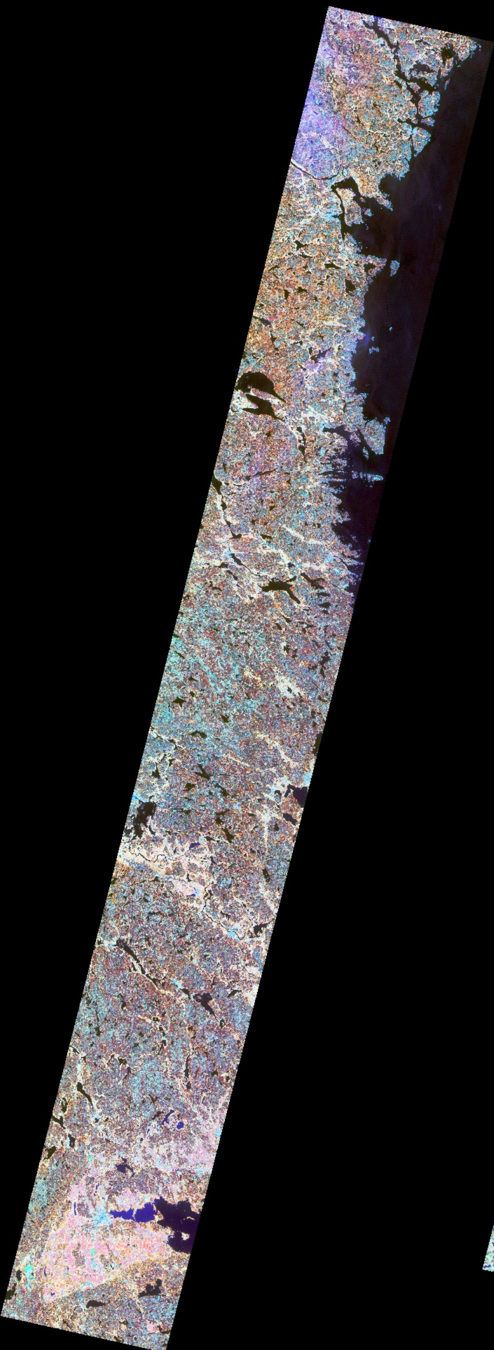
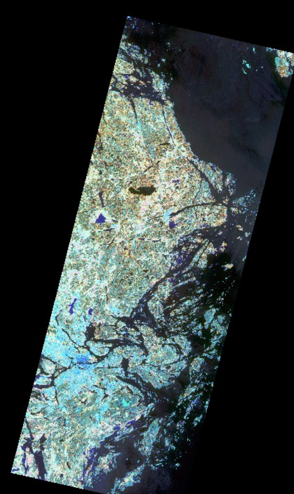
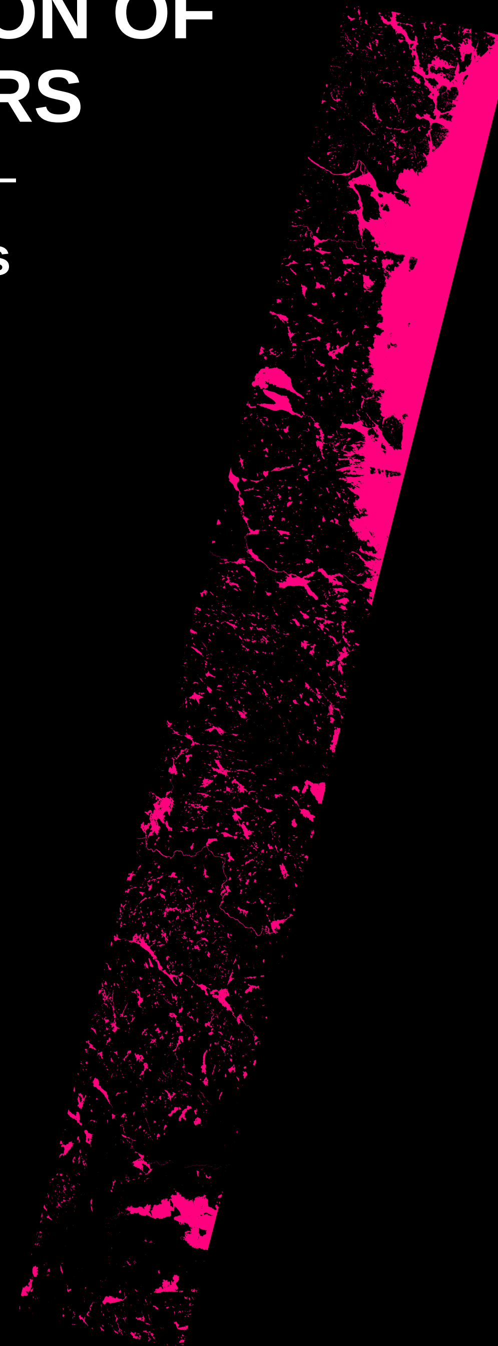
TECHNOLOGICAL DEVELOPMENTS FOR THE PRODUCTION OF HR WATER LAYERS

Dr Gediminas Vaitkus

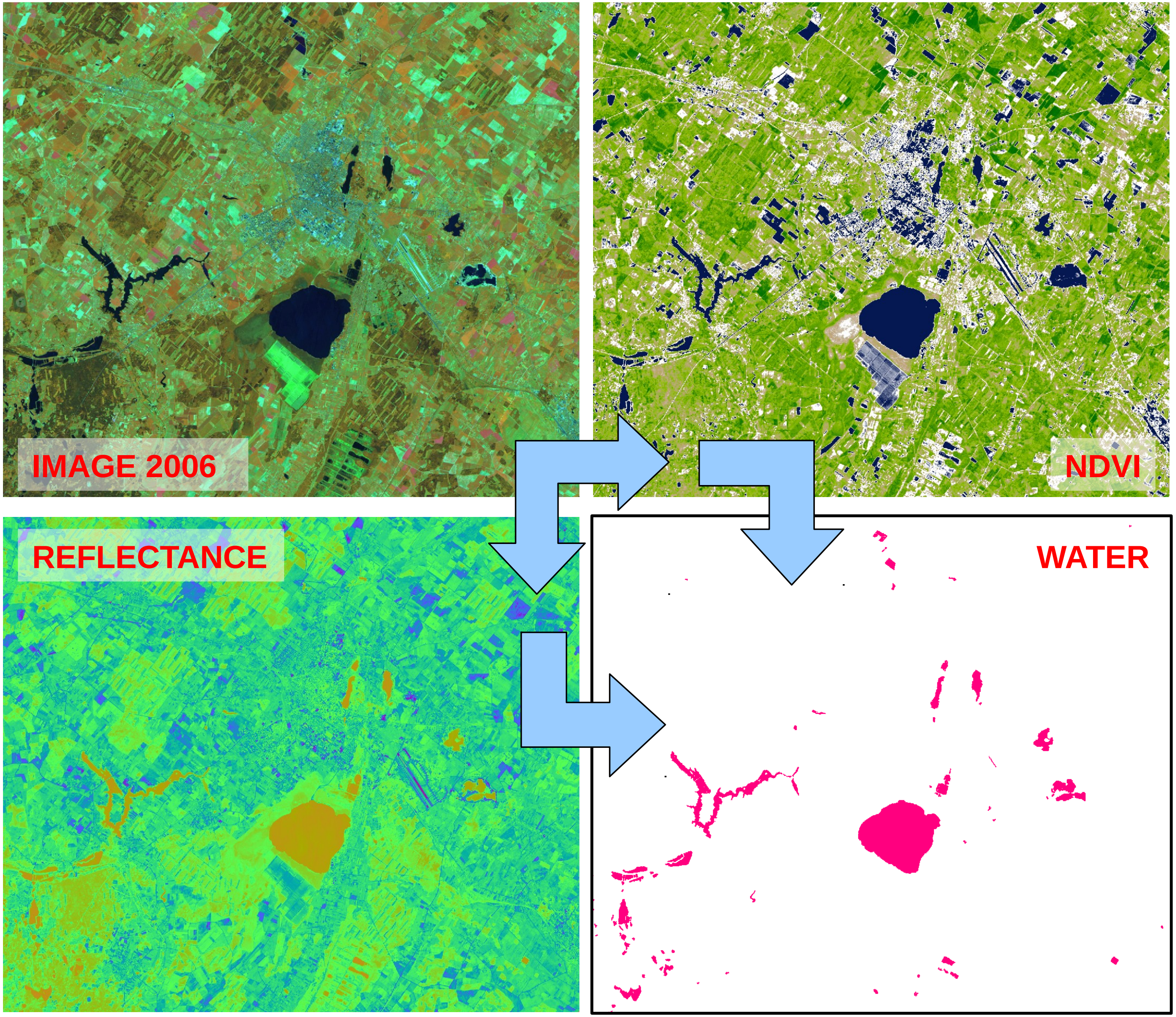
GEOMATRIX uab
gedas.vaitkus@gmail.com





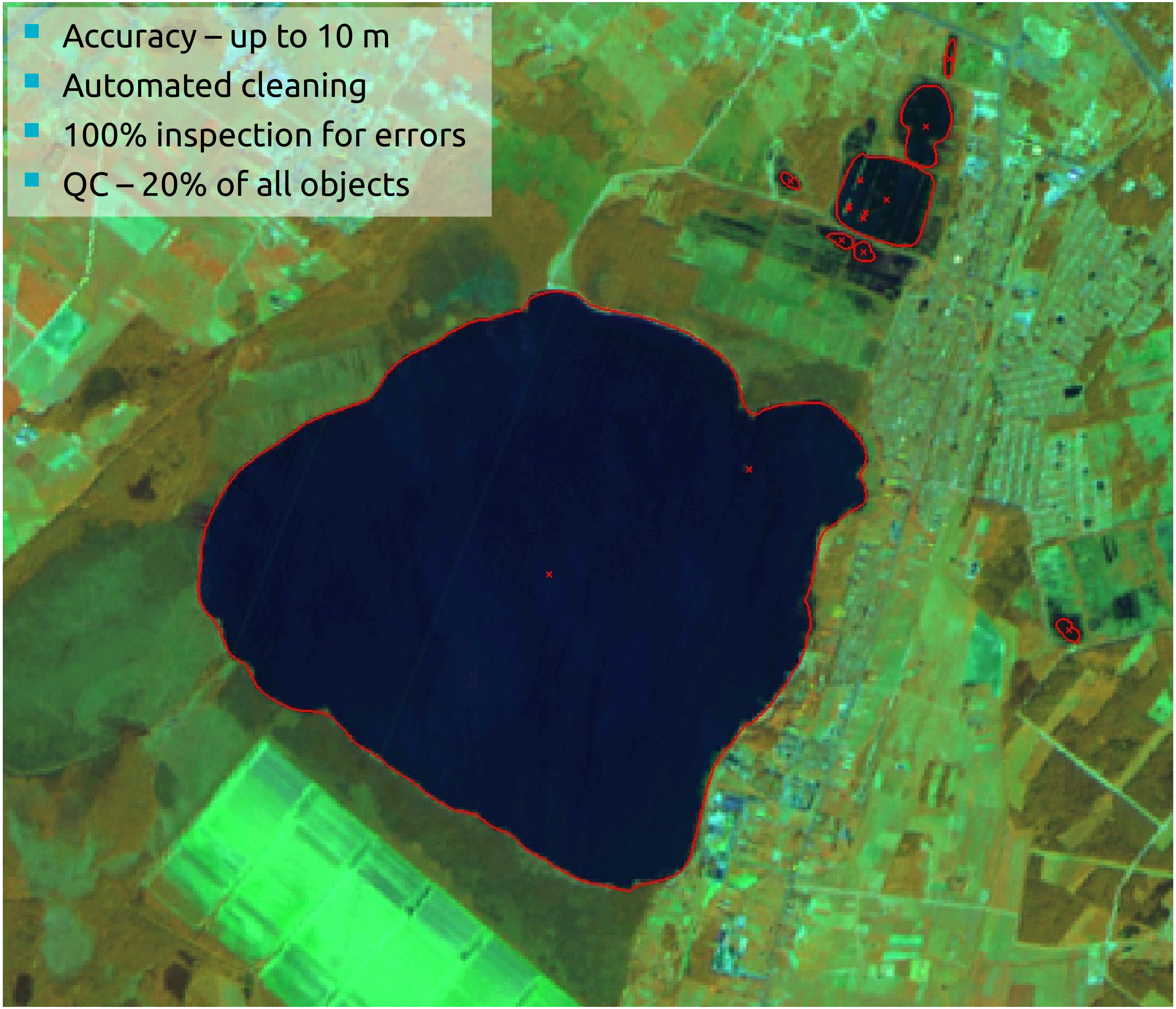


EXTRACTION OF WATER MASK



PRODUCTION SPECIFICATIONS

- Accuracy – up to 10 m
- Automated cleaning
- 100% inspection for errors
- QC – 20% of all objects

