SAPIENZA
UNIVERSITÀ DI ROMA

Earth Observation Data Analysis

HOMEWORK01 - Remote sensing of vegetation from MODIS

Faculty of Information Engineering, Computer Science and Statistics,

Sapienza University, Rome, Italy

Prof. Frank S. Marzano

REZA POURRAHIM - 1859334

UWAMAHIRWE BONAVENTURE – 1850306

HAFIZ MUHAMMAD HASSAN - 1873829

July 2020

1. Download and install SNAP

   Sentinel Application Platform (SNAP) is a common architecture for all Sentinel Toolboxes and its architecture is ideal for Earth Observation processing and analysis due to the following technological innovations: Extensibility, Portability, Modular Rich Client Platform, Generic EO Data Abstraction, Tiled Memory Management, and a Graph Processing Framework.

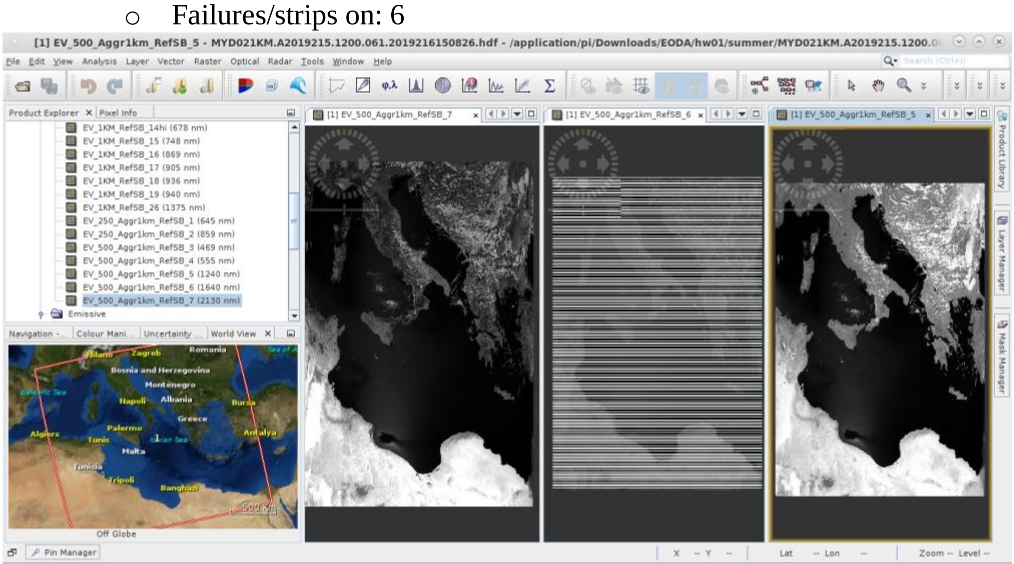
2. Download a first MODIS image at 1-km resolution over Italy or other geographical area of interest during a summer season from https://ladsweb.modaps.eosdis.nasa.gov/search (select MYD021KM*.*).

   From https://ladsweb.modaps.eosdis.nasa.gov/search, we've selected .hdf file (MYD021KM.A2019215.1200.061.2019216150826.hdf) of Summer 2019 (2019-08-03 12:00:00) over Italy. First we start by click on the product and explore the Metadata and Bands, and we see that first_line_time is 03-AUG-2019 12:00:00.000000 and last_line_time is 03-AUG-2019 12:05:00.000000.

3. Perform data quality check

   We check the channel image quality and we can see that we have the quality of images are different and we see failures or strips in some of them(we show some of them in the pictures in the below). In the Product Explorer > Bands > RefSB we can check:

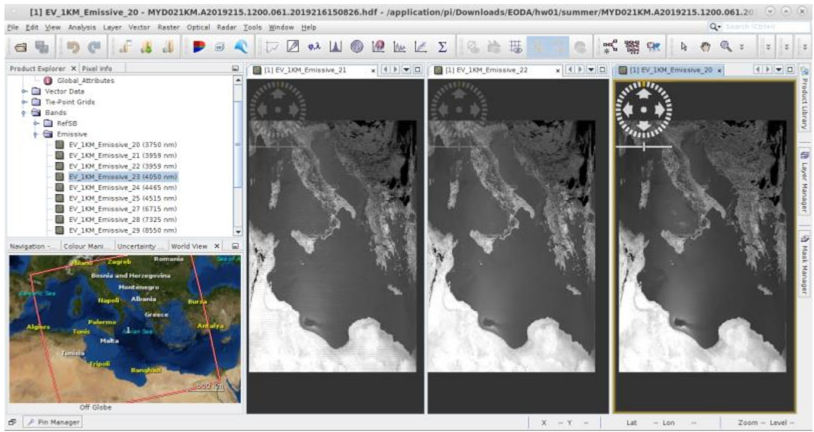
   - Reflective channels Aggr1km
     - OK for: 1-2-3-4-5-7
     - Failures/strips on: 6



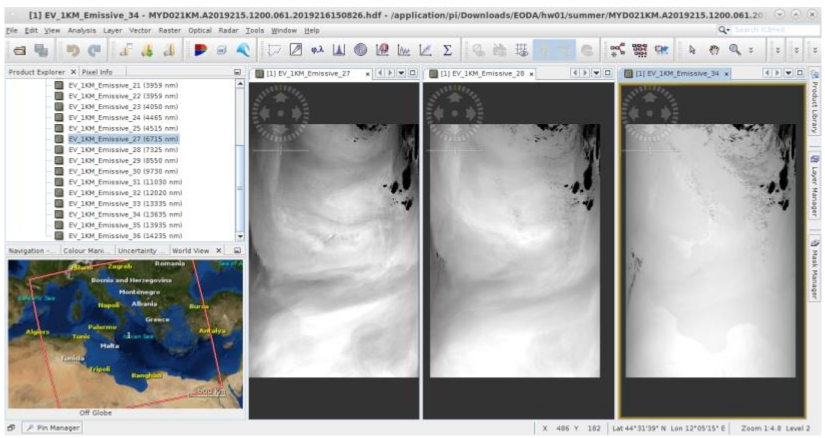
Reflective channels Aggr1km 7 – 6 - 5

   Also, in the Product Explorer > Bands > Emissive we can check:

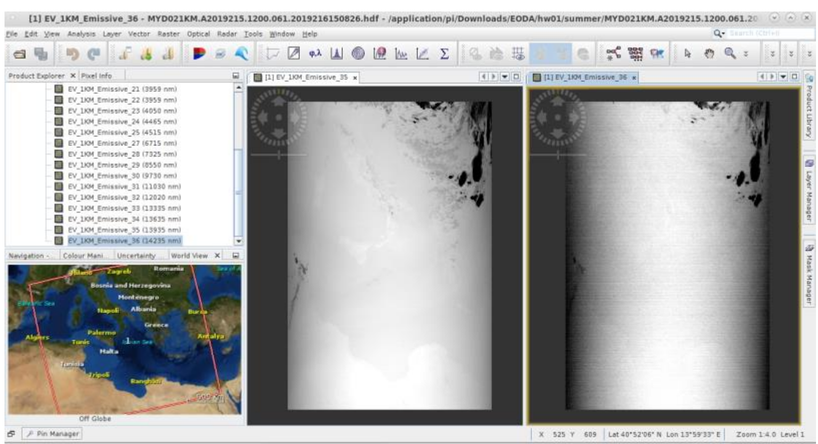
   - Emissive channels EV
     - OK for: 20-21-22-23-24-25-29-30-31-32-33
     - Failures/strips on: 27-28-34-35-36



Emissive channels EV 21 – 22 – 20



Emissive channels EV 27 – 28 - 34



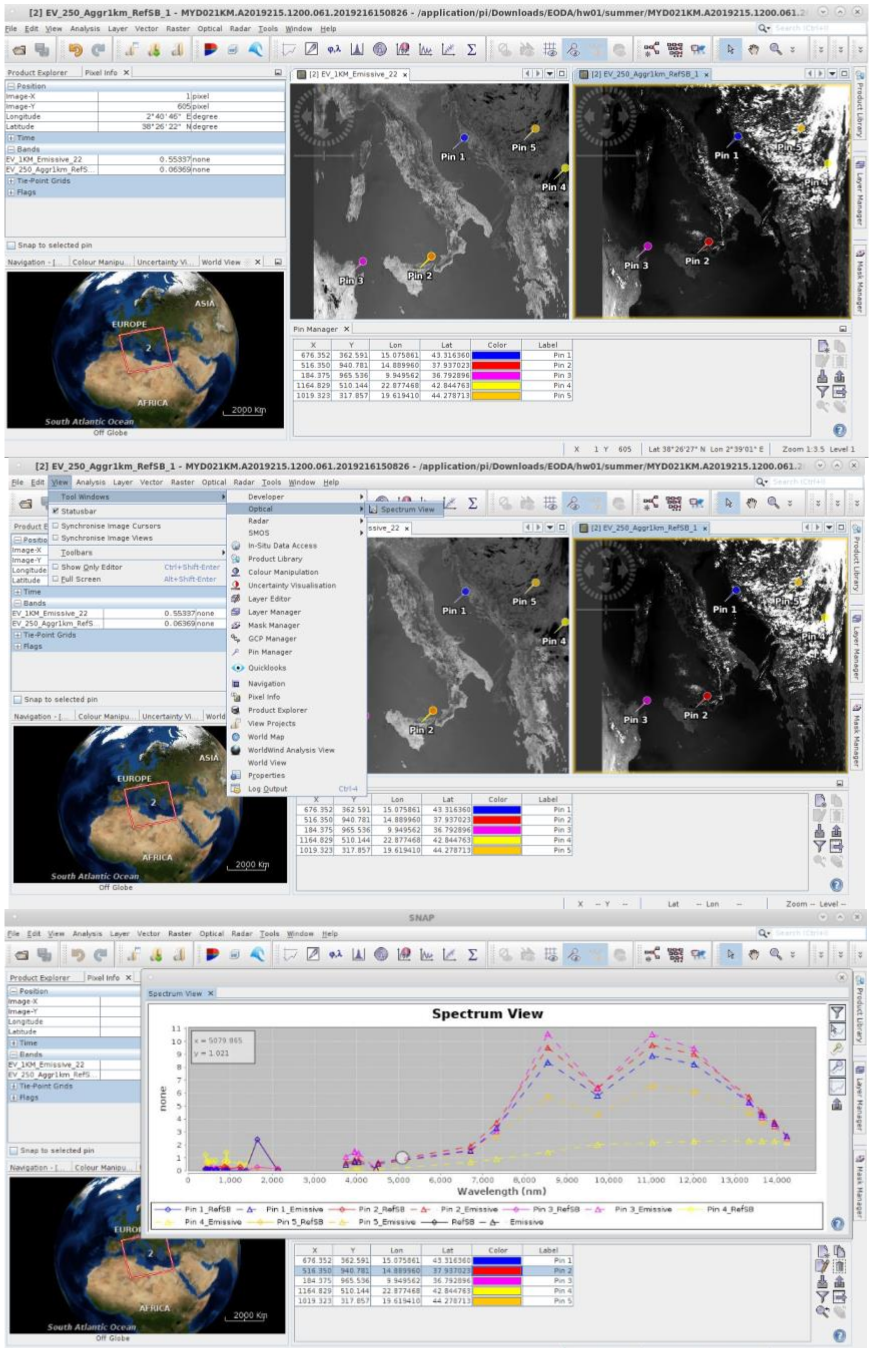
Emissive channels EV 35 - 36

4. <u>Perform and display data analysis by spectrum, histogram and profile tools</u>
   We choose 2 channel with good quality and place 5 pins on some good features to show them. We change the color of each pin to be more visible and use Tile Window to show all the pins and

