


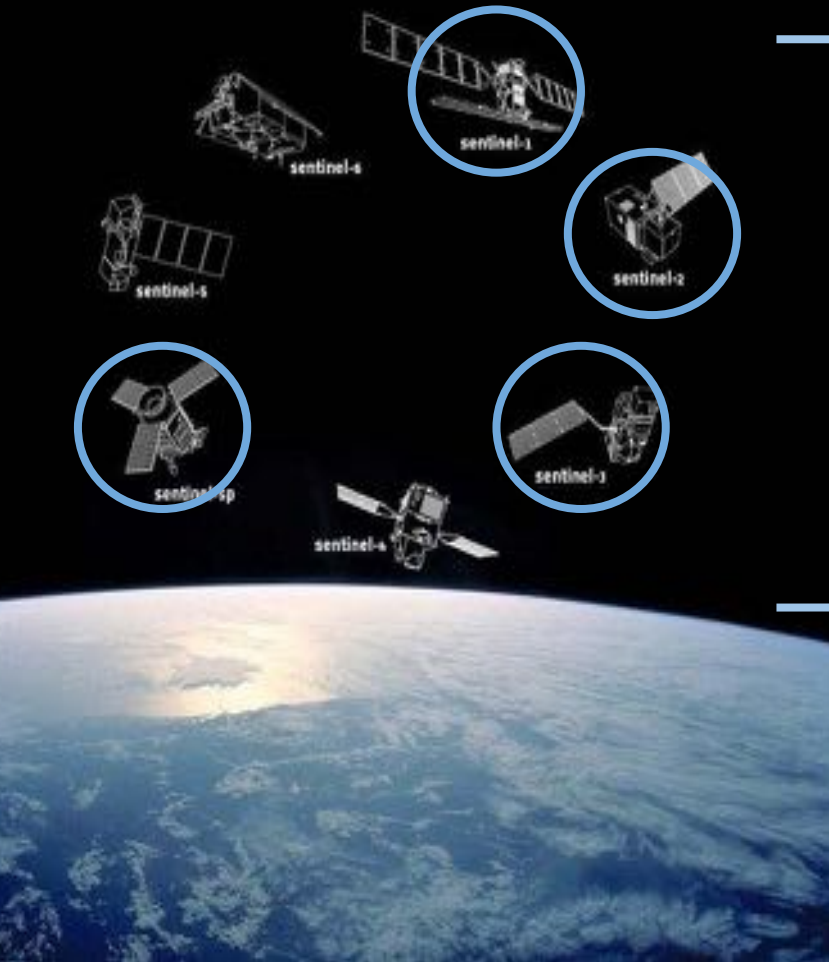
# A Walk-Through on Machine Learning Techniques for Sentinel Big Data Fusion

Sara Aparício  
ESA ESRIN, EOP- $\Phi$

**$\Phi$ -Week** - 12<sup>th</sup> November 2018

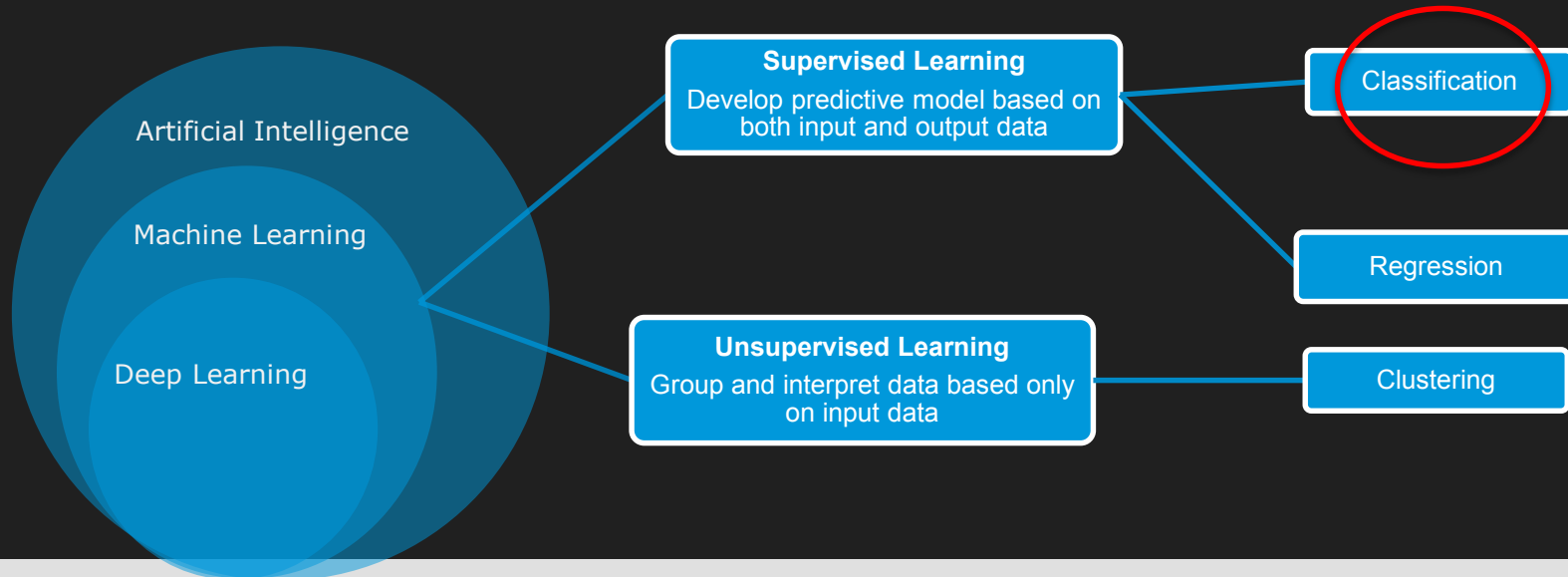
 @\_SaraAparicio\_

The world's most comprehensive suite of dedicated EO missions.

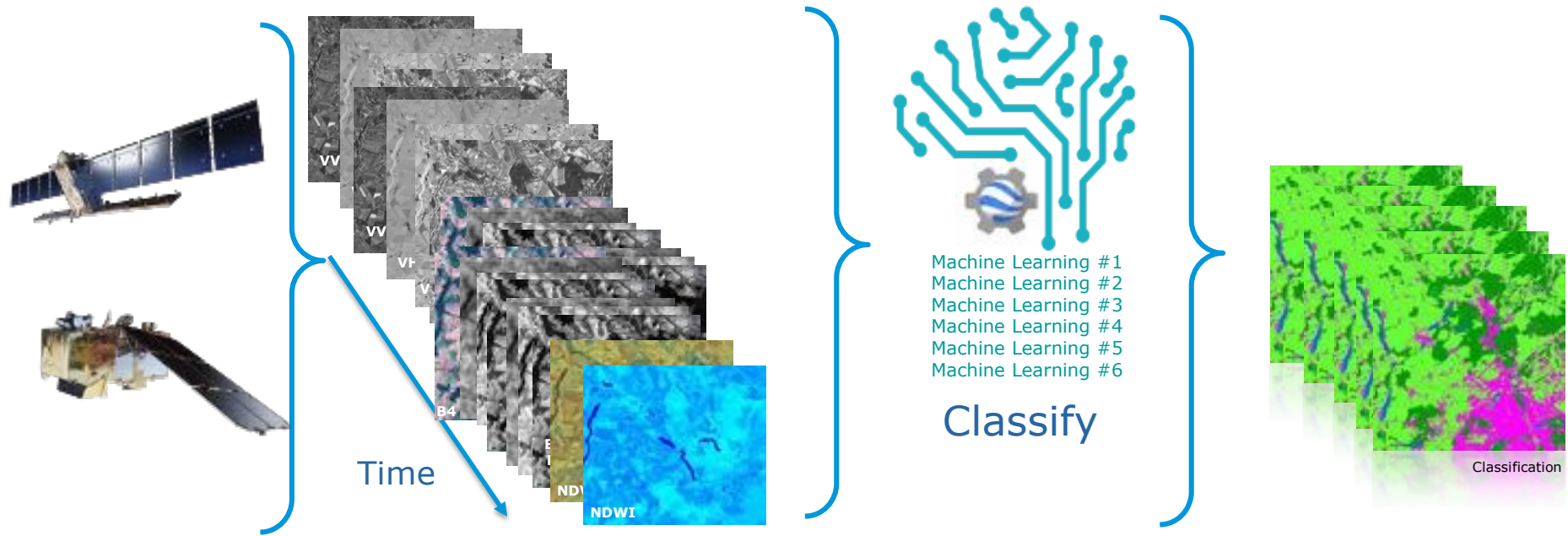


# Major trends

- a) The advent of cloud computing & increase computing power,
- b) proliferation of open-access satellite data streams,
- c) growing use of machine-learning algorithms



