Trans-Atlantic Training 2022

Prague, Monday the 5th of September 2022

Vegetation classification with Sentinel-1 data

P.-L. Frison, C. Lardeux,

pierre-louis.frison@univ-eiffel.fr

Data and context

The aim is to use radar (Sentinel-1) and optical (Sentinel-2) remote sensing data together or separately to produce the cartography of the different vegetation types. The study area is the Cayenne region, in French Guiana. The first step will be to pre-process the satellite data. Then we will perform a supervised classification based on the Random Forest algorithm to produce vegetation maps.

The data used for learning and validating the classifications were obtained during *in situ* inventories. The Sentinel data used are detailed in the following table:

	Acquisition dates	Spatial resolution	wavelengths	Polarisation	Product Level
Sentinel-1	2017/04/25 2017/07/06 2017/10/10 2018/01/14	~ 20 m	5.6 cm (C Band)	VV/VH	GRDH
Sentinel-2	2018/02/06	4 bands: 10 m	430 nm – 1600 nm		L1C

In the first part, the data are processed to enable them to be used by the classification algorithm.

The second part is dedicated to the classification for the production of the vegetation map.

I. *DATA PRE-PROCESSING*

The Sentinel-1 radar data were acquired in 2 distinct polarisations: VV and VH. From these two polarizations, we will produce three bands per acquisition: VV, VH, and VH/VV, filtered to reduce speckle noise.

The Sentinel-2 optical data have 13 bands acquired in different wavelengths. We will use the 4 bands acquired at 10 m to which we will add the NDVI.

I. 1. Sentinel-1:

The Sentinel-1B data were acquired in IW (Interferometric Wide) mode and were downloaded from https://search.asf.alaska.edu/ in GRDH format.

They need to be orthorectified (in the UTM zone 32N geographic projection, EPSG : 32622) and calibrated (to convert the digital number DN into a physical quantity: the radar backscatter coefficient σ^0) over the area of interest defined by the ***study_area.shp*** vector layer. The resulting images are then stacked (each pixel of each image corresponds to the same position on the Earth).

Then, a spatio-temporal filter[1] is applied to reduce the speckle noise inherent in radar images.

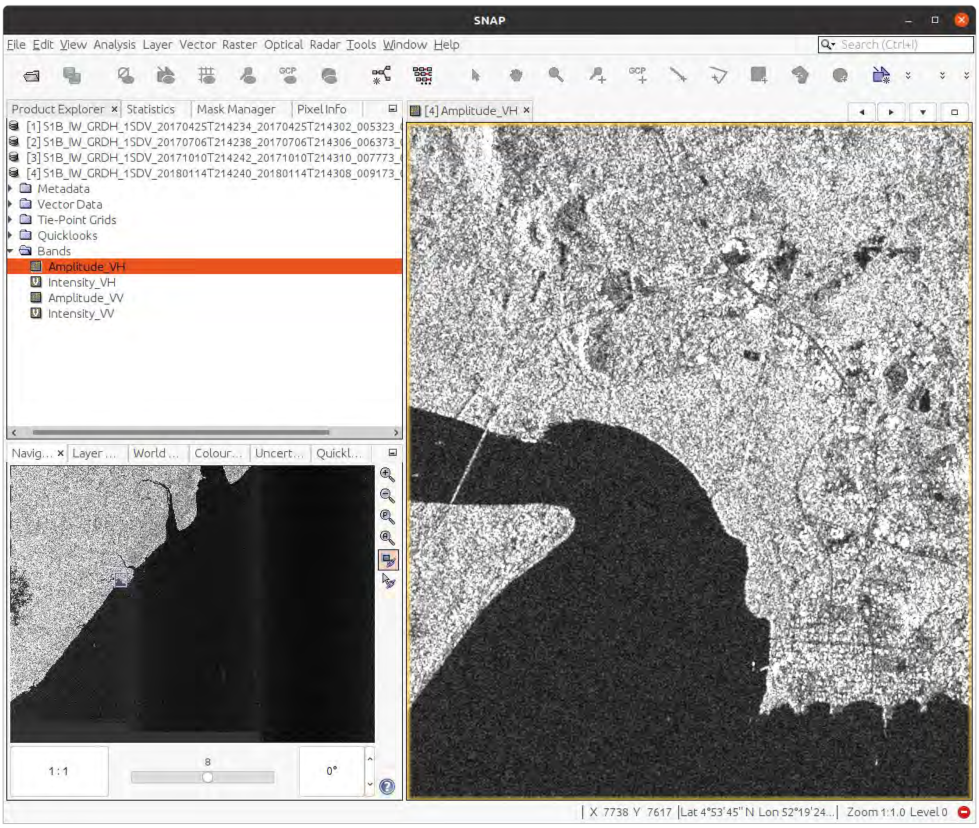
The first pre-processing steps will be carried out with SNP. They consist in:

- Calibration and Ortho-rectification of the 4 S1 Acquisitions
- Spatio-temporal filtering from speckle the VV and VH polarizations

The last pre-processing, namely, the adjunction of the third channel (the VH/VV ratio) will be made with a QGIS model.