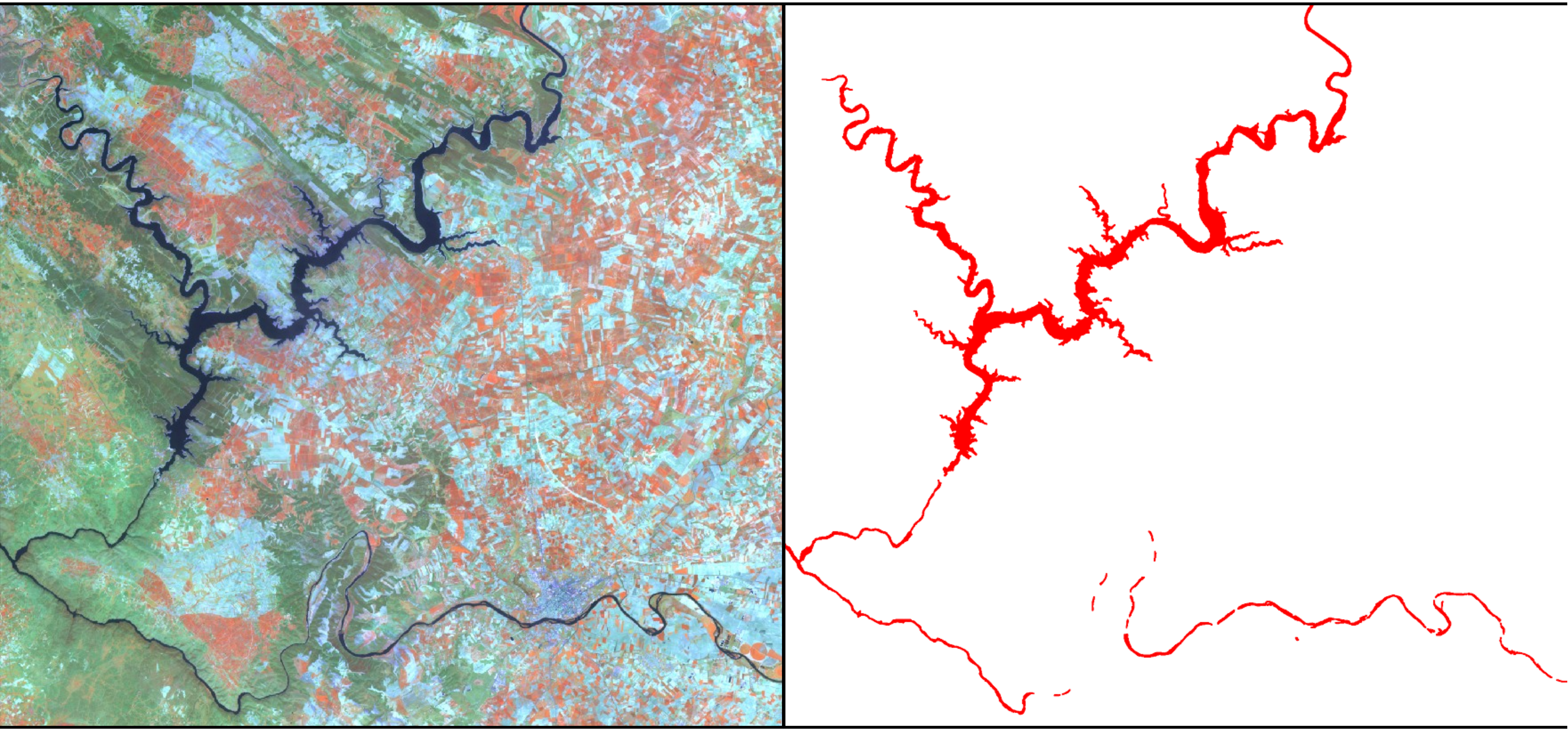


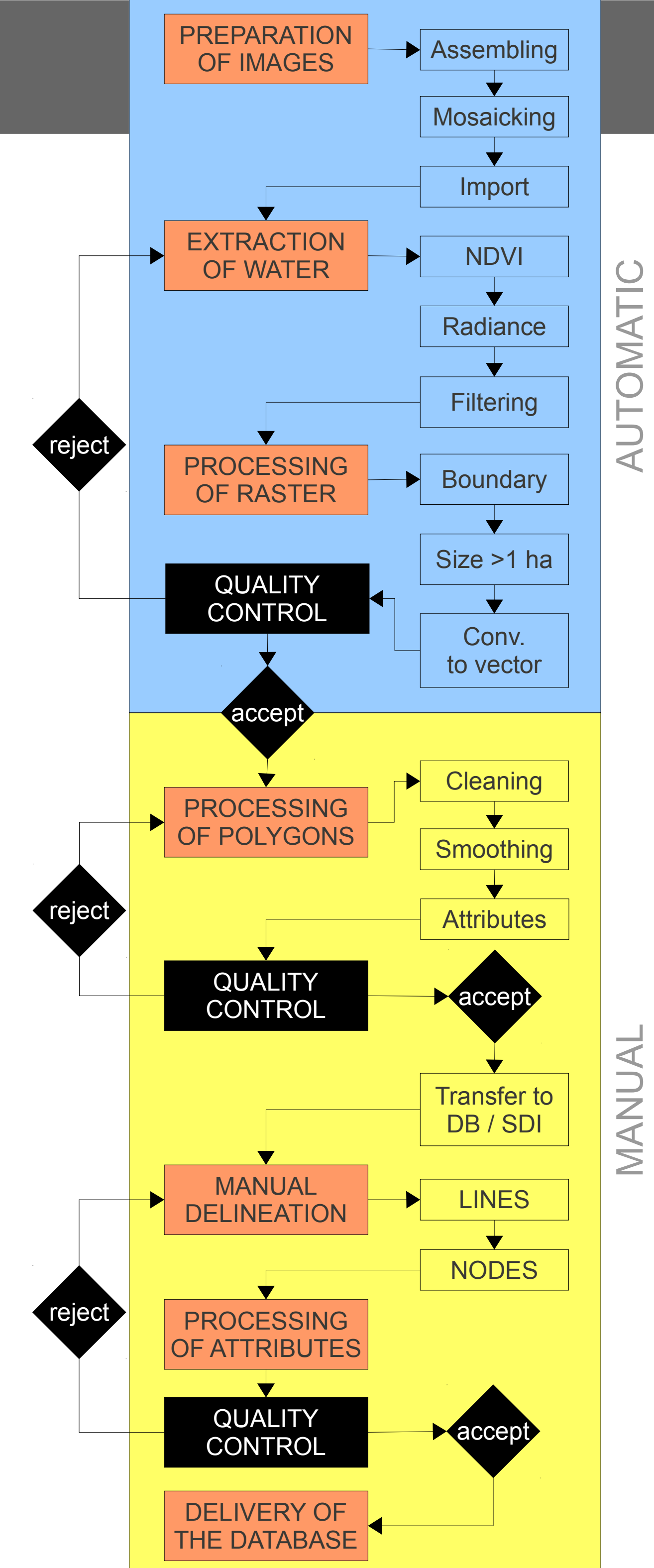
GRDA – Implementation of an Initial GMES Service for Geospatial Data Access *(DG-ENT: 2010-2012)*

- Open tender (DG-ENT) for GMES reference data production
- Project objective – produce seamless high-resolution DEM and hydrography reference datasets covering EEA38 territory (~6 mln km^2), integrating the existing hydrography datasets and improving their geometry and topology
- Project consortium:
  - INDRA Espacio (ES) – project coordinator
  - INTERMAP (DE) – production of EU-DEM
  - **AGI** – production of hydrography reference layers
    - **G.Vaitkus** (founder of GEOMATRIX uab) created automated processing scripts, performed automated extrction of water bodies, and computing of watersheds
- Project supervised by a joint working group of EC, GMES Bureau, JRC, EEA ir Eurostat representatives

CONCEPT : A REFERENCE DATABASE

AUTOMATIC PRODUCTION OF RASTER WATER MASK
FROM IMAGE 2006: ~2000 separate images processed





PRODUCTION WORK-FLOW

PRODUCTION OF HYDRO- LAYERS:

- Polygons — extracted from IMAGE-2006 MOSAIC COV1 (reference) and COV2 (supplement)
- Lines — delineated from IMAGE-2006 MOSAIC and ancillary data (topo maps, etc.)
- Points — generated from network topology (lines)

QA/QC PROCEDURES:

- Internal QA/QC – ISO 9001:2000
  100% of processed images inspected
  Iterative processing to reach compliance to specs.
  20% of all vector objects checked for each unit

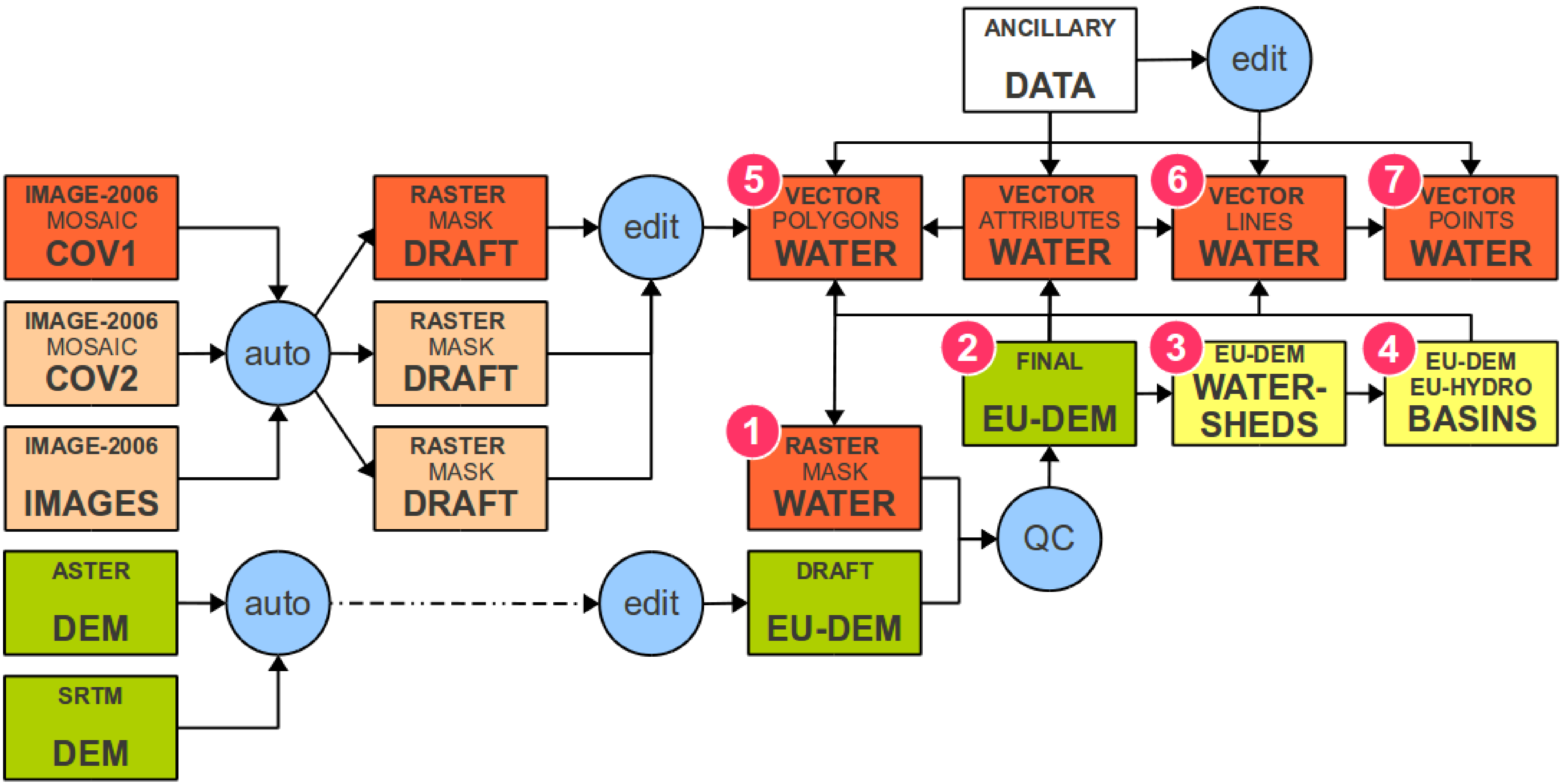
- External QA/QC by project partners
  100% “large” hydro-polygons checked during the data integration process

TECHNOLOGY:

- Use of FOSS GIS tools
  GRASS GIS, GDAL, etc. scripts for production of polygons
  Quantum GIS for manual editing and QC
  PostgreSQL/PostGIS for database storage