# Sentinel data fusion with Machine Learning Techniques



Multisensor/multitemporal data

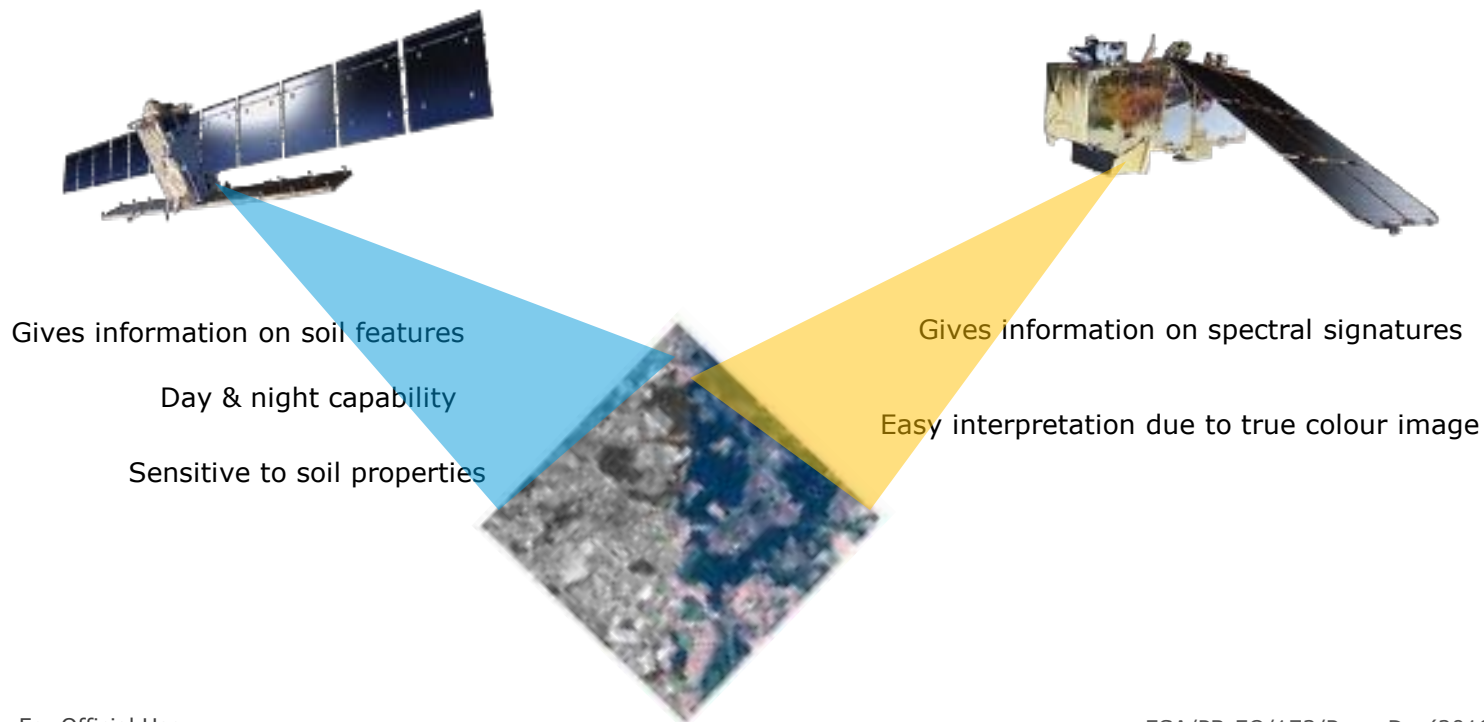
Best technique?  
Best band selection?





# Motivation I – Merging different datatypes

A promising direction of machine learning in Earth Observation is its pairing with data fusion



# Motivation 2 – AI4EO the white paper



This document addresses elements of what needs to be done at European level, some of the potential of artificial intelligence in data and the data supply chain and vice versa.

**Classification/ Recognition**  
- updating land-cover maps

It captures the recommendations of a community-led AI4EO workshop held at ESA/ESRIN (see list in Annex 1). The aim is to assess progress in the development of techniques to the world of EO and to explore the potential value of a concerted **Research and Innovation (R&I)** effort on this topic at **European** level.

**Detection**  
- large scale in automatic basis  
- use SAR in machine learning

The white paper will be a *dynamic* report capturing the community. It will be reviewed at a workshop at ESA/ESRIN on November 14th, 2018.....(....)

**Data Fusion**  
- Merging diverse EO data

# Objectives/Goal



1. To develop a **pixel-based classification, reproducible, scalable with a machine learning-based approach of large-area mapping**/land cover of high resolution (10m) based on a multi-sensor & multi-temporal approach;
2. To evaluate the **additive value** of open-access satellite optical and radar variables, processed using cloud computing, to a topographic baseline model.
3. To explore/understanding/address efficiency of Google Earth Engine to effectively execute **big data workflows** using machine learning techniques on Google Earth Engine (and accuracy) for multi-temporal land use mapping.





# Google Earth Engine

Sentinel-1 (GRD)  
 Sentinel-2 (TOA)  
 Sentinel-3 (OLCI)  
 Sentinel-5  
 Elevation  
 Climate  
 Topography  
 LANDSAT DATA  
 MODIS DATA  
 PROBA-V



The screenshot shows the Google Earth Engine web interface. On the left is the **Script manager** with a tree view of assets. The center is the **Code Editor** containing JavaScript code for a machine learning classification script. On the right is the **Inspector Console** showing a bar chart of the script's output.

Script manager

Code Editor

Console

The screenshot shows the **Map Output Panel** of Google Earth Engine. It displays a satellite image of a forested area with a semi-transparent map overlay showing a multi-class classification result. The map uses a color palette where green represents forest, yellow and red represent other land cover types, and blue represents water.

Map Output Panel

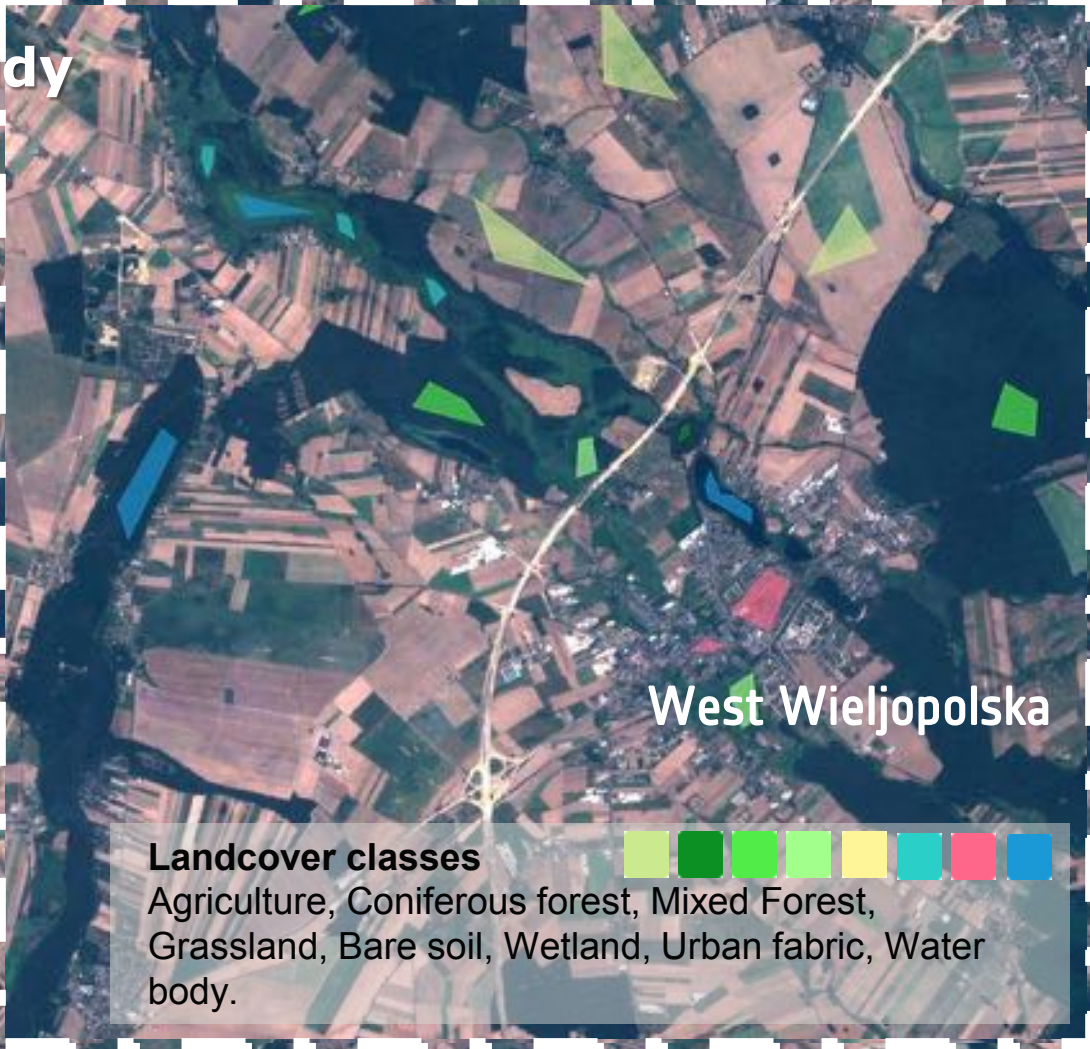




An aerial photograph of a dry, cracked landscape, likely a salt flat or a desert. The ground is a mix of orange, yellow, and brown tones, with prominent white and light-colored mineral deposits or salt crusts. The terrain is uneven and textured, with various shades of earthy colors. A central black bar with rounded corners and a white border contains the word "METHODOLOGY" in white, bold, uppercase letters.

