



GLOBCOVER First End Users Meeting

Classification & Validation

*UCL - Christelle Vancutsem, Jean-Francois Pekel,
Pierre Defourny*

Synoptics-Vixel - Leon Schouten



Classification concept

Key idea: Combine the high spatial consistency of classes obtained from multispectral composite(s) with the great land cover discrimination provided by temporal profiles analysis

Principles:

- Regionally-tuned approach based on **22 equal-reasoning area**
- Multispectral composites and **reflectance** time series
- Typology** defined and documented using **FAO Land Cover Classification System** and as much as possible compatible with EEA CORINE Land Cover product and GLC2000.
- International experts inputs for the classification algorithm calibration

Challenges:

- ⇒ Consistent but regionally-tuned classification processing strategy
- ⇒ Automatic and repeatable processing chain
- ⇒ Building on GLC2000 experience and previous LC products



Legend : 23 LCCS classes

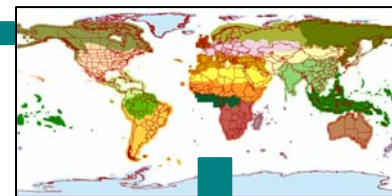
(46 including regional classes)

Global Globcover legend (Level 1)	
	A11 - Cultivated Terrestrial Areas and Managed Lands
1	Cultivated and Managed areas
2	Mosaic cropland (50-70%) / vegetation (grassland/shrubland/forest) (20-50%)
3	Mosaic vegetation (grassland/shrubland/forest) (50-70%) / cropland (20-50%)
	A12 - Natural and Semi-natural Terrestrial Vegetation
4	Closed to open (>15%) broadleaved evergreen or semi-deciduous forest (> 5m)
5	Closed (>40%) broadleaved deciduous forest (>5m)
6	Open (15-40%) broadleaved deciduous forest/woodland (>5m)
7	Closed (>40%) needle-leaved evergreen forest (>5m)
8	Closed (>40%) needle-leaved deciduous forest (>5m)
9	Open (15-40%) needle-leaved deciduous or evergreen forest (>5m)
10	Closed to open (>15%) mixed broadleaved and needleleaved forest
11	Mosaic forest or shrubland (50-70%) and grassland (20-50%)
12	Mosaic grassland (50-70%) and forest or shrubland (20-50%)
13	Closed to open (>15%) shrubland (<5m)
14	Closed to open (>15%) herbaceous vgt (grassland, savannas or Lichens/Mosses)
15	Sparse (<15%) vegetation
	A24 - Natural and Seminatural Aquatic Vegetation
16	Closed to open (>15%) broadleaved forest regularly flooded (permanently or temporarily), fresh or brakish water
17	Closed (>40%) broadleaved forest or shrubland permanently flooded, saline or brackish water
18	Closed to open (>15%) grassland or shrubland or woody vgt regularly flooded or on waterlogged soil, fresh, brakish or saline water
	B15-Artificial Surfaces and Associated Areas
19	Artificial surfaces and associated areas (Urban areas >50%)
20	Bare areas
	B28-Inland Waterbodies
21	Water Bodies
22	Permanent Snow and Ice
23	No Data (burnt areas, clouds,...)

Classification Steps

Step 0: A priori stratification

⇒ Split the world in 22 equal-reasoning regions from ecological and remote sensing point of view



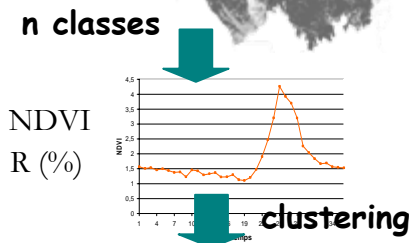
Step 1: For each region, per-pixel sup- & unsupervised classifications

⇒ Homogeneous land cover classes



Step 2: Per-pixel temporal characterization

⇒ Robust temporal metrics computed per-object from bi-weekly multispectral reflectance composites and NDVI



Step 3: Per-object classification algorithm

⇒ Consistent unlabelled spectro- temporal classes



Step 4: Labelling rule-based procedure

⇒ Based on best existing products and LCCS land cover classes



Step 5: Calibration

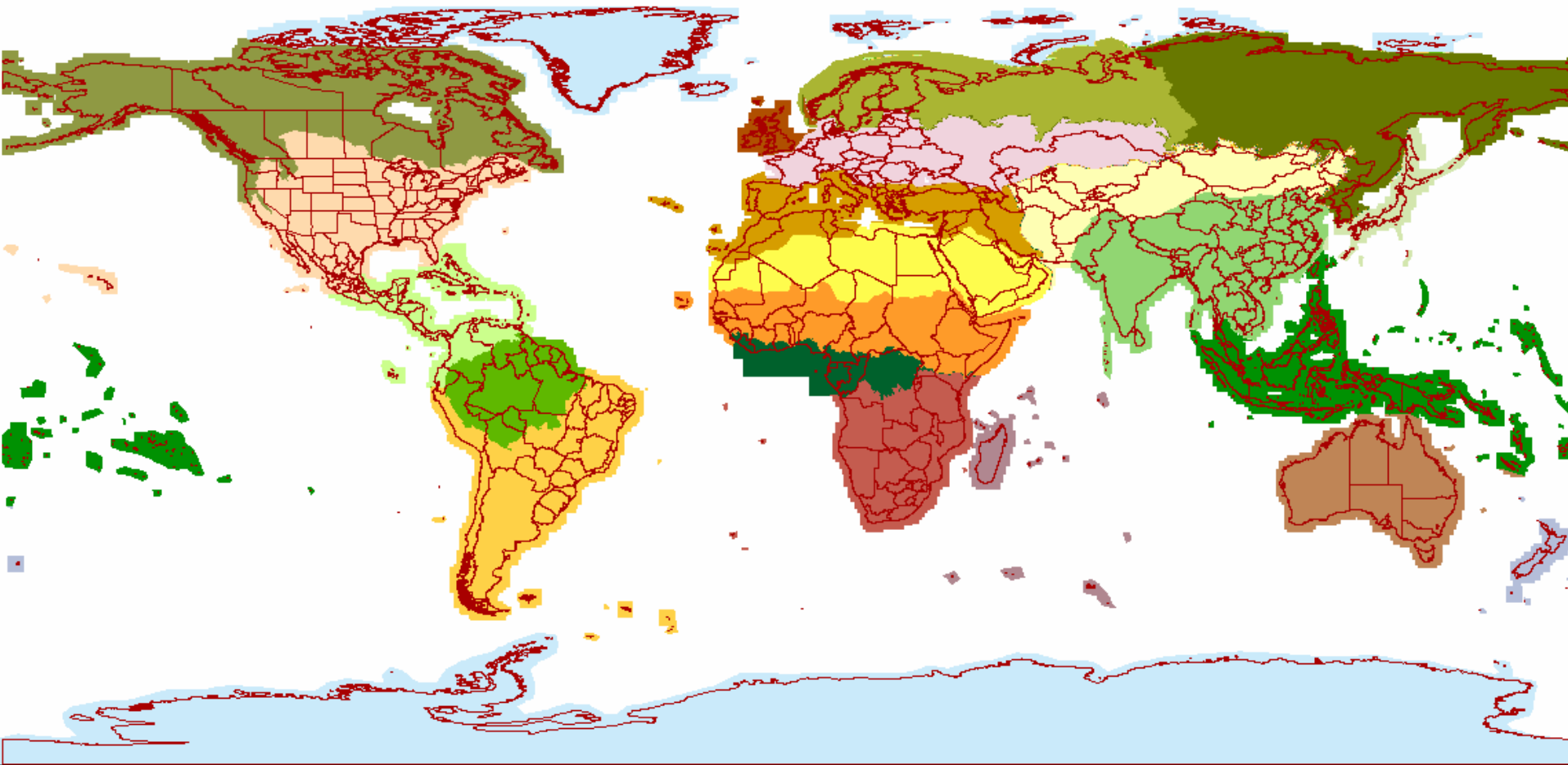
⇒ MERIS specific labelling thanks to interactive calibration by a network of international experts