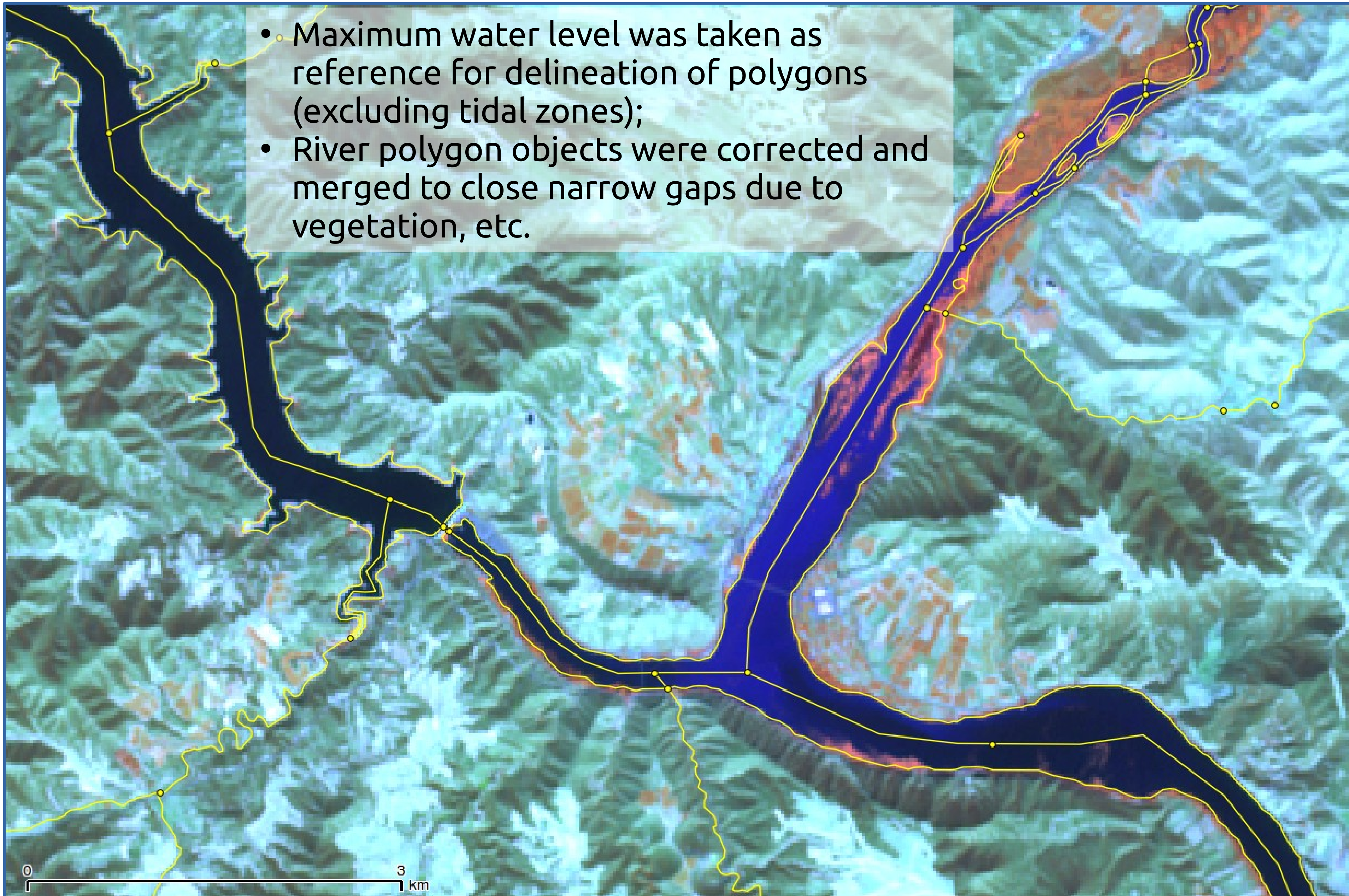
GEOMATRIX UAB
Extreme GEO-computing

GRDA PRODUCTION SEQUENCE



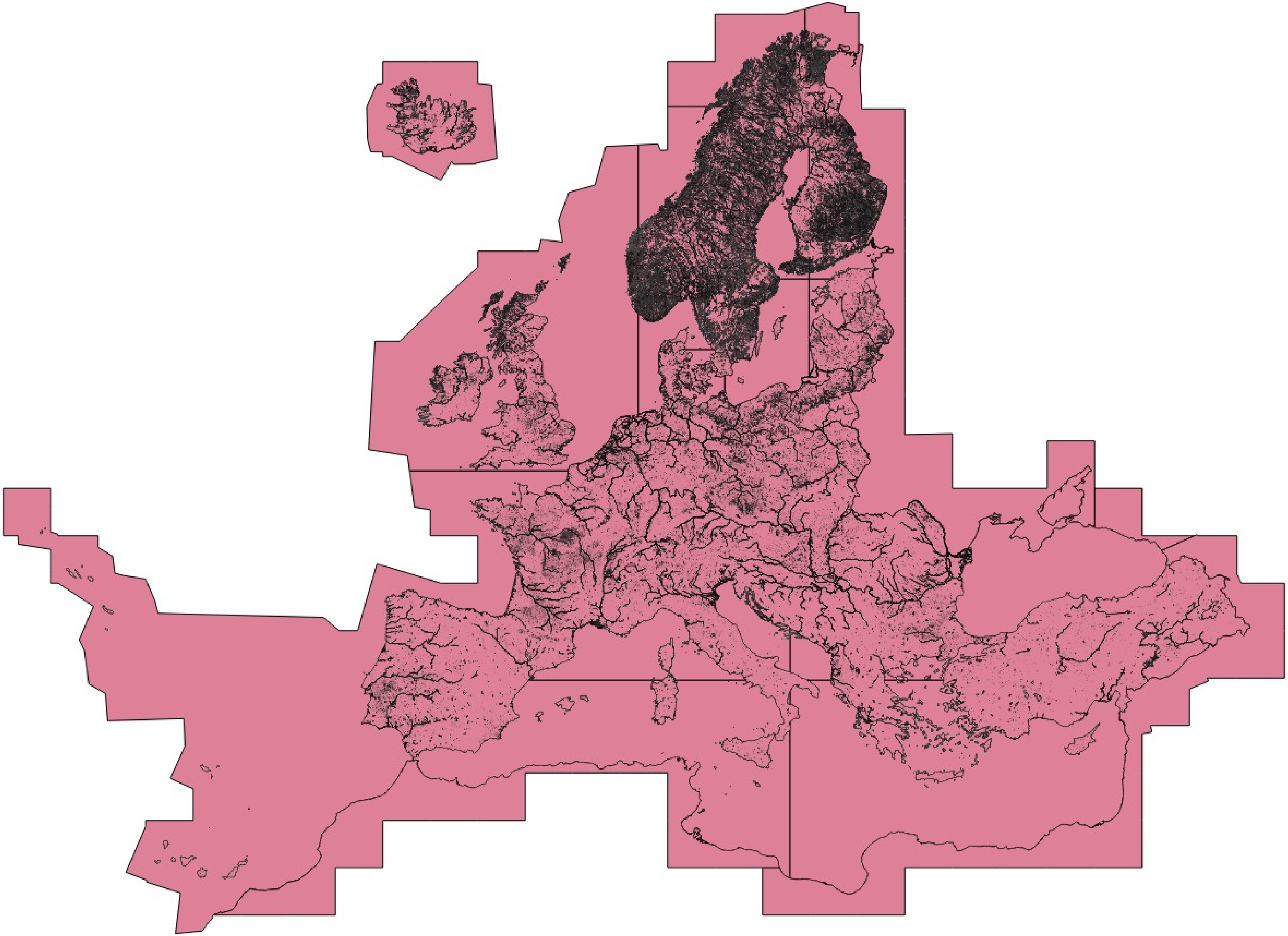
THE EU-HYDRO PRODUCT

- Maximum water level was taken as reference for delineation of polygons (excluding tidal zones);
- River polygon objects were corrected and merged to close narrow gaps due to vegetation, etc.



Geometry corresponds to IMAGE-2006 reference, has 3 layers with topology

EU-HYDRO UPDATE 2015

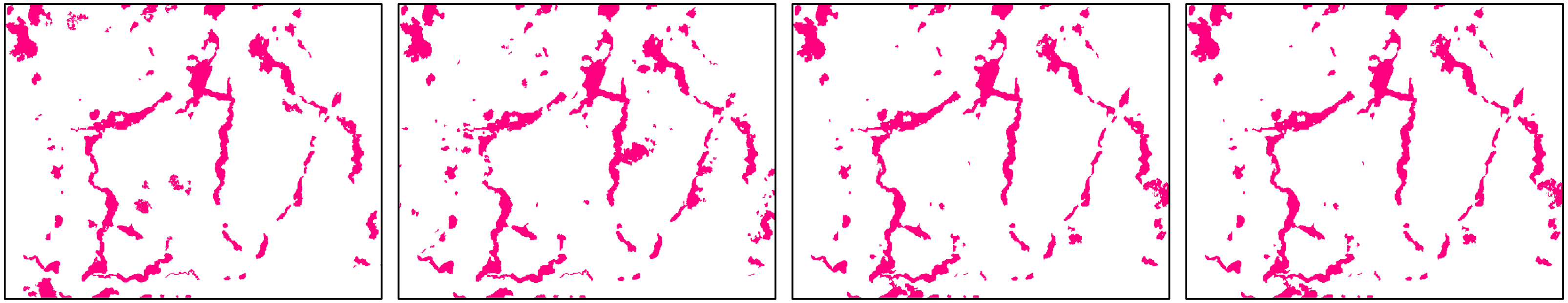


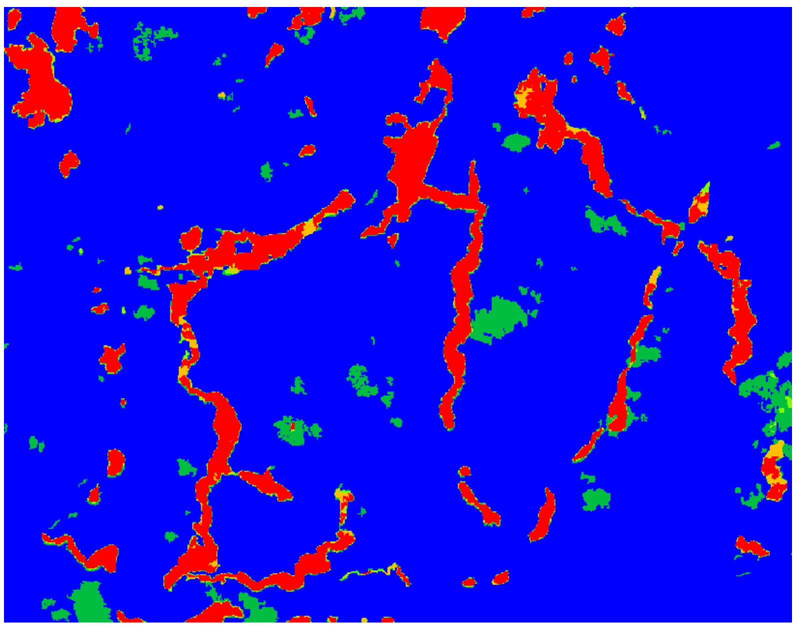
GIO-LAND – GMES Initial Operations Land – *Lot 6: Water & Wetlands layers (EEA: 2013-2014)*

- Open tender for LOT-6 production
- Project objective – production of seamless high-resolution permanent water and wetland masks covering EEA39 territory (~6.3 mln km^2) by processing all the available HR satellite imagery (IMAGE 2006/2009/2012)
- Project consortium:
  - INDRA Espacio (ES) – project coordinator, production of grassland
  - DLR (DE), EUROSENSE (BE) – production of grassland
  - **GEOMATRIX uab** – production of water and wetlands layers
- Production of water and wetlands layers was carried out by automated parallel computing based on open-source technologies

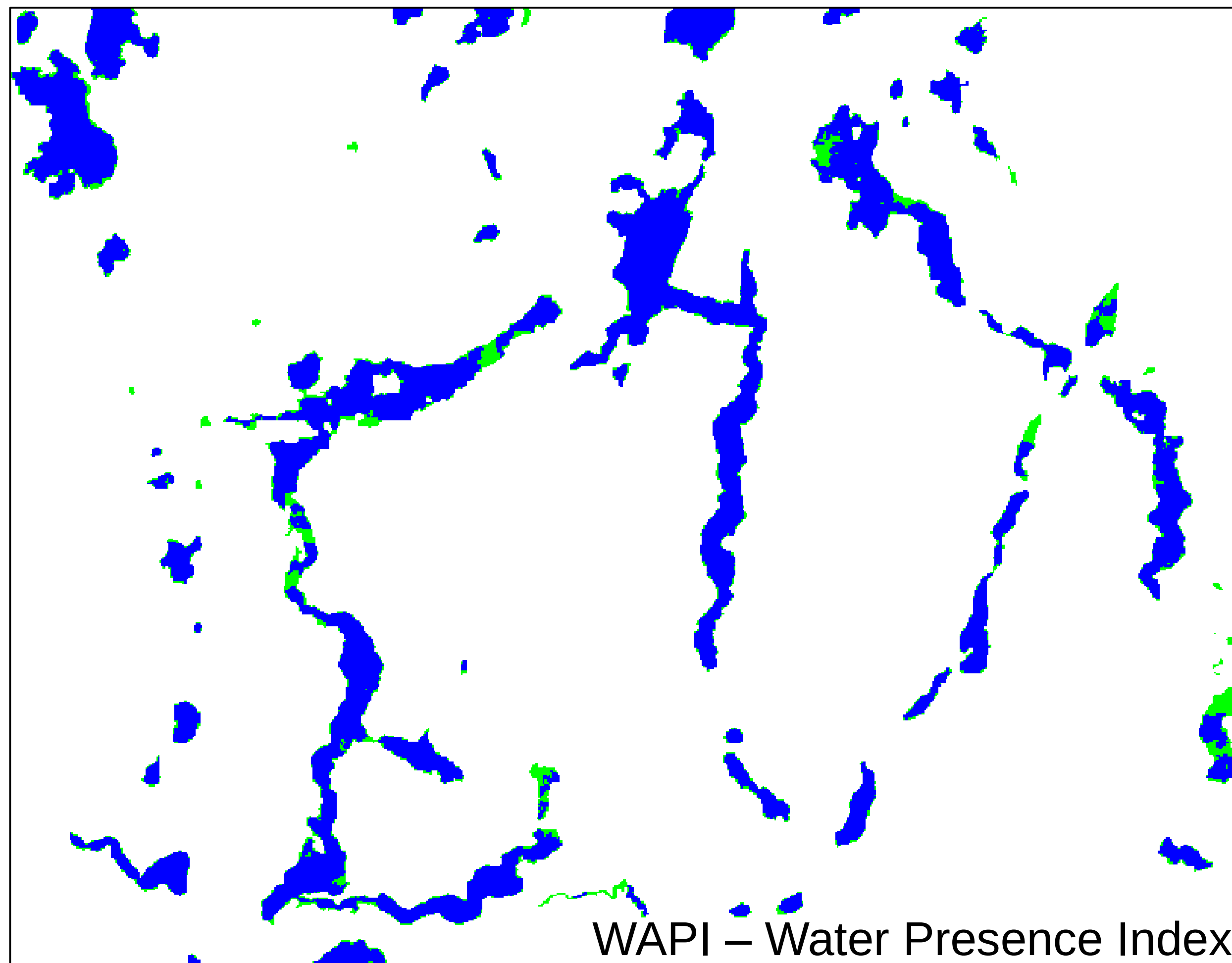
CONCEPT : A STATISTICAL PRODUCT

2006.07.01 2006.09.12 2007.04.25 2007.05.21





SWAM – Sum of Water Masks

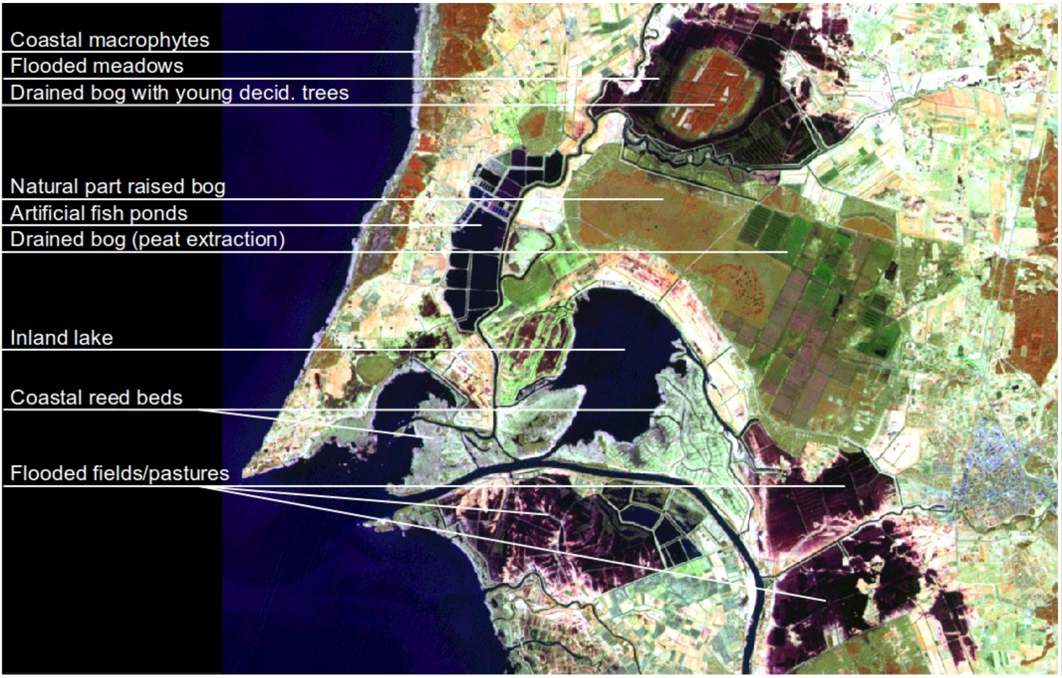


WAPI – Water Presence Index

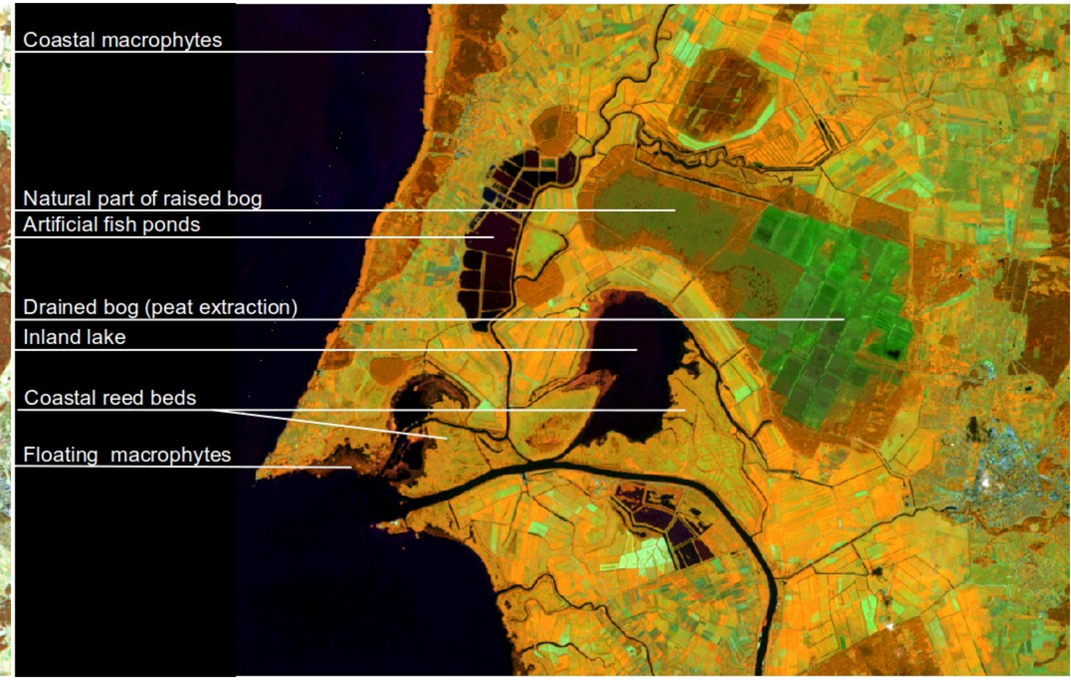
GEOMATRIX UAB
Extreme GEO-computing

TIME-SERIES ANALYSIS

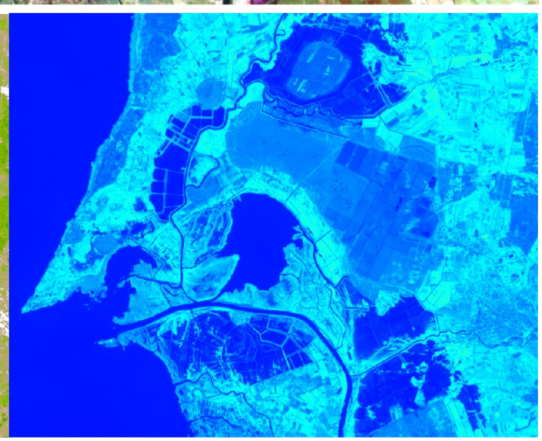
spring (IRS-6: 2009/04/06)