# METHODOLOGY

# Case study



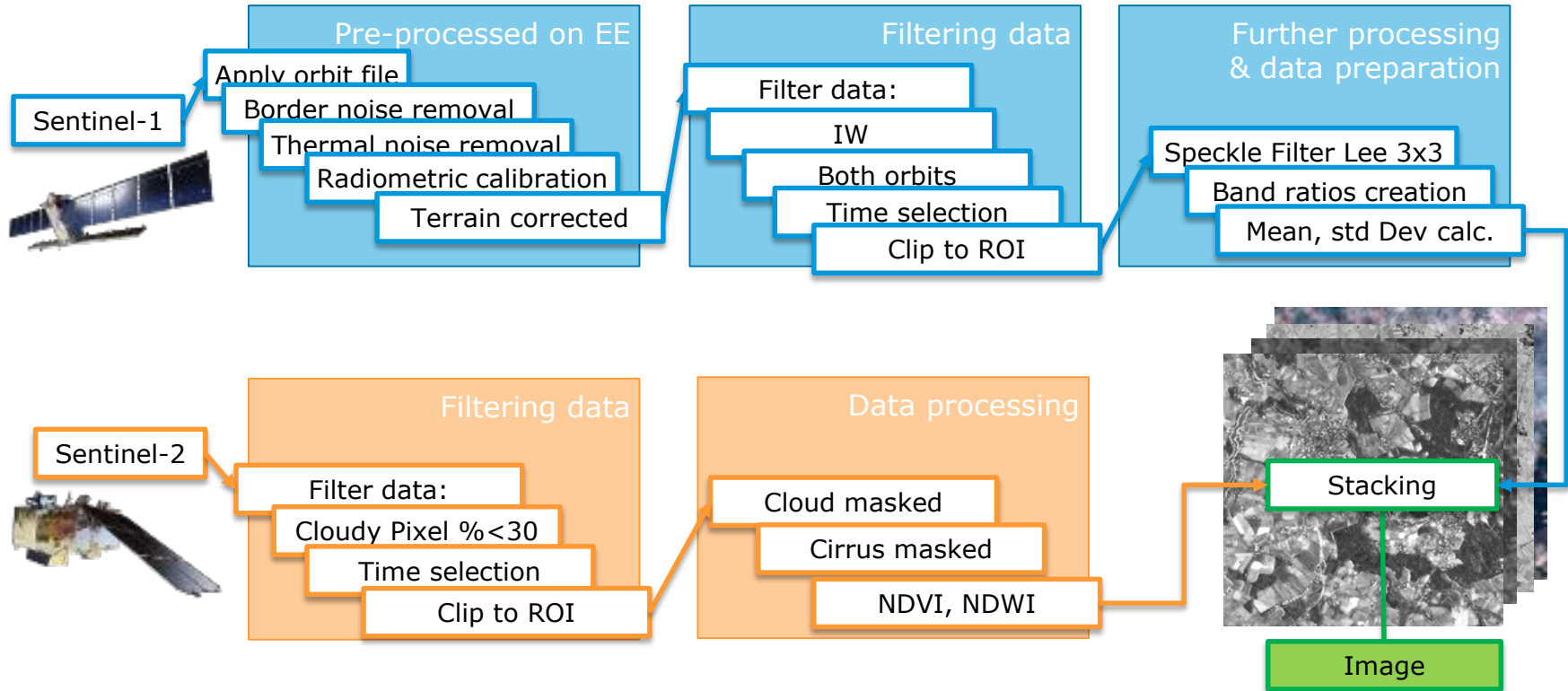
West Wieljopolska

### Landcover classes

Agriculture, Coniferous forest, Mixed Forest, Grassland, Bare soil, Wetland, Urban fabric, Water body.

