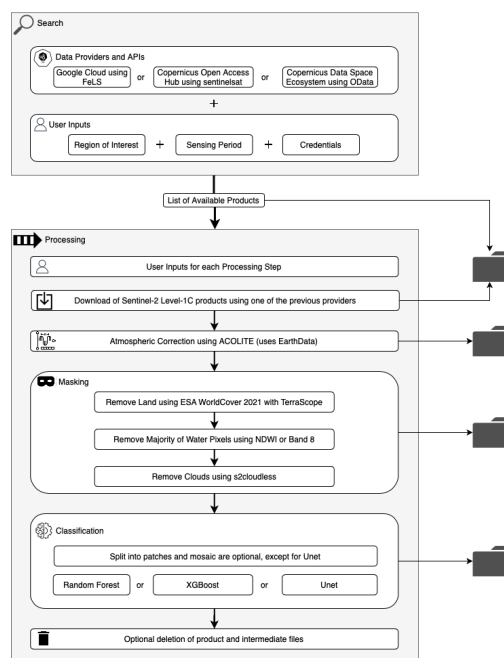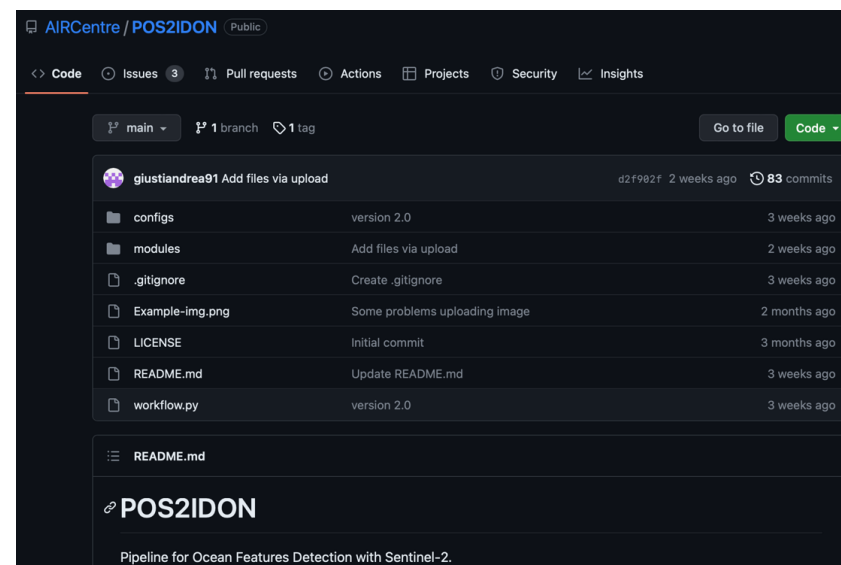**Floating Vegetation** – high reflectance w/ red edge

POS2IDON: Description

POS2IDON is an open-source data pipeline based on Sentinel-2, developed by AIR Centre, for long-term analyses and monitoring of **suspected marine debris accumulations (>10 m)** and **other ocean features**, such as floating organic material.

POS2IDON - Pipeline for Ocean Features Detection with Sentinel-2

POS2IDON: Description

Search and Download

Region of Interest	Sensing Period	COAH/CDSE Credentials	Other	**User Inputs**

↓ Search Sentinel-2 Level 1C

Google Cloud	Copernicus Open Access Hub (COAH)	Copernicus Data Space Ecosystem (CDSE)	**Free Data Provider Platforms**
FeLS https://github.com/vascobnunes/ fetchLandsatSentinelFromGoogleCloud	**sentinelsat** https://github.com/sentinels at/sentinelsat	**OData** AIR Centre developed function based on OData	**Python Interfaces**

↓ Download Products

AIR CENTRE

JuliaEO
Global Workshop on Earth
Observation with Julia 2024

POS2IDON: Description

Atmospheric Correction

Top of Atmosphere
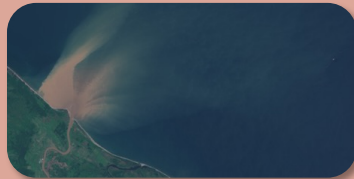
Rayleigh Atmospheric Corrected
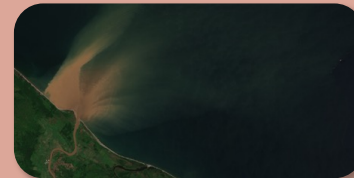
ACOLITE
https://github.com/
acolite/acolite

Includes specific inputs
that can be changed
by more advanced
users inside the
workflow .py file

POS2IDON: Description

Land, Water Masking

Remove Land and leave Water where the classification will be performed. Uses ESA WorldCover 2021.

