Analysis of chlorophyll-a and sediments in the water

Homework 3

Umberto Junior Mele

1388371

Earth Observation

INTRODUCTION

Chlorophyll is the green molecule in plant cells that carries out the bulk of energy fixation in the process of photosynthesis. Besides its importance in photosynthesis, chlorophyll is probably the most-often used estimator of algal biomass in lakes and streams. Its popularity results from several considerations:

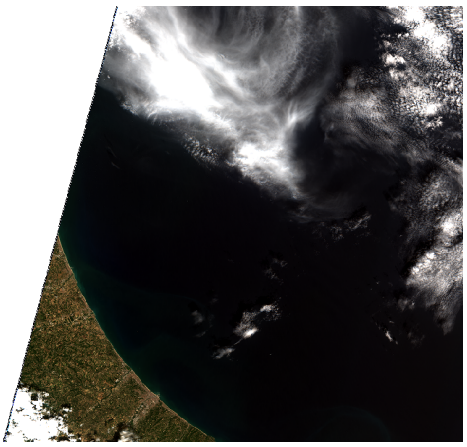
- it is a measure of algal biomass that is relatively unaffected by non-algal substances,
- it is a fairly accurate measure of algal weight and volume, and,
- it acts as an empirical link between nutrient concentration and a number of important biological phenomena in lakes and reservoirs.

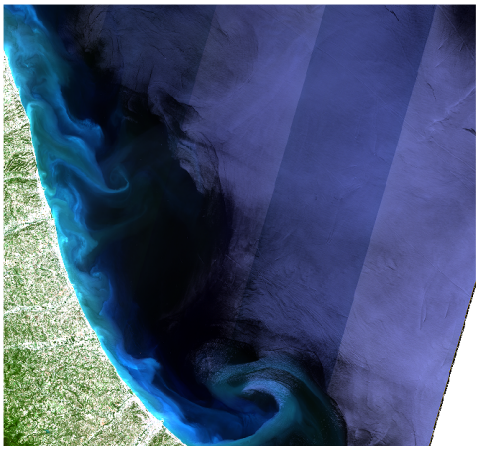
Chlorophyll is also relatively easy to measure. This facility of measurement contributes to its popularity, but the resulting values are far more ambiguous than most are willing to admit.

The aim of this homework is to analyze chlorophyll-a and sediments in the water near the Pescara river, using Sentinel-2 data available at the Copernicus website. The data is downloaded in level 1C format and between the period from 01-08-2016 to 24-08-2016.

Simple RGB of the data

Let's see what we have just downloaded:





These data are really raw, and we need to select a **region of interest (ROI)** and make it clear applying **atmospheric correction** to the image data.

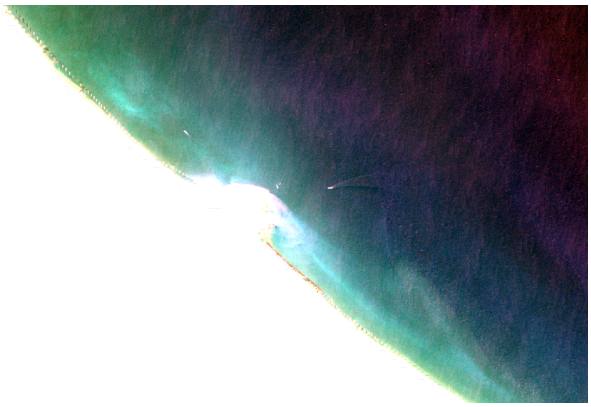
To do this on SNAP we need the external Sen2Cor processor (that is very slow to process).

This are the level 2A image output of Sen2Cor processor and after that has been applied spatial subset selector on the area of interest, close to the river of Pescara. Sentinel-2 data are very good to search chlorophyll-a and sediments on the sea, thanks to a very good resolution up to 10 meter due to the low orbit trajectory, thanks to the 13 bands on VIR