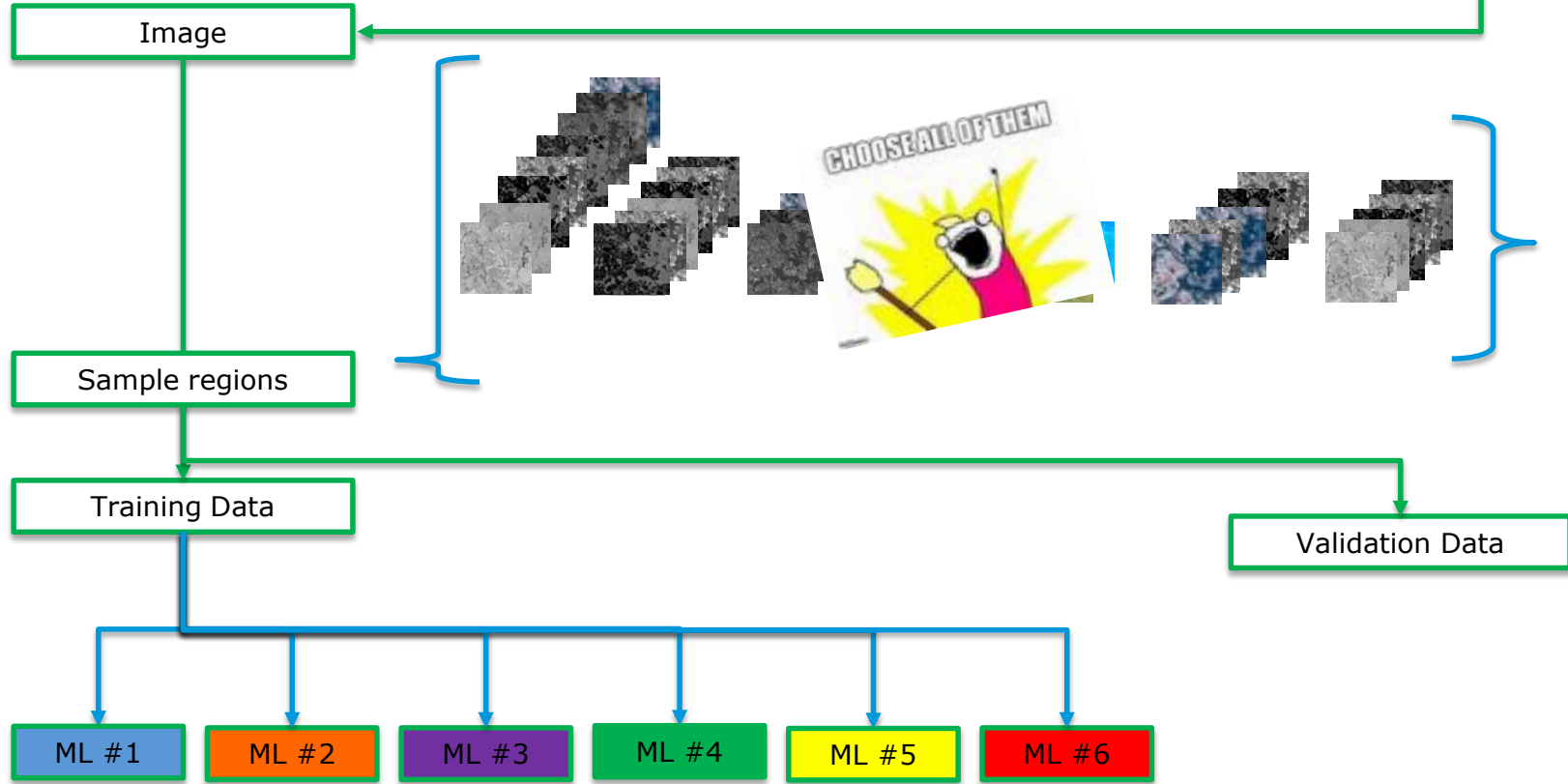


# Workflow – data processing & preparation



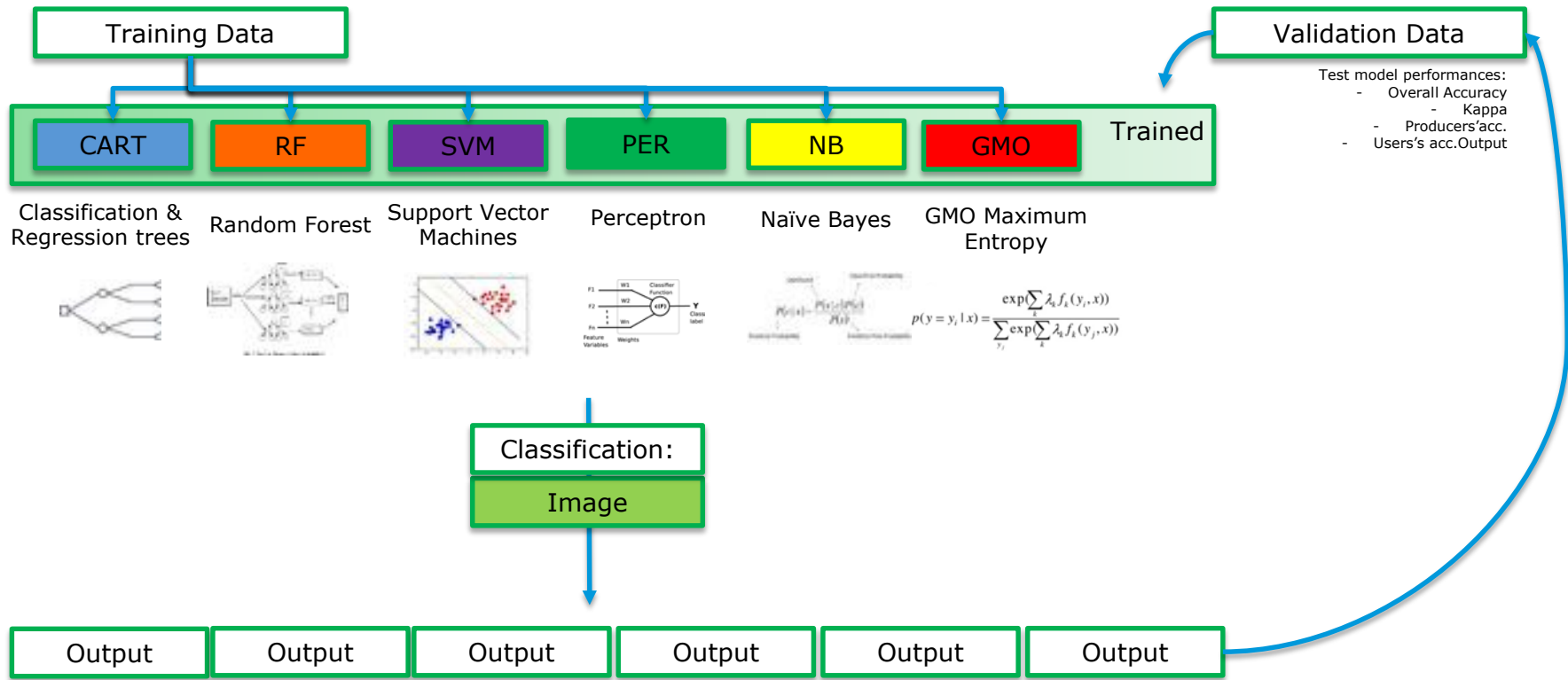



# Workflow – input data





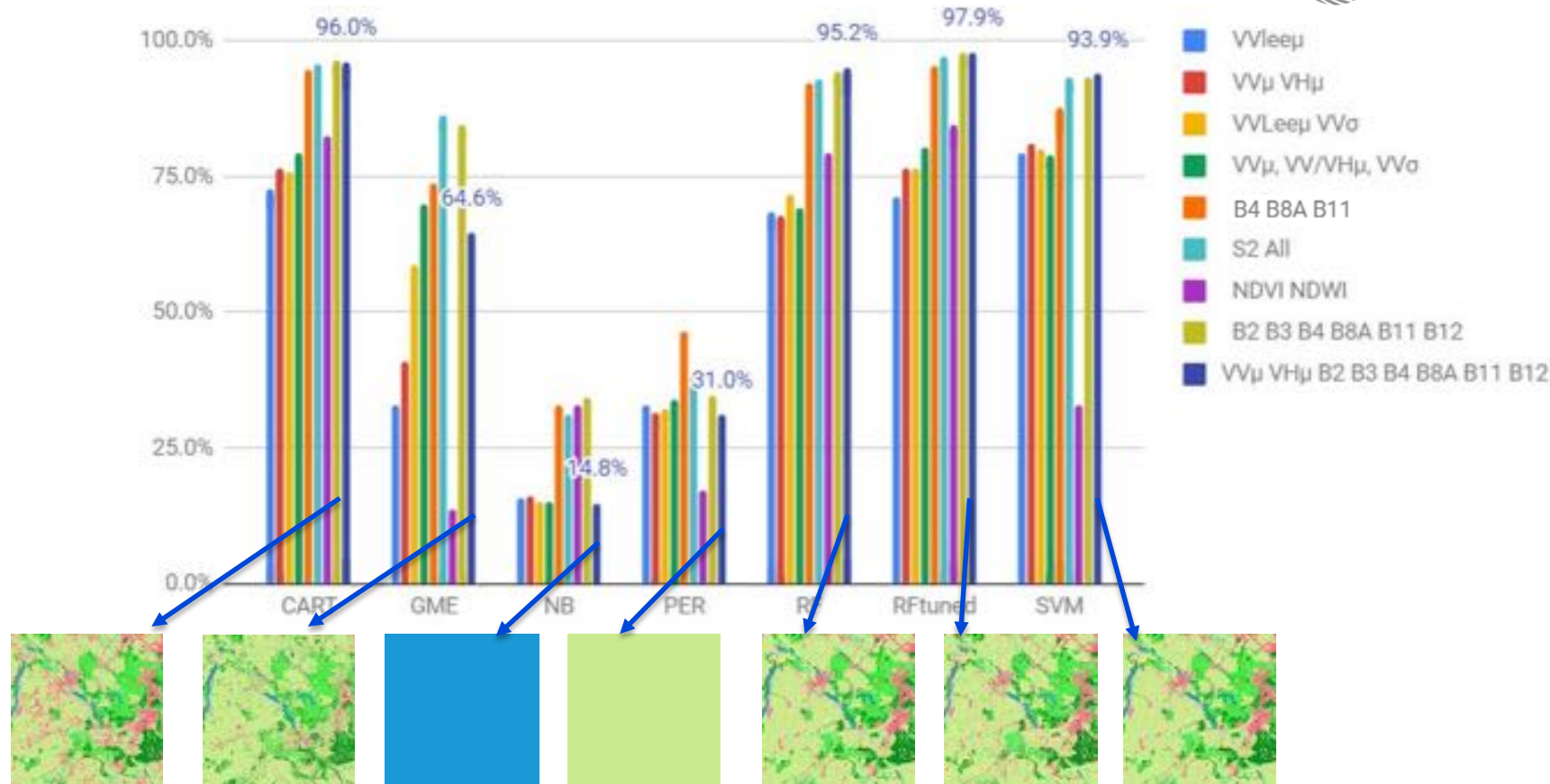
# Workflow – classification



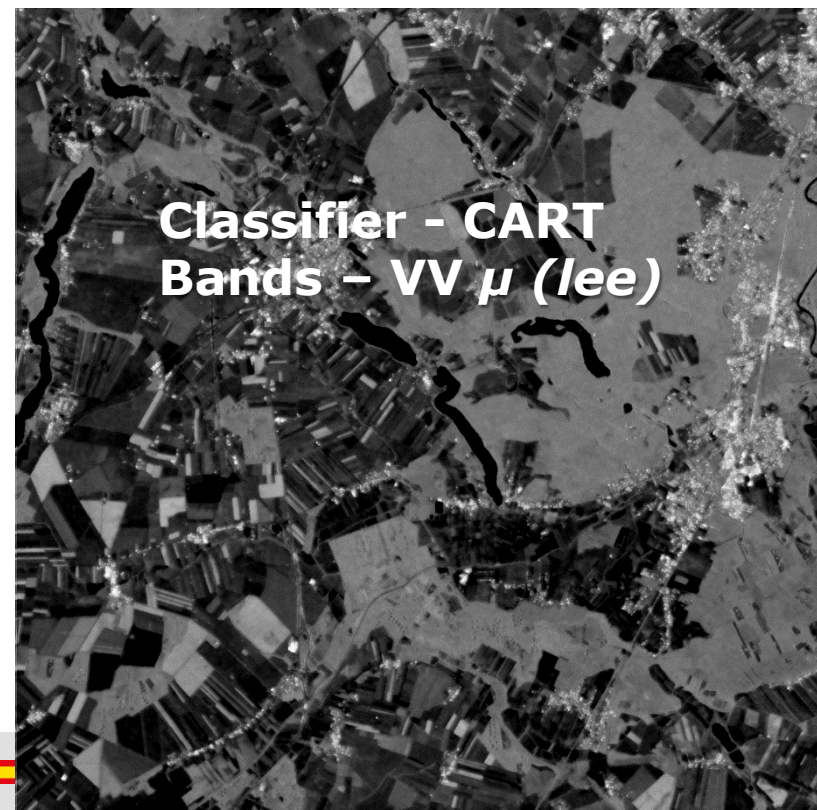
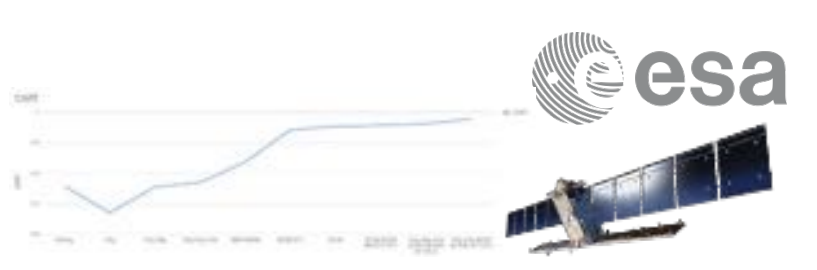
A satellite view of Earth showing a mix of blue oceans, white clouds, and brownish-green landmasses. A dark, rounded rectangular box with a white border is centered horizontally across the middle of the image.

# **RESULTS OVERVIEW**

# Machine Learning performances – a first overview!

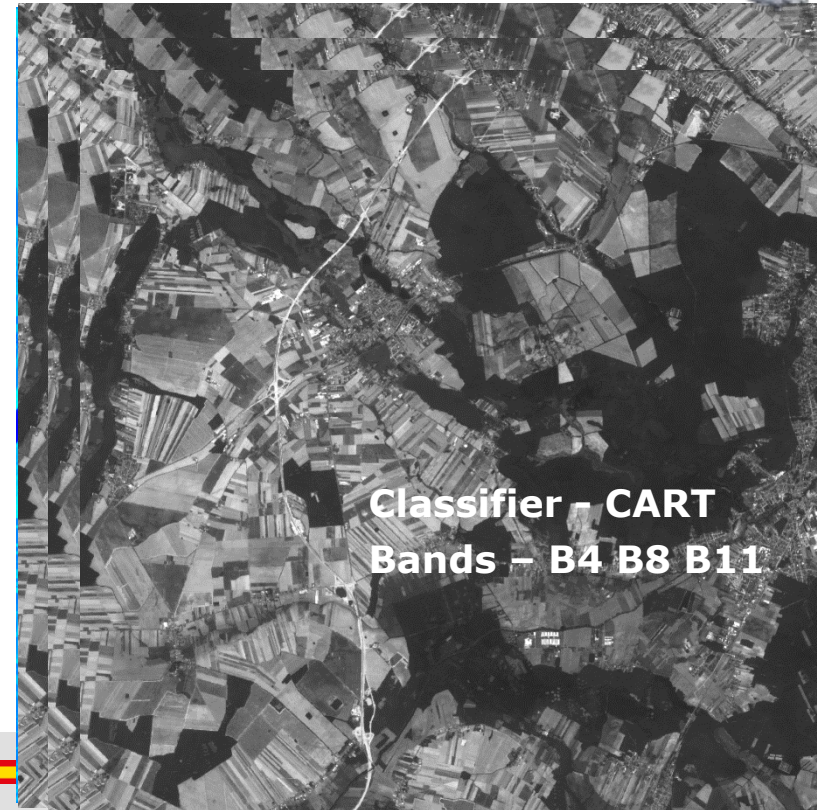
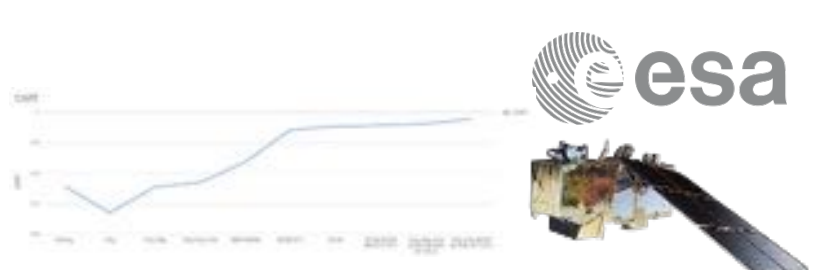
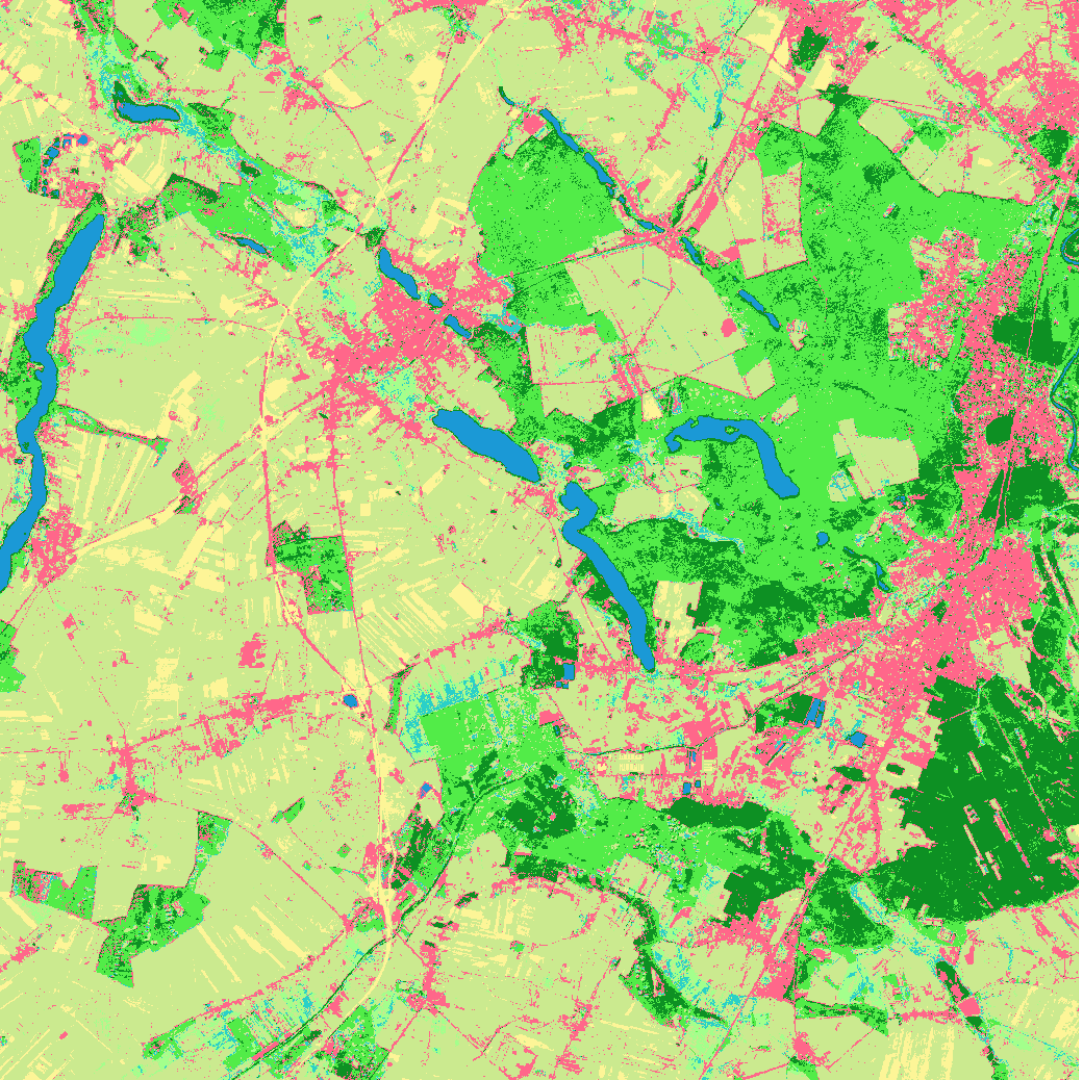






