Satellite Derived Marine Habitat Mapping

Observing underwater habitats through open source software

4EI

Agenda

4E1

4EI

> "

4EI use space data for the betterment of humans and the planet

4EI is **global company** with headquarters in the Middle East and Europe.

We are built on **decades of expertise** in Earth Observation, GIS and remote sensing analytics, as well as the creation of **strong partnerships**.

Our diverse and **passionate team** is the driving force of our success, helping us to always **guarantee trust** and **quality** in everything we do.

4EI capabilities

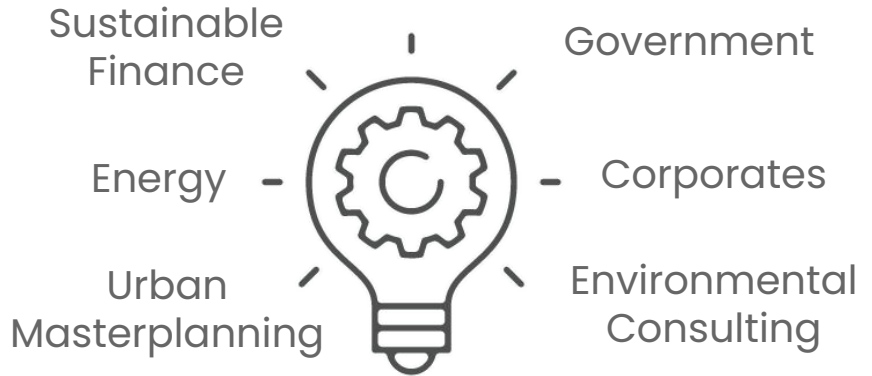**4EI**

REMOTE SENSING

EARTH OBSERVATION

AI & MACHINE LEARNING

GEOSPATIAL MAPPING

DIGITAL EXPERTISE

MOBILE & SOFTWARE

=

Sustainable
Finance

Government

Energy

Corporates

Urban
Masterplanning

Environmental
Consulting

PRODUCT
INNOVATION

Prerequisite software

Anaconda3 (https://www.anaconda.com/download)

Acolite (https://odnature.naturalsciences.be/remsem/software-and-data/acolite)

Git (https://gitforwindows.org/)

Orfeo Toolbox (https://www.orfeo-toolbox.org/)

QGIS (https://www.qgis.org/en/site/)

SNAP - Science Toolbox Exploitation Platform (https://step.esa.int/main/download/snap-download/)

Sklearn package (https://scikit-learn.org/stable/install.html)

Near Shore Bathymetry Plugin (https://github.com/balajiceg/NearShoreBathymetryPlugin)
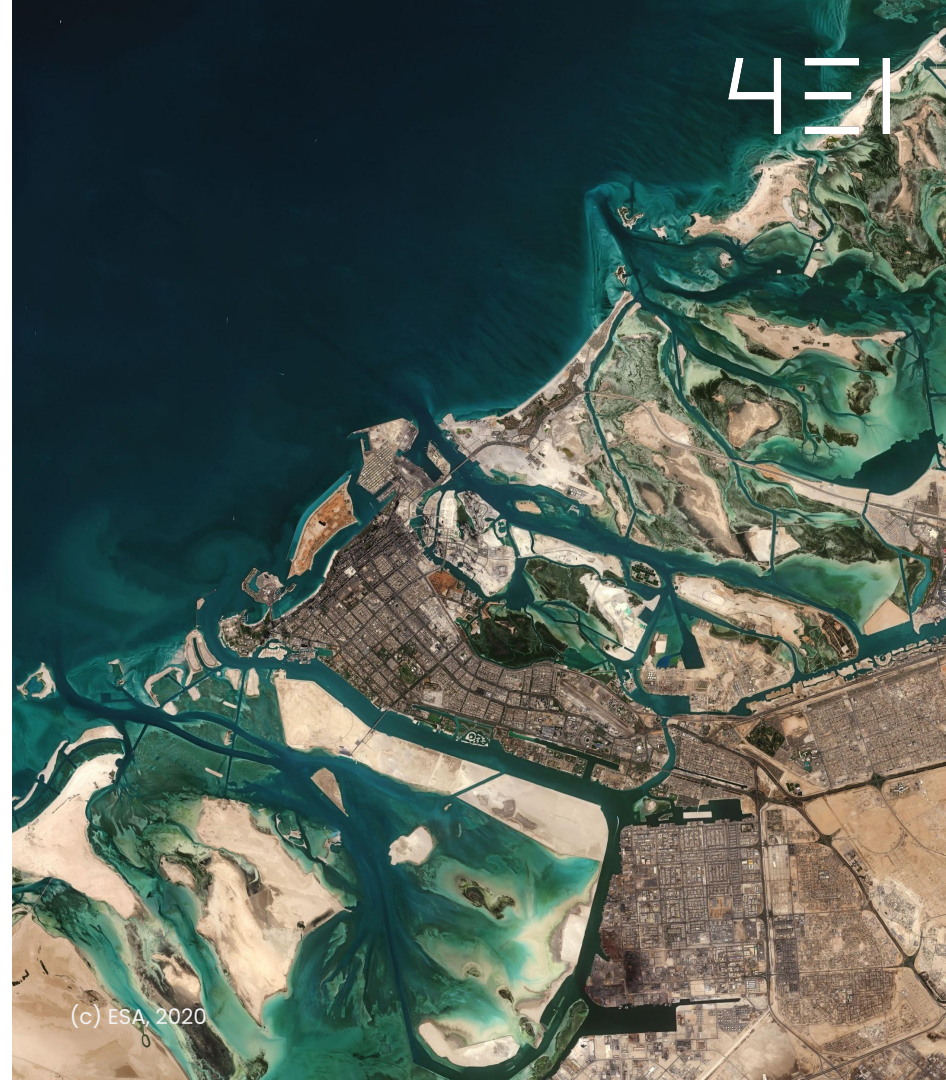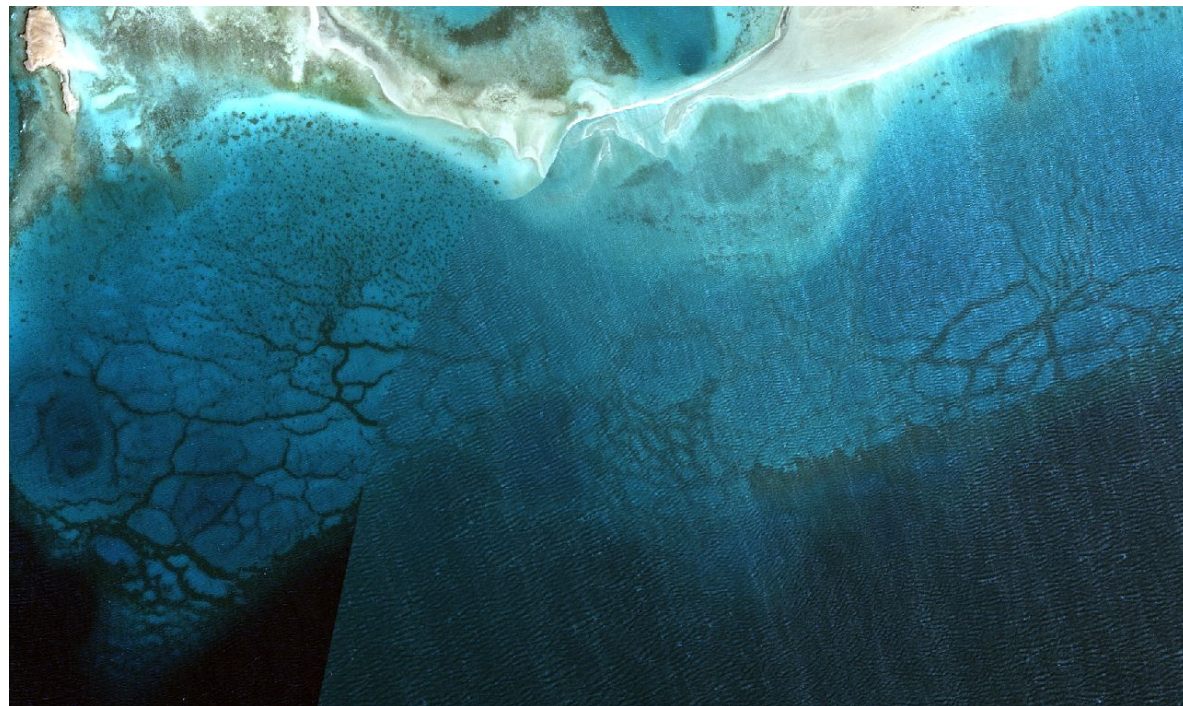
Workflow
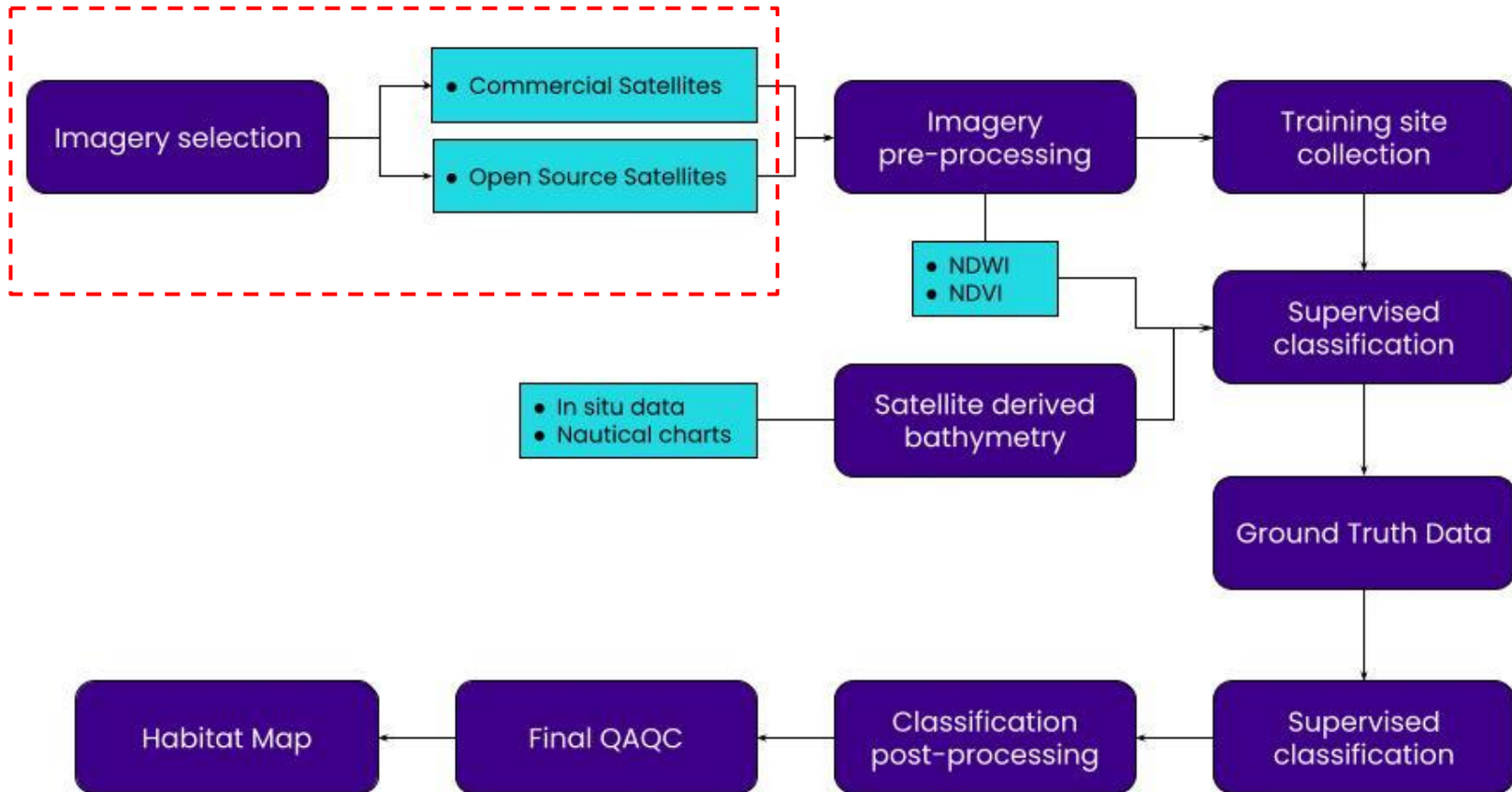
Imagery Selection