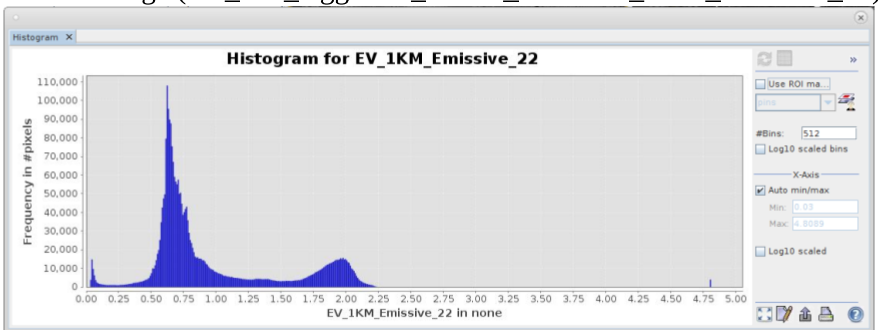
2 images in a better way, and then we plot the spectrum corresponding to these pins, and after that we plot the histogram distribution in order to perform analysis.

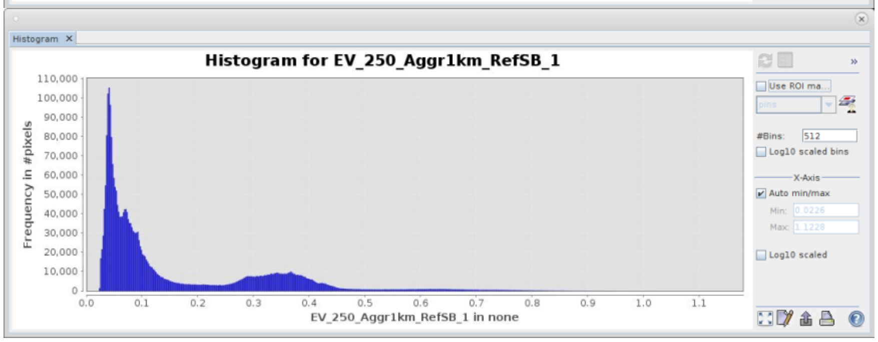
We open:

- EV_250_Aggr1km_RefSB_1 is reflectance (adimensional) at 645 nm
- EV_1KM_Emissive_22 is radiance (W/m^2 sr micron) at 3959 nm



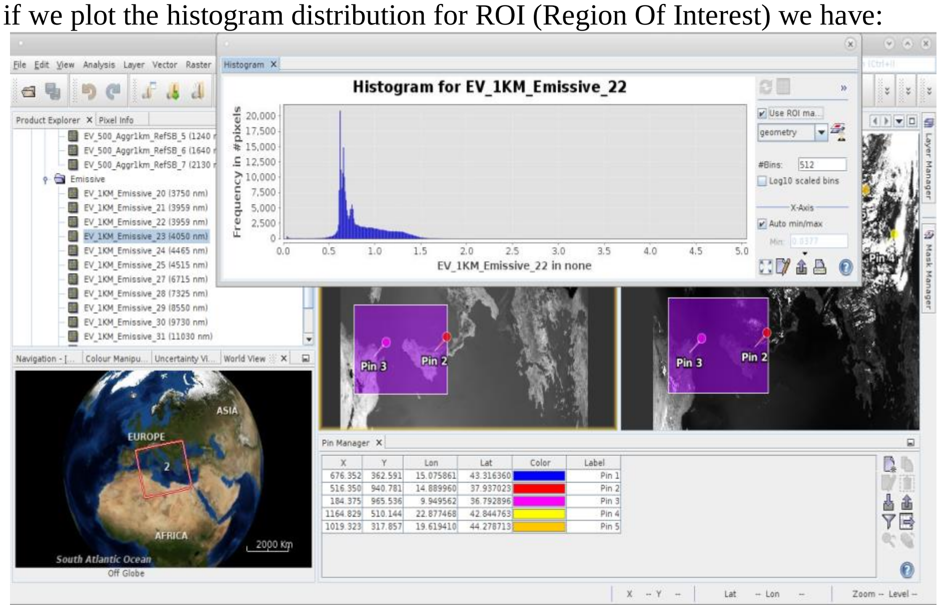
Now we have Spectrum View of our 5 interested pins for both reflective and emission bands. In this plot we can see that those pins that are located on sea or land (darker in Emissive) have lower portion compared with those pin located on clouds for example. Pin number 4 has the lowest portion in Emissive wavelength and Pin number 3 has the highest. From Menu > Analysis > Histogram we plot histogram distribution for both reflective and emissive bands image (EV_250_Aggr1km_RefSB_1 and EV_1KM_Emissive_22):

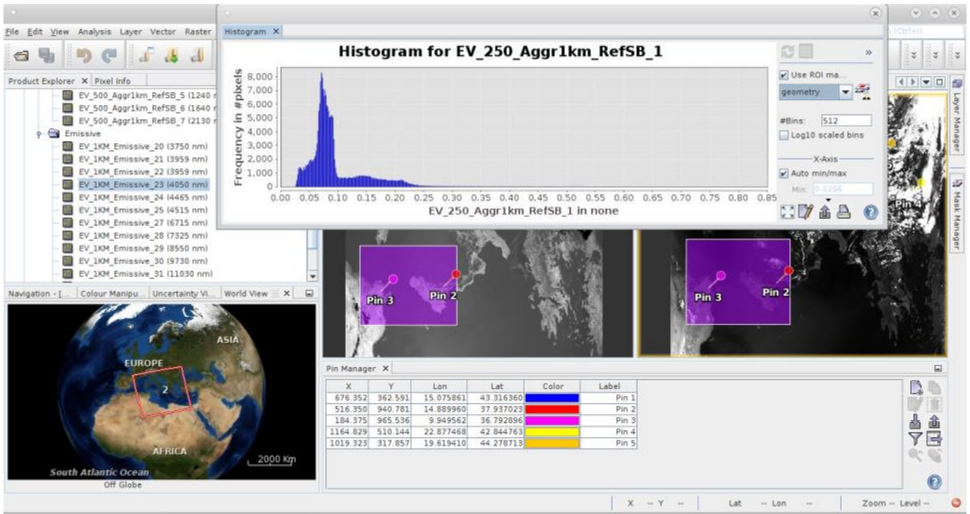




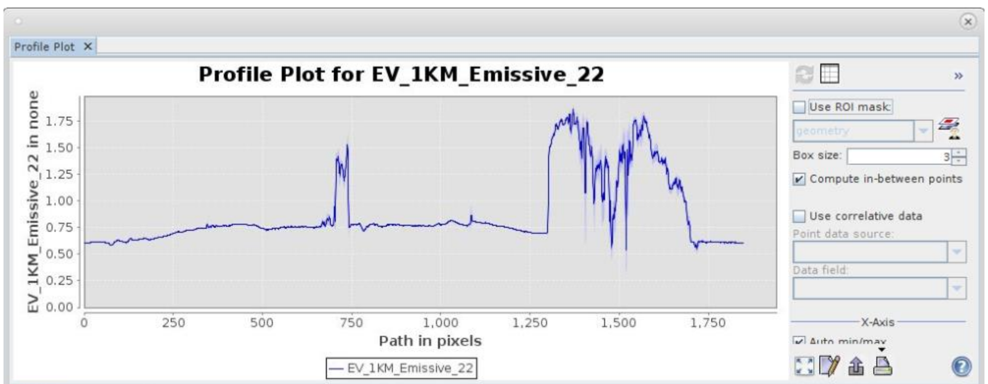
In the histogram plot we can find the frequency in #pixels of each wavelength. For Emissive band histogram distribution we can see that it is distributed between 0.00 and 2.25 and the highest frequencies are mainly between 0.50 and 0.75 and this is due the reason it is radiance. For Reflection band histogram distribution we can see that it is distributed between 0.00 and 1 and the highest frequencies are mainly close to 0 and the other points distributed in a low frequency and this is due the reason it is reflectance.

