Mangrove degradation monitoring training (I) - The Continuous Change Detection and Classification (CCDC) algorithm.

UN-REDD Programme Mangrove Initiative

Introduction

The objective of these training series is the creation of a land cover map for 2020 and a land cover change map between 2015 and 2020 which identifies forest degradation, with special focus on mangroves. Mangroves are very dynamic ecoregions, affected by tidal inundations; therefore, satellite time series analysis are suitable tools for monitoring changes, in concrete we will use the Continuous Change Detection and Classification (CCDC) algorithm.

In this first training we will create a time series of Landsat-8 images for the Ayeyarwady region of Myanmar, identify the land cover change periods and calculate the models of the stable time periods for selected spectral bands and/or indices. Later, the parameters of the models corresponding to the beginning and end dates of our study period will be used to classify land cover maps and create a land cover change map.

For this we will use two new modules in SEPAL: '*Create a CCDC asset from a time-series*', and '*Create a slice of a CCDC asset for a specific date*'.

Background – The CCDC algorithm

The Continuous Change Detection and Classification (CCDC) algorithm, developed at Boston University (Zhu and Woodcock, 2014) is aimed at analysing time series. At each pixel it uses all available satellite observations for a period to simultaneously map land cover and land cover change. To do this, it follows two main steps (Boston University, n.d.):

1. Identification of change points and modelling of stable time segments.

First, the time series is filtered to remove satellite observations affected by clouds, cloud shadows and snow. The resulting time series is then modelled as a Fourier series starting in the first year. Using this initial model, subsequent observations at each pixel are then successively compared against model forecasts, and if there are persistent mismatches between new observations and the model, change points are detected. When no change is detected, new data are appended to the time series and the model is re-fit.

The change points are identified based on model fits across spectral bands and/or indices' values using a change vector metric that integrates differences between observed and predicted reflectance in each spectral band and/or index, weighted by the root mean squared error of the model.

2. Assignment of class labels to each time segment.

Once change points are identified, spectral and temporal information for each stable time segment is used to assign land cover and land use labels to each pixel. To perform this, CCDC uses a supervised classification approach based on Random forest that relies on training data.

This approach has two important advantages:

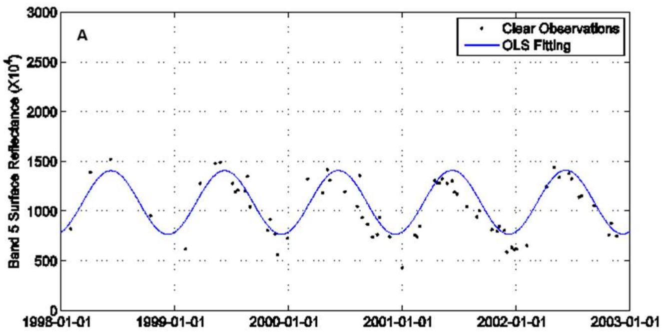
First, data from the entire time segment contributes to the classification of a pixel. Instead of using time series of surface reflectance as the primary inputs, CCDC uses model parameters estimated for each spectral band or index, which has been shown to be highly effective for discrimination of land cover and land use classes.

Second, this approach only estimates a new classification label after a change has been detected, which avoids problems associated with stochastic changes in classification labels that arise when land cover is mapped based on independent time series from each year.

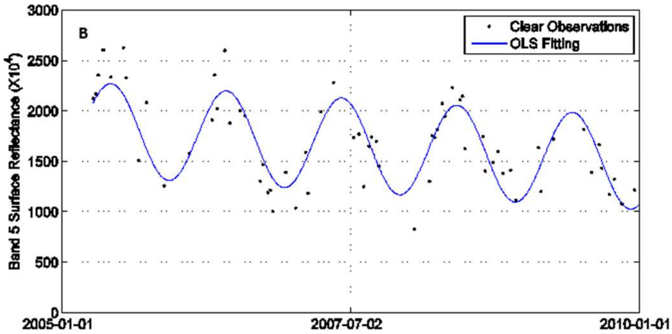
Breakpoints and time segments

Pixel values can fluctuate for several reasons: seasonal (intra-annual), gradual (inter-annual) and abrupt changes (breaks). The CCDC algorithm identifies all three type of changes with the primary purpose of identifying abrupt changes. A time segment is the period between two consecutive land cover changes (if any), bounded by the start and end of the time series.

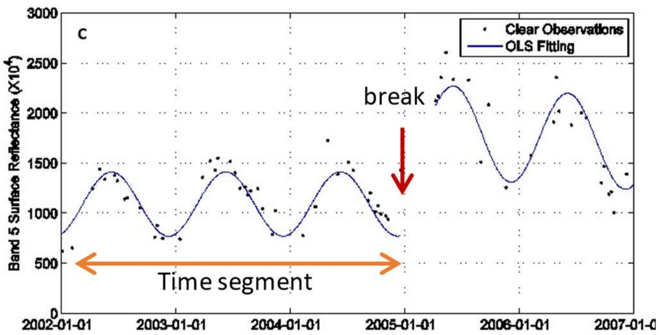
Seasonal



Trend



Break



Examples of type of changes in a time series' reflectance values

In the *'Create a CCDC asset from a time-series'* module in SEPAL we will create a Landsat-8 time series with its breakpoints and time segments according to selected spectral bands and/or indices.

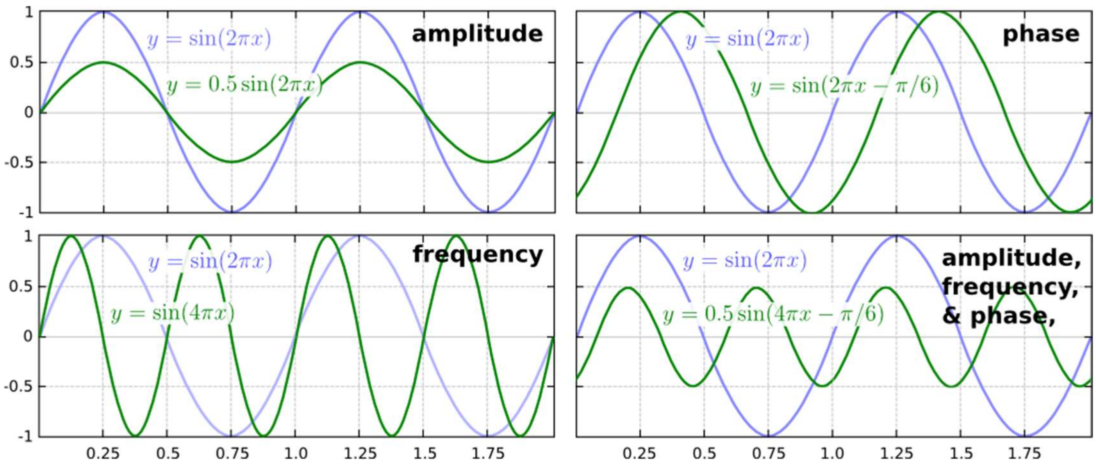
Model extraction of time segments for a particular date

For each time segment, CCDC fits a *harmonic regression model*, which is a linear regression model in which the predictor variables are trigonometric functions of the variable time. Harmonic regression is used in modelling biological phenomena, which tends to have periodic rhythms, like the vegetation greenness along the year.

A harmonic regression function is defined by the following parameters:

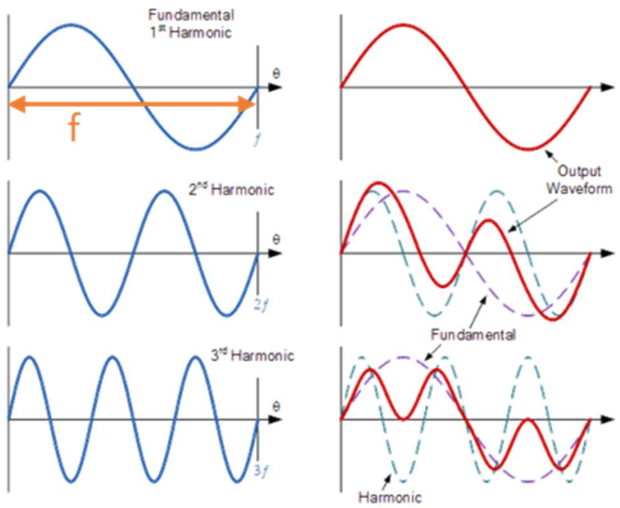
- The *amplitude* is the distance between the resting position and the maximum displacement of points on the wave, stated another way, the vertical distance between a peak or a valley and the equilibrium point.
- The *phase* is related to the distance between the position of the amplitude crest/trough of the wave and the amplitude crest/trough of a wave which starts at t=0 in the equilibrium position.
- The *frequency* is the number wave cycles passing a point per unit time.

The amplitude depends on the intensity of the seasonality and the phase on the location of the seasons along the year (e.g., in which month the humid/dry season starts).



Representation of the parameters of a harmonic regression function

The harmonic regression model can have different *orders* of 'complexity', from 0 order, which is a linear model (in SEPAL from 1st to 3rd order). The 1st order or fundamental is the lowest or base frequency (ƒ), on which the complex waveform is built and as such the periodic time. The shape of the resulting complex waveform will depend on the number and amplitude of the harmonic frequencies present, and on the phase relationship between the fundamental or base frequency and the individual harmonic frequencies. We can see that a complex wave is made up of a fundamental waveform plus harmonics, each with its own peak value and phase angle.



$$Y_1 = A_1(2\pi \mathrm{ft} - \mathrm{P}_1)$$

$$Y_2 = A_2(2\pi \mathrm{ft} - \mathrm{P}_2)$$

$$Y_3 = A_3(2\pi \mathrm{ft} - \mathrm{P}_3)$$

$$Y_\mathrm{T} = \mathrm{Int} + \mathrm{slope}\ (\mathrm{Y}_1 + \mathrm{Y}_2 + \mathrm{Y}_3)$$

Representation of the harmonic regression models of 1st, 2nd and 3rd order

A 1st order harmonic can fit well the satellite observations in our region since it is related to the seasonality and contains most of the information. However, the 3rd order fits even better the observations in case of variations inside the year and season. A 3rd order harmonic regression model has 8 parameters (Y_T): interception, slope, 3 amplitudes (A_1, A_2 and A_3), 3 frequencies and 3 phases (P_1, P_2 and P_3).

In the '*Create a slice of a CCDC asset for a specific date'* module in SEPAL we can extract the coefficients of the models for each band and/or index selected, the RMSE and the value for that specific date.

Gap strategies for breaks

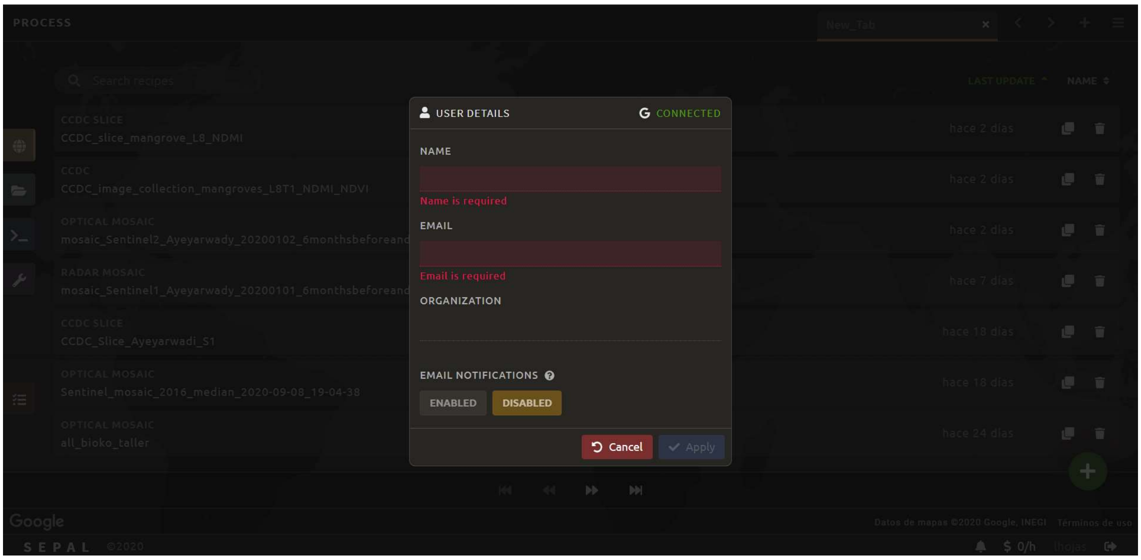
When extracting the parameters for a specific date, there will be some pixels without model because they fall into a break. In SEPAL there are three possible strategies to fill the model for these pixels:

- Interpolation: Generate a new model based on surrounding segments; a kind of mean between the time segment before and after the break. It is recommended when breaks are long, but for short periods the slope between both time segments can be very steep.
- Extrapolation: Generate a new model based on the time segment before or after the gap. Recommended for short breaks.
- Mask

Exercises in SEPAL and Google Earth Engine

Getting ready

Connect your SEPAL account to the Google account you used to create your Google Earth Engine account. You must see CONNECTED in green in the upper right corner of your screen.

