SISSA DATASCIENCE
Machine Learning for the Natural Sciences

S.W.A.M.P.

Scientific Wetlands Analysis and Monitoring Project*

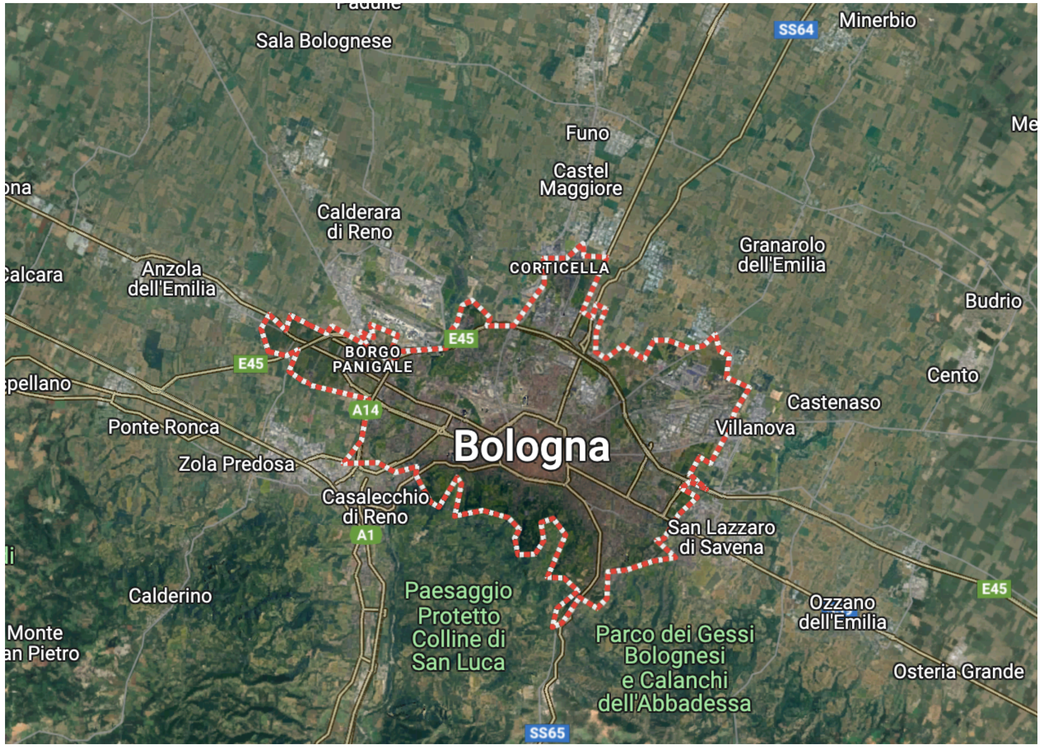
Roberto Trotta and Chiara Moretti
International School for Advanced Study (SISSA), Trieste

Climate Arena Conference, Bologna, Oct 18th 2024

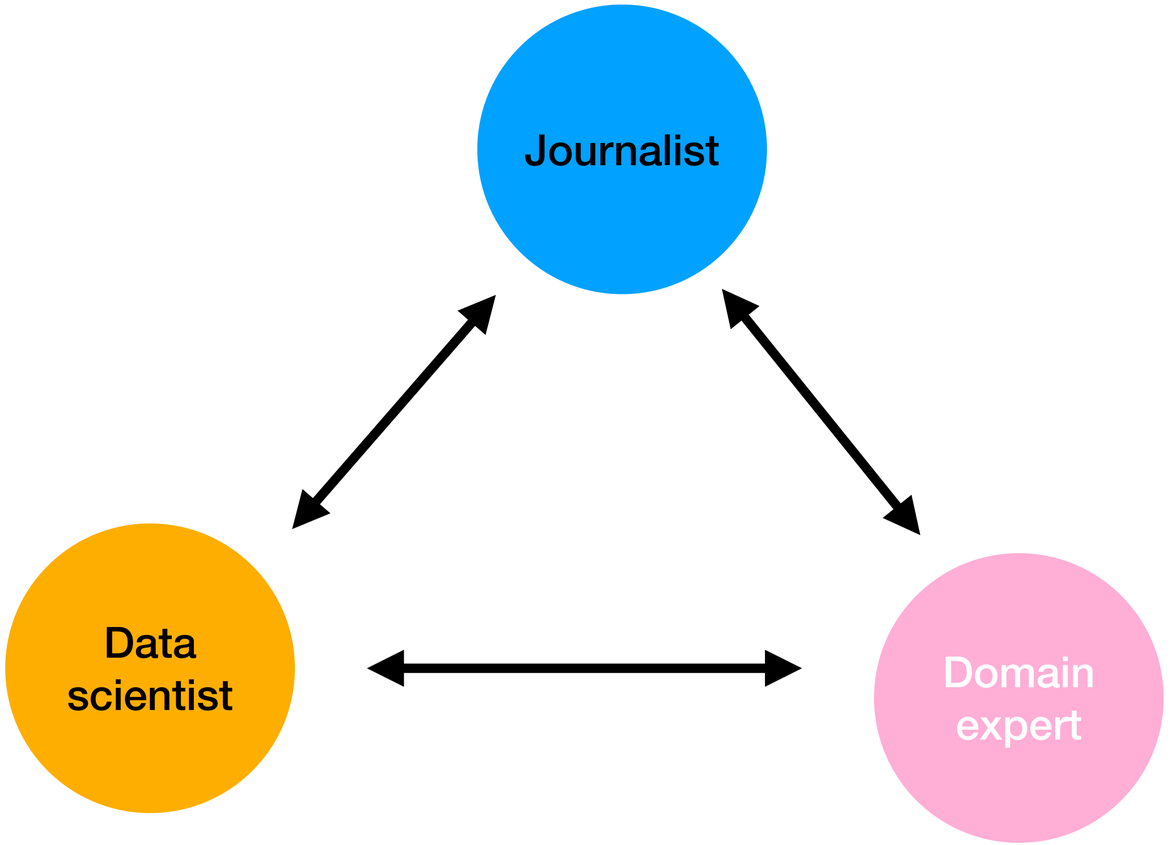
** Thanks, ChatGPT*

Chiara Moretti

“How difficult can it be?!”