Step 6: Independent Validation ⇒ Classification accuracy

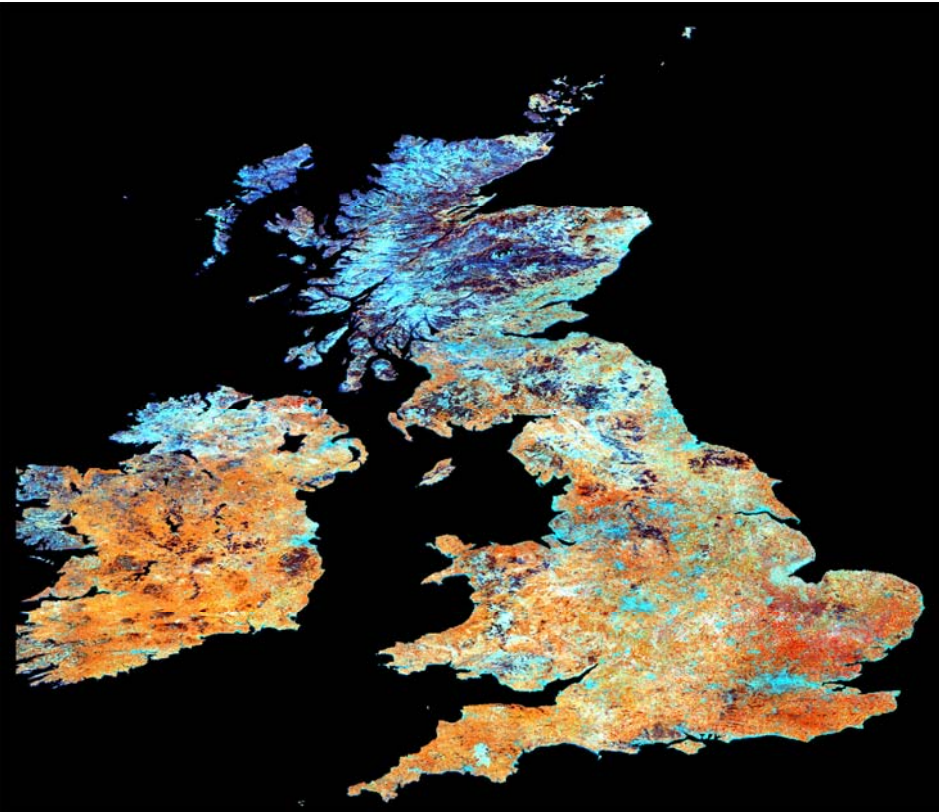
A priori stratification

4 strata to cover Europe



Globcover seasonal products dedicated for each stratum

Great Britain and Ireland



April – June (14,7,5)



September – November (14,7,5)

First End Users Meeting Ispra,
June 20, 2007

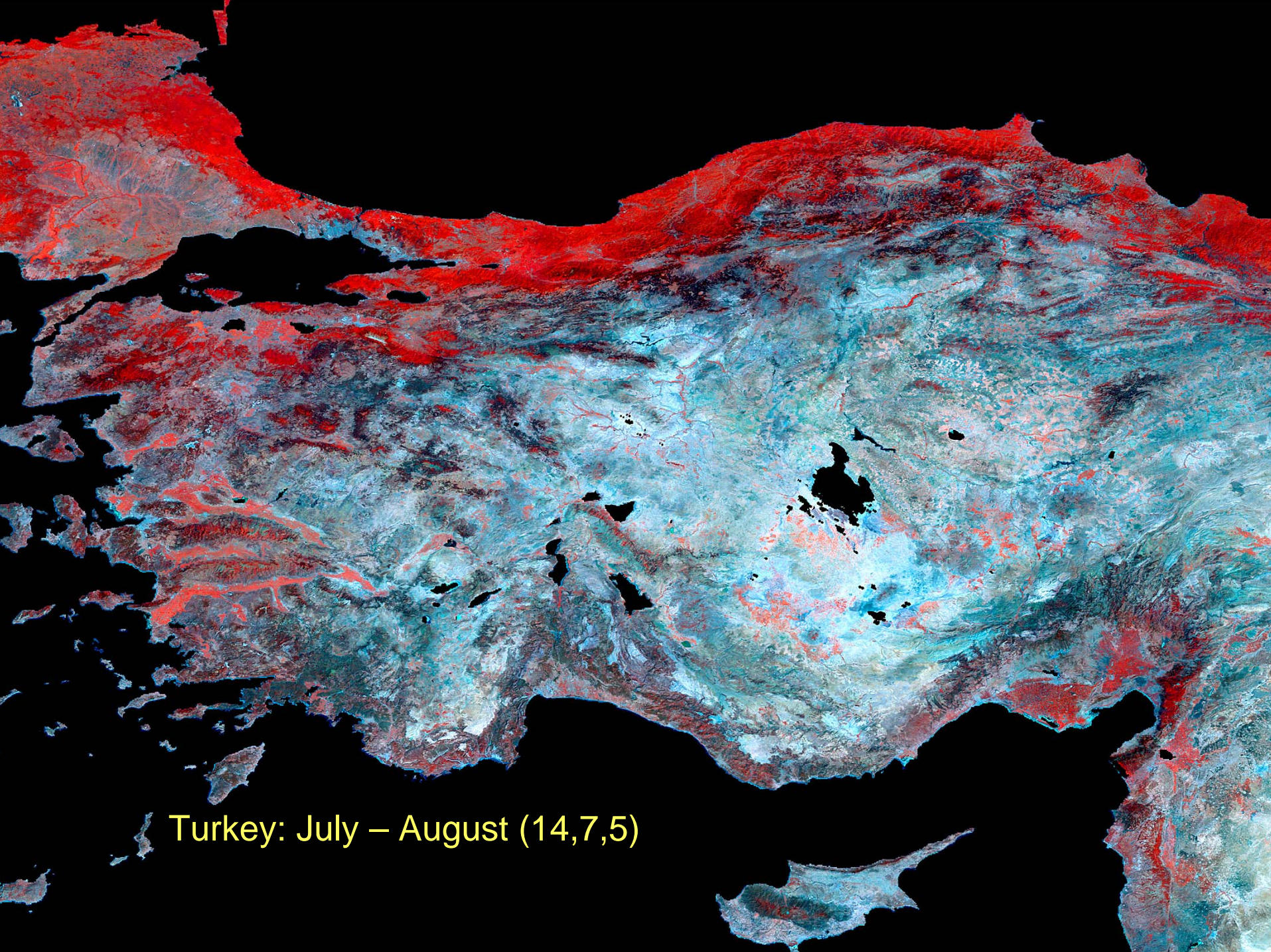
GLOBCOVER seasonal products

Romania : April – June (14,7,5)