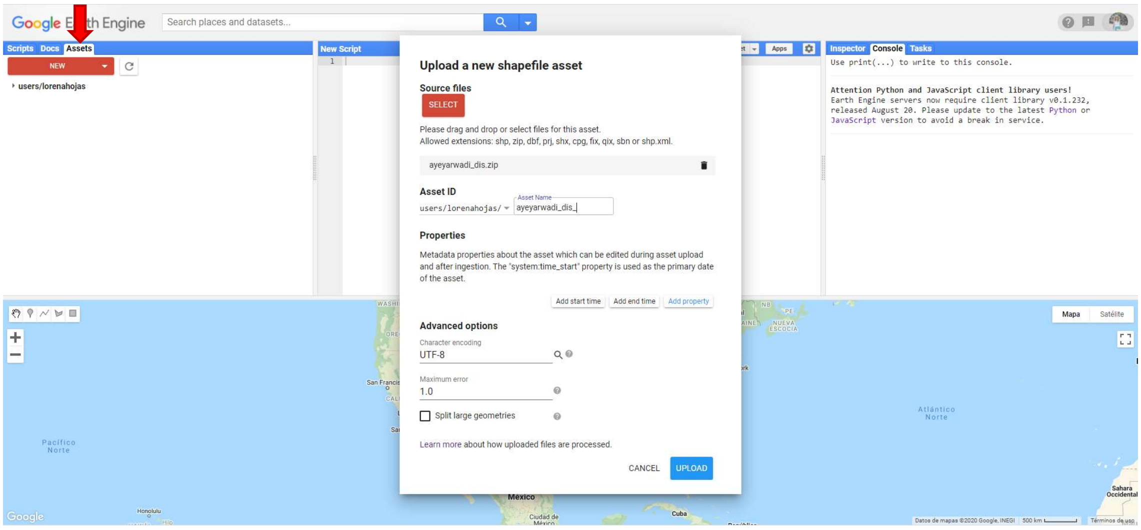


Upload a Shapefile to Google Earth Engine

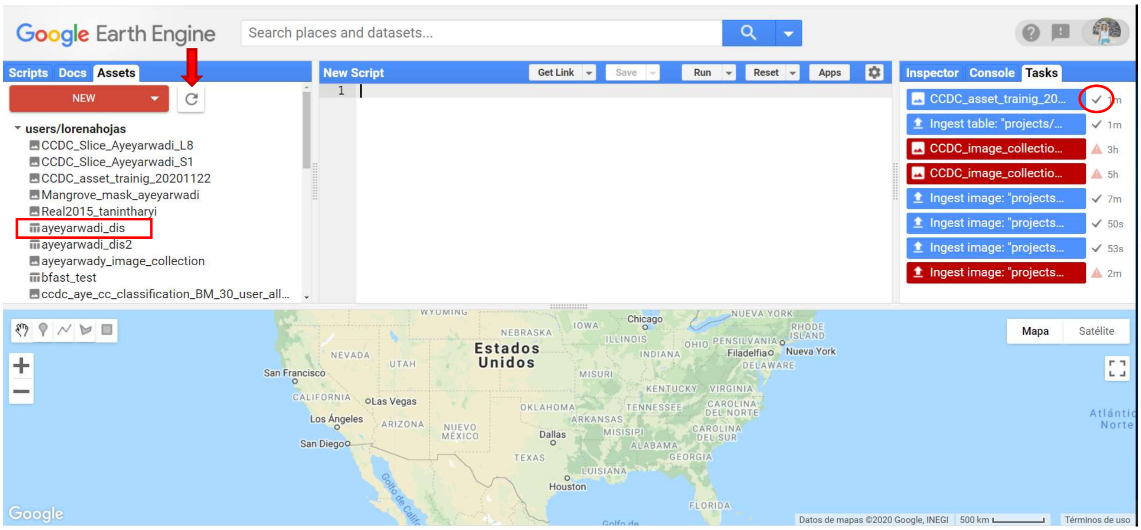
In the Code Editor go to Assets, click the NEW button, then select Shapefiles under the Table Upload section. An upload dialog like the figure below will be presented. Click the SELECT button and navigate to a Shapefile or Zip archive containing your Shapefile on your computer. When selecting a .shp file, be sure to select the related .dbf, .shx and .prj files. When uploading a Zip archive, make sure it contains one Shapefile (set of .shp, .dbf, .shx, prj, etc.). The shp, .dbf, and .shx files are required, the other sidecar files are optional.

Give the table an appropriate Asset ID (which does not already exist) in your user folder. Click UPLOAD to start the upload.

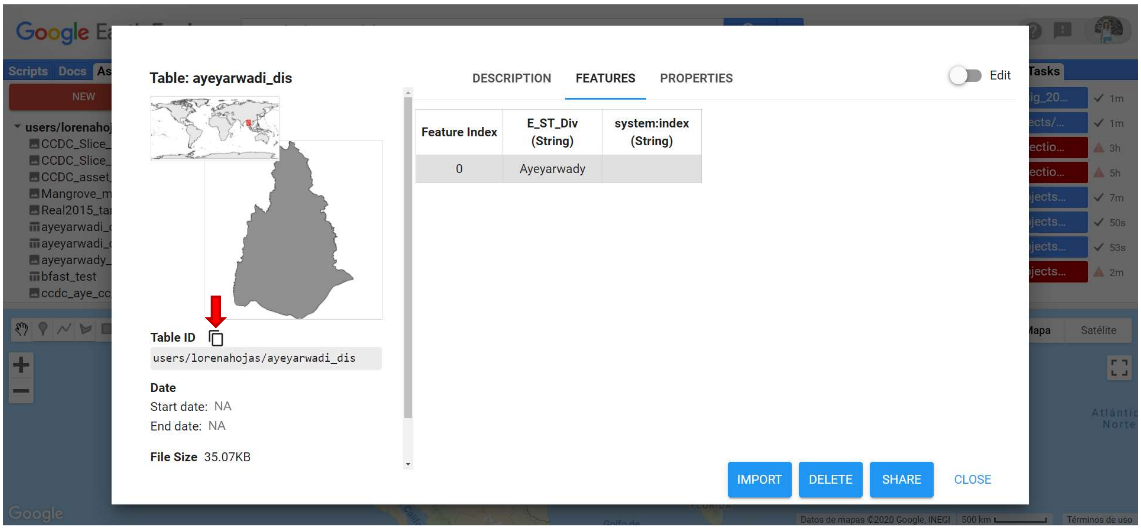
In this example we will upload the Shapefile with the boundary of Ayeyarwady region.



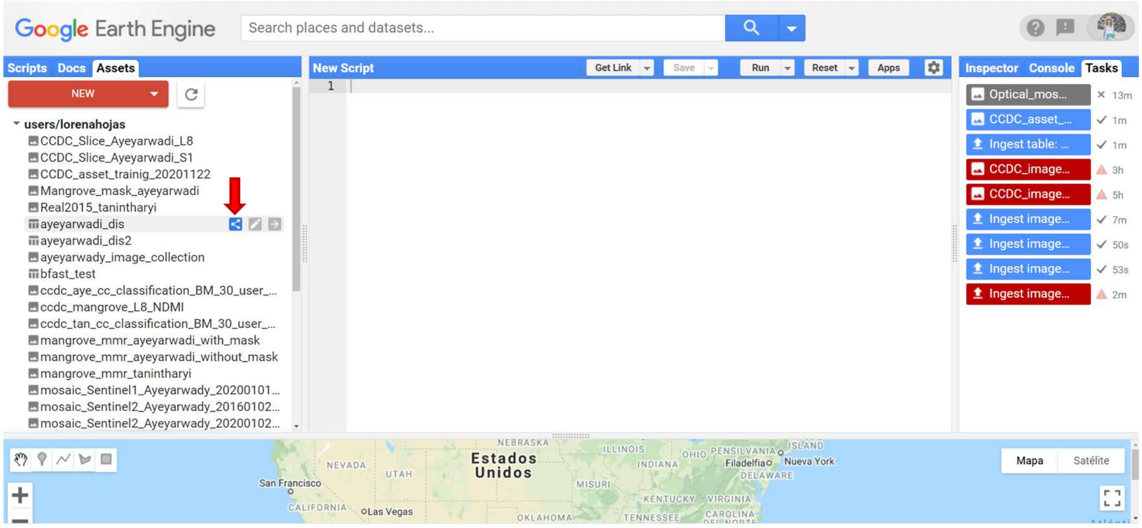
When the upload has finished, in the Tasks panel the check mark will appear next to the name of your asset. If you do not see it in your Assets panel, click on the REFRESH button.



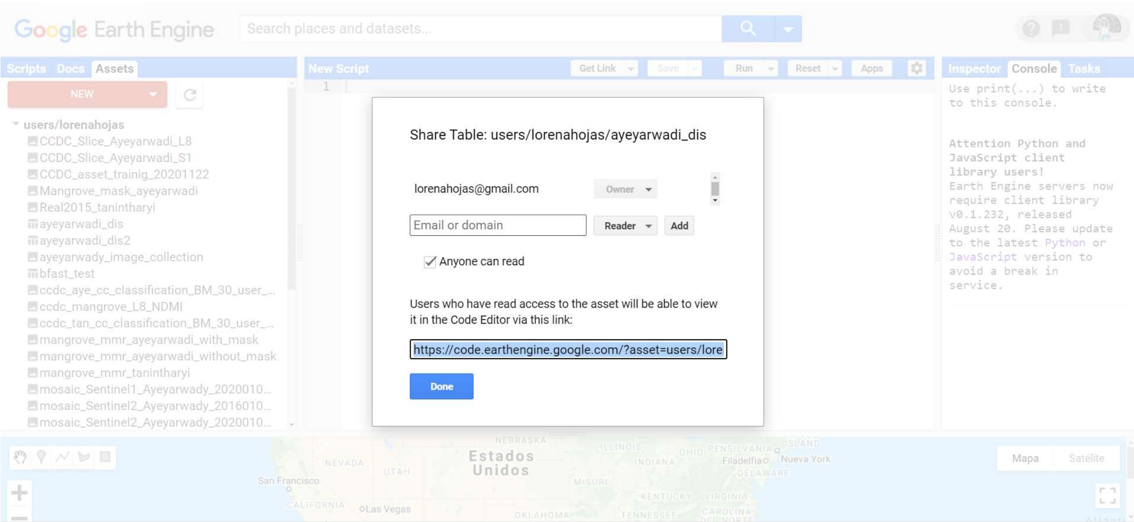
Click on your asset to see its properties. By clicking on the symbol next to Table ID you can copy the asset ID to your clipboard and paste it later in SEPAL. You can also see the columns and rows.



To share your table with other users, click on the SHARE button.



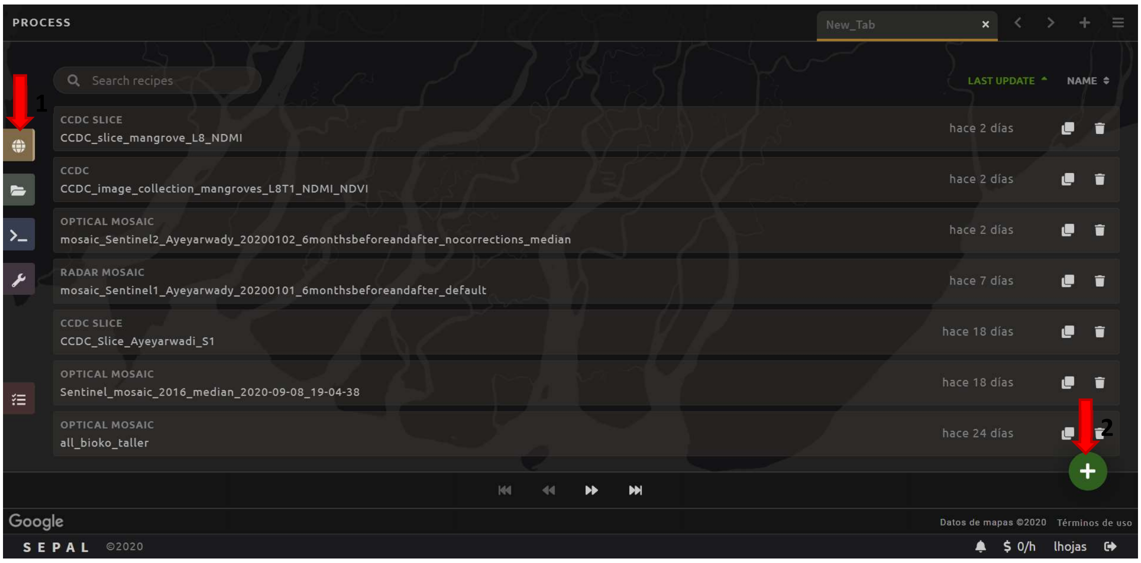
Click on the box ‘Anyone can read’ and press Done.



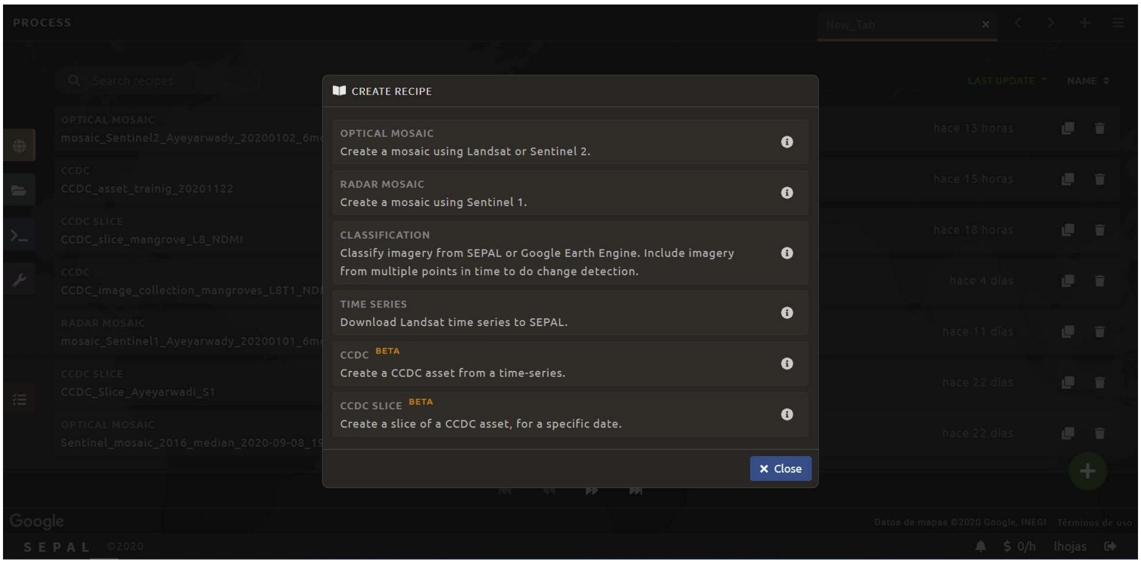
1. Creating Sentinel-2 mosaics in SEPAL

This section shows how the mosaics for the years 2016 and 2020 were created.

