Sapienza Università di Roma

EODA

HW 2: SURFACE MAPPING FROM MSI SENTINEL 2 DATA

Author:
Francesco Romeo

Supervisor:
Frank S. Marzano

August 2020

1 DATA EXPLORATION

In order to accomplish the task of performing surface mapping, sentinel 2 data will be used. Sentinel 2 is a passive system (Optical Remote Sensign), meaning that it works with the fraction of reflected sunlight by targets on earth. In optical remote sensing, multispectral images are available with no distortion effects (as instead for Active system Sentinel-1), however it is limited by clouds and absence of light. Sentinel 2 is polar orbit, sun synchronous (meaning that covers a specific area almost at the same time), multispectral, high resolution missions composed by a constellation of 2 satellites, Sentinel 2A and Sentinel 2B, which are in the same orbit but in opposite positions.
The used products are MultiSpectralImage (MSI), products atmospherically corrected (Level 2 data, i.e BOA data, where the atmospheric effect has been already removed).
Sentinel 2 acquisitions provide products with 13 bands spanning from 443 nm wavelenght to 2190 nm. Every band can be used to better distinguish different targets in the image; for example, by using band 8 which is in NIR (nearinfrared, 842 nm) it is possible to easily see rivers, roads and highways. This due to the fact that in NIR these targets absorb a lot and reflect very few.
For this task two acquisitions has been chosen, representing the Lazio coast area, more precisely in the proximity of Rome. The acquisitions date back to March 2^{nd} and September 13^{th} of 2019. Both products belong to the same area, and this can be checked by the fact that the tile, reported in the name of the products, is the same (T32TQM). Below are reported the two acquisitions:

- **Winter:** S2A_MSIL2A_20190302T100021_N0211_R122_T32TQM_20190302T142827.SAFE
- **Summer:** S2B_MSIL2A_20190913T100029_N0213_R122_T32TQM_20190913T142855.SAFE

1.1 For both summer and winter MSI images, combine B2, B3, B4 channels to create an RGB image and zoom in to observe visible details.

In order to get an RGB composite is necessary to combine red (band 4), green (band 3, 560 nm) and blue (band 2, 490 nm). The RGB image can be returned by using the **RGB Image Window** operator implemented in SNAP. Figures 1 and 2 show the RGB images for both summer and winter acquisitions zoomed in Fiumicino area, from which it is possible to distinguish the Tevere outlet, roads, agricultural fields, urban and vegetation areas.

Figure 1: Winter acquisition

Figure 2: Summer acquisition

1.2 For both summer and winter MSI images, use L2 MSI products for scene classification masks (cloud shadows, vegetation, bare soil, water, snow and ice, ...) and build/visualize an RGB using bare soil, vegetation and water respectively.

Instead of assigning a band to return an RGB composite, it is also possible to assign masks. Bare soil, vegetation and water masks are already present in the products under the section Masks. In order to return an RGB composite using the masks, it has been used the **RGB image window** where for the red it has been assigned the bare soil mask, for the green the vegetation mask and for the blue the water mask. Figures 3 and 4 show the results of the composite. As it can be seen from the Figures, it is possible to distinguish the areas with water, vegetation and bare soil contents. Significantly differences in vegetation can be distinguished from summer and winter even using this simple image processing.

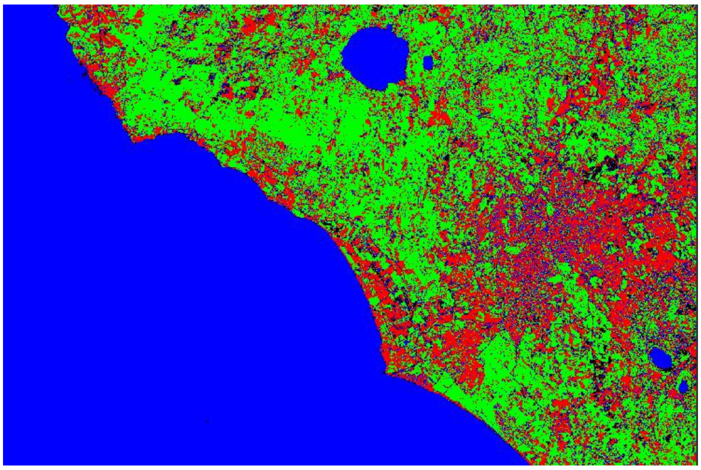

Figure 3: RGB using masks for Winter acquisition

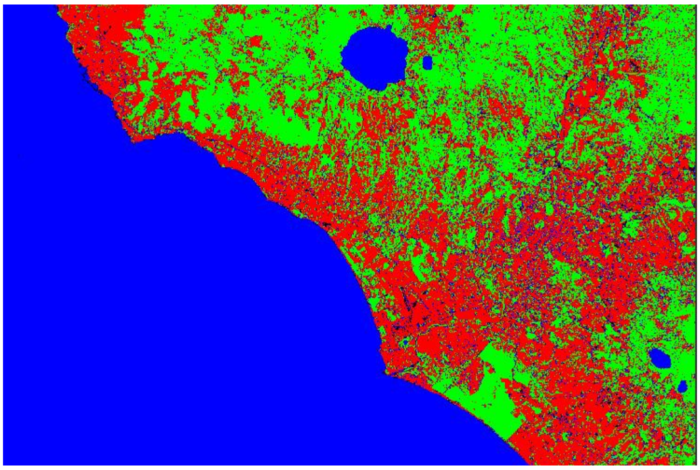

Figure 4: RGB using masks for Summer acquisition

1.3 For both summer and winter MSI images, using SNAP Graph Builder tool, subset around a lake/river and resample all B2, B3, B4 channels at 10 m and visualize them.

Before subsetting around a certain area of interest, it is necessary to perform the resampling operation because it is fundamental that all the bands will be at the same spatial resolution. In order to do that a graph builder has been used in order to create a workflow used to resample the data at 10m resolution and to subset around the Bracciano lake.
Figure 5 shows the workflow.

Figure 5: Workflow to resample and subset around Bracciano lake

- **Read:** takes in input the products
- **Resample:** it is used to resample all the bands at 10m resultion. For bands that were at 20 or 60 m the Nearest upsampling method has been used.
- **Subset:** it is used to subset around the Bracciano lake using the following polygon shape: POLYGON ((12.13038325207106 42.17778178213108, 12.355422725033566 42.18319422900609, 12.360862109468073 42.05232777553797, 12.136284015699527 42.04693995549425, 12.13038325207106 42.17778178213108))
- **Write:** writes on disk the output product.

Figures 6 and 7 show the resampled and subsetted RGB images around Bracciano lake.

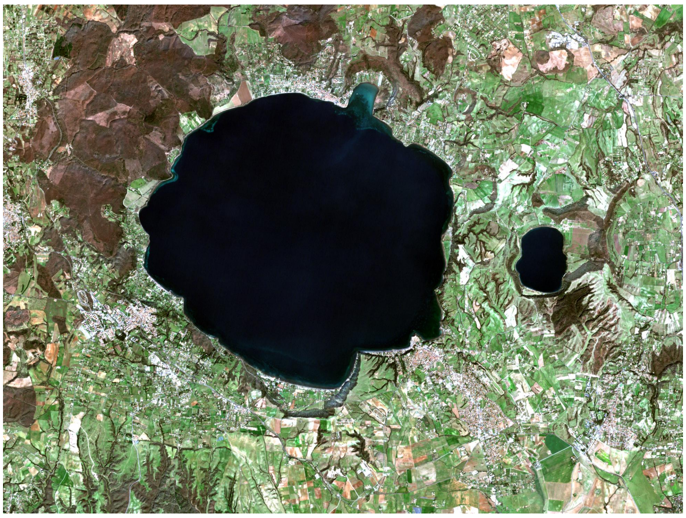

Figure 6: RGB Winter acquisition

Figure 7: RGB Summer acquisition

2 WATER AND VEGETATION NORMALIZED INDEX

2.1 For both summer and winter MSI resampled images, compute NDWI on winter product and compute NDWI2 showing that NDWI puts in evidence water content in the leaves while NDWI2 can be used to build a "water mask" estimate chlorophyll-a (Chl-a) and total suspended sediments (TSS).

In this part, two indexes will be computed to put in evidence the water content. More precisely, NDWI (normalized difference water index) will be the index that evidences the water content in the leaves, whereas the NDWI2 (second normalized difference water index) will evidence the inland water surface. Below is reported how NDWI and NDWI2 are computed:

$$NDWI = \frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}} \quad (1) \qquad NDWI2 = \frac{\rho_{Green} - \rho_{NIR}}{\rho_{Green} + \rho_{NIR}} \quad (2)$$

The NDWI index algorithm, developed by Gao 1964, can be computed by using the **NDWI operator** implemented in SNAP that can be found under the path: OPTICAL/THEMATICC LAND PROCESSING/WATER RADIOMETRIC INDECES. As MIR source band the B12 band has been used and for NIR source band the B8 band.
As it is possible to see from Figure 10, after a color manipulation (shown in Figure 8), the white points are the one to which a very high NDWI value is associated. It could seem unusual that the highest values are widely present in the middle of the lake but this depends on the fact that Bracciano lake is famous to have very long seaweeds that reach the water surface but also water lily that cover wide water surface areas. So, because this index measures the water content in the leaves, it becomes very useful to monitor and measure the vegetation status in a region during different periods of the year. In fact, by comparing the NDWIs of both summer and winter acquisitions, it is possible to see how the vegetation changes around the lakes. This change is also supported by the RGB images (Figures 6 and 7) reported above.