The images were inspected by a remote sensing analyst to ensure that it was of **high enough quality** for **habitat mapping**. This included several quality control and assurance steps to confirm the selected imagery conformed to the **following parameters**:

a) Coverage

b) Date

c) Cloud Cover

d) Waves

e) Turbidity

f) Sunglint
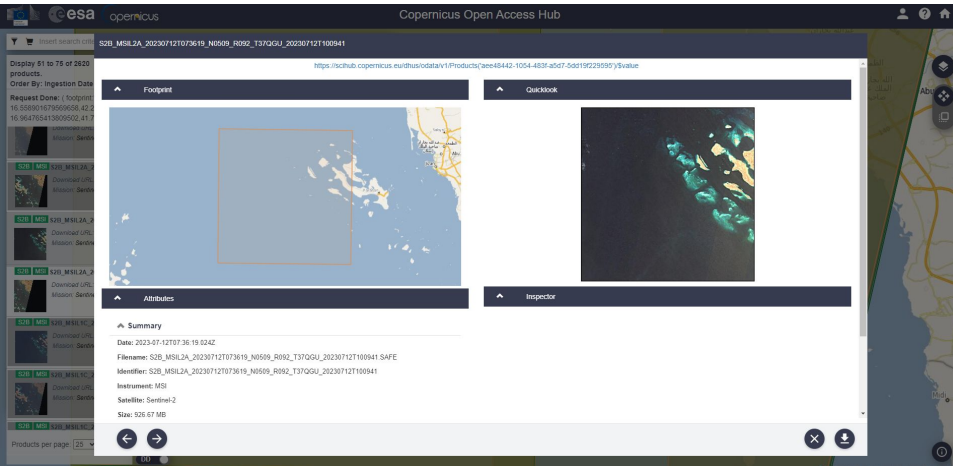
Imagery Assessment

Airbus Pleiades

NASA images courtesy LANCE/EOSDIS MODIS Rapid Response Team, GSFC.

