Earth Observation Data Analysis - Homework 03 - 2018

Andrea Pisani

1. **Download and install SNAP (if needed, you can use a Virtual Machine to exploit the resources available within ESA Cloud ToolBox facility at http://eogrid.esrin.esa.int/cloudtoolbox after registering)**

To complete all the tasks of the homework I downloaded and installed SNAP on my PC.

2. **Download a MSI/S2 image over Central Italy (Pescara river estuary, 42.46° N, 14.21° E) within the period 5-12 August 2016 (ndr: period was extended in class) from https://scihub.copernicus.eu or download MSI/S2 imagery provided by the professor. You can also choose a different target area (e.g., Tiber river estuary in Italy, Po river delta in Italy, …).**

The **level-1C** products used for the homework are acquired between the **1st** and the **24th of August**.

The first product (2016/08/01) is used to introduce and perform all the algorithms requested in the tasks, while the other products are considered in the last part (time series analysis). For the atmospheric correction I will show the image with the most evident improvement from the level-1C product and the level-2A product.

In the chosen period the **Sentinel-2A** satellite only was in low earth orbit, with no possibility of choosing a certain tile during the download phase in the Copernicus portal.

However, the full images (with more tiles) have been download successfully and the tile **T33TVH** containing the Pescara river has been exploited to compute the algorithms.

For each product the full name assigned in Copernicus is listed with the sensing date.

- S2A_OPER_PRD_MSIL1C_PDMC_20160801T171032_R079_V20160801T095532_20160801T095532
  Sensing date: 2016/08/01 (Figure 1)

Figure 1 – L1C RGB image 2016/08/01

- S2A_OPER_PRD_MSIL1C_PDMC_20160814T232331_R122_V20160814T100032_20160814T100604
Sensing date: 2016/08/14 (Figure 2)

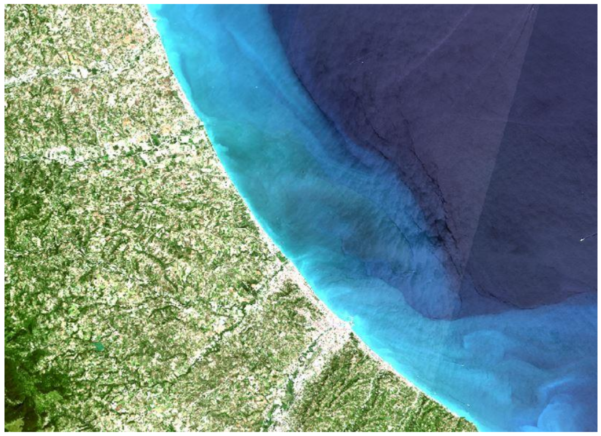


Figure 2 - L1C RGB image 2016/08/14

- S2A_OPER_PRD_MSIL1C_PDMC_20160825T031525_R079_V20160821T095032_20160821T095527
Sensing date: 2016/08/21 (Figure 3)

Figure 3 - L1C RGB image 2016/08/21

- S2A_OPER_PRD_MSIL1C_PDMC_20160825T191127_R122_V20160824T100032_20160824T100607
Sensing date: 2016/08/24 (Figure 4)

Figure 4 - L1C RGB image 2016/08/24

3. Perform data quality check

All the products listed in point 2 have been inspected and all the 12 bands in each of them are usable to perform the tasks of the homework.

The RGB images of the 2016/08/14 and the 2016/08/24 have diagonal periodical stripes (one stripe can be seen in the right part of Figure 2 and Figure 4) but not within the area of the Pescara river. However, possible effects in the final results will be considered.

4. Apply atmospheric correction to image data if needed and compare the results with and without this correction

To apply the atmospheric correction to the products I used the Sen2Cor tool developed by ESA for SNAP.

Sen2Cor is a processor based on Python for Sentinel-2 level 2A product generation and formatting. It performs the atmospheric, terrain and cirrus correction of TOA level 1C input data to generate **level 2A data**.

The processor creates BOA, optionally terrain and cirrus corrected reflectance images. Its output product format is equivalent to the Level 1C User with three different resolutions [1]:

- 60 m **-** Bands 1, 9 and 10
- 20 m - Bands 5, 6, 7, 8A, 11, 12
- 10 m - Bands 2, 3, 4, 8

I used the version 2.5.5 of the Sen2Cor processor on each tile T33TVH of each product downloaded from Copernicus: the algorithm has been applied **3 times** for each resolution (60 m, 20 m and 10 m) to obtain the L2A products.

The installation was performed following the ESA instructions [2]. Instead of using the standalone installer, I applied the atmospheric correction on each tile using the **Sen2Cor plugin** available in the SNAP environment. In Figure 5 there's a screenshot showing the Sen2Cor computation.

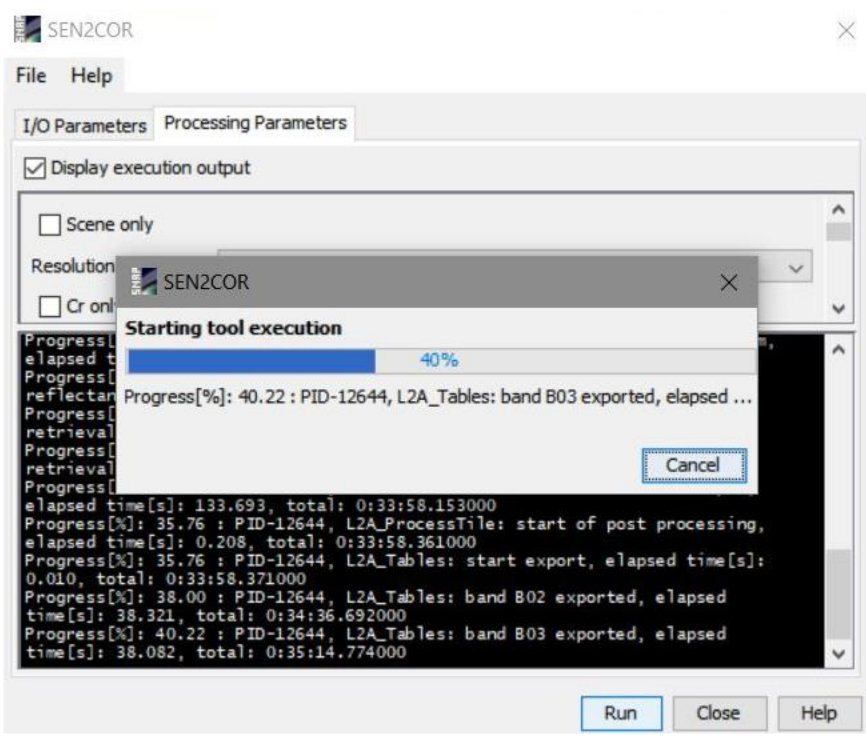


Figure 5 - Sen2Cor plugin in the SNAP environment

The results of the Sen2Cor correction are not always so evident in every image. In Figure 6 the **before/after** atmospheric correction of the 08/21 product is shown, in which the removal of the water vapor is more evident.

As it can be easily seen the L1C product on the left is brighter than the L2A product on the right.

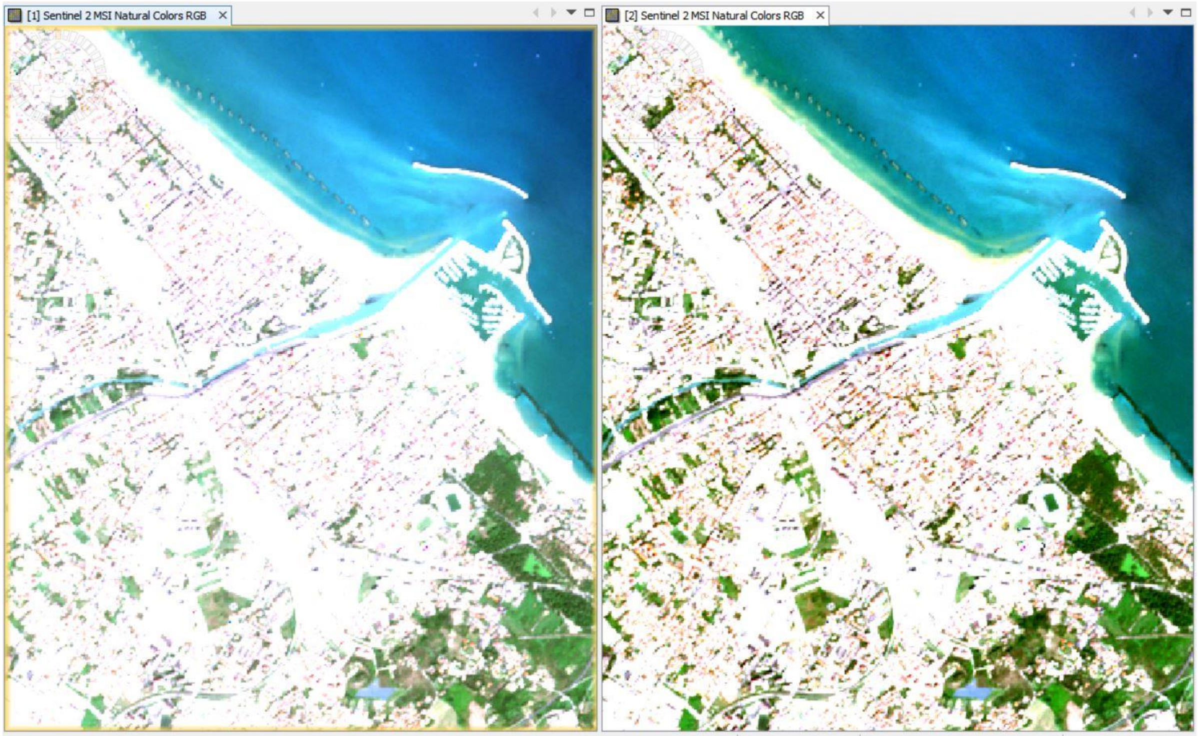


Figure 6 - Before (L1C on the left) and after (L2A on the right) the Sen2Cor computation. Product acquired the 2016/08/21

5. Select a ROI (Region of Interest) around the Pescara river estuary

From now on, I will work on the product acquired the **2016/08/01**. To select a specific region, it's mandatory to perform a resample of the picture. The **S2 resampling tool** has been used to obtain 10 m of resolutions in each of the 13 channels. The Figure 7 shows the resampling tool and the ROI chosen to perform the tasks of the homework. From now on, I'll use the 2016/08/01 L2A product for each task.