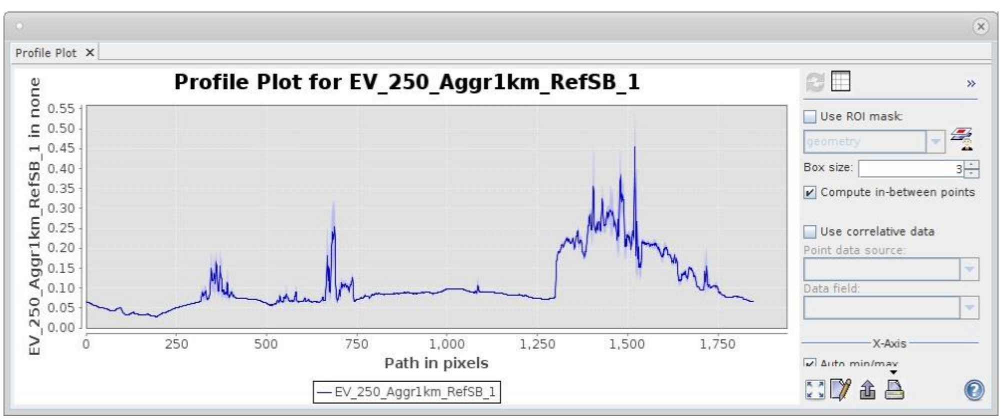
And if we plot the histogram distribution for ROI (Region Of Interest) we have:



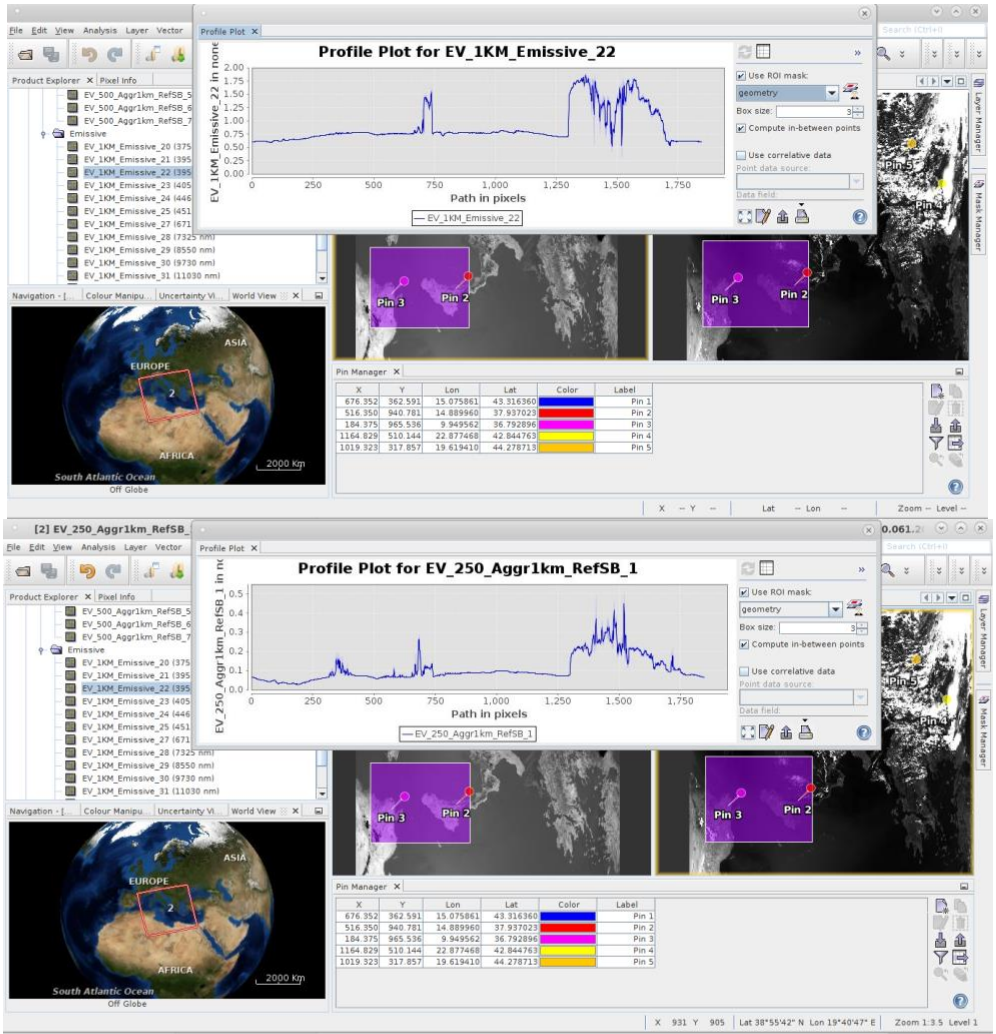


From Menu > Analysis > Profile we can open the profile plot:

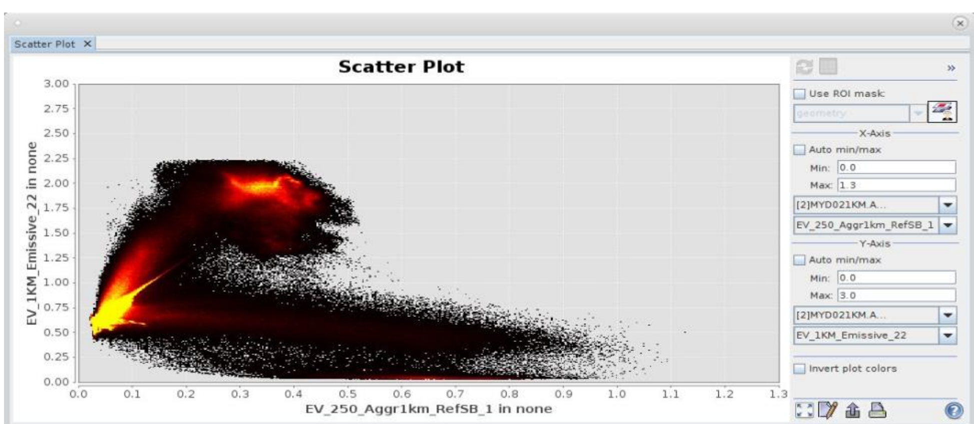




And if we want to profile plot for ROI (Region Of Interest) we have: (it is almost the same with last ones)

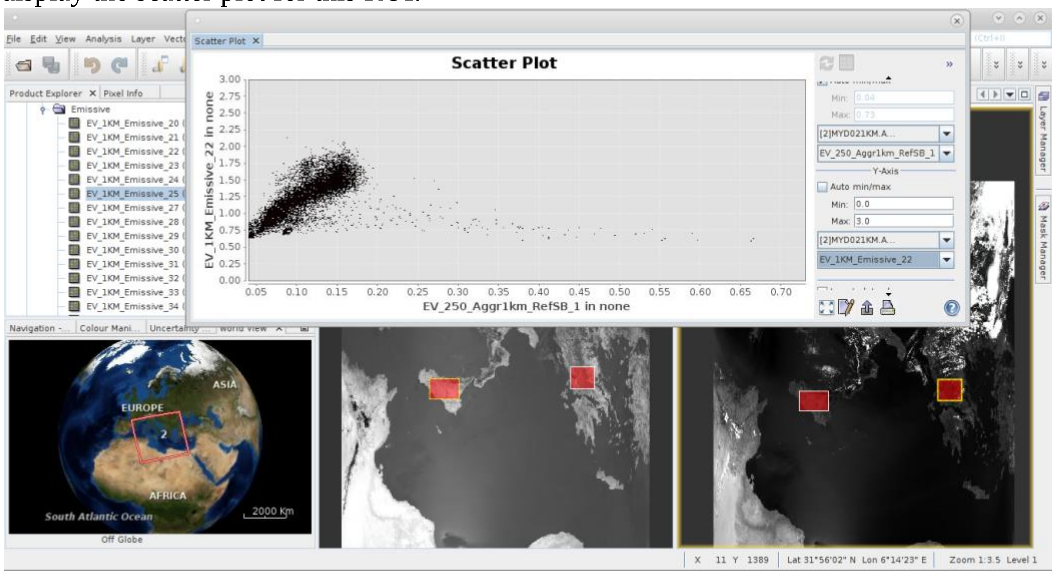


5. <u>Perform and display channel data correlation of the whole image</u>
   For channel data correlation between reflective and emissive, we use the scatter plot (Menu > Analysis > Scatter Plot):

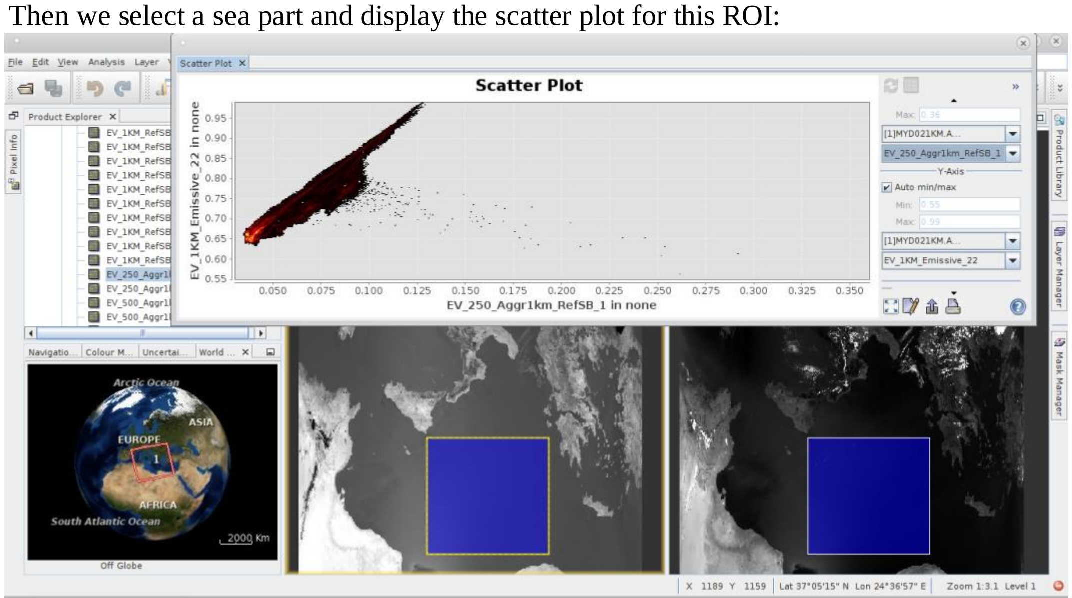


We can clearly see in this plot the correlation of the whole image. The yellow part correspond the high density of pixel frequency. The low reflective and high emitting is related to land and low reflective and medium to high emitting is related to sea and high reflective and low-to-medium emitting is related to clouds.