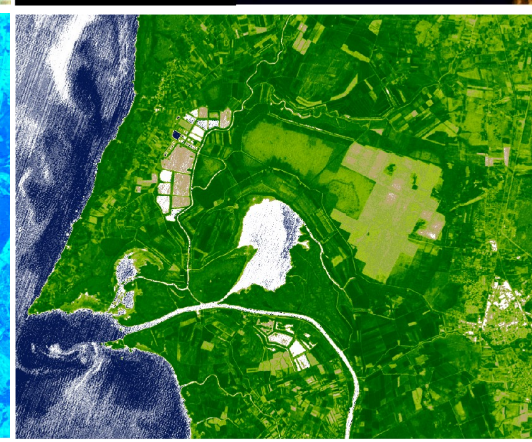
summer (IRS-6: 2008/08/14)



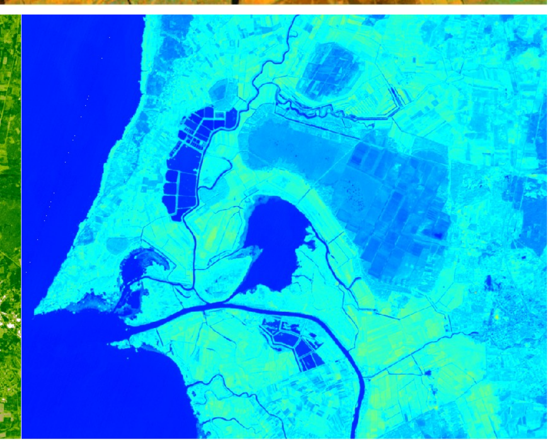
Vegetation Index (NDVI)



Water Index (NDWI)



Vegetation Index (NDVI)

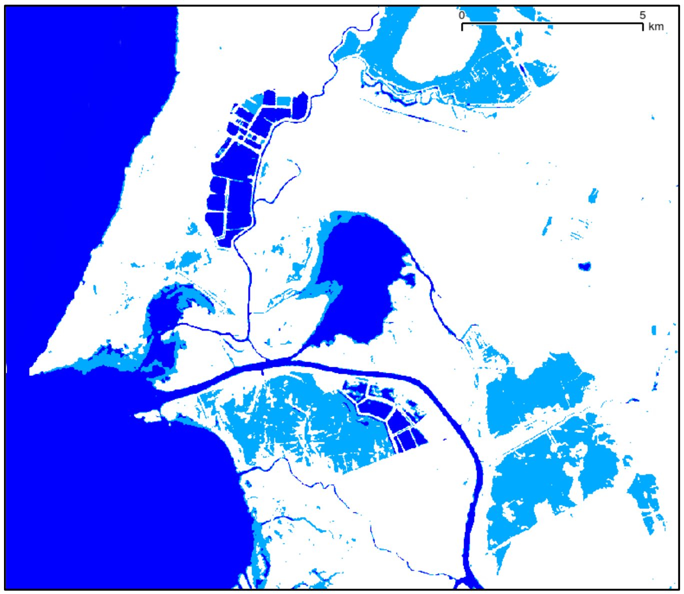


Water Index (NDWI)

GEOMATRIX UAB
Extreme GEO-computing

PERMANENT & TEMPORARY OBJECTS

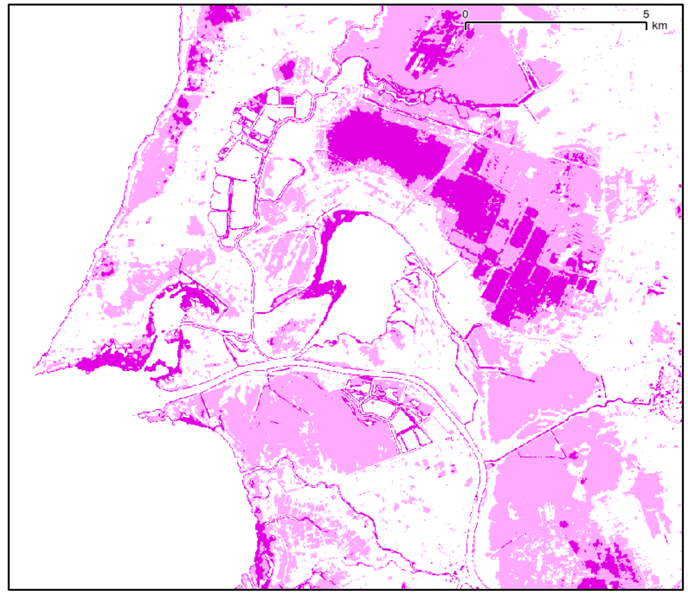
WATER



Dark color – “permanent” water
{WAPI => 0.5}

Light color – “temporary” water
{WAPI < 0.5}

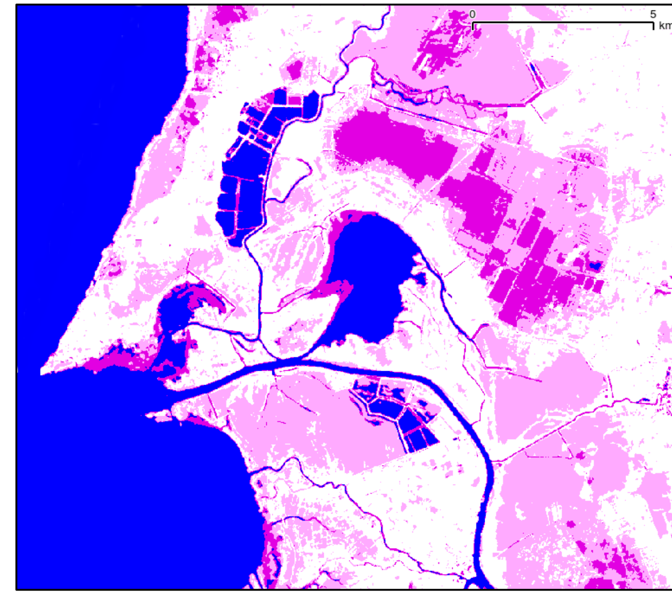
WETLANDS



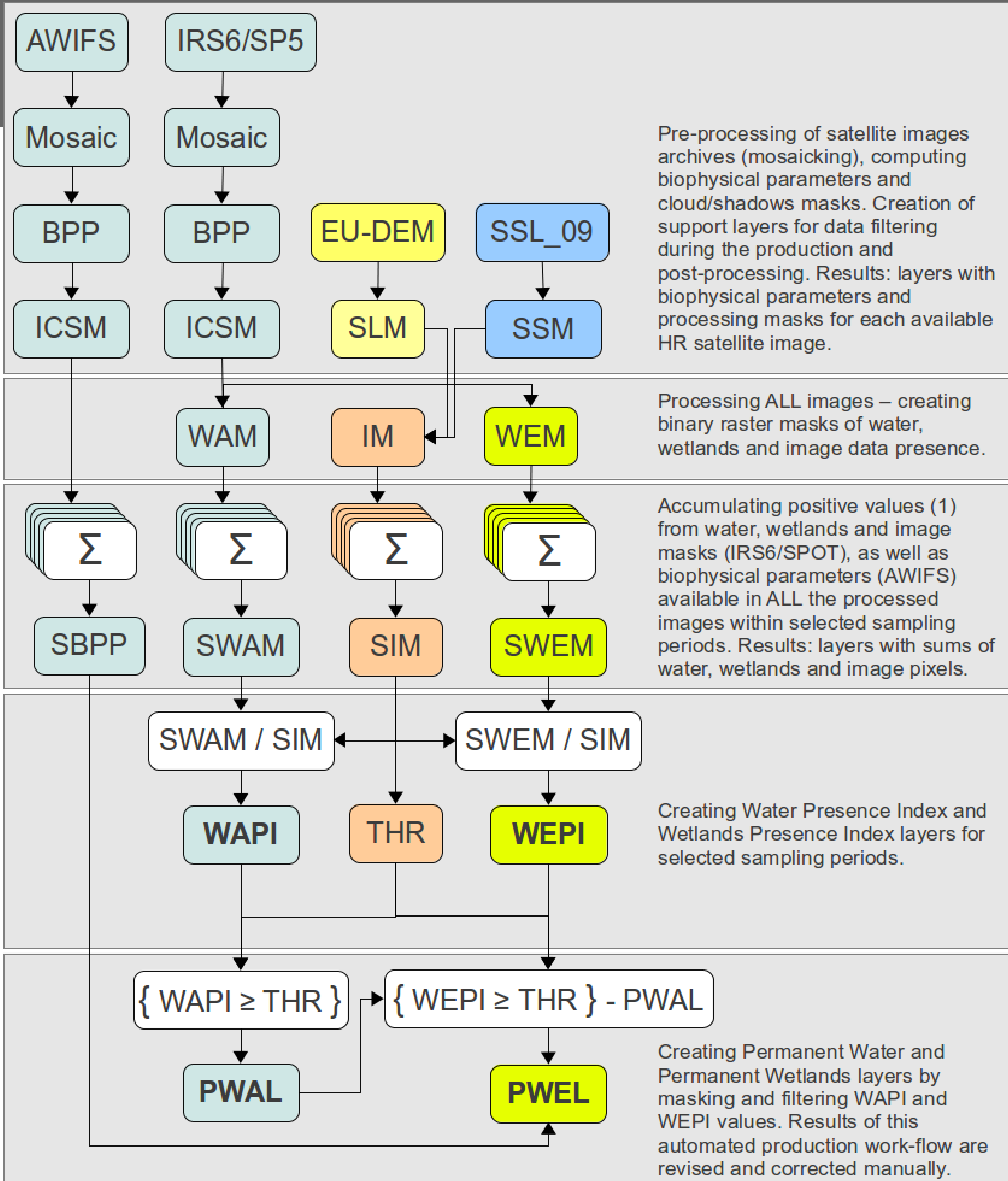
Dark color – “permanent” wetland
{WEPI => 0.5} <u>OR</u>
{WAPI < 0.5 <u>AND</u> WEPI => 0.5}

Light color – “temporary” wetland
{WEPI < 0.5} <u>OR</u> {WAPI < 0.5}

COMBINED LAYER

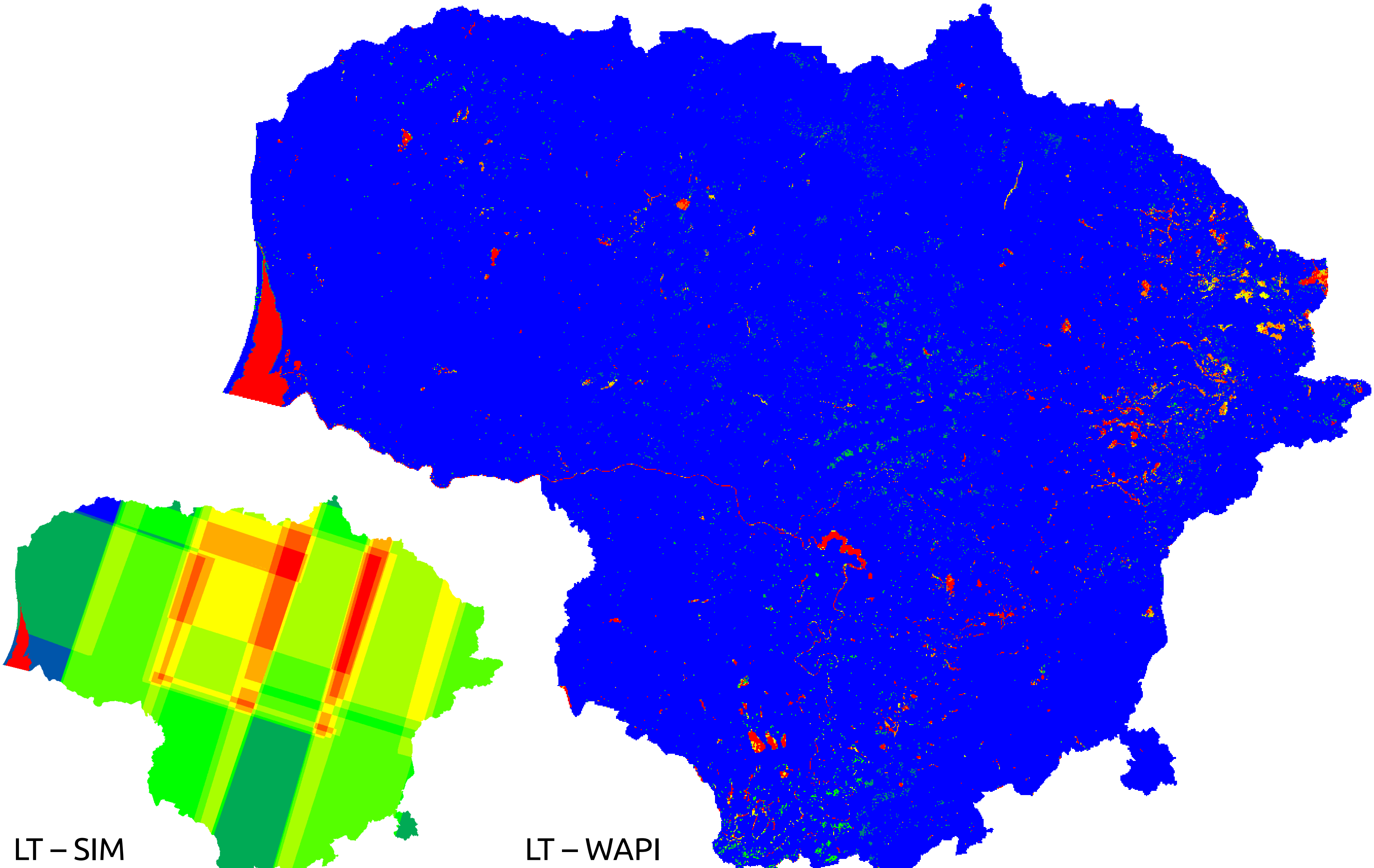


PRODUCTION WORK-FLOW



IRS6/SP5 – High Resolution satellite images – IRS6/SPOT4,5 (IMAGE2006/2009/2012);
AWIFS – Medium Resolution satellite images (AWIFS time-series);
BPP – Biophysical parameters (Total Radiance, NDVI, NDWI, WI);
SBPP – Summary layers for BPP (minimum, maximum, range);
EU-DEM – The standard EU-DEM product (30 m resolution, EPSG:3035);
SML – Slope Mask computed from EU-DEM;
SSL_09 – Soil sealing dataset (imperviousness 2009 product - 20 m resolution, EPSG:3035);
SSM – Soil Sealing Mask (>20 %) computed from SSL_09;
WAM – binary raster mask indicating presence of water In a separate image;
WEM – binary raster mask indicating presence of wetlands (water + adjacent wetlands vegetation) in a separate image;
IM – binary raster mask covering non-null values of a separate image;
SWAM – sum of water mask values from all processed images;
SWEM – sum of wetlands mask values from all processed images;
SIM – sum of image masks of all processed images;
WAPI – Water Presence Index (SWAM / SIM);
WEPI – Wetlands Presence Index (SWEM / SIM);
PWAL – Permanent Water Layer ({ WAPI ≥ THR});
PWEL – Permanent Wetlands Layer ({ WEPI ≥ THR} - PWAL);
THR – threshold value of WAPI and WEPI for computing PWAL and PWEL.