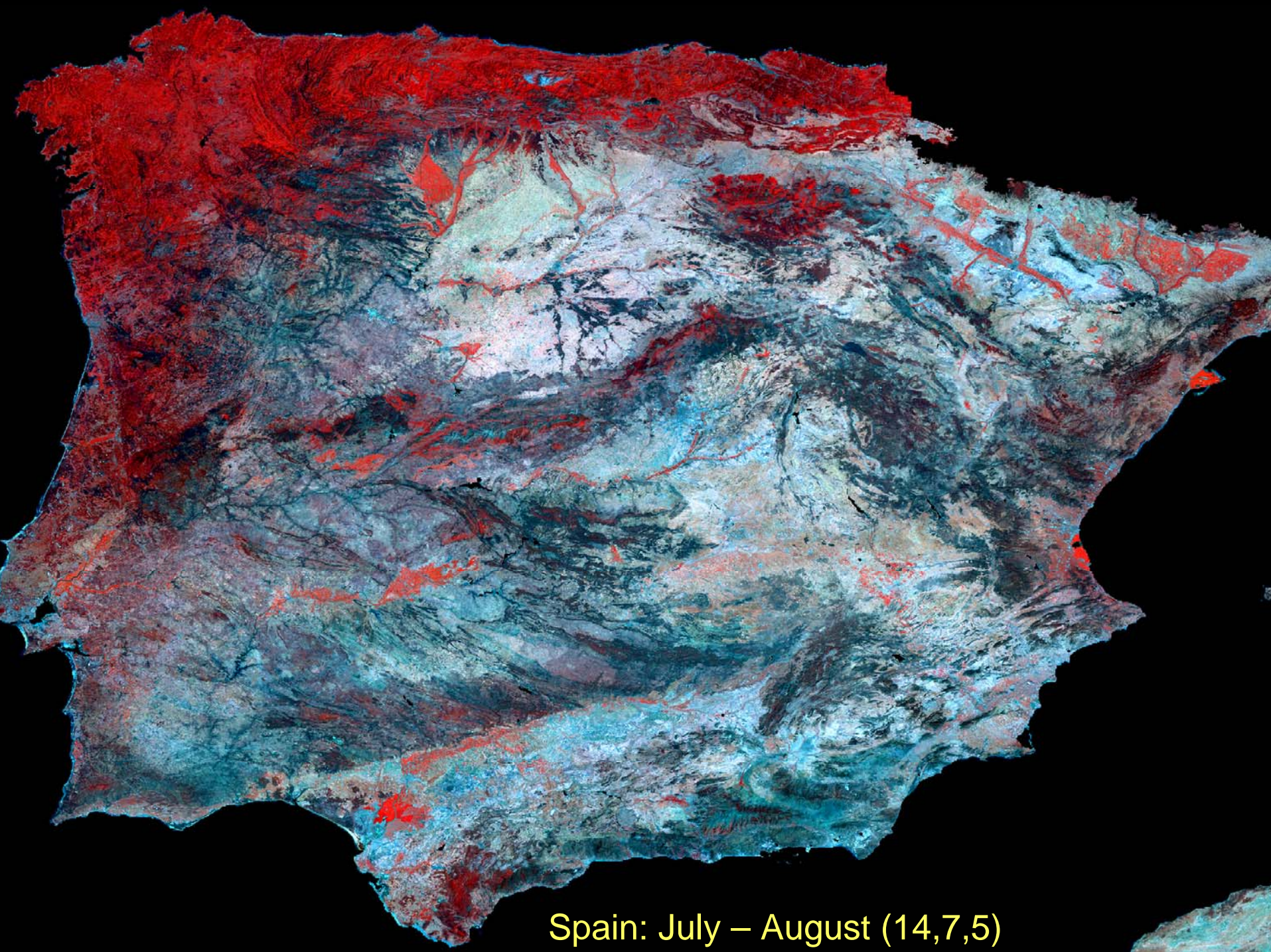




France - Belgium
April – June
(14,7,5)



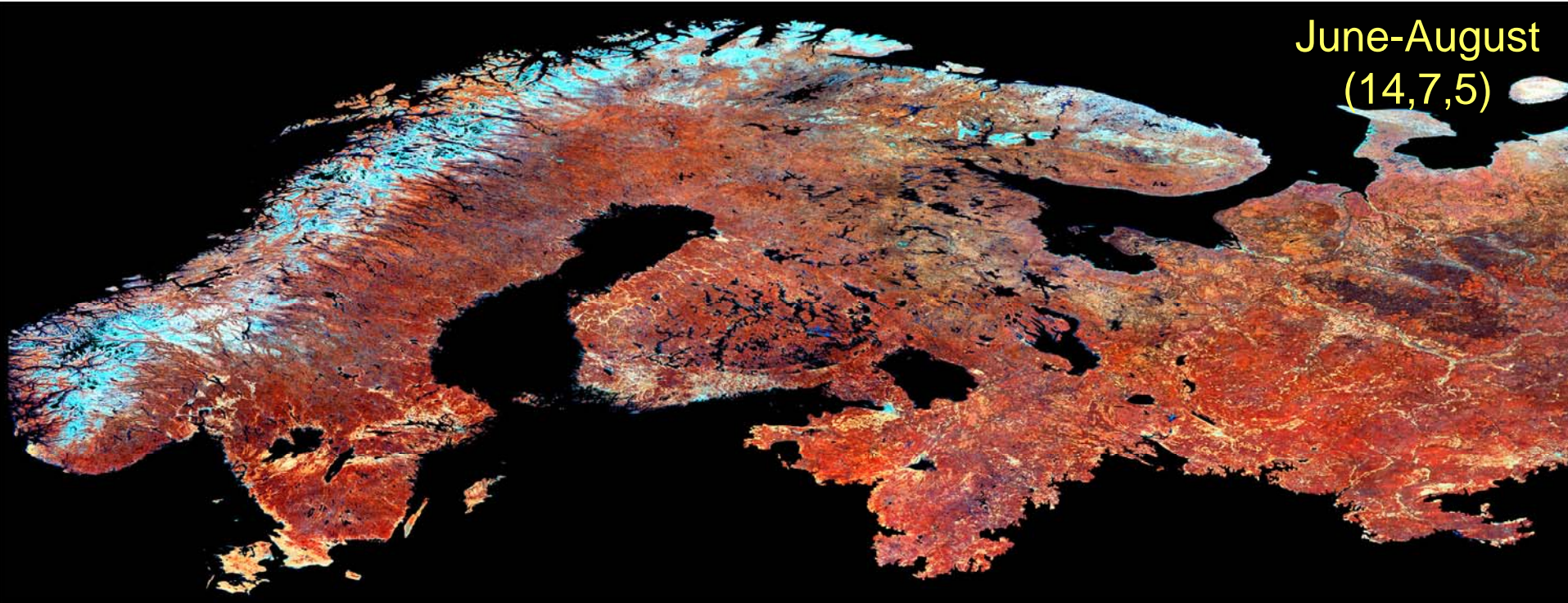
Turkey: July – August (14,7,5)



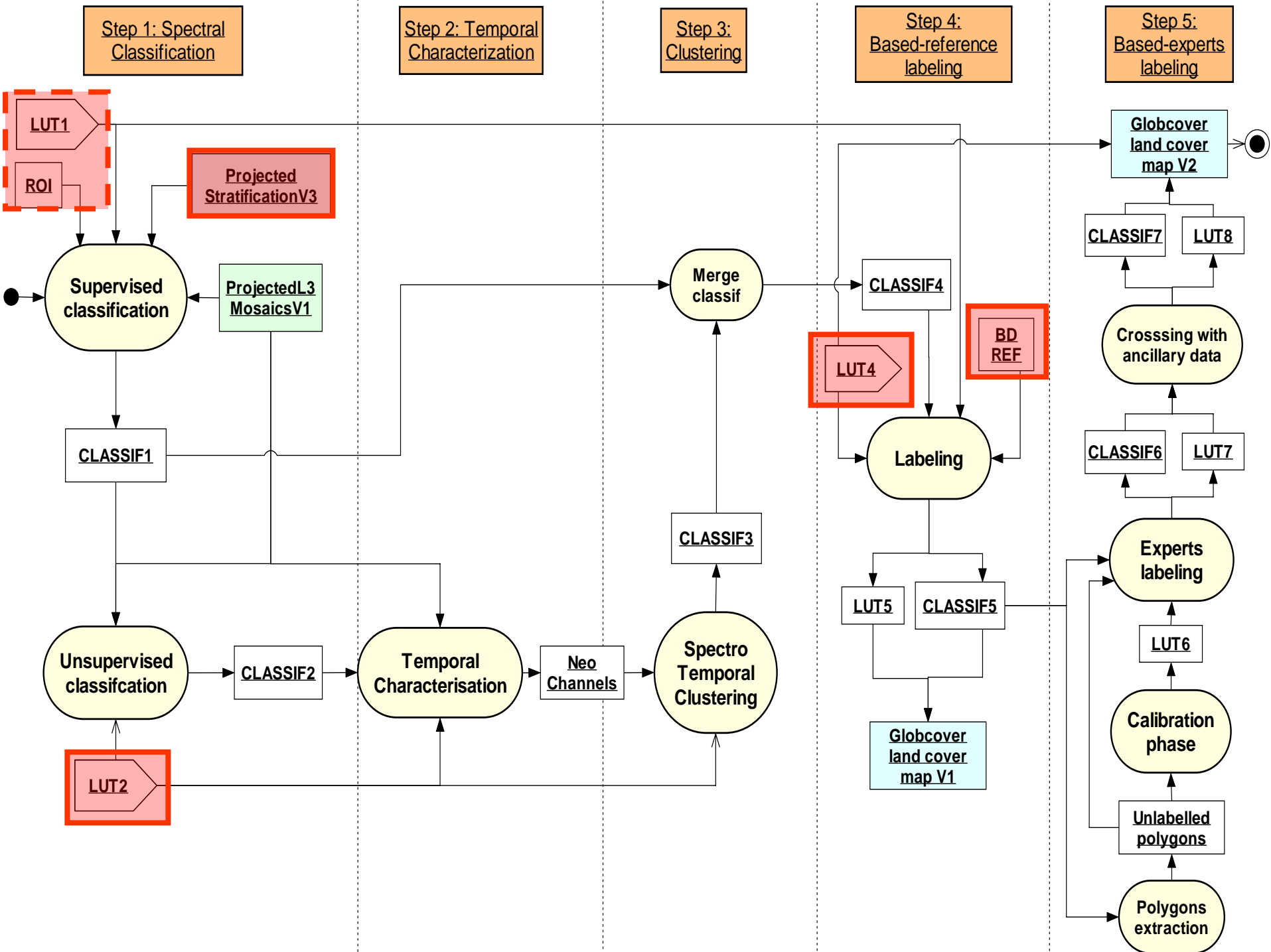
Spain: July – August (14,7,5)