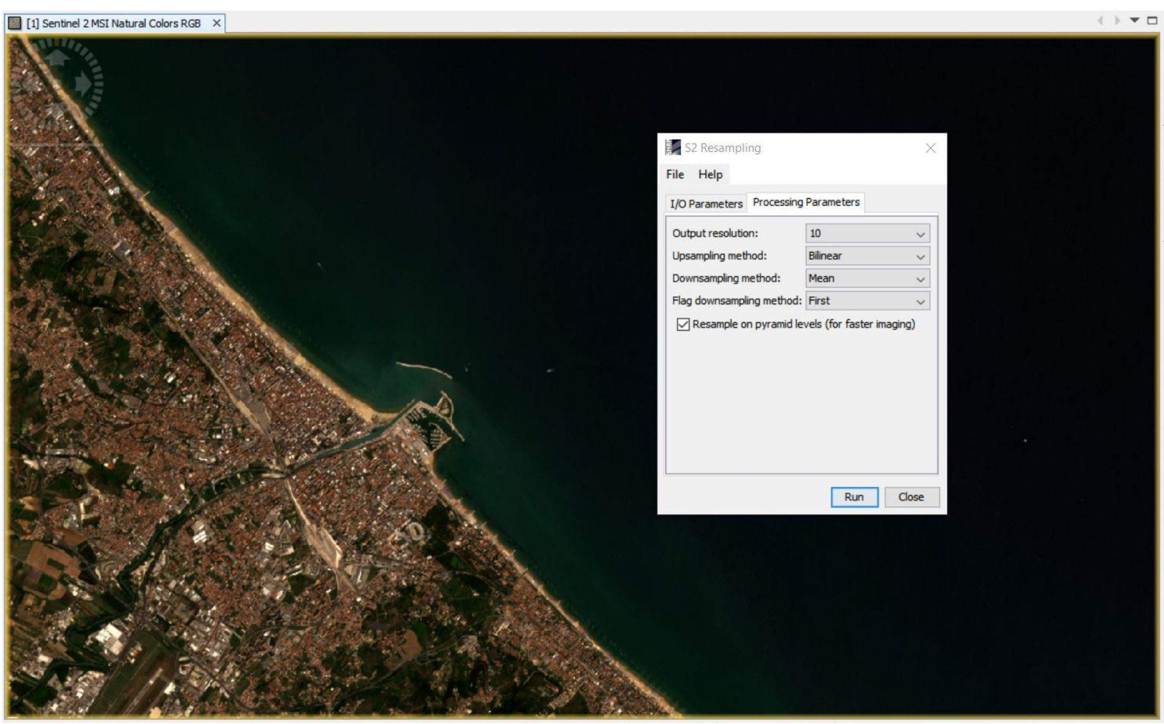


Figure 7 - Resampling of the channels to obtain 10 m of resolution

In Figure 8 the specific area of interest with respect to the entire tile is shown.

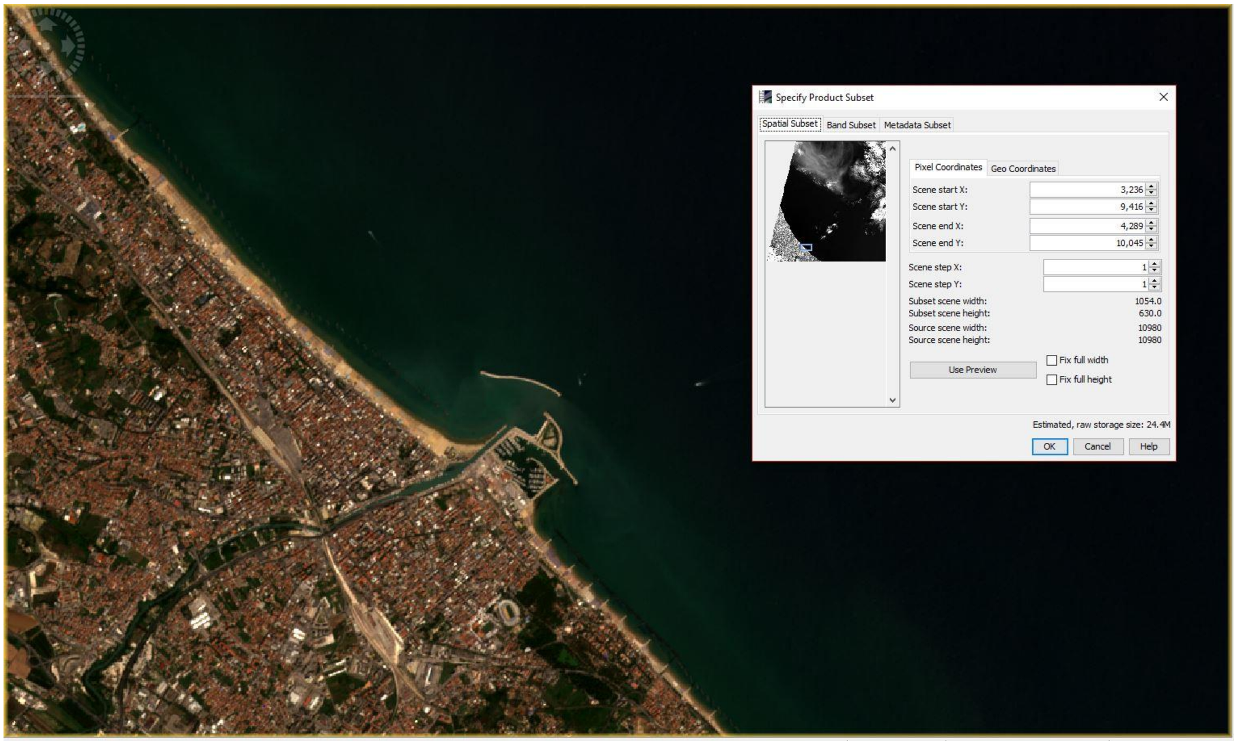


Figure 8 - subset near the Pescara river

The aim of the homework is the analysis of chlorophyll-a and sediments in the water near the Pescara river. First, I created a mask to remove the terrain. The final mask that I will use to apply all the algorithm is the one shown in Figure 9. The sea seems brighter with respect to the original image only because it's darker than the terrain: SNAP automatically changes the **color scale** according to the mean value of the brightness of all the pixels.

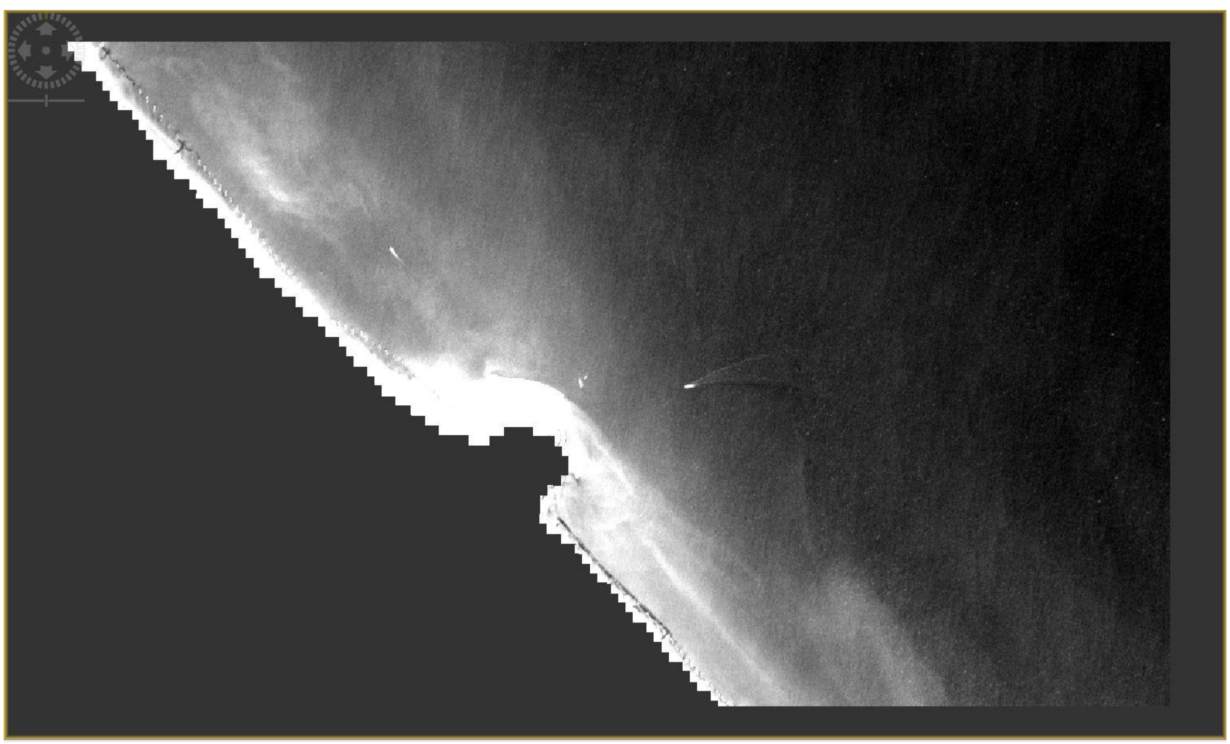


Figure 9 - Channel 3 (560 nm) of the final mask

6. Perform and display channel data correlation of selected ROI

The channels that I will use for the estimation of the **chlorophyll-a (Chl-a)** and also for the total suspended sediments will be B1 (443 nm), B2 (490 nm), B3 (560 nm), so I will display the scatterplot between those bands in this analysis.

As it can be seen in **Errore. L'origine riferimento non è stata trovata.** the channel 1 and 2 have a similar pixel frequency corresponding to the same pixel brightness. This isn't a surprise, because the two bands have a small distance between each other (443 nm and 490 nm).

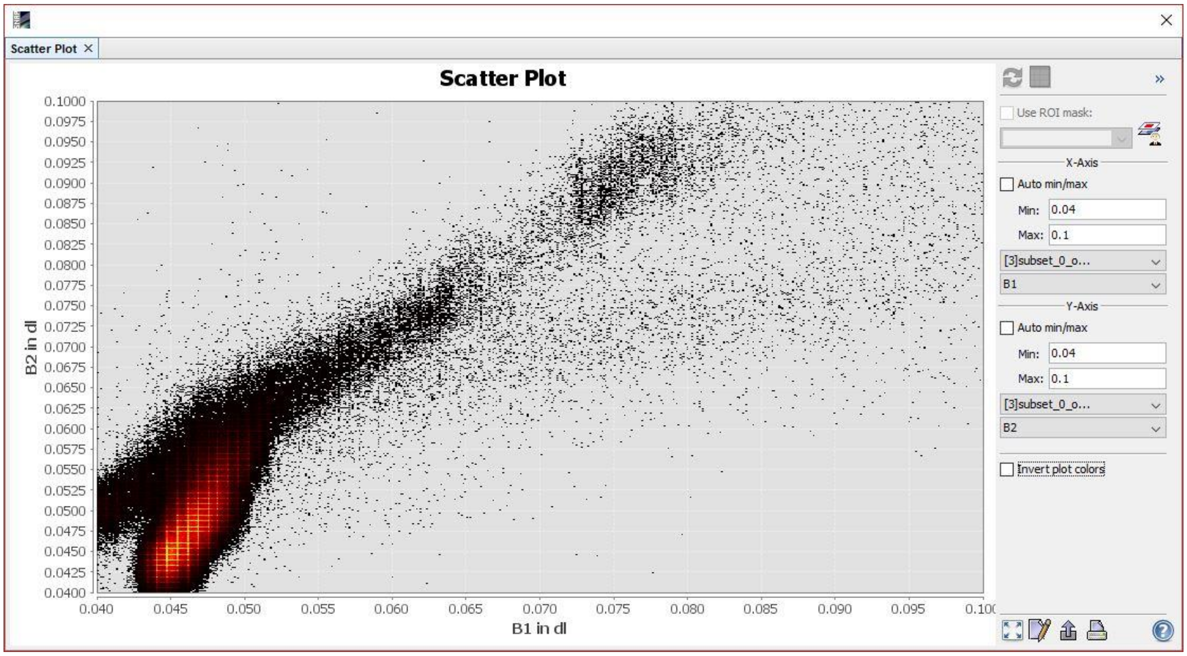


Figure 10 - B1 vs B2 scatterplot

In the case of B1 vs B3 the pixel frequency is not really matching the pixel intensity. As a matter of fact, there is a notable difference between those 2 bands.