I.1.1. Calibration and Ortho-rectification

First launch the SNAP toolbox and open the four Sentinel-1 products (.zip format) located in ***FrenchGuiana/data/S1/zip***

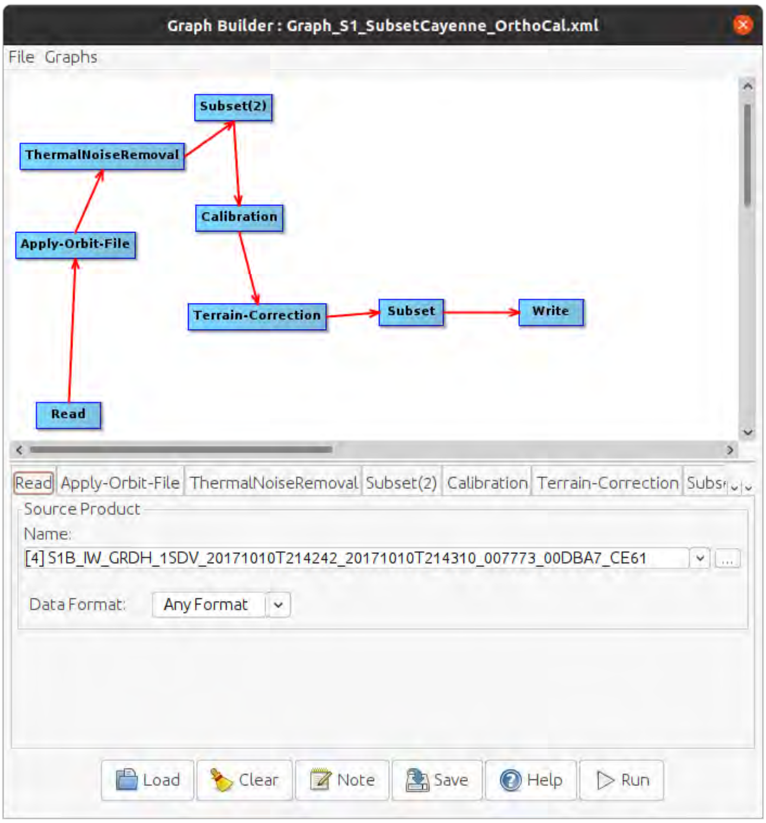


Original Sentinel-1 GRDH products opened in SNAP. They are in “image” geometry, i.e. lines and columns of the image are horizonthal/vertical

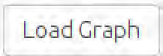
---

[1] S. Quegan, J. J. Yu, 2001: Filtering of Multichannel SAR images, IEEE Transactions on Geosciences and Remote Sensing, vol. 39, N° 11, 2373-2379.

Open the Graph Builder (***Tools →Graph Builder***) and load (***File → Load Graph***) the graph ***FrenchGuiana/codes/snap/Graph_S1_SubsutCayenne_OrthoCal.xml***. You can see the whole processing chain allowing the calibration and ortho-rectification of a Sentinel-1 GRD product.



After a refinement of the satellite position and the removal of thermal noise, the data are cropped over the study area (cf ***Subset*** tab), calibrated and then orthorectified (cf ***Terrain-Correction*** tab) from a DEM over the area of interest.

In order to apply this processing to the four Sentinel-1 products open the *Batch Processing tool* (***Tools → Batch Processing***). Click on the button in order to load all the images already opened in SNAP. Then click on 

and load the ***FrenchGuiana/codes/snap/Graph_S1_SubsutCayenne_OrthoCal.xml*** graph.

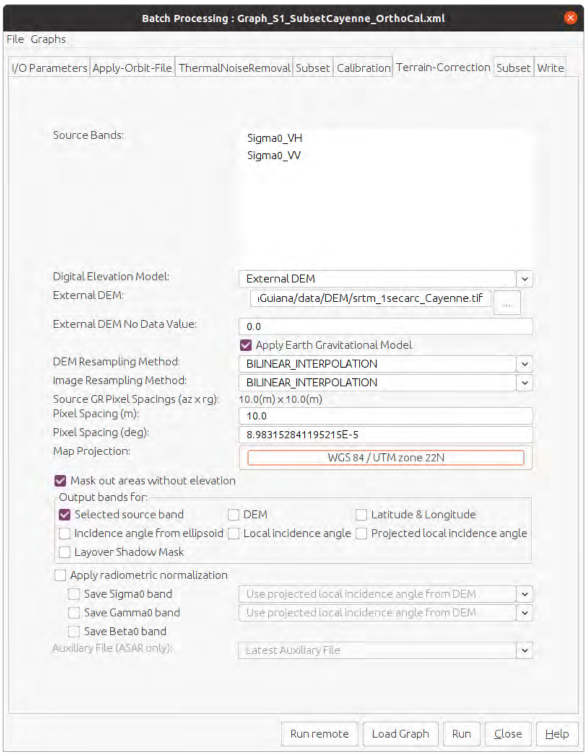
You can see as many tabs as there are bricks in the processing ***Graph_S1_SubsutCayenne_OrthoCal.xml*** chain.

Some modifications has to been made:

In the ***Terrain-Correction*** tab:

- Select ***External DEM*** in the ***Digital Elevation Model*** drop-down menu
  In ***External DEM*** load ***FrenchGuiana/data/DEM/srtm_1secarc_Cayenne.tif***[2]
- Click on the ***Map Projection*** button, the ***Predefined CRS*** (Coordinates Reference System), ***Select…***, and load the ***EPSG:32622 – WGS 84 / UTM Zone 22N***

[2] This DEM file has been downloaded from USGS Earth Explorer web site https://earthexplorer.usgs.gov/



In the ***Write*** tab, enter the Directory ***FrenchGuiana/data/S1/Snap/OrthoCal***

Then ***Run*** the Batch Processing.

Once the Processing has been completed, The four products cropped over the study area are open automatically in SNAP

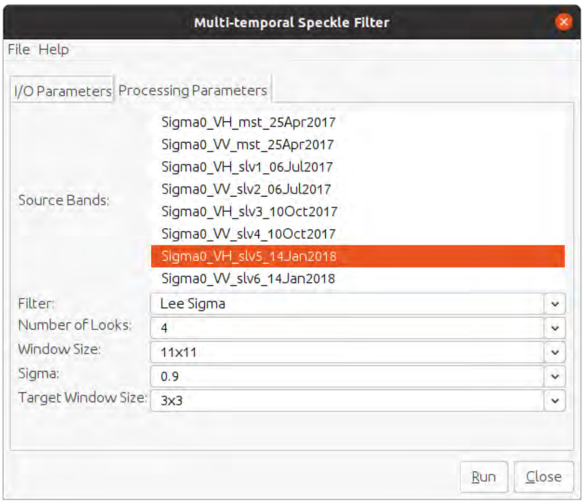
Close the first 4 products that we don't need any more.