NASA Earth Observatory image by Joshua Stevens, using MODIS data from LANCE/EOSDIS Rapid Response
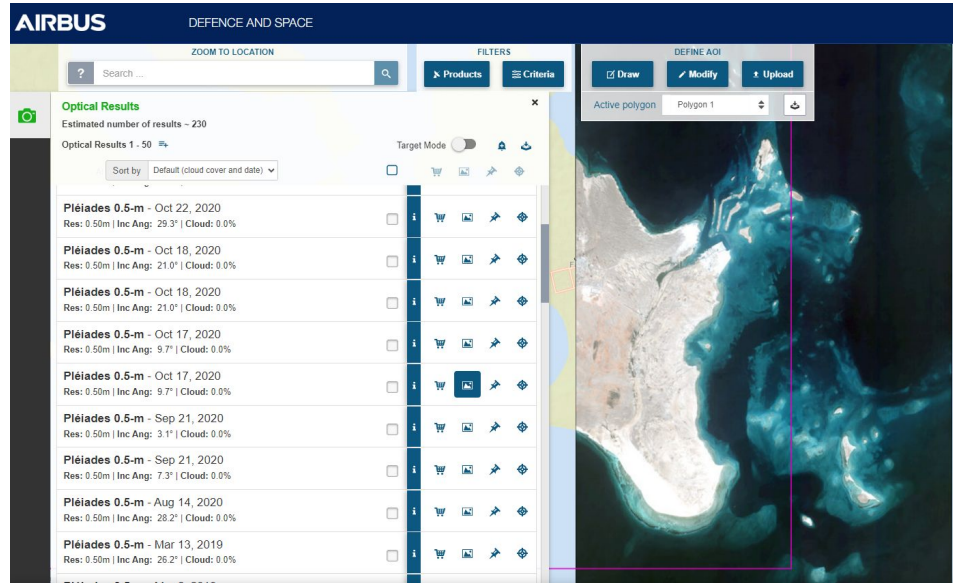
Satellite Imagery Options
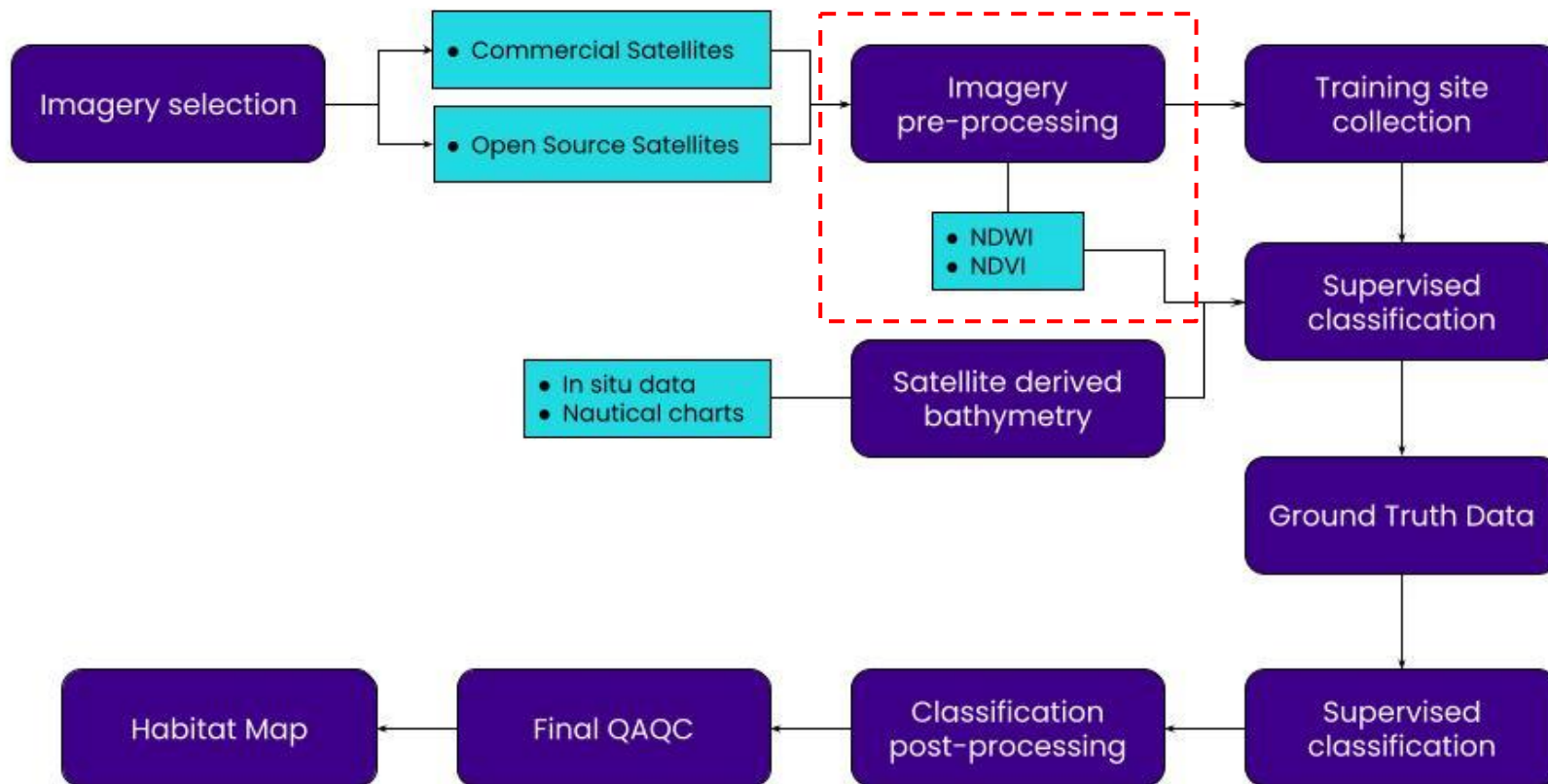

Copernicus Open Access Hub
https://scihub.copernicus.eu/dhus/#/home

Airbus
https://www.intelligence-airbusds.com/en/4871-ordering

Workflow

Imagery Processing

4EI performs orthorectified with atmospheric compensation. This ensures that the atmosphere did not contribute to any of the habitat classifications
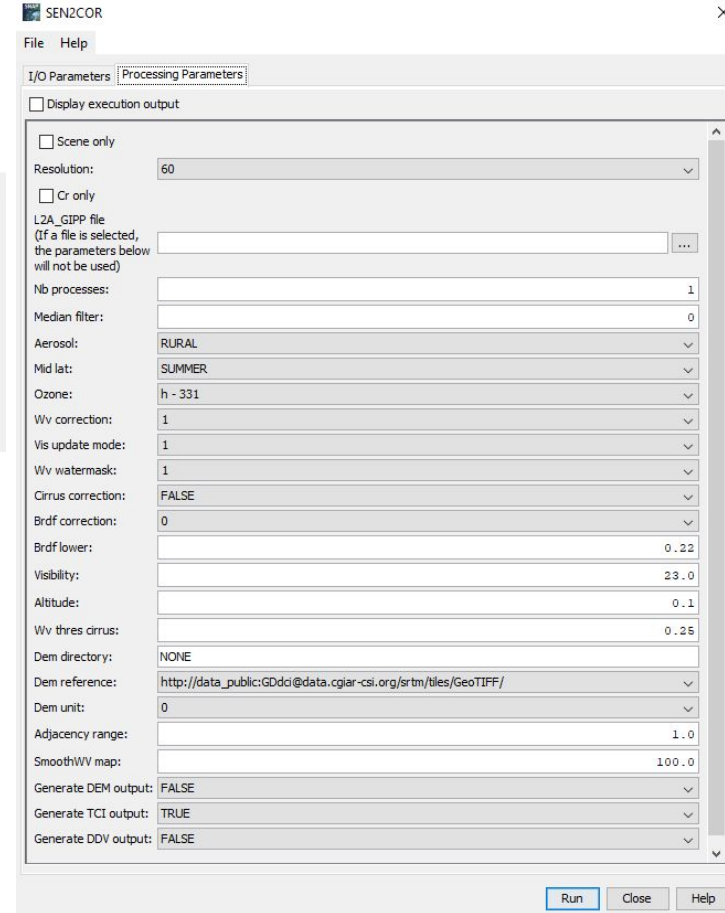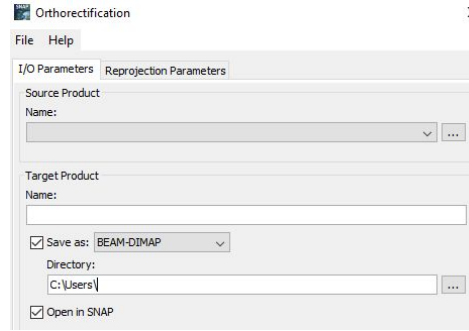
- Orthorectification

Sen2Cor plugin on the SNAP software.