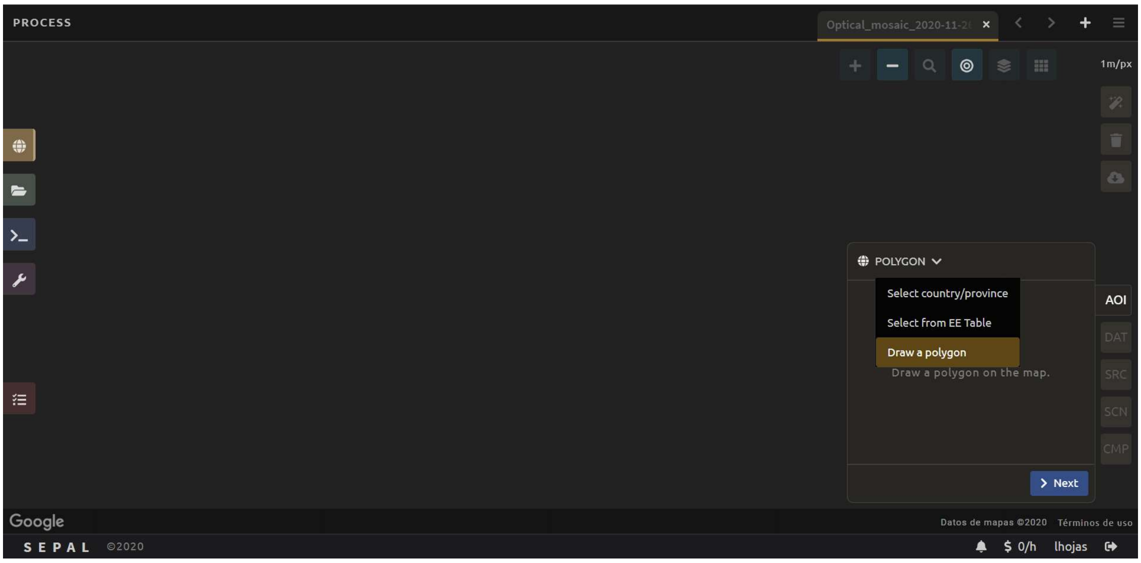
Go to process (1). Click on create new recipe (2).



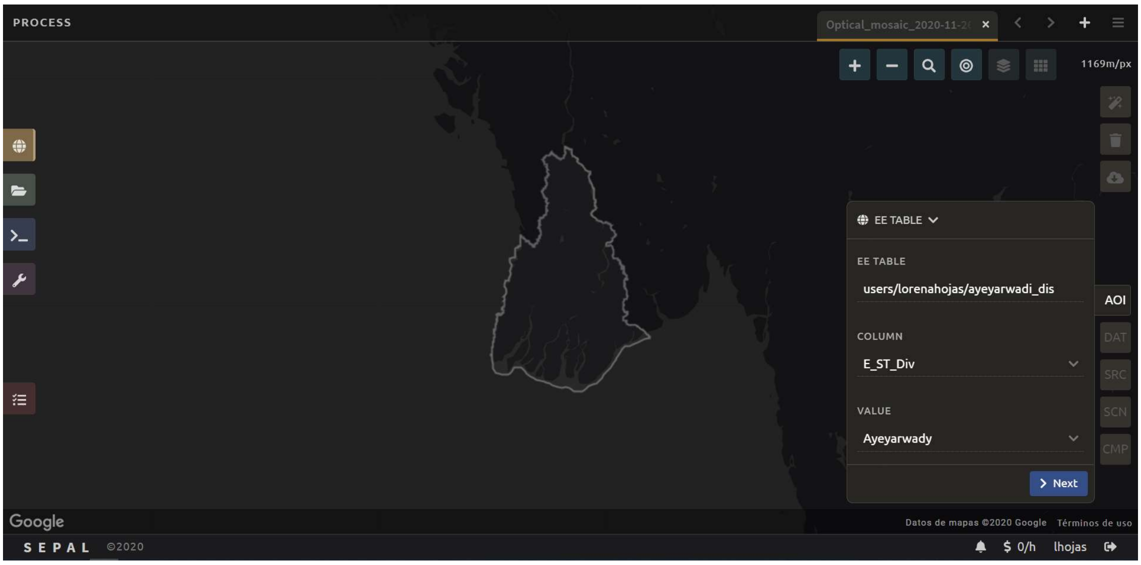
Select Create a mosaic using Landsat or Sentinel 2.



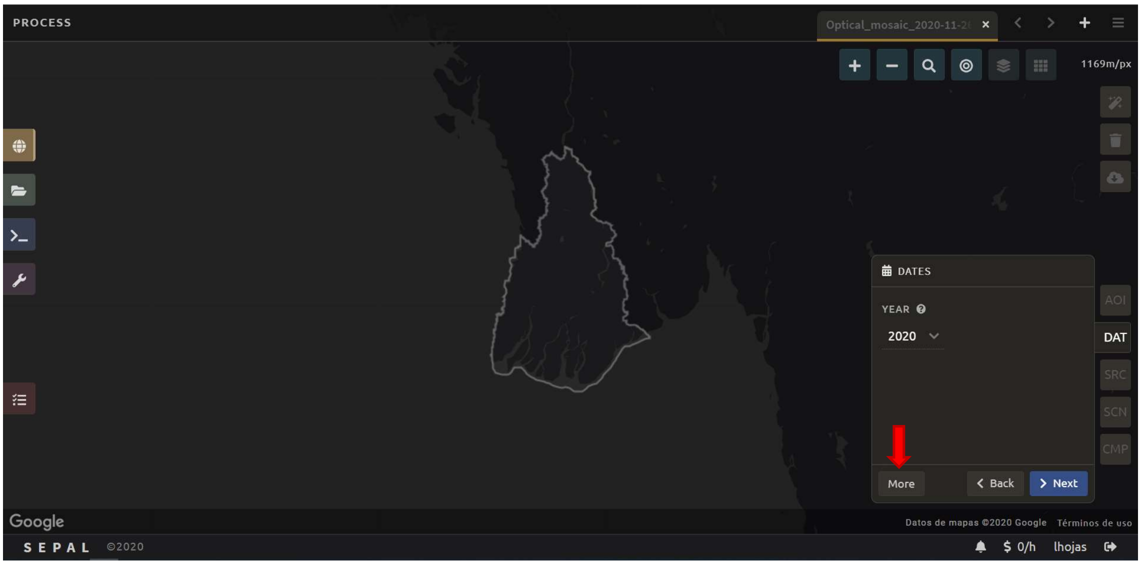
AOI: You can draw a polygon, select Myanmar from the country list or paste an EE Table (a Google Earth Engine Table)



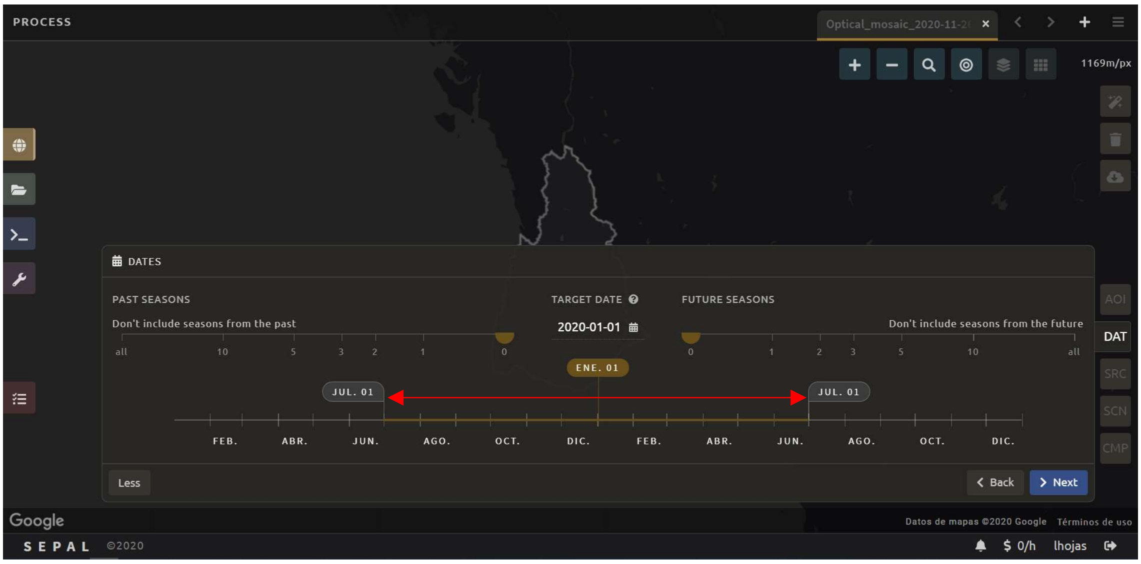
You can paste de table ID with the Ayeyarwady boundary from the previous section.



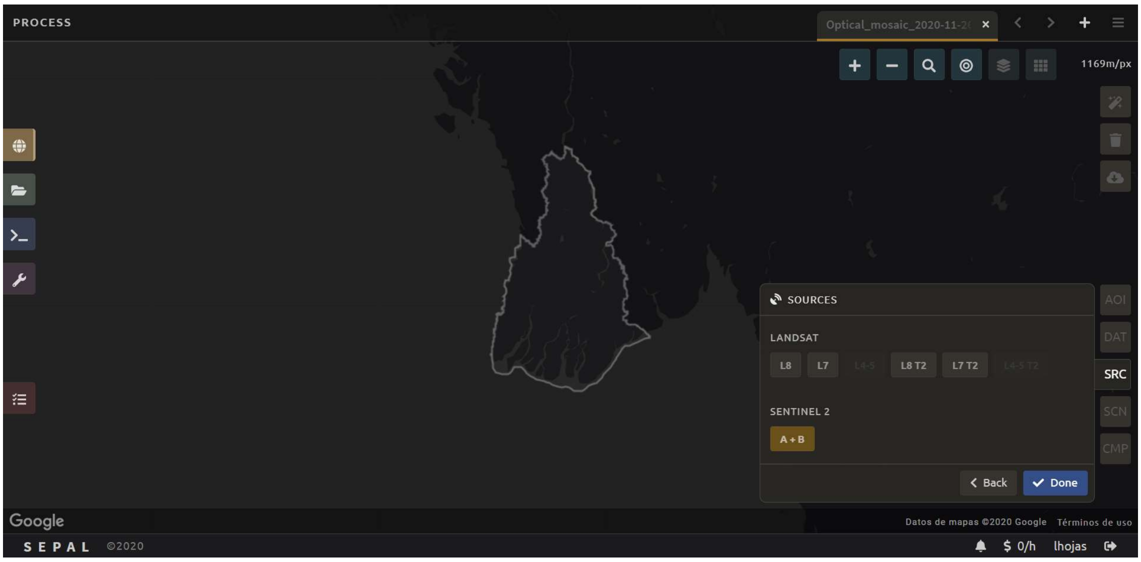
Date: select 2020 for the first mosaic (you can repeat the exercise selecting 2016). Click on more.



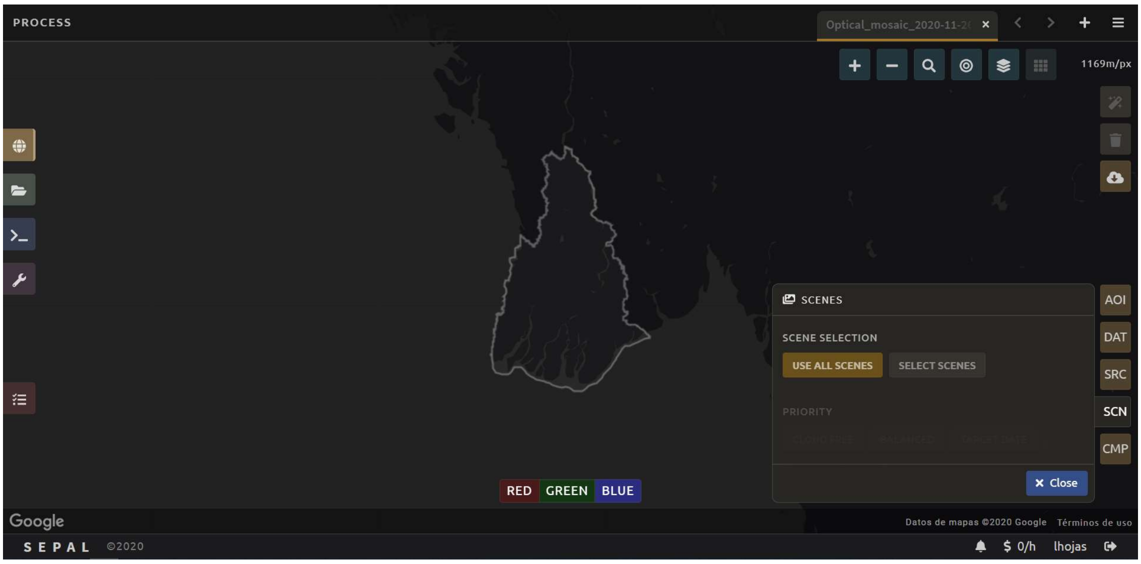
The target date is the date in which pixels in the mosaic should ideally come from, in both mosaics select the 1st of January. To get the full free-cloud cover of the region, images from 6 months before and after were considered. Alternatively, the period 4 month later + 2 months before, the dry season, also cover the full region.



Sources: Sentinel-2 A and B.



Scenes: all scenes were used.