I. 1. 2: Speckle FIltering

In order to us a spatio-temporal filter, we first need to realize a coregitration of the products. To do so: ***Radar → Coregistration***. Click on the button in order to load the 4 images still opened in SNAP (Calibrated and Ortho-rectified products).
Click on the ***Write*** tab and enter ***FrenchGuiana/data/S1/Snap*** as the output Directory.
Run the Coregistration
Once the processing has been completed, the coregistred product (***Subset_Subset_....Orb_NR_Cal_TC_Stack***) is automatically opened in SNAP.

☛ *Optional:* You can process the temporal average (***Radar → Coregistration → Stack Tools → Stack Averaging***) to compare speckle noise with an original GRD product.

To apply the spatio-temporal speckle filter, go to ***Radar → Speckle Filtering → Multi-temporal Speckle Filter***, then select the Coregistered file as ***Source Product***, and enter as output Directory ***French/Guiana/data/S1/Snap***. In the Processing Parameters tab, fulfill fields as follow:



Then ***Run*** the filter.

Once the processing has been completed, the spatio-temporally filtered product (***Subset_Subset_....Orb_NR_Cal_TC_Stack_Spk***) is automatically opened in SNAP.

I. 1. 3. VH/VV polarization channel

The last step is to build the third polarization channel: the VH/VV polarization channel from the VV and VH speckle filtered polarization channels. This will be done with QGIS software and its model builder.

Launch QGIS.

Then, in ***Processing → Graphical Modeler***, go to ***Model → Open Model*** and load the model ***FrenchGuiana/codes/qgis_model/PolarCompoCol.model3***.

This model creates the VH/VV polarization channel, in dB scale. In the aftermath, we transform also the VV and HH polarization channels in decibel scale, better suited for visual observation. And a virtual file, composed of these 3 channels is created to generate a Color Composite Image with the following code:

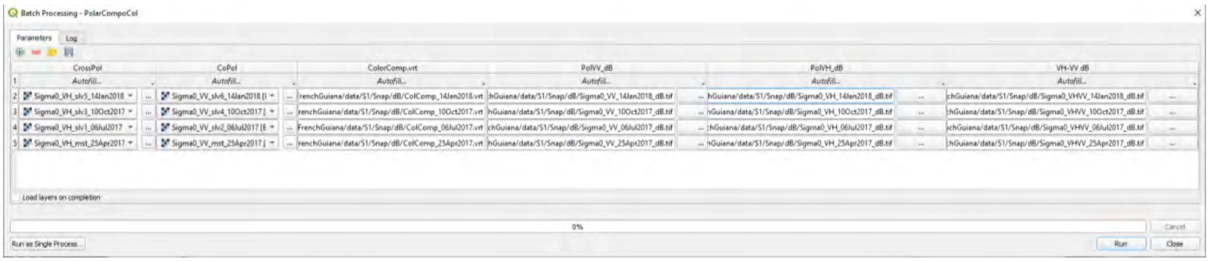
R : polarisation ratio VV/VH

V : VH Polarisation

B : VV Polarisation

In order to create these files for each of the 4 Sentinel-1 acquisitions, Click on ***Run as batch process…***

Then, fulfill fields as follow:



For the first line:
Crosspol: FrenchGuiana/data/S1/Snap/Subset_....TC_Stack_Spk.data/Sigma0_VH_slv5_14Jan2018.img
Copol: FrenchGuiana/data/S1/Snap/Subset_....TC_Stack_Spk.data/Sigma0_VV_slv6_14Jan2018.img
ColorComp.vrt: FrenchGuiana/data/S1/Snap/ColComp_14Jan2018.vrt
PolVV dB: FrenchGuiana/data/S1/Snap/Sigma0_VV_14Jan2018_dB.tif
PolVH dB: FrenchGuiana/data/S1/Snap/Sigma0_VH_14Jan2018_dB.tif
VH-VV dB: FrenchGuiana/data/S1/Snap/Sigma0_VHVV_14Jan2018_dB.tif

the 2nd line:
Crosspol: FrenchGuiana/data/S1/Snap/Subset_....TC_Stack_Spk.data/Sigma0_VH_slv3_10Oct2017.img
Copol: FrenchGuiana/data/S1/Snap/Subset_....TC_Stack_Spk.data/Sigma0_VV_slv6_10Oct2017.img
ColorComp.vrt: FrenchGuiana/data/S1/Snap/ColComp_10Oct2017.vrt
PolVV dB: FrenchGuiana/data/S1/Snap/Sigma0_VV_10Oct2017_dB.tif
PolVH dB: FrenchGuiana/data/S1/Snap/Sigma0_VH_10Oct2017_dB.tif
VH-VV dB: FrenchGuiana/data/S1/Snap/Sigma0_VHVV_10Oct2017_dB.tif