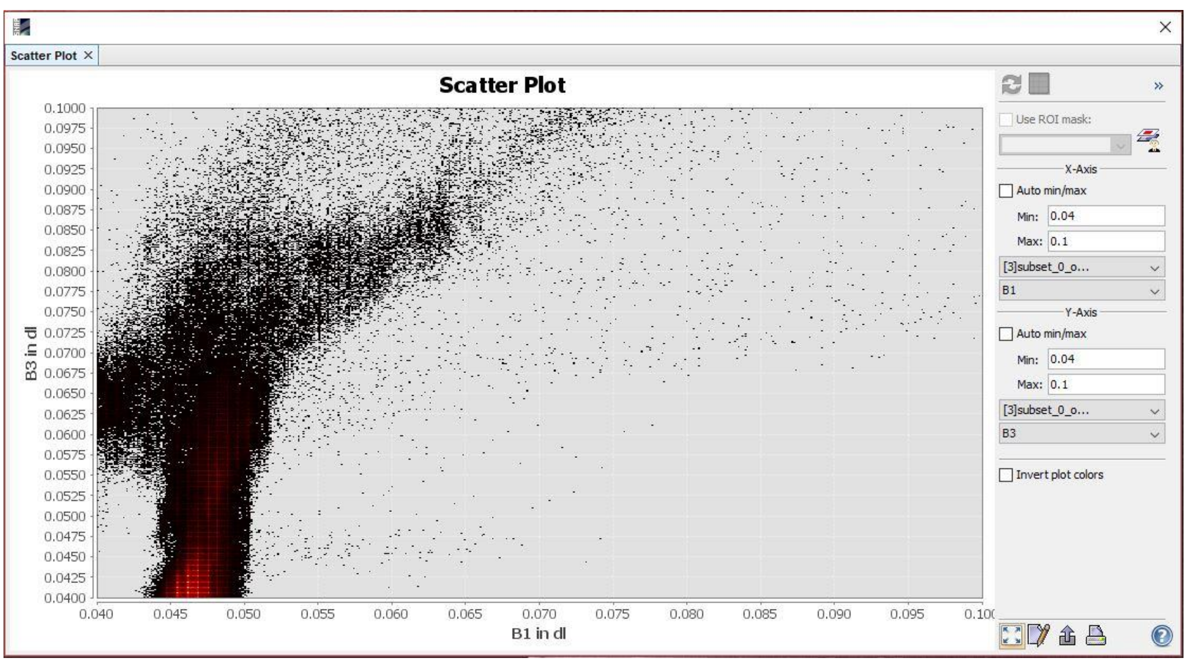


Figure 11 - B1 vs B3 scatterplot

The same argument can be used in the analysis of B2 vs B3, because as I said before the first two channel tends to have a similar behavior.

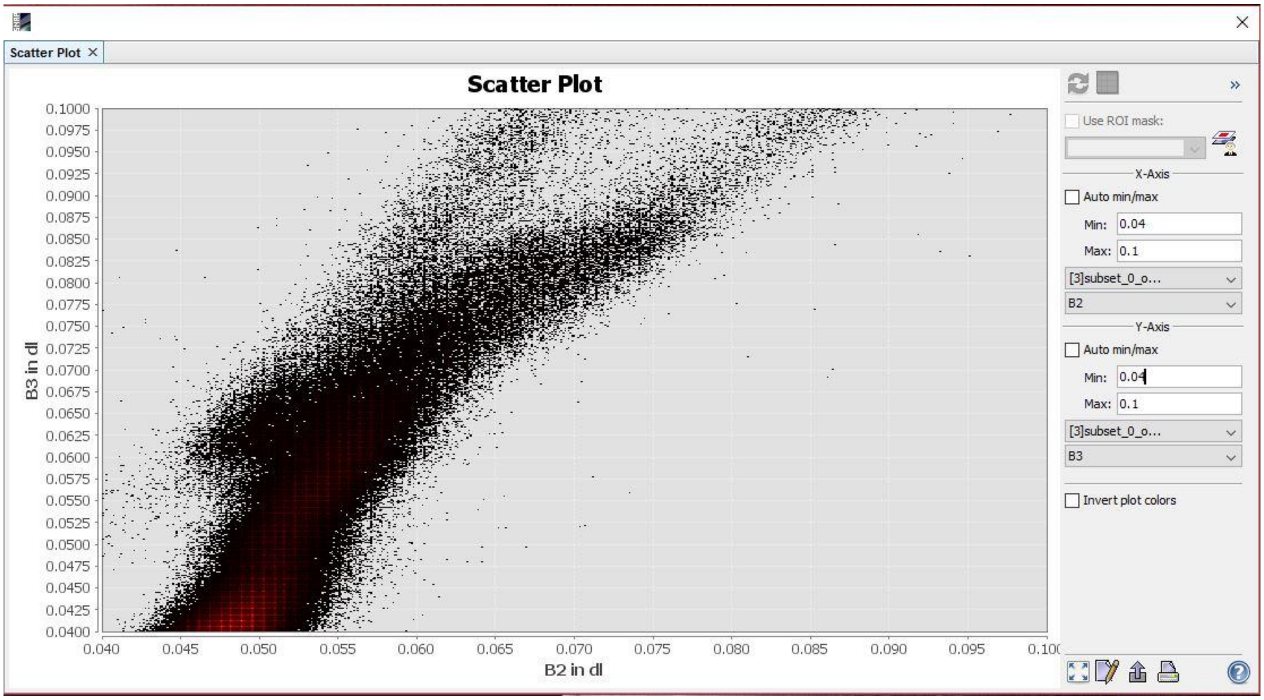


Figure 12 - B2 vs B3 scatterplot

7. Implement at least 3 regressive algorithms to estimate chlorophyll-a (Chl-a) using SNAP processing tool

The first regression formula implemented to estimate the Chl-a is the **NASA-GSFC algorithm**, which returns the near-surface concentration of Chl-a in mg/m^3, calculated using an empirical relationship derived from in situ measurements of Chl-a and blue-to-green band ratios [3]. The formula implemented is a fourth-order polynomial relationship between a ratio of bands and Chl-a:

$$\log_{10}(\mathrm{chlor_a}) = \mathrm{a_0} + \sum_{i=1}^{4} \mathrm{a_i}\left(\log_{10}\left(\frac{\mathrm{R_{rs}}(\lambda_{\mathrm{blue}})}{\mathrm{R_{rs}}(\lambda_{\mathrm{green}})}\right)\right)^{\mathrm{i}}$$

For the MSI of the Sentinel 2 satellite the bands B1 (443 nm), B2 (490 nm) and B3 (560 nm) are provided to retrieve Chl-a. In Figure 13 each pixel represents the variable **MBR**, which indicate the maximum bandwidth ratio.

$$\mathrm{MBR} = \frac{\max(\mathrm{B1}, \mathrm{B2})}{\mathrm{B3}}$$

The idea behind the MBR resides on the physical nature of the Chl-a, a specific form of chlorophyll in oxygen photosynthesis that has a peak of **absorption** around 443 nm (B1 and B2 essentially) and less absorption around 560 nm (B3 band). The MBR is constructed as a blue to green band ratio and has high value where the Chl-a is present.

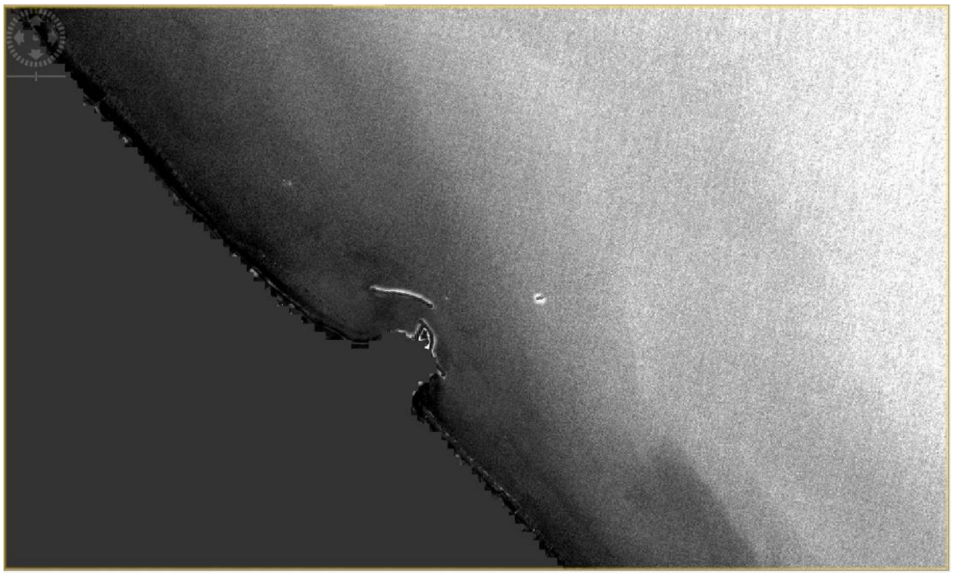


Figure 13 - MBR

In the first algorithm I will use the coefficients of the sensor **MERIS** (algorithm OC4E provided in the table of coefficients of the NASA OceanColor website [3]). The values are:

$$\mathrm{a}_0 = 0.3255, \quad \mathrm{a}_1 = -2.7677, \quad \mathrm{a}_2 = 2.4409, \;\; \mathrm{a}3 = -1.1288, \;\; \mathrm{a}_4 = -0.4990$$

The expression is shown in Figure 14.

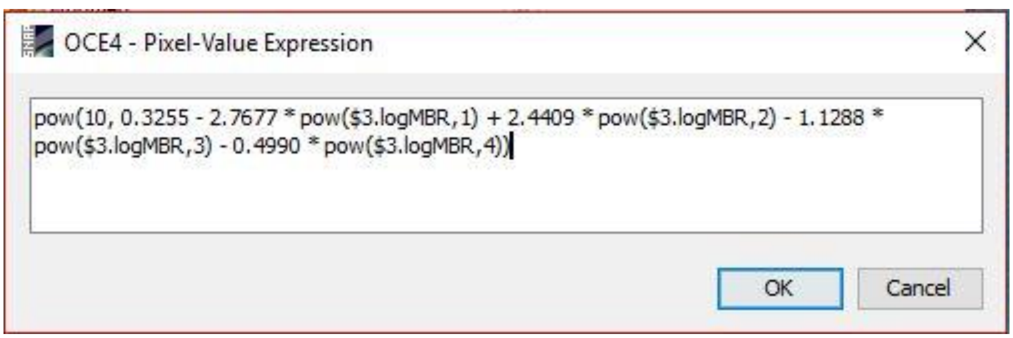


Figure 14 - OCE4 pixel-value expression

For the second algorithm I will use the coefficient of the sensor **OLI/Landsat 8** (algorithm OC4E provided in the table of coefficients of the NASA OceanColor website [3]).

$$a_0 = 0.2412, \quad a_1 = -2.0546, \quad a_2 = 1.1776, \;\; a3 = -0.5538, \;\; a_4 = -0.4570$$

The expression is shown in Figure 15.

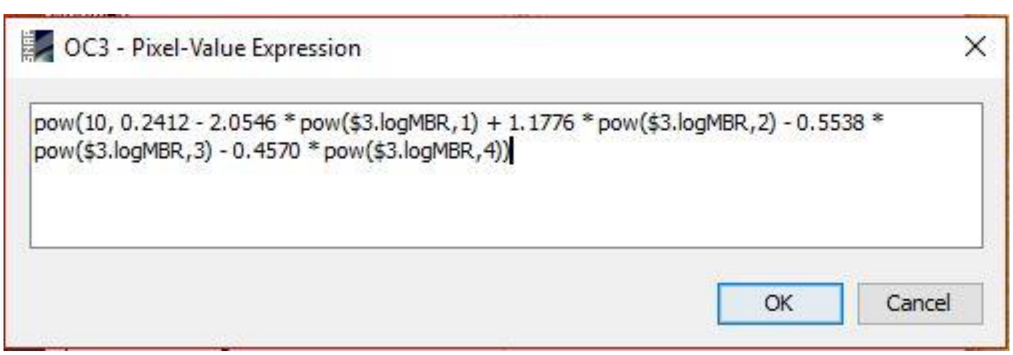


Figure 15 - OC3 pixel-value expression

For the third algorithm I will use 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 $\mu g/l$.

$$\mathrm{SICP_{chla}} = 1.7541 \cdot e^{-1.547 \cdot \mathrm{MBR}}$$

The expression is shown in Figure 16.