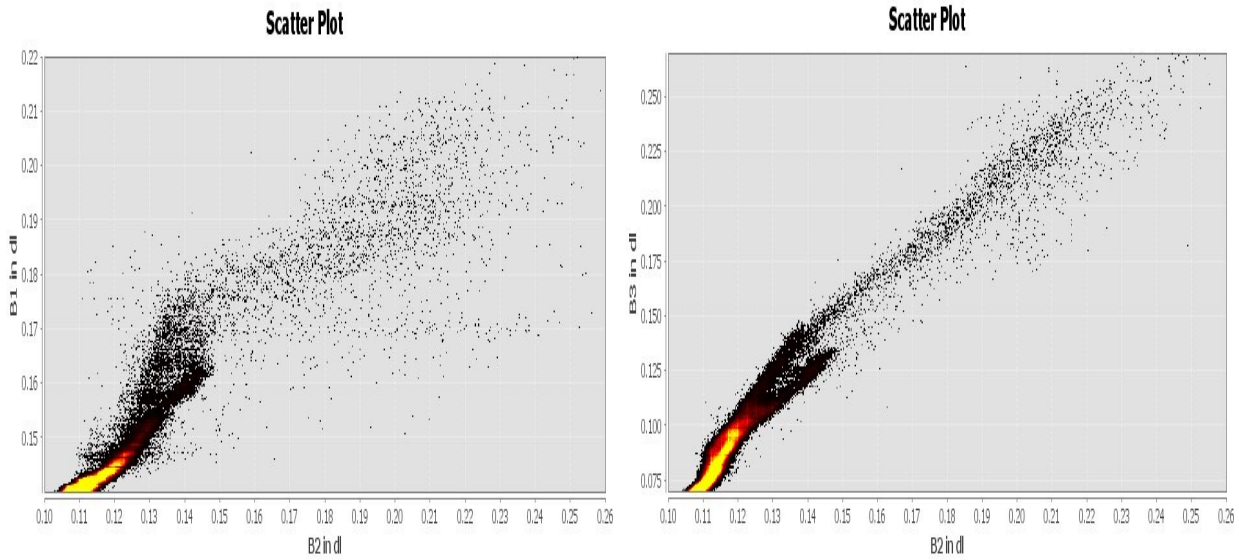
and on IR , and least but not last thanks to free data accessibility every 5 days on the same area.