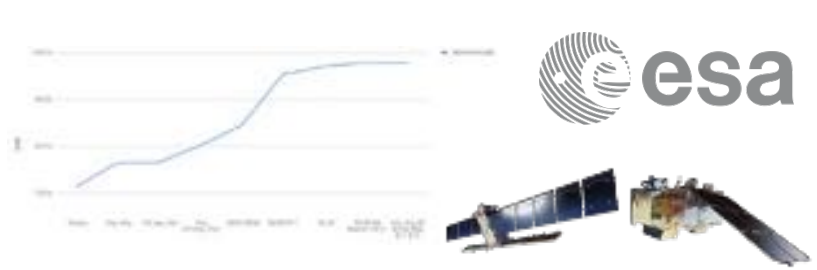
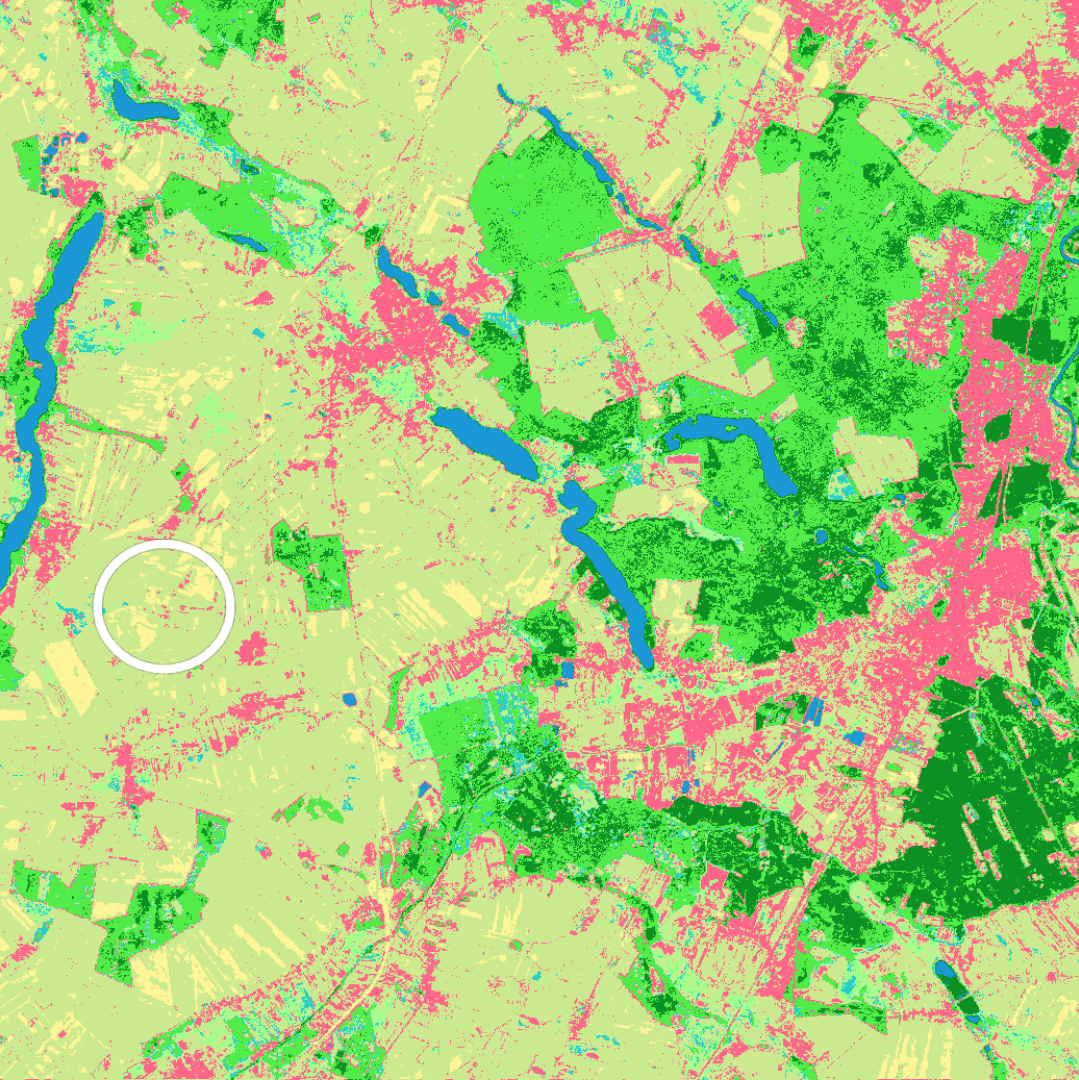
**Classifier - CART**  
**Bands - VV  $\mu$  (lee)**





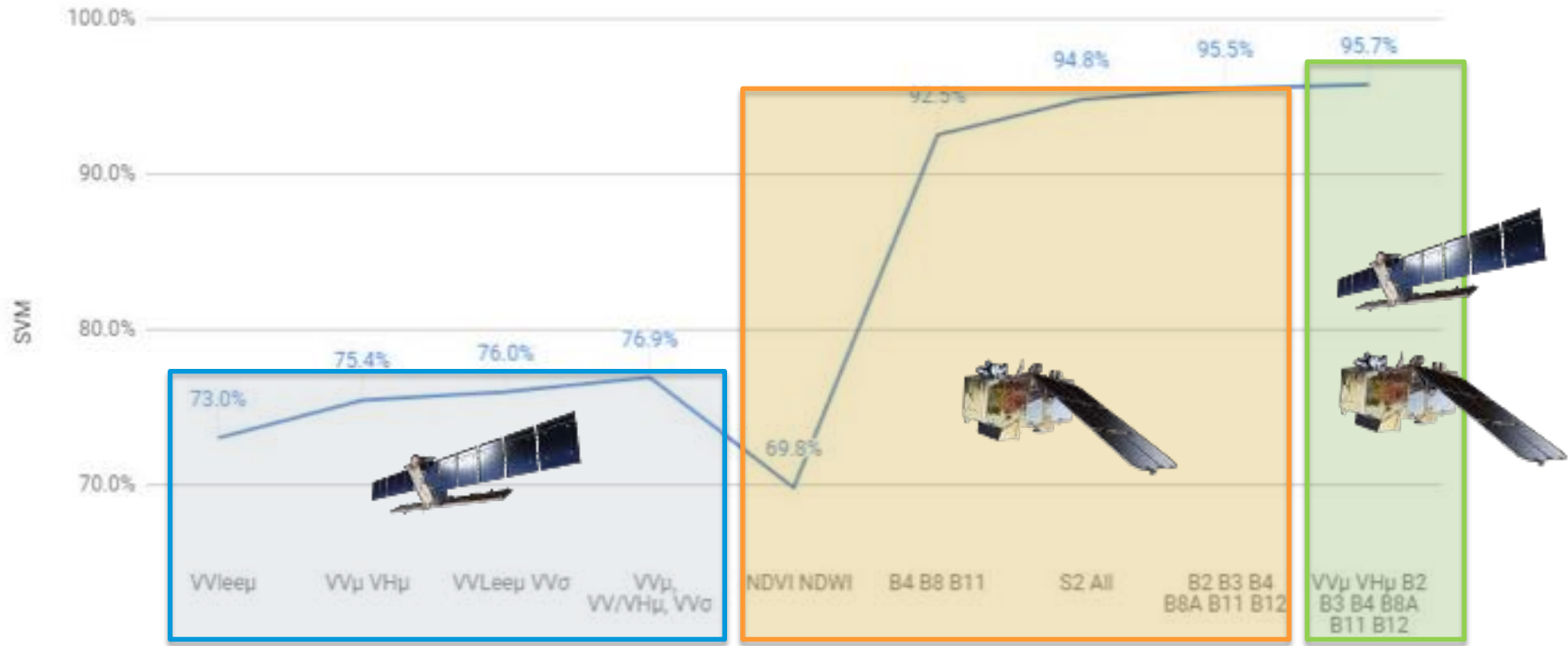
**Classifier - CART**  
**Bands - B4 B8 B11**