Remove some water pixels depending on thresholds to improve classification time. Uses Band 8 or Normalized Difference Water Index (NDWI).

MASKS

POS2IDON: Description

Cloud Masking
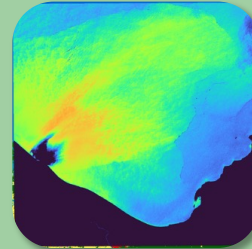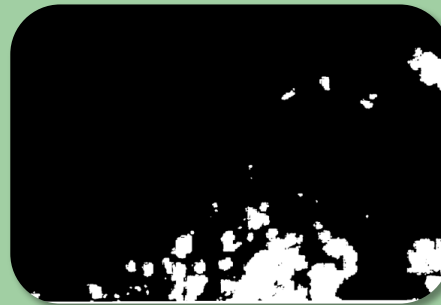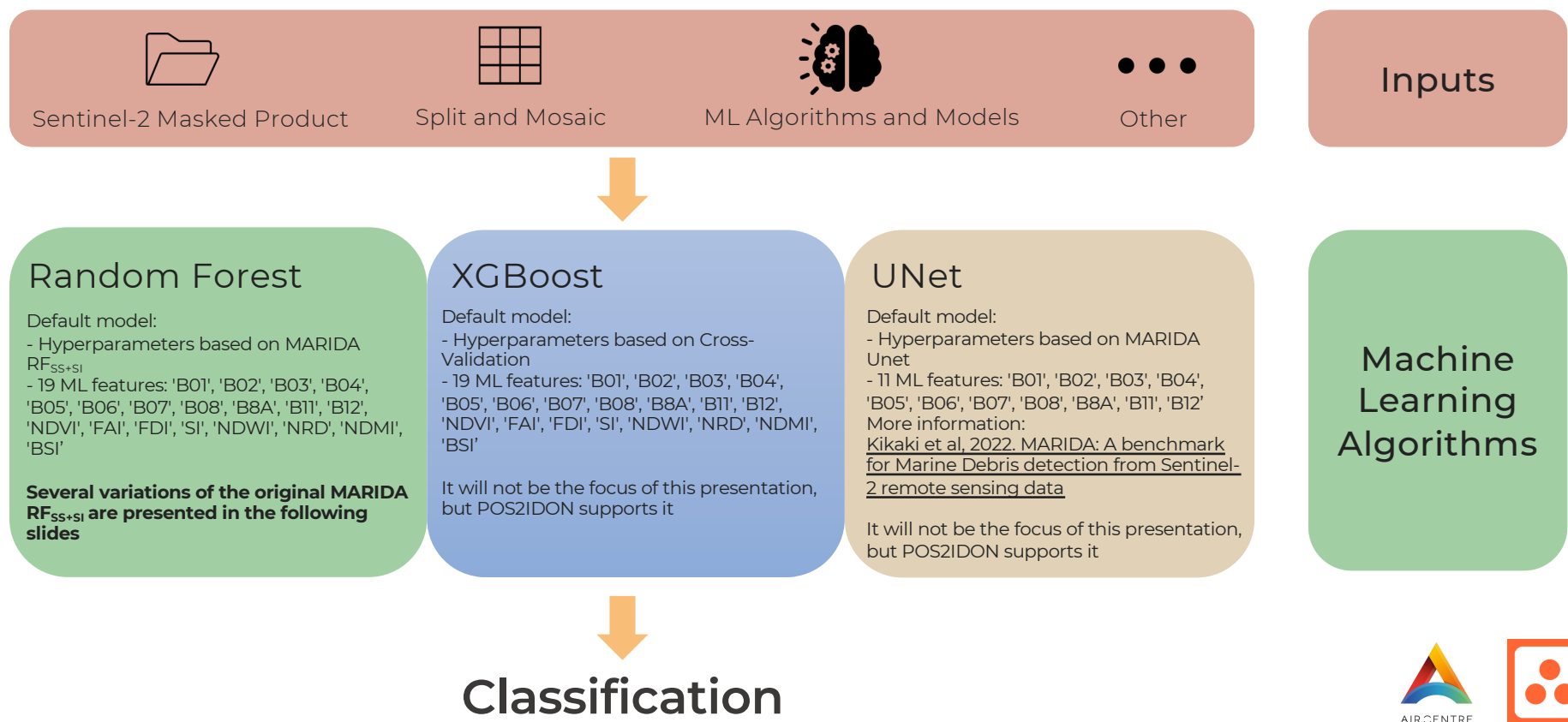
Remove Clouds using s2cloudless (https://github.com/sentinel-hub/sentinel2-cloud-detector).