PROBLEM – IMAGE AVAILABILITY

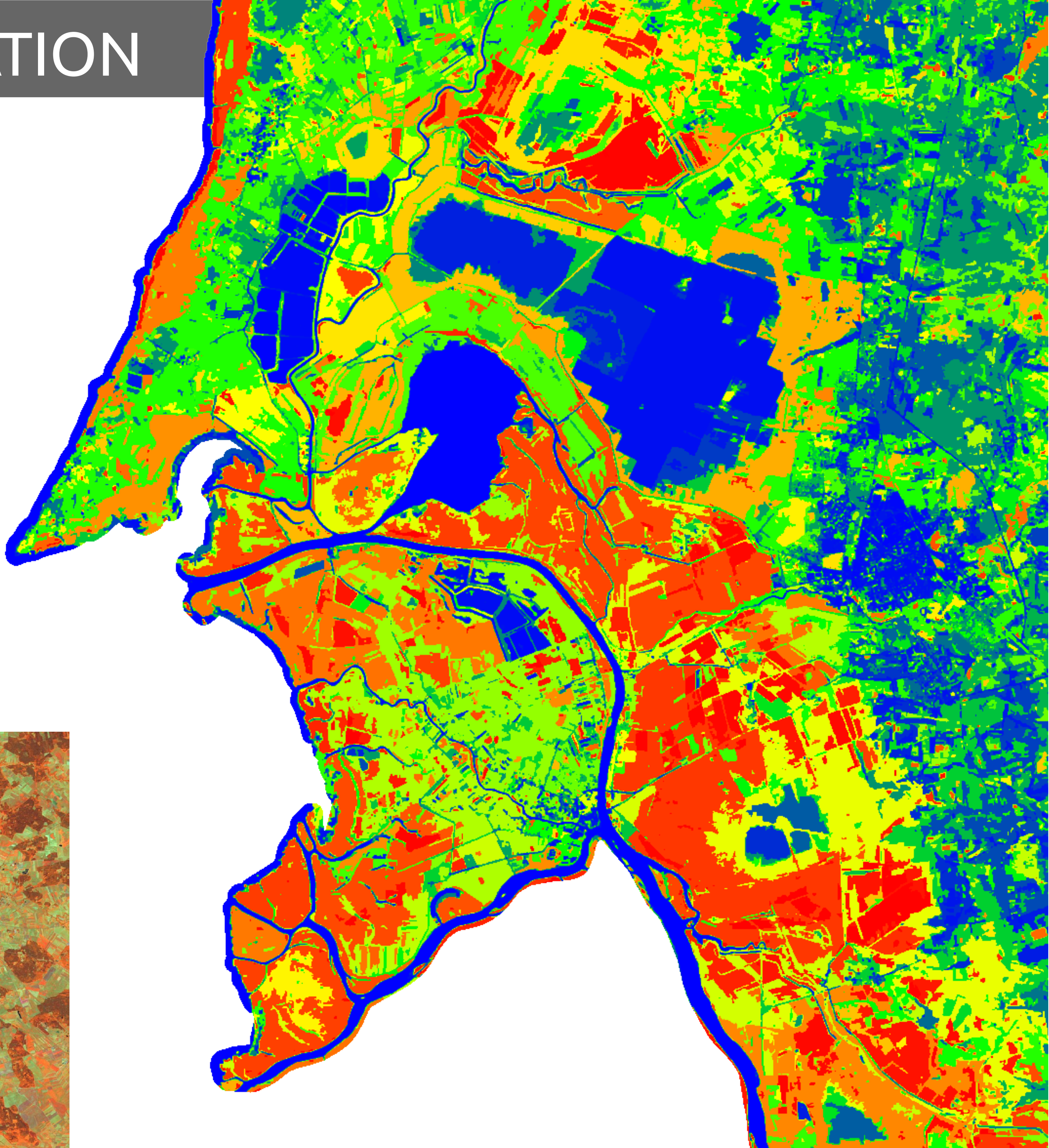


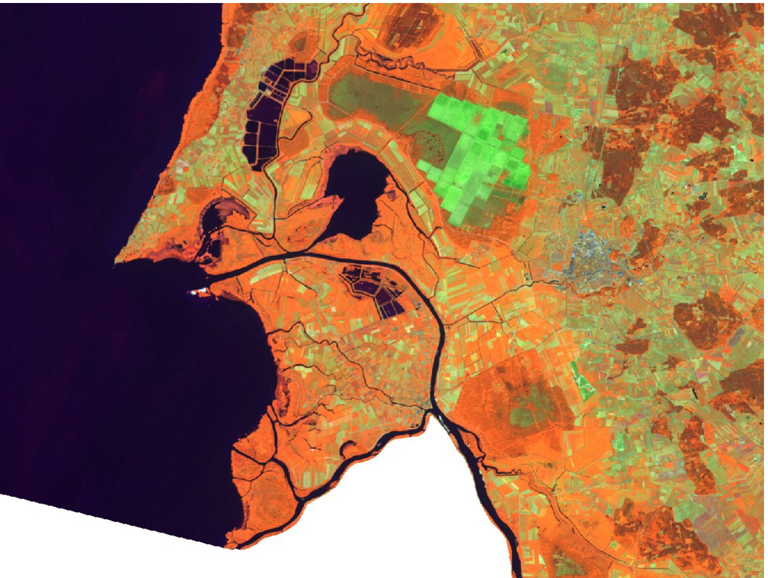
LT – SIM

LT – WAPI

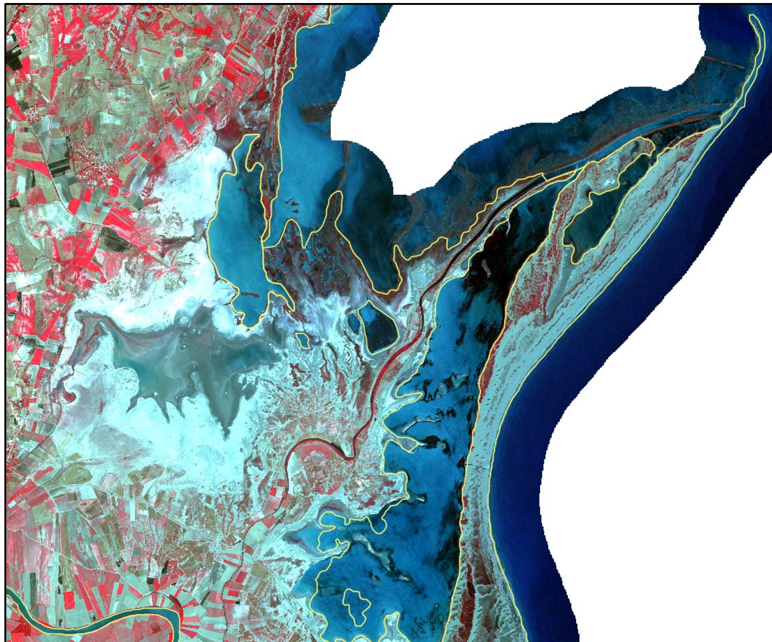
BPP SEGMENTATION

In order to compensate for insufficient availability and quality of IRS-6 images in moderate and northern latitudes, a new method of BPP segmentation was developed and non-water-related wetland segments were integrated into the HRL Wetland layer. This will be done in some of the southern European countries during the enhancement phase.

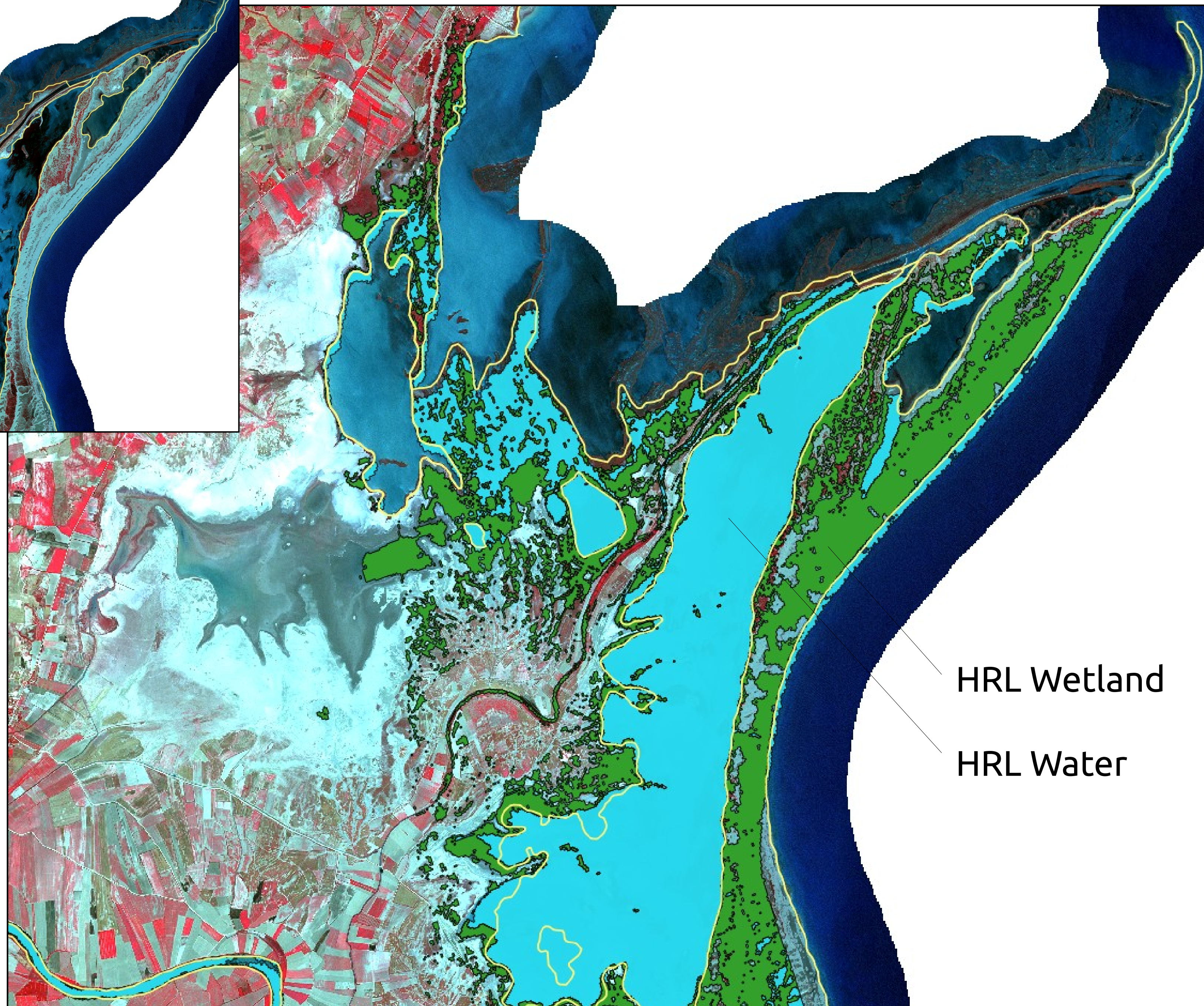




HRL WATER & WETLAND LAYERS



VHR SPOT 5
2011/06/24
Turkey



RECENT DEVELOPMENTS

Consolidation of technological developments resulted in an integrated EO imagery pre-processing and analysis technological process:

- Highly-efficient, completely automated and production-ready technological solution;
- Based exceptionally on the Open-Source software and platforms;
- Focused on multi-purpose GMES/Copernicus downstream production operations;
- Capable of handling both HR and VHR EO Data products.

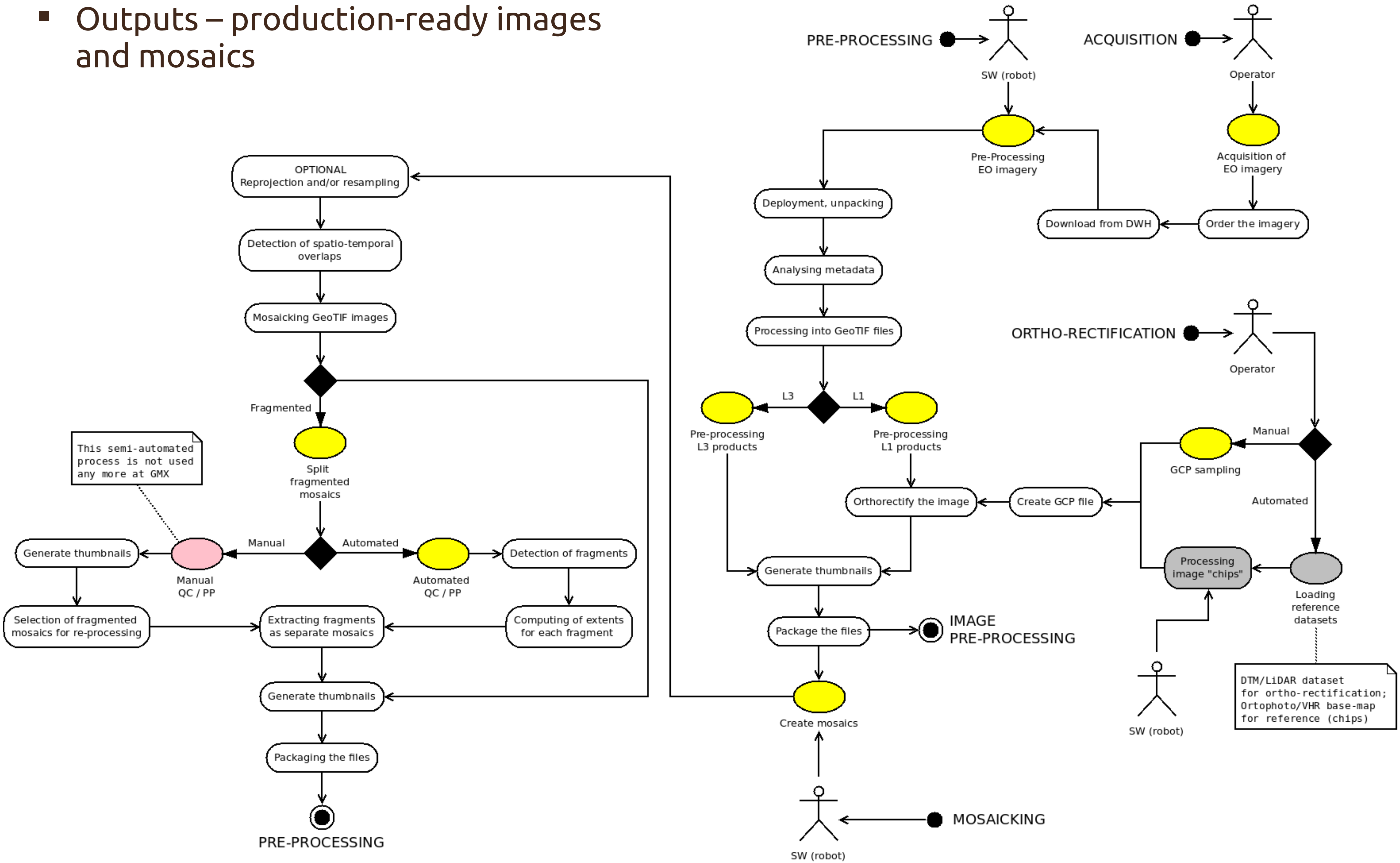
Operational testing carried out in summer 2014 on pan-European coverages of:

  - HR (5 m) RapidEye (ESA DWH_MG2_CORE_01 Coverage_2)
  - VHR (2.5 m) SPOT 5/6 (ESA DWH_MG2b_CORE_03 Coverage_2)

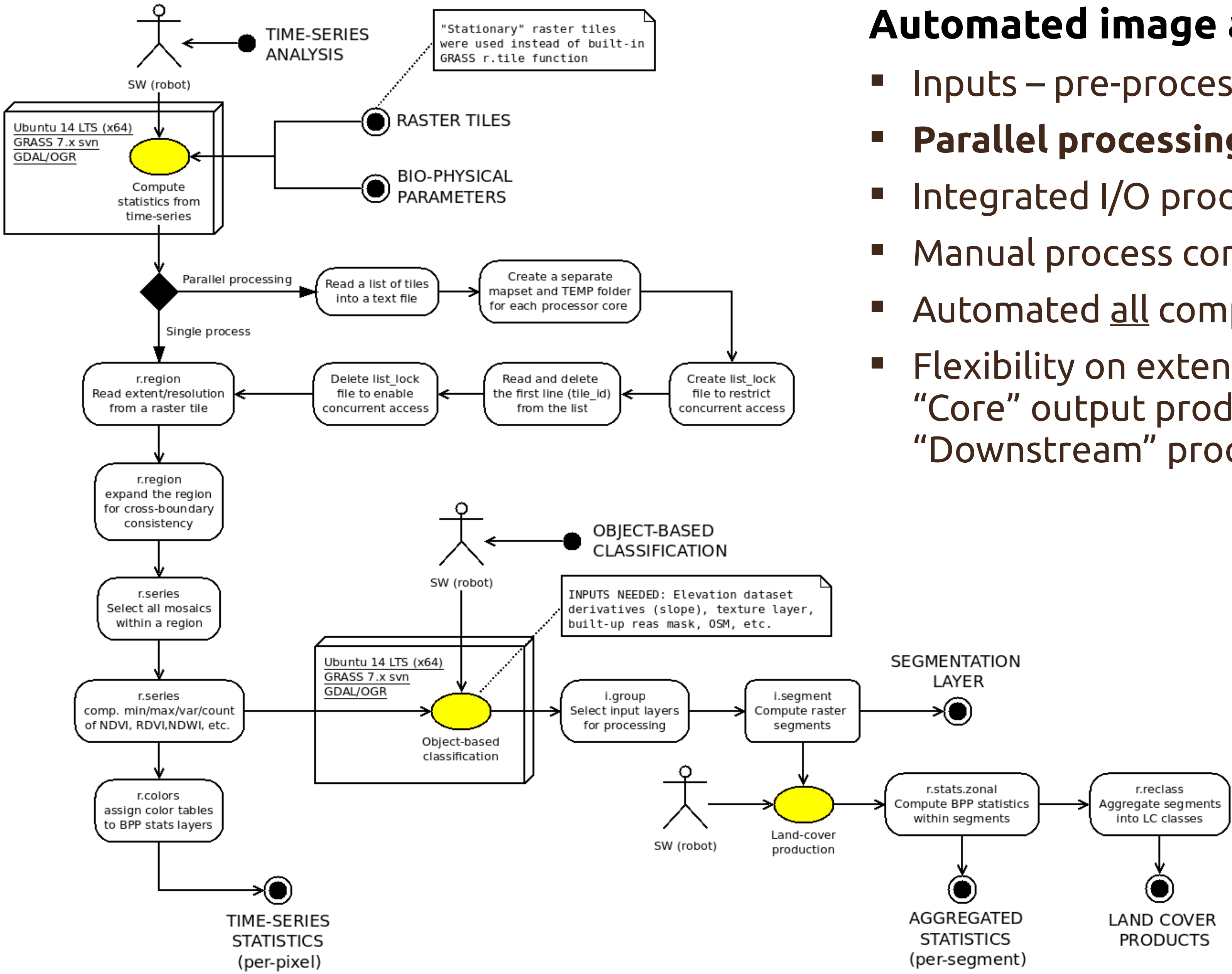
STEP 1 : PRE-PROCESSING

Automated image pre-processing:

- Outputs – production-ready images and mosaics



STEP 2 : ANALYSIS



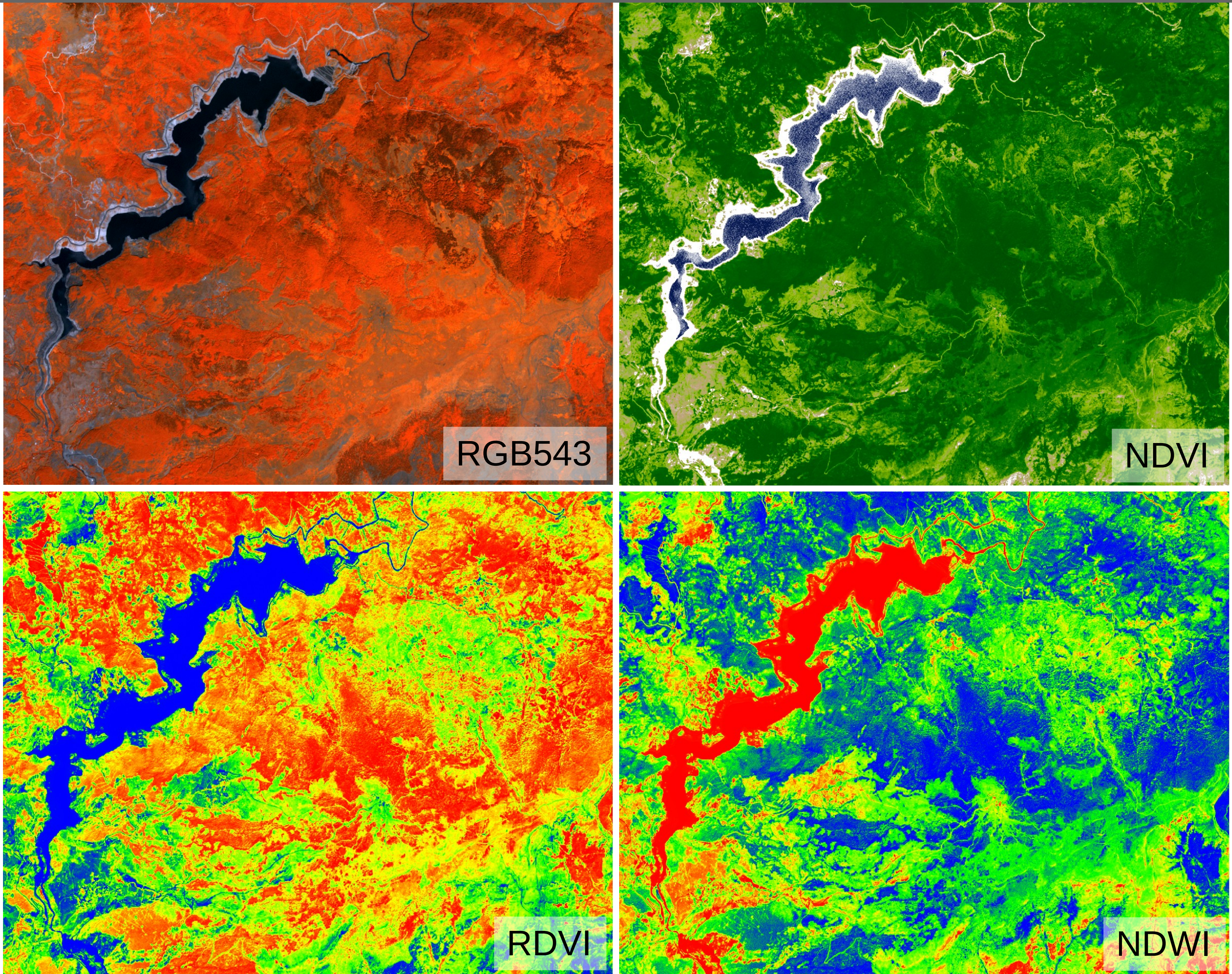
Automated image analysis:

- Inputs – pre-processed mosaics;
- **Parallel processing** in tiles;
- Integrated I/O products chain;
- Manual process control & QC;
- Automated <u>all</u> computing steps;
- Flexibility on extensions from “Core” output products to “Downstream” processes.

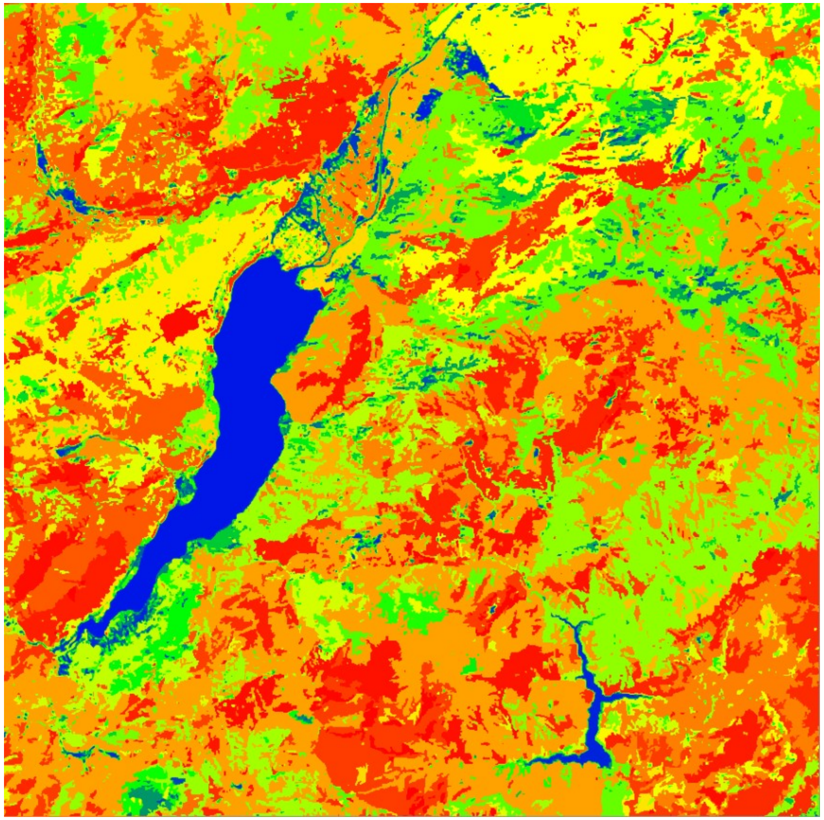
CONCEPT : PRE-CLASSIFIED OBJECTS

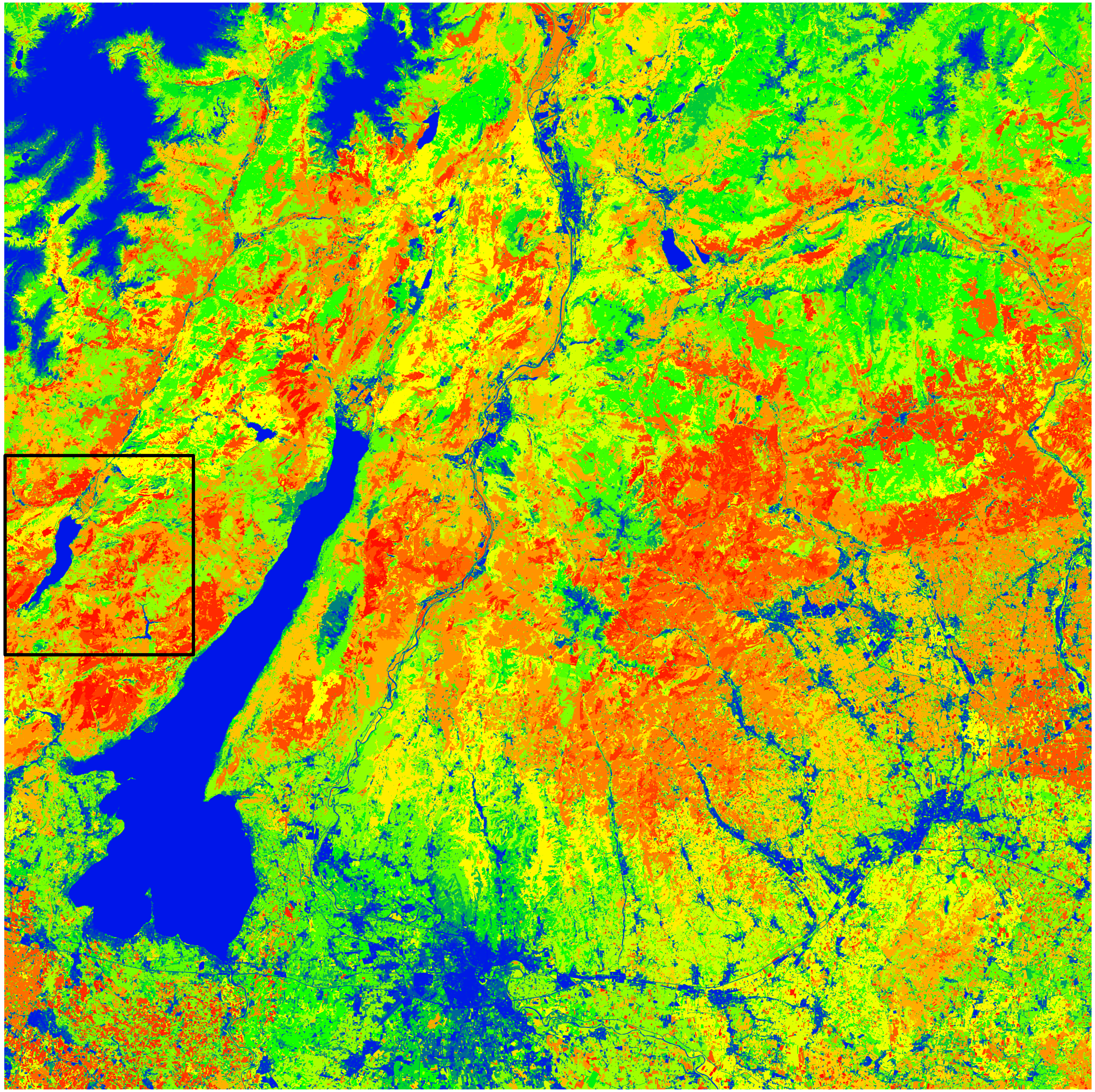
- Mosaicking of single-pass images into disconnected regions in order to reduce the number of iterations and avoid discrepancies of statistics;
- Computing of a broad range of biophysical parameters for the analysis of land cover structure and assessment of environmental conditions;
- Correction of the originally pan-sharpened VHR satellite images by computing piece-wise smooth approximation of each spectral band;
- Object-based classification of satellite images implemented as automated parallel processing work-flow and tested on pan-European coverage of HR RapidEye and VHR SPOT-5 imagery;
- Zonal statistics (per “object”) computed from a series of biophysical parameters, DTM and ancillary datasets are used for extraction of land cover classes by conditional filtering.