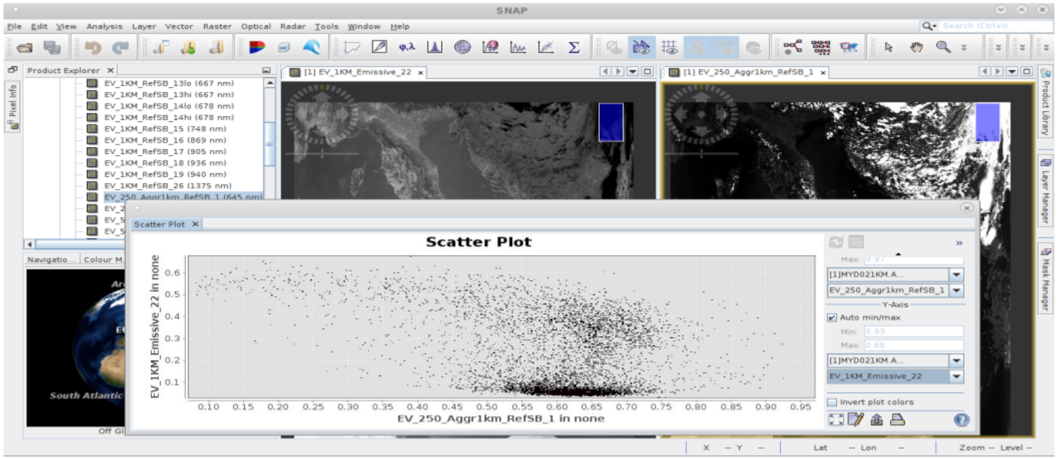
6. <u>Perform and display channel data correlation of selected ROI (Region of Interest)</u>
   We select a land part of image to check if our hypothesis in the last question is correct or not. We display the scatter plot for this ROI:



In the above plot we can see the low reflective and high emitting that we know it is related to the land region.
Then we select a sea part and display the scatter plot for this ROI:



In this plot we can see low reflective and medium to high emitting. And then we select a clouds part and display the scatter plot for this ROI:



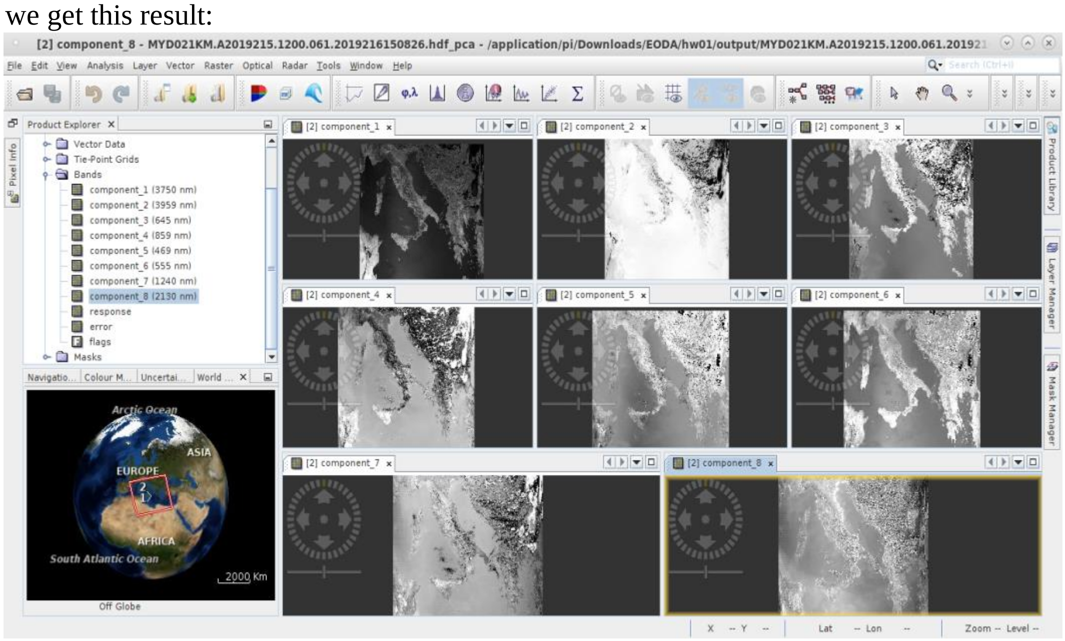
Here we can clearly see that the low emissive with high reflective is related to clouds region. Therefore, our concluding in the previous question is correct.

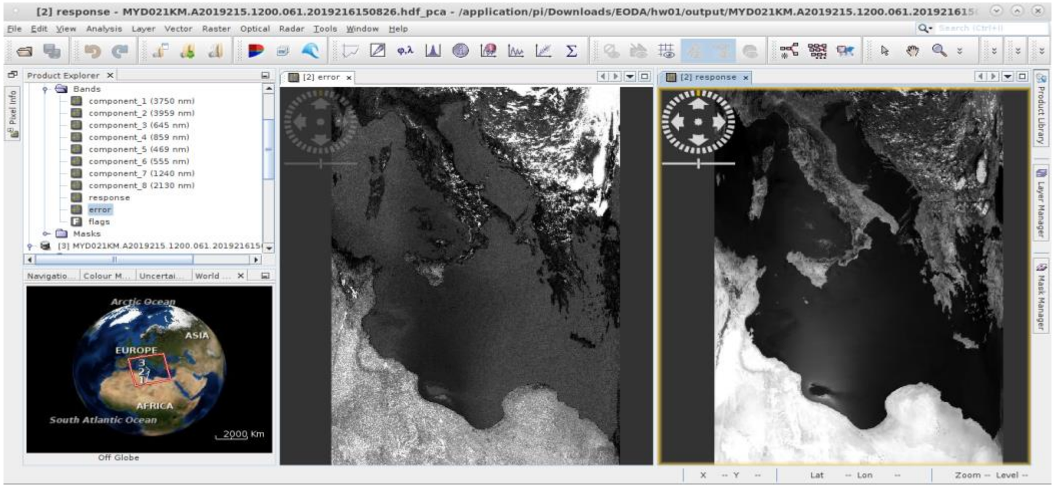
7. Perform and display principal component analysis

   For principal component analysis we need high quality images, so we check the quality of images one more time:

   - Reflective channels Aggr1km
     - OK for: 1-2-3-4-5-7
     - Failures/strips on: 6
   - Emissive channels EV
     - OK for: 20-21-22-23-24-25-29-30-31-32-33
     - Failures/strips on: 27-28-34-35-36

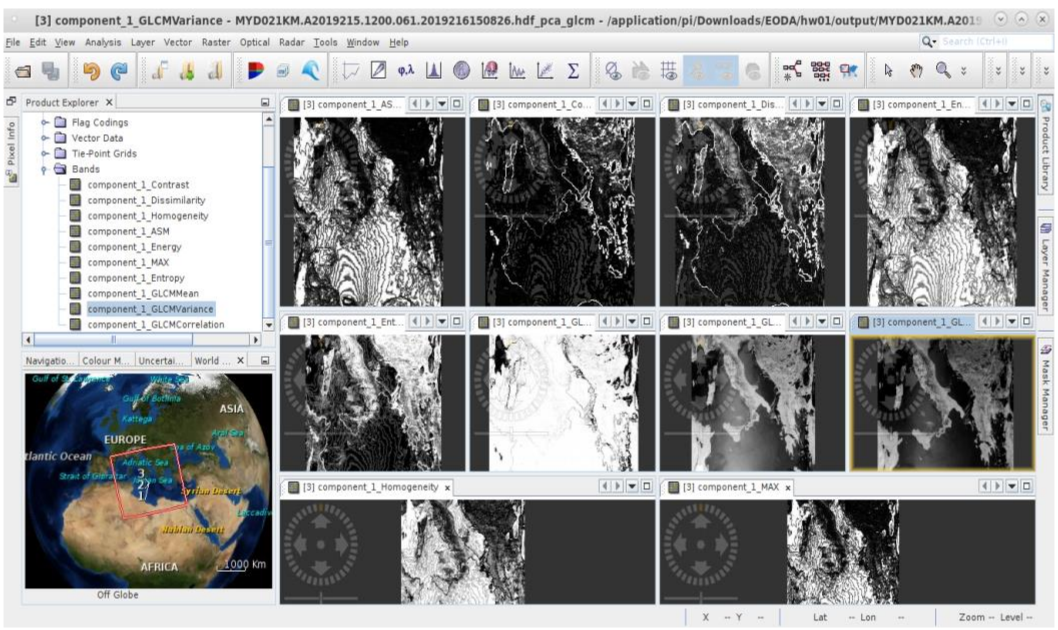
   Then from Menu > Raster > Image Analysis > Principal Component Analysis we perform PCA analysis. For this purpose we have to specify the bands in processing parameters tab. We decide to select 1-2-3-4-5-7 from reflective and 20-22 from emissive bands. After perform the analysis we get this result:





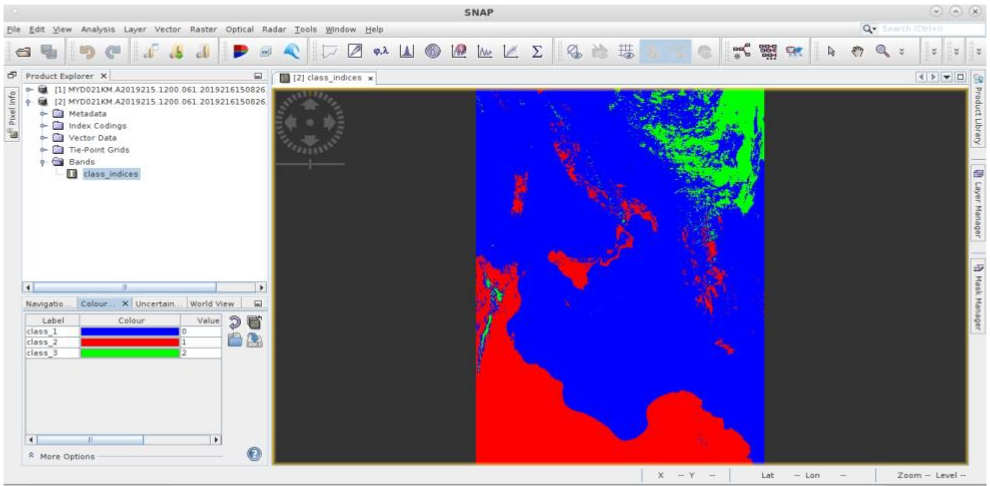
We can see in the above picture of our result, except component number 2 that approximate noise features, all others have good information. Components number 1, 3, 4, and 7 have more contrast and information(details).