MASKS

POS2IDON: Description

Sentinel-2 Masked Product	Split and Mosaic	ML Algorithms and Models	• • • Other	**Inputs**

Random Forest

Default model:
- Hyperparameters based on MARIDA RF$_{SS+SI}$
- 19 ML features: 'B01', 'B02', 'B03', 'B04', 'B05', 'B06', 'B07', 'B08', 'B8A', 'B11', 'B12', 'NDVI', 'FAI', 'FDI', 'SI', 'NDWI', 'NRD', 'NDMI', 'BSI'

Several variations of the original MARIDA RF$_{SS+SI}$ are presented in the following slides

XGBoost

Default model:
- Hyperparameters based on Cross-Validation
- 19 ML features: 'B01', 'B02', 'B03', 'B04', 'B05', 'B06', 'B07', 'B08', 'B8A', 'B11', 'B12', 'NDVI', 'FAI', 'FDI', 'SI', 'NDWI', 'NRD', 'NDMI', 'BSI'

It will not be the focus of this presentation, but POS2IDON supports it

UNet

Default model:
- Hyperparameters based on MARIDA Unet
- 11 ML features: 'B01', 'B02', 'B03', 'B04', 'B05', 'B06', 'B07', 'B08', 'B8A', 'B11', 'B12'
More information:
Kikaki et al, 2022. MARIDA: A benchmark for Marine Debris detection from Sentinel-2 remote sensing data

It will not be the focus of this presentation, but POS2IDON supports it

Machine Learning Algorithms

Classification

POS2IDON: Models

Class	RF			XGBoost		
	IoU	Rec	F1	IoU	Rec	F1
MD	0.68	0.83	0.81	0.73	0.84	0.84
DenseS	0.92	0.95	0.96	0.93	0.96	0.96
SpS	0.82	0.90	0.90	0.85	0.90	0.92
NatM	0.73	0.80	0.85	0.75	0.81	0.86
Ship	0.69	0.75	0.82	0.76	0.81	0.86
Clouds	0.93	0.96	0.96	0.97	0.98	0.98
MWater	0.94	0.97	0.97	0.97	0.99	0.98
SLWater	1.00	1.00	1.00	1.00	1.00	1.00
Foam	0.59	0.74	0.74	0.66	0.78	0.80
TWater	0.99	1.00	1.00	0.99	1.00	1.00
SWater	0.93	0.96	0.96	0.96	0.97	0.98
Cyanob	1.00	1.00	1.00	1.00	1.00	1.00
Noctil	0.99	0.99	0.99	0.99	1.00	1.00
Phaeoc	0.86	0.94	0.92	0.91	0.96	0.95
Average	0.86	0.91	0.92	0.89	0.93	0.94

Random Forest and XGBoost ML models show satisfactory metrics. An improved U-NET model is under testing.

Models were trained with an extended version of the MARIDA spectral library, totalizing 14 classes, including Marine Debris, Floating Macroalgae, Ships, Foam, Clear and Turbid Water and Phytoplankton Blooms.

However, challenges are always present in unseen data...

Many other reflective/bright features like plastics exist, foam, clouds, ships. Difficult to differentiate between these features
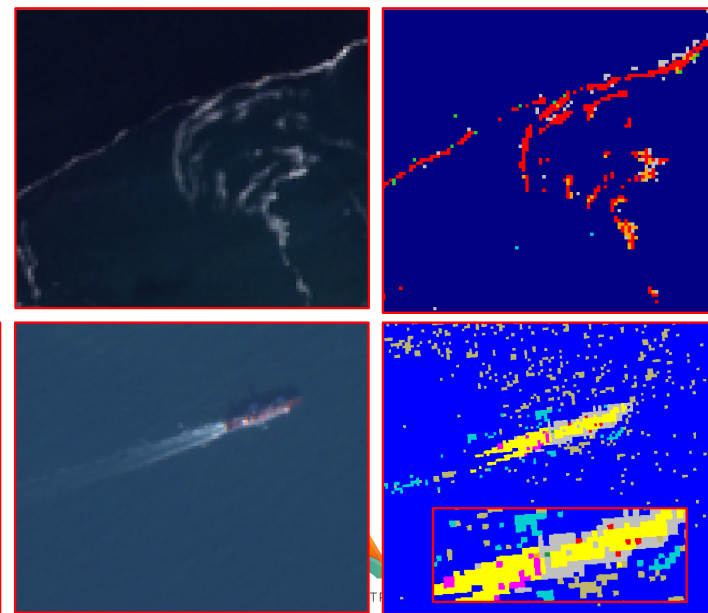
- Marine Debris
- Dense Sargassum
- Sparse Sargassum
- Natural Organic Material
- Ship
- Clouds
- Foam