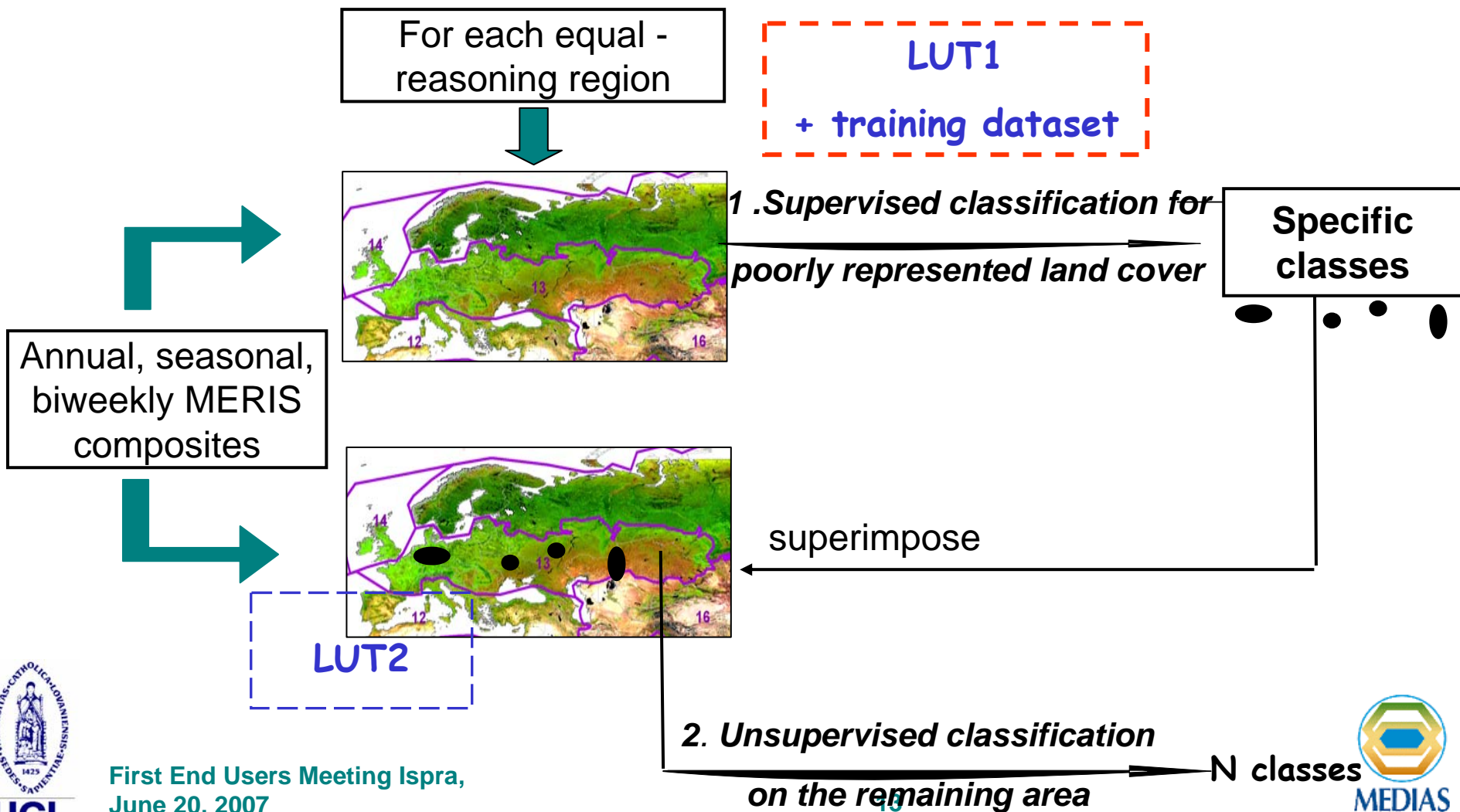
Scandinavia
April–October
(14,7,5)



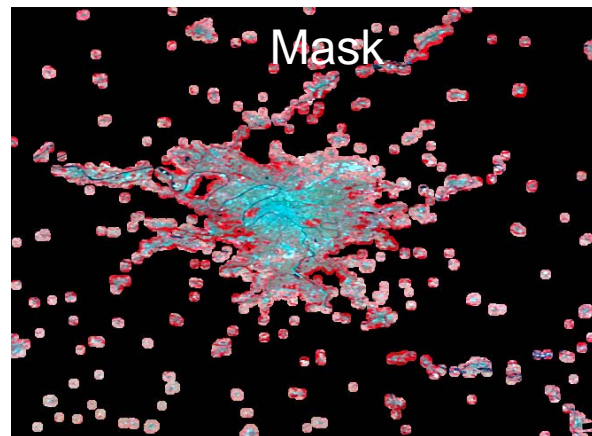
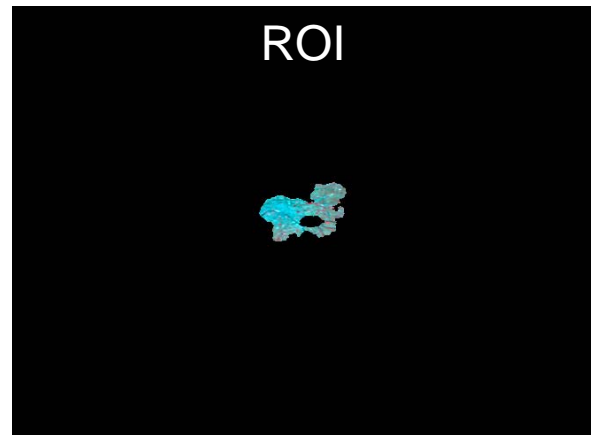
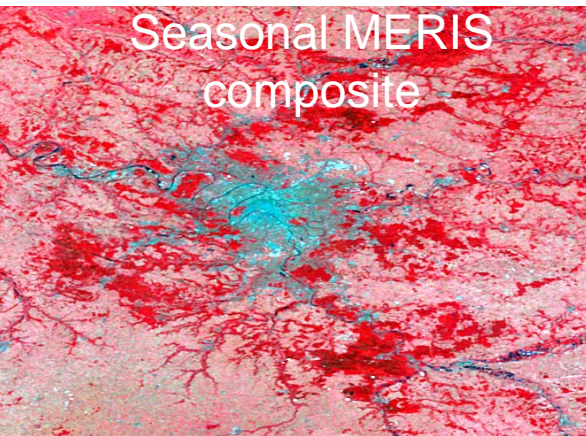
June–August
(14,7,5)



Step 1 - Per-pixel Classification using supervised and unsupervised algo.



Supervised classification approach

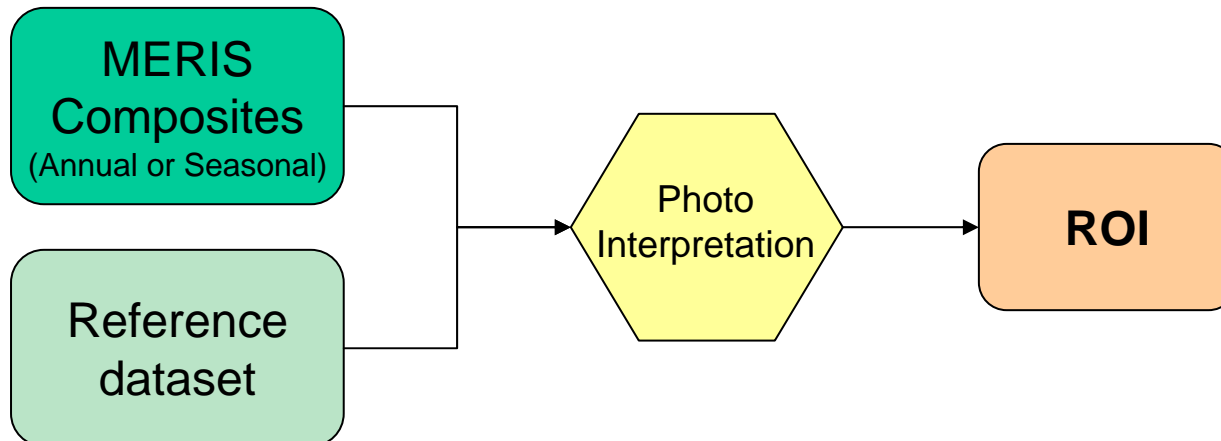


Supervised classification

Training set delineation (ROI)

For each region and for each poorly represented land cover class:

- Labelling reference dataset used to localise this land cover classes
- ROI visually delineated based on MERIS syntheses (seasonal or annual)