Figure 16 - SICP pixel-value expression

8. Apply the Chl-a retrieval algorithms around the Pescara river estuary and compare their results

In Figure 17 the results of the OC4E algorithm can be seen. Using this formula, the water near the coast is clearly not transparent.

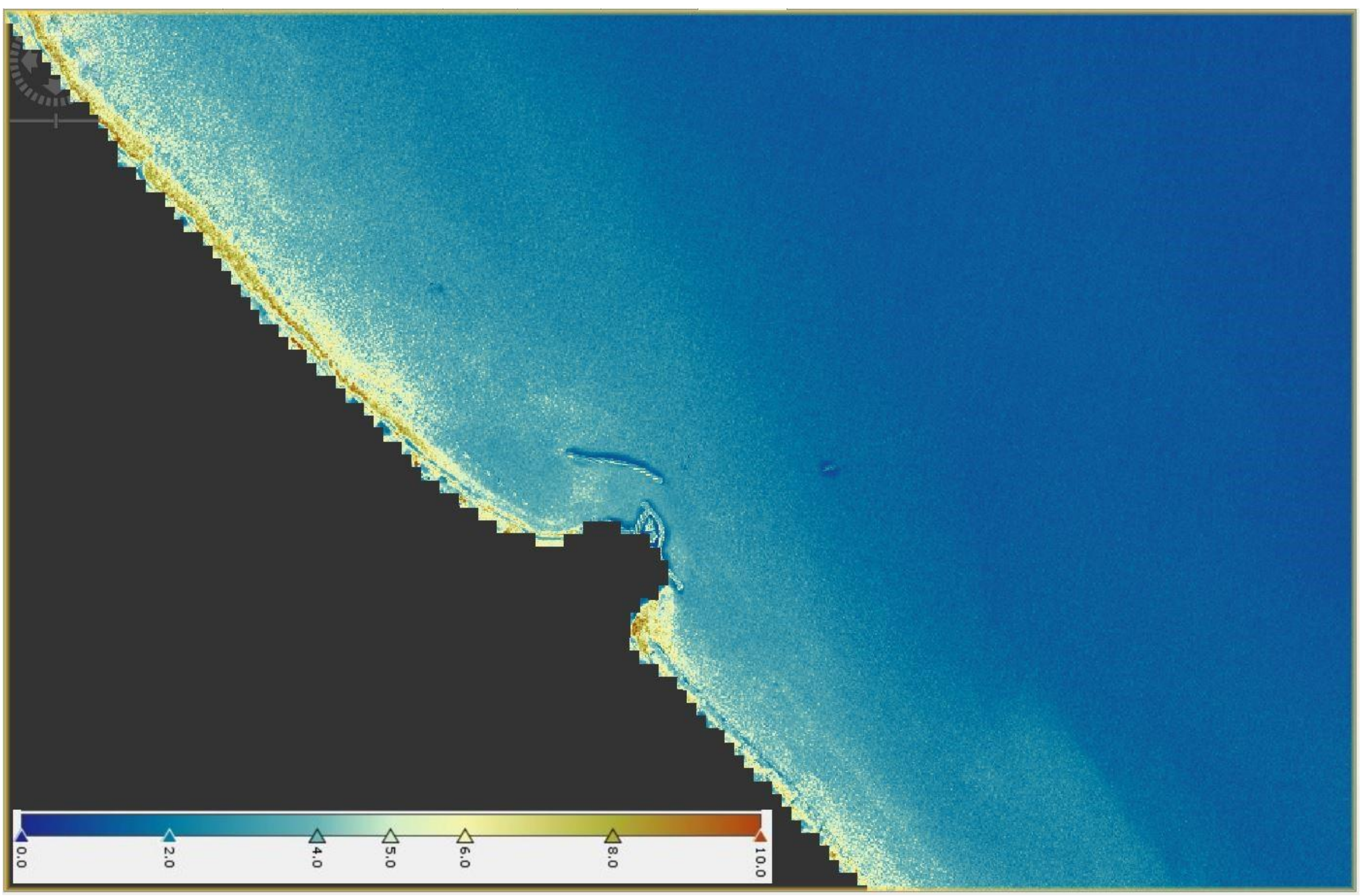


Figure 17 - OCE4

In the Figure 18 statistics of the OCE4 algorithm are shown.

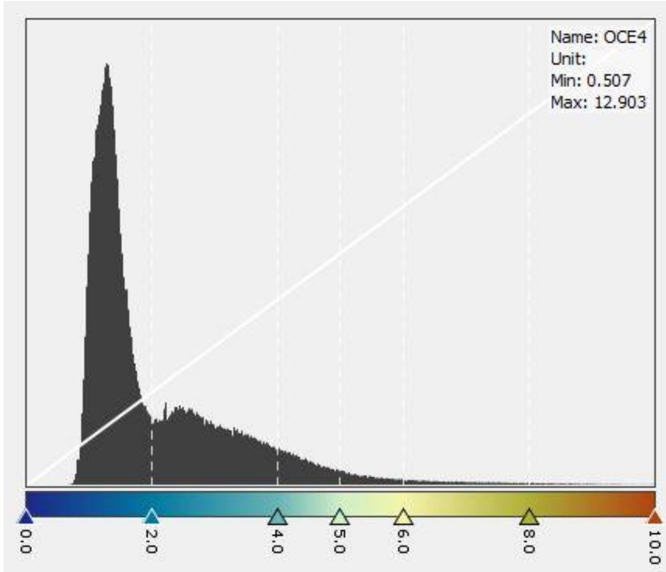


Figure 18 - OCE4 statistics

Around the area of the Pescara river and toward the coasts, there are pixels with value greater than 5, sometimes also over 8. This means that there should be a **significant amount** of Chl-a.

In Figure 19 the results of the OC3 algorithm can be seen. With respect to the previous algorithm, the prediction of the presence of Chl-a is more optimistic and the values are smaller.

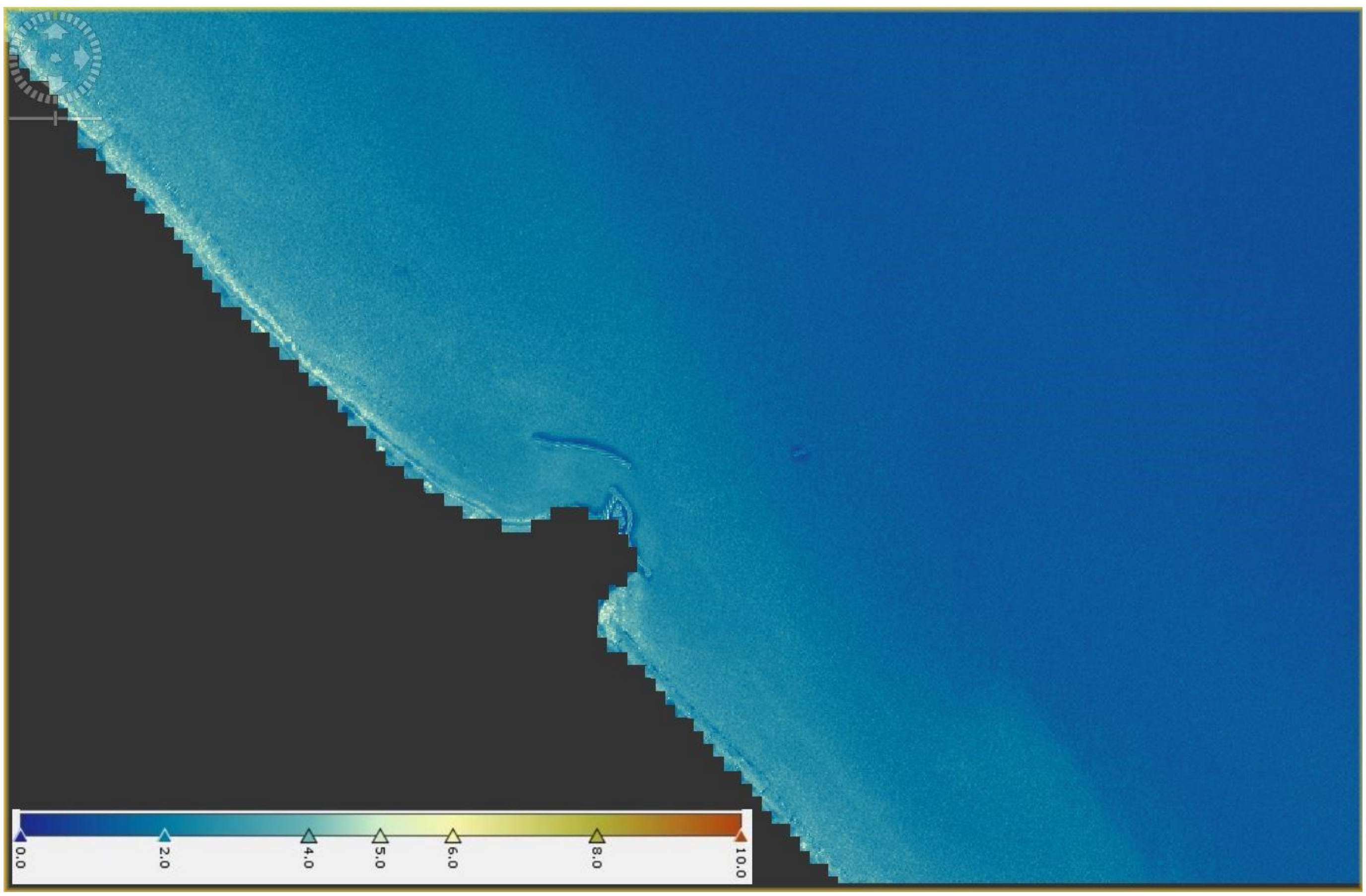


Figure 19 - OC3

In the Figure 18 statistics of the OC3 algorithm are shown. In average, the values of the pixels tend to be concentrated around the mean (lower variance) and there are few pixels reaching a high value (the maximum value is 6.11).

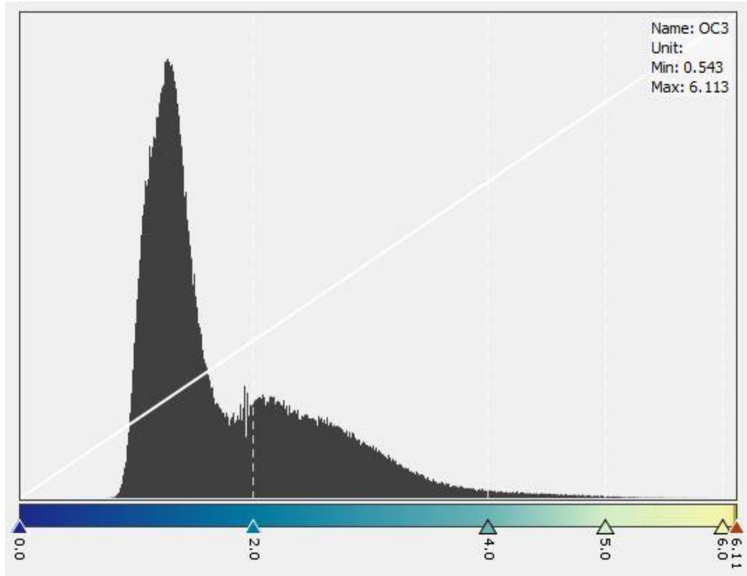


Figure 20 - OC3 statistics

For the results of the SICP algorithm I used a different scale, because the maximum value is around 0.7.