Julia UNET integration into Python

- The UNET model was developed with the Flux.jl package.

- The integration of the model structure into POS2IDON main code in Python is assured by JuliaCall from PythonCall.jl.

- This allows the POS2IDON user to have more machine learning model options and take advantage of Julia's processing capabilities.
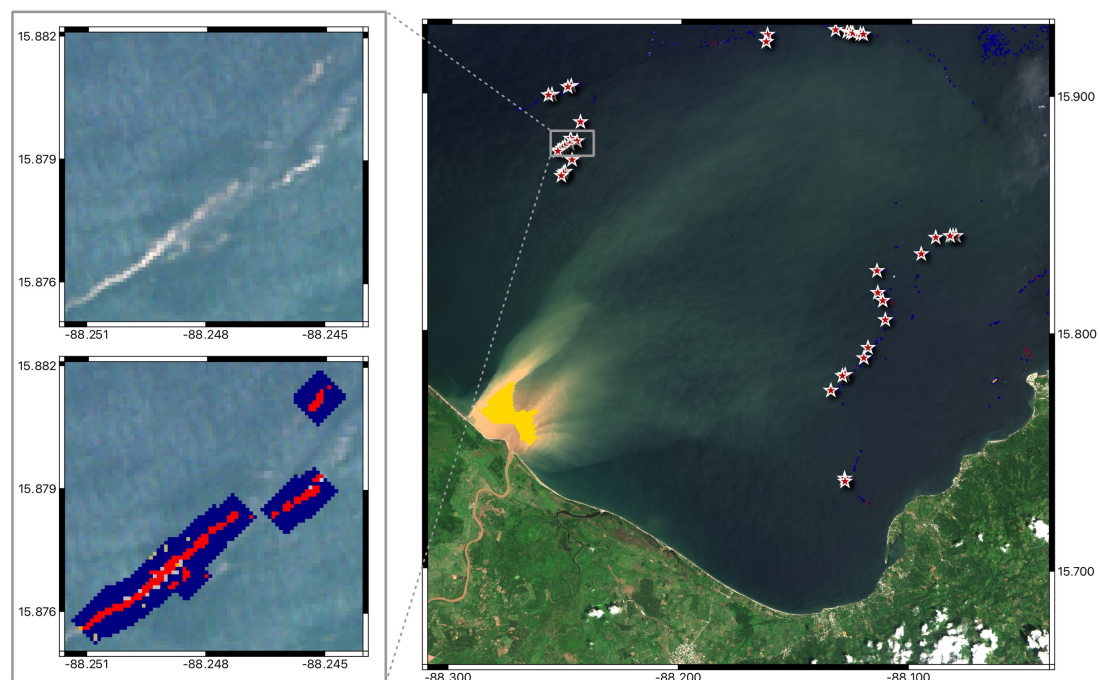
POS2IDON: Applications (1)

Post-disaster management after flooding events in heavily polluted regions to direct clean-up activities and assess the input of debris into the ocean. Next figure shows POS2IDON applied to a major plastic debris event in Honduras Gulf in 2020/09/18.

Likely presence of MD (red pixels and dots) detected along the river front. Red stars with white borders are clusters of at least 10 MD pixels at 100 meters distance.

POS2IDON: Applications (2)

Long-term seasonal analysis taking advantage of regular Sentinel-2 imagery (5 days) since 2018, useful to better understand patterns and trends in marine debris. POS2IDON is used to analyse Honduras Gulf for all 2020 (72 images).