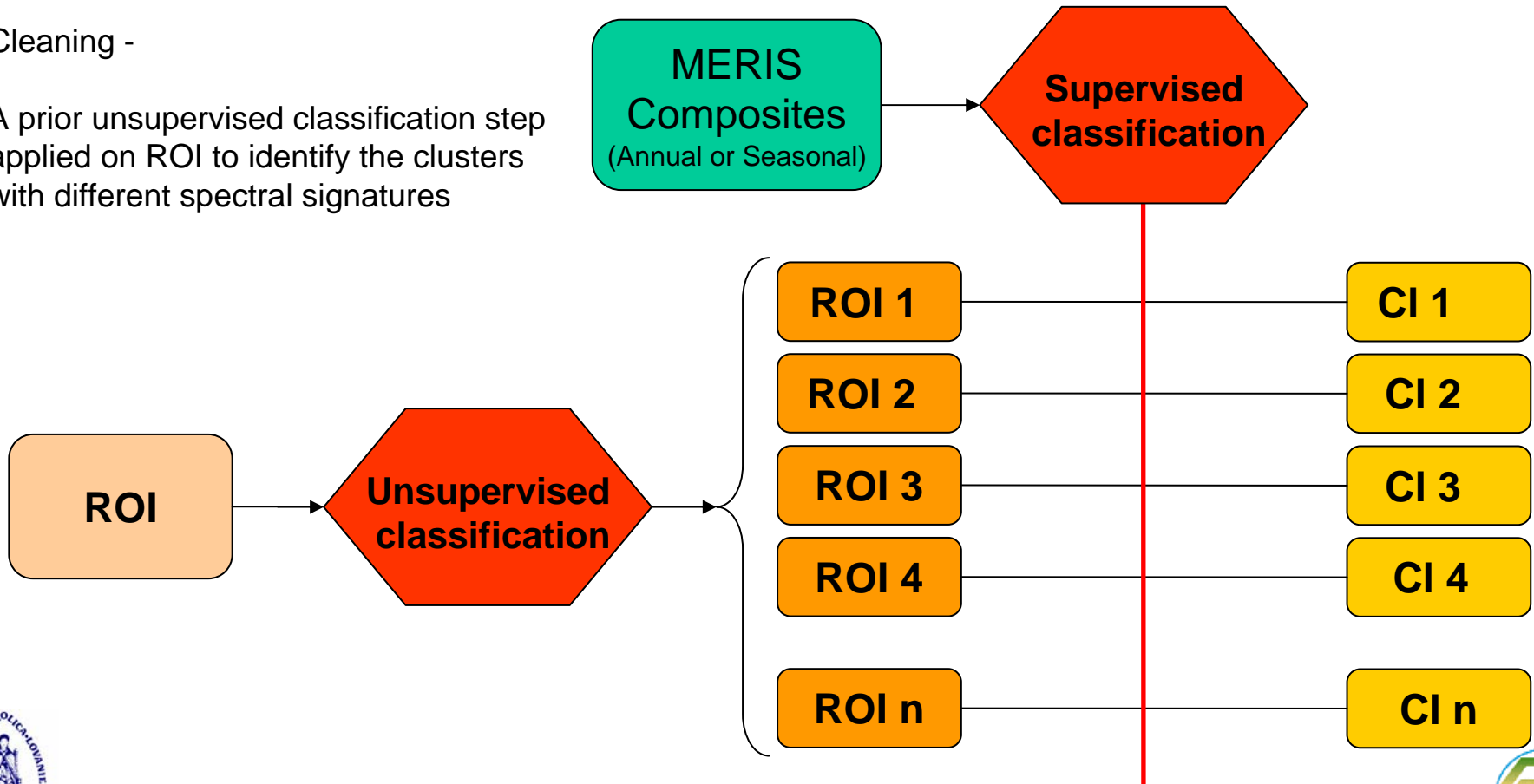


Supervised classification

Training set delineation (ROI)

Cleaning -

A prior unsupervised classification step applied on ROI to identify the clusters with different spectral signatures

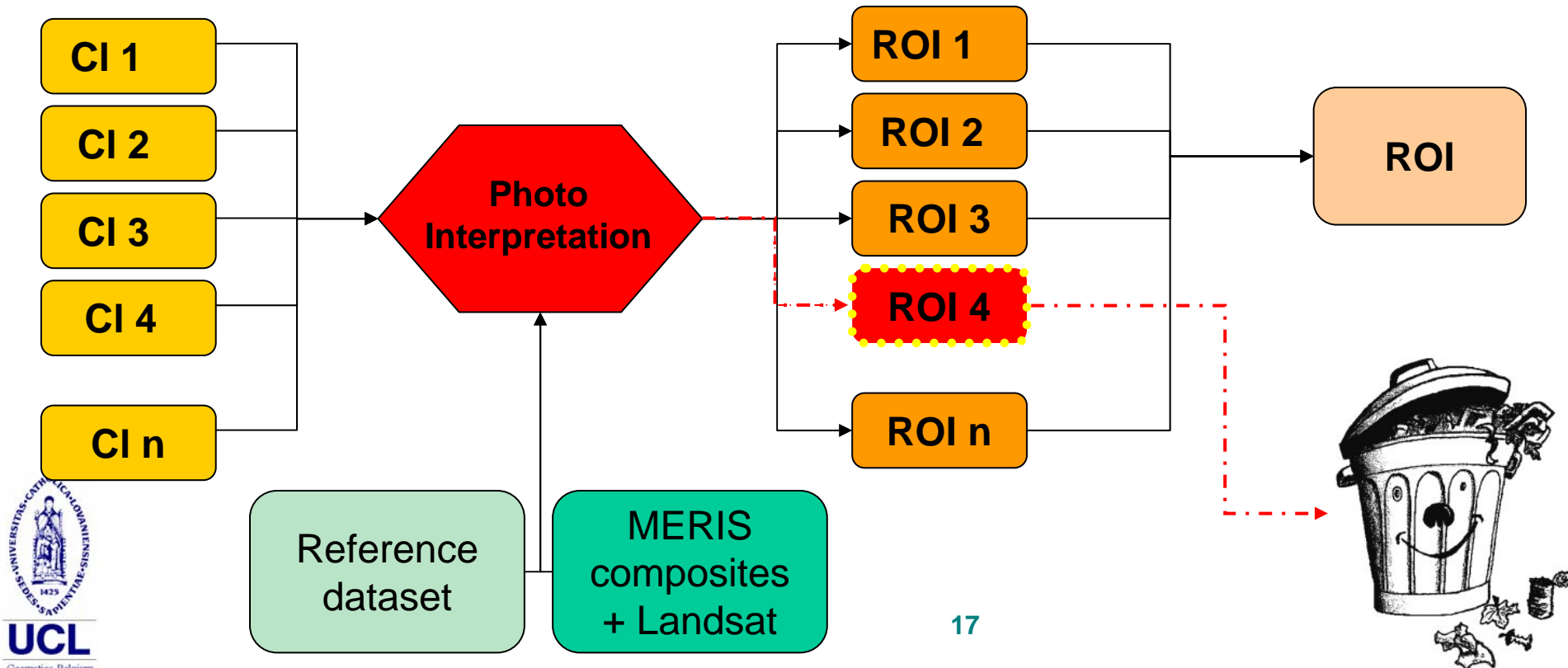


Supervised classification

Training set delineation (ROI)