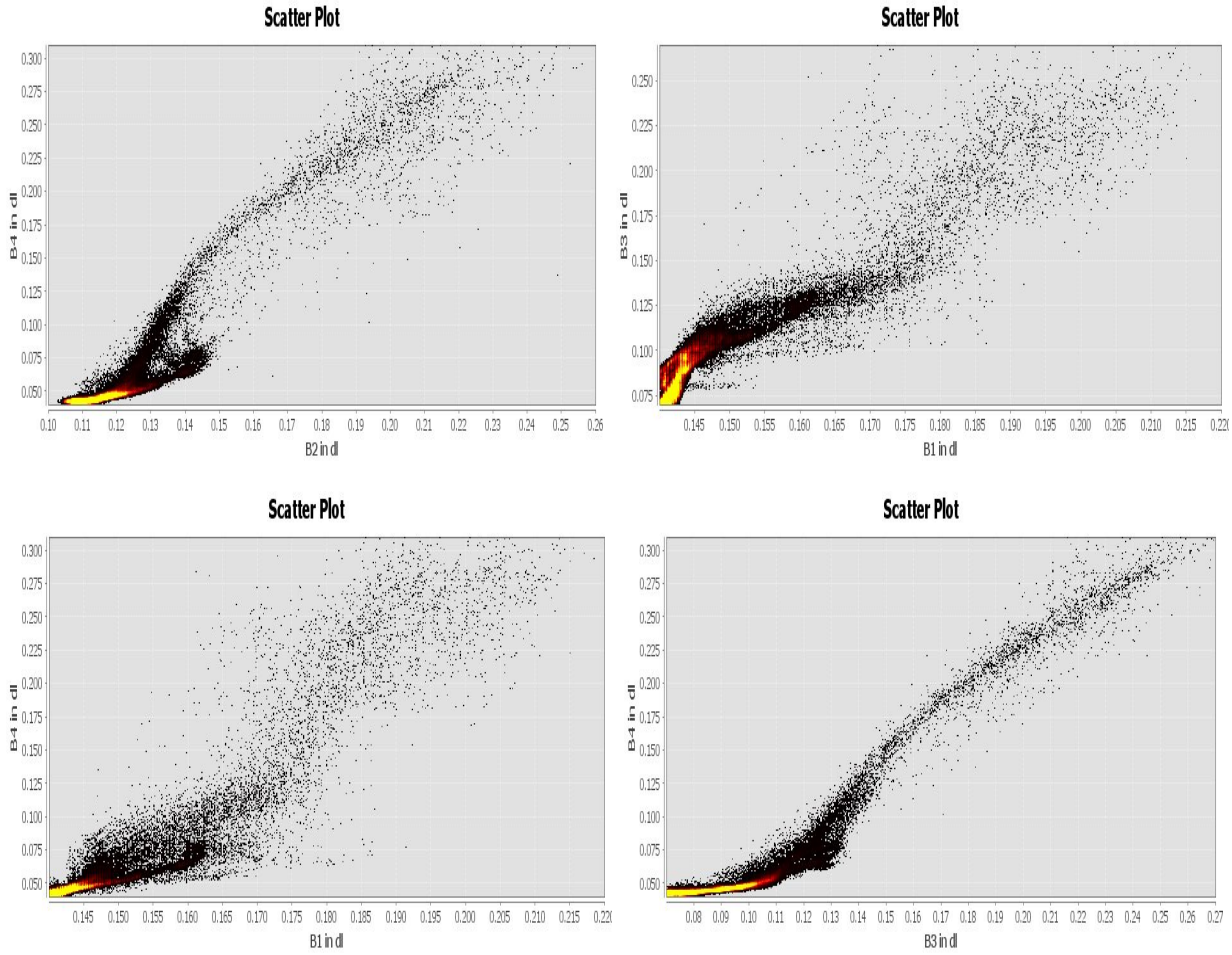
Analysis of selected ROI

Fundamental for a regression model is to check the multi-collinearity of the available data, with this aim we now are going to investigate correlation. So, after that is made a mask on sea:



I plotted the correlation of the first four bands to visualize and better understand a regression model on it.





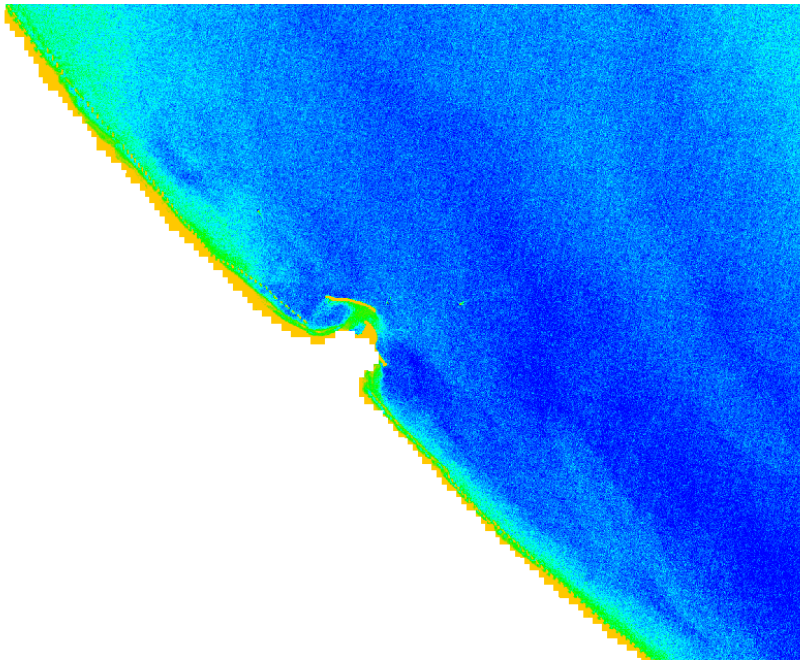
It’s clear that this channels have high correlation since the wavelengths are close each other, but still is possible to reach some good information on it since these are the four bands that can be used for RGB plot precisely B4 for red, B2 for green and B3 for blue.

Implement 4 regressive algorithms to estimate chlorophyll-a (Chl-a)

These algorithms returns the near-surface concentration of chlorophyll-a (chlor_a) in mg/m^3, calculated using an empirical relationship derived from in situ measurements of chlor-a and remote sensing reflectances in the blue-to-green region of the visible spectrum. The implementation is contingent on the availability three or more sensor bands spanning the 440 - 670 nm spectral regime. The algorithms is applicable to all current ocean color sensors.

First algorithm:

$$CI = R(\lambda_{green}) - [R(\lambda_{blue}) + (\lambda_{green} - \lambda_{blue})/(\lambda_{red} - \lambda_{blue}) * (R(\lambda_{red}) - R(\lambda_{blue}))]$$

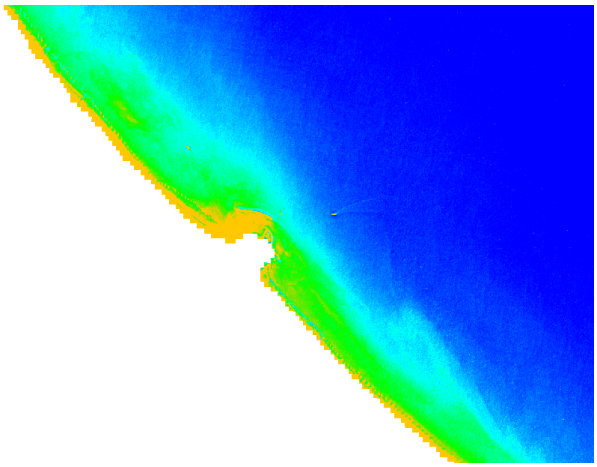


This algorithm doesn't underline all the shades and is quite rude on changing values…