BIO-PHYSICAL PARAMETERS



OBJECT-BASED CLASSIFICATION

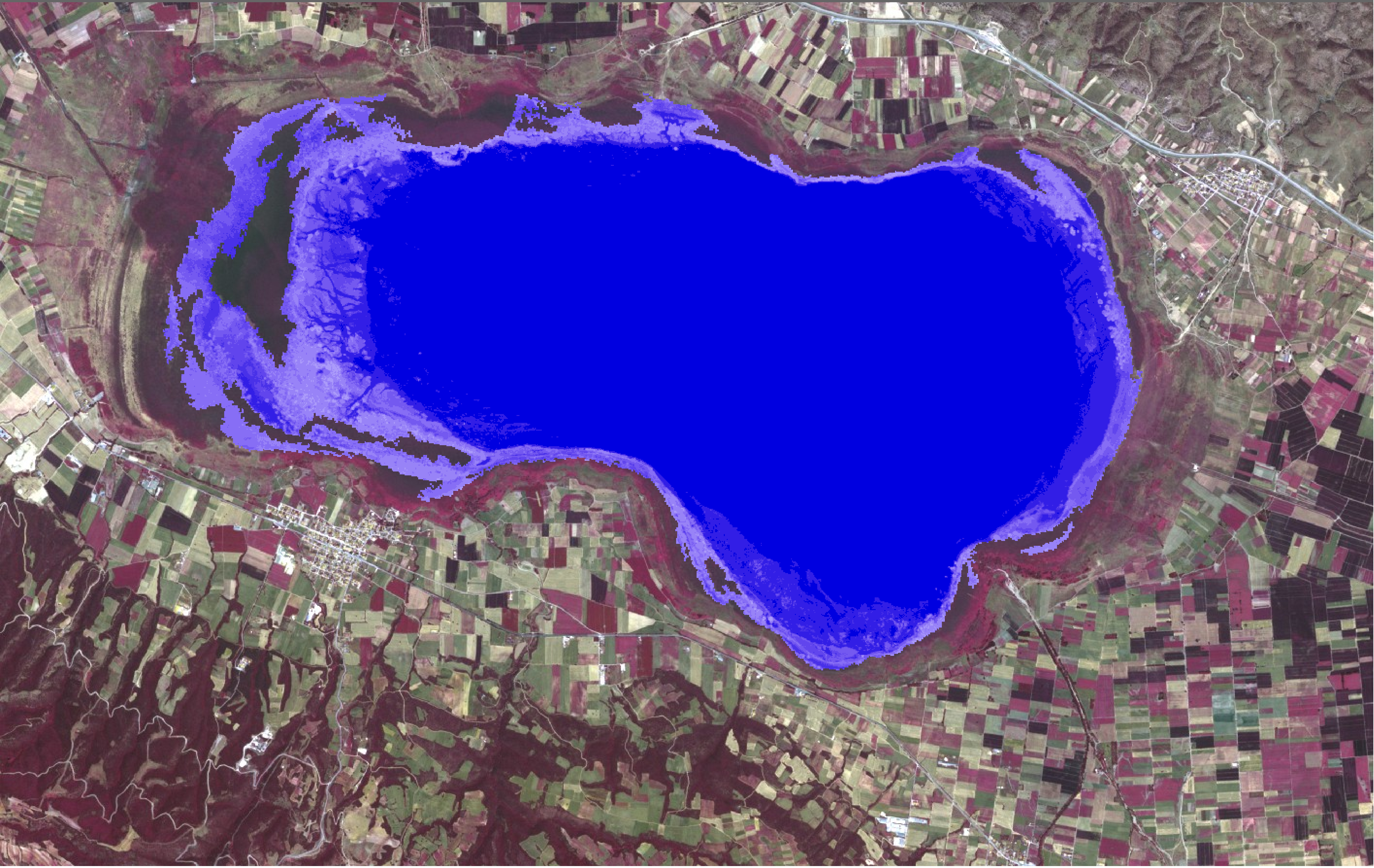




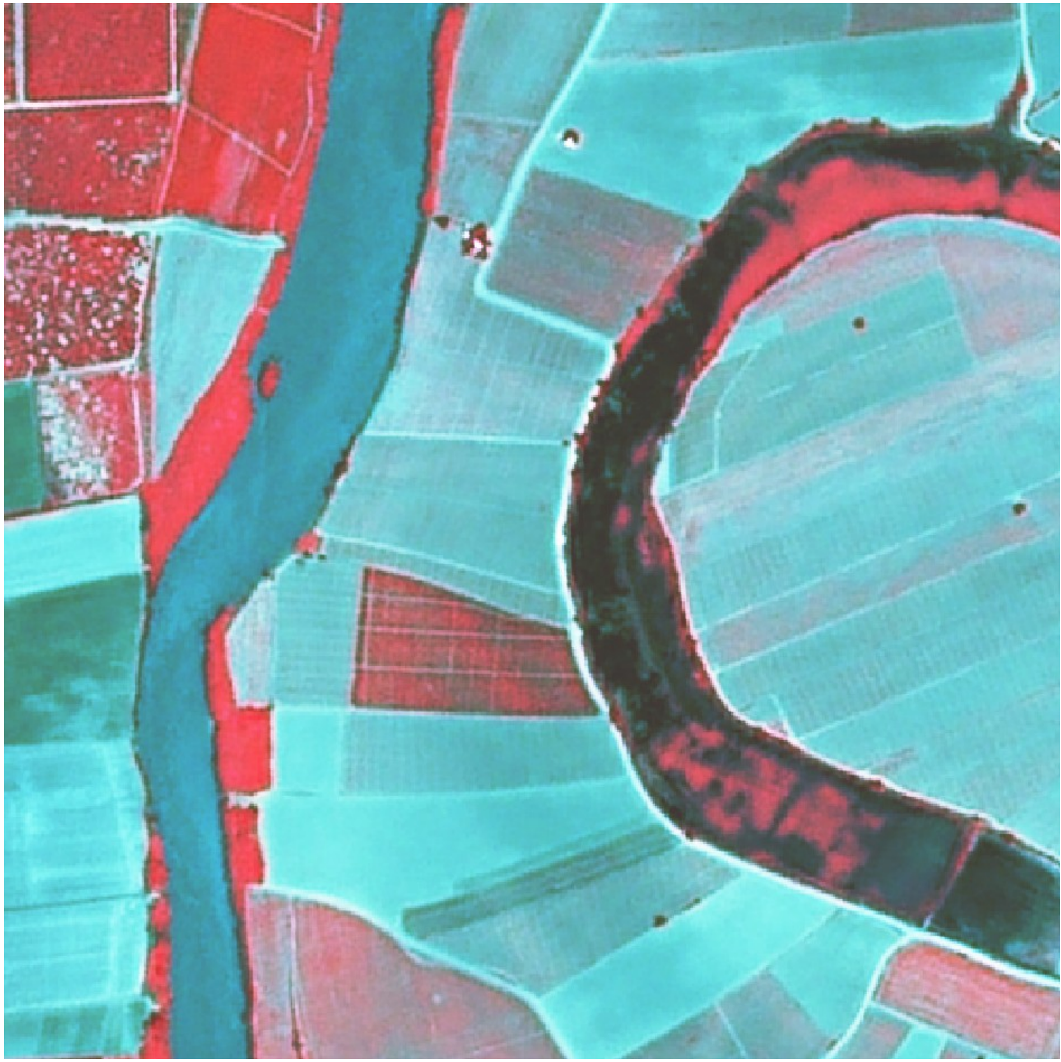
Automated object-based classification of combined layers with BPP and ancillary datasets produce standard “pre-classified” VHR land cover products useful for “core” LC analysis and mapping.

<u>Examples:</u> Biophysical LC of the Northern Italy (RapidEye imagery).

A COMBINED WATER OBJECT



CORRECTION OF VHR IMAGERY



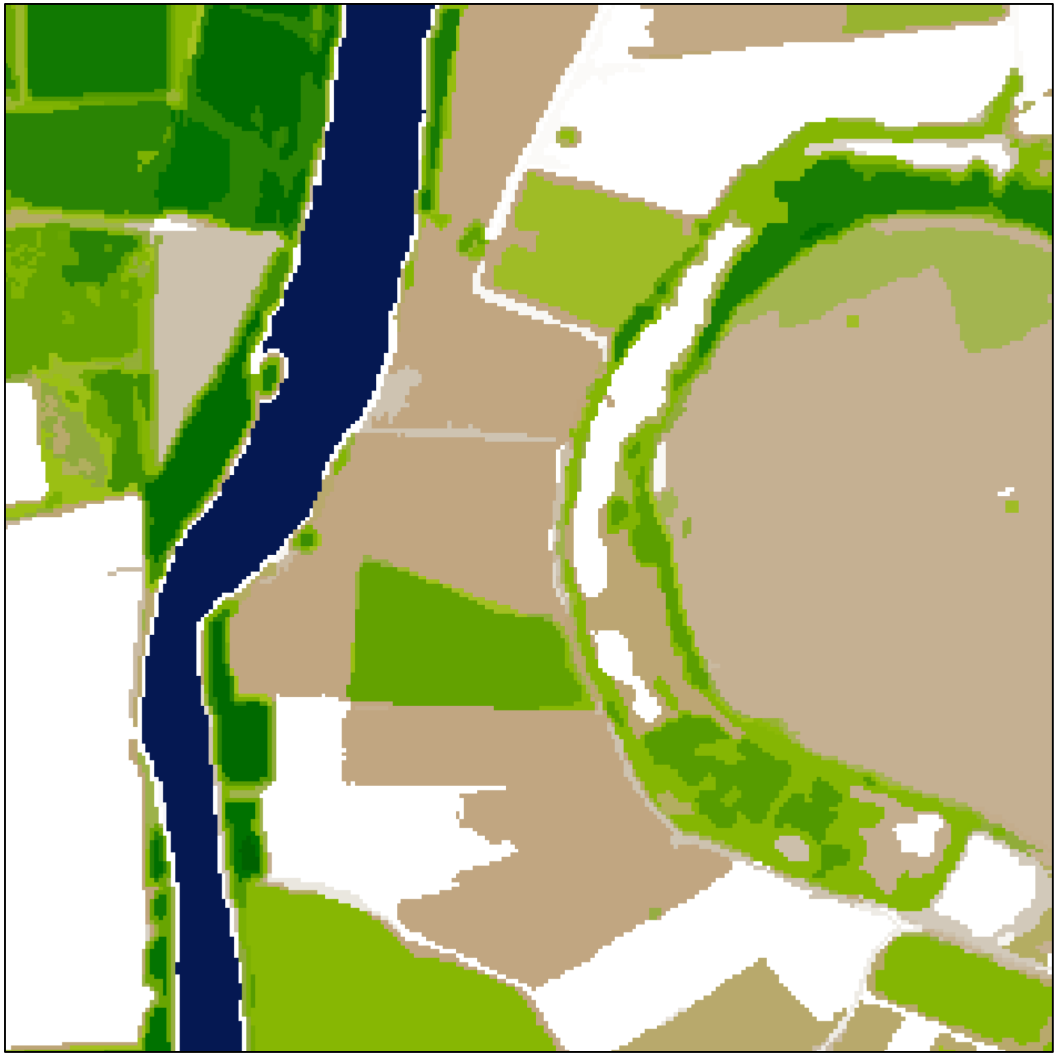
VHR SPOT 5 color-balanced RGB:123 composite image of Turkey taken on 2011/06/24 (DWH_MG2b_CORE_03): due to pan-sharpening, the image contains a significant amount of digital noise on homogeneous surfaces (especially water)

After piece-wise smooth approximation (segmentation) of the input raster map, most of the digital noise is removed on heterogeneous surfaces, yet the natural discontinuities are preserved

PROCESSING OF CORRECTED DATA

NDVI computed from the approximated image of Turkey taken on 2011/06/24.

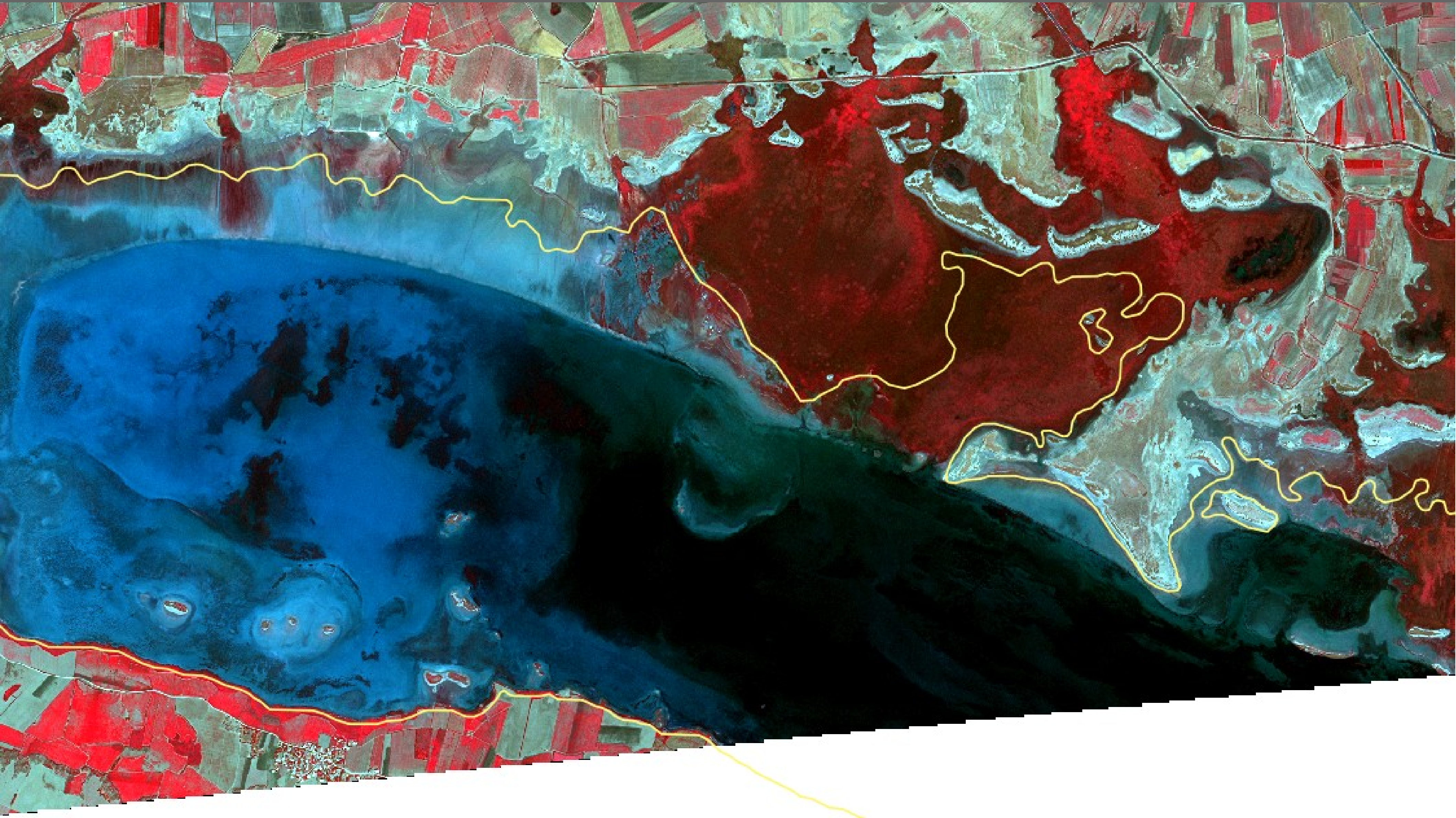
After piece-wise smooth approximation, the NDVI layer does not contain any traces of the digital noise, present in the original image.