The magic ingredients



Data Access

SISSA DATASCIENCE
Machine Learning for the Natural Sciences

Copernicus
Europe's eyes on Earth

The Copernicus Data Space Ecosystem

- Largest EO data offering in the world, with discovery and download capabilities
- A set of data processing tools to extract objective information and conduct public, private or commercial activities
- A thriving ecosystem to offer data, services and applications from public, commercial and scientific service providers
- A service to benefit institutional users, research, commercial sector as well as to every citizen of our planet

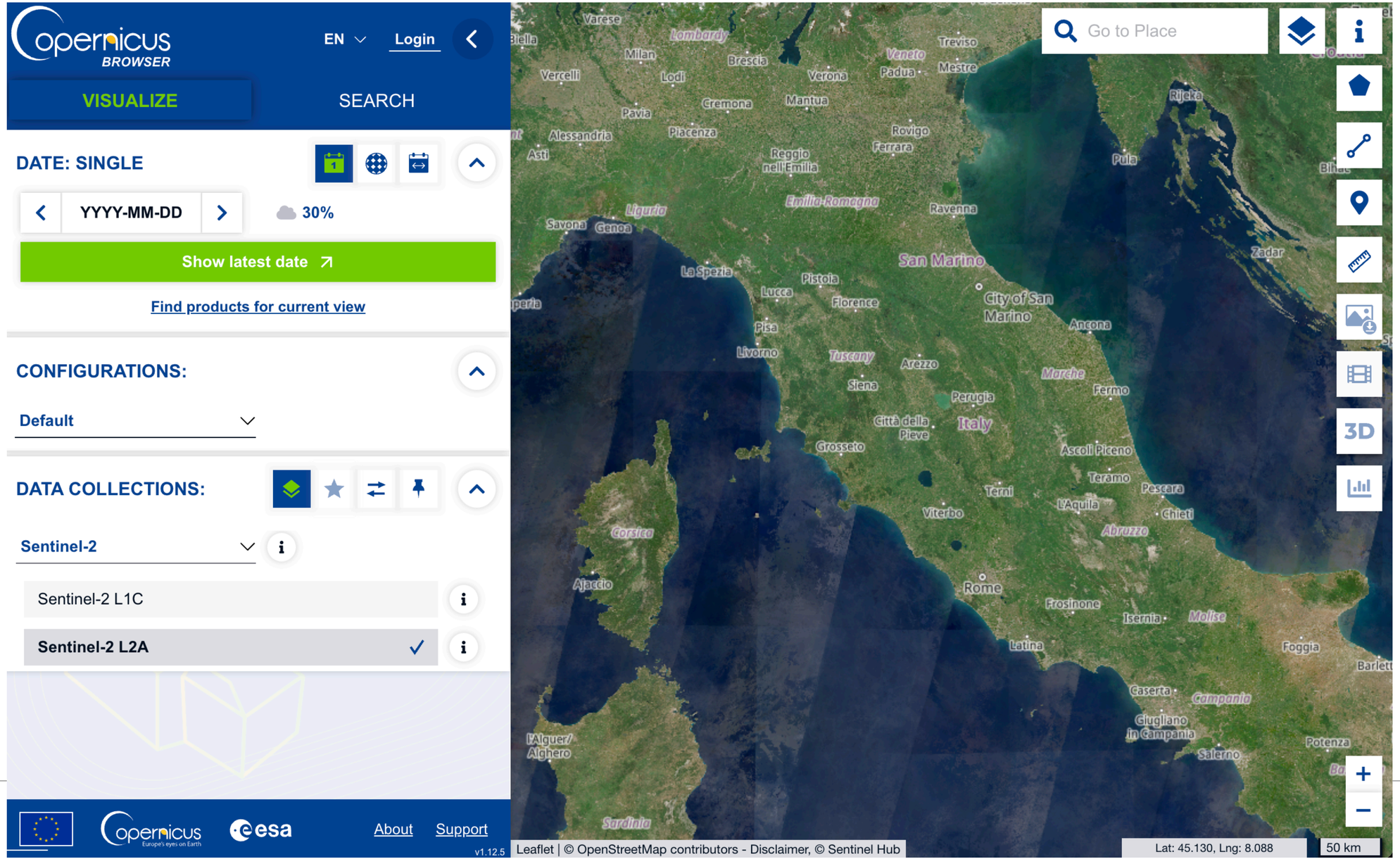
Driven by data

The primary goal of the service is to ensure instant data availability to users. The full data archive acquired by the Copernicus Sentinel satellites will be available online and can be accessed in real-time.

- Sentinel-1 SLC and GRD
- Sentinel-2 L1C and L2A
  (reprocessed to Collection 1, in parallel with results becoming available)
- Sentinel-3
- Sentinel-5P
- Copernicus Contributing Missions

The data will be available free of charge via designated quotes for individual use. Users that wish to build large scale operations can use practically unlimited resources available under commercial terms.

Browsing is for free… the rest, not



In the end, it took some hard DS skills…

SISSA DATASCIENCE
Machine Learning for the Natural Sciences

CREODIAS is therefore offering commercial cloud services in the ecosystem on top of this full archive of data. This allows for extended scaling, performance and flexibility on top of the free and open offering.