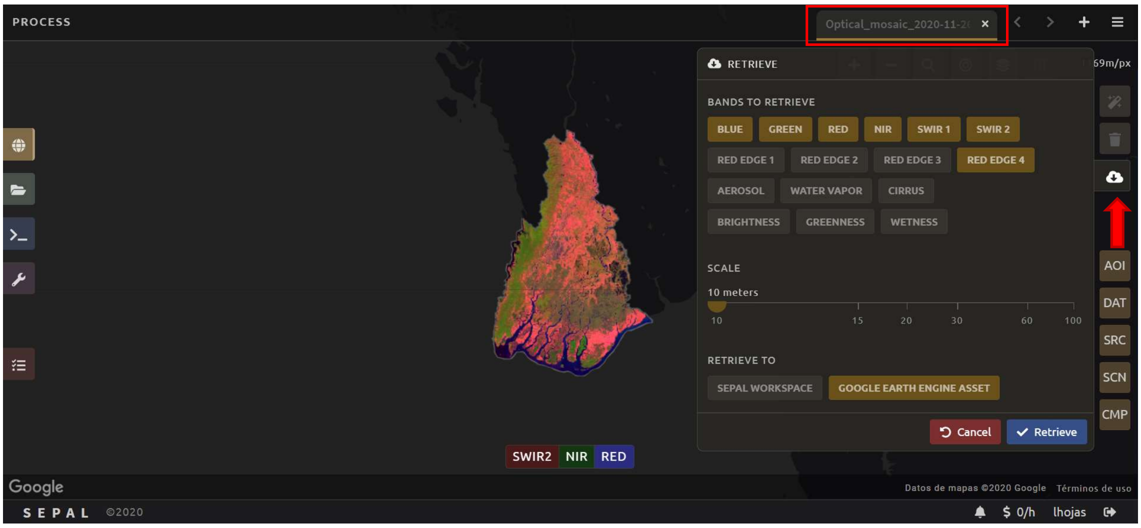
Composite options: the options in the image below were selected.

The scenes were not corrected to atmospheric surface reflectance (SR), neither the bidirectional reflectance distribution function (BRDF) effects.

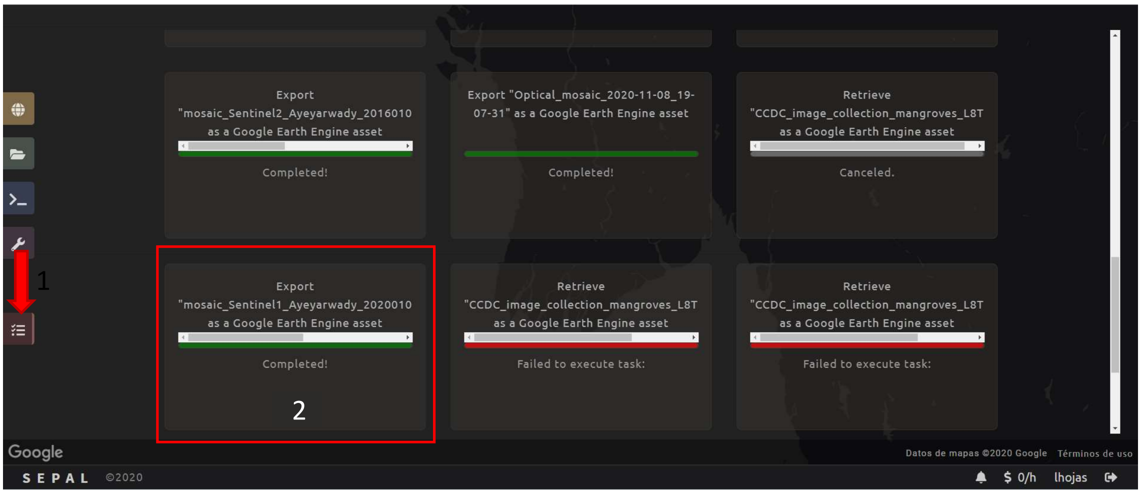
No pixel filter was applied. If the target day is very important, the filter based on day of the year could be used.

You can preview the mosaic with several band combinations.

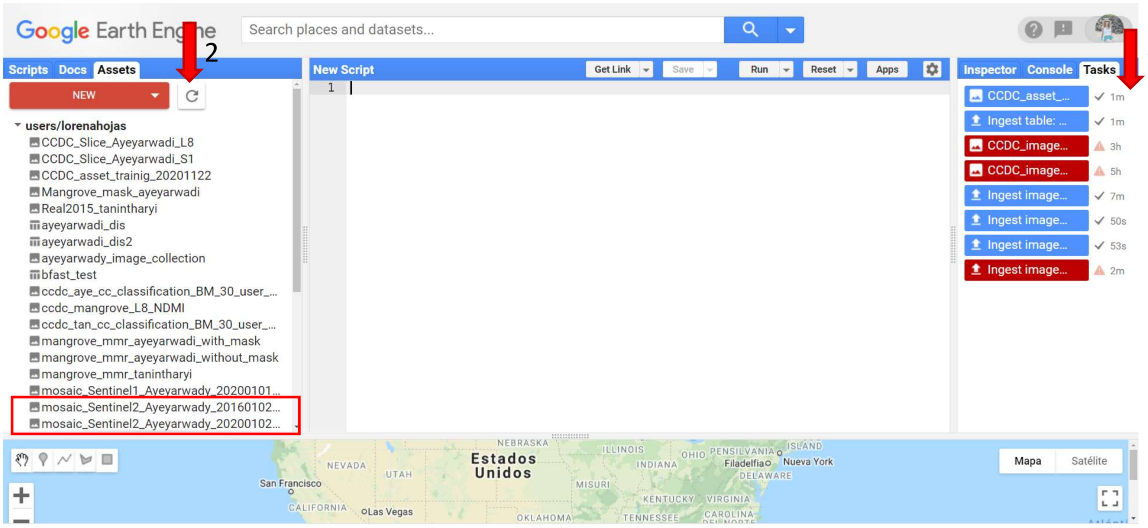
Select the bands you want to retrieve. Give a name to your recipe.



The task symbol will become a spinning wheel while your mosaic is being exported (1). When the mosaic has been exported into your Google Earth Engine assets successfully, the processing bar will appear in green, and below it will be written Completed! (2). The task symbol will appear again as in the picture below (1).



In the Tasks panel of the code editor of your Google Earth Engine account the check mark will appear next to your asset name (1). If you don't see it in the asset panel, click on refresh asset cache (2). The name of the asset will be the same you gave to your recipe.



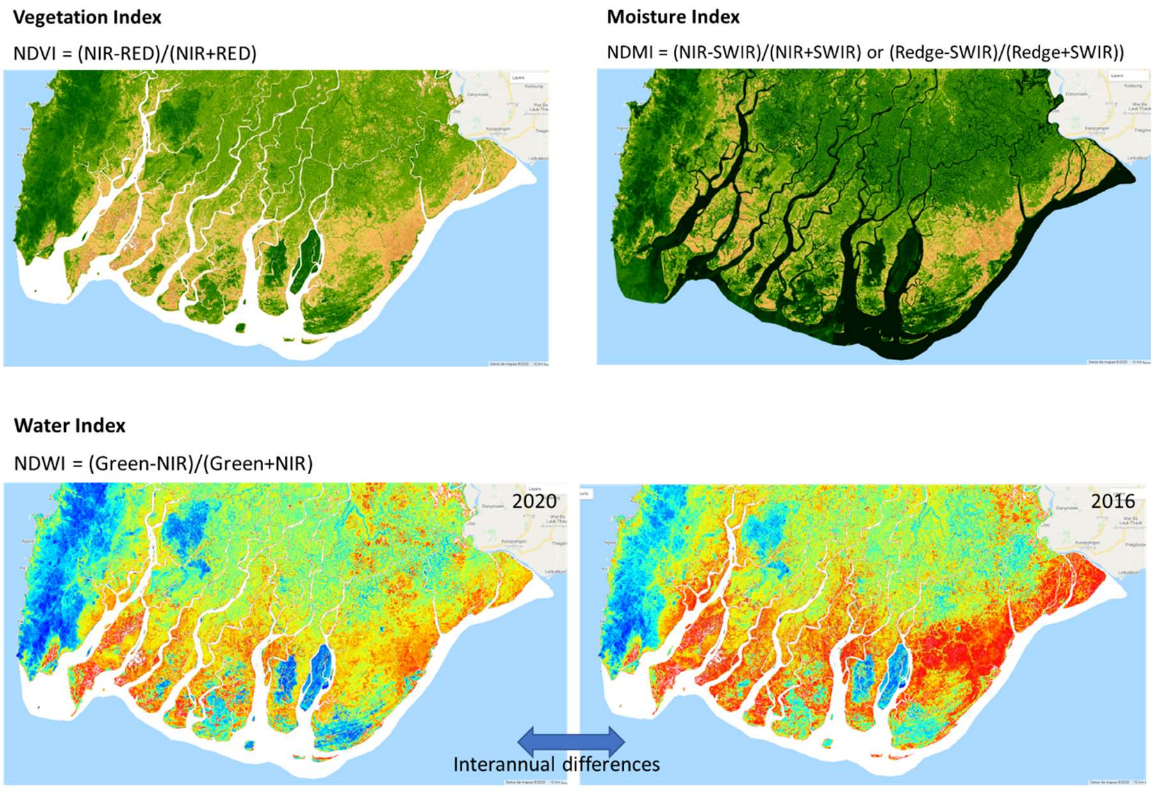
2. Calculate and visualize vegetation indices in Google Earth Engine

Spectral indices or combinations of them have been used to map mangroves and differentiate them from other types of evergreen tropical forest and permanent water bodies. For example, (Chen et al., 2017) use NDVI, which is good for monitoring canopy coverage (leaf area index), the moisture index (NDMI), sensitive to the water content in plants and its background, and the water index (NDWI), for monitoring permanent water bodies.

Copy and paste the following URL link to your address bar:
https://code.earthengine.google.com/358c22977b768839736fe7eb8d388f5f

The code that appears in the Google Earth Engine code editor allows you to:

1. Import and visualize the Sentinel-2 mosaics created in the previous exercise.
2. Calculate and visualize vegetation indices: NDVI, NDMI, NDWI, EVI…
3. Visualize additional data:
    - TPC project's Myanmar mangrove extent 2015
    - Global Mangrove Watch (2010)
    - NASA SRTM Digital Elevation Data (30m)
    - JRC Global Surface Water (30m)



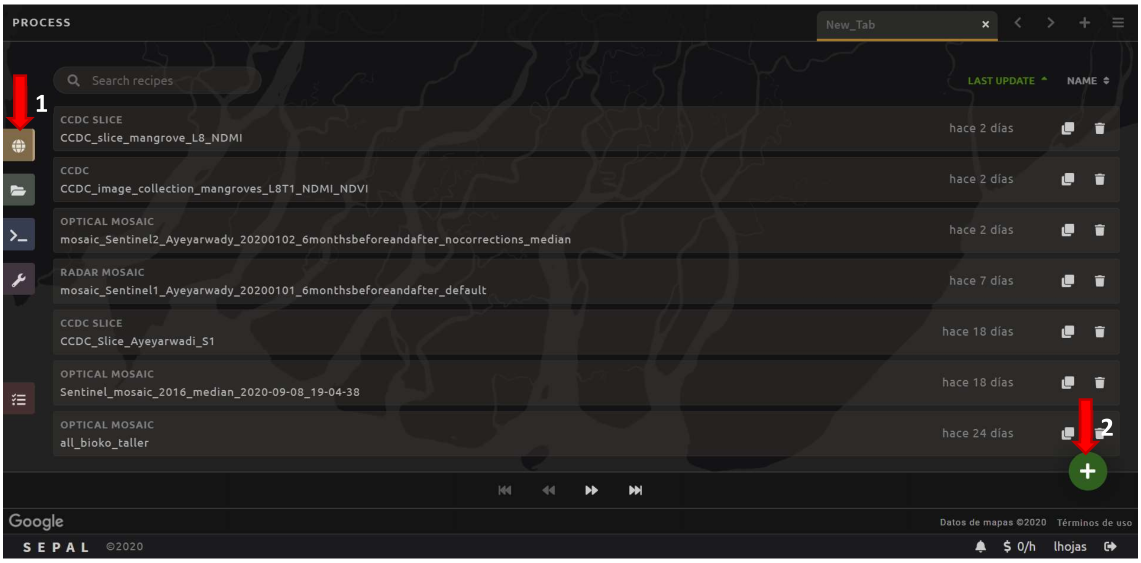
Example of visualization of the NDVI, NDMI and NDWI indices in Google Earth Engine

Exercise: Select areas of changes for study according to the Sentinel-2 mosaics of 2016 and 2020, which vegetation index/indices seem/s better for breakpoint detection and classification of mangrove? And for other type of vegetation (e.g., evergreen hardwood)?