Segmentation of the NDVI layer computed from the approximated image of Turkey taken on 2011/06/24.

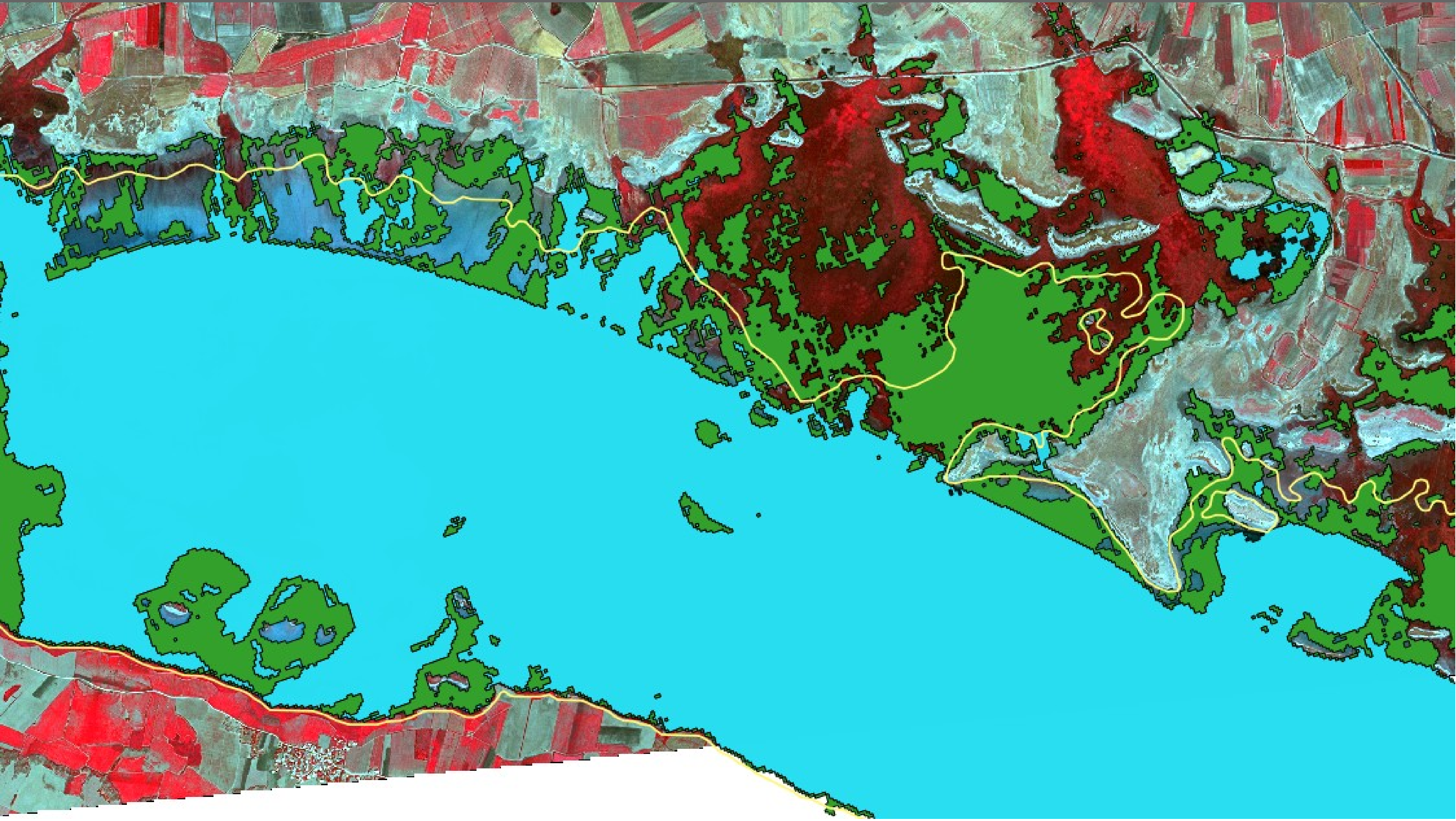
This pre-classified product can be used as input for extraction of water objects.

HR / VHR COMPARISON



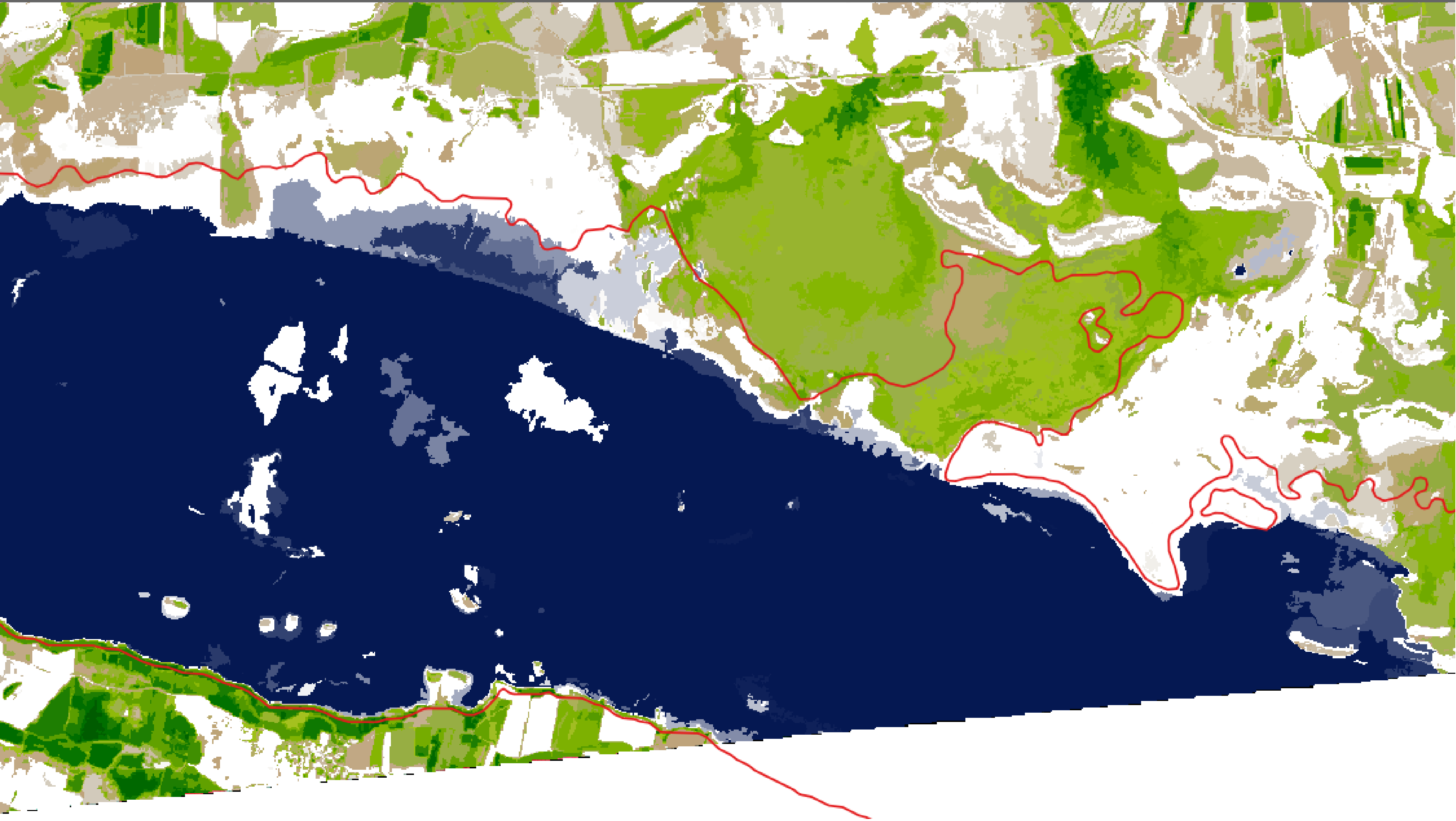
VHR SPOT 5 image of Turkey taken on 2011/06/24 (enhanced RDB 123 color composite; source: DWH_MG2b_CORE_03). Yellow line – old GRDA coastline.

HRL WATER & WETLAND



GIO-Land products: HRL Permanent Water (blue) and HRL Wetland (green) displayed on top of an enhanced VHR SPOT 5. Yellow line – old GRDA coastline.

VHR PRE-CLASSIFIED



VHR pre-classified BPP dataset produced from VHR SPOT 5 image of Turkey (2011/06/24). Red line – old GRDA coastline.

GEOMATRIX UAB
Extreme GEO-computing

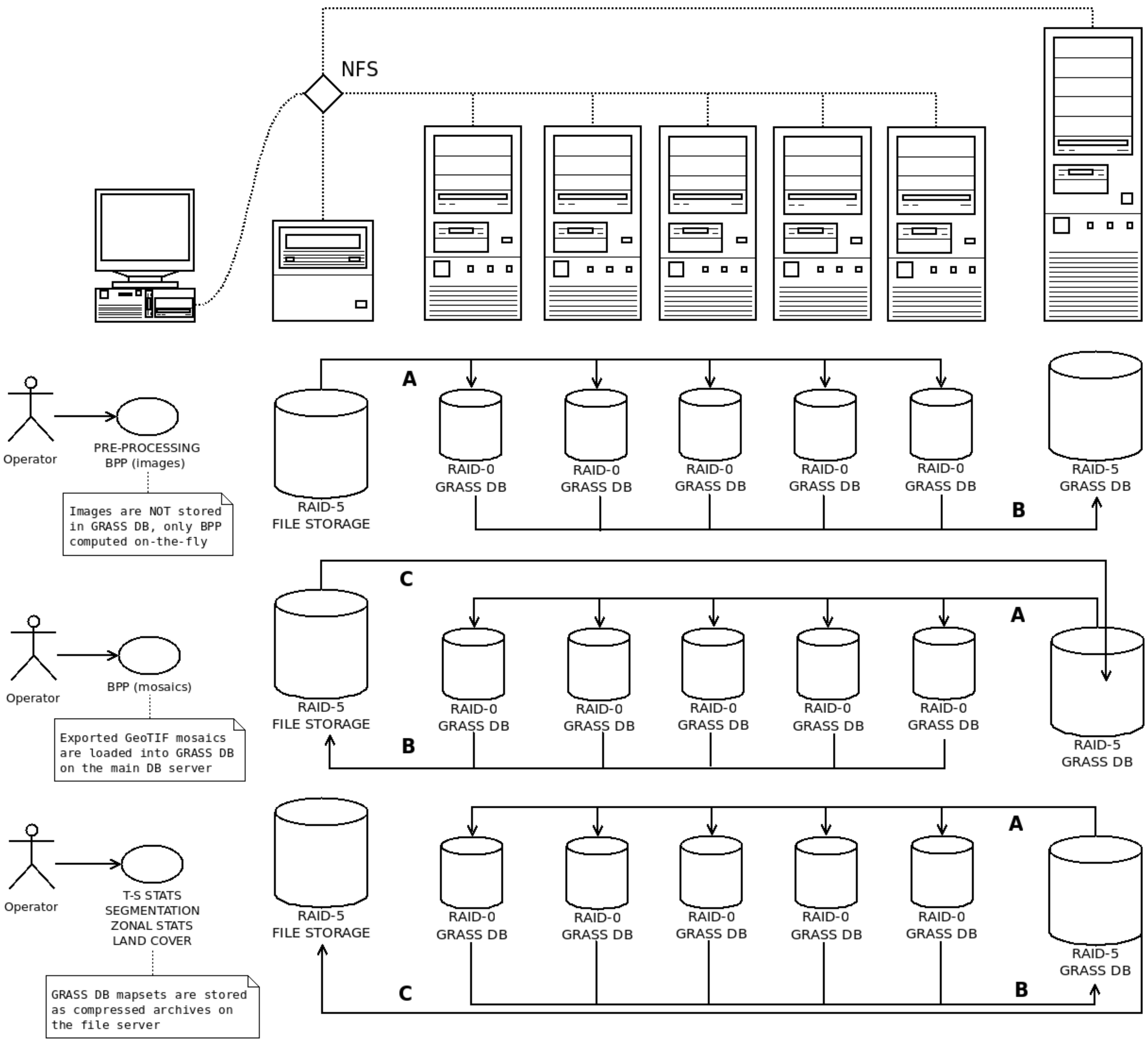
OPERATIONAL TESTING

A part of FP7 LAMPRE project tasks on technological developments:

- 7 processing workstations (i7, 32 Gb RAM, 4 Tb RAID-0, Ubuntu 14.04 LTS);
- 2 production servers:
  - GRASS DB server (i7, 16 Gb RAM, 26 Tb RAID-5, Ubuntu 14.04 LTS);
  - File/FTP/GeoNode server (i5, 16 Gb RAM, 16 Tb RAID-5);
- ~23,000 HR RapidEye images with 5 m pixel resolution and 12-bit data depth (IMAGE2012, Coverage-2) were used;
- ~5,000 VHR SPOT5/6 images with 2.5 m pixel resolution (DWH_MG2b_CORE_03) were used;
- EU-DEM elevation dataset was used;
- Parallel processing with GDAL & GRASS software implemented and tested.

GEOMATRIX UAB
Extreme GEO-computing

DATA HANDLING ISSUES



- NFS service was used for inter-connected BIG Data storage;
- On-the-fly transfer from storage through processors to the main GRASS DB;
- Only final products are stored in the system;
- No data is kept on the processing servers (vulnerable 4x RAID-0 spinning disk arrays);
- Archiving and removal of GRASS DB mapsets also allows easy transfer of large databases.

Source: FP7 LAPMPRE D4.3

CONSOLIDATED USE-CASES

Use-Cases:	Outputs:
Acquisition of EO images	Original EO datasets
Pre-processing of EO images (with manual ortho-rectification)	Production-ready GeoTIF images (packaged with metadata and previews)
Mosaicking of EO images	Unique (spatio-temporal) mosaics: - Satellite images (multi-spectral) - Derivative products (single-band)
Computing bio-physical parameters	Multiple biophysical products (multi-spectral data transformed into standard parametric data layers)
Computing time-series statistics	Layers of statistical values (statistical values computed from parametric data layers) – input for change detection
Image segmentation and discontinuity detection	Piece-wise smooth approximation of each spectral band – removal of pan-sharpening “noise”
Object-based classification	Segmentation layer (raster)
Zonal statistics for the objects	Aggregated statistics from ancillary datasets
Land-cover extraction	Conditional filtering into Land Cover classes