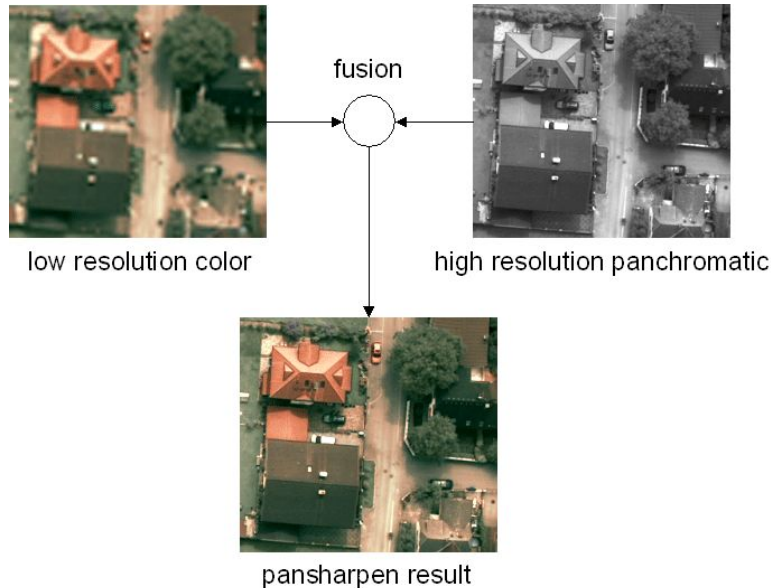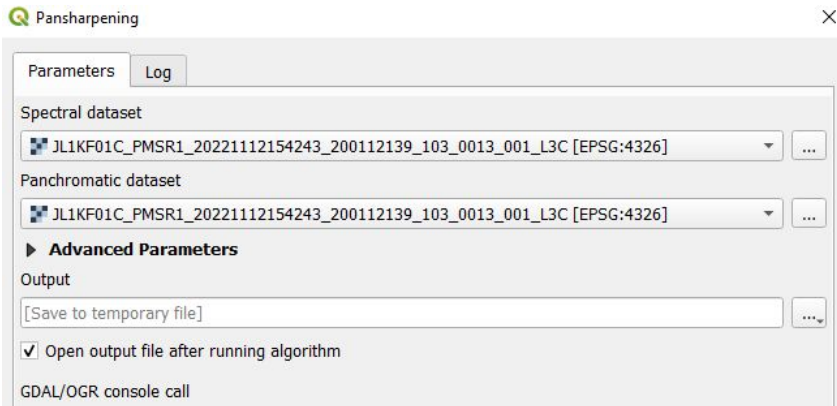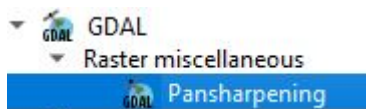- Atmospheric Correction

Sen2Cor plugin on the SNAP software.

Acolite: generic processor developed at RBINS for atmospheric correction and processing for coastal and inland water applications.

Pansharpening

- Imagery pansharpening

This process refers to the creation of a single high resolution colour image by merging the low resolution multispectral imagery with the higher resolution panchromatic imagery (Ahmad, 2021).


fusion

low resolution color

high resolution panchromatic

pansharpen result

Roland Perko, Computer Vision for Large Format Digital Aerial Cameras

4ΞI


```
GDAL
  Raster miscellaneous
    Pansharpening
```

Pansharpening

Parameters   Log

Spectral dataset
JL1KF01C_PMSR1_20221112154243_200112139_103_0013_001_L3C [EPSG:4326]

Panchromatic dataset
JL1KF01C_PMSR1_20221112154243_200112139_103_0013_001_L3C [EPSG:4326]

▶ **Advanced Parameters**

Output
[Save to temporary file]

✔ Open output file after running algorithm

GDAL/OGR console call

Satellite Imagery Indices
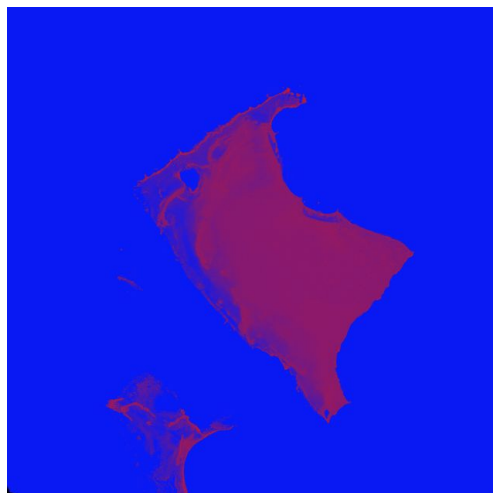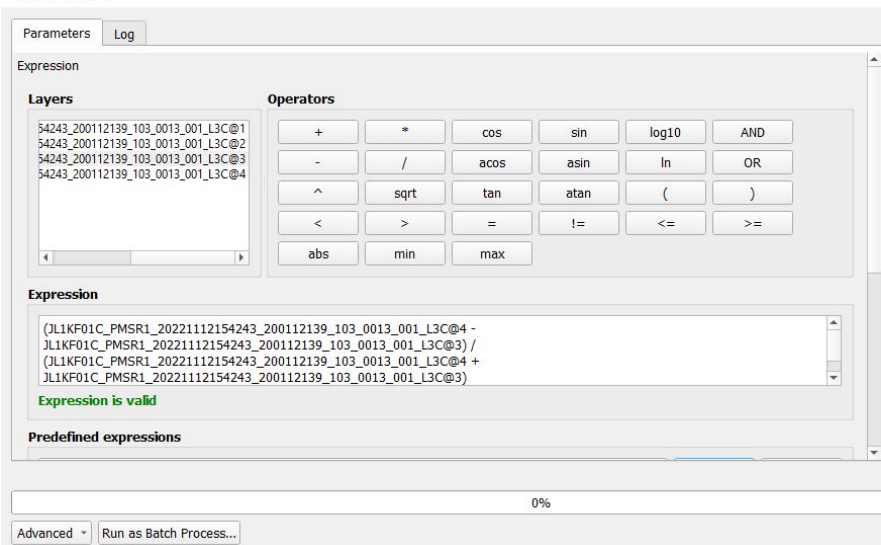
- <u>Index thresholding</u>

This pre-processing technique provides the supervised classification algorithm with additional detail, allowing for a more accurate habitat classification. Two spectral indices were derived and are explored below.
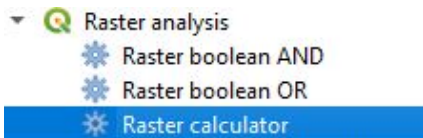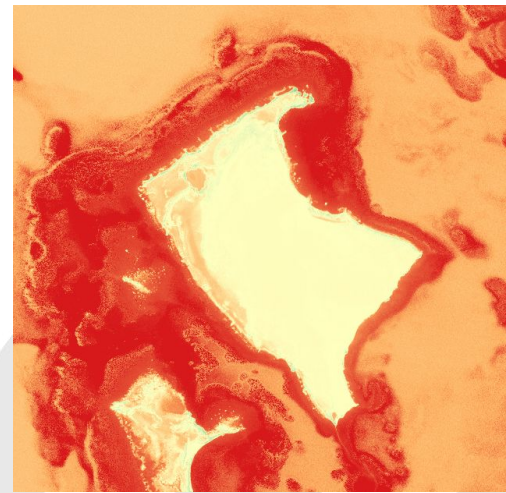
$$NDWI = \frac{(Green - Near\ Infrared)}{(Green + Near\ Infrared)} \qquad NDVI = \frac{(Near\ Infrared - Red)}{(Near\ Infrared + Red)}$$

NDWI

NDVI

SDB (optional)

Bathymetry **helps with identifying specific habitats** and similarly to index thresholding, this technique provides the supervised classification algorithm with additional detail, allowing for a more accurate habitat classification.