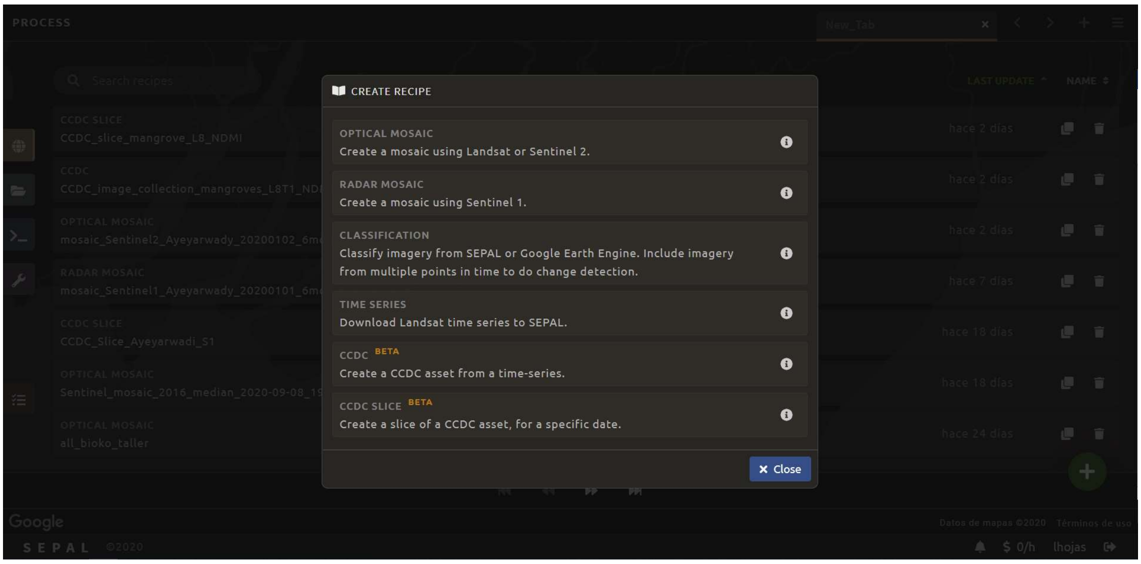
The mosaics target the same day (1st January) of different years. There are differences due to interannual changes (e.g., 2020 has been wetter than 2016). Therefore, time series analysis is better for mapping changes.

3. Creating a CCDC asset from a time series of Landsat-8 in SEPAL

Go to process (1). Click on create new recipe (2).



Select 'Create a CCDC asset from a time-series'.

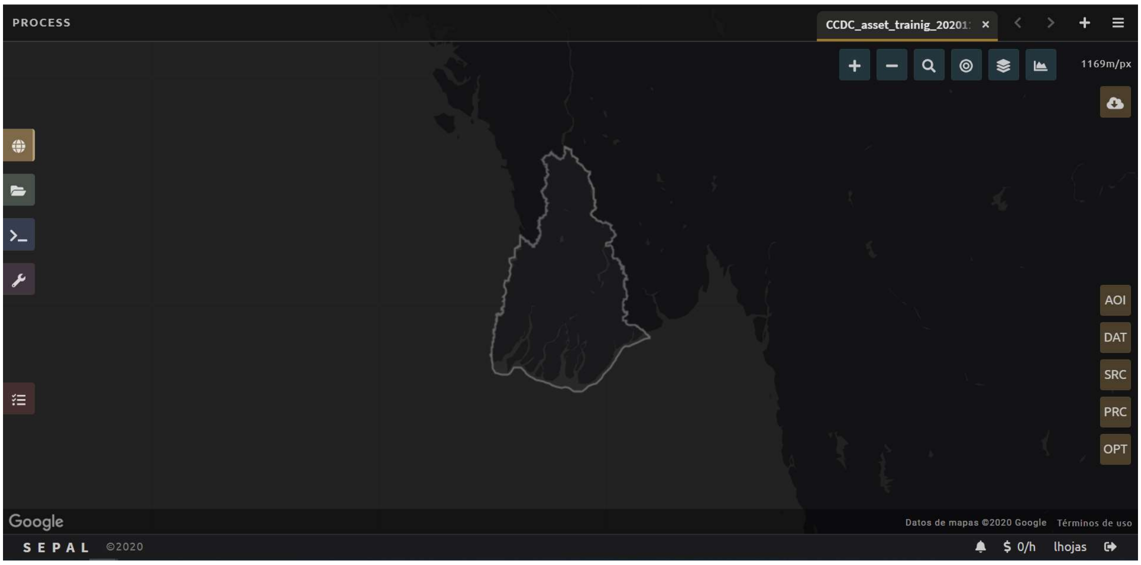


Options at the bottom right:

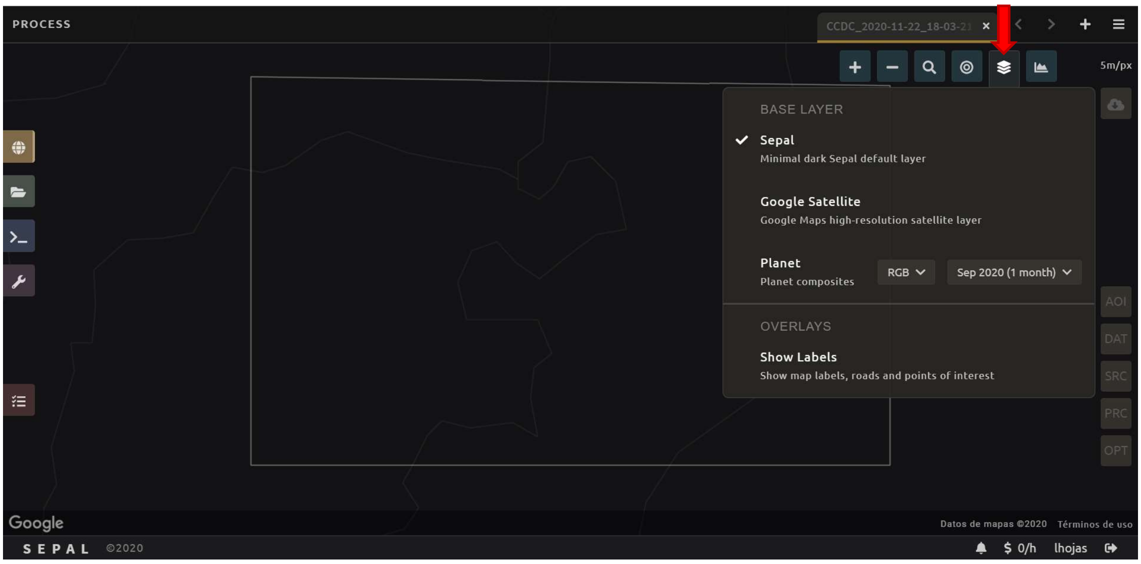
Area of interest: Select draw a polygon.

Right click at the vertices of your polygon (5 times for a rectangle, the last one on your first vertex to close it). If you are not satisfied, doble click on your polygon to delete it and create a new one.

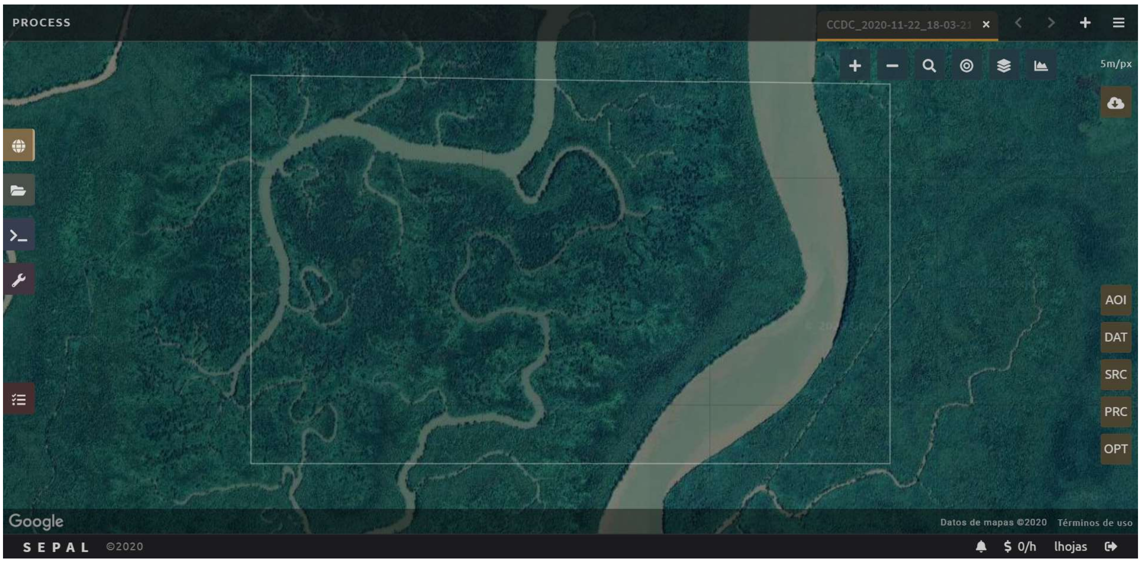
Another possibility is select from EEtable: Paste an ID table from Google Earth Engine. For example, for the region of Ayeyarwady you can paste *users/lorenahojas/ayeyarwadi_dis*.



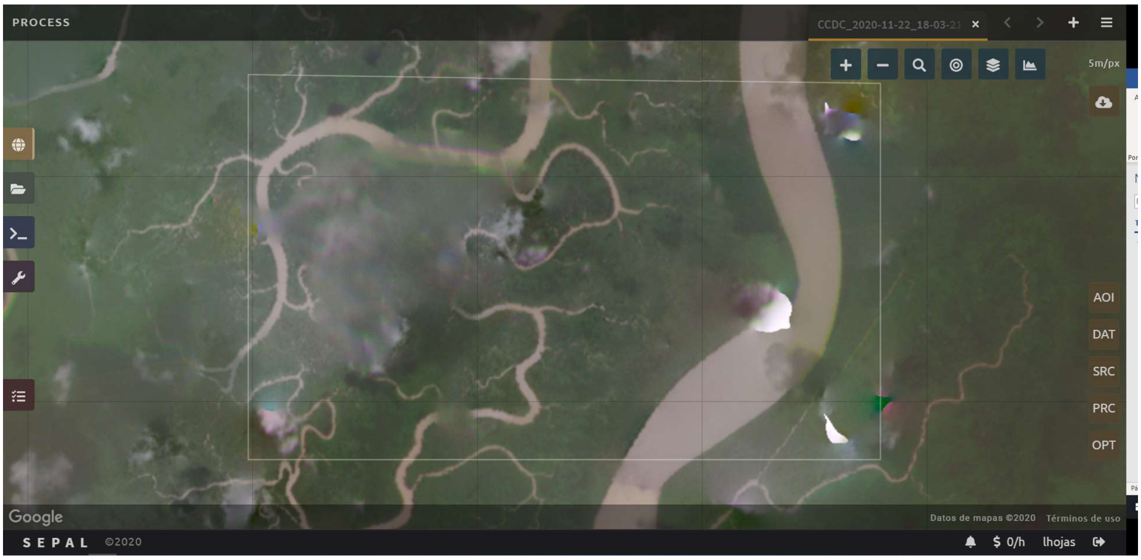
Base layer: apart from the SEPAL default layer you can select Google Maps high resolution satellite imagery or Planet composites (4,7m spatial resolution), in RGB or Color-Infrared.



Google Maps

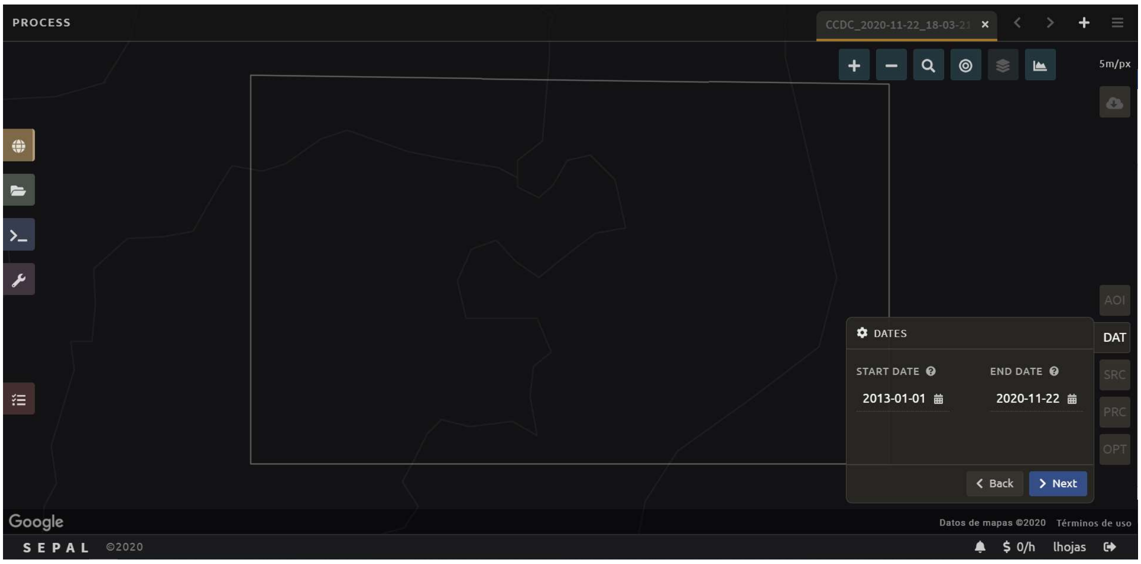


Planet composites. There are monthly composites from September 2020 and historical 6-month composites from December 2018.



