Research and User Support for Sentinel Core Products

The RUS Service is funded by the European Commission, managed by the European Space Agency and operated by CSSI and its partners.

Authors would be glad to receive your feedback or suggestions and to know how this material was used. Please, contact us on training@rus-copernicus.eu

Cover images produced by RUS Copernicus

The following training material has been prepared by Serco Italia S.p.A. within the RUS Copernicus project.

Date of publication: January 2018

Version: 1.2

Suggested citation:


This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

DISCLAIMER

While every effort has been made to ensure the accuracy of the information contained in this publication, RUS Copernicus does not warrant its accuracy or will, regardless of its or their negligence, assume liability for any foreseeable or unforeseeable use made of this publication. Consequently, such use is at the recipient’s own risk on the basis that any use by the recipient constitutes agreement to the terms of this disclaimer. The information contained in this publication does not purport to constitute professional advice.
# Table of Contents

1. Introduction to RUS ................................................................. 4
2. Crop mapping – background ..................................................... 4
3. Training .................................................................................. 4
   3.1 Data used ........................................................................ 4
   3.2 Software in RUS environment ............................................. 4
4. Register to RUS Copernicus ..................................................... 5
5. Request a RUS Copernicus Virtual Machine ............................ 6
6. Step by step ............................................................................ 10
   6.1 Data download – ESA SciHUB .......................................... 10
   6.2 SNAP – open and explore data ......................................... 12
   6.3 Resampling ..................................................................... 13
   6.4 Subset ............................................................................ 14
   6.5 Import vector data ............................................................ 15
   6.6 Reproject ....................................................................... 16
   6.7 Mask ............................................................................. 17
   6.8 Random Forest Classification ............................................ 18
7. Extra steps ................................................................................ 21
   7.1 Multi-temporal Random Forest classification ..................... 21
      7.1.1 Data download ........................................................ 21
      7.1.2 SNAP Open data ..................................................... 22
      7.1.3 Batch Processing ..................................................... 22
      7.1.4 Random Forest Classification .................................. 25
    7.2 Export to QGIS ................................................................. 28
8. Further reading and resources .................................................. 30
1 Introduction to RUS

The Research and User Support for Sentinel core products (RUS) service provides a free and open scalable platform in a powerful computing environment, hosting a suite of open source toolboxes pre-installed on virtual machines, to handle and process data derived from the Copernicus Sentinel satellites constellation.

In this tutorial we will employ RUS to identify and map different crops in an agricultural area near Seville (Spain) using Sentinel-2 satellite-borne optical data.

2 Crop mapping – background

The Guadalquivir Marshes represent one of the last territories occupied by man in the Spain’s region of Andalucia. Settlements were not definitive until the 1940 decade, when the expansion of agriculture, especially rice, led a complete transformation of the area. Nowadays, 35.000 ha are used to grow rice with an average production rate of 8.500-10.000 kg/ha. Other crops such as wheat, cotton, sugar beet, or sunflower can be found as well.

Reliable information on crops is required to improve agricultural management and reduce the environmental impact of this activity. Different methods can be used to gather this information but satellite earth observation techniques offer a suitable approach based on the coverage and type of data that are provided. The imagery data from Sentinel satellites enables a new approach for agriculture monitoring. The combination of their temporal, spatial, and spectral resolutions together with relevant analysis can lead to improvements of the decision-making process.

3 Training

Approximate duration of this training session is one hour.

The Training Code for this tutorial is LAND01. If you wish to practice the exercise described below within the RUS Virtual Environment, register on the RUS portal and open a User Service request from Your RUS service > Your dashboard.

3.1 Data used

- One Sentinel-2A level 2A image acquire on 01 June 2017 [downloadable at @ https://scihub.copernicus.eu/]
  S2A_MSIL2A_20170601T110651_N0205_R137_T30STG_20170601T111225
- Pre-processed data stored locally
  @/shared/Training/LAND01_CropMapping_Seville/AuxData/

3.2 Software in RUS environment

Internet browser, SNAP + Sentinel-2 Toolbox + QGIS
4 Register to RUS Copernicus

To repeat the exercise using a RUS Copernicus Virtual Machine (VM), you will first have to register as a RUS user. For that, go to the RUS Copernicus website (www.rus-copernicus.eu) and click on Login/Register in the upper right corner.

Select the option **Create my Copernicus SSO account** and then fill in ALL the fields on the Copernicus Users’ Single Sign On Registration. Click Register.

Within a few minutes you will receive an e-mail with activation link. Follow the instructions in the e-mail to activate your account.