The bathymetry layer produced is a relative layer since it does not depict absolute depths

- Near Shore Bathymetry Plugin (https://github.com/balajiceg/NearShoreBathymetryPlugin)

It uses the **Log Ratio Transformation method**, using a bottom albedo-independent Bathymetry algorithm developed by Stumpf and Holderied (2003).
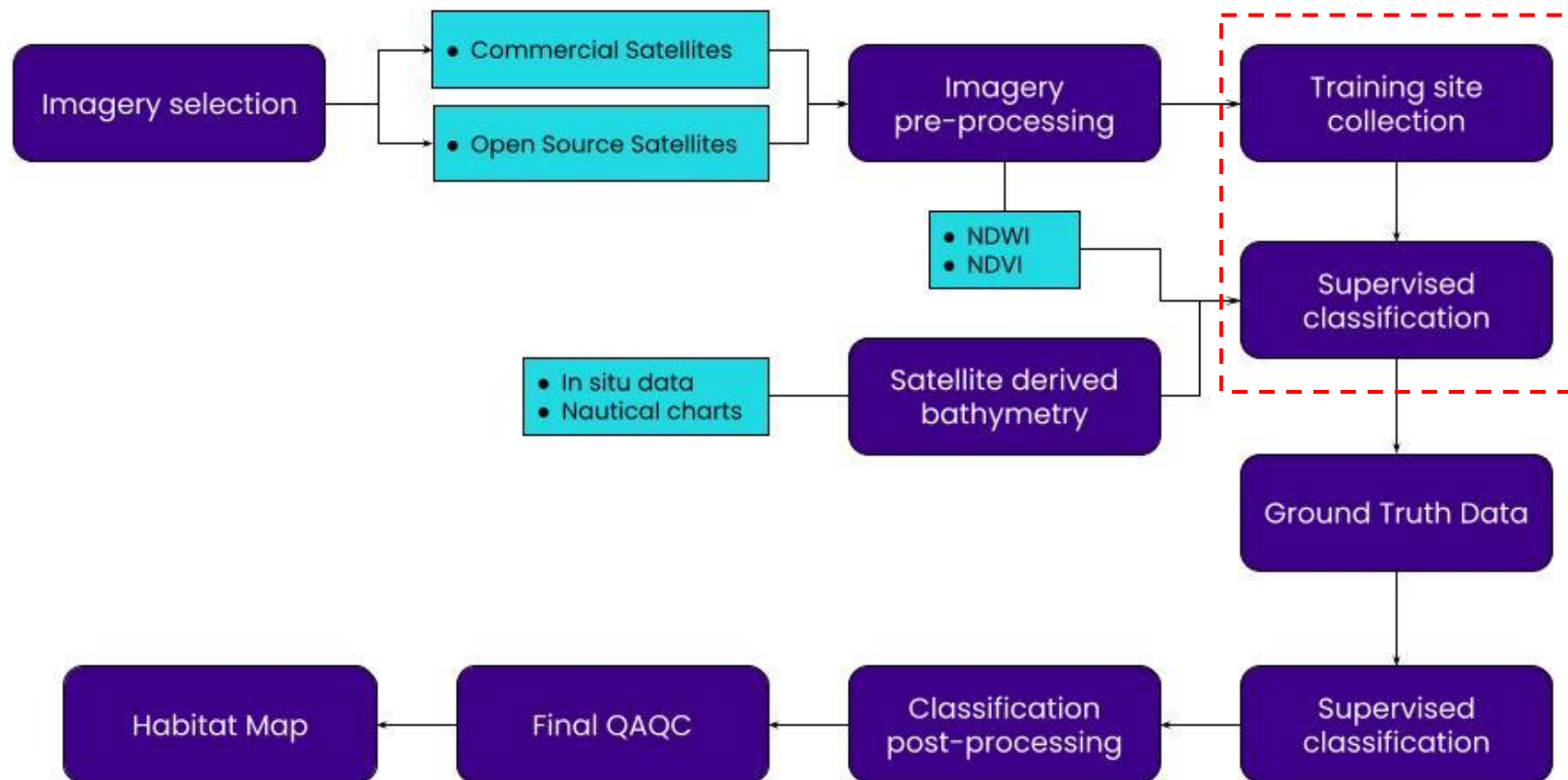
This bathymetry is reasonable only **between depths of 2 to 20m**. This tool is similar to the Relative Water Depth tool in ENVI.

Data required for the tool are:

- Blue & Green Bands -for computing the relative depths.
- OPTIONAL: Red, NIR, SWIR bands - for creating mask to extract water body.
- Ground truth: actual depth information data.
- MetaDataFile(MLT) OPTIONAL: for converting DN value to TOA.

(c) ESA, 2004

4EI

Workflow

4ΞI

Imagery selection → Commercial Satellites / Open Source Satellites → Imagery pre-processing → NDWI, NDVI → Training site collection → Supervised classification

In situ data, Nautical charts → Satellite derived bathymetry

Training site collection → Supervised classification → Ground Truth Data → Supervised classification → Classification post-processing → Final QAQC → Habitat Map

Training Data

4EI reviews the satellite imagery and collected polygons (training areas) over identifiable classes in order to train the classification model. Training sites are areas in an image that are representative of each of the land cover classes that are to be defined.


(c) ESA, 2018
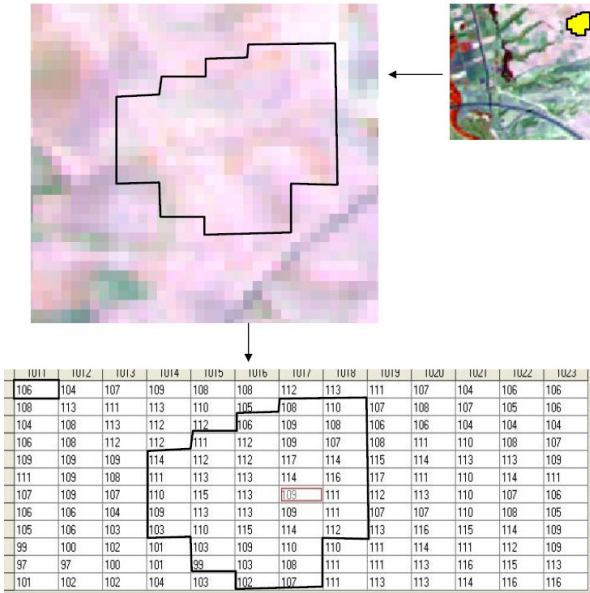
Supervised Classif.

A **random forest classification algorithm** is used to produce the draft marine habitat map.

Random forest is a supervised learning algorithm. The algorithm builds **a number of decision trees** ("The Forest"). Each decision tree in the forest considers a random subset of features when forming "questions" and only has access to a random set of the training data points.