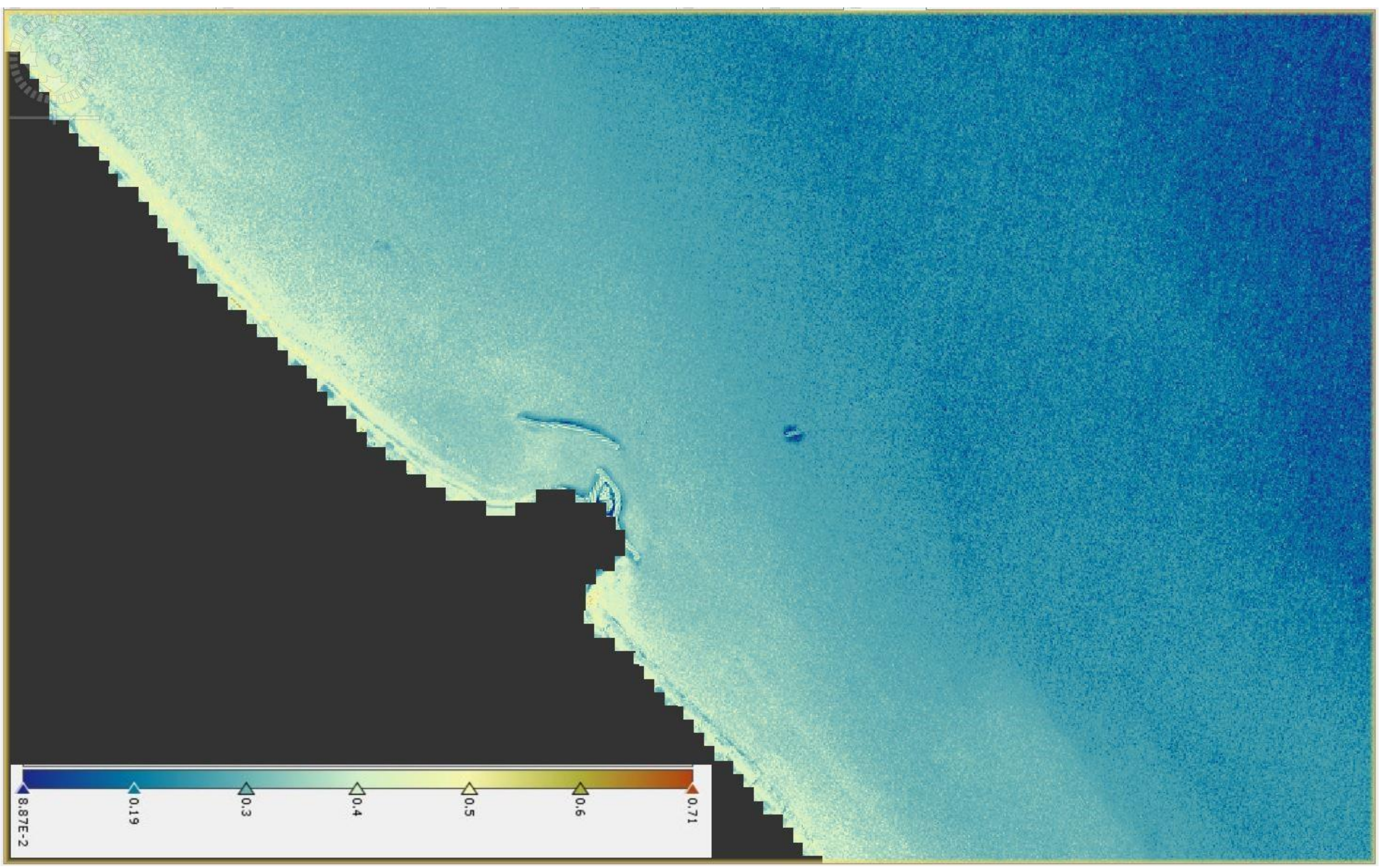


Figure 21 - SICP

In the statistics of the OCE4 algorithm are shown.

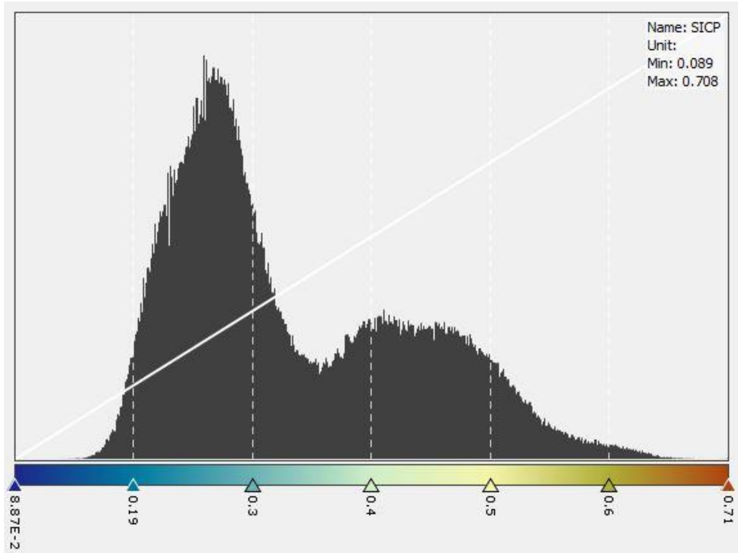


Figure 22 - SICP statistics

To compare the values of Chl-a predicted using the SICP algorithm with the previous formulas, I should convert them in mg/m^3, because they are measured in in $\mu g/l$, but the unit of measure are equivalent. The values predicted are lower with respect to the previous ones, and their distribution is smoother over the see.

All the 3 algorithms predict a higher value of Chl-a **near the coast** with respect to the rest of the sea.

9. Implement at least 2 regressive algorithms to estimate total suspended sediments (TSS) using SNAP

The formula used to estimate the total suspended sediments are retrieved empirically using a linear regression model. Those results are based on the relation between **the TSS and the NIR reflectance**. The first formula (I will call it TSS1) uses the band B2 and B4 and is shown in Figure 23.

Figure 23 - TSS1 formula

Another TSS algorithm is implemented, using also the B3 channel. The TSS2 formula is showed in Figure 24.

Figure 24 - TSS2 formula

10. Apply the TSS retrieval algorithms around the Pescara river estuary and compare their results

The results of the implementation of the TSS1 algorithm are shown in Figure 25.

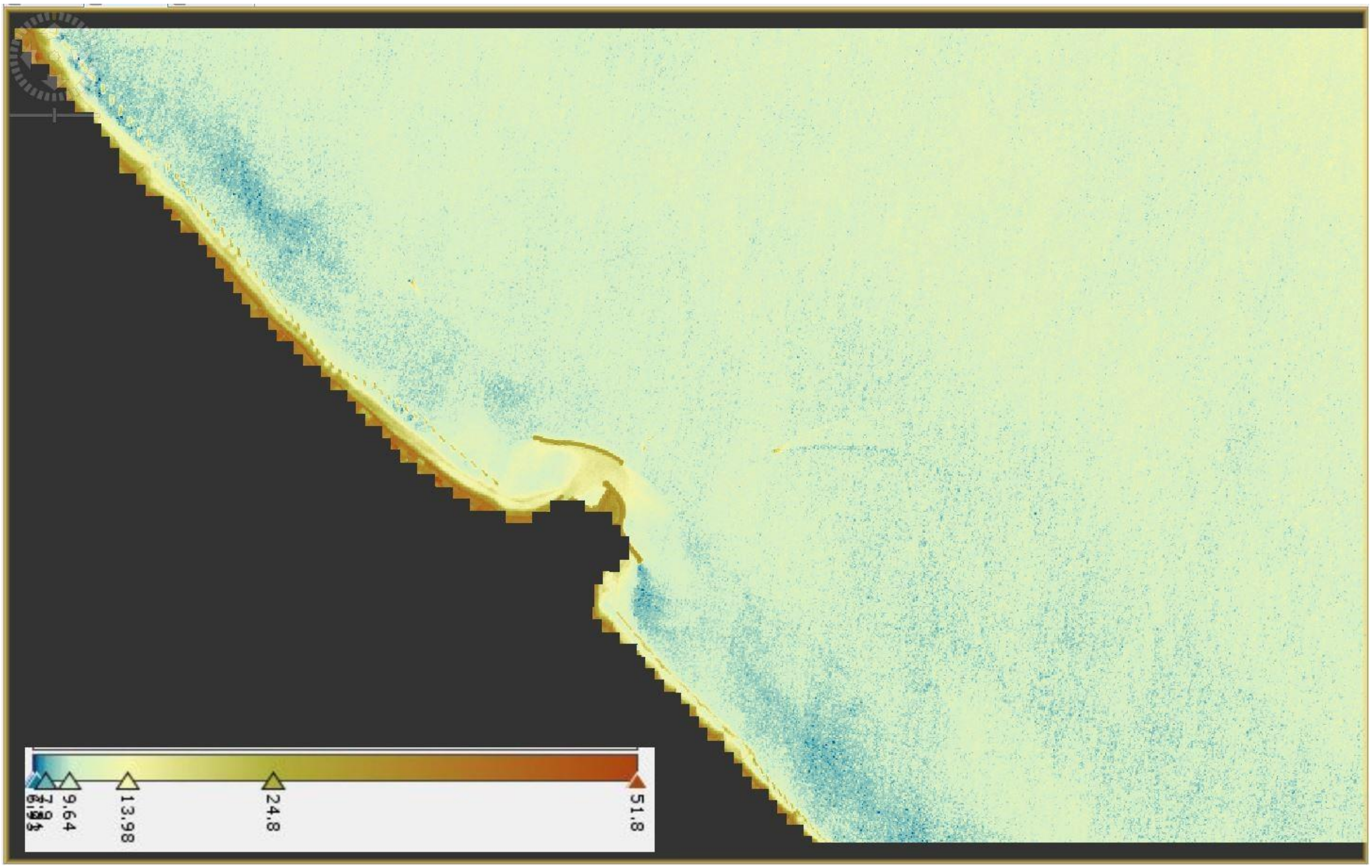


Figure 25 - TSS1 results

The values near the coast are very high, but they are mostly associated to the beach or to the boats and they **do not** indicate a high presence of sediments. However, near the coast the pixels have high values and given their distribution it is reasonable to think that the concentration of sediments is high.

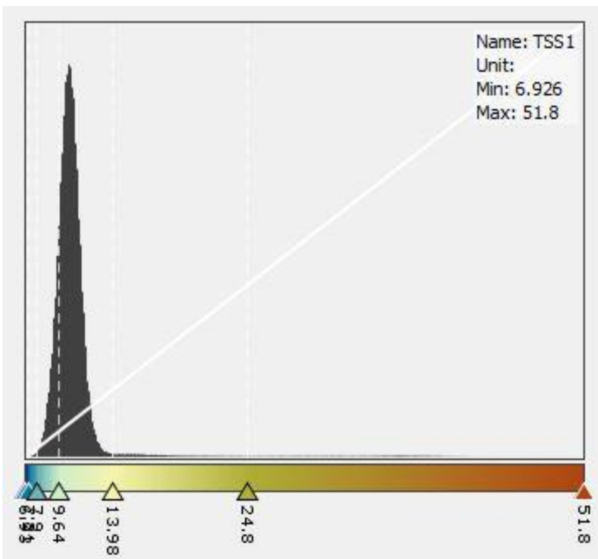


Figure 26 - TSS1 statistics

In Figure 27 the results of the TSS2 algorithm are showed. The values are similar to the previous ones, but maybe more realistic, because the pixels of the sea far from the coast have a lower value with to respect of those of TSS1. I think that is more reasonable to have less sediments far from the coast and I will use the TSS2 algorithm on the final analysis.