Dryer seasons (winter and spring) have less debris detection, in agreement with rivers being the major source of pollution.
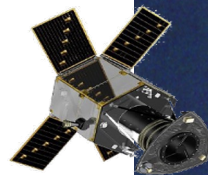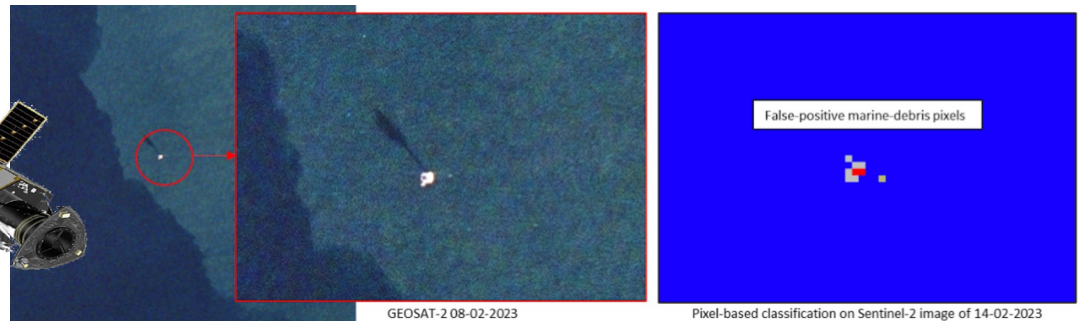
Winter (Dec-Feb) Spring (Mar-May)

Summer (Jun-Aug) Fall (Sep-Nov)

mm/h — 2020
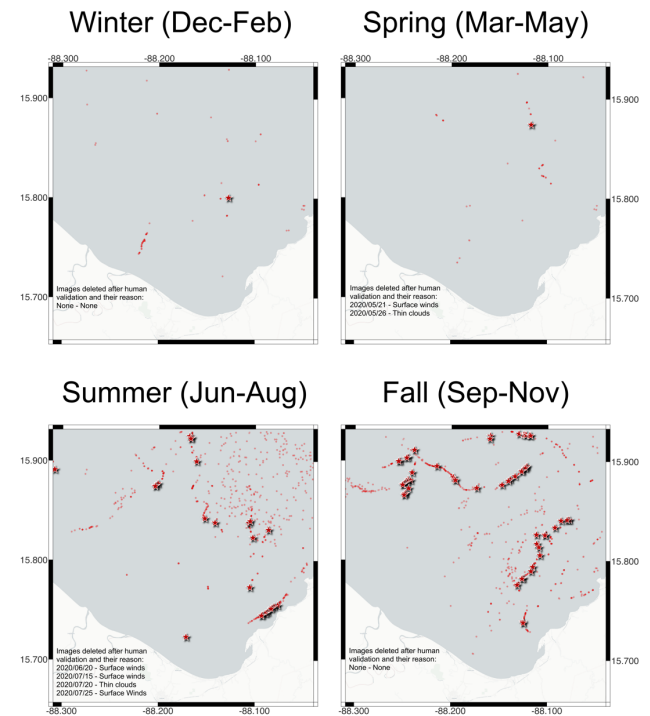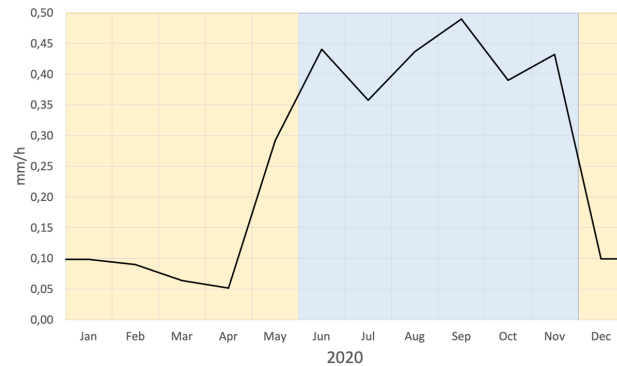
Improvements since last JuliaEO

Events in Literature and News

GEOSAT-2 very-high resolution

False-positive marine-debris pixels

GEOSAT-2 08-02-2023

Pixel-based classification on Sentinel-2 image of 14-02-2023

ALKOR, Geomar

2023-06-30 ALKOR in-situ

2023-06-30 GEOSAT-2 RGB

Field Campaign June 2023 (LABPLAS project)