Cleaning -

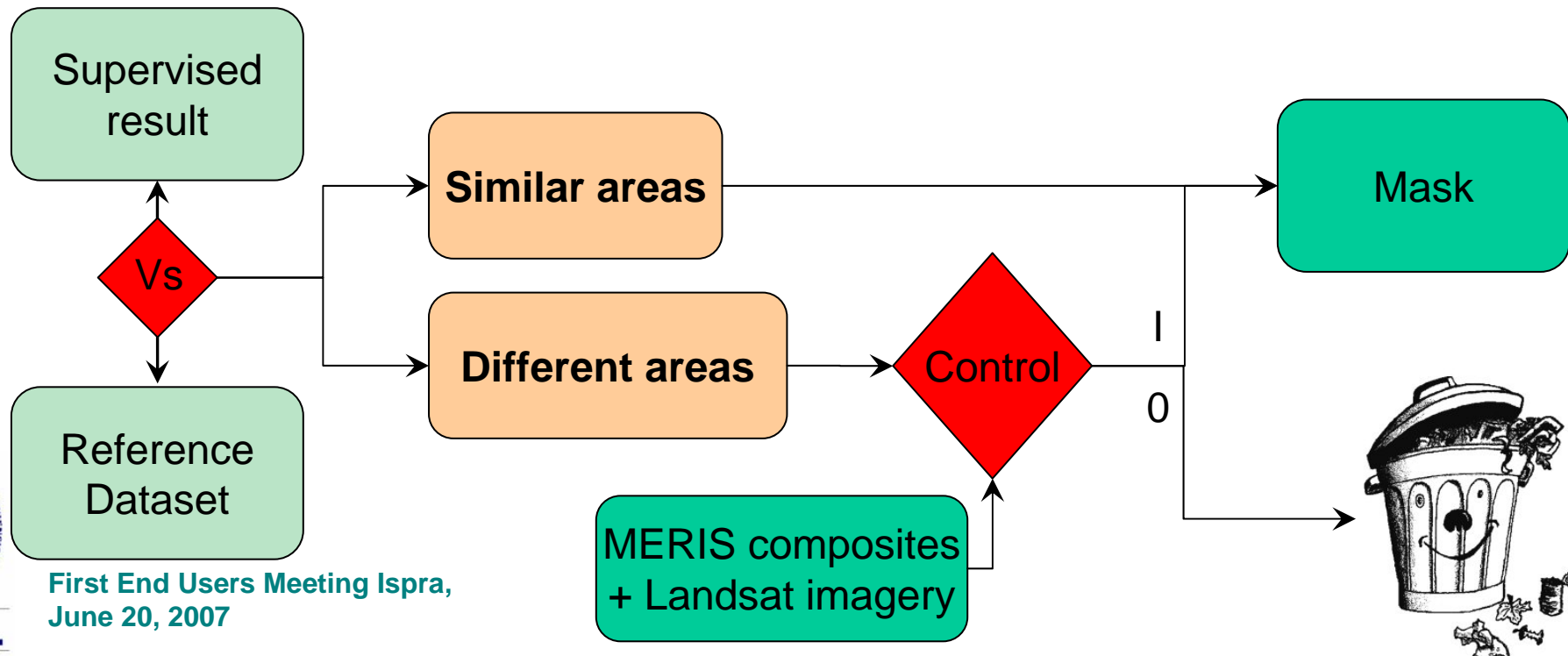
- Identification of erroneous pixels not corresponding to the land cover based on MERIS and Landsat images. If these pixels come from a specific sub ROI x, this one is removed from the ROI dataset
- Supervised classification process applied a second time and cleaning of ROI dataset again if necessary



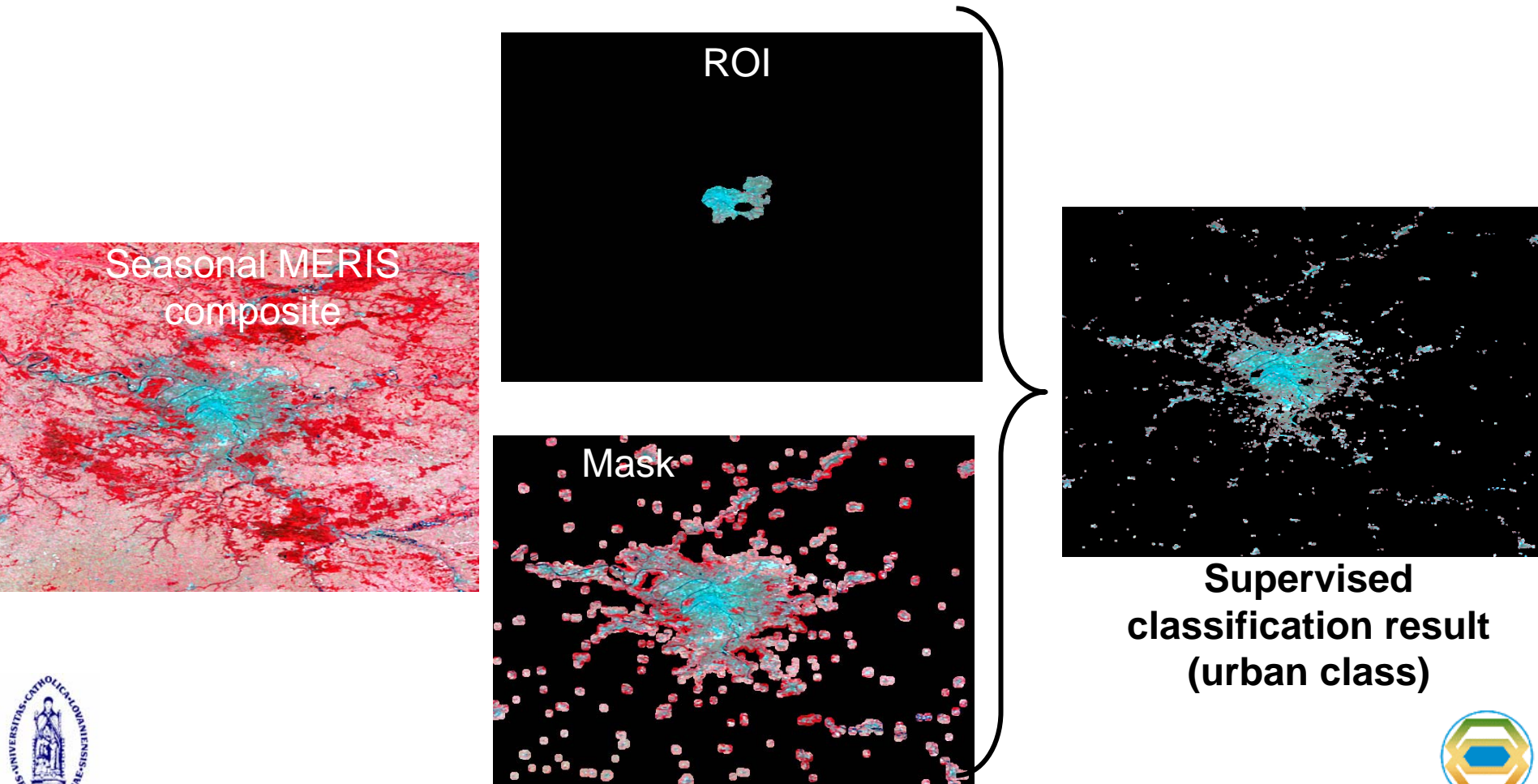
Supervised classification MASK delineation

For all the land cover types concerned by the supervised classification step:

- Comparison of supervised classification results to the dilated land cover reference dataset
- Screen out of the difference between the 2 products based on MERIS and Landsat images
- Cleaning of the products and merged to produce the MASK