Second algorithm, called OC4:

$$log_{10}(chlor-a) = a_0 + \sum_{i=1}^{4} a_i \, (log_{10}(R(\lambda_{blue}) / R(\lambda_{green})))^i$$

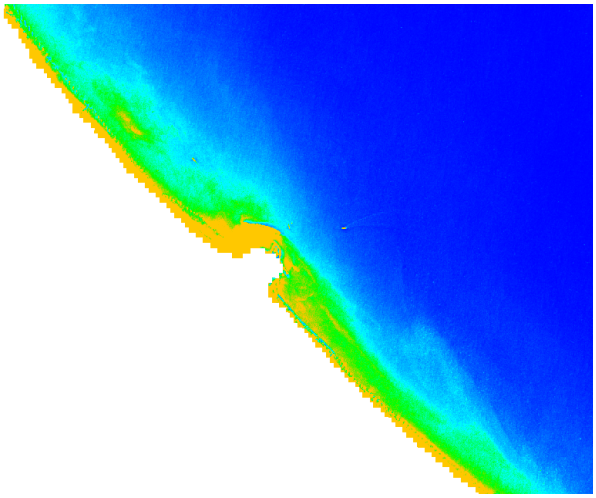
$$a_0 = 0.3272\,,\ a_1 = -2.9940,\ a_2 = 2.7218,\ a_3 = -1.2259,\ a_4 = -0.5683$$



Instead this other ones do a good job

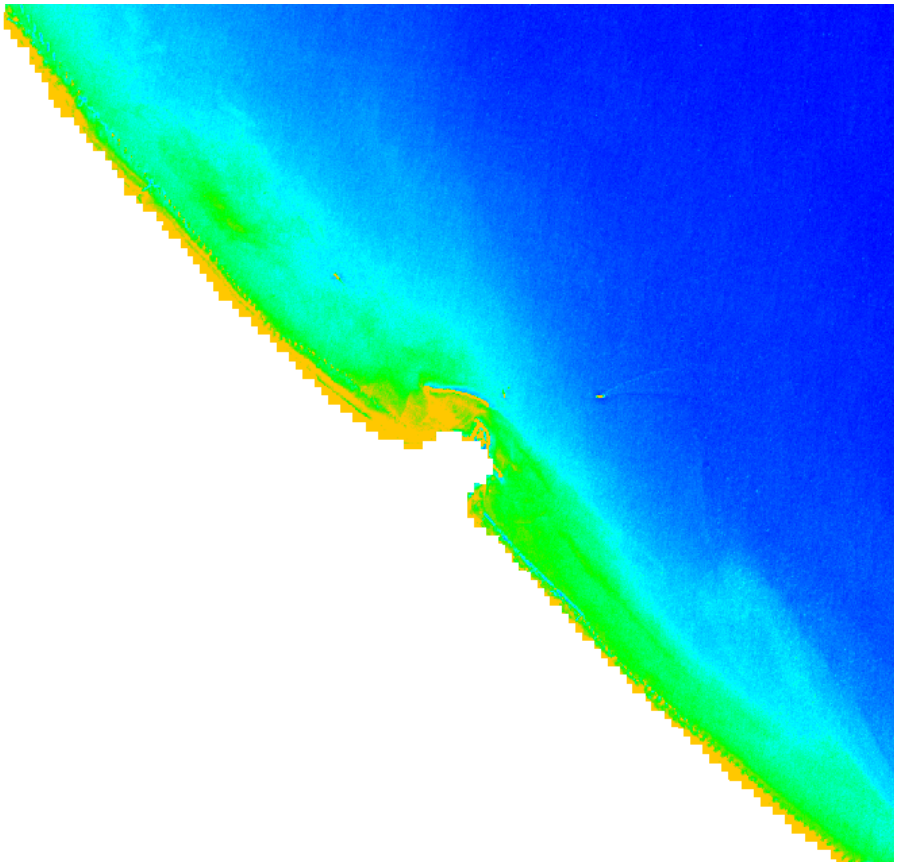
Second algorithm, called OC3M (same structure):

$$a_0 = 0.2424\,,\ a_1 = -2.7423,\ a_2 = 1.8017,\ a_3 = -0.0015,\ a_4 = -1.2280$$