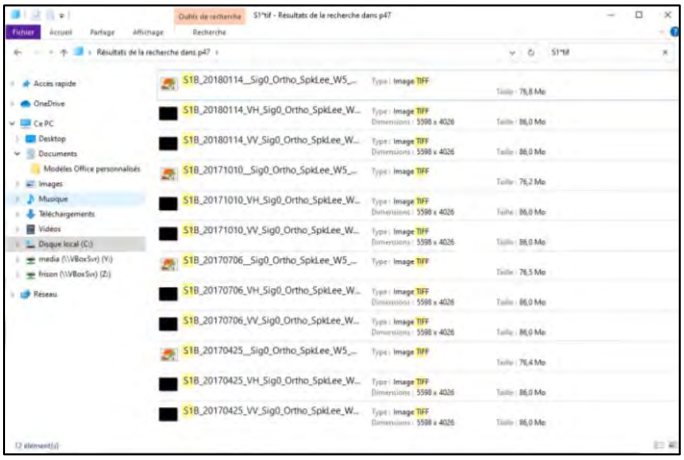
the 3rd line:
Crosspol: FrenchGuiana/data/S1/Snap/Subset_....TC_Stack_Spk.data/Sigma0_VH_slv1_10Oct2017.img
Copol: FrenchGuiana/data/S1/Snap/Subset_....TC_Stack_Spk.data/Sigma0_VV_slv2_10Oct2017.img
ColorComp.vrt: FrenchGuiana/data/S1/Snap/ColComp_06Jul2017.vrt
PolVV dB: FrenchGuiana/data/S1/Snap/Sigma0_VV_06Jul2017_dB.tif
PolVH dB: FrenchGuiana/data/S1/Snap/Sigma0_VH_06Jul2017_dB.tif
VH-VV dB: FrenchGuiana/data/S1/Snap/Sigma0_VHVV_06Jul2017_dB.tif

the 4th line:
Crosspol: FrenchGuiana/data/S1/Snap/Subset_....TC_Stack_Spk.data/Sigma0_VH_slv1_25Apr2017.img
Copol: FrenchGuiana/data/S1/Snap/Subset_....TC_Stack_Spk.data/Sigma0_VV_slv2_25Apr2017.img
ColorComp.vrt: FrenchGuiana/data/S1/Snap/ColComp_25Apr2017.vrt
PolVV dB: FrenchGuiana/data/S1/Snap/Sigma0_VV_25Apr2017_dB.tif
PolVH dB: FrenchGuiana/data/S1/Snap/Sigma0_VH_25Apr2017_dB.tif
VH-VV dB: FrenchGuiana/data/S1/Snap/Sigma0_VHVV_25Apr2017_dB.tif

I.1.4: Creating a multi-date, multi-polarization filtered radar image

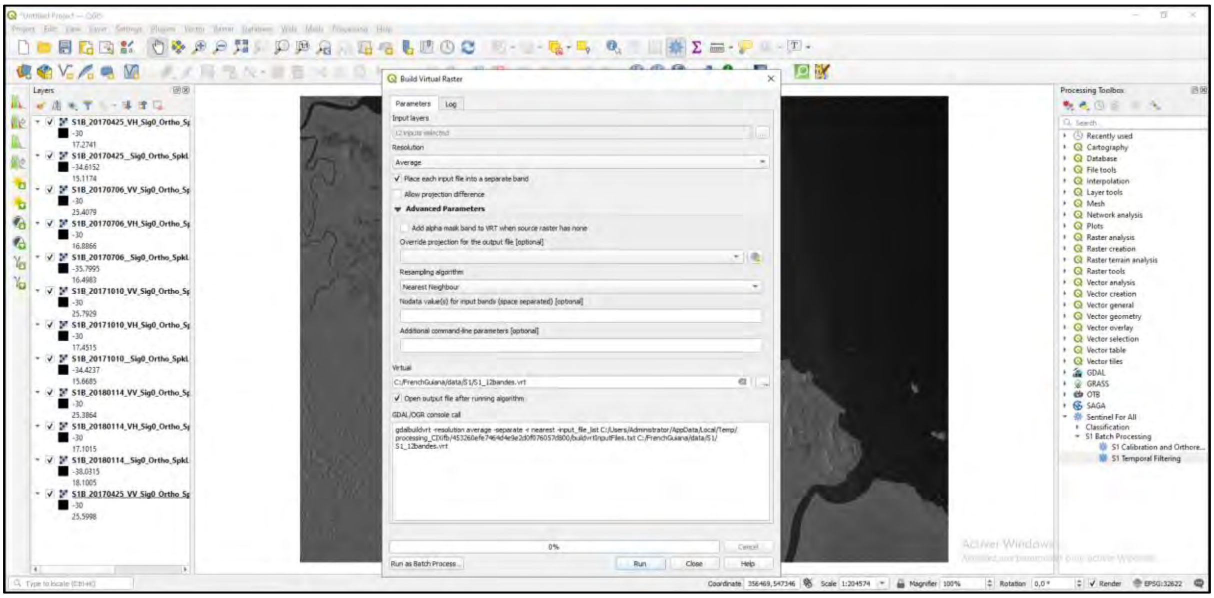
Finally, we need to create the image that will be used to classify the vegetation. This is a 12-bands speckle filtered image, consisting of 4 dates x 3 polarisations. We will create a virtual image:

1) Close all open layers in qgis.
2) In a file manager, go to the directory ***data\S1\ Filt_Spck\p47*** and type in the search bar at the top right ***S1*.tif***.



Select the 12 files and drag and drop them into QGIS

Then ***Raster → miscellaneous → Build Virtual Raster*** and fill in the fields as follow :



Input layers : Select the 12***.tif*** files

Virtual : ***data\S1\S1_12bandes.vrt***

Don't forget to check the box ***place each input file in a separate band.***

You have just created a Sentinel 1 virtual radar image containing 12 bands which you will use for vegetation classification.