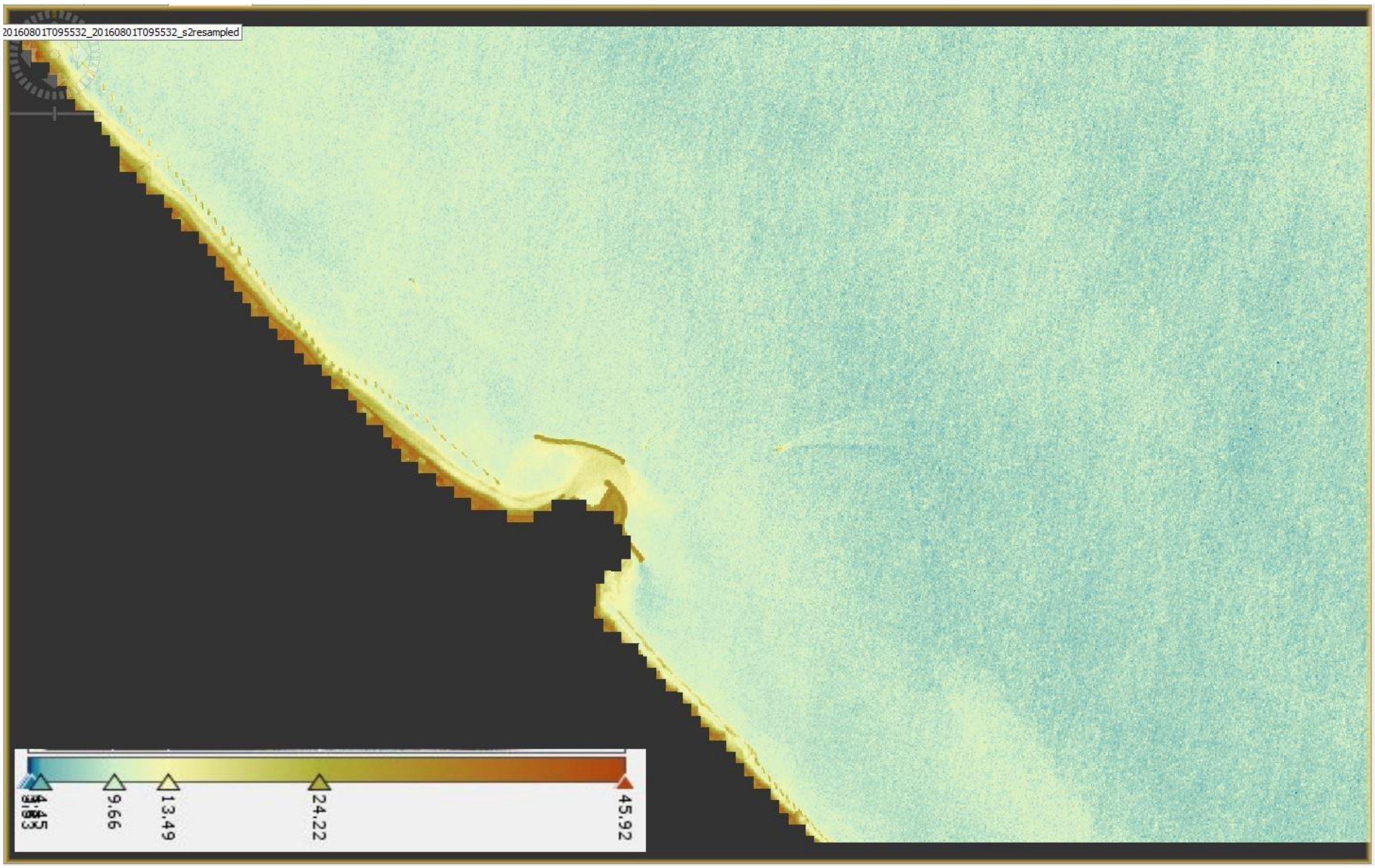


Figure 27 - TSS2 results

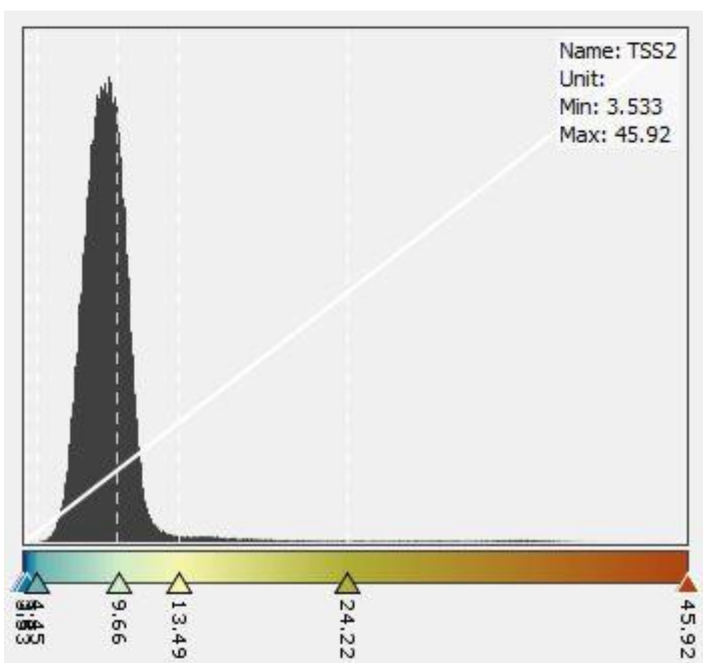


Figure 28 - TSS2 statistics

11. **Download all MSI/S2 images over Central Italy (Pescara river estuary, 42.46° N, 14.21° E) within the period 5-12 August 2016 from https://scihub.copernicus.eu. You can also choose a different time interval on a different target area (e.g., Tiber river estuary in Italy, Po river delta in Italy, ...).**

The downloaded products were displayed in point 2, using the extended time range from 5-24 August.

12. **Apply the chosen Chl-a and TSS retrieval algorithm to MSI image time series around the Pescara river estuary and show/discuss the time evolution of the retrieved parameters.**

All the images showed in point 2 are used here in the final analysis. The algorithms implemented are the **SICP** to retrieve an estimate of the presend of Chl-a and the **TSS2** to estimate the presence of suspended sediments. I used the graph builder tool to choose the exact **same area** tover the 4 products proceeding in this way: first I create a subset in a specific area around the Pescara river estuary and then I mask out the land. In Figure 29 there is an example of the graph, while in Figure 30 the final products are showed (RGB image) in increasing order of date from left to right and from top to bottom.

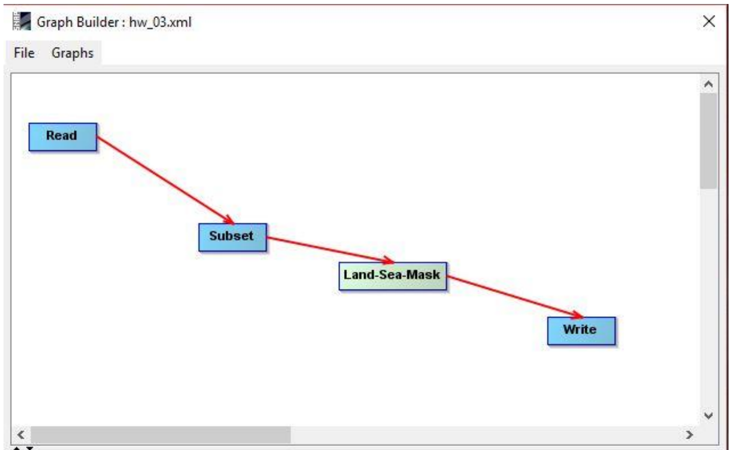


Figure 29 - graph builder

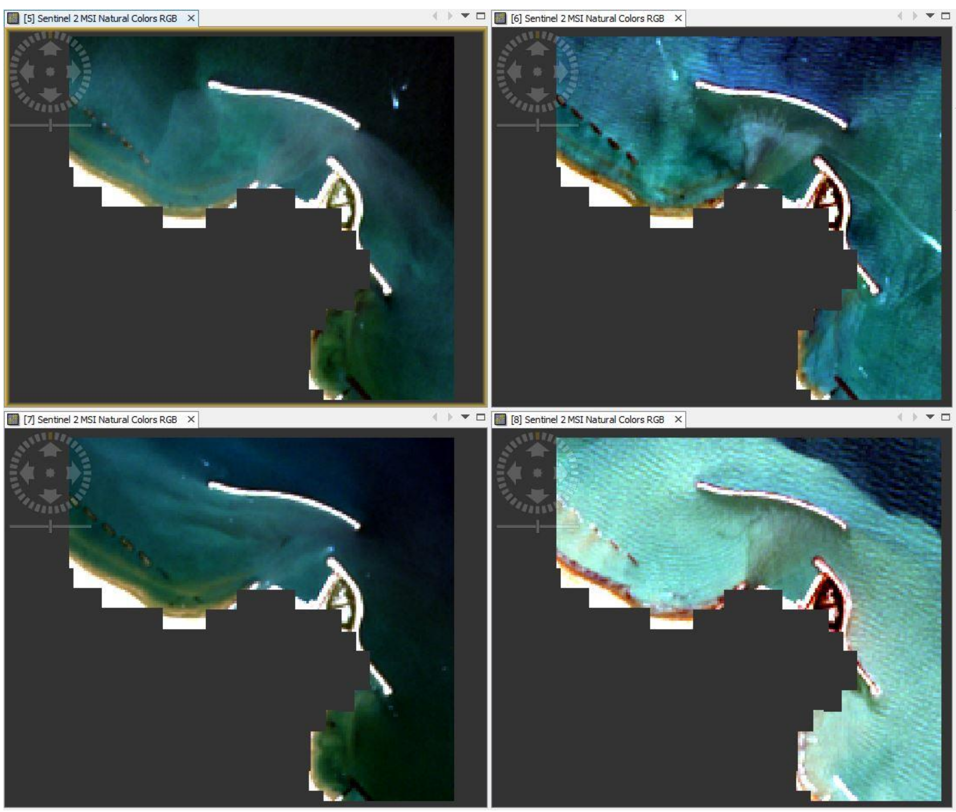


Figure 30 - results of the graph

In the final version of the graph I implemented the SICP (BandMaths and BandMaths2) and the TSS2 (BandMaths3 and BandMaths4), merged the BandMerge tool.