This increases diversity in the forest leading to **more robust overall predictions** and the name 'random forest'. When it comes time to make a prediction, the random forest will take a **majority vote for the predicted class**. The final output of the algorithm is a raw classified raster.
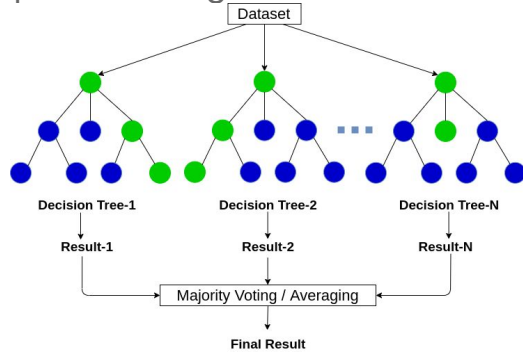
Source: Random Forest vs Decision Tree | Which Is Right for You?

Random Forest Parameters

4ΞI

- <u>Number of Trees</u>

By building forests with a **large number** of trees we can create a **more robust** aggregate model with less variance, at the cost of a **greater training time**.

- <u>Node criteria split method</u>

They do this by using a specific metric (a **measure of impurity**) which for classification is **Gini** or **Entropy**.

- <u>Maximum Depth of the Individual Trees</u>

The **deeper the tree, the more splits** it has and the more information about the data it takes into account.

- <u>Number of random features to consider at each split</u>

A **small value** (less features considered when splitting at each node) will **reduce the overall variance**, at the cost of higher individual tree bias. The value should be set accordingly to how many informative or quality features you have.
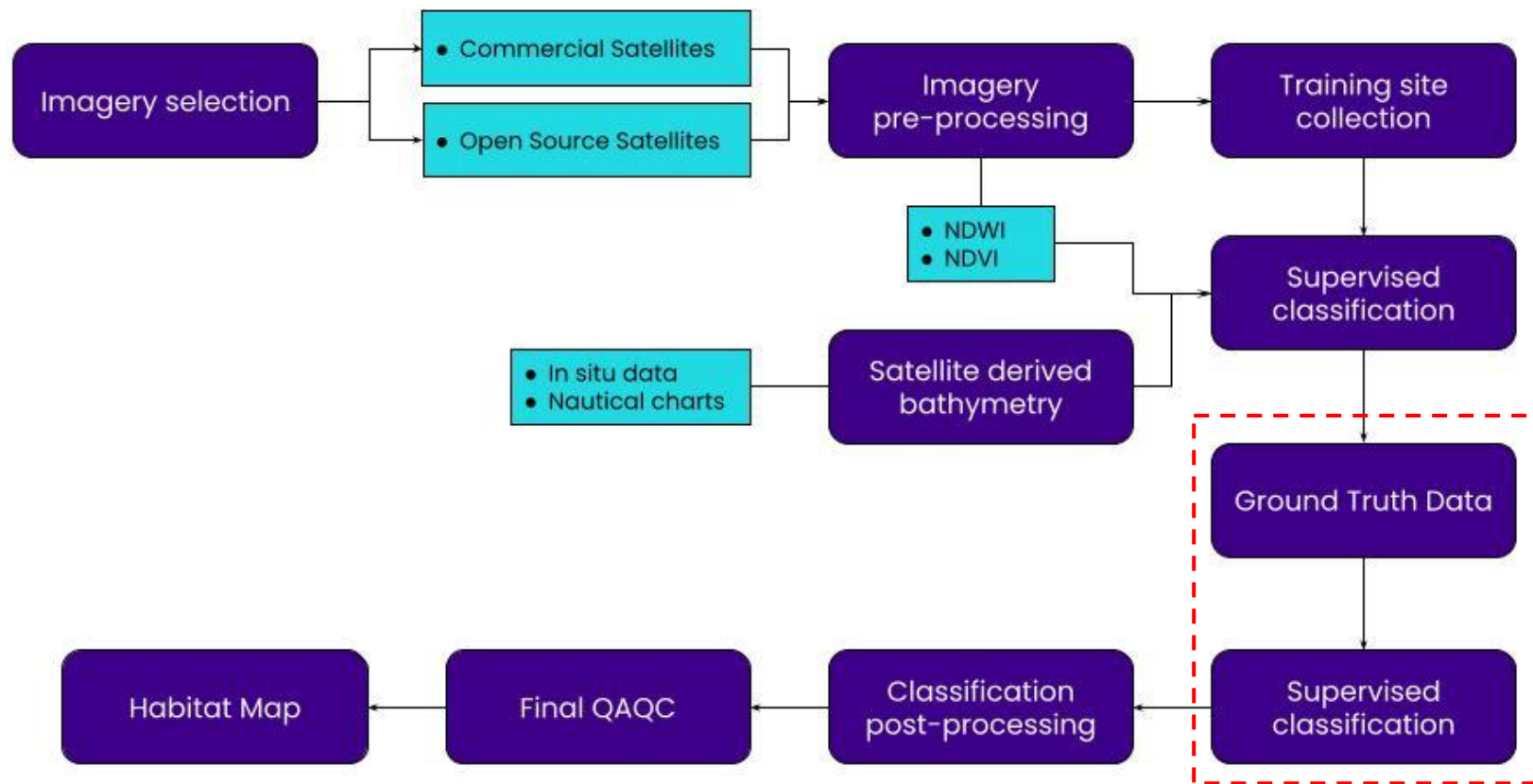
- <u>Size of the Bootstrapped Dataset</u>

This is what **percentage of the training data** should be used to train each individual tree.
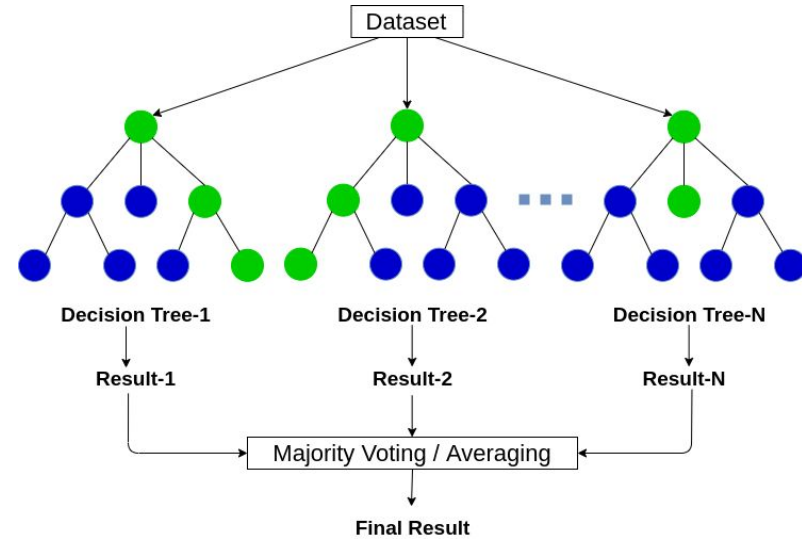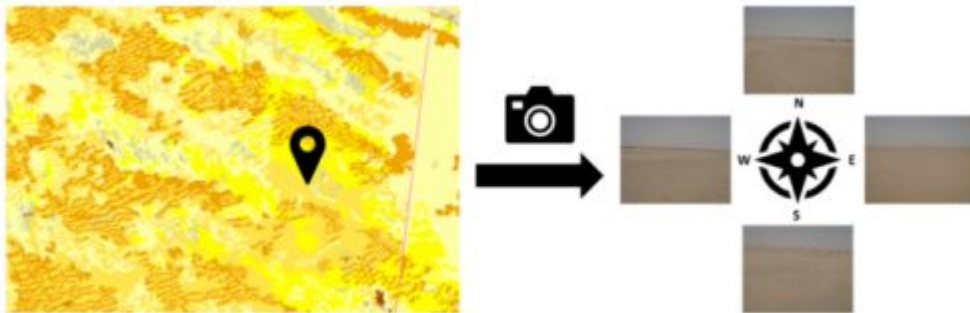