1 month free trial; then you can apply for 5,000€ worth of credits as a researcher.

OUR TOOL:

Python code to interact via API + automated computation of multi-band indexes for water presence and water turbidity

Able to specify target region via geojson cut-out

Able to specify date range (currently limited to 2018 onwards, using Sentinel; Landsat goes back to 2000)

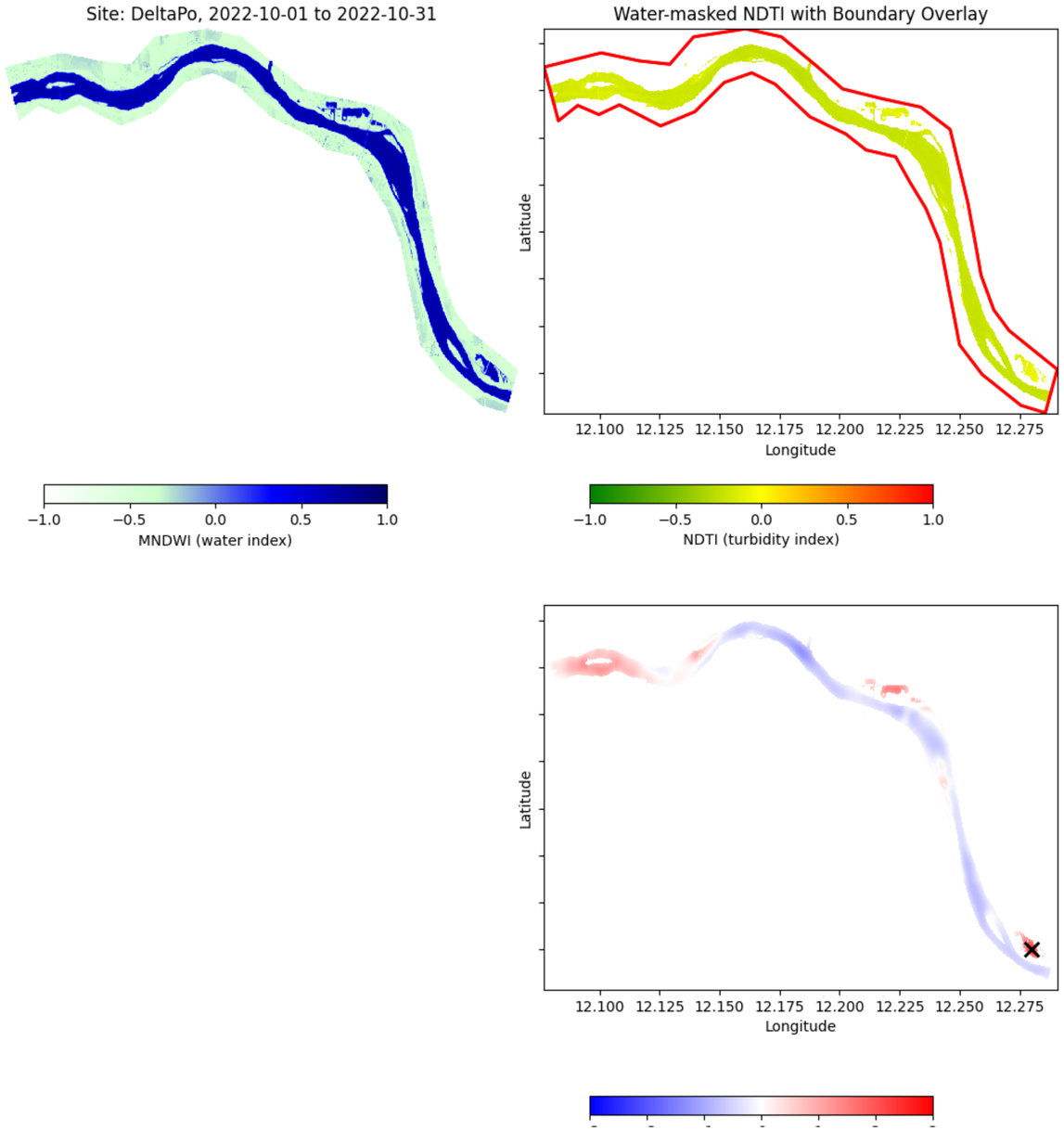
In [ ]:

```
# Define the request parameters for SCL data
def create_request(date_range):
    return SentinelHubRequest(
        data_folder='./temp_data/',
        evalscript="""
            //VERSION=3
            function setup() {
                return {
                    input: ["B03", "B04", "B11", "SCL"],
                    output: [
                        { id: "bands", bands: 3, sampleType: SampleType.FLOAT32 },
                        { id: "scl", bands: 1, sampleType: SampleType.UINT8 }
                    ]
                };
            }
            function evaluatePixel(sample) {
                return {
                    bands: [sample.B03*10000, sample.B04*10000, sample.B11*10000],
                    scl: [sample.SCL]
                };
            }
        """,
        input_data=[
            SentinelHubRequest.input_data(
                data_collection=DataCollection.SENTINEL2_L2A,
                time_interval=date_range,
                mosaicking_order='leastCC',
                maxcc=0.05  # Maximum cloud coverage of 5%
            )
        ],
        responses=[
            SentinelHubRequest.output_response('bands', MimeType.TIFF),
            SentinelHubRequest.output_response('scl', MimeType.TIFF)
        ],
        bbox=bbox,
        size=bbox_to_dimensions(bbox, resolution=10),
        config=config
    )
```