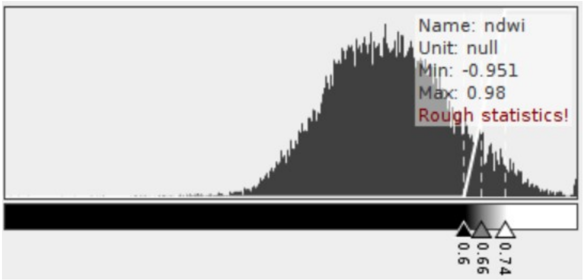

Figure 8: Color manipulation NDWI winter

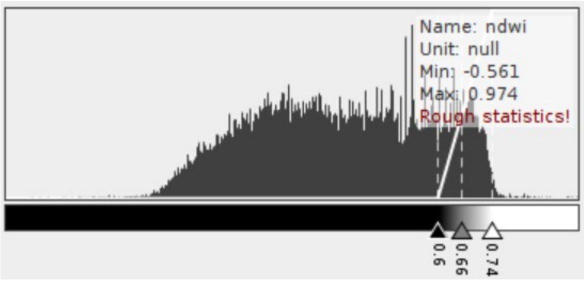

Figure 9: Color manipulation NDWI summer

Figure 10: NDWI winter

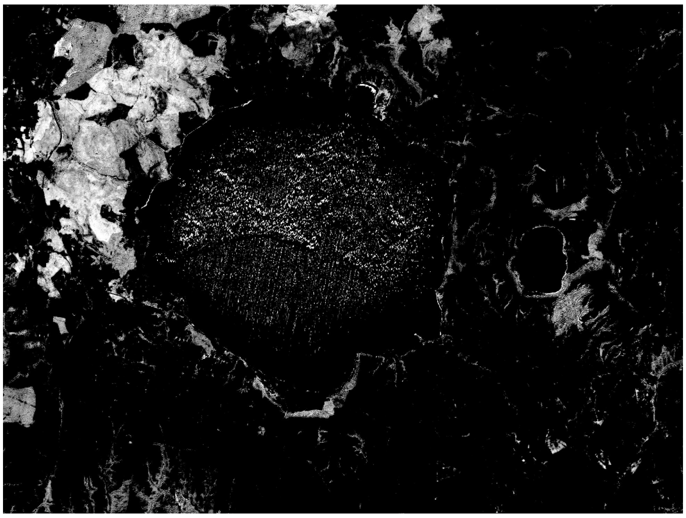

Figure 11: NDWI summer

In SNAP is also present the **NDWI2 operator** that can be used to compute the NDWI2 index responsible to put in evidence the water content in the lake, that will be useful to compute the water mask. The operator can be found by following this path in SNAP: OPTICAL/THEMATICC LAND PROCESSING/WATER RADIOMETRIC INDECES. The NDWI2 index algorithm, developed by McFeeters in 1996, takes in input the resampled data, and for the Green source band it uses the B3 band and B8 band for NIR source band.
As shown by figures 12 and 13, this index put in evidence, with high values, the water content in both lakes, Bracciano and Martignano.

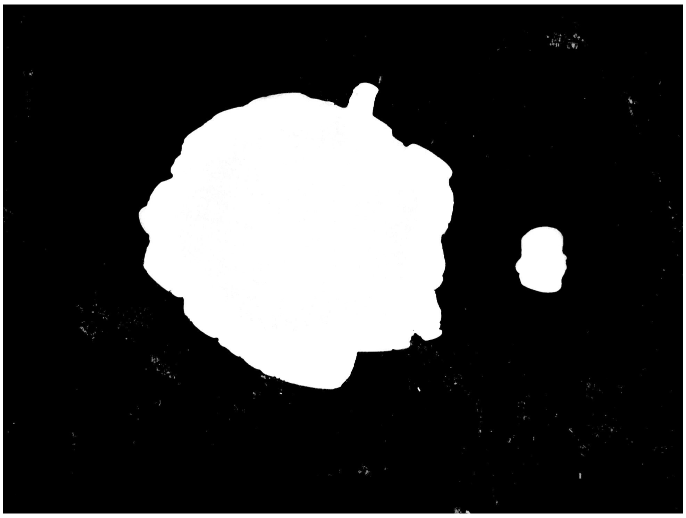

Figure 12: NDWI2 winter

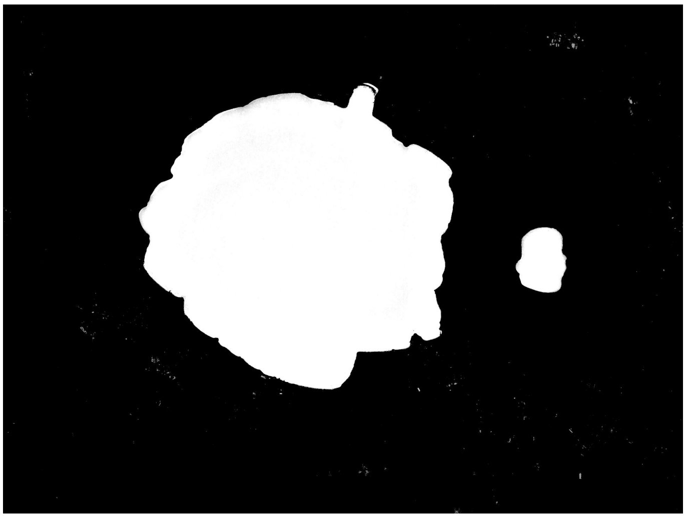

Figure 13: NDWI2 summer

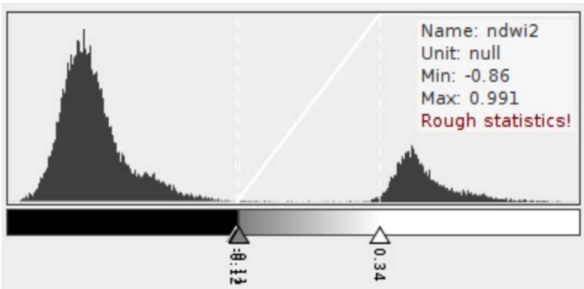

Figure 14: Color manipulation NDWI2 winter

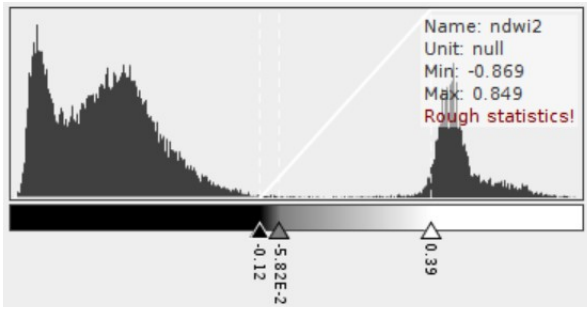

Figure 15: Color manipulation NDWI2 summer

As said before, the NDWI2 index can be used to build up a binary water mask. In order to do that, it is necessary to use the **Band Maths** operator implemented in SNAP, that allows to create a band math expression. Below the binary water mask expression is reported: *if NDWI2 > 0.3 then 1 else 0*. The threshold 0.3 is used because it allows to highlight the lakes. As Figures 16, 17, 18 and 19 show, once the cursor (red dot) is outside the water the associated value is 0 whereas when the cursor is in the lake the associated value is 1.

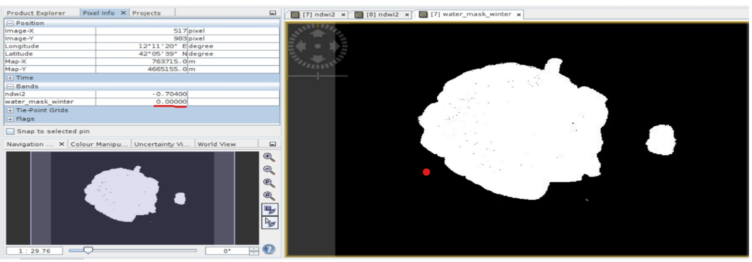

Figure 16: NDWI2 water mask 0 winter

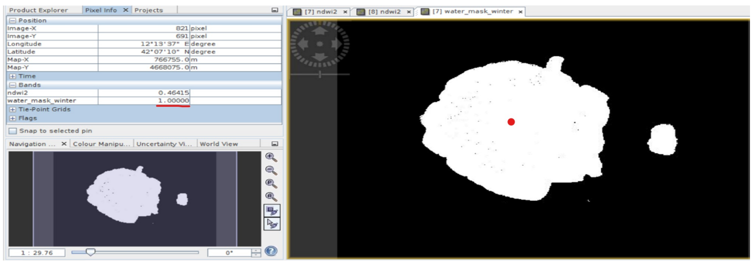

Figure 17: NDWI2 water mask 1 winter

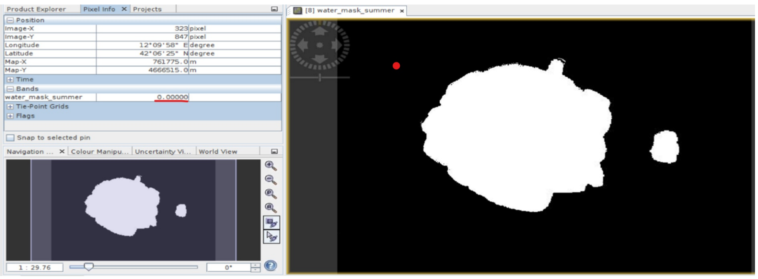

Figure 18: NDWI2 water mask 0 summer

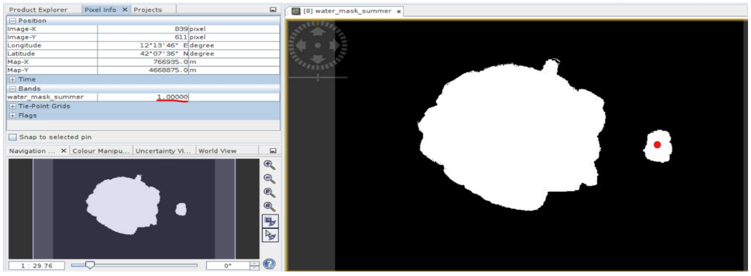

Figure 19: NDWI2 water mask 1 summer

Once the water masks have been computed on the products, the estimation of the chl_a and TSS has been performed. As it'll be explained in the 3^{rd} paragraph, the chl_a is a measure of chlorophyll concentration whereas the TSS is made up of inorganic materials, bacteria, algae and any other particles larger than 2 microns. These two measures have been computed on both summer and winter products around Bracciano and Martignano lakes. Figures below show respectivelly ch_a and TSS:

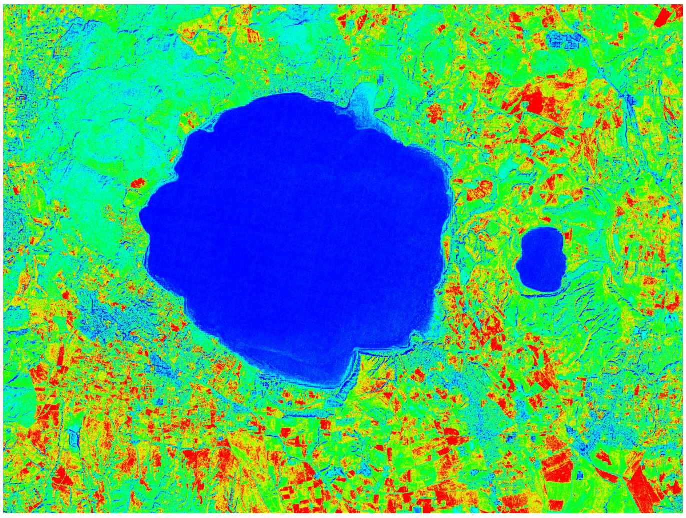

Figure 20: chl_a Winter

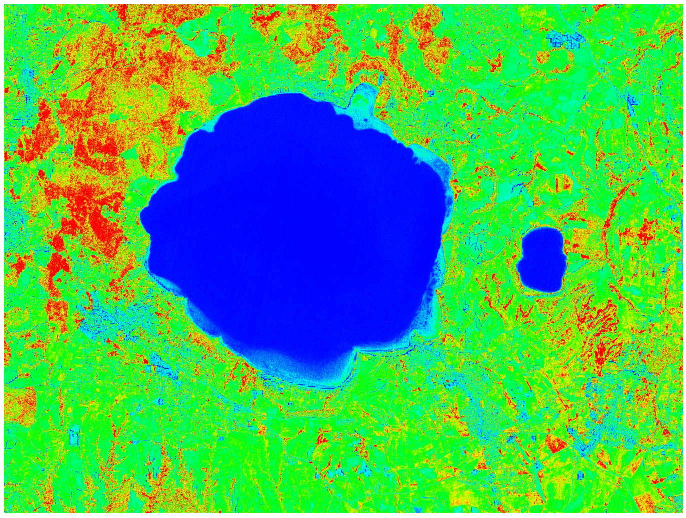

Figure 21: chl_a Summer

For chl_a images the **wew_water_chl** color pallete has been used. Red areas in the two figures above represent pixels with a high chl_a index value (> 5). By comparing these two figures with the RGB images (Figures 6 and 7) it is possible to see how the index reflects the change in the vegetation around the two lakes. In particular, it is possible to see how the index highlights wooded areas at north-west of the Bracciano lake and nearby Martignano lake.

On the other hand, the two figures below show the estimated TSS index on the two products. Here, in order to better highlight differences in the index, a more suitable color palette has been used, the **wew_water_tsm**.Low values of TSS are expressed in blue, whereas high and very high values are expressed in red. As it can be seen from the Figures, high values of TSS are areas that have a low chl_a value, so areas with a higher concentration of sediments.

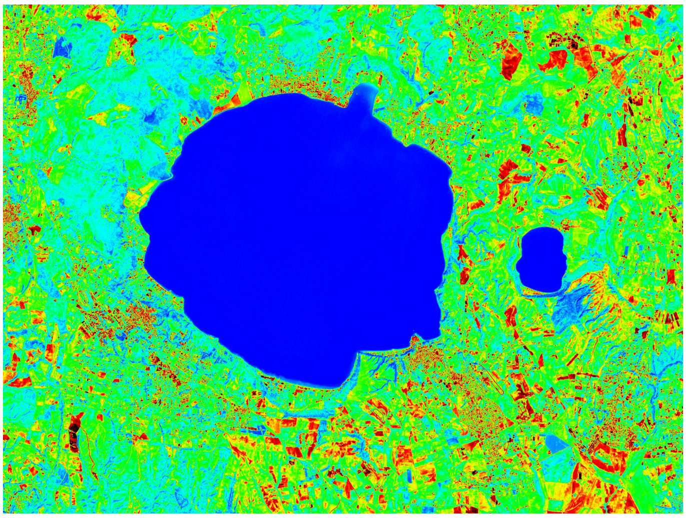

Figure 22: TSS Winter

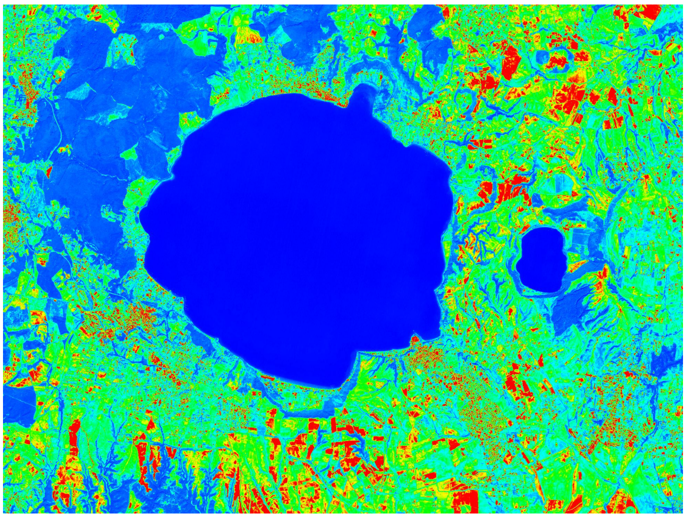

Figure 23: TSS Summer

2.2 Compute the NDVI on winter and summer resampled MSI products showing the difference between winter and summer NDVI products in terms of histograms.

The NDVI (normalized difference vegetation index) index is one of the most famous index in remote sensing. It provides an indication of the amount and state of the vegetation. The index spans from values between -1 and 1; lower values identify unhealthy vegetation, whereas values close to 1 represent healthy vegetation (green). The index has been computed on both winter and summer acquisitions in order to show how the healthy state vegetation, around Bracciano and Martignano lakes, changes during the seasons. The NDVI index can be computed by calling the **NDVI operator** implemented in SNAP that can be found by following this path: OPTICAL/THEMATICC LAND PROCESSING/VEGETATION RADIOMETRIC INDECES. It takes in input the resampled product and for the Red source band it uses the B4 band whereas, for the NIR source band it uses the B8 band. Figures 24 and 25 show respectively NDVI for both summer and winter. The used color palette is the **meris_veg_index**, where the green represents very high NDVI values and yellowish very low NDVI values.