# Behaviour to data input

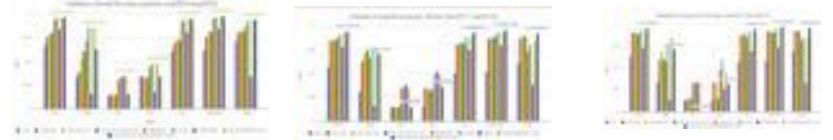
Averaged from best performing models



# Improving model performances...



**Changing temporal coverage**  
- updating land-cover maps

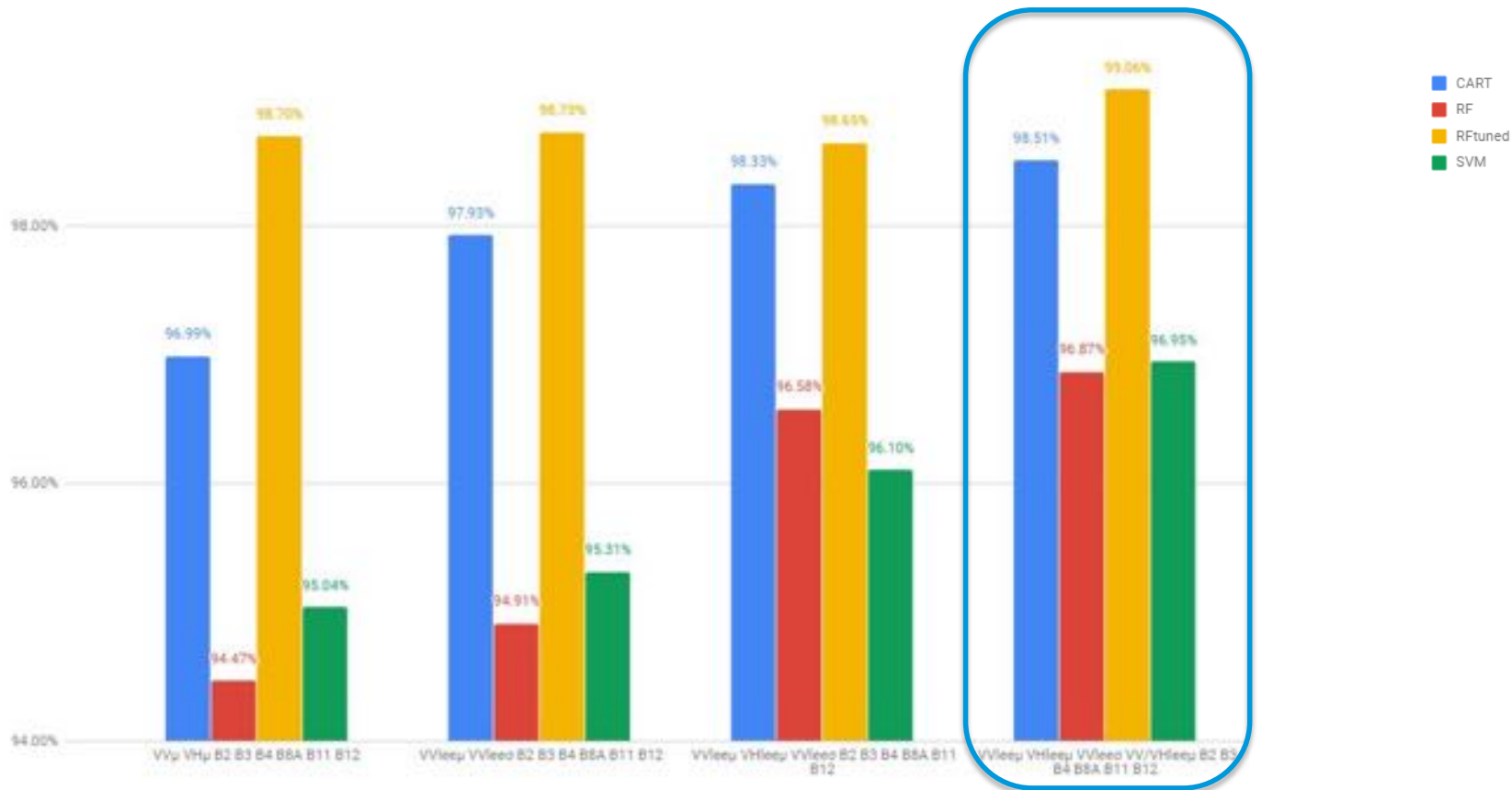


**Changing training data**  
In size or coverage

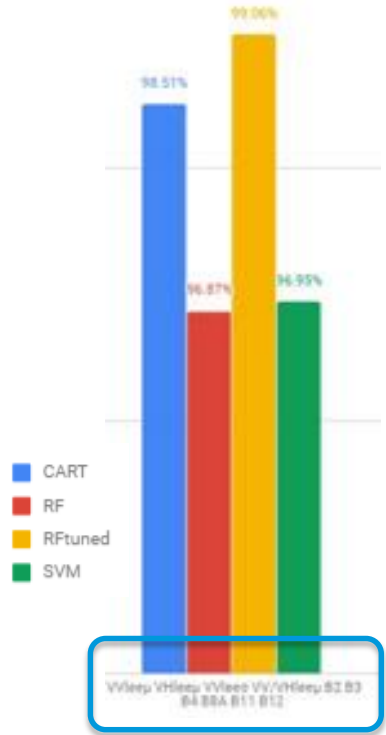
**Trying new band selection**  
Different combinations of S1+S2



# Best performing models with S1+S2 variations



Other acc



Random Forest (Tuned)

CART

Support Vector Machines

Random Forest



# Further accuracy metrics...

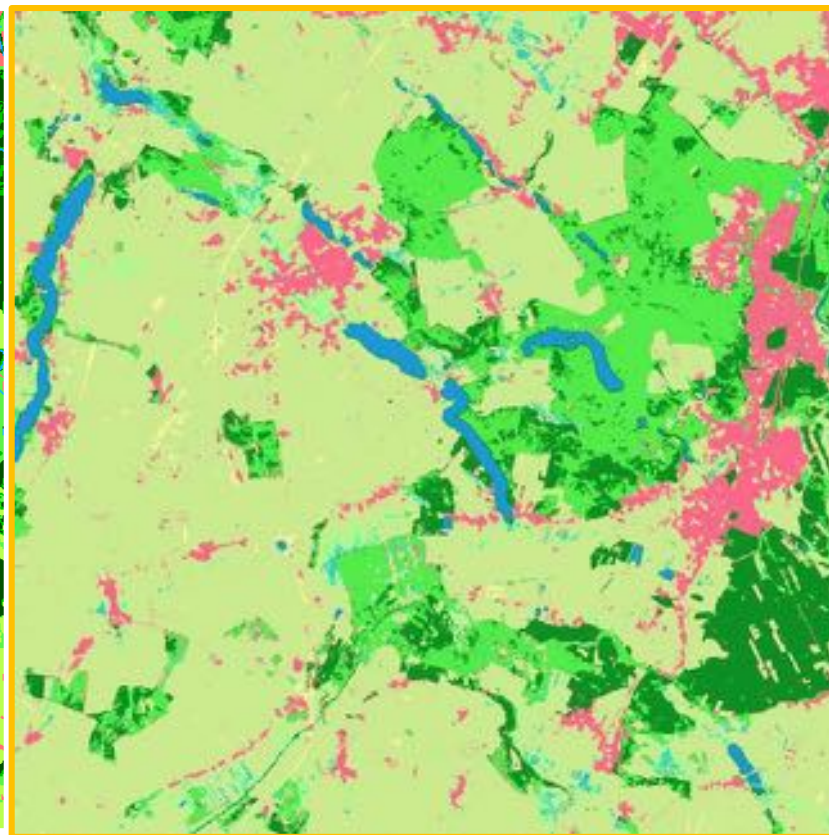
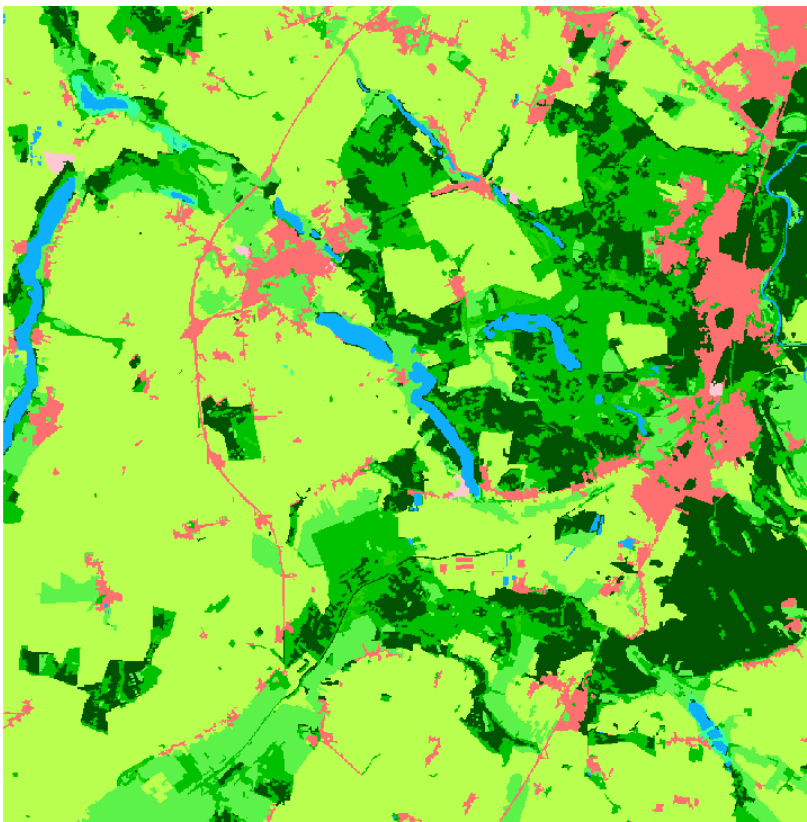
	Reference									
Classified	Water body	Wetland	Urban Fabric	Agriculture	Coniferous Forest	Mixed Forest	Grassland	Bare soil	PA	UA
Water body	193	0	0	0	0	0	0	0	1.000	1.00
Wetland	0	29	0	1	0	0	0	0	0.967	0.94
Urban Fabric	0	0	239	2	0	0	0	0	0.992	1.00
Agriculture	0	1	0	428	0	1	0	0	0.995	0.99
Coniferous Forest	0	0	0	0	154	2	0	0	0.987	0.99
Mixed Forest	0	0	0	0	2	157	0	0	0.987	0.98
Grassland	0	1	0	1	0	1	60	0	0.952	1.00
Bare soil	0	0	0	0	0	0	0	5	1.000	1.00