You can now return to https://rus-copernicus.eu/, click on Login/Register, choose Login and enter your chosen credentials.
Upon your first login you will need to enter some details. You must fill all the fields.

5 Request a RUS Copernicus Virtual Machine

Once you are registered as a RUS user, you can request a RUS Virtual Machine to repeat this exercise or work on your own projects using Copernicus data. For that, log in and click on Your RUS Service → Your Dashboard.
Click on **Request a new User Service** to request your RUS Virtual Machine. Complete the form so that the appropriate cloud environment can be assigned according to your needs.

If you want to repeat this tutorial (or any previous one) select the one(s) of your interest in the appropriate field.

Complete the remaining steps, check the terms and conditions of the RUS Service and submit your request once you are finished.
Further to the acceptance of your request by the RUS Helpdesk, you will receive a notification email with all the details about your Virtual Machine. To access it, go to Your RUS Service → Your Dashboard and click on Access my Virtual Machine.
Fill in the login credentials that have been provided to you by the RUS Helpdesk via email to access your RUS Copernicus Virtual Machine.

This is the remote desktop of your Virtual Machine.
6 Step by step

6.1 Data download – ESA SciHUB

In this step we will download the Sentinel-2A level 2A image from the Copernicus Open Access Hub using the online interface.

Go to https://scihub.copernicus.eu/

Go to Open HUB. If you do not have an account please sign up in the upper right corner, fill in the details and click register.

You will receive a confirmation email in the account you have specified: open the email and click on the link to finalize the registration.

Once your account is activated – or if you already have an account – log in.
Switch the rectangle drawing mode to pan mode by clicking on the icon in the lower left corner of the map (Green rectangle) and navigate to the Seville city area (approximate area – blue rectangle).

Switch now to drawing mode (box) and draw a search rectangle approximately as indicated below. Open the search menu by clicking to the left part of the search bar (≡) and specify the parameters below. Press the search button (√) after that.

**Sensing period:** From 2017/06/01 to 2017/06/01

**Check Mission:** Sentinel-2

**Product type:** S2MSI2Ap
In our case the search returns 4 results depending on the defined search area. Download scene: S2A_MSIL2A_20170601T110651_N0205_R137_T30STG_20170601T111225

Move the downloaded scenes (desktop, /home/rus/Downloads) to the following path and unzip it by right clicking on it and using Extract Here.

Path: shared/Training/LAND01_CropMapping_Seville/Original

6.2 SNAP – open and explore data

In Applications -> Processing open SNAP Desktop; click Open product ( ), navigate to the following path and open the product: MTD_MSIL2A.xml from the .SAFE folder.

Path: shared/Training/LAND01_CropMapping_Seville/Original/

The opened product will appear in Product Explorer window. Right click on the product and select Open RGB Image Window to create and visualize an RGB composition image. Check that B4, B3, and B2 are selected for the Red, Green, and Blue channels respectively. Click OK.
6.3 Resampling

Since the Sentinel-2 images have different pixel sizes depending on the spectral band (See NOTE 1), it is necessary to resample the product and equalize the different spatial resolutions. This step is generally required for any further processing.

Click Raster -> Geometric Operations -> Resampling
In the I/O Parameters tab, make sure that the source product starts with index [1]. Set the Output folder to the following path:

Path: shared/Training/LAND01_CropMapping_Seville/Processing/

Use this folder to save all following processing products. Go to the Resampling Parameters tab and select By reference band from source product. Select B2 as the reference band. All other settings remain set to default values and click Run.

To visualize the output product, right click on it. Select Open RGB Image Window and click OK.

6.4 Subset

Since our Area of Interest (AOI) is quite small and there is no need to process the whole image we start with sub-setting the scene to a more manageable size (See NOTE 2). This will reduce the processing time in further steps and is recommended when the analysis is focused only over a specific area and not the complete scene. In the Product Explorer window, select the second resampled product and go to Raster -> Subset.

In the Subset menu, set the extent of the AOI to the following pixel coordinates and click OK.

Scene start x: 0  Scene start Y: 7651  Scene end X: 5689  Scene end Y: 10978
The subset product will be created immediately. Before continuing, save it to prevent errors. Right click on the subset product (index [3]) and select *Save Product*. Set the output folder to the following path and save it.

Path: *shared/Training/LAND01_CropMapping_Seville/Processing/*

To visualize the output product, right click on it. Select *Open RGB Image Window* and click OK.