Test case: river Po delta

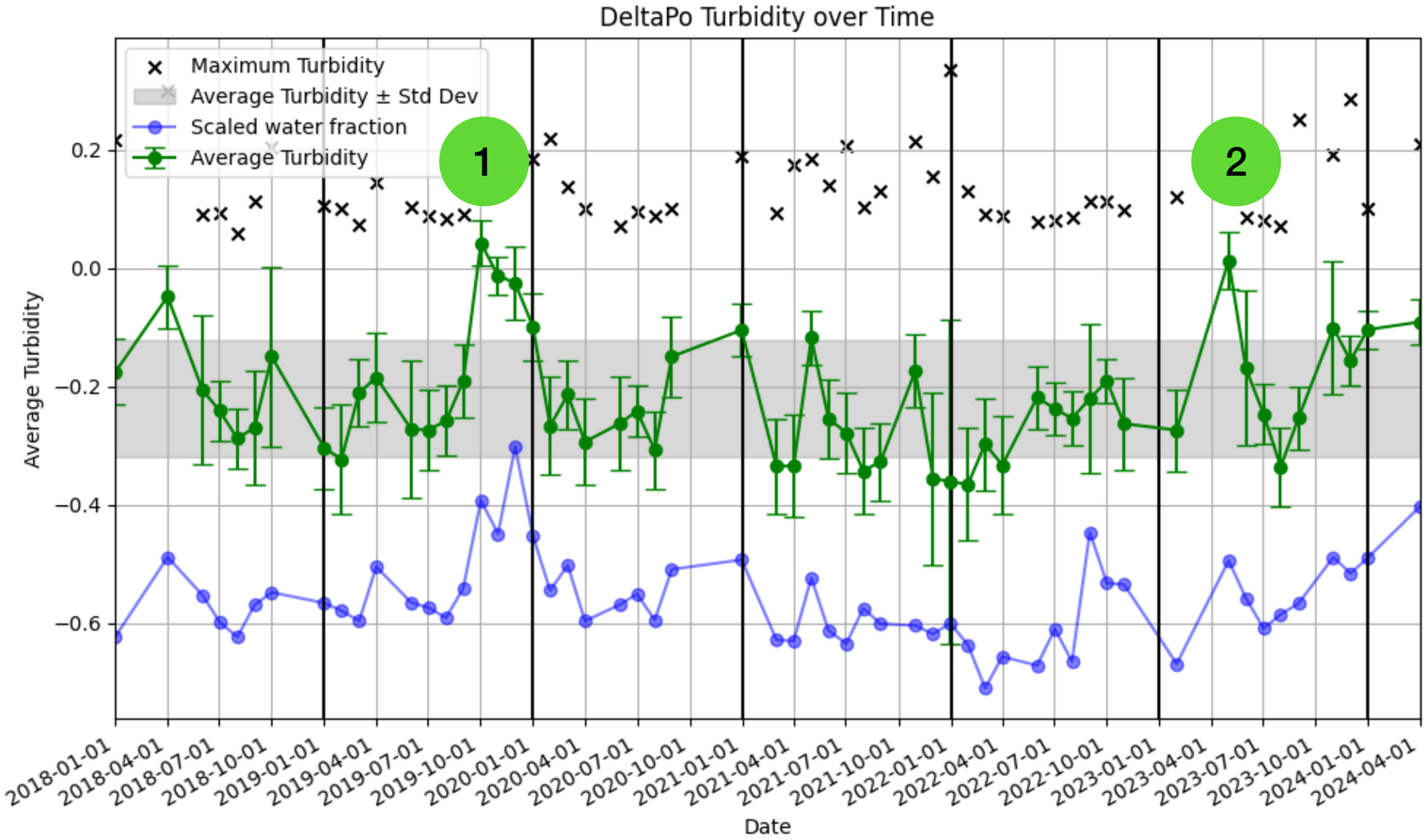
Cut-out images

- Top left: cropped map of the area in question, clipped along the geojson polygon; the scale indicates the **values of the water index.**
- Top right: **turbidity** index (geojson polygon in red).
- Bottom right: deviations of the turbidity index from the average (higher values in red). This helps identify where turbidity is higher than the average (for this specific image).



Delta del Po

- The green line indicates turbidity over time (missing data are skipped);
- The gray horizontal band represents the standard deviation – points above the band indicate exceptional turbidity.
- Thick vertical lines mark the beginning of each year; thin lines every 3 months.
- The blue line indicates the water fraction; values below the average may be due to residual cloud cover, while values above the average suggest the possibility of flooding.
- The crosses mark the maximum turbidity value – likely irrelevant.