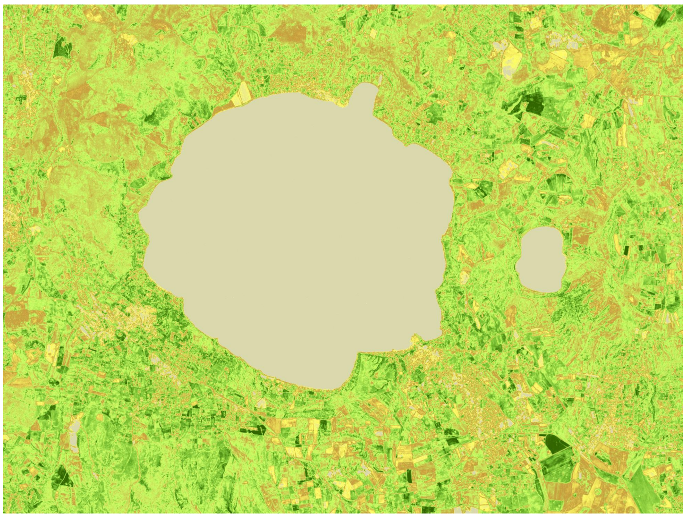

Figure 24: NDVI winter

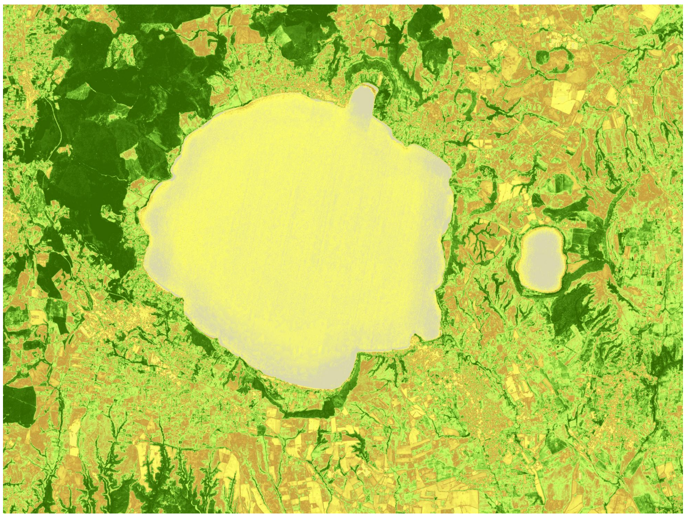

Figure 25: NDVI summer

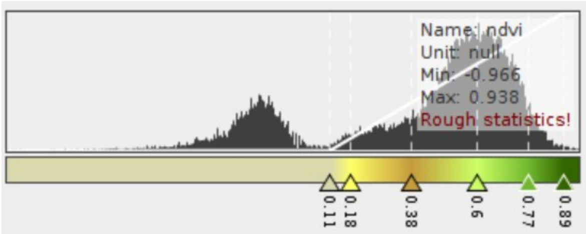

Figure 26: Color manipulation NDVI winter

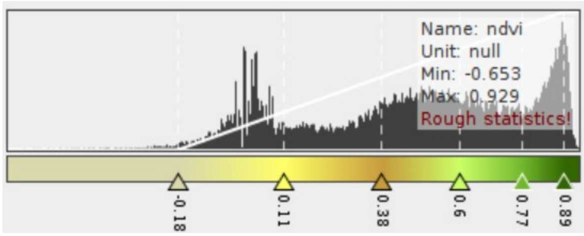

Figure 27: Color manipulation NDVI summer

As it can be seen from Figures 24 and 25, there is a significant difference of the distribution of the NDVI values on the two products. In particular, Figure 26 shows that NDVI values are mostly concentrated among 0.38 and 0.77 but there are also significant values that are below zero, where some of them are close to -1. Instead, Figure 27 shows that values are almost positive and the highest picks are visible for NDVI values of around 0.89.

3 SEA CHLOROPHYLL-A AND SEDIMENT ESTIMATION

For this part, what will be computed is the maximum chlorophyll index (MCI) that is used as indicator to identify the presence of chlorophyll in the water. i.e it is an indicator of presence of red tides, algae blooms or most in general aquatic vegetation. The **MCI algorithm** exploits the height of the measurement in a certain spectral band above a baseline. Since Sentinel 2 products have been used, the bands B4 (665 nm), B5 (705 nm) and B6 (740 nm) are the ones used by the algorithm. The formula to compute the MCI is the one reported below:

$$MCI = L_2 - k(L_1 + (L_3 - L_1)\frac{\lambda_2 - \lambda_1}{\lambda_3 - \lambda_1})$$

Where the indexes 1 and 3 indicate the baseline bands (B4 and B6) and index 2 the peak wavelength (B5) and k is set to 1.005 in order to reduce the effect of thin clouds. What is also important to highlight is that there are 3 main types of water leaving reflectance. The first is the **Clean water** that has the maximum reflectance in the blue part ($\simeq$ from 400 nm to 480 nm) and the lowest reflectance is in yellow and red ($\simeq$ from 600 nm to 700 nm). The second type is the **Higher Phytoplankton** which has the maximum reflectance in the green color ($\simeq$ from 570 nm to 590 nm). The last type is the **Opaque Water**, which is a combination of phytoplankton and dead organic and inorganic matter, it has more reflectance in green and red.

The sea chlorophyll-a and sediment estimation analysis will be performed on the Delta del Po. Two Data acquisitions have been used, one for the winter (February 23^{rd}, 2019) and one for the summer (July 23^{rd}, 2019). Both acquisitions belong to the same area, and this can be verified by looking at the tile file of the products and check that are the same (T32TQQ); moreover, they have been acquired by the same satellite (Sentinel 2A) and are both Level 2 data (i.e BOA data, where the atmospheric effect has been already removed). Below are reported the two acquisitions:

- **Winter:** S2A_MSIL2A_20190223T101021_N0211_R022_T32TQQ_20190223T134548.SAFE
- **Summer:** S2A_MSIL2A_20190723T101031_N0213_R022_T32TQQ_20190723T125722.SAFE

3.1 Select a subset ROI within the summer and winter MSI image subsetting a coastline near a river estuary/Delta.

Before starting any kind of analysis is necessary to resample and subset the product around the cost line and the Delta. This is necessary due to the fact that the aim is to see chlorophyll and sediments in water but also to speed up the computation by working on a smaller area. For the resampling it has been used, as for the Bracciano lake, the 10 m resolution and for the ROI (region of interest) the Delta del Po. The resample has been performed by calling the **Resampling operator** and for the subsetting the **Subset**

operator, both are implemented in SNAP.
Figures 28 and 29 show the resampled and subsetted products around Delta del Po in winter and summer.

Figure 28: Delta del Po winter

Figure 29: Delta del Po summer

3.2 Implement at the EmpReg regressive algorithms (see below) to estimate chlorophyll-a (Chl-a) and total suspended sediments (TSS) using SNAP formula processing tool

The empirical Regression Algorithm (EmpReg) is a blue-to-green reflectance maximum band ratio (MBR) model. The formula to compute the reflectance MBR is reported below:

$$r_{MBR} = \frac{R_{wlB1}, R_{wlB2}}{R_{wlB3}}$$

B1 (centered at 442 nm) and B2 (497 nm) are the MSI (Multi Spectral Imager) blue spectral bands whereas, B3 (centered at 560 nm) is the green spectral band. The numerator of r_{MBR} returns the maximum value among the water leaving reflectance at B1 and B2. High values of MBR correspond to a low Chl-a concentration and the other way around. Blue and Green spectral bands have been preferred to the red-edge wavelengths that are more effective in water where there is a higher chl-a but are also more sensible to reflectance effects due to the coast proximity. For that reason, blue and green spectral bands should provide more accurate results.
The expression for the optimal empirical regressive (EmpReg) retrieval algorithm is reported below:

$$\hat{C}_{Chla} = a_1 * exp(-a_2 * r_{MBR})$$

where $a_1 = 59.795 mg/m^3$ and $a_2 = 4.559$.
[Source: Frank S. Marzano et al., IEEE TGRS, <<Coastal water remote sensing from Sentinel-2 satellite data using physical-statistical and neural-network retrieval approach>>, 2020].

Instead, regarding the TSS (total suspend sediments) computation, another exponential model using the water-leaving reflectance at spectral band B4 (665 nm) has been used. The formula to compute the TSS is the one reported below:

$$\hat{C}_{TSM} = b_1 * exp(-a_2 * R_{wlB4})$$

where $b_1 = 5.499 g/m^3$ and $b_2 = 9.04$.
[Source: Frank S. Marzano et al., IEEE TGRS, <<Coastal water remote sensing from Sentinel-2 satellite data using physical-statistical and neural-network retrieval approach>>, 2020].

Both The Chl-a and TSS indexes have been computed in SNAP by using the **band math operator**.

3.3 Compare estimated chlorophyll-a (Chl-a) and total suspended sediments (TSS) for a winter and summer cases by making the image differences.

Once both Chl-a and TSS have been computed in SNAP, the images have been exported and the results are displayed below. Figures 30 and 31 show the estimated Chl-a for summer and winter acquisitions:

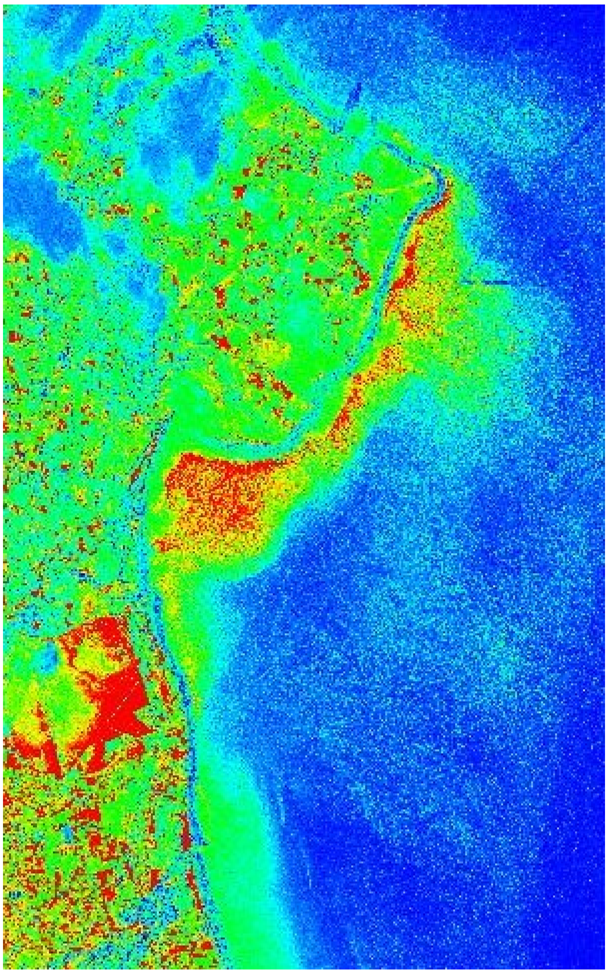

Figure 30: Chl-a winter

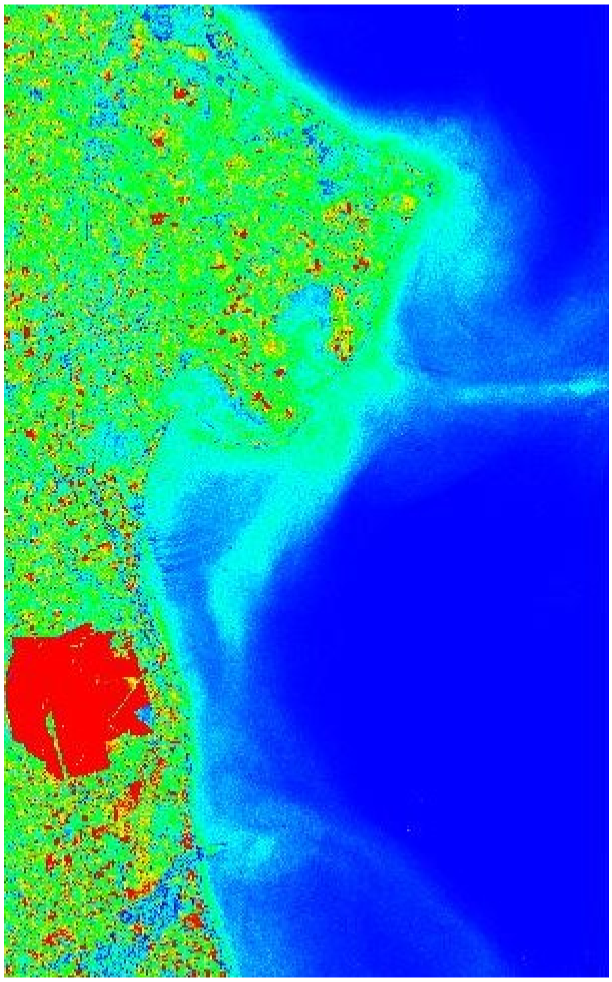

Figure 31: Chl-a summer

In order to better highlight the estimated chl-a index, the **wew_water_chl** color palette has been used. Low values of chl-a are represented in blue whereas high and very high values are expressed in yellow and red. As it can be seen from the figures above, most of the high values are around the coast line and river Delta for winter acquisition. Whereas in summer, the aquatic vegetation tends to be more present in the hinterland, because of the presence of paddy fields which are widely cultivated thanks to fertility of the ground; in fact, the paddy field cultivation represents an important business for the north of Italy. For that reason, this kind of index could play a key role in decision making in fact, it can be used to understand which is the best period for harvesting, in this case rice harvesting.

Regarding the estimated TSS index instead, Figures 32 and 33 show the results. Also here, in order to better highlight differences in the values index, a more suitable color palette has been used, the **wew_water_tsm**. Low values of TSS are expressed in blue, whereas high and very high values are expressed in red and black. As it can be seen from the Figures, areas with a high TSS index are the ones that present a very low value of chl-a, showing the lack of water vegetation and algae blooms and so a higher concentration of sediments.

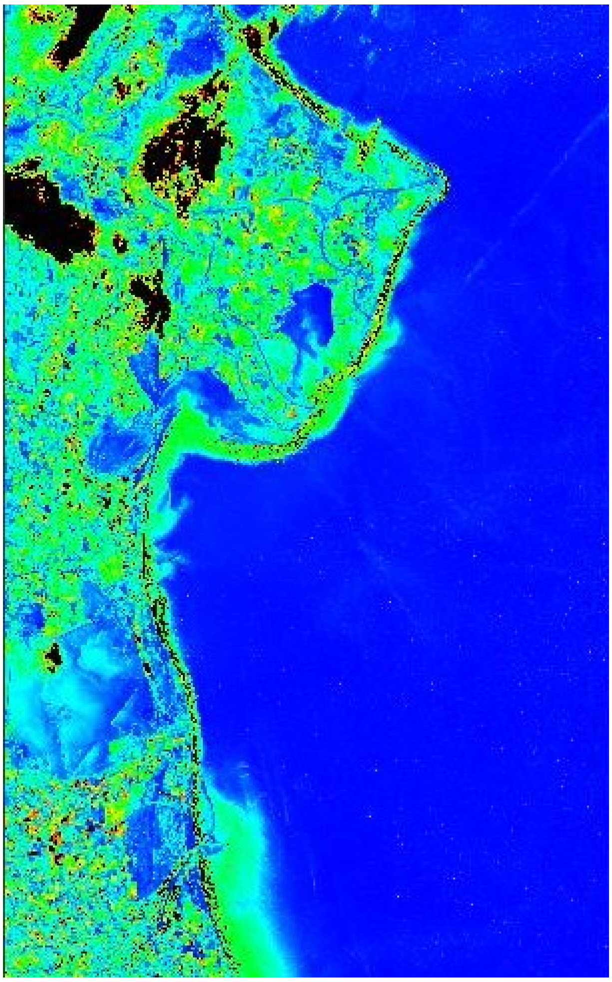

Figure 32: TSS winter

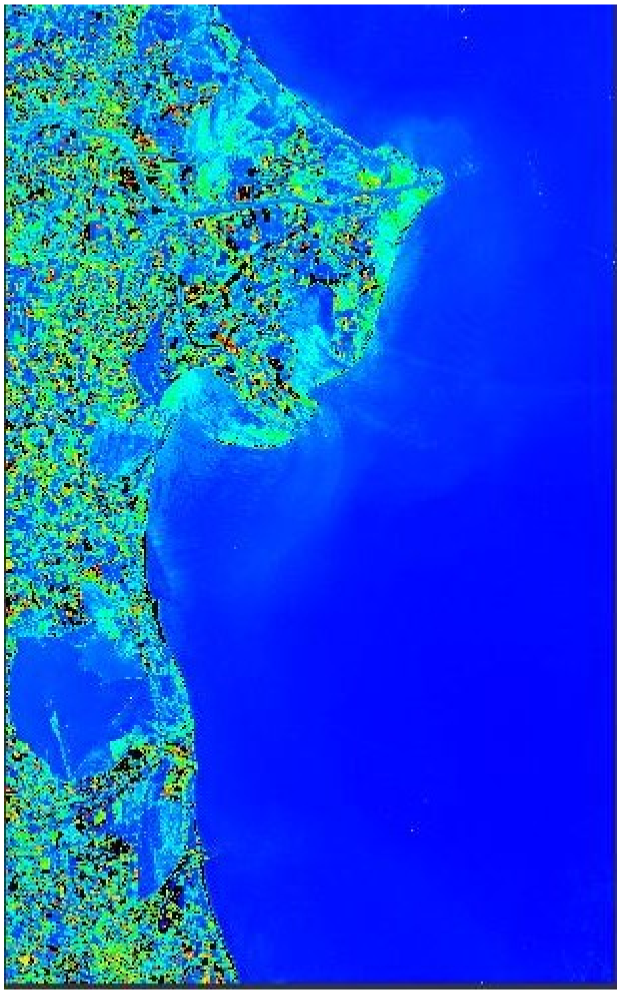

Figure 33: TSS summer

3.4 Compare estimated chlorophyll-a (Chl-a) with the one from SNAP MCI (Maximum Chlorophyll Index) plugged-in algorithms.

In order to compute the MCI (Maximum Chlorophyll Index) for both winter and summer products, the **S2 MCI Processor**, implemented in SNAP and available under the path: OPTICAL/THEMATIC WATER PROCESSING/S2 MCI PROCESSOR) has been used. The processor takes in input the resampled and subset product. Instead, the processing parameters have been set as described below:

- **Preset:** S2MSI L2 MCI (because the products are level 2 data)

- **Lower baseline band:** B6. It is the lower band defining the baseline.
- **Upper baseline band:** B4. It is the upper band defining the baseline.
- **Signal band name:** B5. It is the band for which the baseline height is calculated.
- **Mask expression:** scl_water. It is the ROI-mask expression that has to be evaluated during processing. It is used to define pixels of interest. This means that the computation is computed only on the water and land is masked out.

Once the process is completed, the MCI is calculated and the Figures below show the results for winter and summer acquisitions.