### 6.5 Import vector data

To prepare the data for the classification, the shapefiles of the study area and training areas must be imported. Click on *Vector -> Import -> ESRI Shapefile*. Navigate to the following path, select all the shapefiles and click *Open*. Click *No* in all the Import Geometry dialogs.

Path: *shared/Training/LAND01_CropMapping_Seville/AuxData/*
Once the vector data have been imported, do not forget to save the changes. Right click on the subset product (index [3]) and click on Save Product. The Vector Data folder of the subset product should look like the following image. Expand the product and open the Vector Data folder to check it.

6.6 Reproject

Before doing the classification, the product has to be reprojected due to software requirements. Click on Raster -> Geometric Operations -> Reprojection.

In the I/O Parameters tab make sure that the selected source product is the subset product (index [3]). Set Output folder to the following path:

Path: shared/Training/LAND01_CropMapping_Seville/Processing/
Move to the Reprojection Parameters tab. Select Custom CRS and choose the UTM / WGS 84 (Automatic) projection. All other settings remain set to default values. Click Run.

To visualize the output product, right click on it. Select Open RGB Image Window and click OK.

6.7 Mask

The last step before running the classification will be to mask out the pixels that are not inside the study area. To perform this process, click on Raster -> Mask -> Land/Sea Mask. Make sure the reprojected product (index [4]) is selected as input. Set the Output folder to the following path:

Path: shared/Training/LAND01_CropMapping_Seville/Processing/

Go now to the Processing Parameters tab. Select from B1 to B12 as source bands. Uncheck the option Use SRTM 3sec, and select Use Vector as Mask. Open the drop-down menu and choose Study_Area as mask, then click Run.

To visualize the output product, right click on it. Select Open RGB Image Window and click OK.
6.8 Random Forest Classification

For this exercise, the Random Forest classification algorithm will be used (See NOTE 3).

NOTE 3: The Random Forest algorithm is a machine learning technique that can be used for classification or regression. In opposition to parametric classifiers (e.g. Maximum Likelihood), a machine learning approach does not start with a data model but instead learns the relationship between the training and the response dataset. The Random Forest classifier is an aggregated model, which means it uses the output from different models (trees) to calculate the response variable.

Decision trees are predictive models that recursively split a dataset into regions by using a set of binary rules to calculate a target value for classification or regression purposes. Given a training set with $n$ number of samples and $m$ number of variables, a random subset of samples $n$ is selected with replacement (bagging approach) and used to construct a tree. At each node of the tree, a random selection of variables $m$ is used and, out of these variables, only the one providing the best split is used to create two sub-nodes.

By combining trees, the forest is created. Each pixel of a satellite image is classified by all the trees of the forest, producing as many classifications as number of trees. Each tree votes for a class membership and then, the class with the maximum number of votes is selected as the final class.

More information about Random Forest can be found in Breiman, 2001.

Click on Raster -> Classification -> Supervised Classification -> Random Forest Classifier

To properly visualize the product later, go to Layer -> Layer Manager and make sure that the Vector Data is not selected. Now, Click on the symbol. Navigate to the following path and select the following product:
Move to the Random-Forest-Classifier tab and set the following parameters:

- Uncheck the Evaluate classifier option
- Set the number of trees to 500
- Select all the shapefiles as training vectors except the Study_Area
- Select all the bands (B1 to B12) as feature bands

Click now on the Write tab, set the Output folder to the following path (create the RandomForest folder if not created yet), and specify the output name according to the number of input images used: RF_${num_images}. Finally, click Run.

Path: /shared/Training/LAND01_CropMapping_Seville/Processing/RandomForest/

Before visualizing the classified output, the valid pixel expression has to be changed. By default, only the pixels with a confidence threshold above 0.5 are displayed. To change this parameter, expand the Random Forest product (index [6]), open the Bands folder, right click on the LabeledClasses file, and click on Properties.
Delete the expression $\text{Confidence} >= 0.5$ and click Close.

Double click on the LabeledClasses file to open the classification image. You can change the colours by clicking on the Colour Manipulation tab located in the lower left corner or by clicking on View -> Tool Windows -> Colour Manipulation. Select your own colours or click on the Import colour palette icon ( ). Navigate to the following path and select the file Colour_Palette_SNAP.cpd

Path: /shared/Training/LAND01_CropMapping_Seville/AuxData/Colour_Palette/

To properly visualize the new product, go to Layer -> Layer Manager and make sure that the Vector data is not selected.
7 Extra steps

7.1 Multi-temporal Random Forest classification

7.1.1 Data download

To improve the classification result obtained in the previous chapter, the number of input images can be increased. In that way, the algorithm will have more information to take into account and better output can be produced.