I. 2. SENTINEL-2

The Sentine -2 data
data\S2\zip\S2B_MSIL1C_20180206T140049_N0206_R067_T22NCL_20180206T172306.zip
has been doanload from https://earthexplorer.usgs.gov/. It corresponds to the level 1C product. These products are already orthorectified and contain the following 13 spectral bands

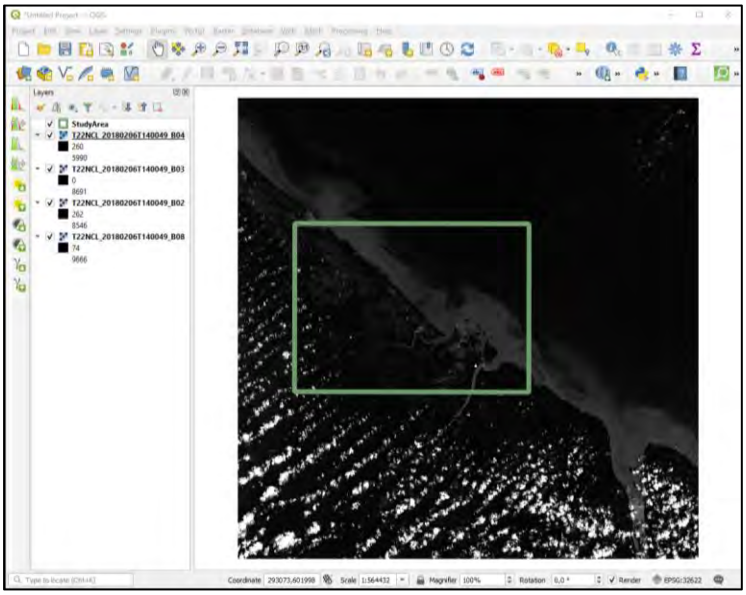
Band	**Band range (nm) Band Center (nm)**	**Spatial resolution (m)**	**Purpose in L2 processing context**
B1	433-453 / 443	20	Atmospheric Correction
B2	458-523 / 490	10	Blue-Sensitive to Vegetation Aerosol Scattering
B3	543-578 / 560	10	Gree-Green peak, sensitive to total chlorophyll in vegetation
B4	650-680 / 665	10	Red-Max chlorophyll absorption
B5	698-713 / 705	20	Vegetation classification
B6	734-748 / 740	20	Vegetation classification
B7	765-785 / 783	20	Vegetation classification
B8	785-900 / 842	10	Vegetation classification
B8a	855-875 / 865	20	NIR-Used for water vapour absorption reference
B9	930-950 / 945	60	Water vapour absorption atmospheric correction
B10	1365-1385 / 1375	60	Detection of thin cirrus for atmospheric correction
B11	1565-1655 / 1610	20	MIR-Snow / ice / cloud discrimination / Soils detection
B12	2100-2280 / 2190	20	AOT determination

We will only process the 4 bands at 10m spatial resolution (B2-4 +B8) to which we will add an NDVI channel.

1.2.1 Creating a 5 bands image over the study area

a) Loading the images into QGIS

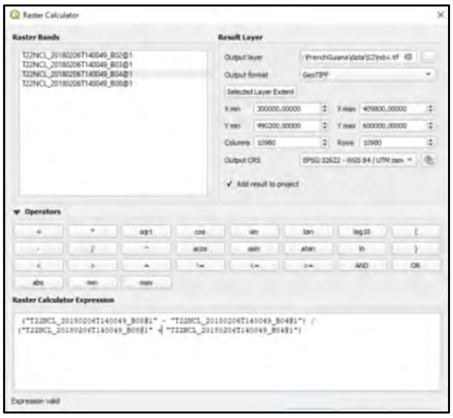
Décompress the file
S2B_MSIL1C_20180206T140049_N0206_R067_T22NCL_20180206T172306.zip in the same folder
Then in a file manager, go to the repertory
data\S2\zip\S2B_MSIL1C_20180206T140049_N0206_R067_T22NCL_20180206T172306.SAFE\GRANULE\L1C_T22NCL_A004812_20180206T140045\IMG_DATA then drag and drop the 4 bands ***T22NCL_20180206T140049_BXX.jp2*** (***XX*** for ***02***, ***03***, ***04*** et ***08***) into QGIS.



The 4 Sentinel-2 bands loaded into QGIS as well as the polygon delimiting the study area. They are already in the geographic projection UTM zone 22 N - EPSG: 32622.

b) ***Creating the NDVI band (**$NDVI = \frac{\rho_{IR} - \rho_R}{\rho_{IR} + \rho_R}$*, where ρ_{IR} and ρ_R are the reflectances of Red and InfraRed channels respectively).

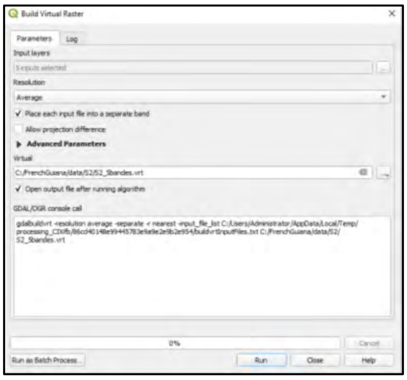
For this ***Raster → Raster Calculator*** then fill in the fields as folow :



In the ***output layer*** field specify : ***data\S2\ndvi.tif***

c) Creating a 5 bands image (B2, B3, B4, B8 and NDVI) cropped over the study area

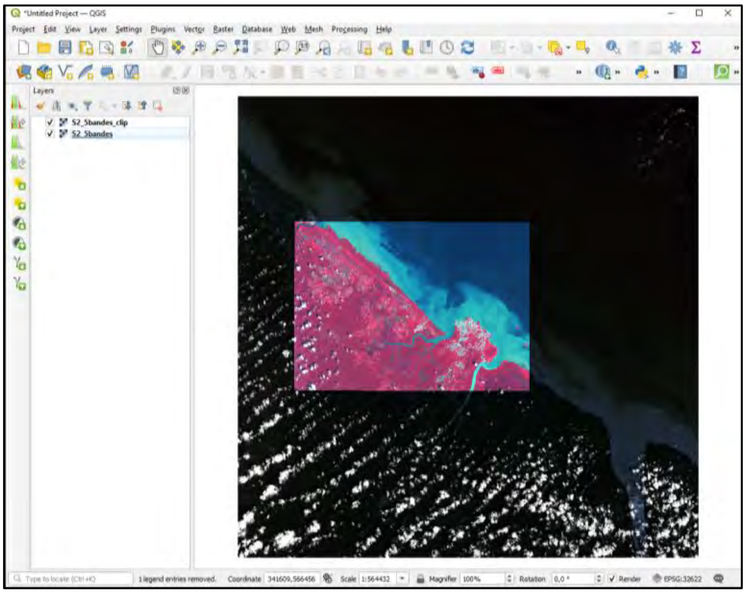
Raster → miscellaneous → Bulld Virtual Raster