Also, we perform Texture analysis (Menu > Raster >Image Analysis > Texture Analysis > GLCM) in order to get more information about our PCA and comparative information about them. The result is shown below:



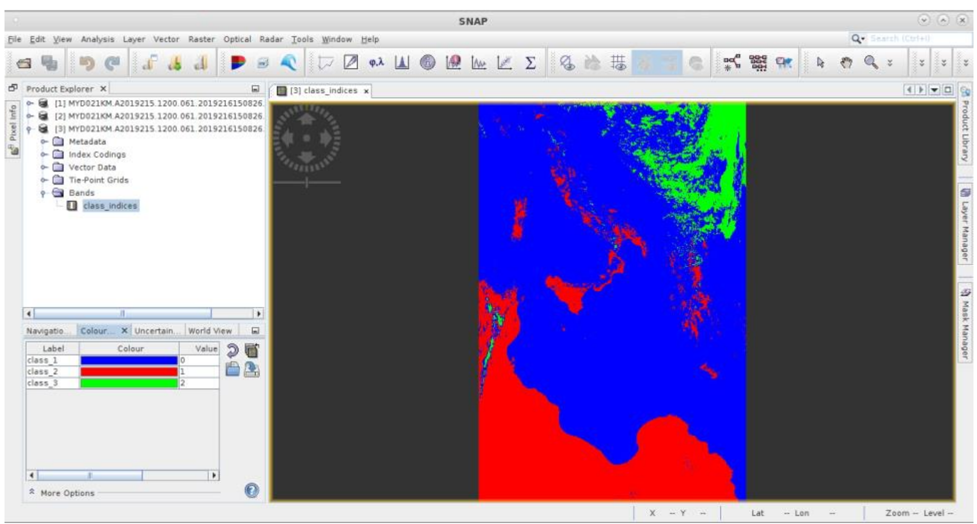
8. Perform, display and interpret unsupervised classification with at least 3 classes (sea, land, cloud)
   For unsupervised classification (from Menu > Raster > Classification > Unsupervised > K-Means Cluster Analysis) we decide to use K-Means Cluster Analysis and for at least 3 classes of sea, land without clouds, clouds over land. For this purpose and in order to get good results, we decide to apply our K-Means algorithm with different source bands as input and also, with 3 clusters and 30 iterations and random seed of 31415. The results we've obtained are shown below. We consider that Blue is for the Sea, Red is for the Land, Green is for the Cloud:

- Using all 38 bands:



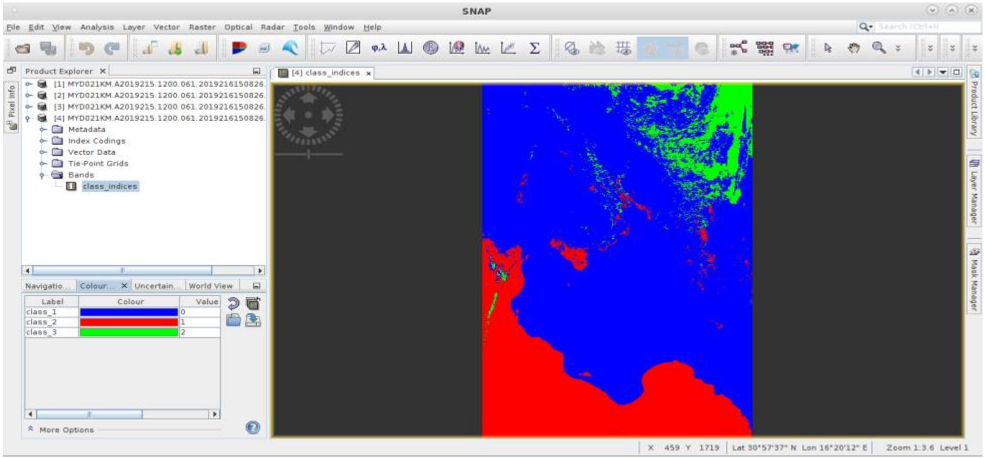
As we can see with this bands, our result and classification is not that accurate, especially in detection of Italy's land. The right above clouds and left down land and sea sections are detected good enough. However, we it could not detect the clouds over Italy.

- Using just Ok bands (without using failures/strips) that are :
    - Reflective channels Aggr1km
    - OK for: 1-2-3-4-5-7
- Emissive channels EV
    - OK for: 20-21-22-23-24-25-29-30-31-32-33



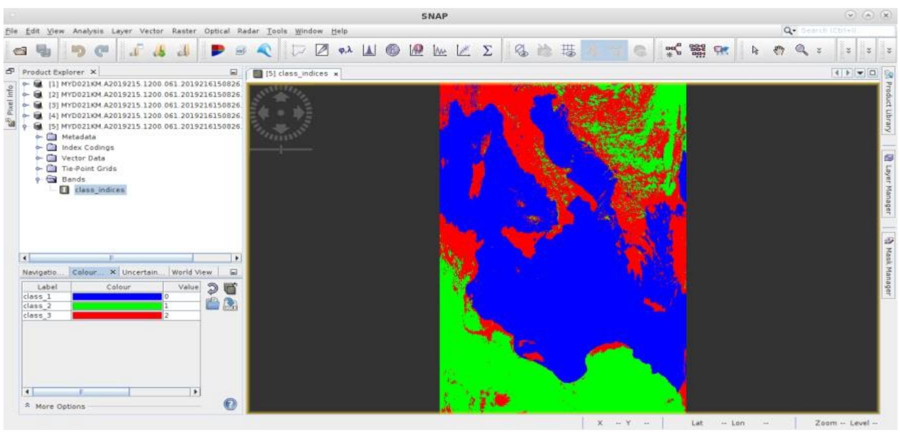
As we can see here the result did not change or get better with compared of all bands in input.

- Using just high quality bands that are:
    - 1-2-3-4-5-7 from reflective bands
    - 20-22 from emissive bands



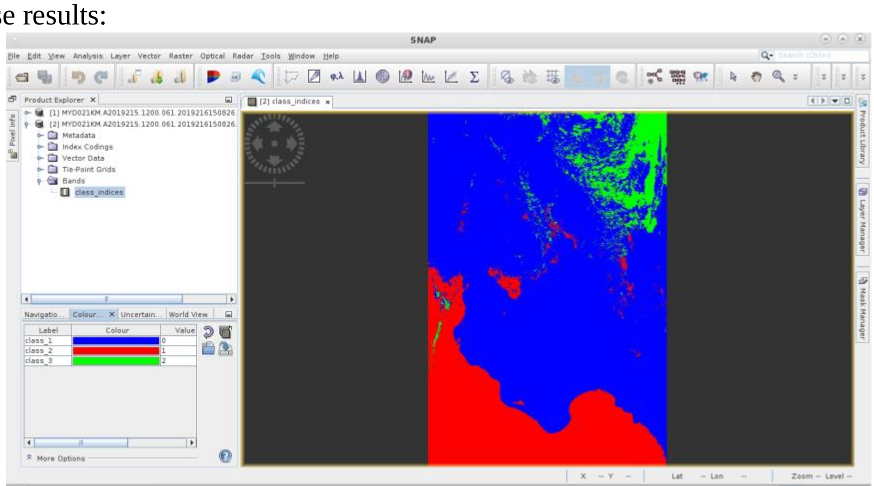
In this result we can clearly see that it do better in detection of clouds over Italy but the land of Italy itself is still do not detected. All other sections is the same with our last results.

- Using 1-2-3-4-5-7 from reflective bands and number of iterations equal to 50:

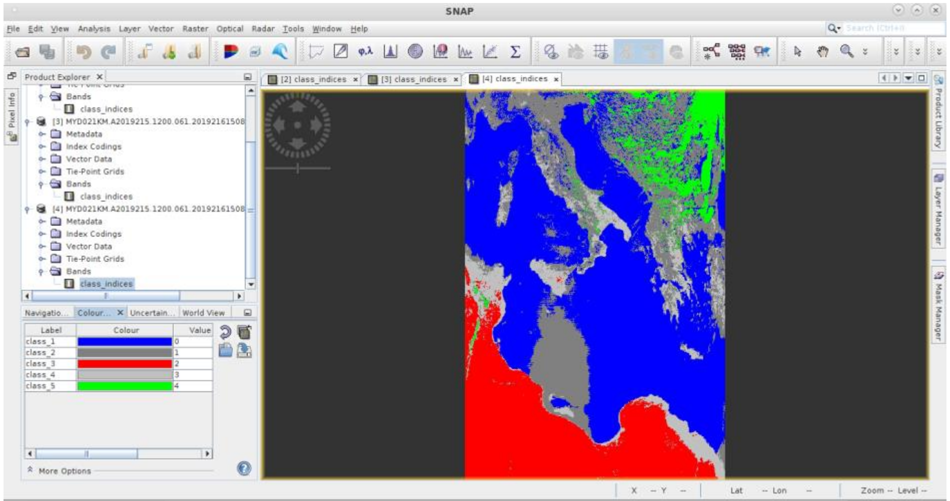


As we can see in this result, our algorithm did his job very good and detect land, sea, and cloud over lands very well. The quality and accuracy in this result is in its highest state.