Go back to chapter 6.1 Data download – ESA SciHUB and follow the steps to download the Sentinel-2 images corresponding to the following dates:

- **2017_06_11** → S2A_MSIL2A_20170611T110621_N0205_R137_T30STG_20170611T111012
- **2017_06_21** → S2A_MSIL2A_20170621T110651_N0205_R137_T30STG_20170621T111222
- **2017_07_01** → S2A_MSIL2A_20170701T111051_N0205_R137_T30STG_20170701T111746
- **2017_07_11** → S2A_MSIL2A_20170711T110651_N0205_R137_T30STG_20170711T111223
- **2017_07_21** → S2A_MSIL2A_20170721T110621_N0205_R137_T30STG_20170721T112025
- **2017_07_31** → S2A_MSIL2A_20170731T110651_N0205_R137_T30STG_20170731T111220

Do not forget to move the downloaded products (desktop, /home/rus/Downloads) to the following path and unzip them by right clicking to each one and then selecting Extract Here.

Path: shared/Training/LAND01_CropMapping_Seville/Original/

Go back to SNAP and close all the images. Right click on any of the products and select Close All Products. Click No when asked to save the product.

Due to software requirements, the Sentinel-2 Reader Masks have to be disabled to prevent errors. Click on Tools -> Options and select the S2TBX tab. At the Sentinel-2 Reader Masks tab, uncheck all the options and click OK.
7.1.2 SNAP Open data

Go to *File -> Open product* ( ![File -> Open product](image) ) and navigate to the path indicated below. In each folder, open the file *MTD_MSIL2A.xml*. Start the process with the folder on top and continue successively to open all the images in chronological order. Your *Product Explorer* window should have the same content as the one shown on the figure below.

Path: `shared/Training/LAND01_CropMapping_Seville/Original/`

7.1.3 Batch Processing

Before running the multi-temporal Random Forest Classification, the images must be pre-processed as before. However, since we have more inputs to analyze, we will use the Batch Processing option included in SNAP. This feature allows the execute graphs including processing tools to a set of inputs.

Click on *Tools -> Batch Processing* or click on the Batch Processing icon ( ![Batch Processing icon](image) ).

Click on the *Add Opened* icon ( ![Add Opened icon](image) ) to load all the images in the Batch Processing. Press the refresh button ( ![Refresh button](image) ) to update the metadata information, and uncheck the *Keep source product name* option.

Once the Sentinel-2 images are loaded, the graph containing the pre-processing tools have to be included as well. Click on *Load Graph*, navigate to the following path and select the file *Pre_Processing.xml*.

Path: `shared/Training/LAND01_CropMapping_Seville/AuxData/Batch_Processing_Graph/`
Go to the Resample tab and choose the option *By reference band from source product*. Open the drop-down menu and choose B2 as reference band. All other settings remain set to default values.

Click on the Subset tab and set all the bands from B1 to B12 as source bands. Select *pixel coordinates* and specify the following values. All other settings remain set to default values.

X: 0       Y: 7651       Width: 5689       Height: 3327
Click on the `Import-Vector` tab. Navigate to the following path and select the `Study_Area.shp` file.

Path: `shared/Training/LAND01_CropMapping_Seville/AuxData/`

Click on the `Reproject` tab and select the `Custom CRS` option. At the `Projection`, open the drop-down menu and select the option `UTM / WGS 84 (Automatic)`. All other settings remain as default.
Click on the Land-Sea-Mask. Select from B1 to B12 as source bands. Uncheck the option *Use SRTM 3sec* and select *Use Vector as Mask*. At the drop-down menu choose *Study_Area_1* as mask. All other settings remain set to default values.

Finally, click on the *Write* tab and set the output directory to the following path. Then, click *Run*. If not created yet, add the *Batch_Processing* folder within the *Processing* directory.

Path: *shared/Training/LAND01_CropMapping_Seville/Processing/Batch_Processing/*

### 7.1.4 Random Forest Classification

After pre-processing the images, the random forest classification can be applied. First, select all the Sentinel-2 original images (files in the *Product Explorer* window with index [1] to [7]), right click and select *Close All Products*. Click *No* when asked to save the product.

Select now the first product (index [8], date 2017_06_01) and click on *Vector -> Import -> ESRI Shapefile*. Navigate to the following path, select all the shapefiles except the *Study_Area.shp* shapefile and click *Open*. Click *No* in all the *Import Geometry* dialog that will appear.