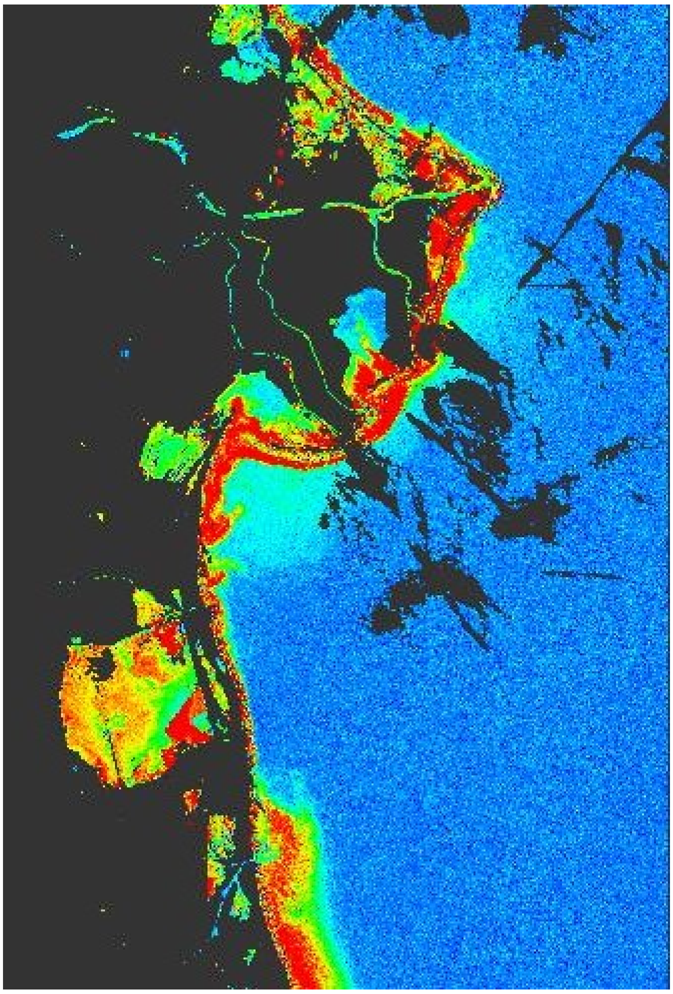

Figure 34: MCI winter

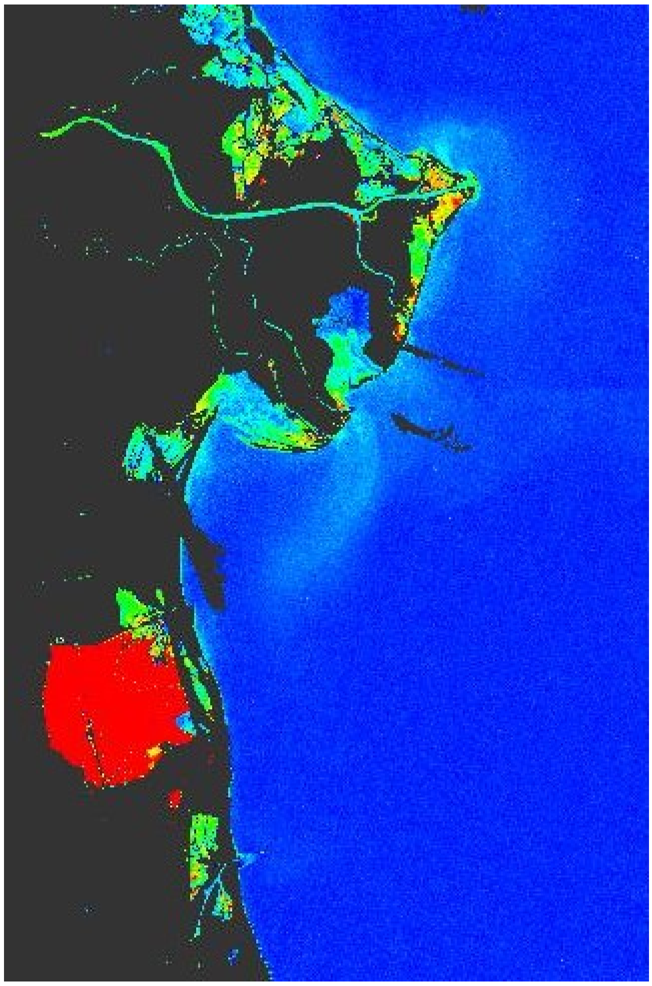

Figure 35: MCI summer

As it has been done for the estimated chl-a in section 3.3, the **wew_water_chl** color palette has been used. However here the images show some black areas, this due to the fact that a water mask has been used in order to mask out the land. As it can be noticed by comparing images of estimated chl-a and the MCI, the areas representing the high concentration of water vegetation are almost the same. However here, due to the fact that the land is masked out, it is easier to distinguish paddy fields and river with their concentration of chlorophyll.

3.4.1 Compare estimated chlorophyll-a (Chl-a) with the one from SNAP C2RCC neural-network and MCI (Maximum Chlorophyll Index) plugged-in algorithms. [EXTRA POINT]

The **C2RCC neural-network algorithm** implemented in SNAP works with rower data, more precisely level 1 data; i.e TOA data where the atmospheric effect has not been removed. For that reason, two new products have been downloaded from CREODIAS repository of Earth Observation satellites, covering the same areas in the same period but taking in consideration the atmospheric effect. The L1 products are the ones reported below:

- **Winter:** S2A_MSIL1C_20190223T101021_N0207_R022_T32TQQ_20190223T122359.SAFE
- **Summer:** S2A_MSIL1C_20190723T101031_N0208_R022_T32TQQ_20190723T121220.SAFE

Once the products have been downloaded, it has been possible to run the **C2RCC algorithm** to compute the MCI. The algorithm can be found under the fallowing path in SNAP: OPTICAL/THEMATIC WATER PROCESSING/C2RCC PROCESSORS/S2-MCI. The Algorithm requires the setting of different parameters like *Valid-pixel expression*, which defines the arithmetic expression to identify valid pixels for which compute the MCI (the not valid pixels instead will be masked out and will be represented with a black color), *Salinity* ,which is the value used as water salinity for the scene, *Temperature*, which is the value used as water temperature for the scene, *Ozone* (the ozone value), *Air Pressure at Sea level*, *Elevation*, which is used as elevation for the scene. For this particular purpose almost all predefined parameters have been left as default one. Figure 36 shows the set parameters.

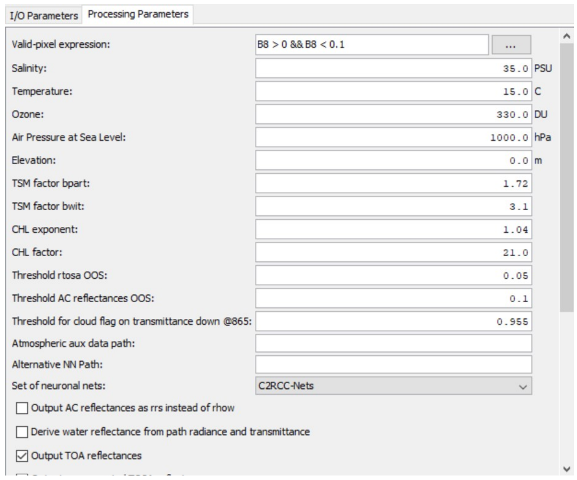

Figure 36: C2RCCS2 MCI PROCESSOR

Figures 37 and 38 show the MCI for winter and summer acquisition after have run the C2RCC algorithm. Also here, the **wew_water_chl** color palette has been used, and the masking operation has been performed in order to exclude land areas; in this way it is better to distinguish paddy fields and river with their concentration of chlorophyll. As it can be noticed by comparing images of estimated chl-a and the MCI, the areas representing the high concentration of water vegetation are almost the same but here it is possible to see how the atmospheric effect have a strong impact on the MCI. In fact, red areas (pixels with a very high MCI) are reduced and only some of them are maintained. Moreover, the presence of atmospheric effect, have a strong impact also in open sea by returning all this area with quite high MCI values.