JuliaEO
Global Workshop on Earth
Observation with Julia 2024
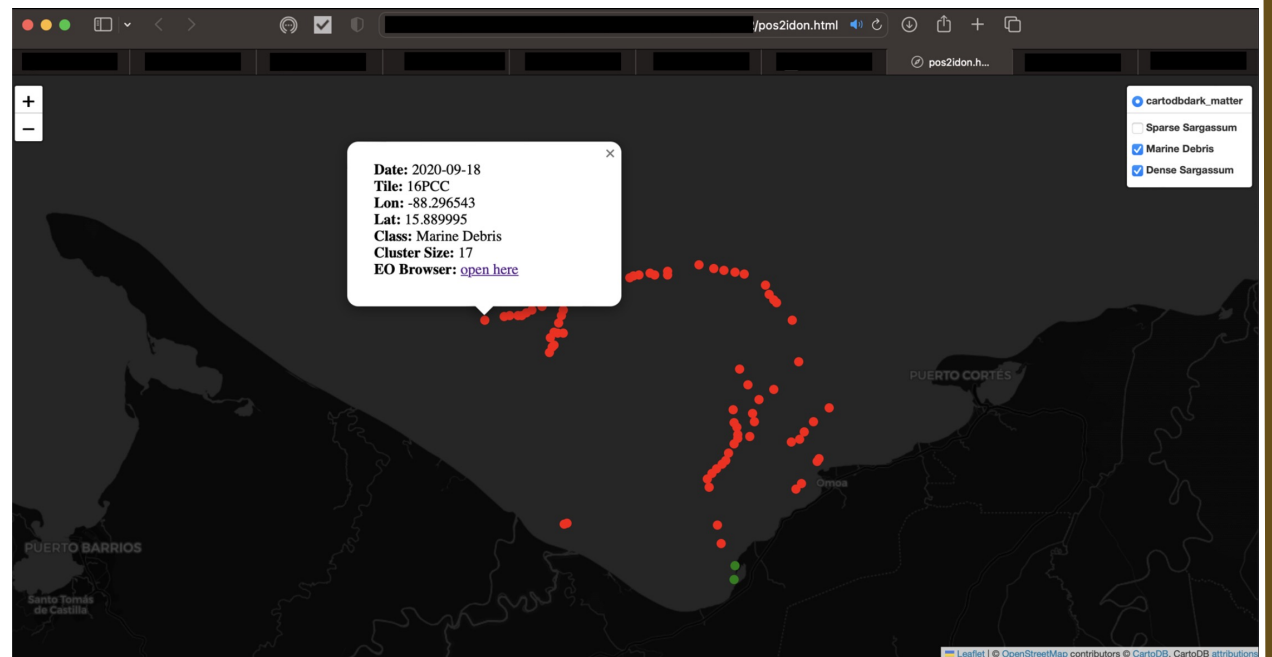
Improvements since last JuliaEO

- Support for the new Copernicus Data Space Ecosystem for downloading Sentinel-2 products.

- From JuliaEO23 a collaboration start to train a UNET model in Julia to improve GPU classification speed.

- Integration of the UNET model developed in Julia into the POS2IDON source code in Python;

- Web app prototype to show POS2IDON results (using Python folium, **how can Julia facilitate this web development?**);

Improvements since last JuliaEO

- Visualization and analysis of long-term data with density and occurrence maps.

- GEOTIF may not be the most efficient way to save POS2IDON results during long-term analysis (**alternative formats?**).