Path: *shared/Training/LAND01_CropMapping_Seville/AuxData/*
Once the vector data has been imported, do not forget to save the changes. Right click on the product (index [1]) and click on Save Product. The Vector Data folder of the subset product should look like the following image. Expand the product and open the Vector Data folder to check it.

Click on Raster -> Classification -> Supervised Classification -> Random Forest Classifier

On the Product-Reader tab, click on the Add Opened icon ( ) to load all the images. Press the refresh button ( ) to update the metadata information.

You can now decide the number of images you want to use for the multi-temporal Random Forest Classification. The classification can be repeated as many times as desired and with different number of Sentinel-2 images as input. To remove an image from the list, select it and press the icon.

Remember to always include the image from 2017_06_01 (index [1]) since it is the one containing the training vectors.

For this exercise we will run the multi-temporal Random Forest Classification using all the S2 images.

Move to the Random-Forest-Classifier tab and set the following parameters:

- Select the option Train and apply classifier
- Uncheck the Evaluate classifier option
- Set the number of trees to 500
- Select all the shapefiles as training vectors except the Study_Area.shp
- Select all the bands as feature bands
Click now on the Write tab, set the output folder to the following path, and specify the output name according to the number of input images (e.g. ‘RF_7_images’). Finally, click Run (See NOTE 4).

Path: shared/Training/LAND01_CropMapping_Seville/Processing(RandomForest/

To properly visualize the result, go to Layer -> Layer Manager and make sure that the Vector Data field is not selected. Once the classification is done, remember to change the confidence threshold. Expand the Random Forest classification product, open the Bands folder right click on the LabeledClasses file, and click on Properties. Delete the valid-pixel expression Confidence >= 0.5 and click Close.

Double click on the LabeledClasses file to open the classification image. You can change the colours by clicking on the Colour Manipulation tab located in the lower left corner or by clicking on View -> Tool Windows -> Colour manipulation. Select your own colours or click on the ‘Import colour palette’ icon ( ). Navigate to the following path and select the file Colour_Palette_SNAP.cpd

Path: shared/Training/LAND01_CropMapping_Seville/AuxData/Colour_Palette/

NOTE 4: Due to a SNAP issue, it is possible that an error will appear. In case you see the message Error: [Nodeld: Random-Forest-Classifier] Cannot select feature band B1 in more than one product you need to rename all the bands of 6 out of 7 products (do not change the first one, the image from 20170601, since it contains the training vectors). For this, expand each product, open the Bands folder, right click on each band, and click Properties. As an example, change the name of the first band in the second product (20170611) from B1 (443 nm) to 1a as shown in the image below. Do the same for the remaining bands.

For the third product change the name from B1 (443 nm) to 1b and continue accordingly. Repeat this step for all bands of the remaining products. Remember to save the products after the changes have been made (right click on the product -> Save Product.)
7.2 Export to QGIS

To import the classification raster in QGIS, we will convert it into GeoTIFF format. You can perform this step for the single-date Random Forest classification or for the multi-temporal Random Forest classification.

Expand now the classified product and open the LabeledClasses band.

Click on File -> Export -> GeoTIFF and save the product in the following path with the appropriate name: RF_ + number of images used + _images (e.g. RF_7_images).

Path: shared/Training/LAND01_CropMapping_Seville/Processing/RandomForest/

Minimize SNAP and open QGIS (Applications -> Processing -> QGIS Desktop). Press the Add Raster Layer button ( ). Navigate to the following path and select the Random Forest classification GeoTIFF file. Click Open.

Path: shared/Training/LAND01_CropMapping_Seville/Processing/RandomForest/

The classification product is open as a multiband raster file. To change the visualization, right click on opened file and select Properties.
Select the Style tab on the left panel and choose the following parameters:

**Render type:** Singleband pseudocolor

**Mode:** Equal Interval

**Classes:** 12

**Min / Max:** -1 / 10

Press Classify. You can choose your own colours or use a predefined colour palette. Press the ‘Load Colour Map’ icon ( ), navigate to the following path, select Colour_Palette_QGIS.txt and click OK.

Path: shared/Training/LAND01_CropMapping_Seville/AuxData/Colour_Palette/
THANK YOU FOR FOLLOWING THE EXERCISE!

8 Further reading and resources


FOLLOW US!!!

@RUS-Copernicus

RUS-Copernicus

RUS-Copernicus

RUS Copernicus Training

RUS-Copernicus website

RUS-Copernicus Training website