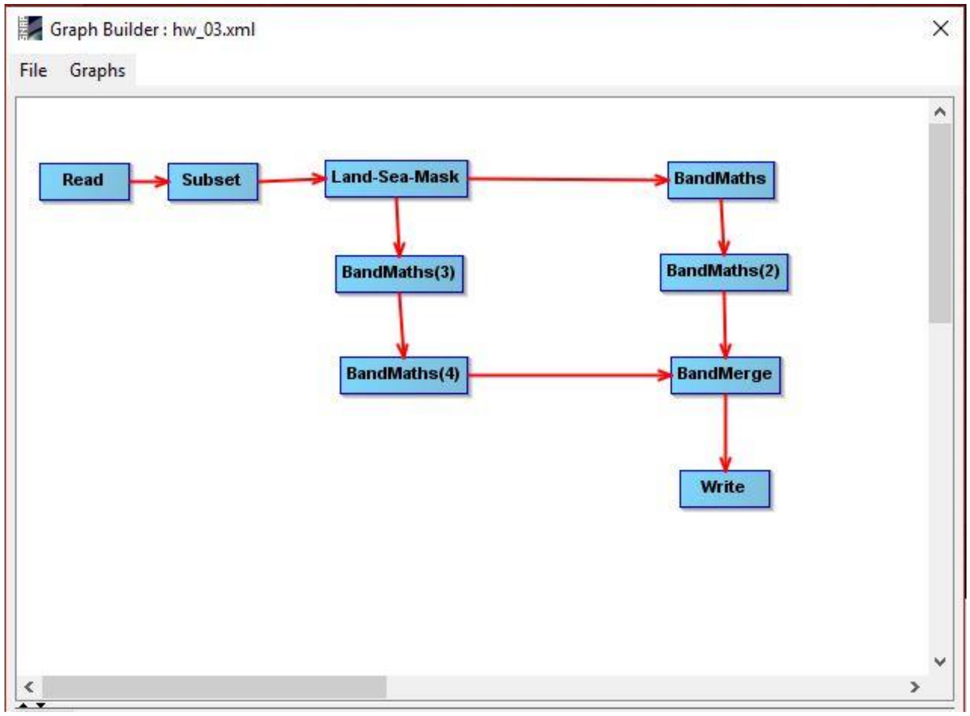


Figure 31 - final graph

To apply the graph on each of the 4 products I used the **batch processing tool**, which allow to iterate the graph for every date.

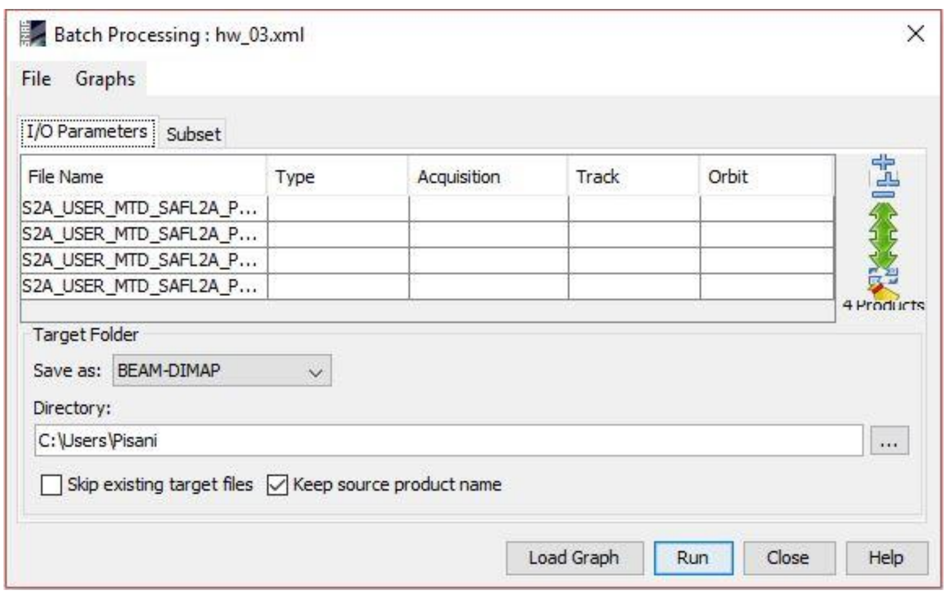


Figure 32 - batch tool

The results of this analysis are showed in Figure 33 and Figure 34.

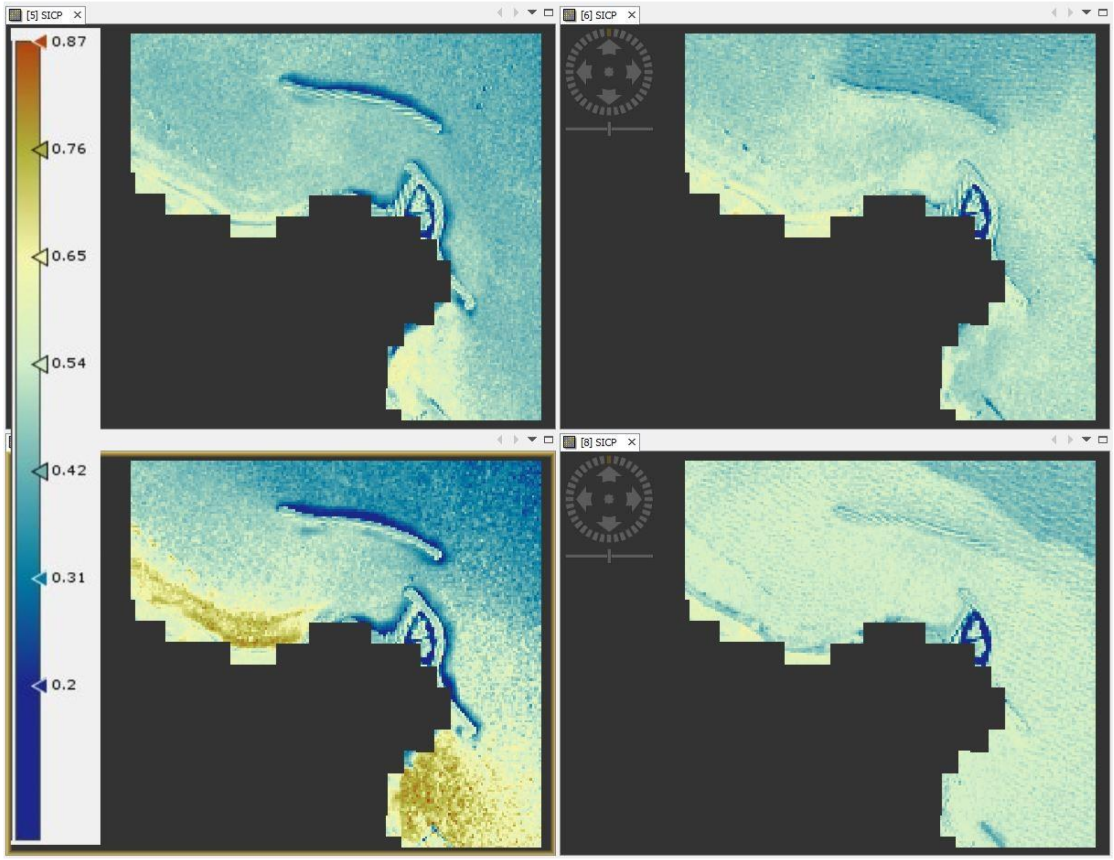


Figure 33 - SICP algorithm applied to all products

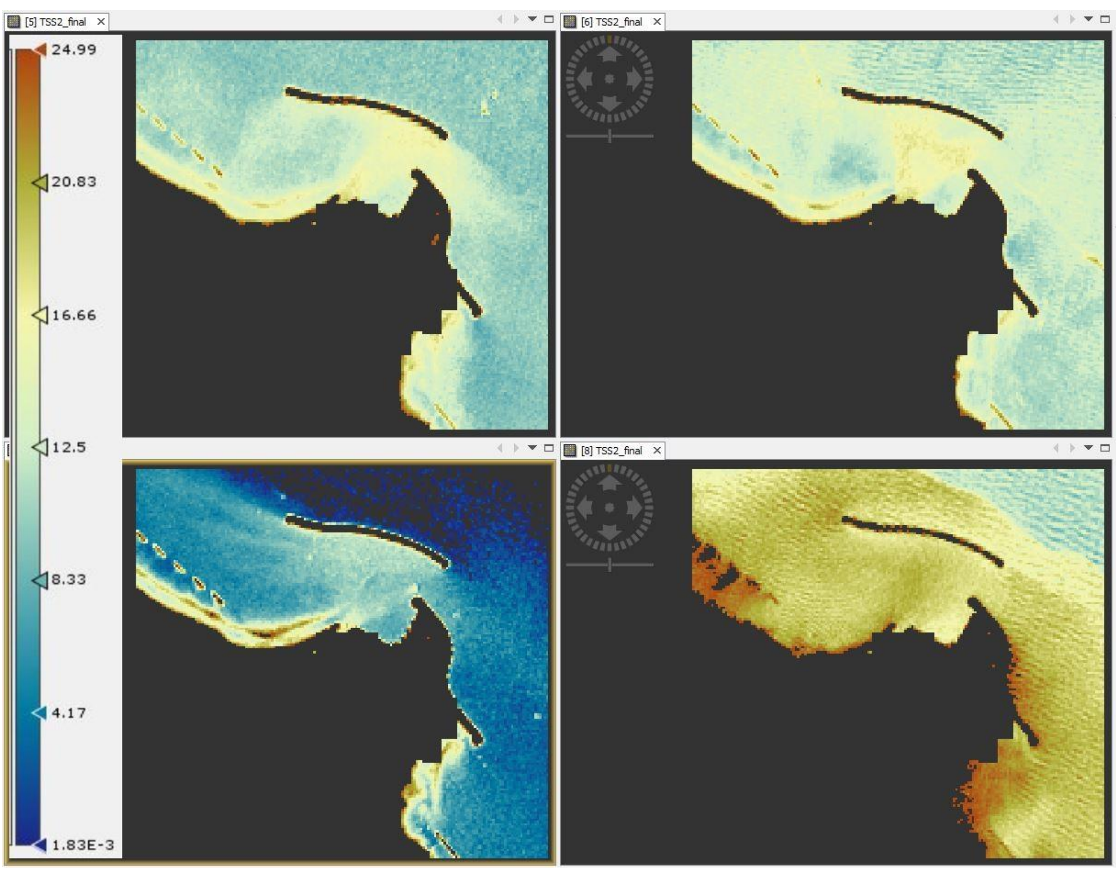
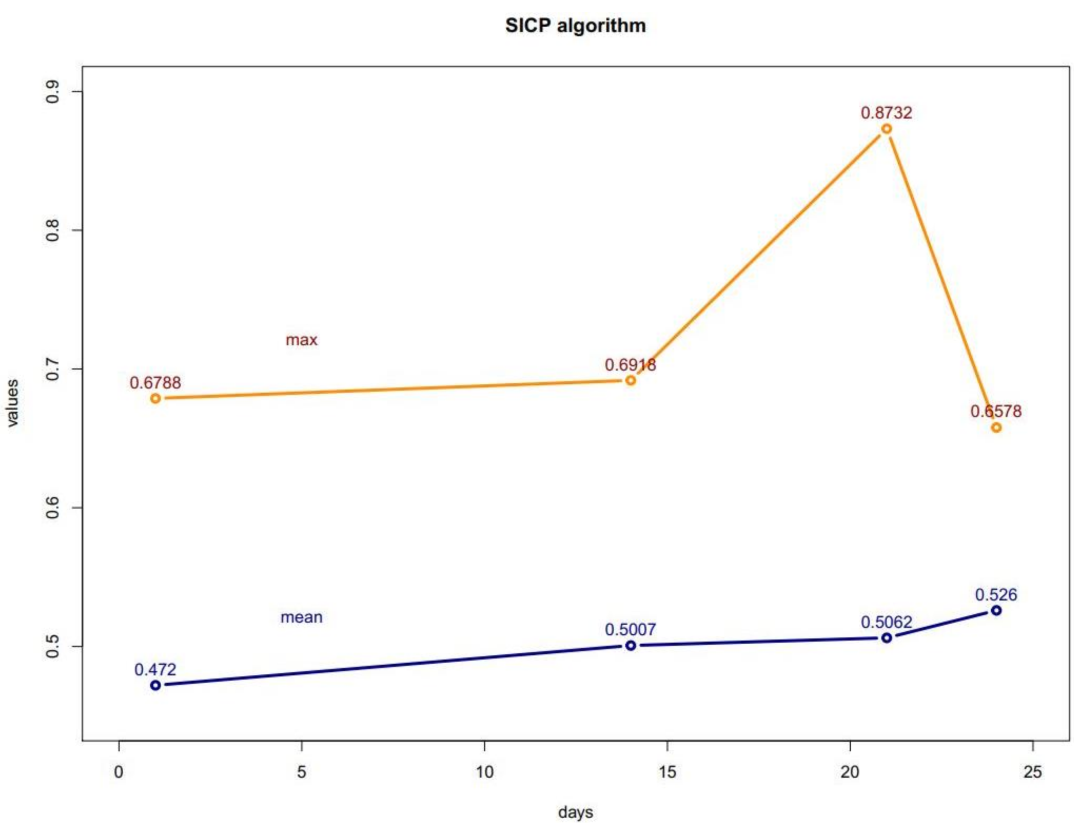


Figure 34 - TSS2 algorithm applied to all products

To analyze the time series results I will use the mean value for each product and show all of them in a single plot. I exported the statistics of every result in a csv file, imported all of them in RStudio and compute a plot using the R programming language.