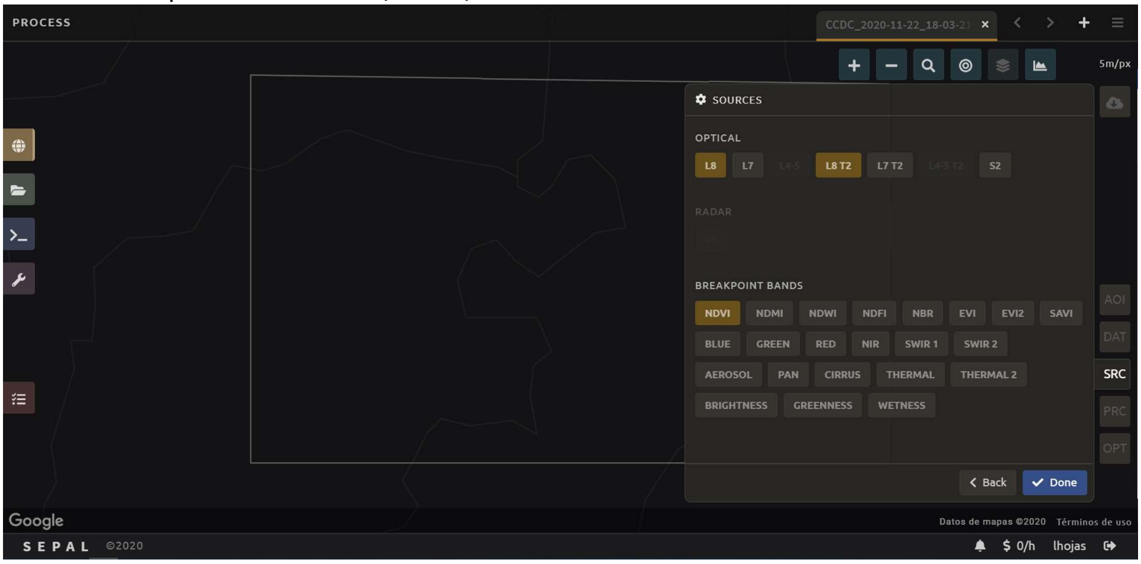
More options at the bottom right:

Dates: 1st January 2013 – today (L8 data is available since 2013)



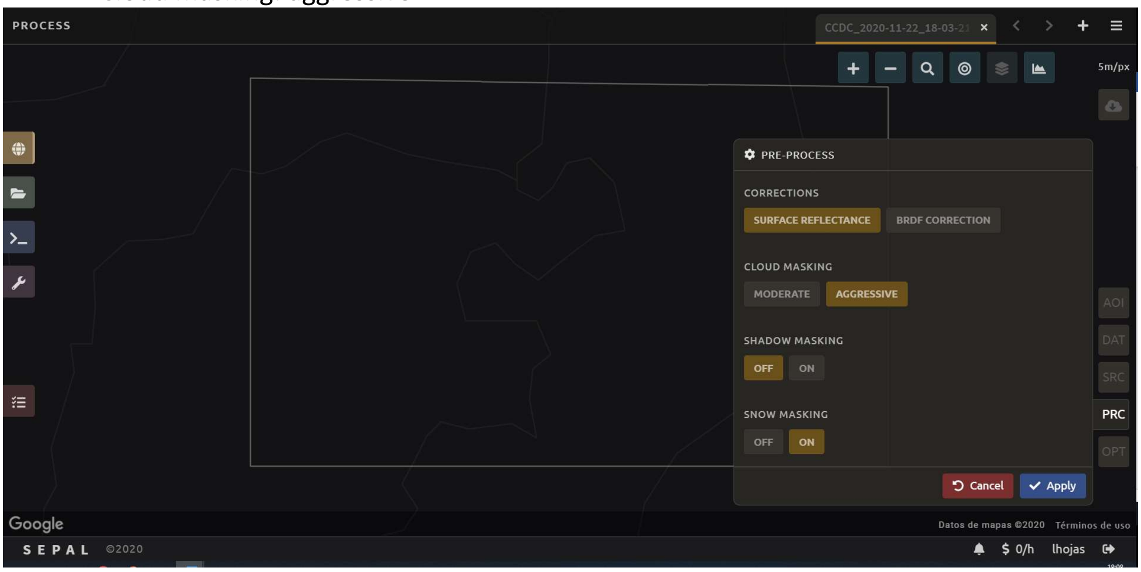
Sources:

- Optical: L8 + L8 T2
- Breakpoint bands: NDVI/NDMI/NDWI.



Pre-process:

- Corrections: surface reflectance
- Cloud masking: aggressive



CCDC Options

- Break detection: moderate

The 'aggressivity' for breakpoint detection depends on several factors, like the minimum number of consecutive observations with outlier values to flag a change and the probability of change. In SEPAL the default value for break detection is already more aggressive than the original algorithm; the first has been decrease from 8 to 6, and the second decreased from 99% to 95%. Too long-time segments can derive in wrong classifications (e.g., with continue changes).

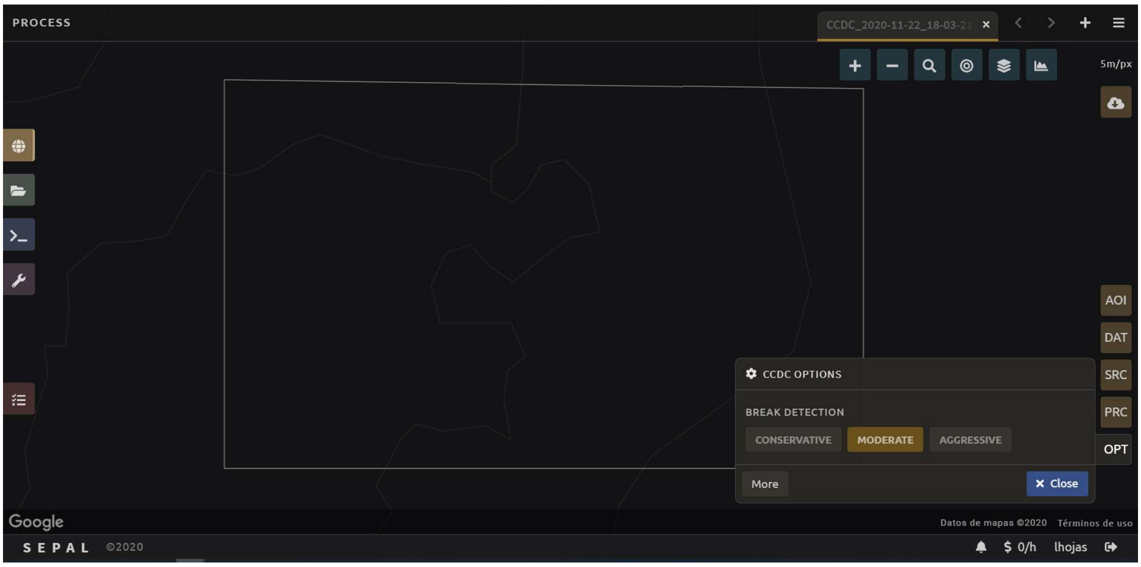
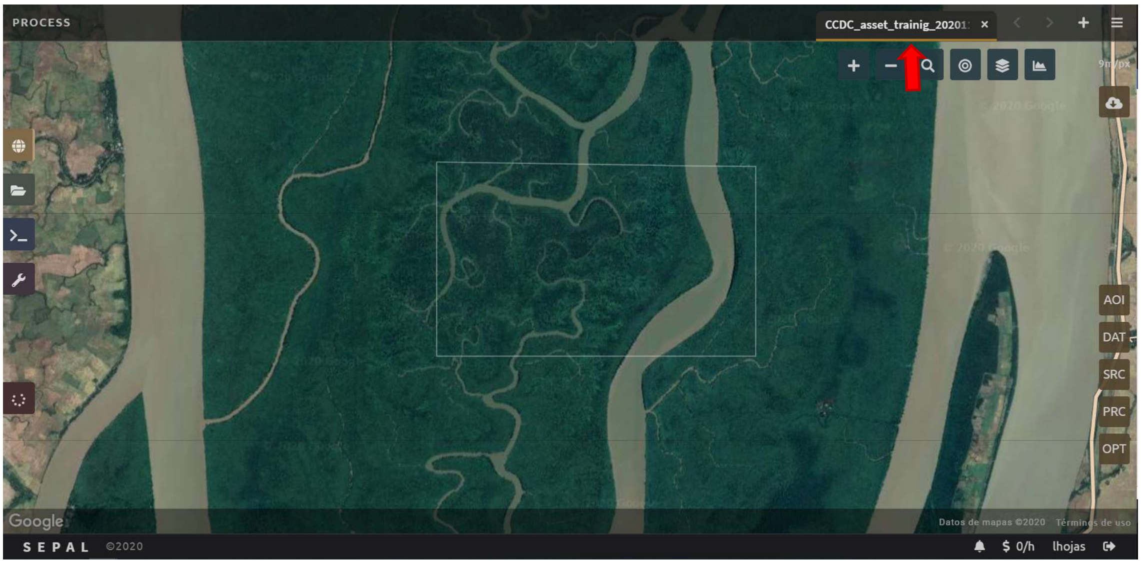


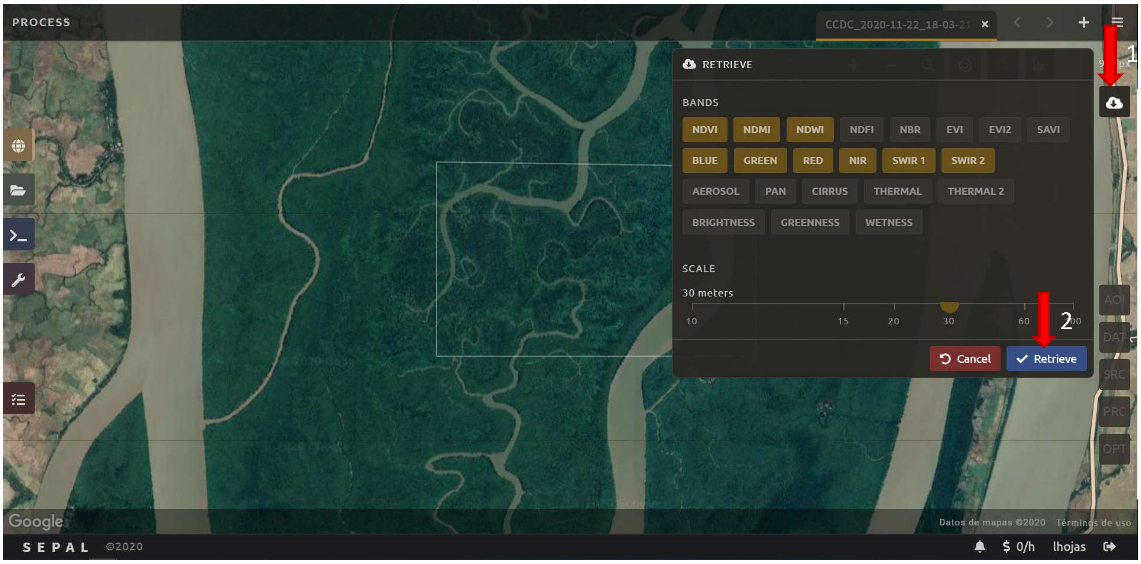
Chart time series for pixels: Click on a point inside your AOI and the time series of the corresponding pixel will be displayed. The chart represents the regression model for each time segment (red and orange lines) for the vegetation index or band selected from the list (in the example below, the NDVI), the breakpoints detected by the NDVI index as selected previously (in the example just before the 2018-04-29) and the real observations (yellow dots).

To select the most suitable indices, you can pay attention to how well the yellow dots fit the model.

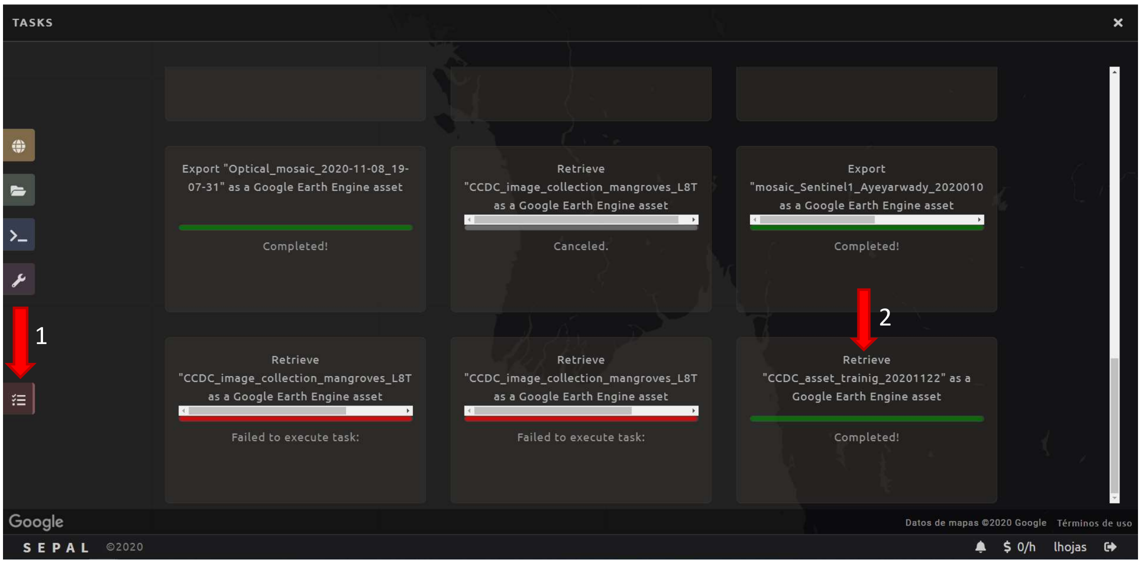
Give a name to save your recipe.



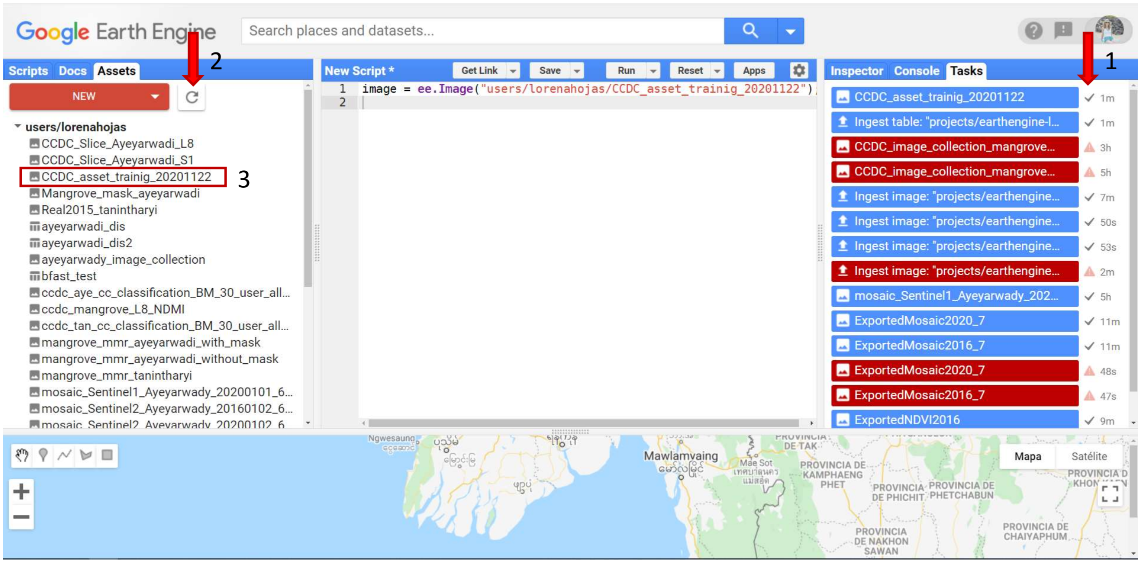
Bands to retrieve: Select the bands for which the models of the stable segments will be saved for later classification.