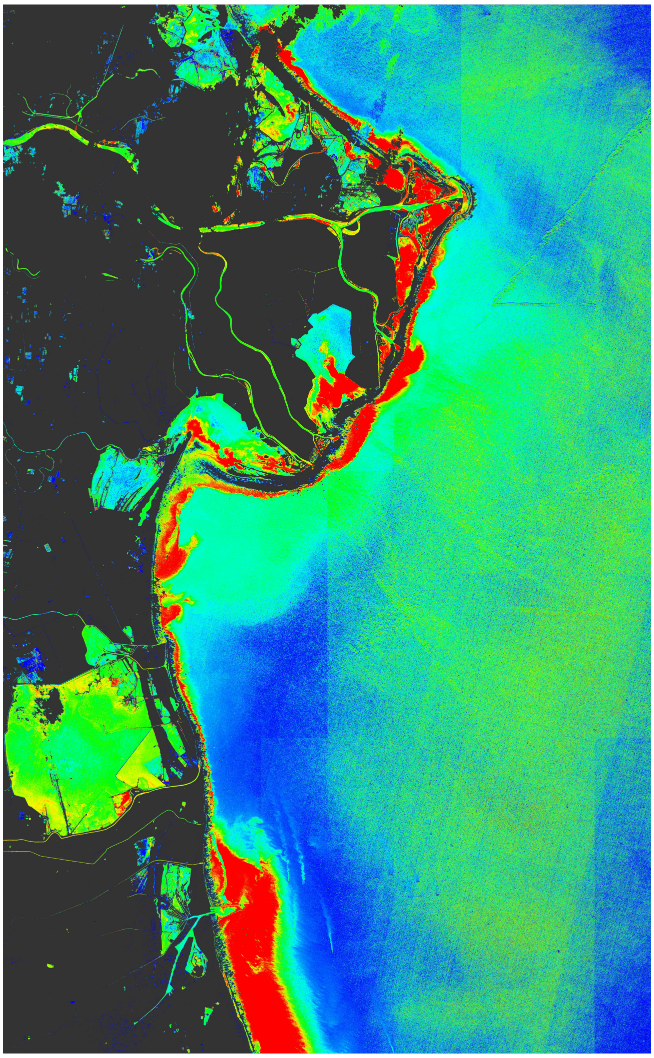

Figure 37: MCI winter

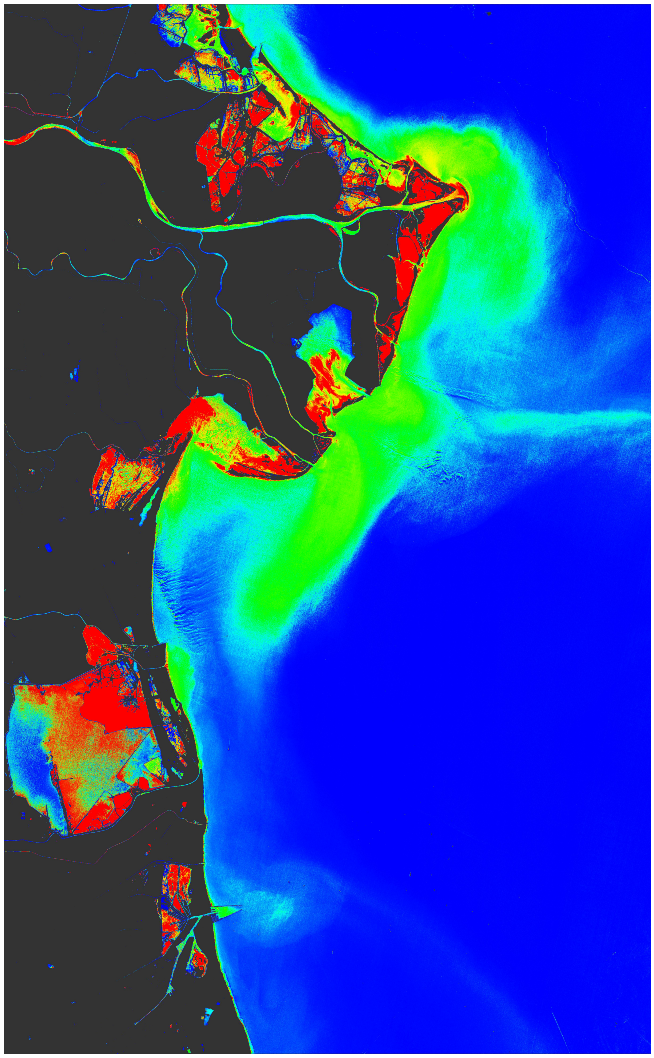

Figure 38: MCI summer

3.5 Download Chl-a and TSS products around the target area in the same summer days but in year 2018, 2019 or 2020 from international Copernicus Marine Service at http://marine.copernicus.eu

By going on Copernicus Marine Service, EUROPEAN SEA SURFACE CHLOROPHYLL CONCENTRATION FROM MULTI SATELLITE OBSERVATIONS (click on the hyperlink to be directed on the Copernicus folder) product has been downloaded and subset around the Delta del Po. Figure 39 shows the chlorophyll analysis and it is possible to see how the areas around the Delta del Po are the ones with a very high concentration of aquatic vegetation; as discovered before after have performed the previous analyses.
Instead, due to the fact that Copernicus does not provide information about TSS, what it has been found is an analysis about **mole concentration of dissolved inorganic matter in sea water** (which are a component of TSS). The product in question is the MEDITERRANEAN SEA BIOGEOCHEMISTRY ANALYSIS AND FORECAST and the analysis has been performed in August 2020 (click on the hyperlink to be directed on the Copernicus folder). Figure 41 shows the subset selected product.

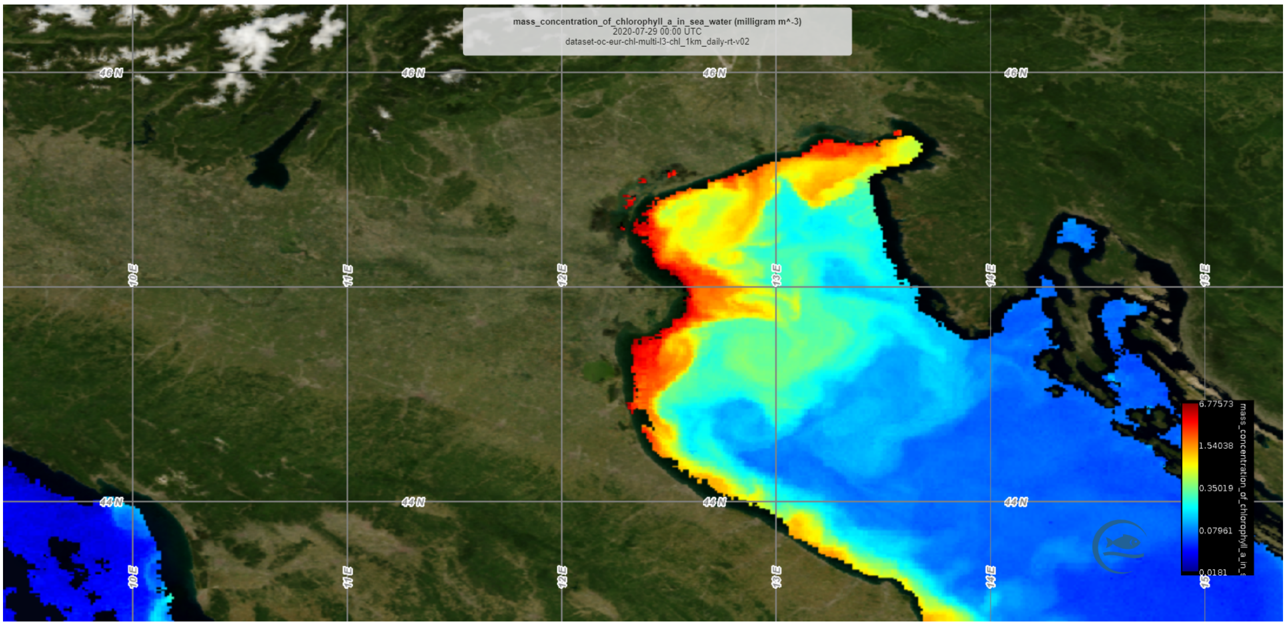

Figure 39: Delta del Po Chlorophyll concentration (Copernicus Analysis)

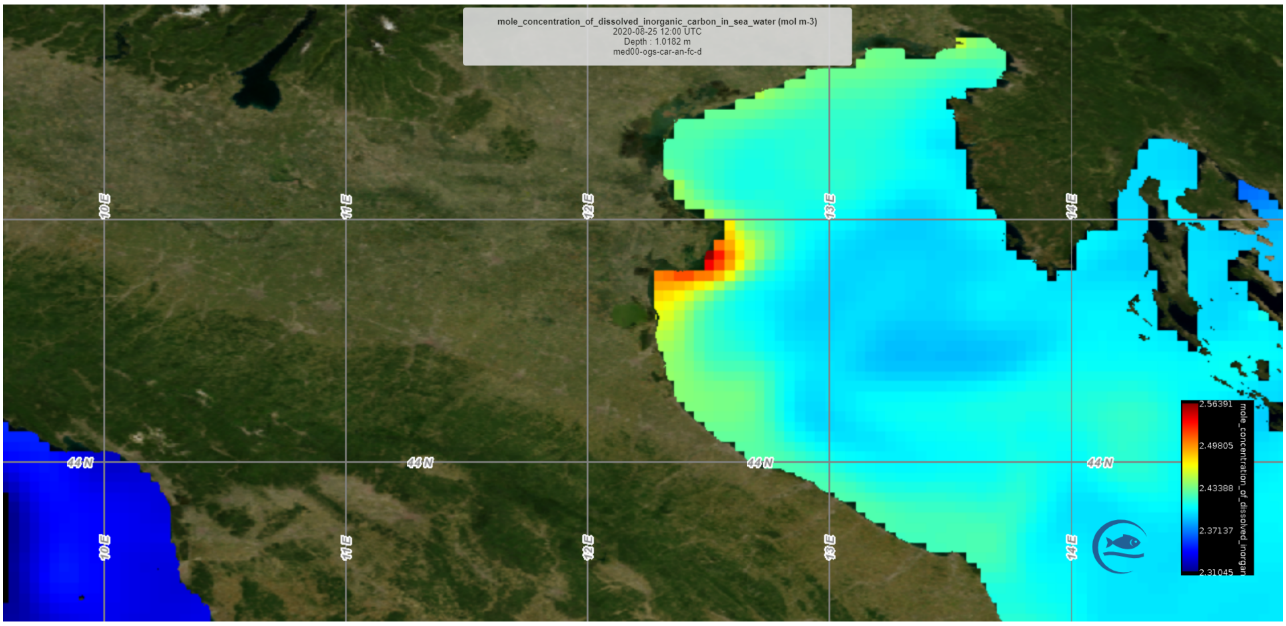

Figure 40: Mole concentration of dissolved inorganic matter in sea water (Copernicus Analysis)

3.6 Discuss the official Copernicus products within the selected areas in terms of spatial resolution and data sources.

The EUROPEAN SEA SURFACE CHLOROPHYLL CONCENTRATION FROM MULTI SATELLITE OBSERVATIONS product comes from periodic analysis performed by the Global Ocean Satellite monitoring and marine ecosystem study group of the Italian National Research Council on the European sea to monitor the aquatic vegetation status during different periods. This product has a 1km spatial resolution and the analysis has been performed 2020-07-29.
The MEDITERRANEAN SEA BIOGEOCHEMISTRY ANALYSIS AND FORECAST are produced by means of the MedBFM model system. MedBFM model is executed by OGS (National Institute of Oceanography and Applied Geophysics) and uses as physical forcing the outputs of the NEMO-OceanVar

model system. The analysis is weekly performed, in fact the figure 40 has been updated the 25^{th} of August. The product has a spatial resolution of 0.042 degree.

4 SCENE CLASSIFICATION

This last part is focused on Scene Classification. Here, what it has been done was to define a certain number of classes and run different supervised classification algorithms to perform classifications.