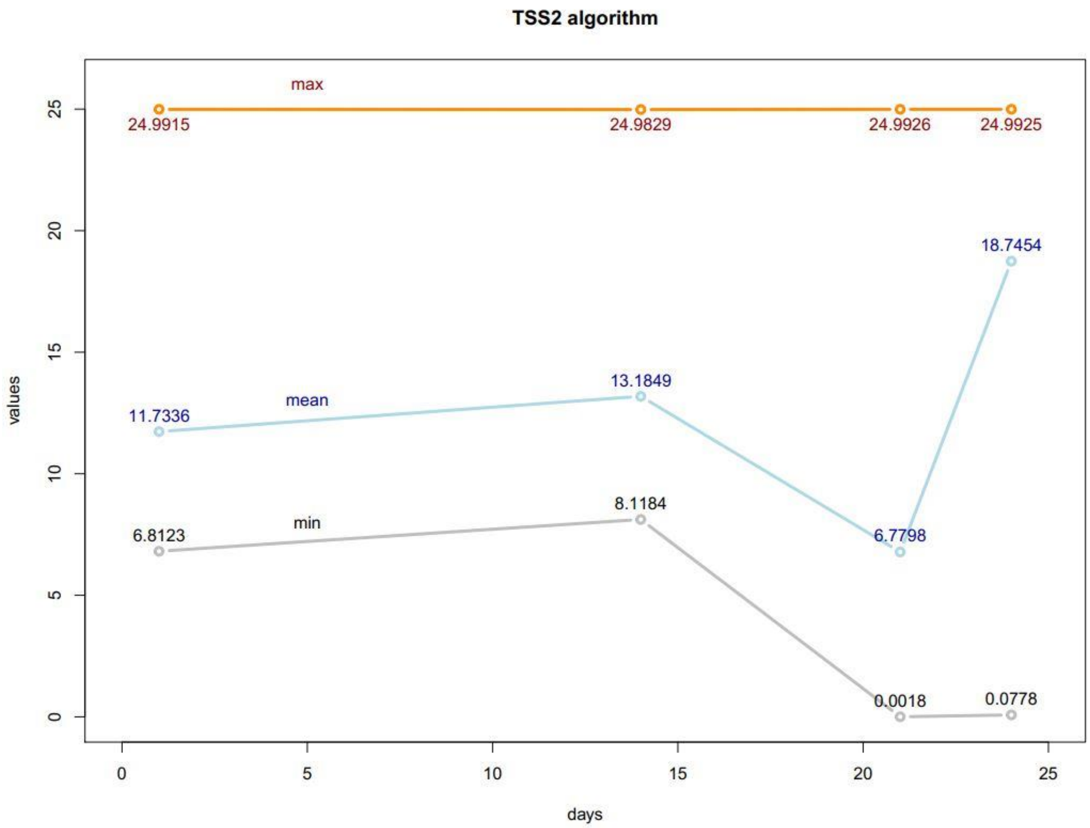
```
#R code for the SICP results plot
#defining vectors
mean_values = c(0.4720, 0.5007, 0.5062, 0.5260)
days = c(01, 14, 21, 24)
max = c(0.6788, 0.6918, 0.8732, 0.6578)

#plot
plot(days, mean_values, type = 'b', ylim= c(0.45,0.9), col="darkblue", lwd=3,
        ylab = "values", xlim=c(0,25), main = "SICP algorithm")
points(days,max, type = 'b', col="darkorange", lwd=3)
text(5,0.52, "mean", col="darkblue")
text(days, mean_values, mean_values, pos = 3, cex = 0.75, col = "darkblue")
text(5,0.72, "max", col="darkred")
text(days, max, max, pos = 3, cex = 0.75, col = "darkred")
```

The execution of the previous chunk of code plot the following image. I decided to show only the mean value of the pixels and the maximum value of the pixels, because the lowest value is the area of shadow near an artificial structure and it's not what we are interested in.



The mean value tends to **grow** from the 1st of August to the 24th of August, while the maximum value has a peak on the 21st of August. This can be seen in the third image (bottom left) where the area near the coast is brown (high presence of Chl-a). From Figure 33 it's clear that the maximum values are the ones that show the Chl-a near the coast.

The code to plot the results of the TSS2 algorithm is similar to the previous one, and the values are obtained using the statistics exported from SNAP.



The maximum values of the TSS2 are distributed around the coast and they show a constant level of sediments. The mean value has a negative peak on the 21st of August, but looking at the image it's clear that the sediments are still present near the coast and only the values far from there are lower.

Following the evolution of the Chl-a and the sediments during August 2016, clearly there's a **growing pattern** in both cases. I made this statement considering the mean value and the fact that even if the sediments are decreasing the 21st of August, the values of the pixels near the coast are high.

13. Download all MSI/S2 Chl-a and TSS products over Central Italy (Pescara river estuary, 42.46° N, 14.21° E) within the period 5-12 August 2016 from Copernicus Marine Service at http://marine.copernicus.eu

I used the web navigator to analyze the results of the Copernicus Marine Service and compare the presence of Chl-a and sediments with my results. Unfortunately for the chosen period (August 2016) there **are no results for the sediments**, which are available only after January 2017. In the next point I will discuss the results of the presence of Chl-a.

14. Compare Chl-a and TSS retrievals with official Copernicus S2 products within the selected period

In Figure 35 the results of the concentration of chlorophyll in the sea water are showed. The period analyzed goes from the 1st of August to the 1st of September. In this range, there are only 2 measurements at the beginning of each month at a depth of -1.47m from the surface.

The difference in the values from my previous results and the official ones may be caused by the analysis of generic chlorophyll instead of chl-a, or to the fact that the area of interest is more extended and takes into account also the water far from the coast.

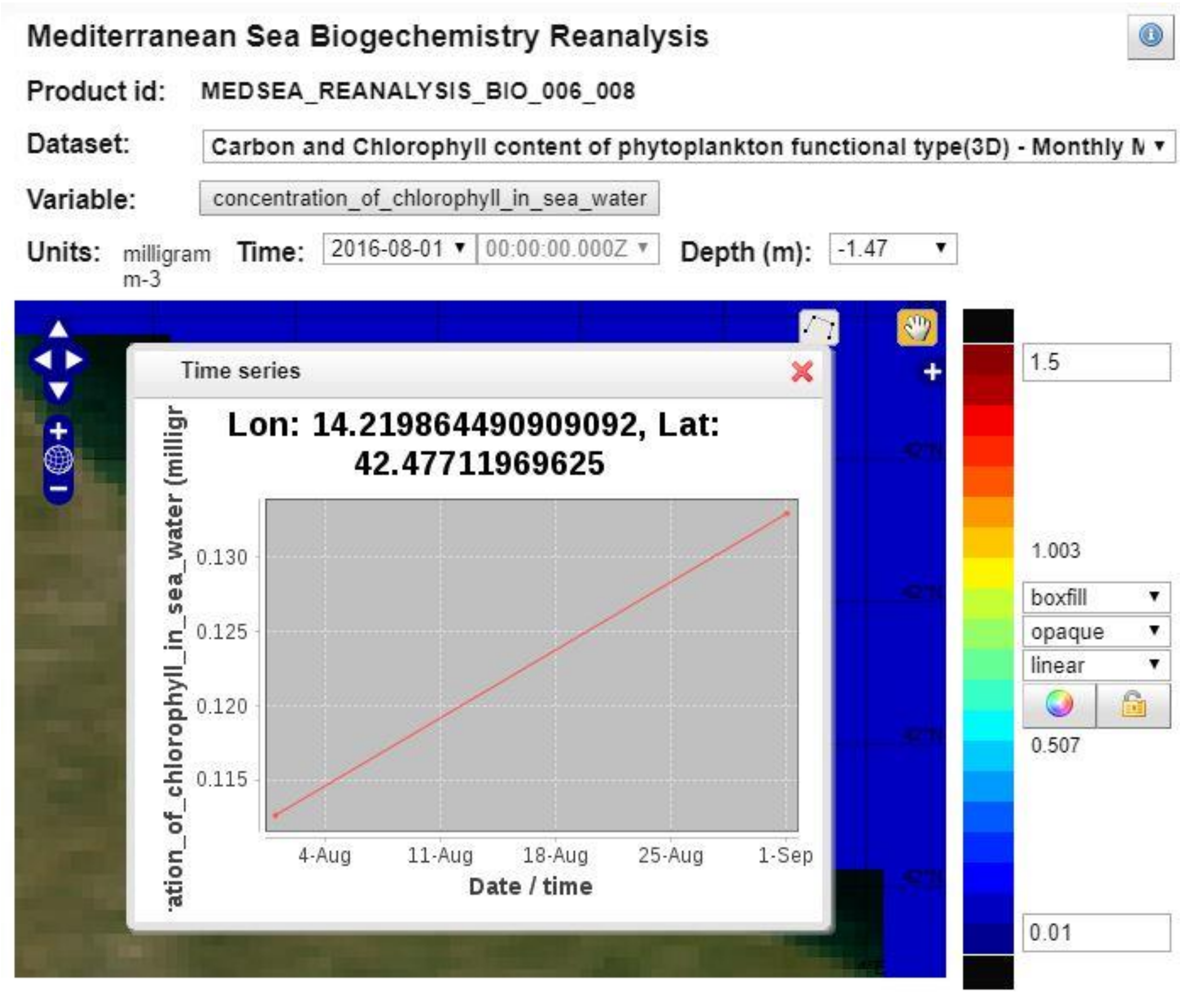


Figure 35 - chlorophyll results from Copernicus Marine Service portal

Despite the difference in the values, the general trend in August 2016 is like the previous analysis. The level of chlorophyll has grown during the month in both the official results and my analysis.

The Copernicus portal result show a growth of chlorophyll from 0.115 to 0.130, more or less **13%**, while in my analysis the growth goes from 0.472 to 0.526 more or less 11%. Despite the different values the overall growth of chlorophyll is similar.

Bibliography

[1] **sentinel-2A satellite sensor,** https://www.satimagingcorp.com/satellite-sensors/other-satellite-sensors/sentinel-2a/

[2] **step esa,** http://step.esa.int/main/third-party-plugins-2/sen2cor/

[3] **NASA OceanColor**, https://oceancolor.gsfc.nasa.gov/atbd/chlor_a/