Store this virtual imagein ***data\S2\S2_5bandes.vrt***

Then ***Raster → Extraction → Clip Raster by Mask Layer*** and fill in the fields as follow :

In ***clipped*** (the output image) specify ***data\S2\S2_5bandes_clip.tif***



Color composite of the S2_5bandes_clip.tif image overlapped to the original Sentinel-2 scene.

II. Classifications of the different types of vegetation

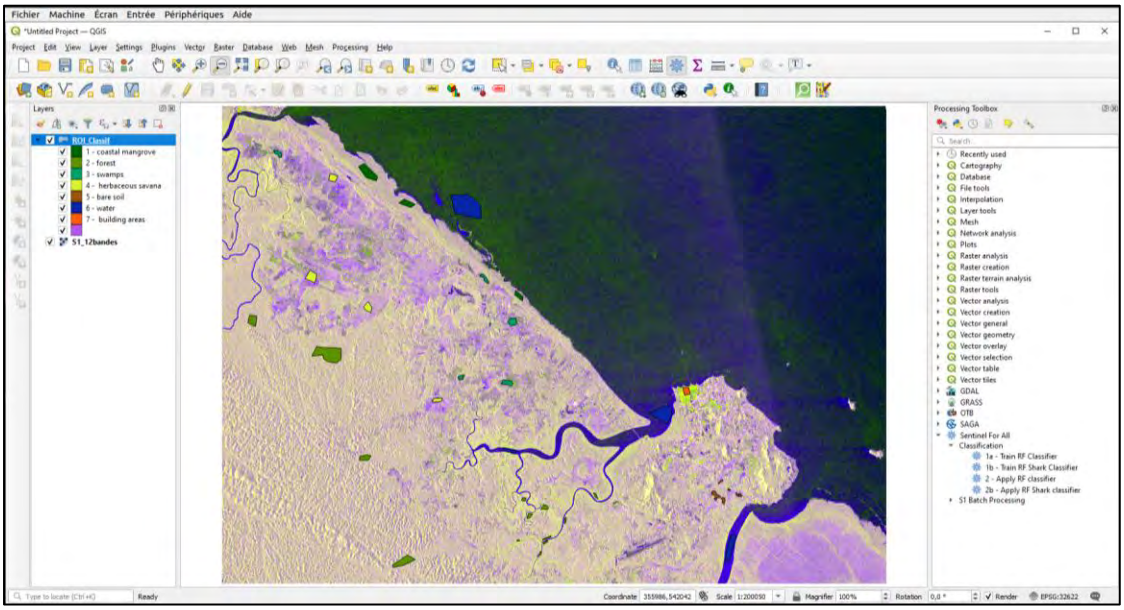
The Sentinel-1 and Sentinel-2 data are now ready to be used in a classification algorithm. We used the RANDOM FOREST supervised algorithm [3,4], as it gives overall good results with limited computing time, even when a large number of bands are involved.

The training and validation polygons have been defined from *in situ* surveys, coupled to Google Earth visual interpretation as well as the Sentinel-1 temporal average image ***VVdB_VHdB_VHdB-VVdB_Temporal_Average_20170425_20180114*** (cf. § I. 1.1).

II.1. classification of the Sentinel-1 data

Within QGIS, load the ***data\S1\S1_12bandes.vrt*** image

Load as well the layer ***classification\vector\ROI_Classif.shp*** containing the training / validation polygons. You can see the 7 vegetation types that we are trying to map.



a) Training the model

In the ***Processing*** panel : ***Sentinel for All → Classification → Train RF Classifier***

[3] https://fr.wikipedia.org/wiki/Forêt_d’arbres_décisionnels
[4] https://en.wikipedia.org/wiki/Random_forest

The output model is a text file that you will name ***classification\model_S1_12bandes.txt***

You will analyse in the *Log* tab of the *Train RF classifier* window the statistics describing the performance of each class :

$$Precision = \frac{Number\ of\ pixels\ correctly\ affected\ to\ the\ class}{Total\ number\ of\ pixels\ affected\ to\ the\ class}$$

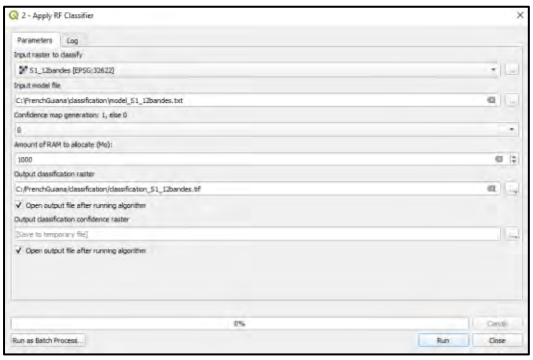
$$Recall = \frac{Number\ of\ pixels\ correctly\ affected\ to\ the\ class}{Total\ Number\ of\ effective\ pixels\ of\ the\ class}$$

$$F - score = \left(\frac{Precision^{-1} + Recall^{-1}}{2}\right)^{-1}$$

The lowest statistics (around 78%) are observed for the Accuracy of class 1 (Mangroves) and the Recall of class 2 (Forests), leading to an F-score of around 85%. If the classification were to be improved, the training polygons for these two classes would have to be refined. The highest statistics are, as is often the case with radar, those of classes 6 (Water) and 7 (Built-up areas) with F-scores of 99%. Generally speaking, (this is obviously relative and depends on the expectations of the problem) we can estimate that the results are not acceptable (training polygons not representative of the class sought) when one of the statistics < 70%.

b) Application of the model to the whole image

The model ***classification\model_S1_12bandes.txt*** was created from the training polygons, so we will apply it to the whole image. For this, ***Sentinel for All → Classification → Apply RF Classifier***



The output classification image is the file ***classification/classification_S1_12bandes.tif***.