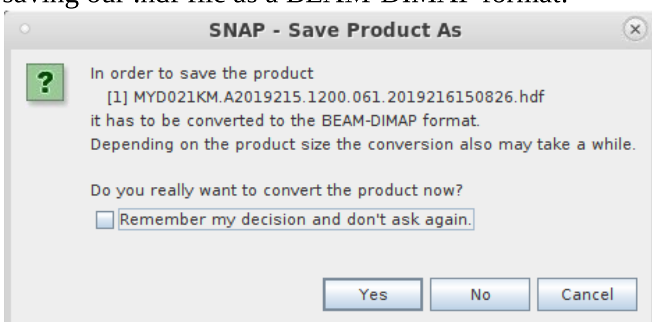
After we've decide to apply K-Means Analysis using the PCA analysis results with component 1 to 8 and number of cluster equal to 3 and 30 iterations, instead of the MODIS channels and we've get these results:



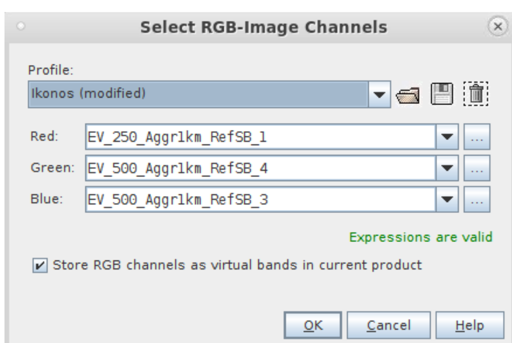
It's result has not better quality or accuracy but if we change the number
of clusters to 5 and use all the source bands, we get this result that is a little noisy but it detected better the classes:

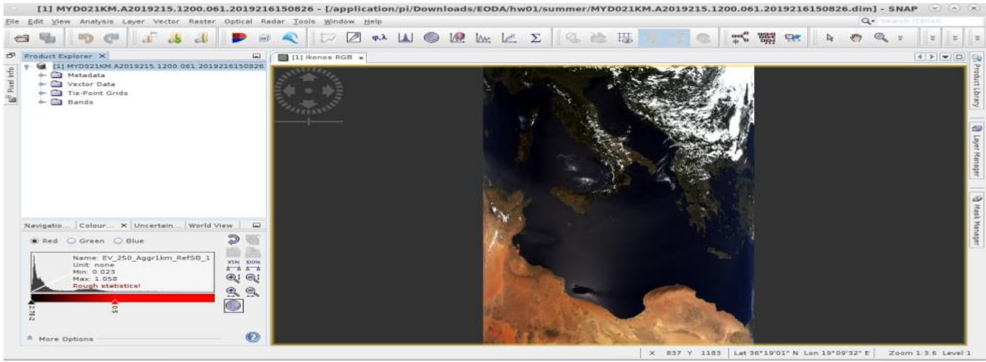


9. <u>Perform, display and interpret supervised classification with at least 4 classes (sea, land, cloud)</u>
First we start by saving our .hdf file as a BEAM-DIMAP format:

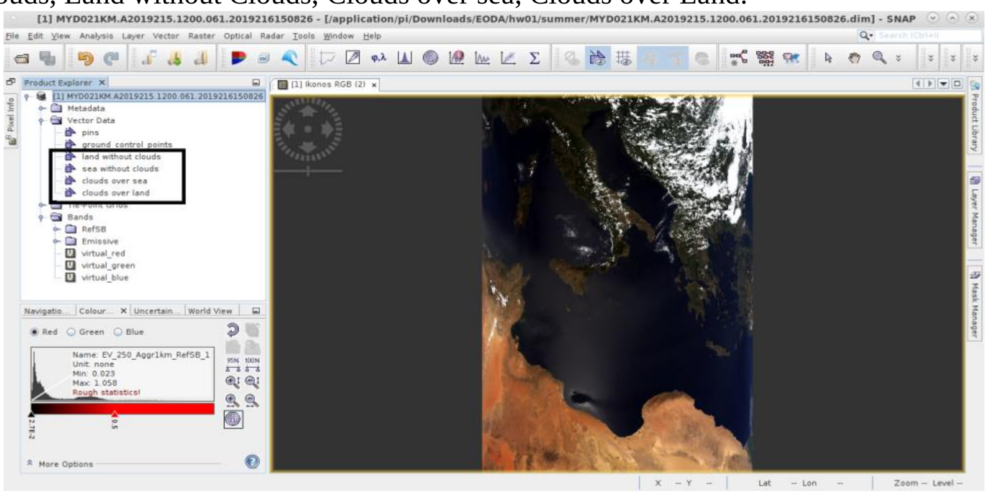


Then we open a RGB image window with bands that is shown below (we select "Store RGB channels as virtual bands in current product" for future use):

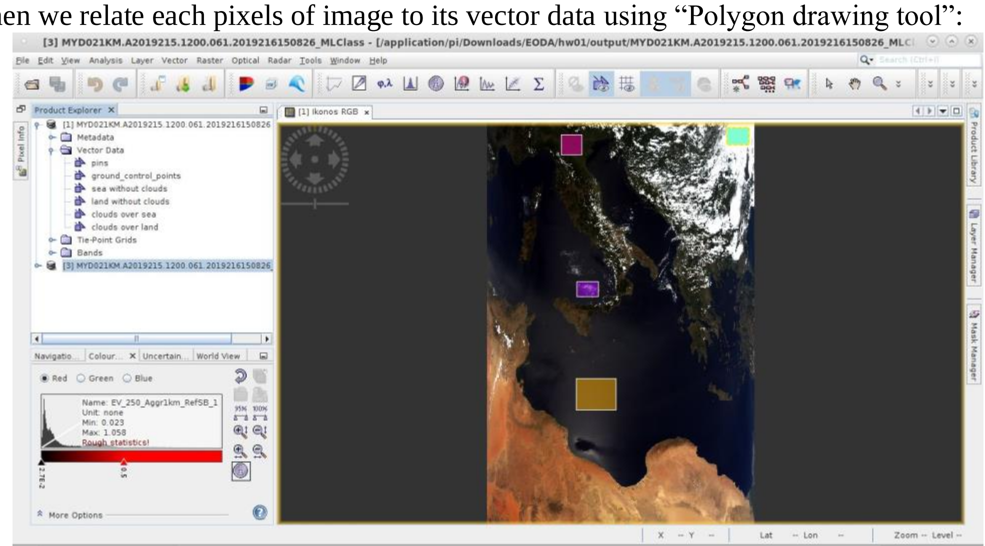




Then we create some Vector Data Container correspond to our 4 classes that are Sea without Clouds, Land without Clouds, Clouds over sea, Clouds over Land.

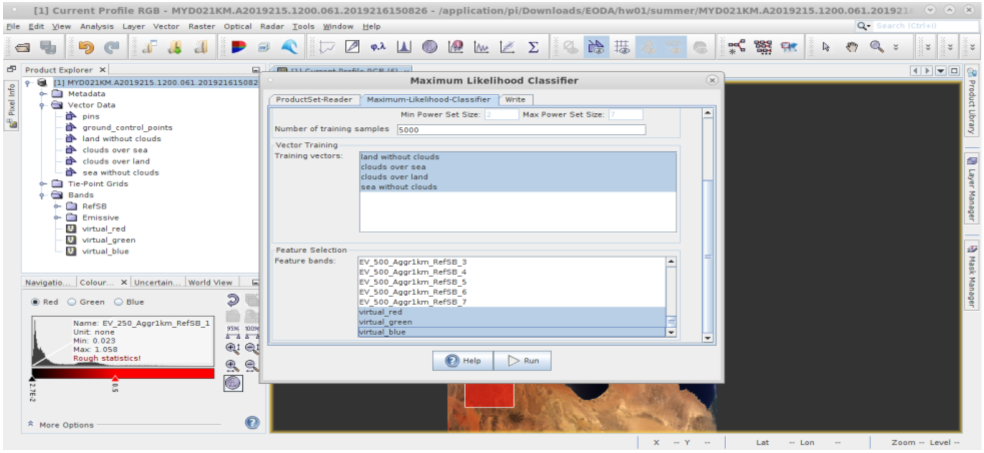


Then we relate each pixels of image to its vector data using "Polygon drawing tool":

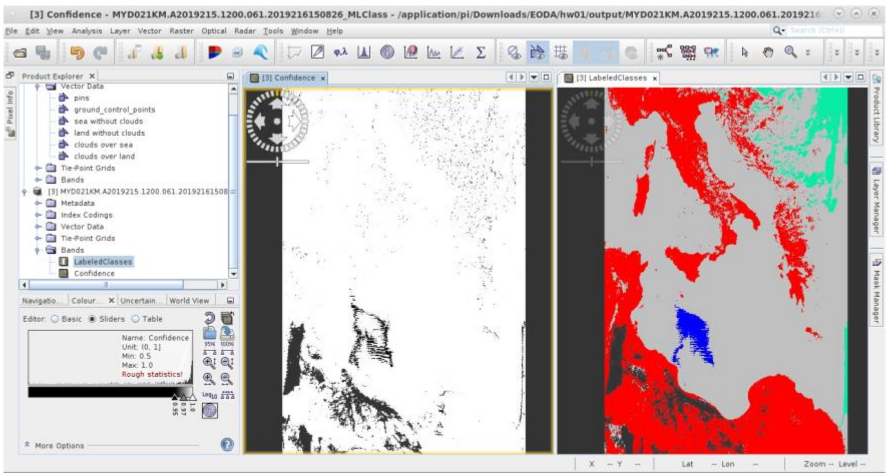


In this step we apply our Supervised Classification Algorithm.

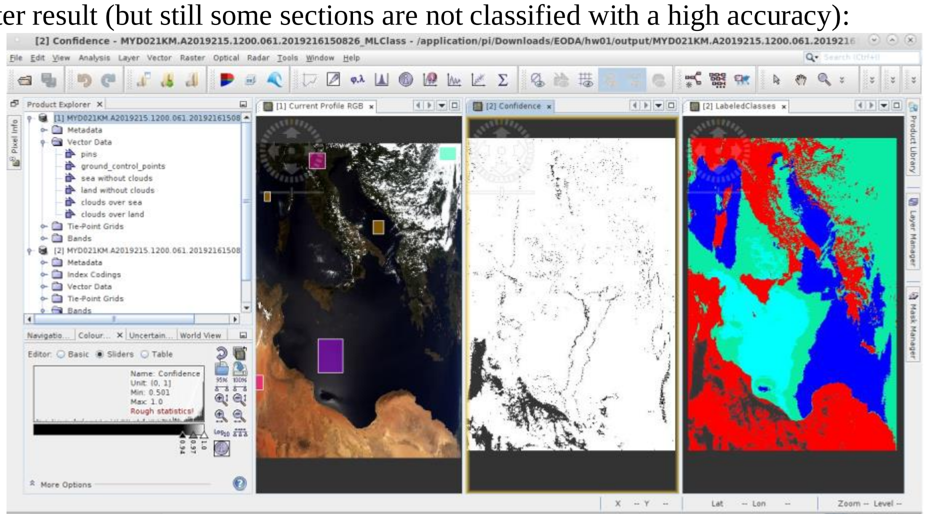
- Maximum Likelihood Classifier:



In our first try the result was not that good:

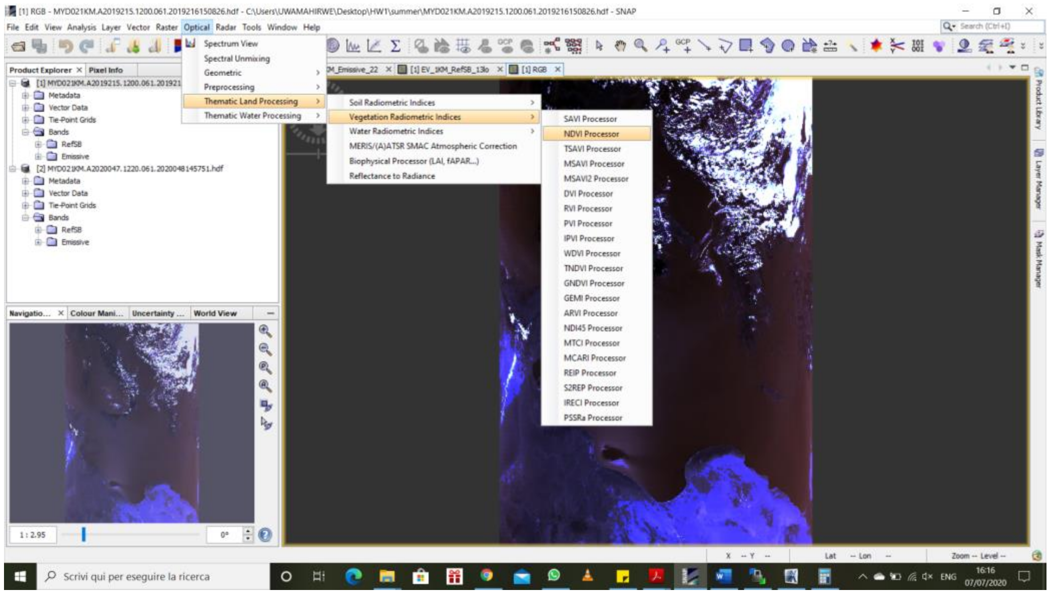


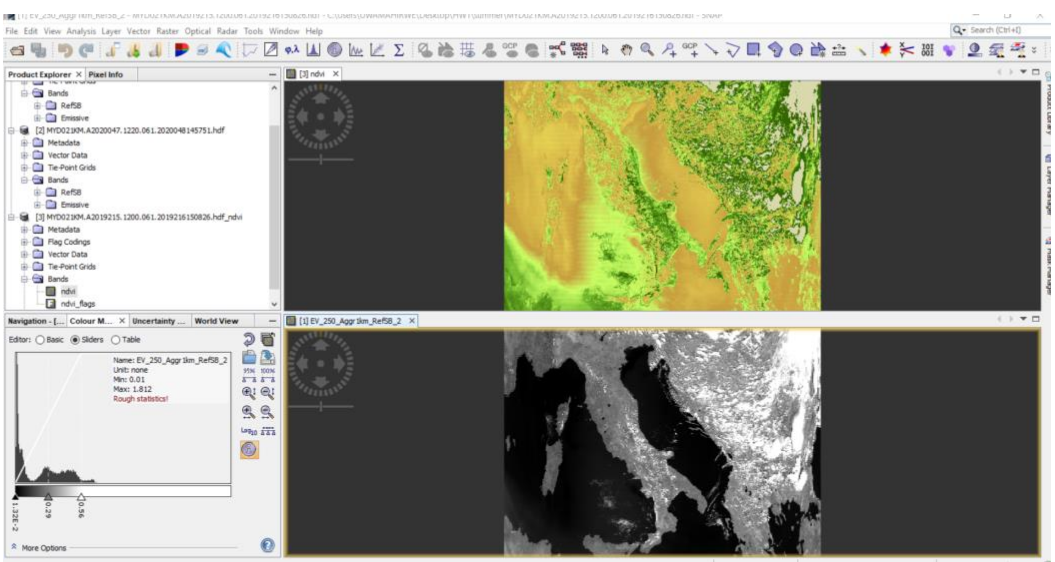
So we change the Polygon using “Polygon drawing tool” again and try one more time and we got a better result (but still some sections are not classified with a high accuracy):



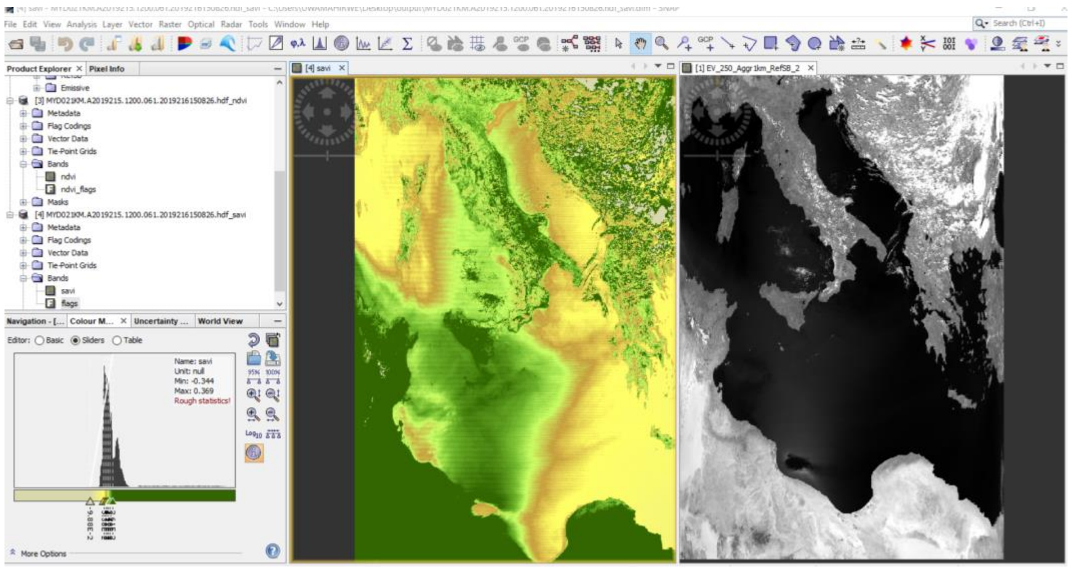
10. <u>Implement at least 3 formulas of 2-band and 3-band vegetation index (VI) using SNAP processing tools</u>

The normalized difference vegetation index (NDVI) is a simple graphical indicator that can be used to analyze remote sensing measurements, typically but not necessarily from a space platform, and assess whether the target being observed contains live green vegetation or not. This tutorial shows how to compute and interpret the results of the NDVI SNAP operator.

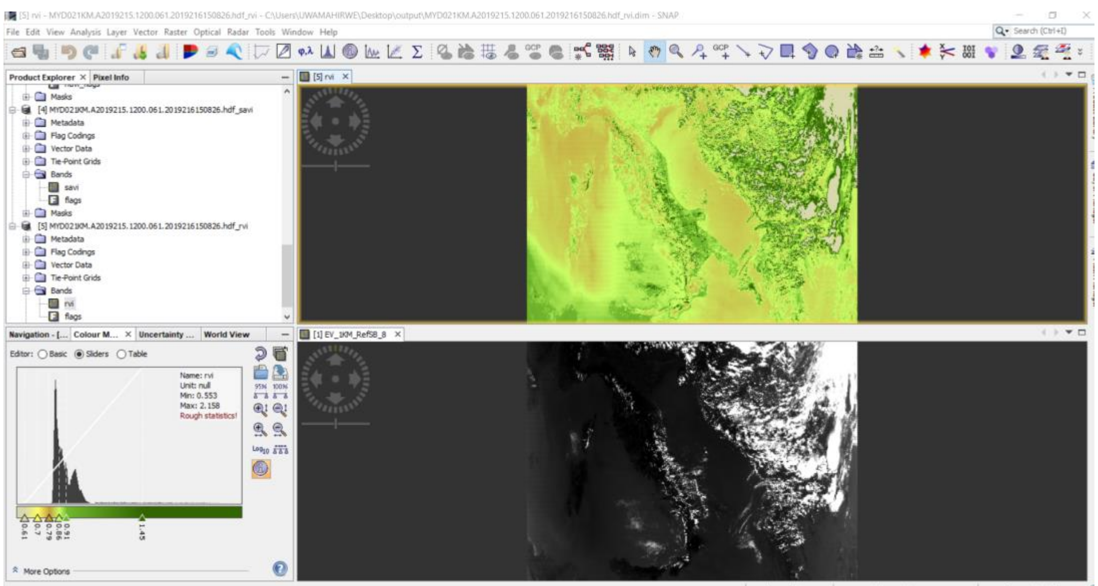




Ratio Vegetation Index (RVI), is simply the amount of healthy vegetation, so the higher. the value of this ratio will be, the healthier vegetation area.

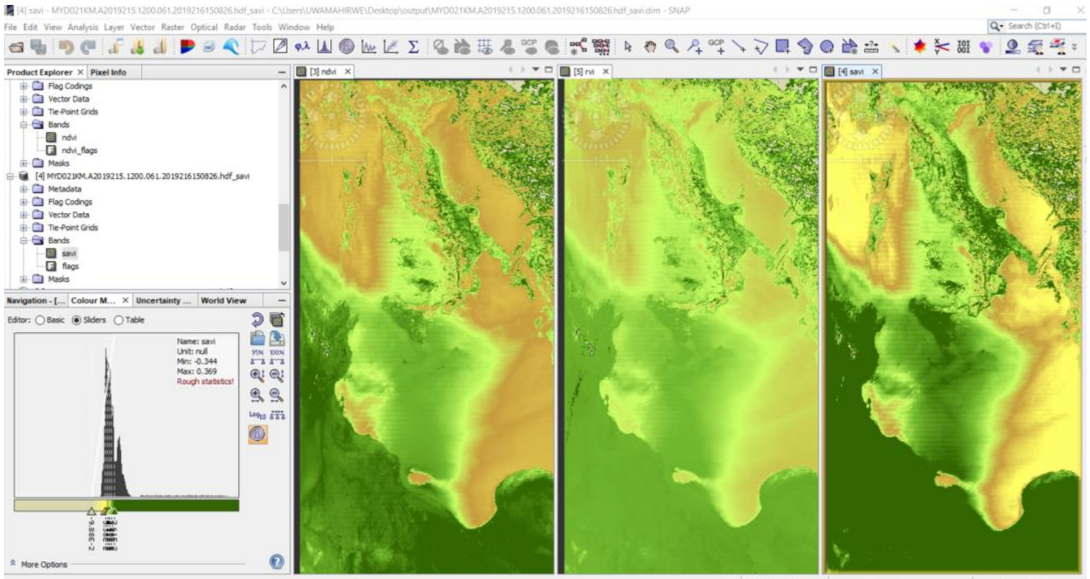


(Soil-Adjusted Vegetation Index (SAVI), Empirically derived NDVI products have been shown to be unstable, varying with soil color, soil moisture, and saturation effects from high density vegetation.