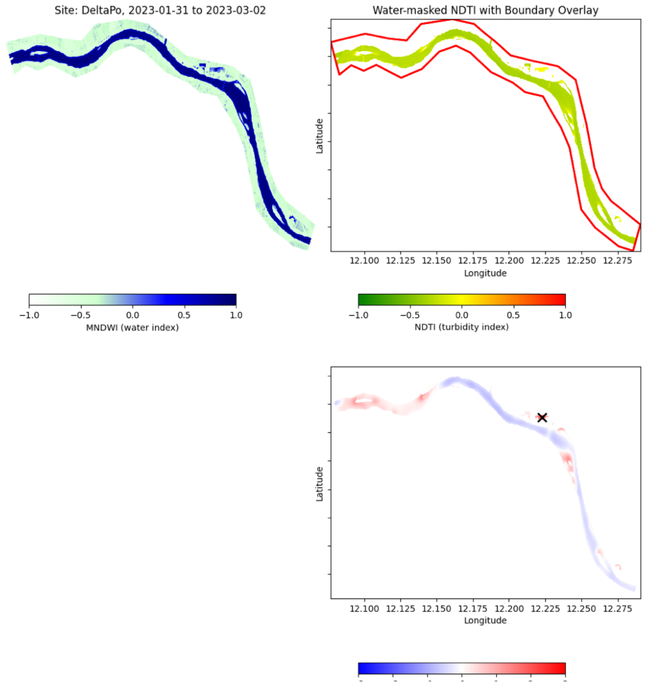


High turbidity correlates with flooding events

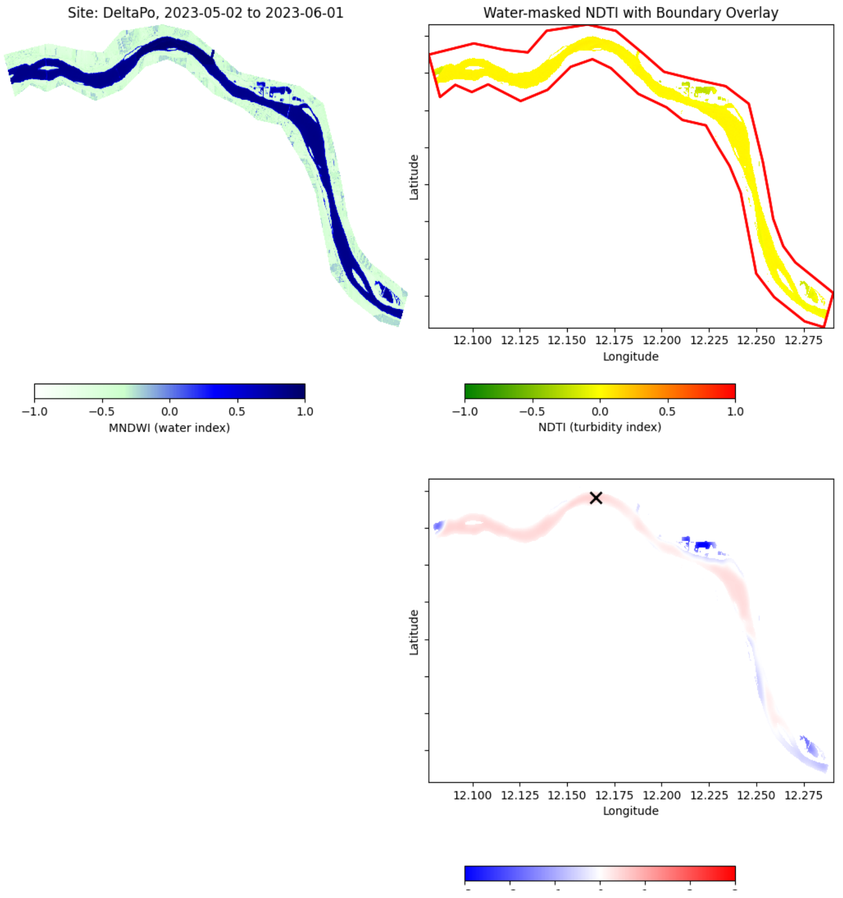
2

SISSA DATASCIENCE
Machine Learning for the Natural Sciences

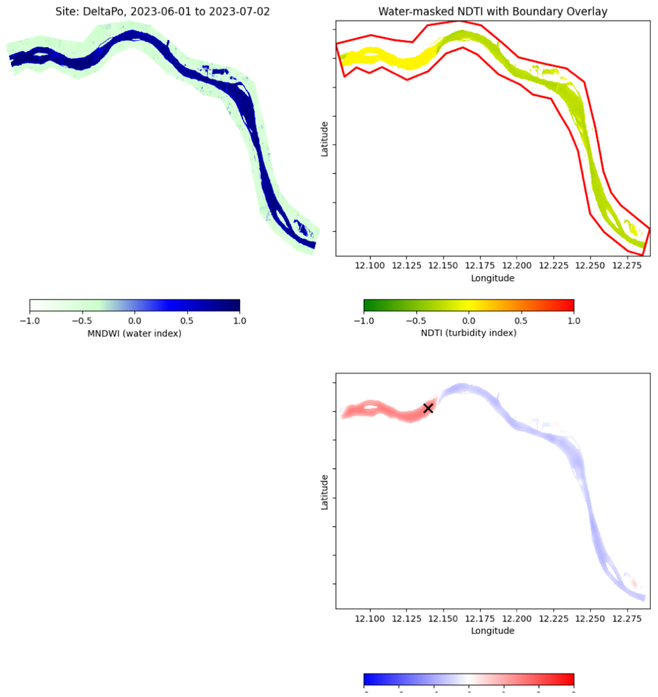
Febbraio 2023



Maggio 2023



Giugno 2023



Example: Cervia Saltworks

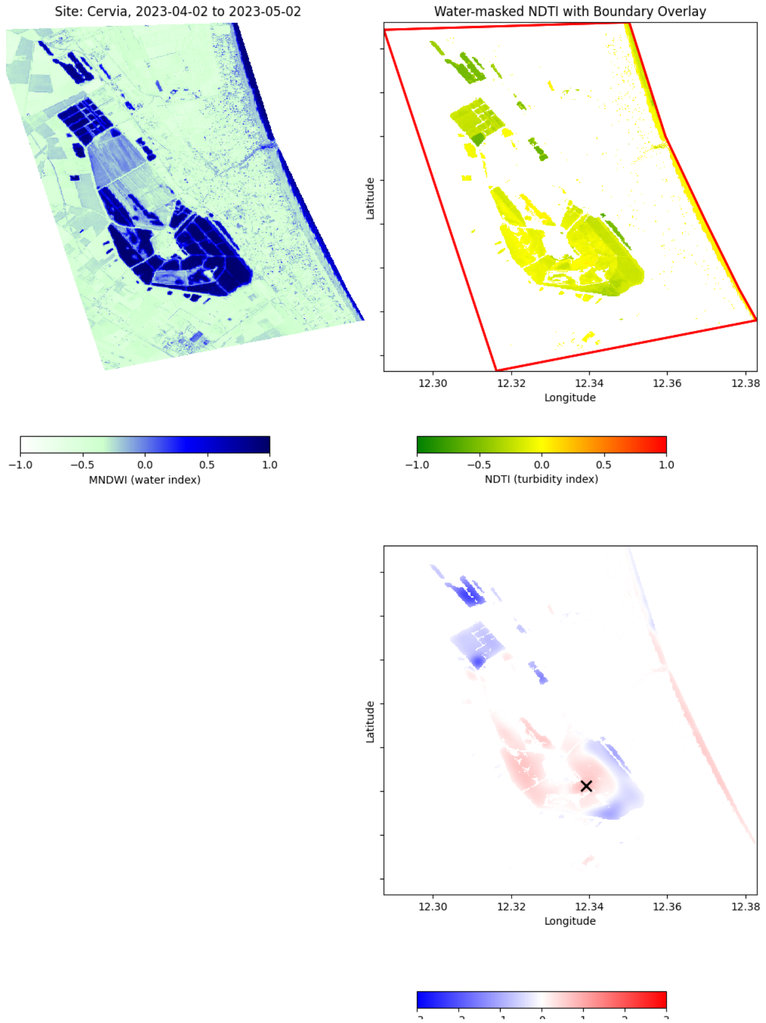
Flooding

A

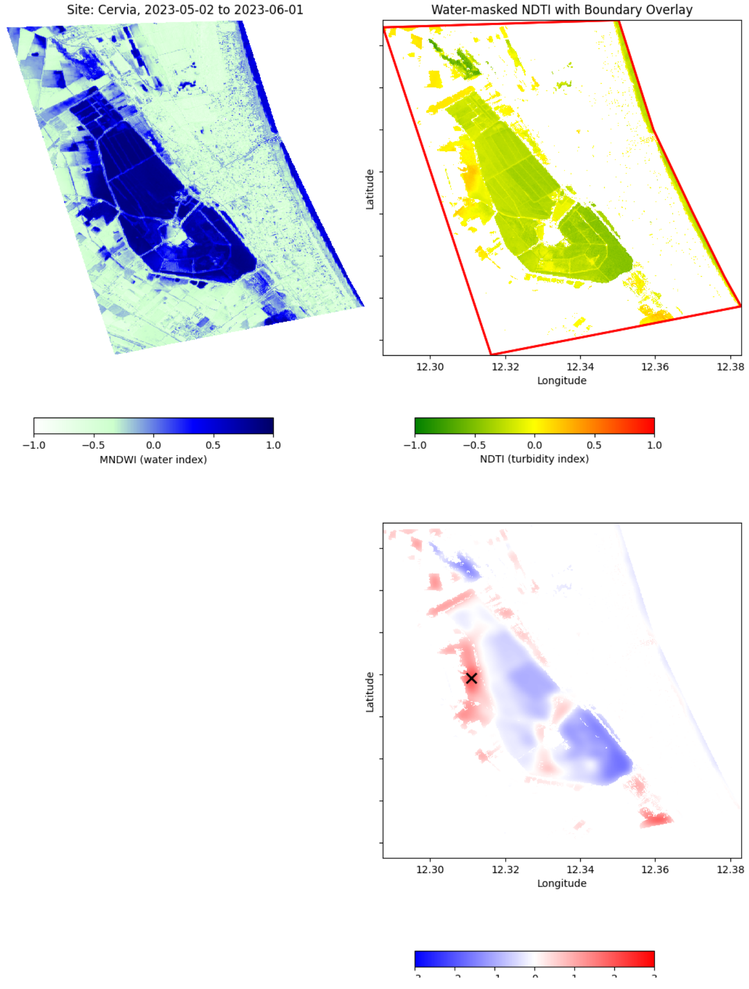
1

SISSA DATASCIENCE
Machine Learning for the Natural Sciences

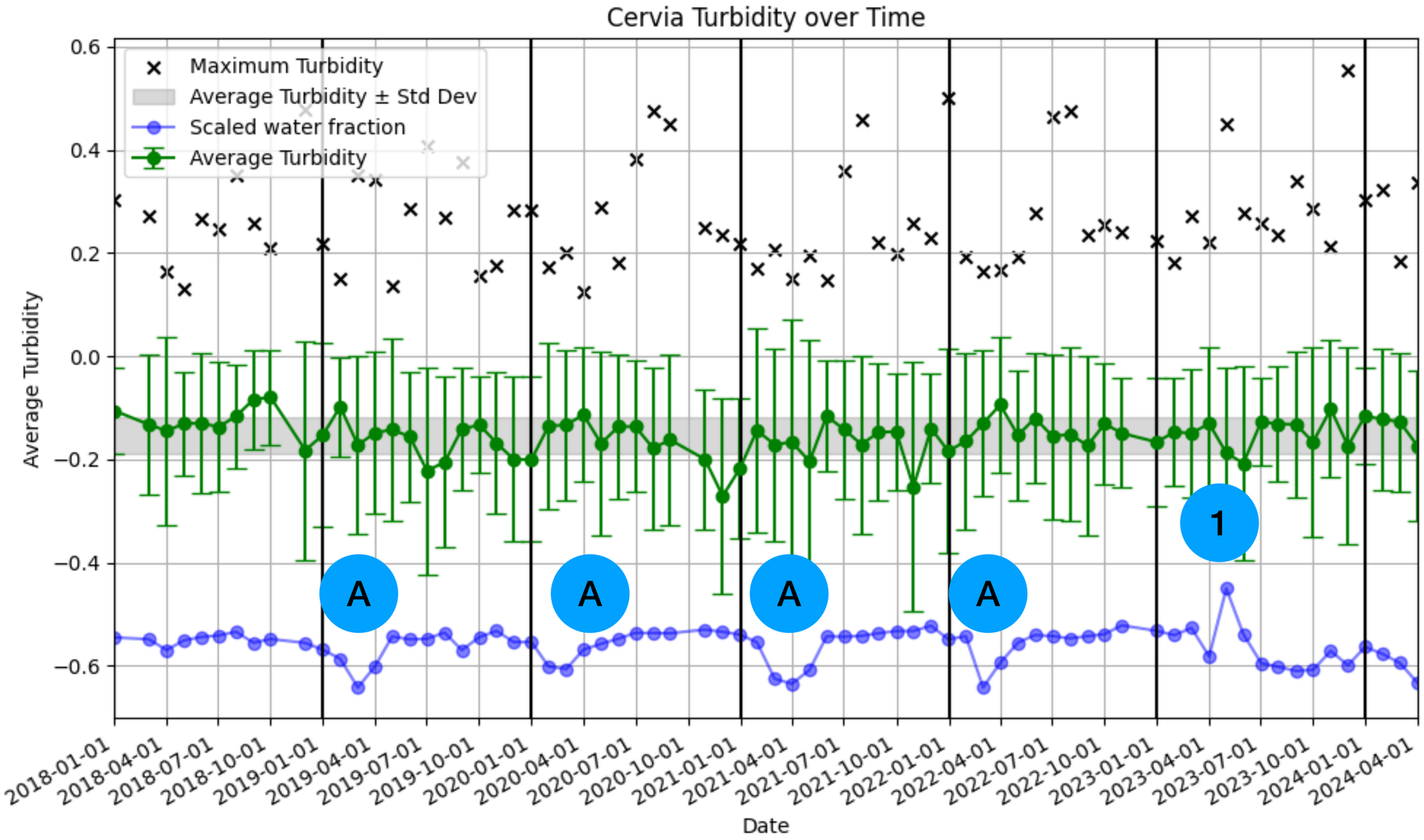
Aprile 2023



Maggio 2023



Cervia Saltworks



A Annual periodicity – reduced water levels in February/ March.
What does this correspond to?
The images seem to show very high salt levels.

1 Flooding, May 2023

The turbidity does not show any particular anomalies – the high variability is presumably due to the variable salt content from one tank to another.

A versatile, powerful tool

- This basic tool can now be fully automated and launched over hundreds or thousands of zones of interest all over Europe: **Fully automated story finder!**
- Some possible developments:
  - Is it possible to measure the variation in temperature/turbidity of rivers downstream from **nuclear power plants and/or data centers**?
  - Correlate with **shipping routes** to check ocean turbidity near busy shipping lanes; the same near **oil platforms**.
  - Investigate **illegal industrial/sewage discharges** near industrial centers or large cities.
  - Changes in river flows downstream from **large dams.**
  - Add your own!

https://datascience.sissa.it/

Contact:

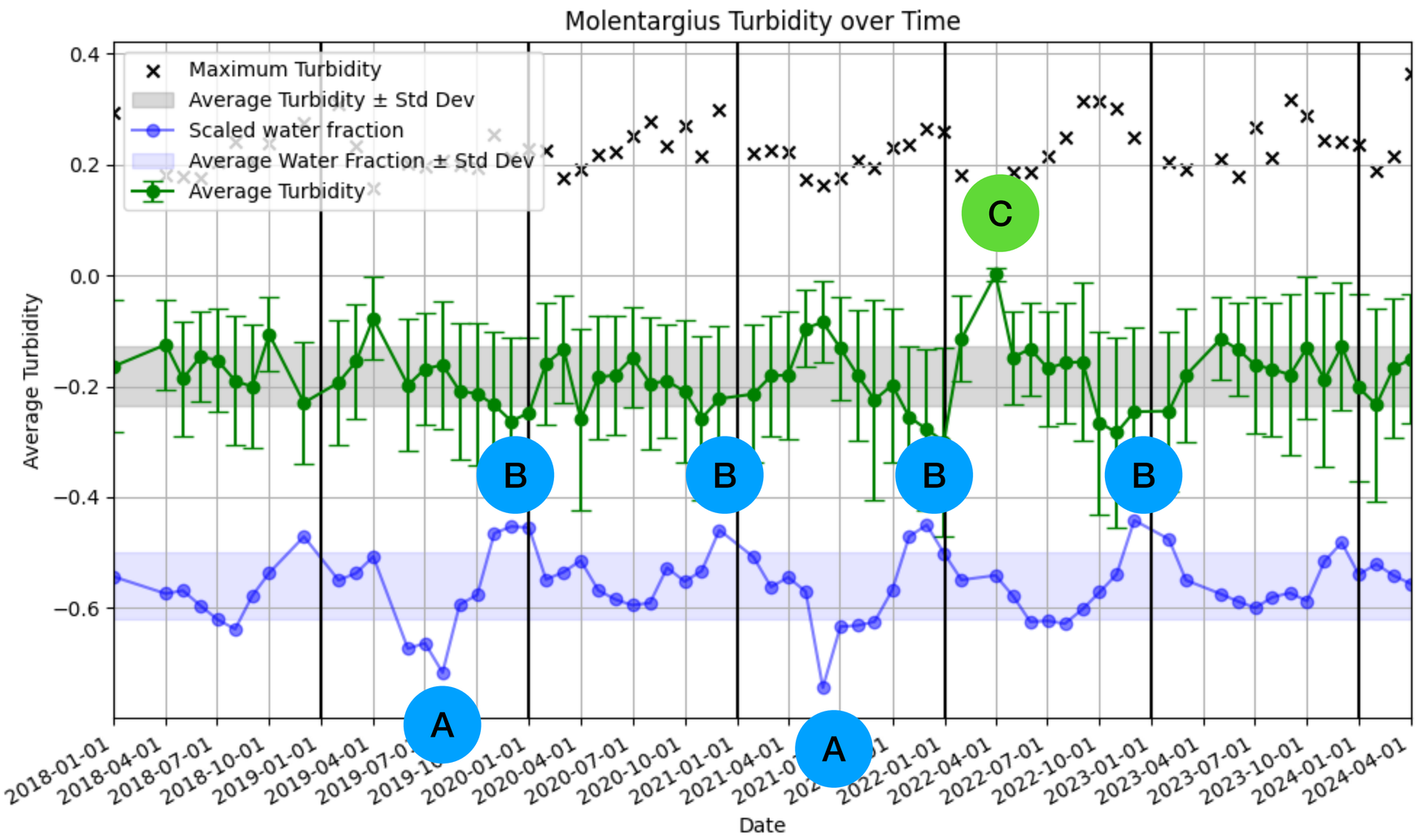
rtrotta@sissa.it

www.robertotrotta.com

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

Analisi - Molentargius

Molentargius



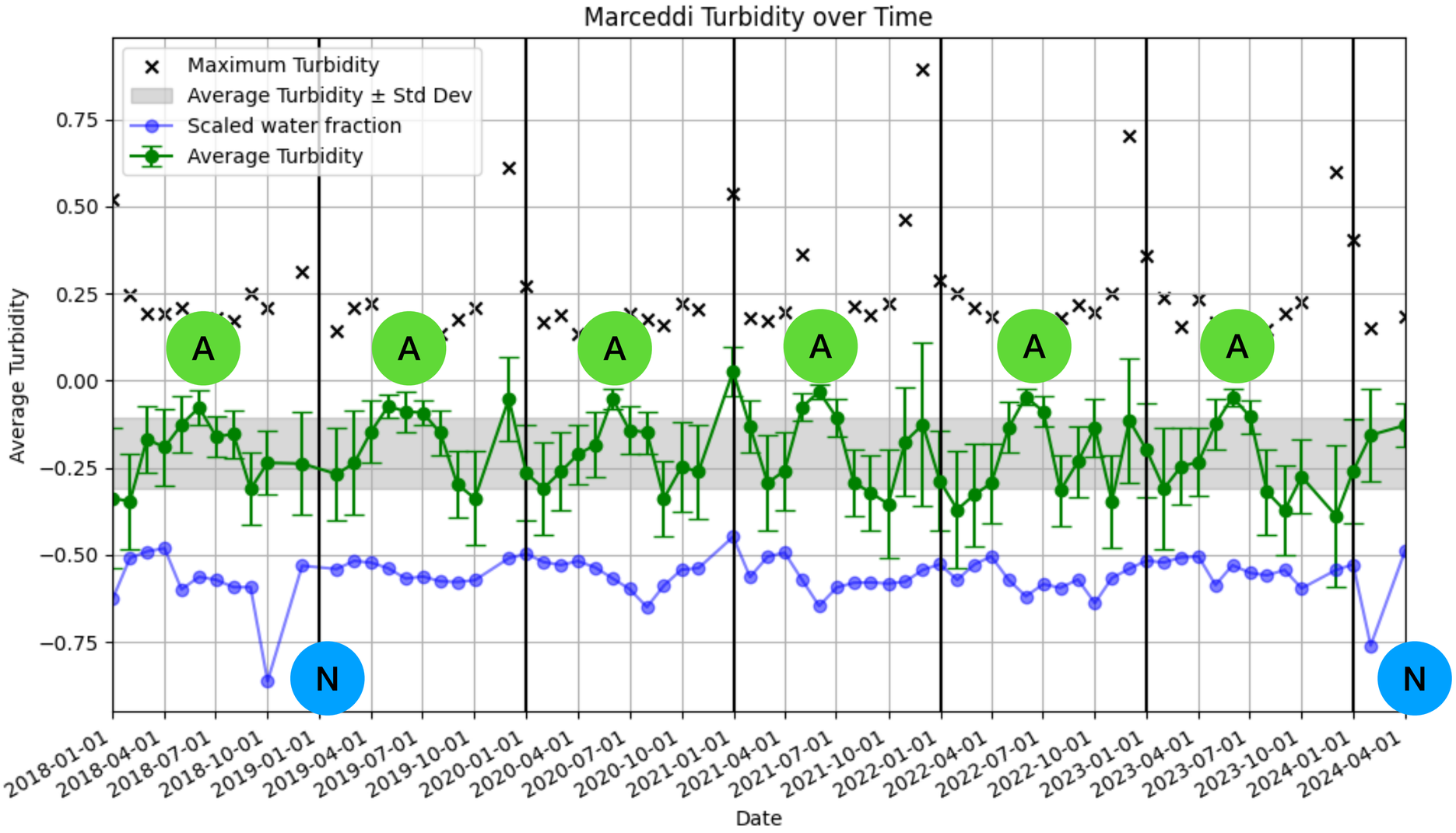
A Periodicità annuale - siccità estiva? O nuvole resiude?

B Periodicità annuale? Aumento acqua in inverno.

C Evento ad alta torbidità aprile 2022. Solo nuvole residue.

Analisi - Marceddi

Marceddi



N Solo nuvole residue

A Periodività estiva della torbidità. Dovuta a che cosa?