Fourth algorithm is the SICP (Sentinel2 Italian Coastal Water Prediction), a regression model obtained by in situ data of Chl-a over the Italian coasts, which returns the near-surface concentration of Chl-a in μg/l.

$$SICP_{chla} = 1.7541 * e^{-1.547 * MBR} \qquad MBR = max(R(\lambda_{green}))/ R(\lambda_{blue})$$

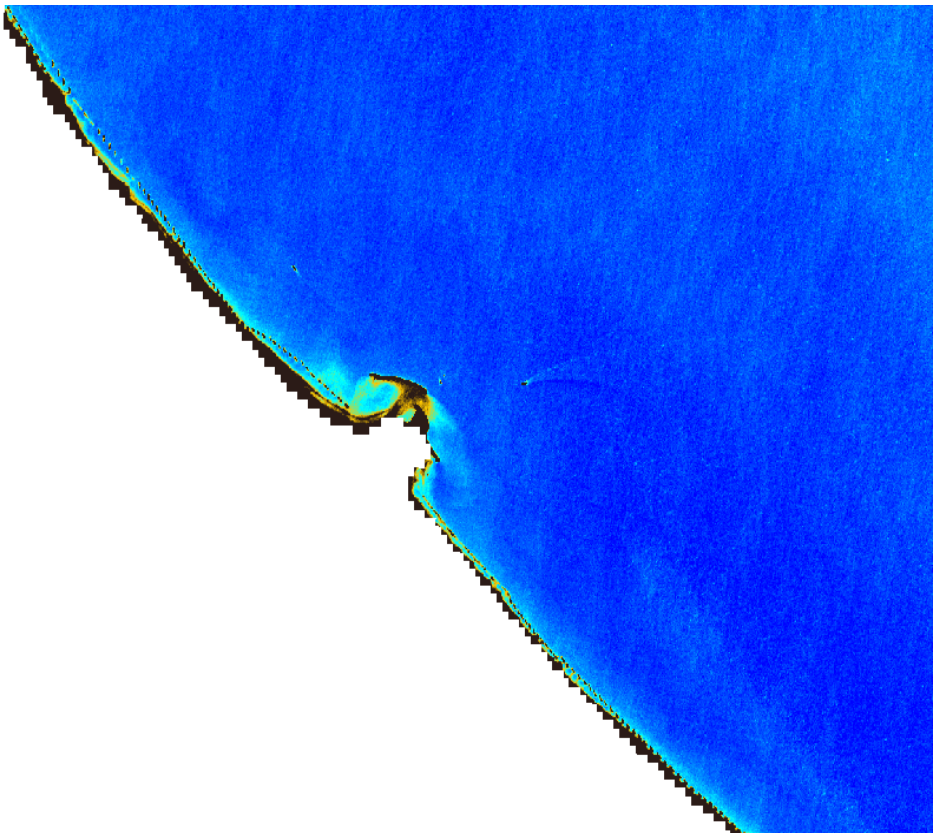


Also this last one very simple model do a good job, and since low complexity is better I think that this last one could be the best choice of them.

Regression model to estimate total suspended sediments (TSS)