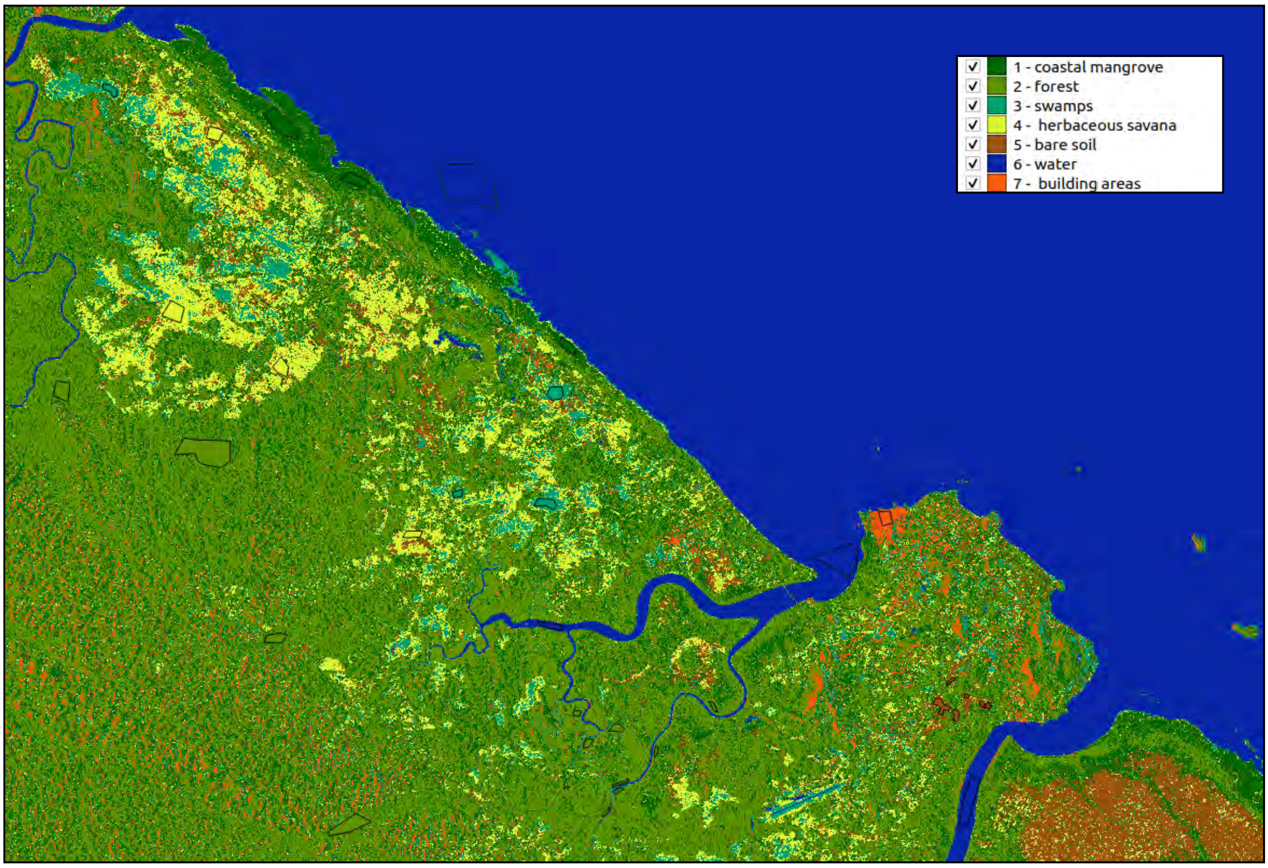
Loa dit into QGIS.

Chargez le dans QGIS. It remains to apply the same colours to the classes as those of the training polygons. To do this, right click on the ***classification_S1_12bandes***, then ***Propriétés*** and click on the *Symbology* tab in the left-hand panel. Click on the *Style* button at the bottom left of the window then load the style ***classification/colours_classif_7cl_tif.qml***.

Comment :

We can remove the isolated points using ***Raster → Analysis → Sieve***... that we apply to the image ***classification_S1_12bandes*** with a threshold value of *20*, and by using ***8-connectedness***.

You will notice that this classification is far from perfect. There are many pixels of the forest class that are classified as mangroves. It is therefore necessary to refine the choice of mangrove polygons and redo the classification model. You can also see the effects of terrain, as the signal is high into the 3 polarisations for the slopes facing the radar. This is the reason why they are classified as built-up areas. This effect can be remedied by applying a slope mask.

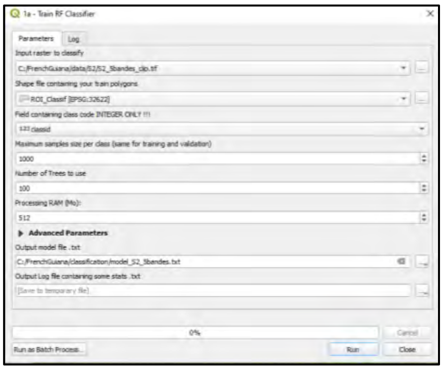


Classification result od Sentinel-1 images (4 dates x 3 polarizations).