Under tasks you can see if your CCDC asset is being created. The task symbol will become a spinning wheel (1). When the CCDC asset has been created successfully, the processing bar will appear in green, and Completed! will be written below (2). The task symbol will appear again as in the picture below (1).



In the task panel of your Google Earth Engine account the check mark will appear next to your asset (1). If you do not see it in the asset panel, click on refresh asset cache (2).



Click on the name (3) to copy the Image ID (*users/lorenahojas/CCDC_asset_trainig_20201122*) for the next exercise.

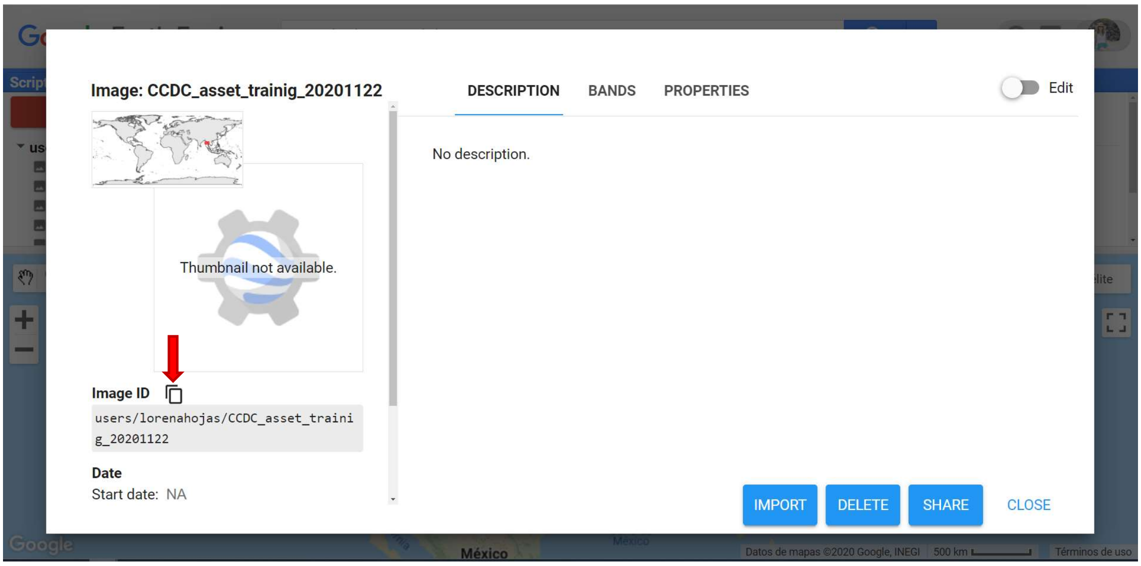
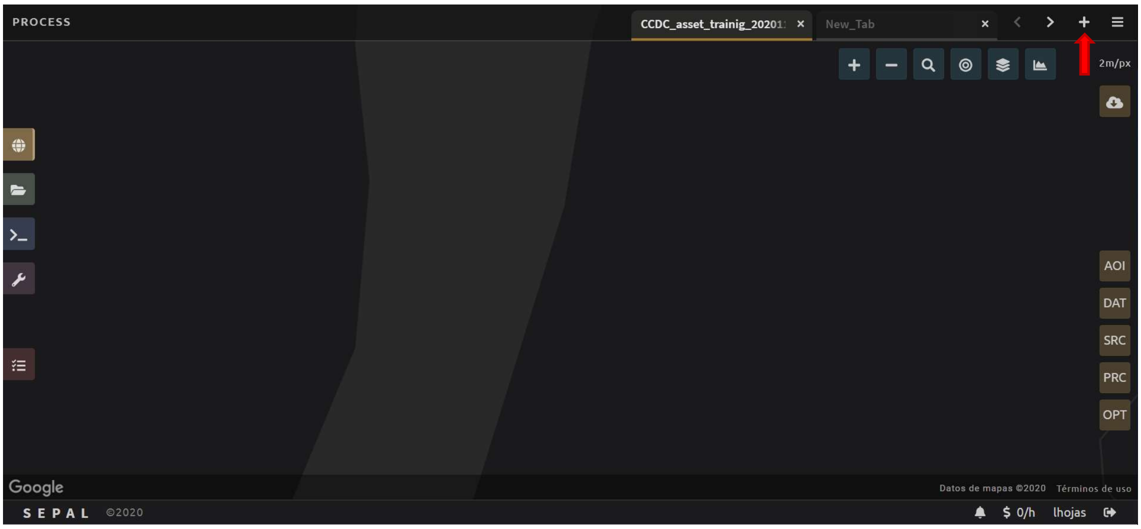


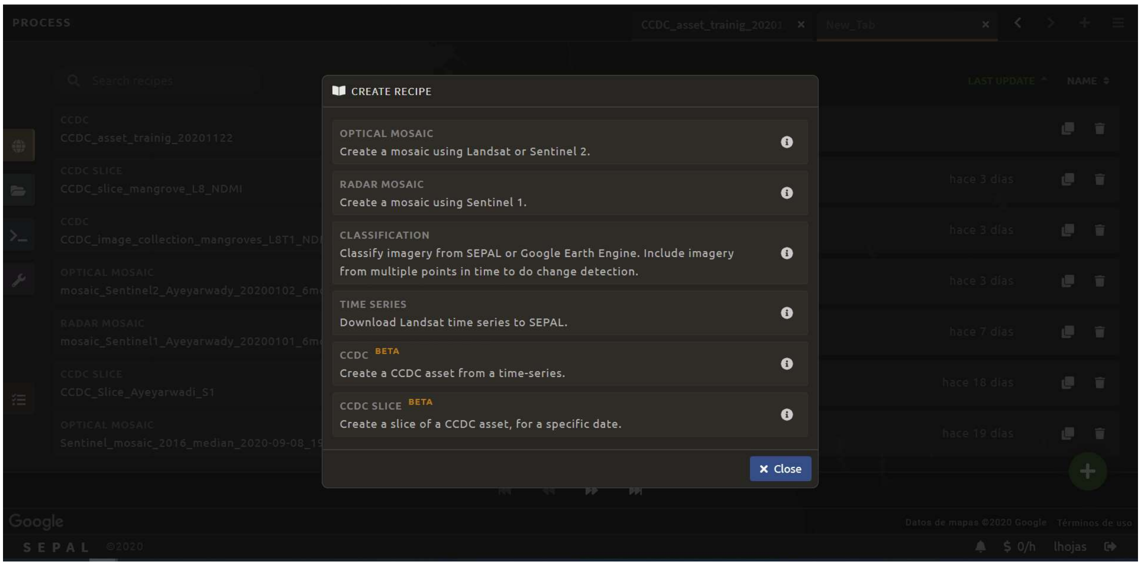
Image: CCDC_asset_trainig_20201122
DESCRIPTION
BANDS
PROPERTIES
Edit
No description.
Thumbnail not available.
Image ID
users/lorenahojas/CCDC_asset_traini
g_20201122
Date
Start date: NA
IMPORT
DELETE
SHARE
CLOSE
México
Datos de mapas ©2020 Google, INEGI
500 km
Términos de uso

4. Create a slice of a CCDC asset for a specific date in SEPAL

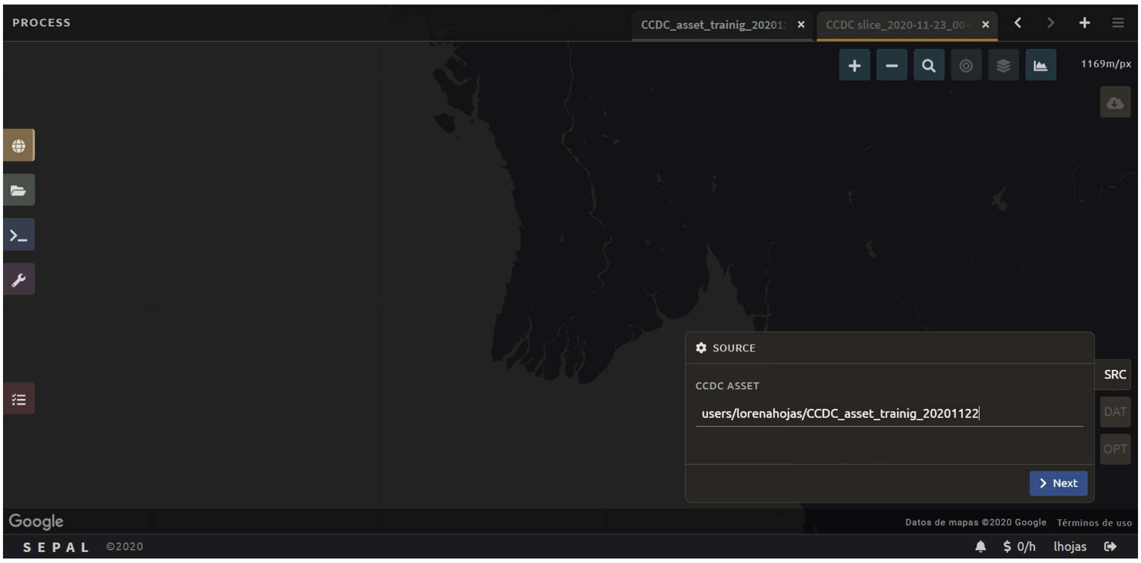
Go again to process. If you do not want to close your previous recipe, you can click on new tab. Select create new SEPAL recipe (cross in green circle at bottom right corner).



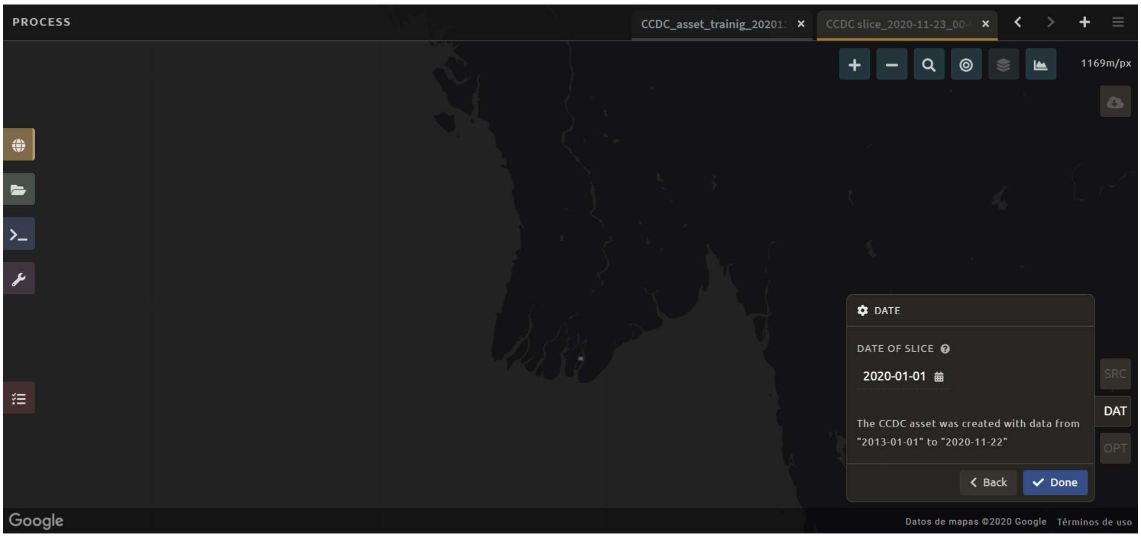
Select 'Create a slice of a CCDC asset for a specific date'.



Source: Paste the CCDC asset ID from your previous exercise.

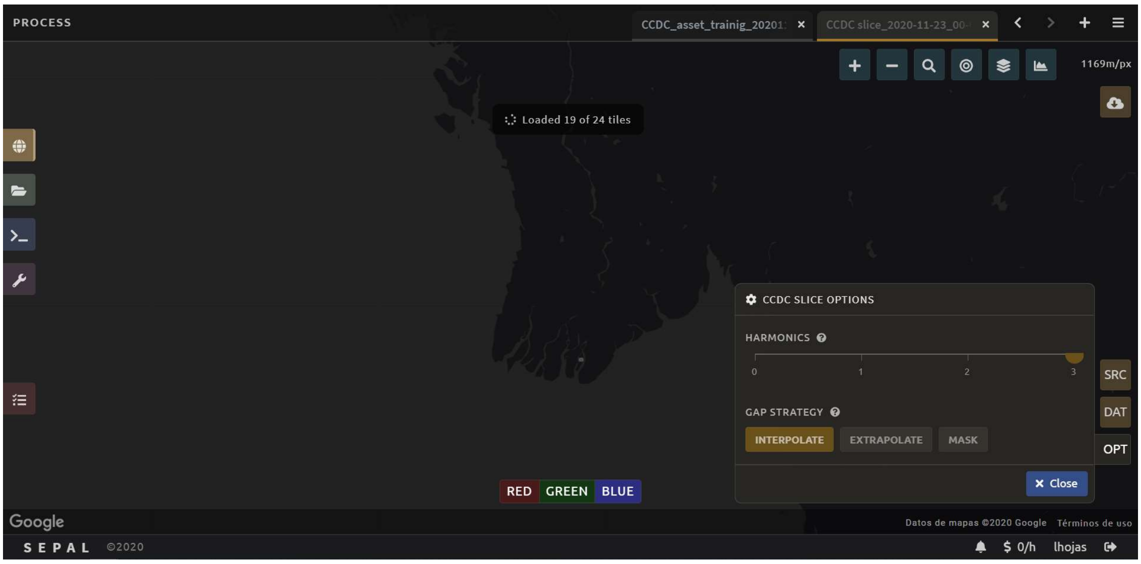


Date: Select a date from the period of your CCDC asset (1st January 2013 - 22nd November 2020), better if close to the date of the collection of your training data.