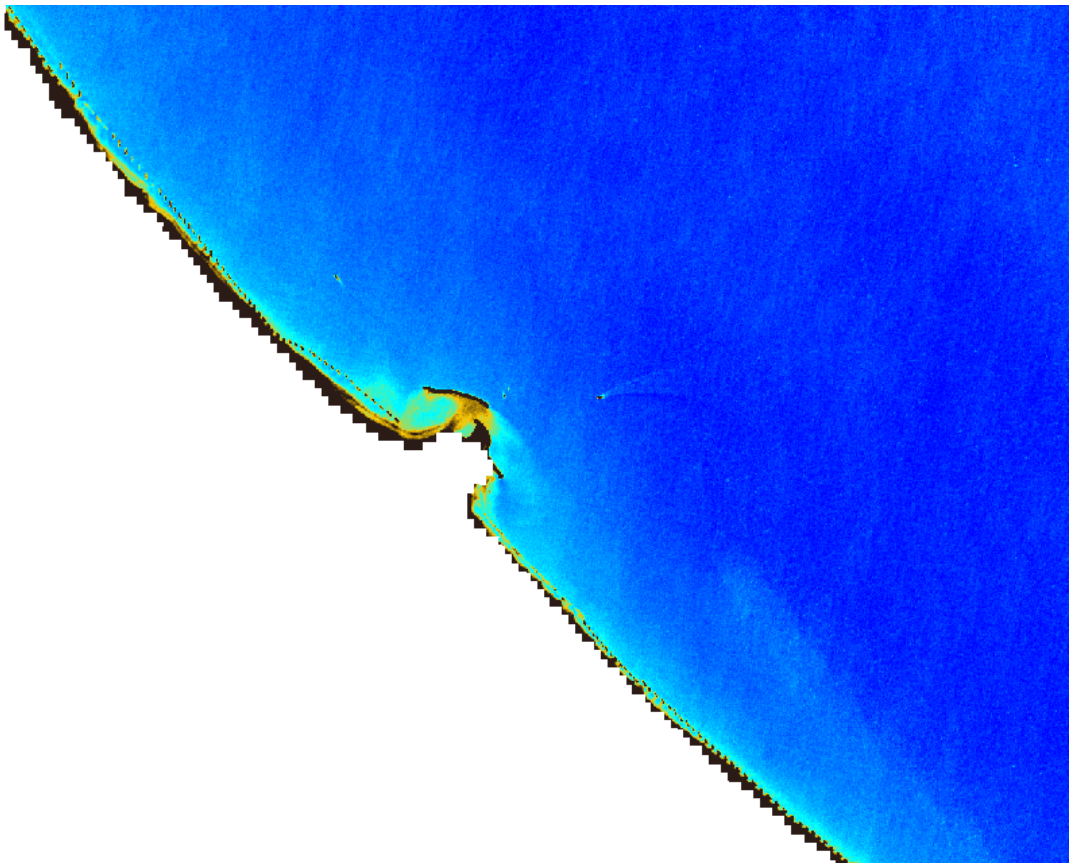
The first model used to estimate TSS is retrieved empirically using a linear regression model, with a single feature represented by the ratio of red reflectance over the green one.

$$TSS_1 = 23.26 * (B4 / B2) + 0.61$$



Second model is still a regression but this time the feature is a smoothed ratio:

$$TSS_2 = 48.35 * ((B3 + B4)/(B3 + B2) - 25.04$$

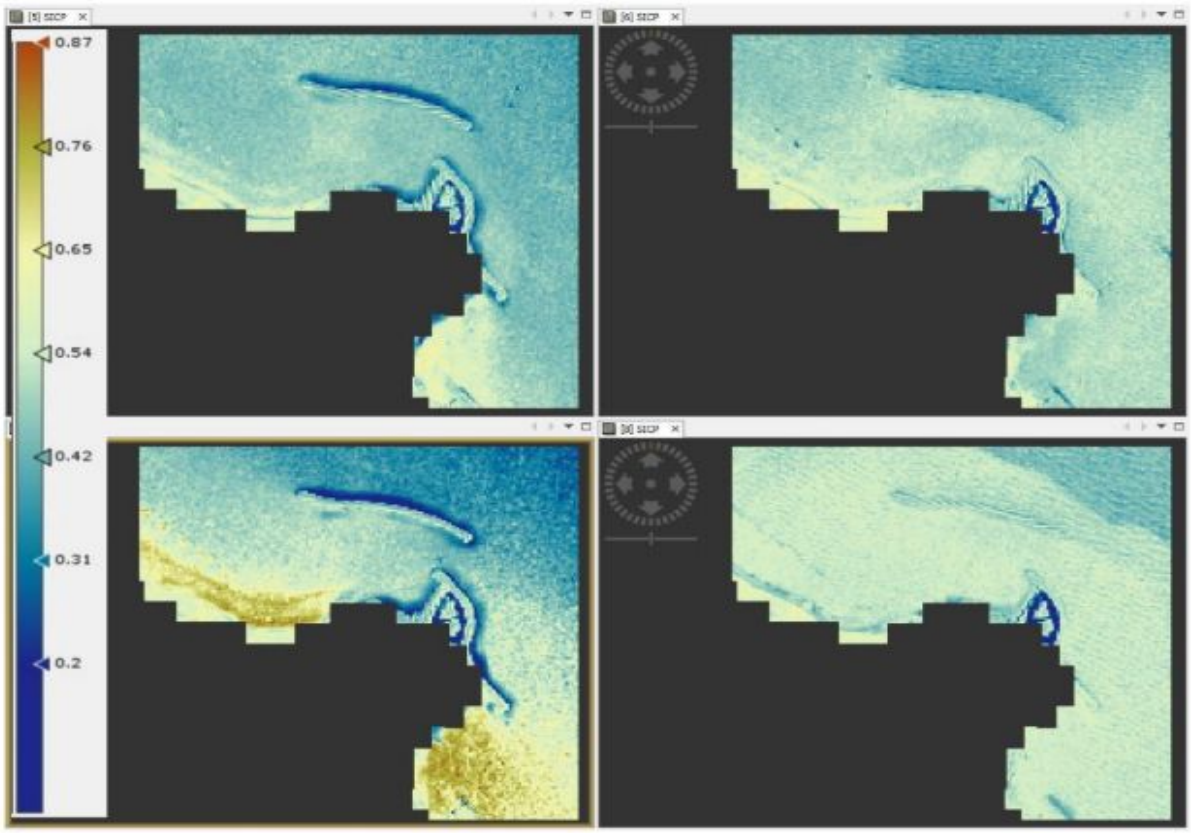


second model has better result in my opinion, since there is the blue reflectance that smooth the ratio.

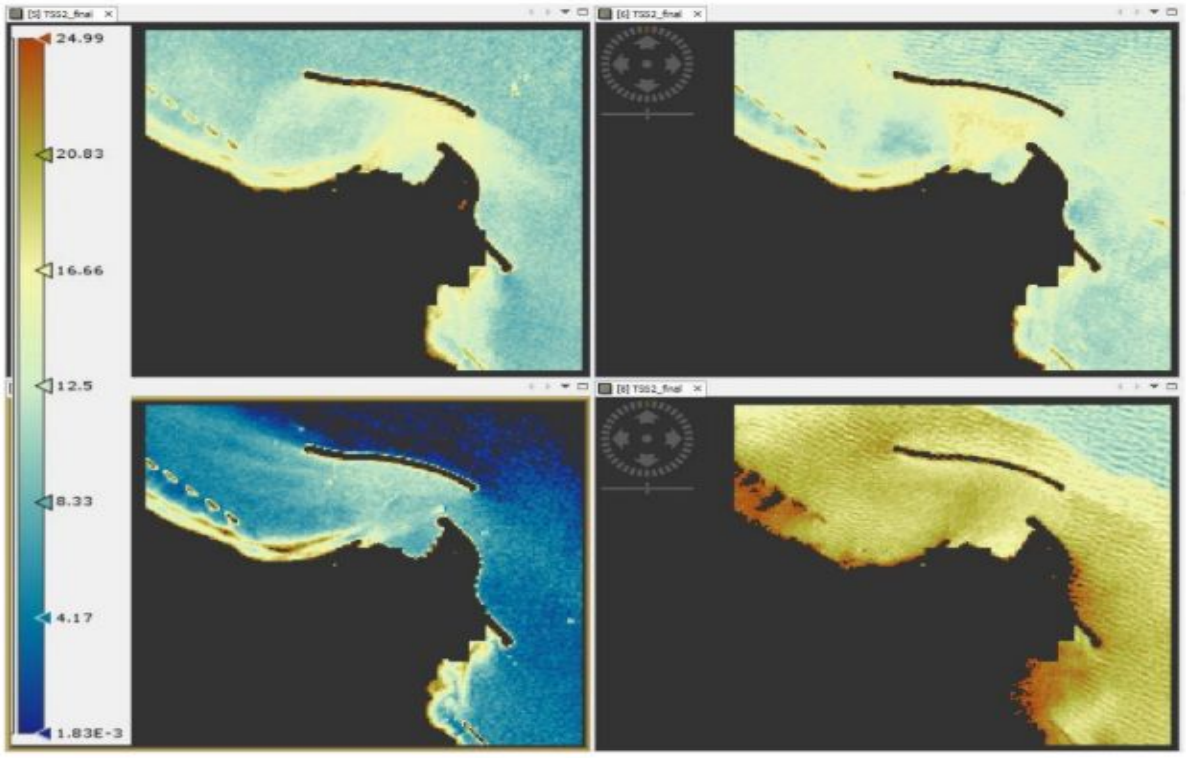
Time series analysis

To make the time series Analysis I downloaded all data available on Copernicus from Sentinel-2 by MSI at level 1c, between the data 01-08-2016 and 24-08-2016 and with cloud coverage small.