II.2 classification of Sentinel-2 data

a) Training the model

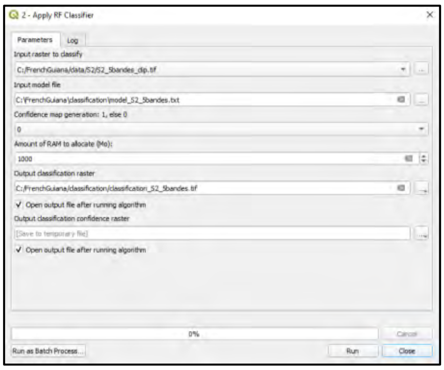
Same as step II.1.a) by replacing the file ***S1_12bandes.vrt*** by ***data/S2/S2_5bandes_clip.tif*** (cf § I.2.1) and by specifying for the output file **classification/model_S2_5bandes.txt**



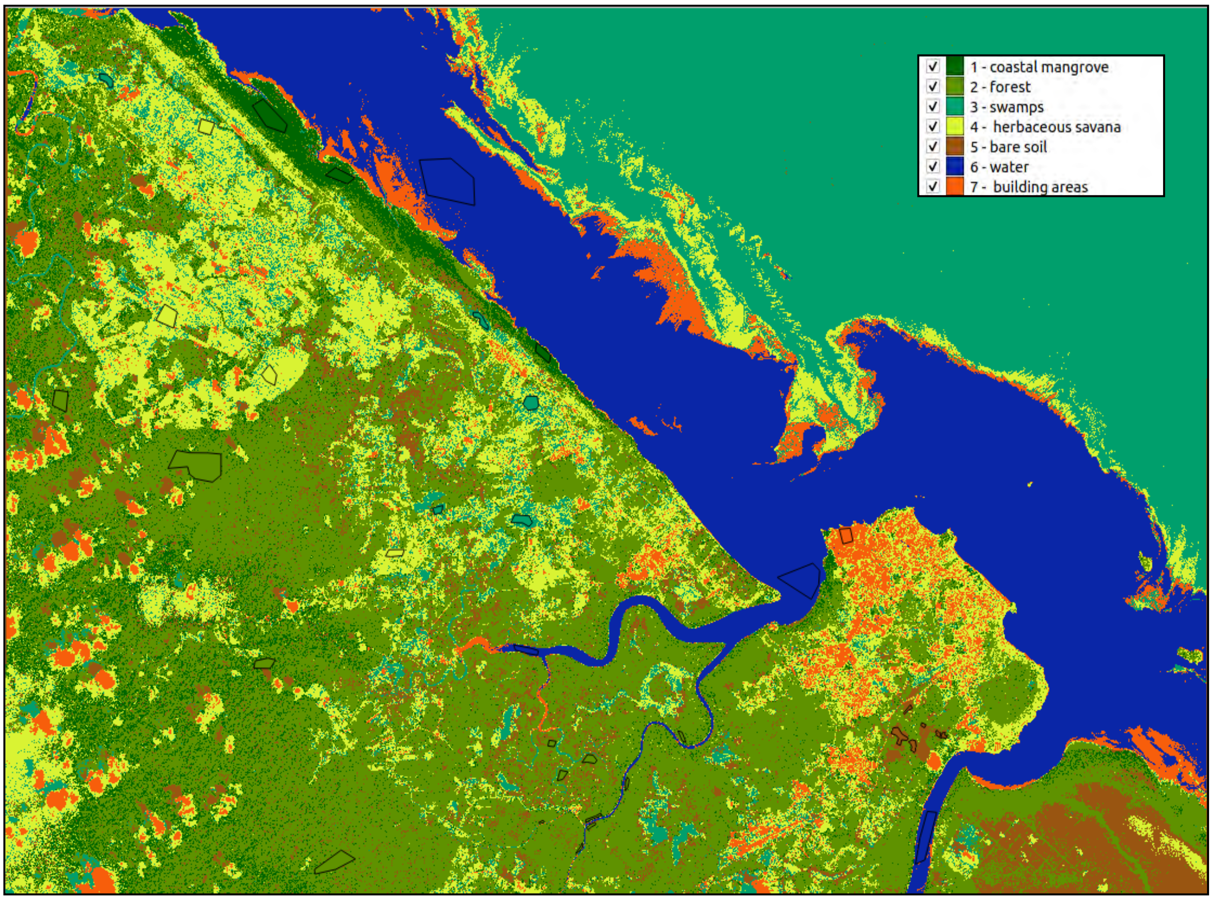
You can note the higher performance statistics (> 80%) than those obtained during the creation of the Sentinel-1 model ***model_S1_12bandes.txt***).

b) Application of the model to the image

Same as step II.1.b) by replacing the input file *S1_12bandes.vrt* by ***data/S2/S2_5bandes_clip.tif***, the classification model *model_S1_12bandes.txt* by ***model_S2_5bandes.txt***, and by specifying for the ouput file **classification/classification_S2_5bandes.tif**



The result (below) clearly shows the differences between the optical and radar data. The effects of terrain on the Sentinel-2 classification are not very noticeable. It also allows better discrimination between mangrove and tropical forest. But.... Obviously, clouds and their shadows have a direct and very disturbing effect on the classification.



Result of the classification from Sentinel-2 data.

III Classification merging Sentinel-1 and Sentinel-2 :

You can repeat the process to create a file containing the 12 (filtered) Sentinel-1 bands and the 5 Sentinel-2 bands. You will have to build it from the 17 bands separately (unfortunately you cannot create this file from *S1_12bands.vrt* and *S2_5bands.tif*).

Vous pouvez recommencer les manipulations pour créer un fichier contenant les 12 bandes Sentinel-1 (filtrées) et les 5 bandes Sentinel-2. Il faudra le construire a partir des bandes de 17 bandes séparément (vous ne pouvez malheureusement pas créer ce fichier à partir de *S1_12bandes.vrt* et *S2_5bandes.tif*.

It's up to you !!!!!