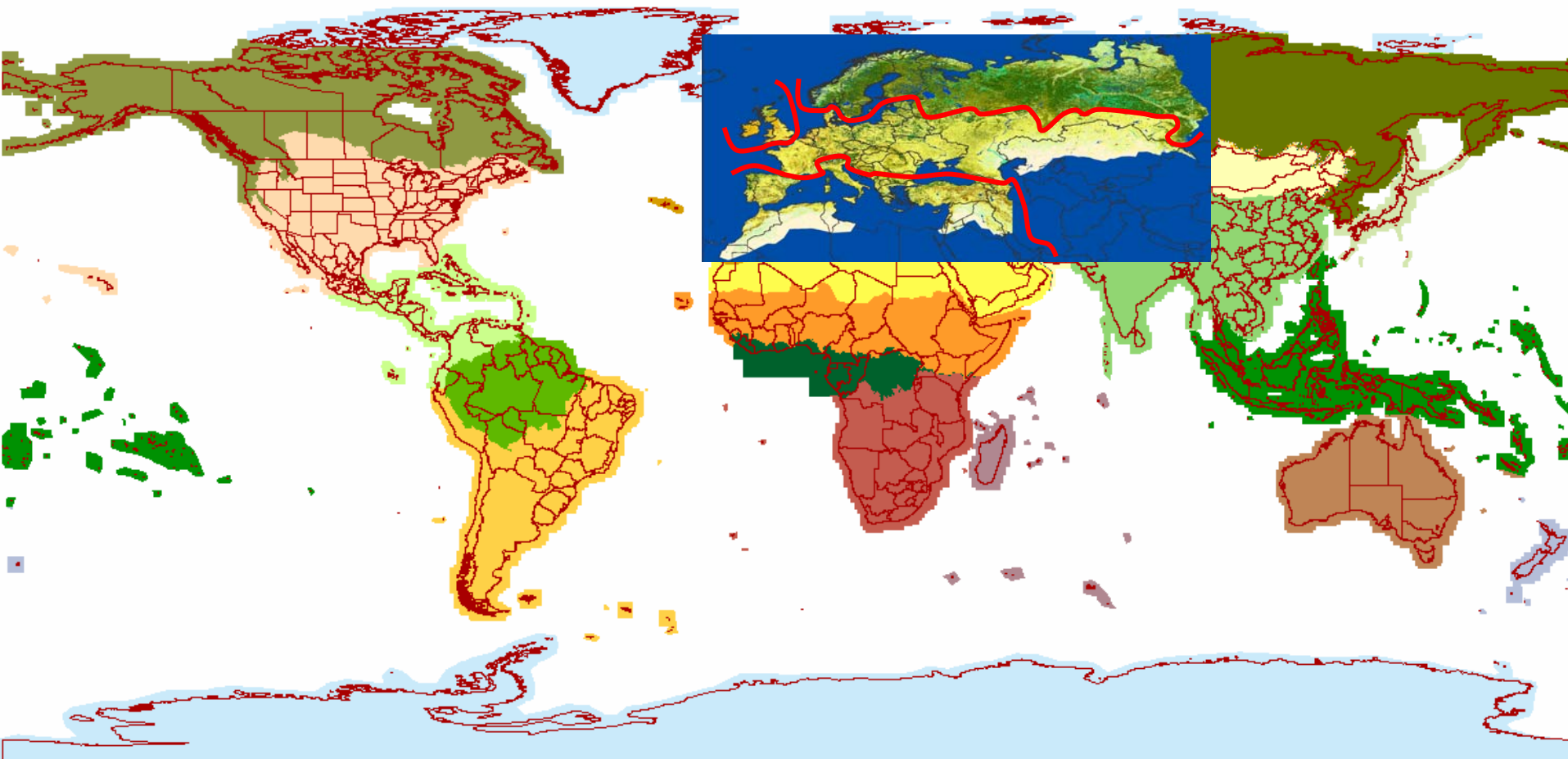


Supervised classification approach



First classification results

4 regions



First classification results

4 regions to cover Europe



First classification results

4 regions

- 10 - Cultivated areas
- 11 - Post-flooding or irrigated croplands
- 14 - Rainfed croplands
- 20 - Mosaic cropland - vgt
- 21 - Mosaic cropland - grassland
- 30 - Mosaic vgt - cropland
- 31 - Mosaic grassland/shrubland - cropland
- 32 - Mosaic forest - cropland
- 50 - Closed broadleaved deciduous forest
- 70 - Closed needle-leaved evergreen forest
- 90 - Open needle-leaved deciduous or evergreen forest
- 92 - Open needle-leaved evergreen forest
- 100 - Closed to open mixed forest
- 101 - Closed mixed broadleaved forest
- 110 - Mosaic forest/shrubland and grassland
- 130 - Closed to open shrubland
- 131 - Closed to open evergreen shrubland
- 140 - Closed to open grassland
- 141 - Closed grassland
- 150 - Sparse vgt
- 151 - Sparse grassland
- 180 - Flooded grass/shrub/woody vgt
- 190 - Urban areas
- 200 - Bare areas
- 210 - Waterbodies

Labelling based on visual interpretation

UCL classification chain

GLOBCOVER Europe product

15 LC classes of Level 1

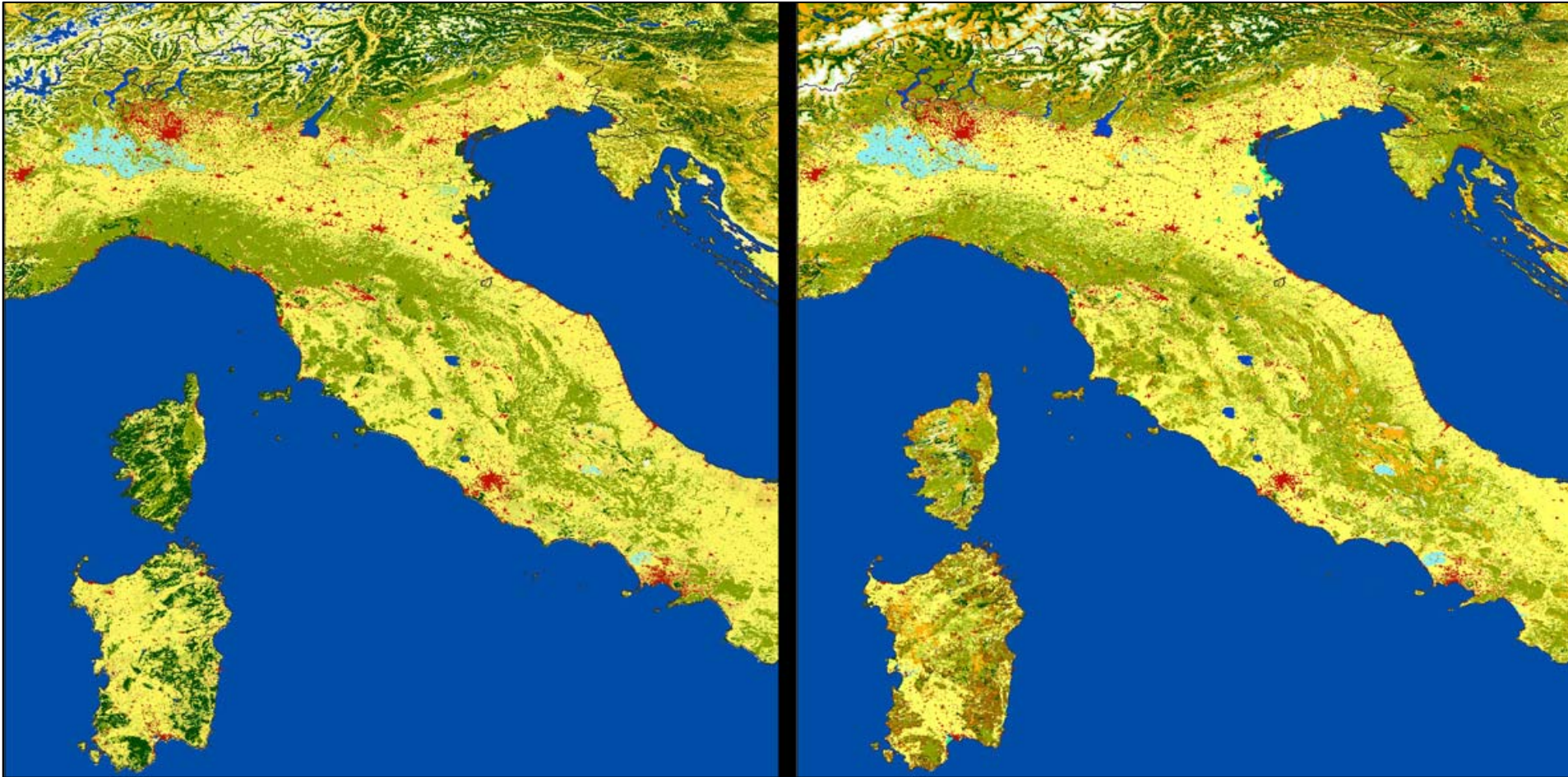
(including 3 supervised classes)

10 LC classes of Level 2

(including 1 supervised class)

Europe classification results

Italia



Globcover 2005-2006 (300 m)

Corine LC 2000 (300 m)