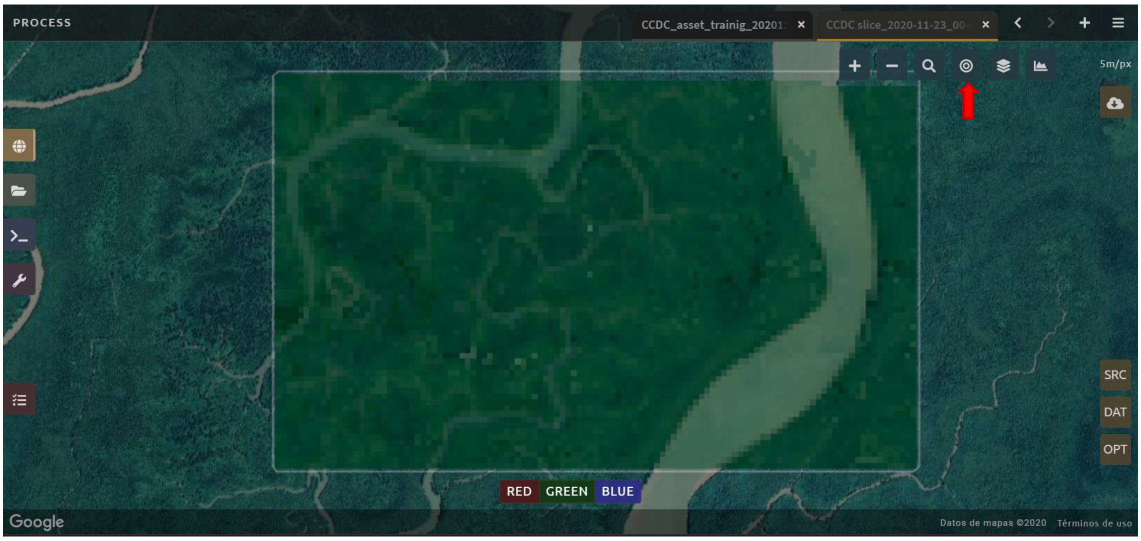


CCDC Slice Options:

- Harmonics from the regression model to use when creating the slice: 3.
- Strategy to use for pixels where there is no CCDC segment (breaks): interpolation. (The break detection was quite aggressive, therefore with interpolation the classification for these pixels will be a compromised).



Zoom to area of interest. Again, as base layer you can choose between SEPAL default layer, Google Maps or PLANET composites.



The CCDC slice is overlaying the base layer. You can unselect it.

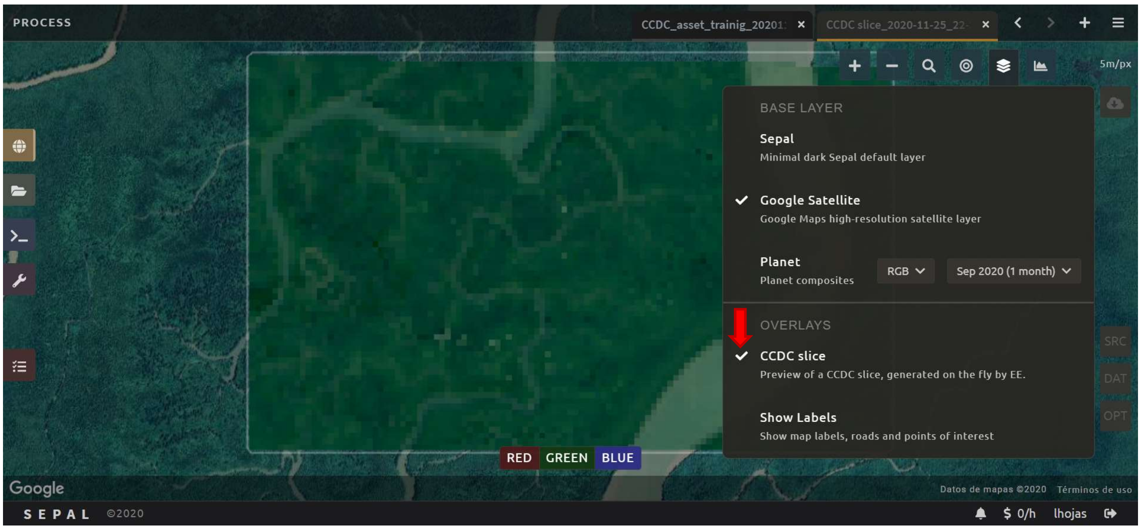
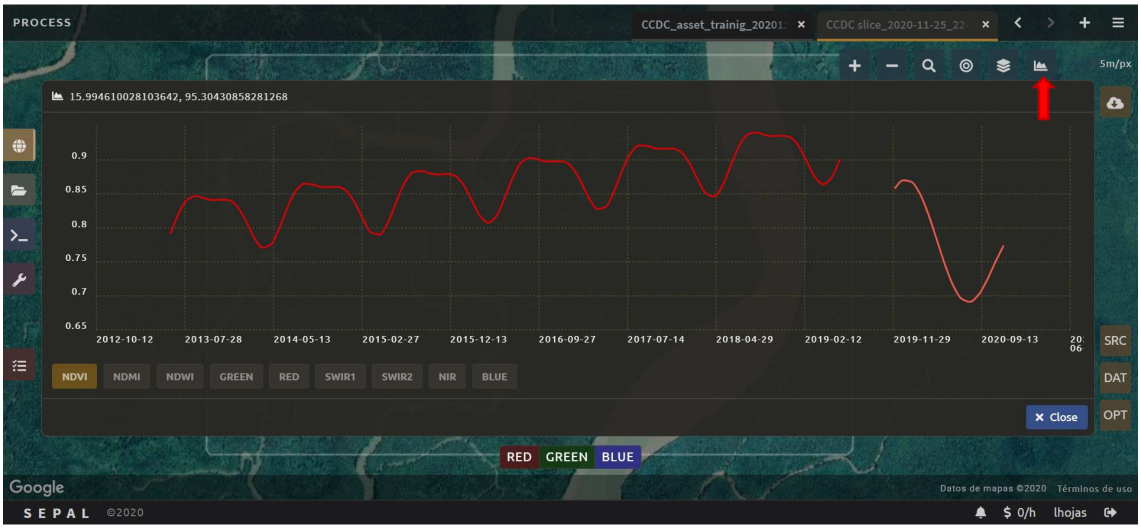


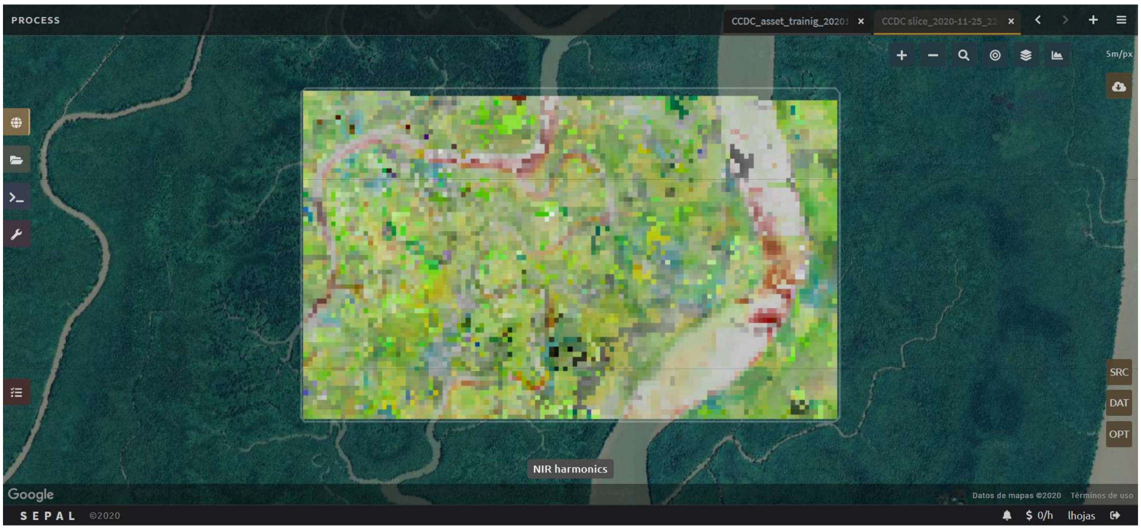
Chart time series for pixels: Click on a point inside your AOI and the time series of the corresponding pixel will be displayed. The chart represents the regression model (3rd order harmonic) for each time segment (red and orange lines) for the vegetation index or band selected from the list (in the example below, the NDVI) and the breakpoints detected by the bands/indices selected in the previous exercise (in the example just before 2019-11-29 by NDVI). Observe that we do no have yellow dots anymore because we do not have single observations anymore but model coefficients.



For the visualization of the slice for the specific date (in this example the 1st of January 2020) you have several options: 1) a band combination in the in the BGR colour model, 2) a vegetation index and 3) the harmonic parameters of one of the models in HSL (hue, saturation, lightness).

When the harmonic parameters are seen in HSL, the phase coefficient of the first order harmonic is represented by the *hue* (colour), the amplitude by the *saturation* (intensity) and the *lightness* (brightness) by the RMSE. The example represents the harmonic parameters of the NIR band model.

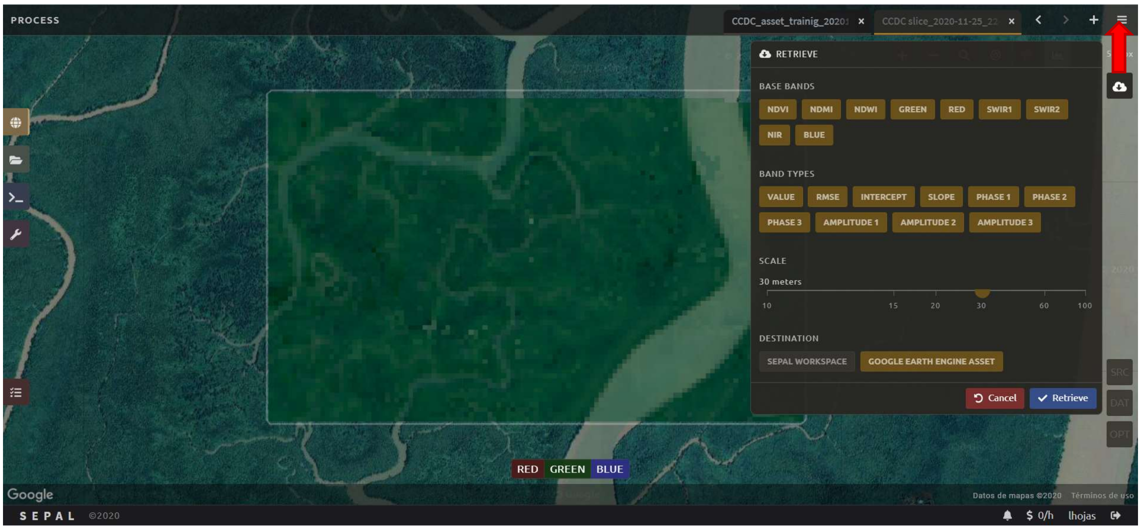
The harmonic representation is difficult to interpret. Areas with similar colour and intensity, represent areas with similar behaviour in the seasonality along the year, its intensity and stability (e.g., dark colour will be more typical of agricultural areas while white of natural forest, because the models are more stable through the year and between the years in natural forest than in agricultural areas).



Hue: phase (1 harmonic)
Saturation: amplitud
Lightness: rmse

Retrieve: Select the bands or indices from your CCDC asset for which the models for the time segment corresponding to the select date will be retrieve. For each band/index model you can choose to retrieve the value, the RMES (root-mean-square error) and/or the coefficients (intercept, slope, phase and amplitude). These parameters will be used later as inputs for the classification of each pixel.

Do not forget to give a name to save your recipe and to easily find the asset in your Google Earth Engine account.



This time you can choose to save it in your SEPAL workspace or as a Google Earth Engine asset. For the moment, there is not a CCDC asset slice classification module yet in SEPAL, so better to export it into your Google Earth Engine account.

Sources

Boston University, N., n.d. Change detection and classification algorithm [WWW Document]. GLanCE Glob. Land Cover Mapp. Estim. URL http://sites.bu.edu/measures/project-methods/change-detection-and-classification-algorithm/ (accessed 5.13.21).

Chen, B., Xiao, X., Li, X., Pan, L., Doughty, R., Ma, J., Dong, J., Qin, Y., Zhao, B., Wu, Z., Sun, R., Lan, G., Xie, G., Clinton, N., Giri, C., 2017. A mangrove forest map of China in 2015: Analysis of time series Landsat 7/8 and Sentinel-1A imagery in Google Earth Engine cloud computing platform. ISPRS J. Photogramm. Remote Sens. 131, 104–120. https://doi.org/10.1016/j.isprsjprs.2017.07.011

Zhu, Z., Woodcock, C.E., 2014. Continuous change detection and classification of land cover using all available Landsat data. Remote Sens. Environ. 144, 152–171. https://doi.org/10.1016/j.rse.2014.01.011