Source: Random Forest vs Decision Tree | Which Is Right for You?
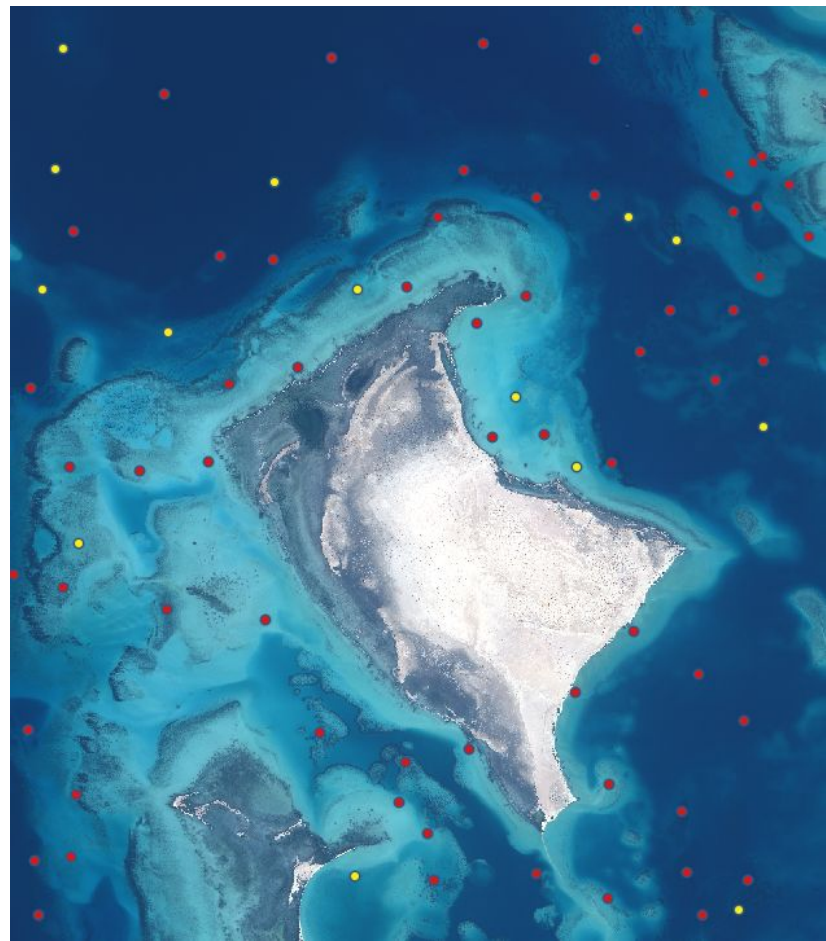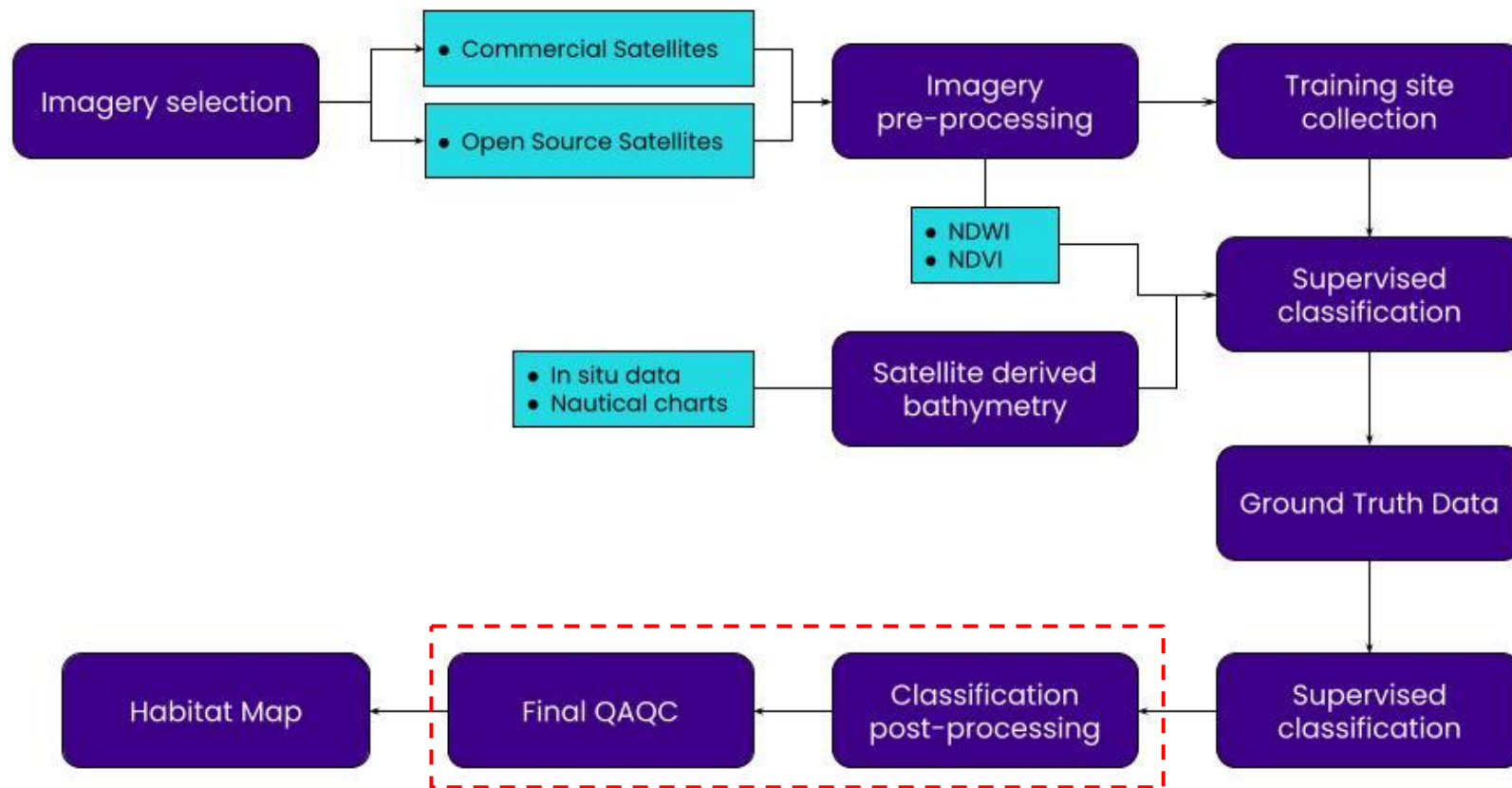
Workflow

Ground Truth Data

🔴 Training

🟡 Validation

A **new random forest classification algorithm** is used to produce the final marine habitat map. The inputs to the algorithm are:

- All bands of the multispectral image
- The derived spectral indices
- Satellite derived bathymetry
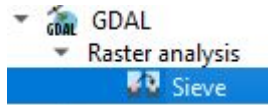- The training data produced from the ground truth data.