- Training new models with an expanded spectral library containing more classes such as algae blooms. However, there is still a problem of lack of in-situ data.
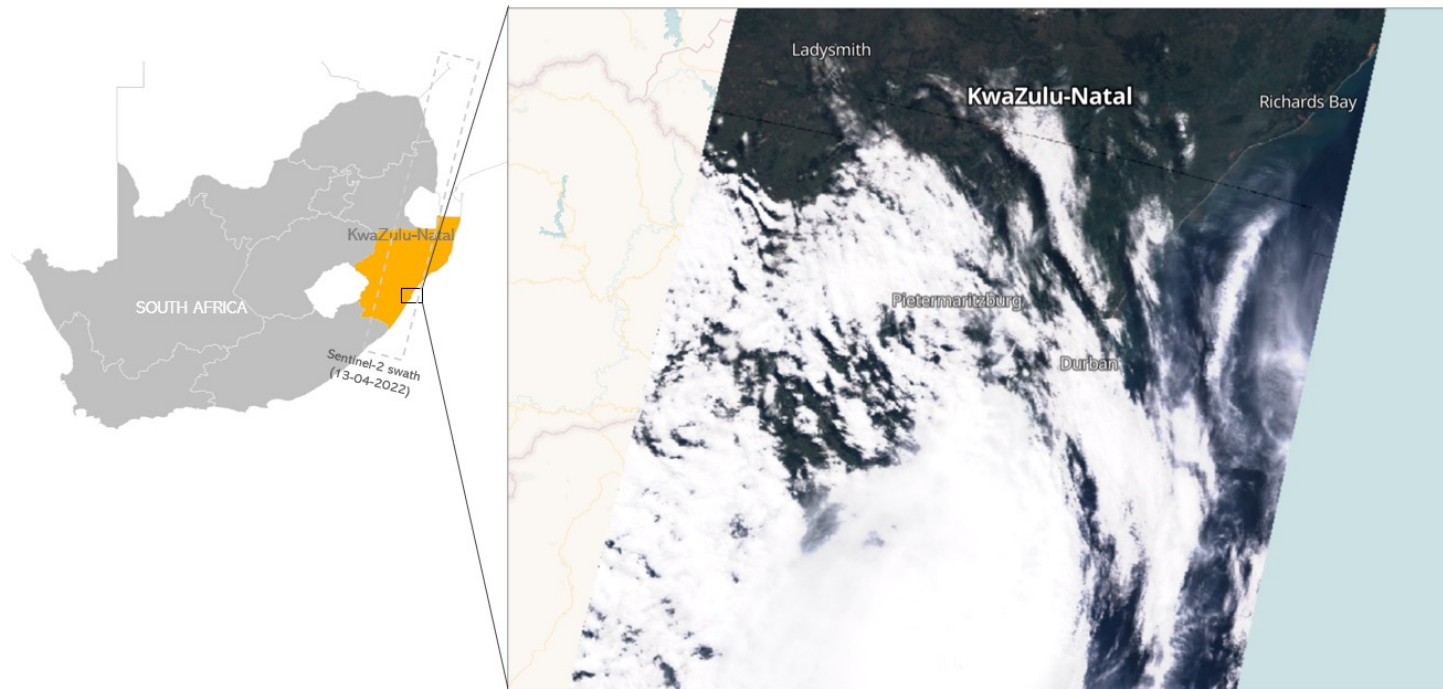

Occurrency Map - Marine Debris
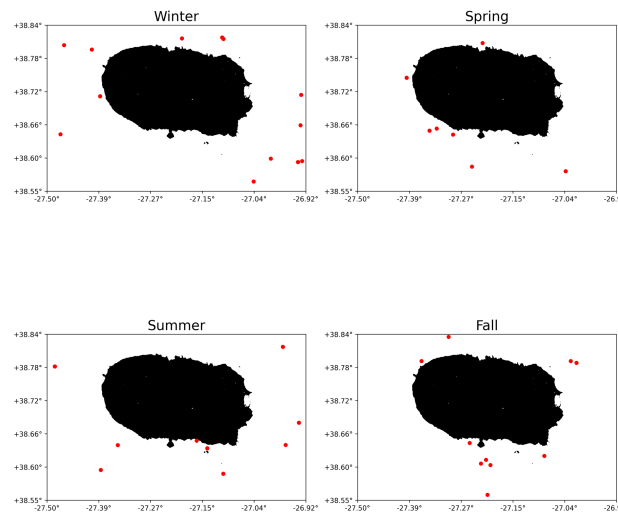
Accumulation Map - Marine Debris
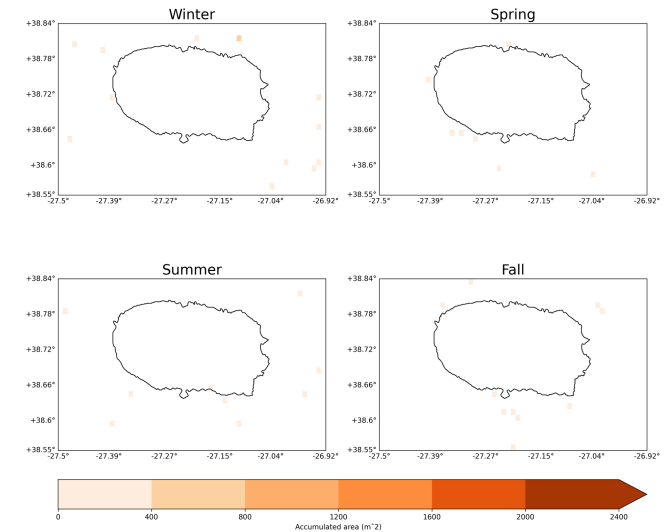
Introduction to Case Study – South Africa Floods 2022

Our case study is an event of litter discharged to the sea following the April 2022 catastrophic floods in KwaZulu Natal, South Africa.