Kappa  
**0.9881**

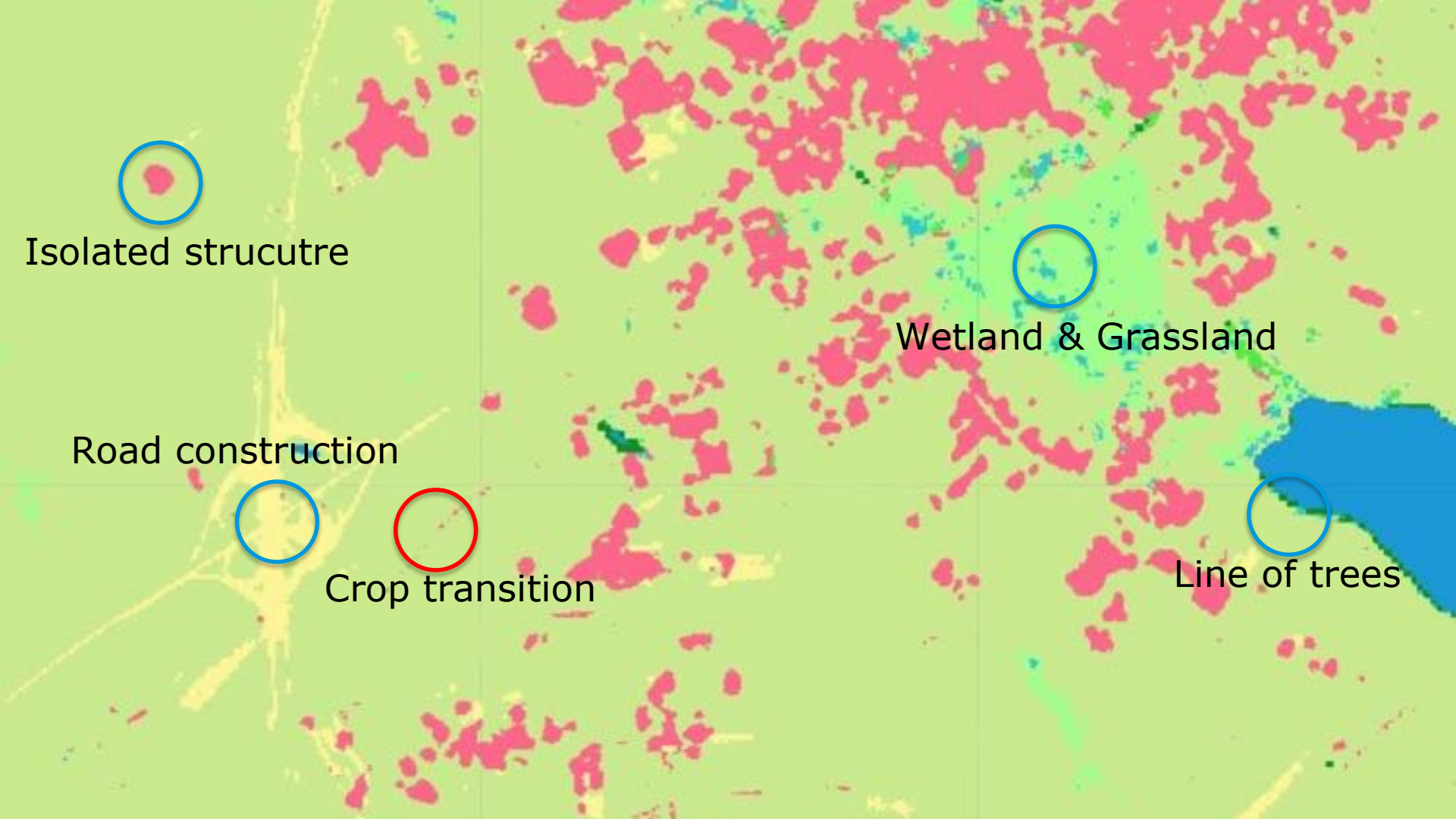
Overall Accuracy  
**0.9906**

# Ground Truth/Reference VS Final model

- waterbody
- wetland
- urbanfabric
- agriculture
- coniferousforest
- mixedforest
- grassland
- baresoil







Isolated structure

Road construction

Crop transition

Wetland & Grassland

Line of trees

# Limitations...bear in mind that....

**So....** Is this then the best model?

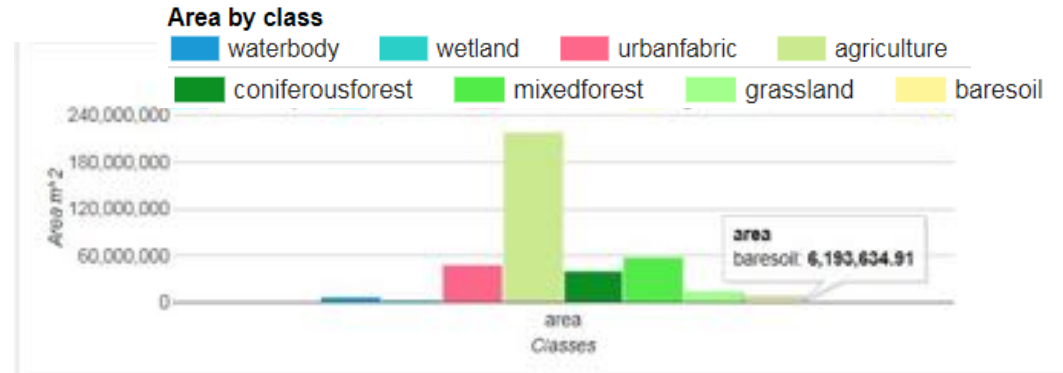
- Training data (plays a **huge** role) affecting:
  - Models performances
  - Smaller size – the more data input...not always is the best
  - Bigger size – more difference in 'additive power' but models behaved + similarly
- Tuning hyperparameters of ML are **done manually**
  - Some models were inserted by default & data was not normalized

A photograph of a rocket launch at night. The rocket is positioned vertically, surrounded by a large plume of bright yellow and orange fire and smoke. Two tall, slender towers are visible on either side of the rocket. The background is dark, and the overall scene is illuminated by the intense light of the launch. A black rounded rectangular box with a white border is overlaid on the image, containing the text "What potential...?".

**What potential...?**

# Example of straightforward applications...

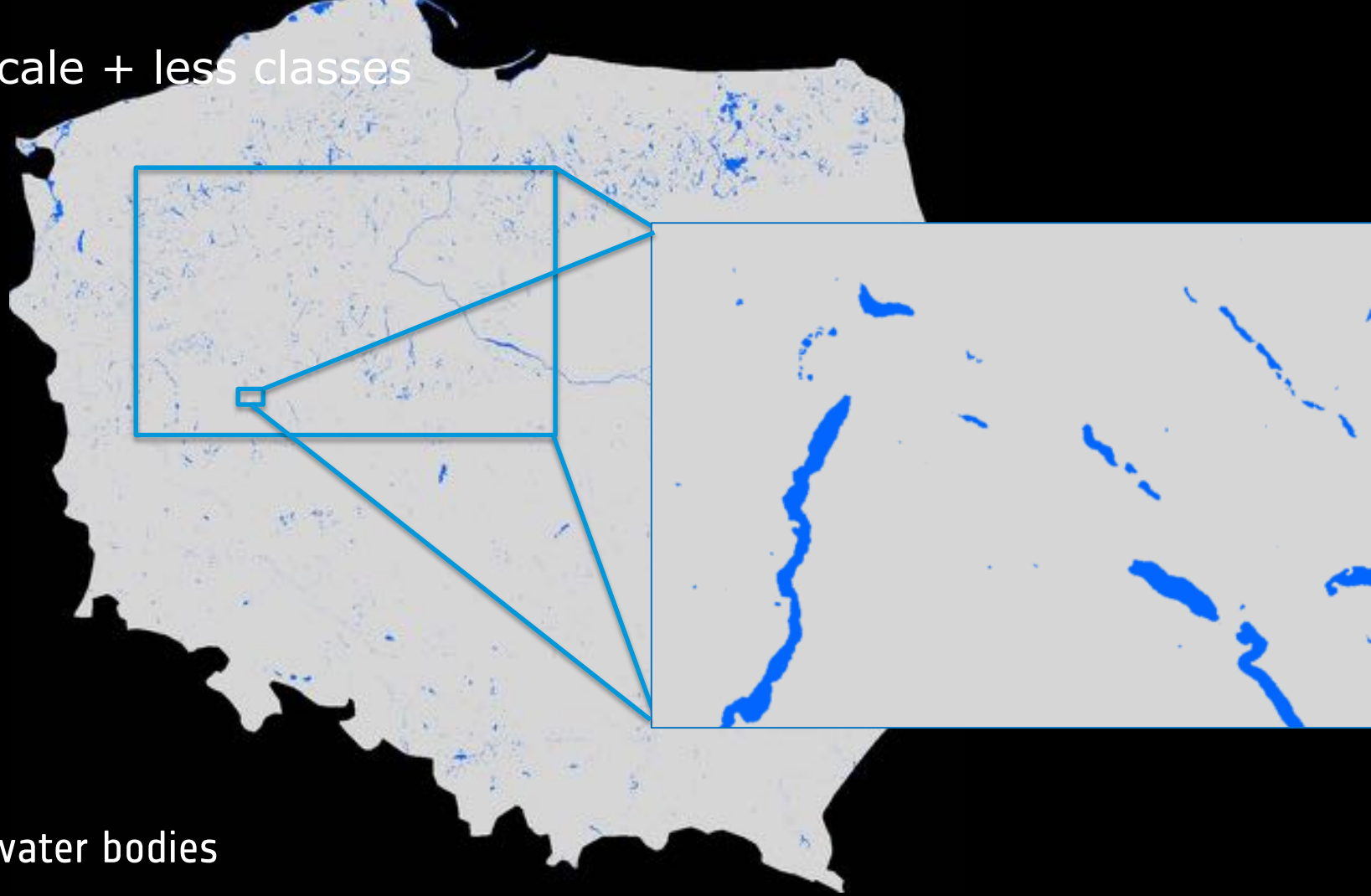
- Analysis of temporal land use and land cover change...