4.1 Using L2 MSI summer image, select scene classification masks, open a subset image and define 4 different polygons over vegetation, (ii) water, (iii) urban and (iv) bare soil. These vectors will be used as training areas for automatic classification.

In order to run supervised learning algorithms it is necessary to have training datasets that can be used by the algorithms to learn how to classify pixels. The first step is to create a number of containers equal to the number of classes that need to be automatically classified. These containers have been created by using the **New Vector Data Container** function. Once the containers have been created, they have been fulfilled with data by drawing rectangles for each classes. Figure 41 shows the shapes used to fulfill the vector containers.

Figure 41: Image Classification Classes

4.2 Reproject the subset image Lat/Lon WGS84 geographiy coordinate system. Use the SNAP Supervised Classification algorithms, such as Maximum Likelihood or Random Forest, to compute a classification.

Before running different classifiers it is required to reproject the image. In order to do that, the **Reprojection operator**, that can be found following this path: RASTER/GEOMETRIC OPERATIONS/, is called in order to repreject the product using latitude and longitude coordinates. The initialization of the Reprojection operator has been set by using the default values. Figure 42 displays the reprojected product.

Figure 42: Image Classification Classes reprojected

Now that the product has been reprojected, it is possible to train different classifiers using the predefined training sets (the training vectors: vegetation, water, urban and bare soil). For this purpose, 4 different classifiers have been chosen, and, for all of them, different configurations have been trained, and good results have been returned by using the following configurations:

- **Random Forest:** n° of training samples = 8000, n° of tress = 80.
- **Knn:** n° of training samples = 8000, n° of neighbours = 10.
- **Maximum Likelihood:** n° of training samples = 8000.
- **Minimum Distance:** n° of training samples = 8000.

4.3 Compare different Classifiers to assess visually the most reliable using the L2 provided mask.

Once the algorithms have been trained, they have been used to perform classifications. Below are reported Figures showing the classifications for each classifier. As it can be immediately noticed, the Maximum Likelihood Classifier doesn't perform very well, it actually does many misclassifications and it is not able to detect the whole vegetation around the lake and bare soil areas. But more important, it is not able to correctly detect the whole water areas in the region in fact, it identifies urban areas inside the Bracciano lake but what actually surprises is that it is not able to detect the Martignano Lake. Random Forest Classifier differently, seems to perform better on classify water areas,vegetation, bare soil and urban areas. The problem is that, if we look in depth it is possible to see that there are different areas that have not been classified at all; these areas are the ones in black. On the other hand, Minimum Distance and KNN Classifiers are the two classifiers that retrieve better results in classifications, and this can be check by comparing them with the RGB image. So they seem to be good candidates to be selected as best classifiers. However, if we take a closer look, it is possible to see that Minimum Distance Classifier performs better on classify vegetation around the lake and urban areas, in particular near Trevignano Romano city. But what is actually more important is that the Minimum Distance Classifier is also able to detect swimming pools in urban areas (Bracciano and Anguillara cities) as highlighted by the red circles in the figure 47. For these reasons, Minimum Distance Classifier is the best classifier.

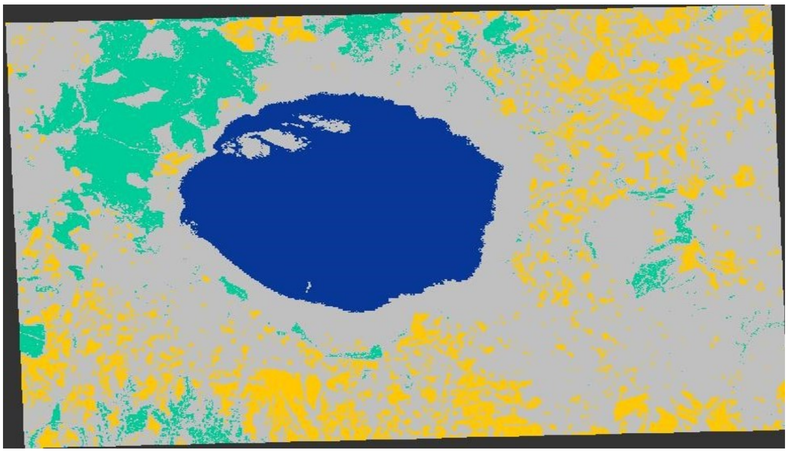

Figure 43: Maximum Likelihood Classifier

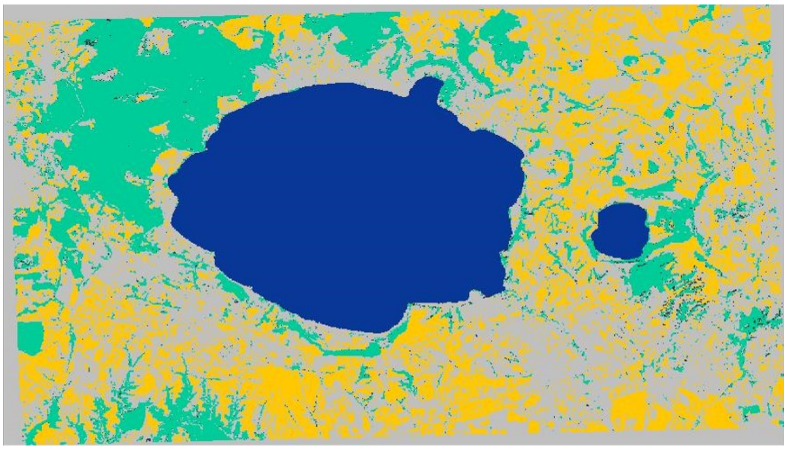

Figure 44: Random Forest Classifier

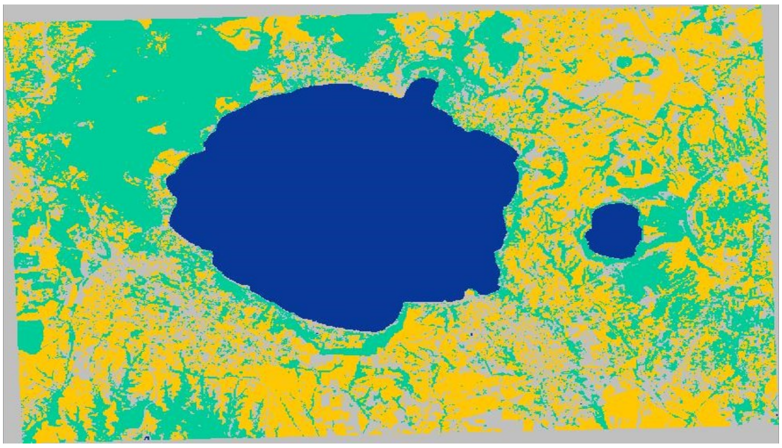

Figure 45: Minimum distance Classifier

Figure 46: KNN Classifier

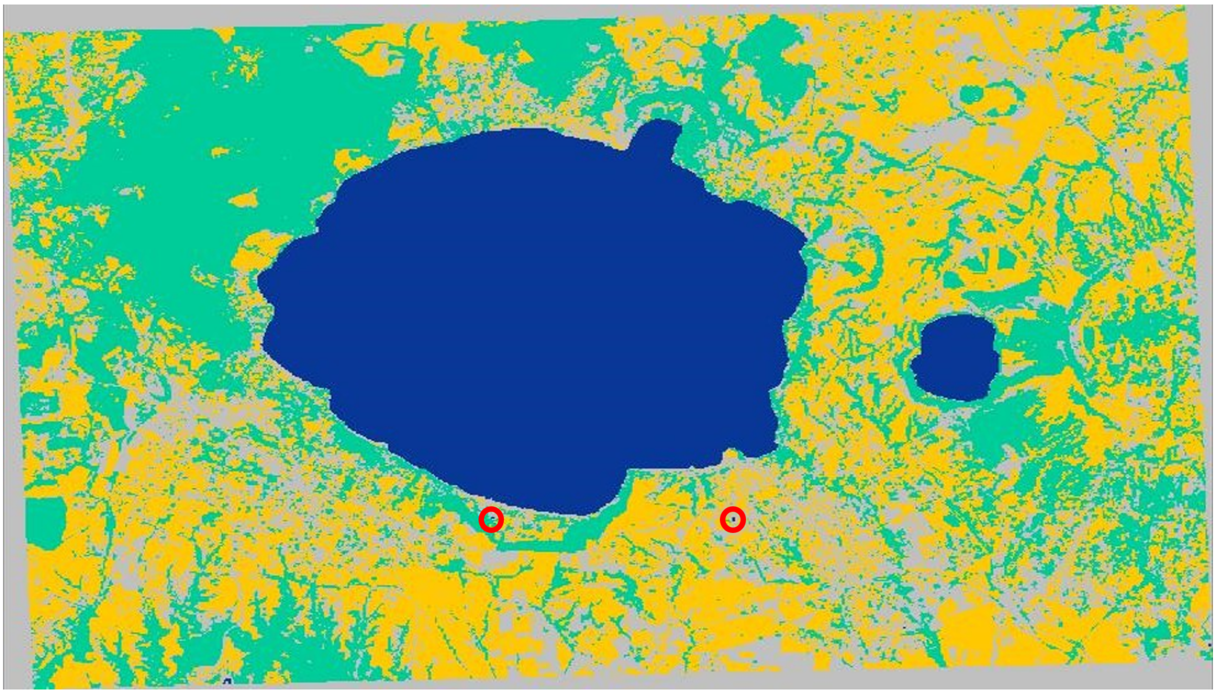

Figure 47: Detected Swimming Pools