Sentinel-2 RGB image on 13-04-2022

Plastic pollution left behind on Durban's beaches after the April 2022 floods.
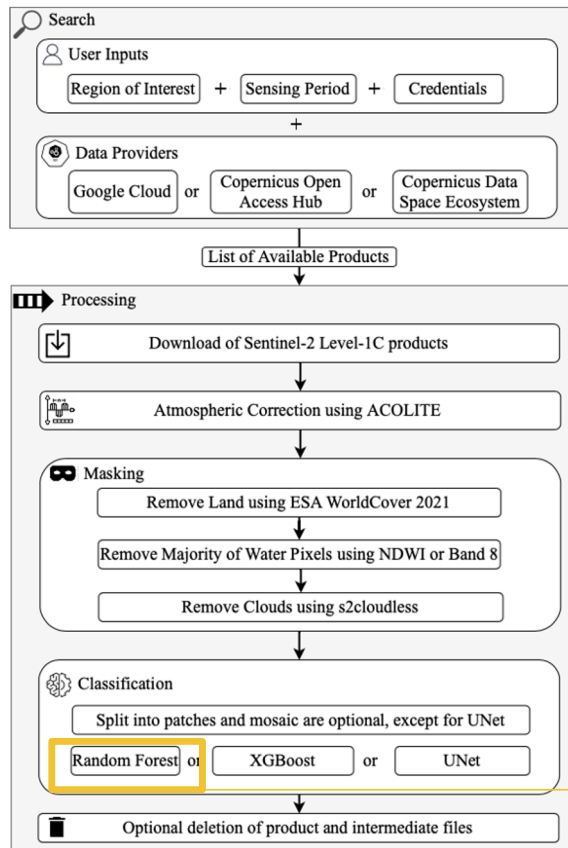
What will be covered?

The Hands-on session is a simplified version of POS2IDON.

You can access the training material and instructions at:

AIRCentre > JuliaEO24 > notebooks > **pos2idon_results_demo**

In addition, a document with basic instructions was sent to you.

POS2IDON architecture.

Focus on this Hands-on session

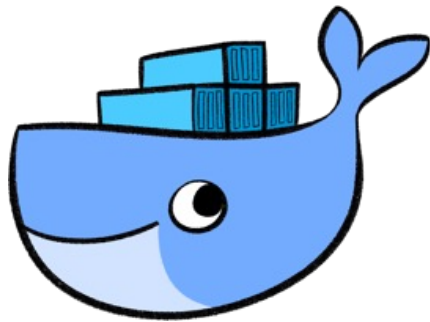Visualization

Classification with RF
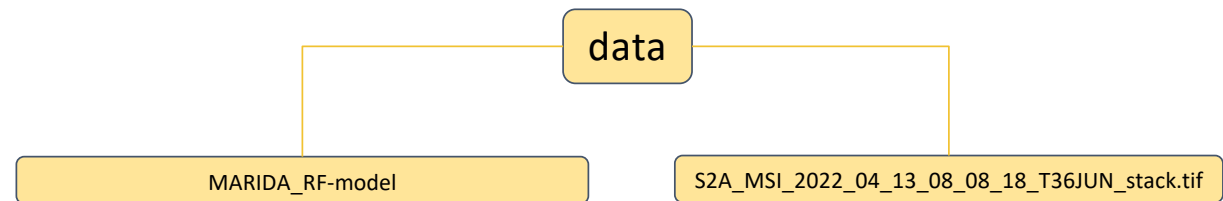
Some questions you might have

Before we start the Hands-on session, here are some important questions answered:

What is Docker?

Docker enables the creation of lightweight, isolated containers for applications and their dependencies. This is valuable because it ensures **consistent environments** and **offers portability**. It streamlines this kind of hands-on sessions by ensuring all participants have a uniform and manageable environment to run a specific project.

Where does the data used in the exercise come from?

```
                    data
         ┌───────────┴───────────┐
 MARIDA_RF-model      S2A_MSI_2022_04_13_08_08_18_T36JUN_stack.tif
```

Previous trained Random Forest model based on the MARIDA spectral signatures library.

Recall for Marine Debris: **91%**

Measures how good your model finds all the positives.

$$Recall = \frac{TP}{TP + FN}$$

Stack processed and outputted by **POS2IDON**.

Includes:

- 11 Sentinel-2 bands atmospherically corrected with ACOLITE and resampled to 10 meters.

- 8 Spectral Indices calculated from the previous bands.

AIRCENTRE

JuliaEO
Global Workshop on Earth
Observation with Julia **2024**

Let's go!