**Attention!**  
This graph moves 😊

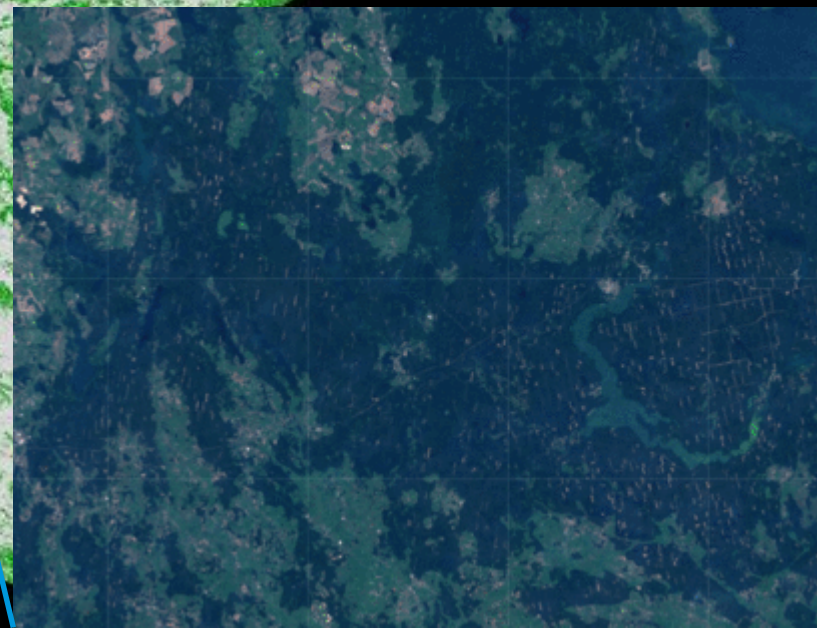
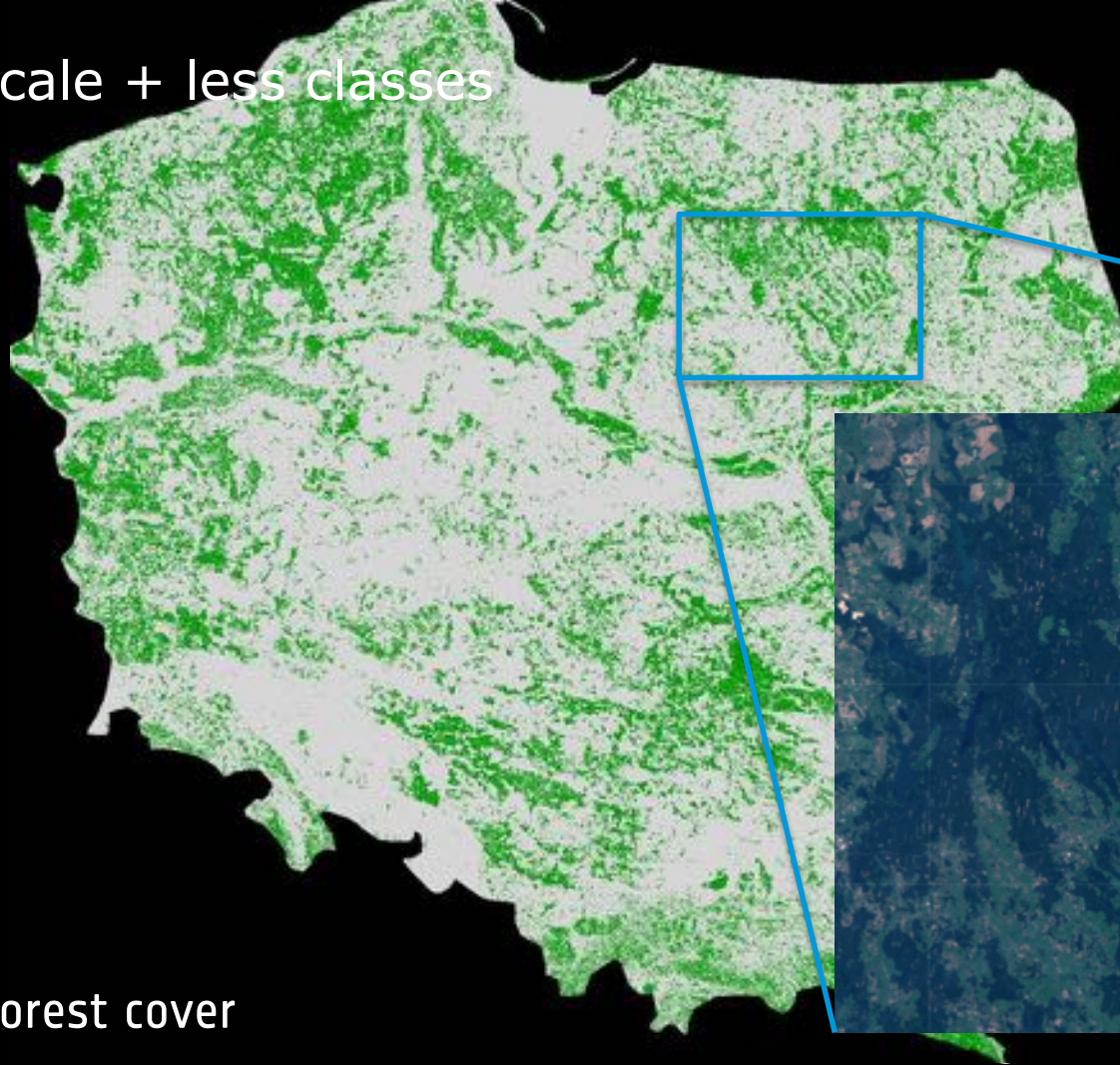


Larger scale + less classes



Mapping water bodies

Larger scale + less classes

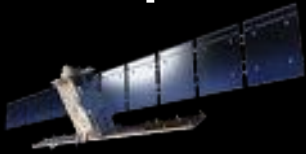


Mapping forest cover

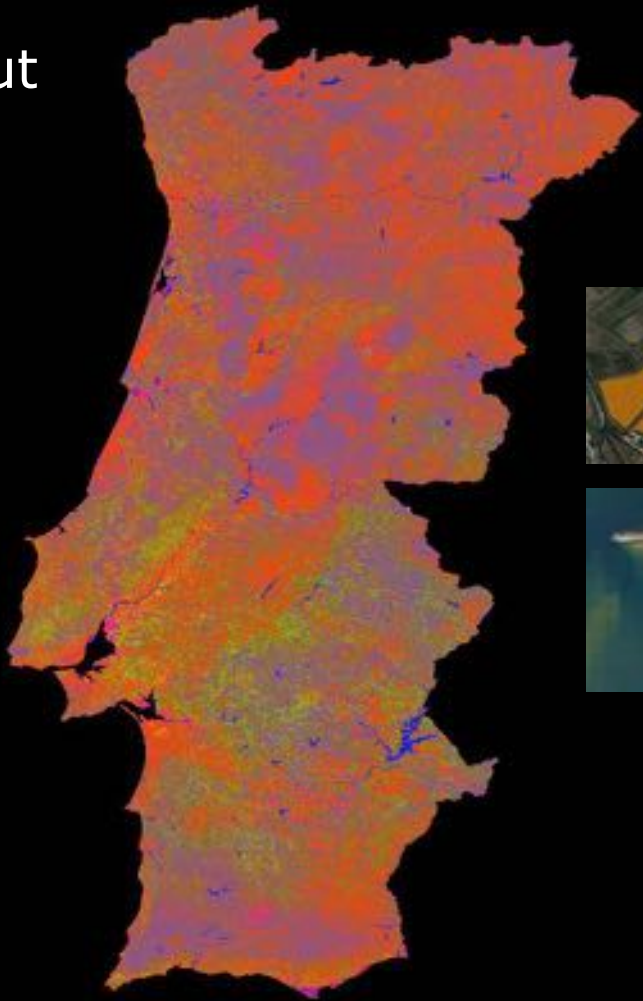
Large scale + data input

**Elevation  
(DEM STRM)**

+



+





# Large scale + data input + classes

CORIN LAND COVER (100m)



**S1 + S2 = + accurate + frequent + higher resolution**



# Overall Conclusions



## Google Earth Engine

- GEE offers **powerful** capabilities in **handling large volumes** of remote sensing imagery
- GEE contains state-of-art machine learning algorithms achieving high accuracies and excellent tool for rapidly prototype AI applications.
- A big limitation is the need of manually tune the machine learning algorithms

## Added value of fusing Sentinel data

- The integration of texture and spectral information for pixel-based classification improves classification accuracy (S2 outperforms S1 alone but together detect finer structures).
- Data: the more, ~~the merrier!~~

## Large scale mapping – further work?

- Results can be used to calculate/estimate use cover and land change dynamics
- Normalize data and test more combinations!



An aerial photograph of a rugged mountain range covered in snow. The terrain is characterized by sharp peaks and deep shadows, creating a high-contrast scene. A dark, rounded rectangular box is superimposed over the center of the image, containing the text "Thank you for your attention!".

**Thank you for your attention!**