11. Apply at least 3 VI formulas to land pixels and interpret their output results and differences

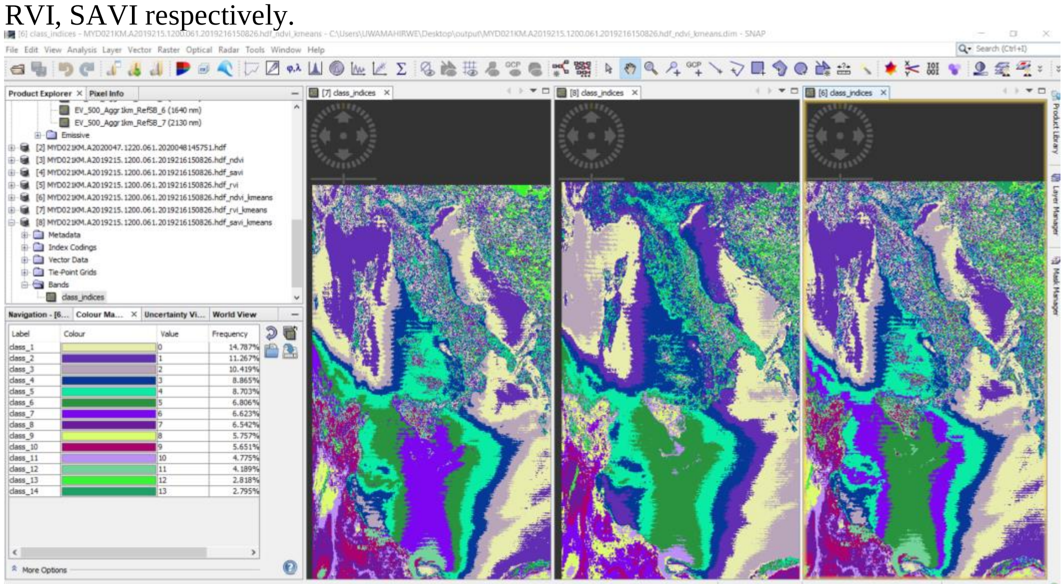
NDV I = (NIR + RED)/ (NIR + RED)
RVI =RED/ NIR
SAVI =((1+L) *(NIR-RED)) / (NIR + RED + L)



The image above shows NDVI, RVI and SAVI Vegetation Index from left to right respectively, as we can see, NDVI showing more details and vegetation area than two other VIs and RVI showing the less detail and vegetation area than other two VIs. In RVI we can easily distinguish the clouds from vegetation areas, while it is trickier to distinguish between clouds and vegetation areas, and it is due to the wavelengths in RVI which is way higher than two other VIs.

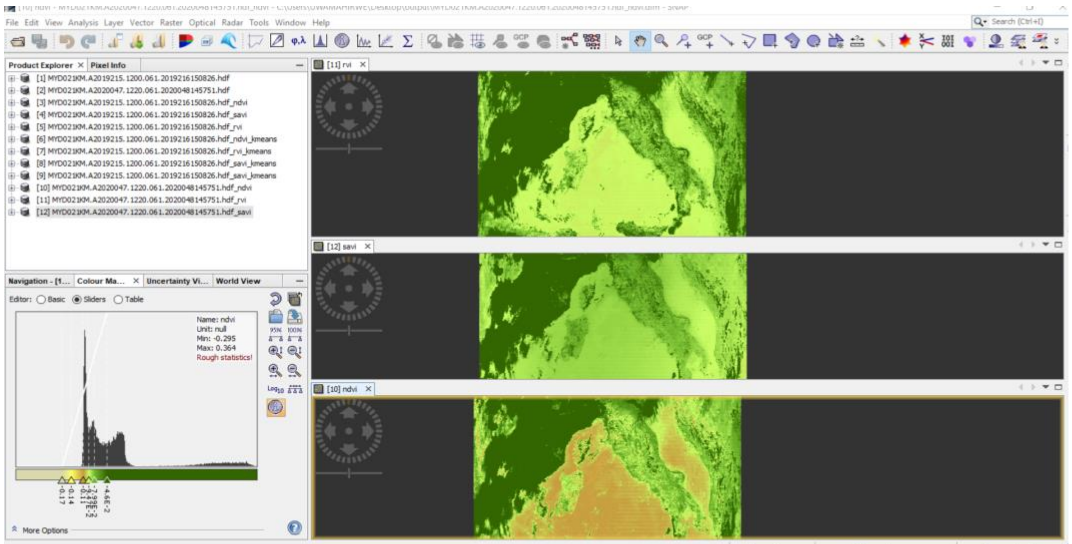
12. <u>Perform an unsupervised classification by using an arbitrary VI index instead of MODIS channel</u>

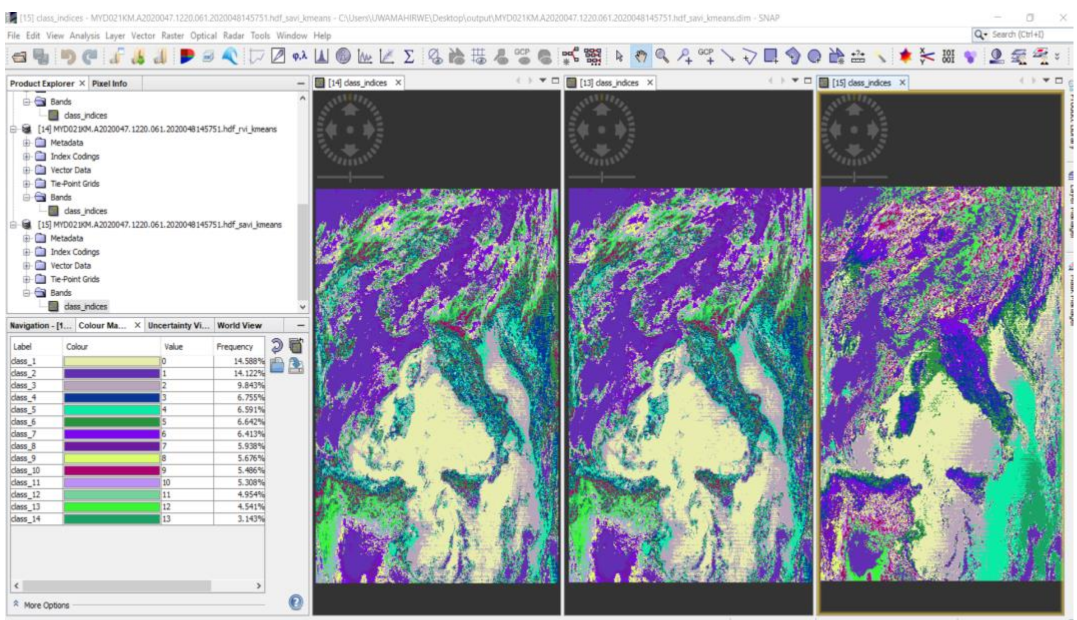
    K-Means from unsupervised classification algorithm, by 50 iterations and 14 clusters and the following results has been obtained. from left to right k-means algorithm been applied on NDVI, RVI, SAVI respectively.



13. <u>Download a second MODIS image at 1-km resolution over the same geographical area of interest during a winter season. Apply the same selected VI index and qualitatively compare the differences.</u>

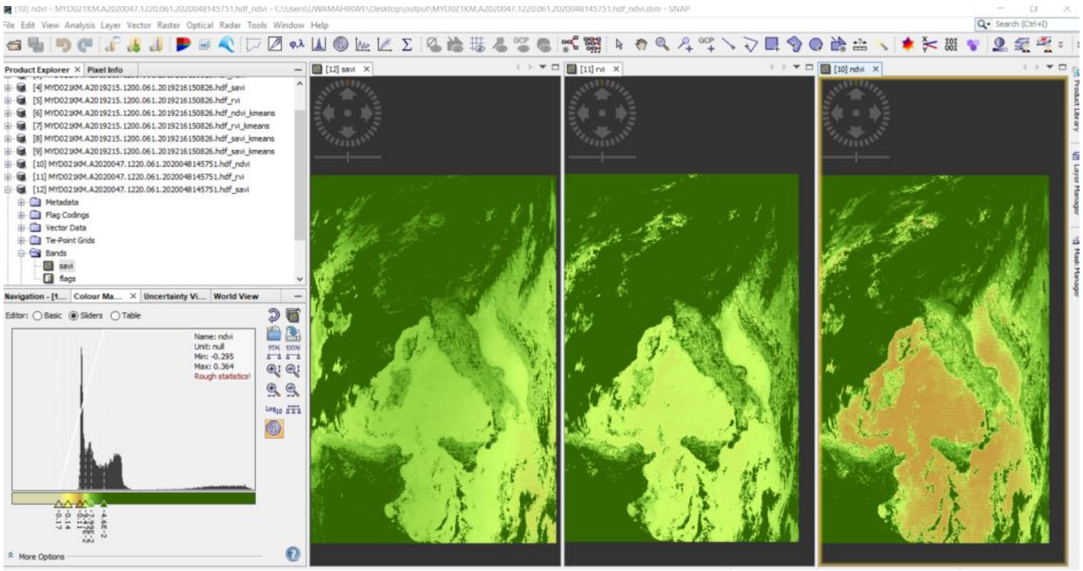
We’ve selected the second MODIS image (MYD021KM.A2020047.1220.061.2020048145751.hdf) of winter 2020 (2020-02-01 09:50:00) over Italy.





14. <u>To perform a quantitative change detection (difference) of the vegetation coverage class by reprojecting the 2 winter and summer MODIS images over the same grid in a selected region of interest (ROI).</u>

The Result is illustrated above are Vegetation Index over same geographical area in winter, from left to right we can see the SAVI, RVI, NVDI respectively. As we can see clearly NDVI and SAVI can detect more vegetation areas while these two methods are not able to distinguish perfectly between cloud and vegetation areas. In the other hand RVI detects less vegetation areas while it can distinguish cloud s perfectly.



Difference between summer and winter:

On the left we can see the result of NDVI in winter and on the right we can find the result of NDVI in summer, clearly the density of VI in winter is more that summer, it can be due to high density of clouds in the selected day of summer in comparison to the winter image.