Workflow

Post-processing

The output classification raster is:

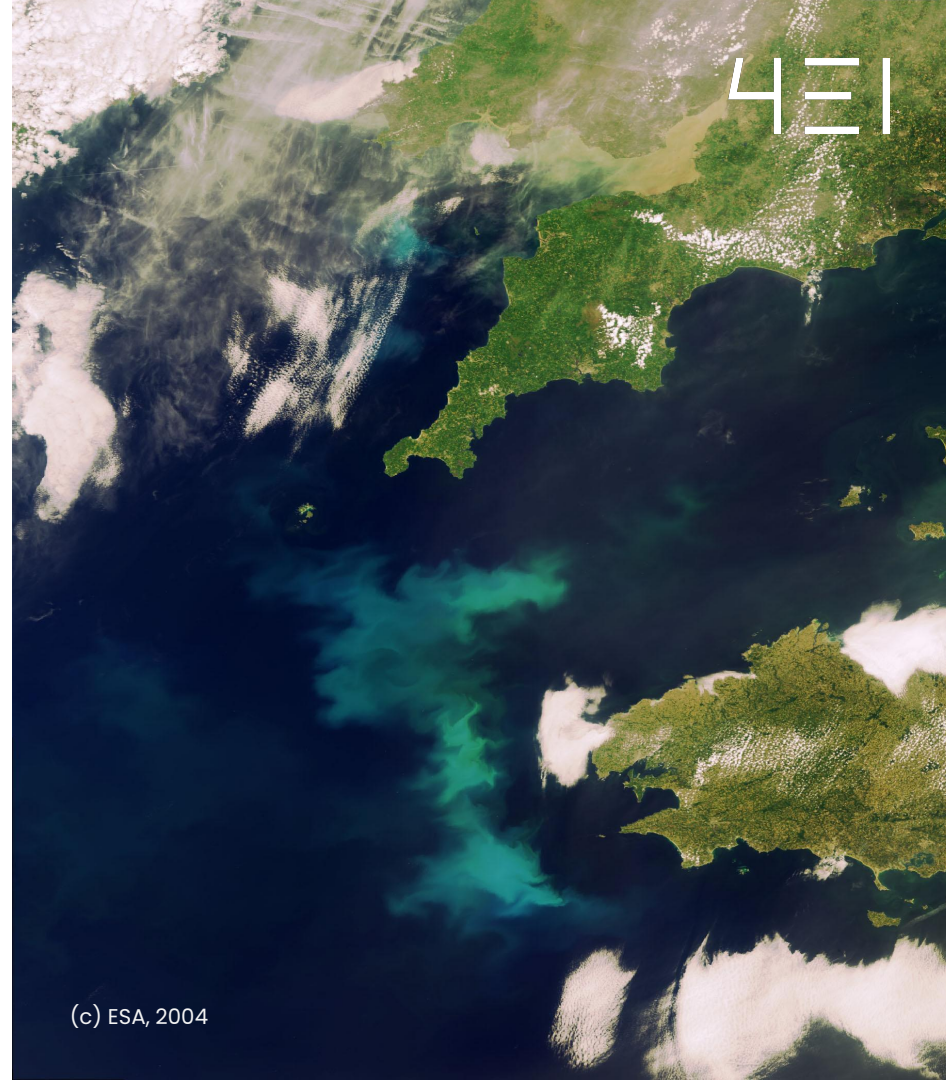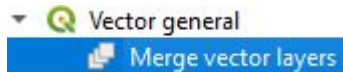- smoothed with a majority filtering algorithm to reduce noise

GDAL
  Raster analysis
    Sieve

- A boundary cleaning function to smooth the transition between classes.

GDAL
  Raster miscellaneous
    Merge

GDAL
  Vector geoprocessing
    Dissolve

- The raster is converted to polygons and simplified to generalise the classifications.
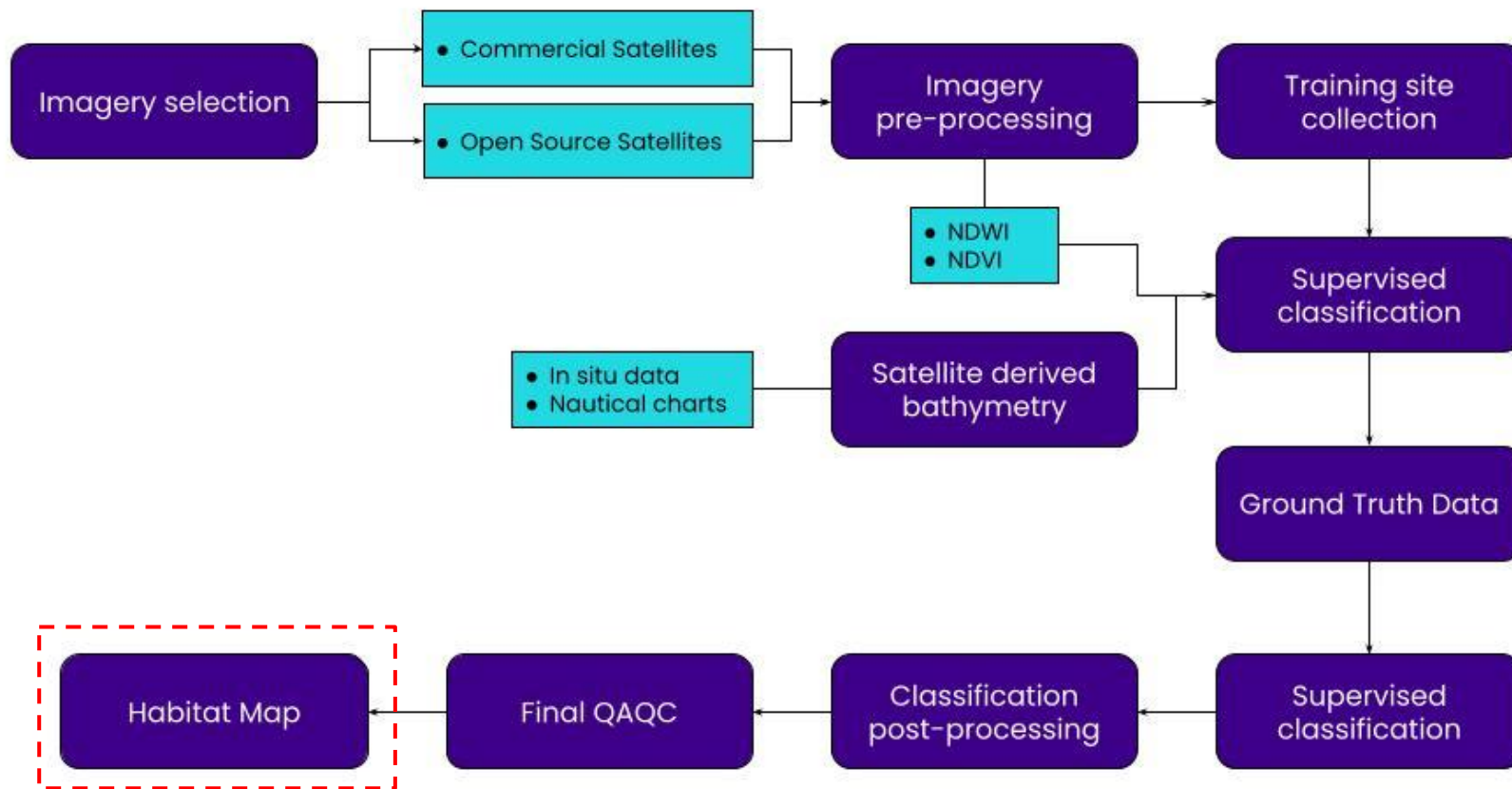
GDAL
  Raster conversion
    Polygonize (raster to vector)

- Any polygons that are deemed small enough to be noise are dissolved into the polygon with which they share the largest area.

Vector general
  Merge vector layers

(c) ESA, 2004

Workflow

Marine Habitat Map

1010 – Intertidal mud with sparse mangroves

1040 – Dense mangroves

11100 – Patch reef

11200 – Fringing reef

14100 – Bare hardground

14200 – Algal hardground

16100 – Shallow sand (mobile)

16200 – Deep sand (stable)

12000 – Coral assemblage

13100 – Sparse seagrass and algal meadow

13200 – Dense seagrass

17000 – Deep open ocean

30000 – Land

431

Thank You

4EI Contact

Javier Garcia Lopez

Project Management, 4EI

+44 117 933 8562

jgl@4earthintelligence.com

4EI

@4earthintelligence

@4_intelligence

4 Earth Intelligence