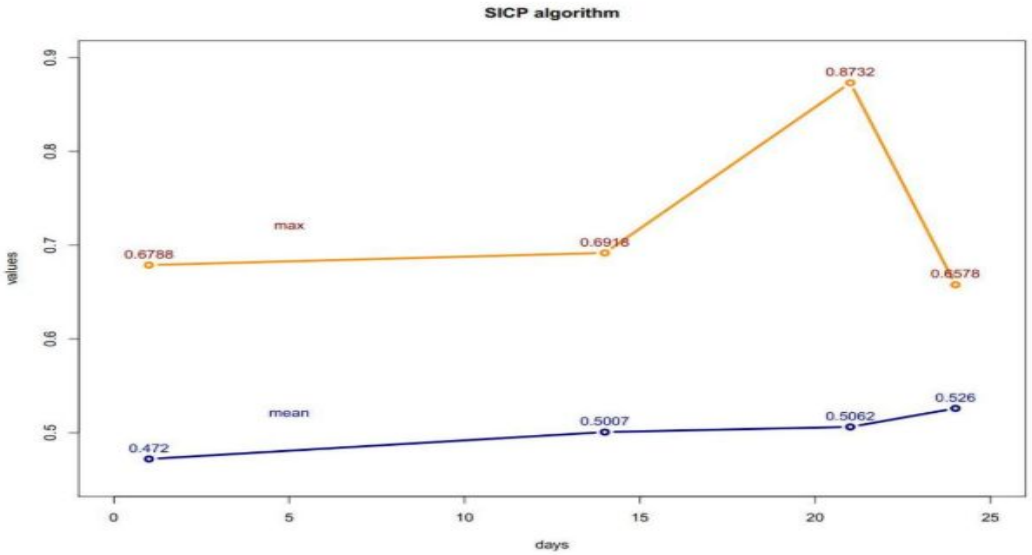
Using SICP regressor to retrieve the estimate of Chl-a and TSS2 model to retrieve the total suspended sediment, we want to estimate this indices during time. We can use the graph builder tool of SNAP to make thighs easier and to be sure that we are doing the same operation on all data. Firstly we need to select the subset and create the sea-mask. After that we implement both regressor mask and the final output is the following for the SICP index:



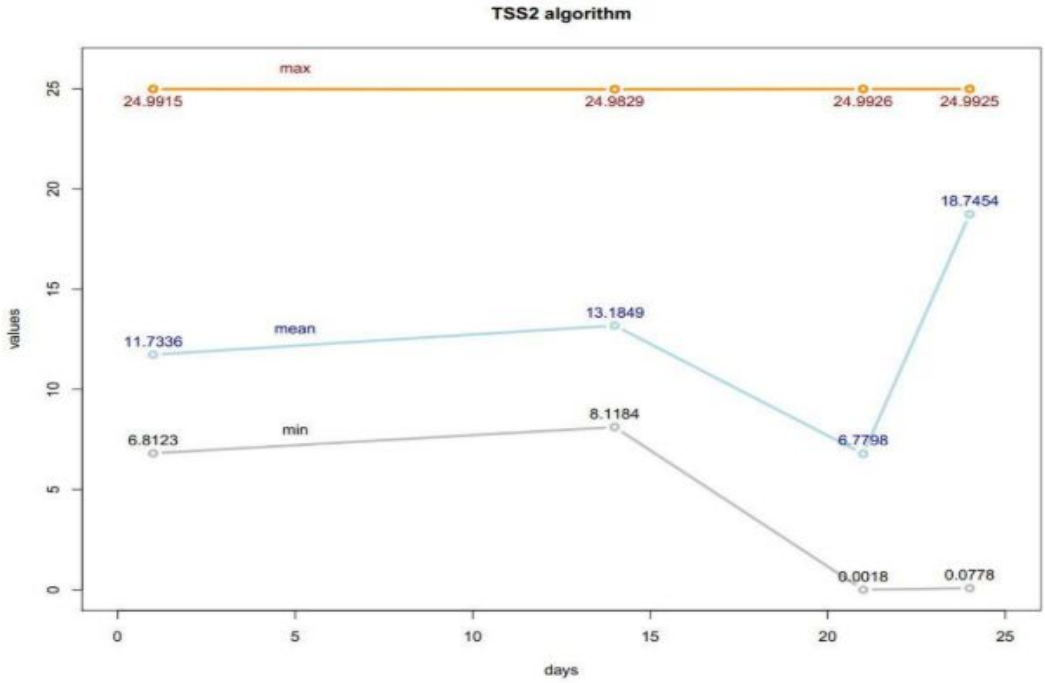
and this one for TSS2:



to visualize it better I made also a plot of the mean and max values for this images, SICP:



and for TSS2, we plot mean, max and min:



RIFERIMENTI

1. Copernicus
2. Sentinel-2 sensors
3. Sen2Cor
4. NASA Ocean Color