First End Users Meeting Ispra,
June 20, 2007

First classification results

Denmark



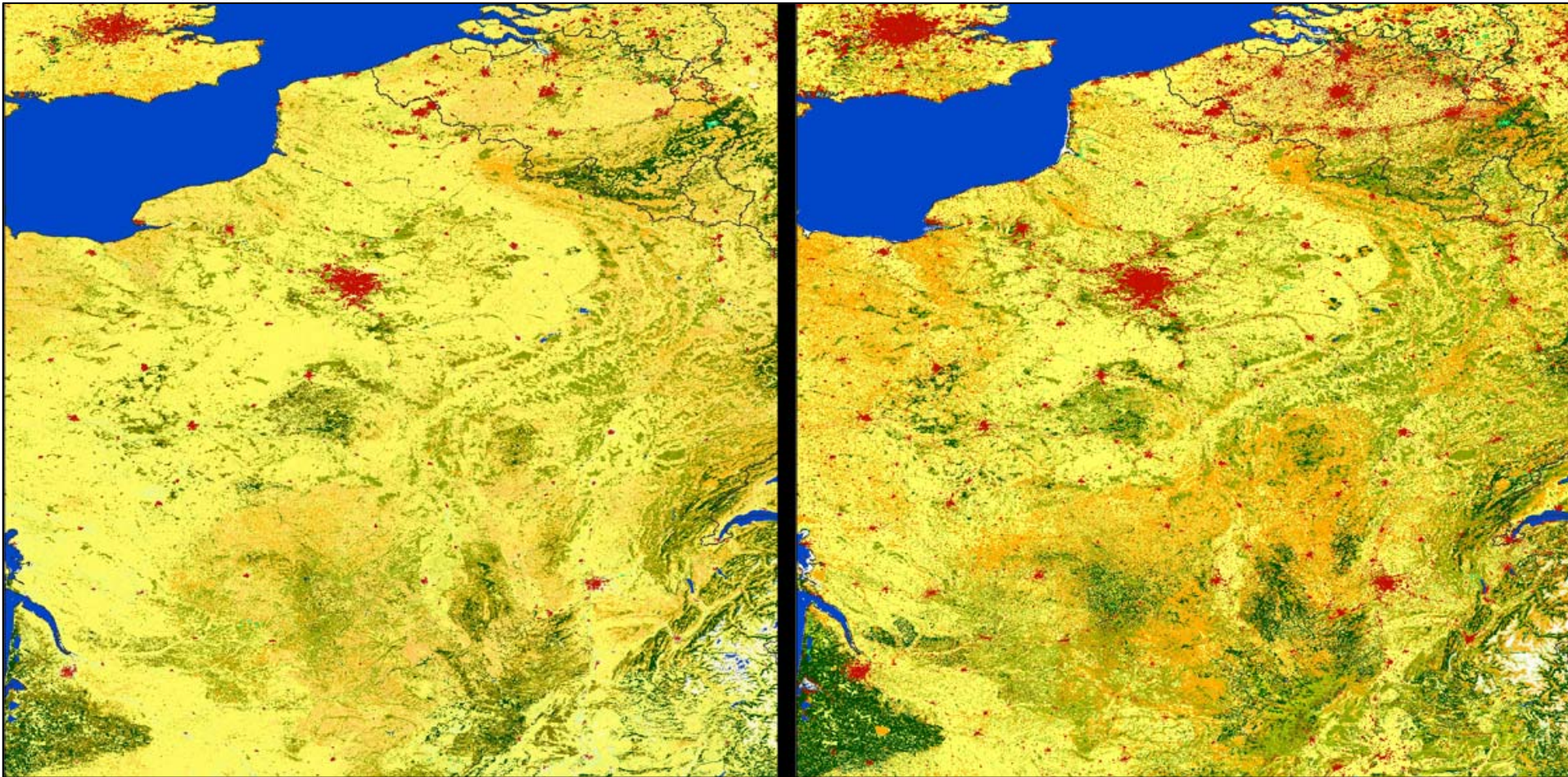
Globover 2005-2006 (300 m)



Corine LC 2000 (300 m)

Europe classification results

France - Belgium

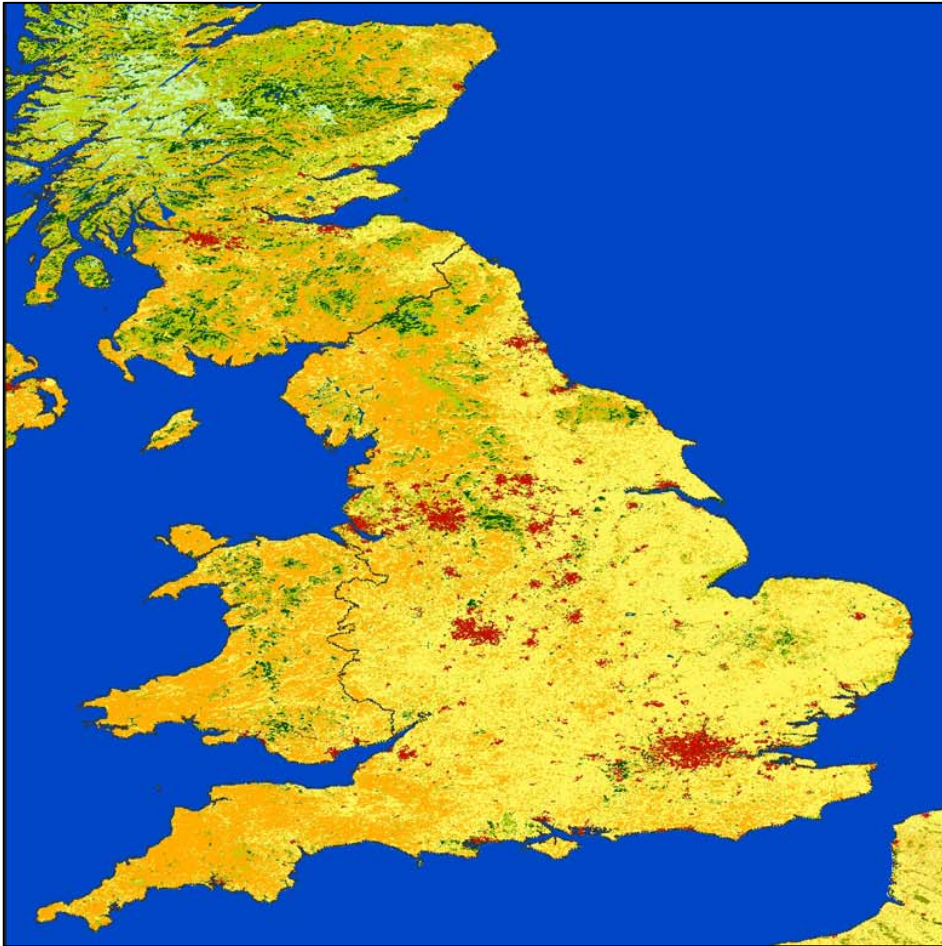


Globcover 2005-2006 (300 m)

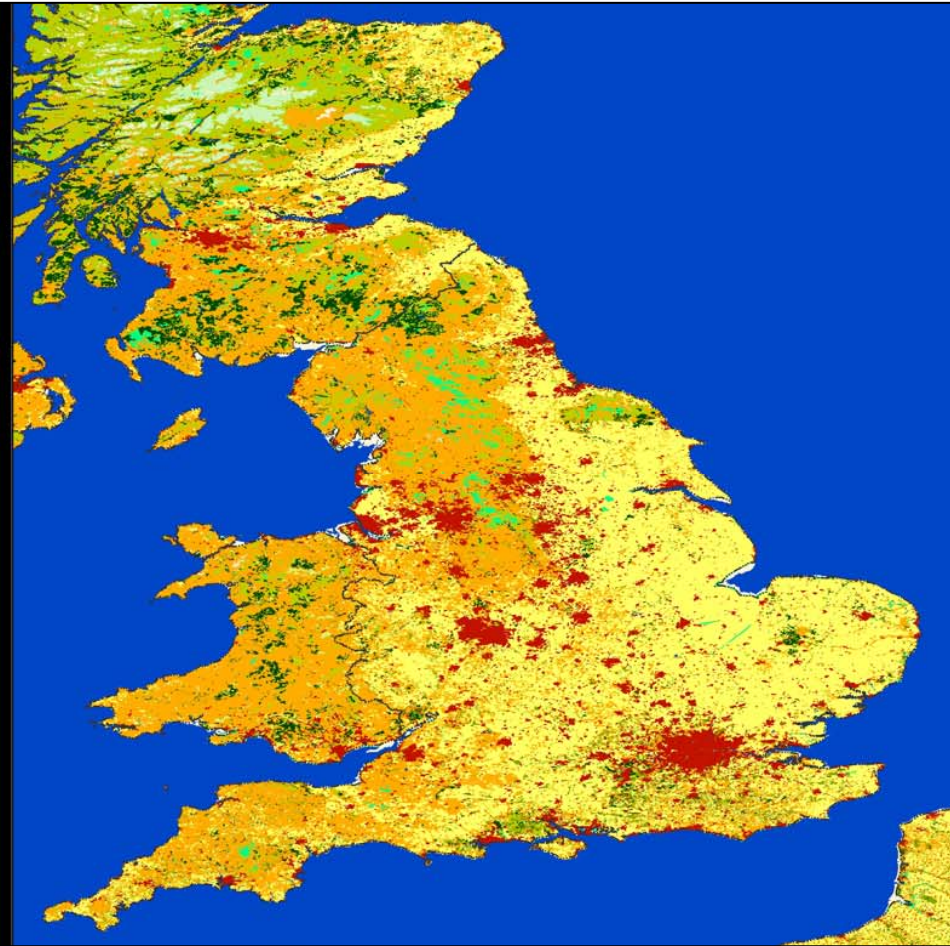
Corine LC 2000 (300 m)

First classification results

UK



Glocover 2005-2006 (300 m)



Corine LC 2000 (300 m)

First classification results

Spain



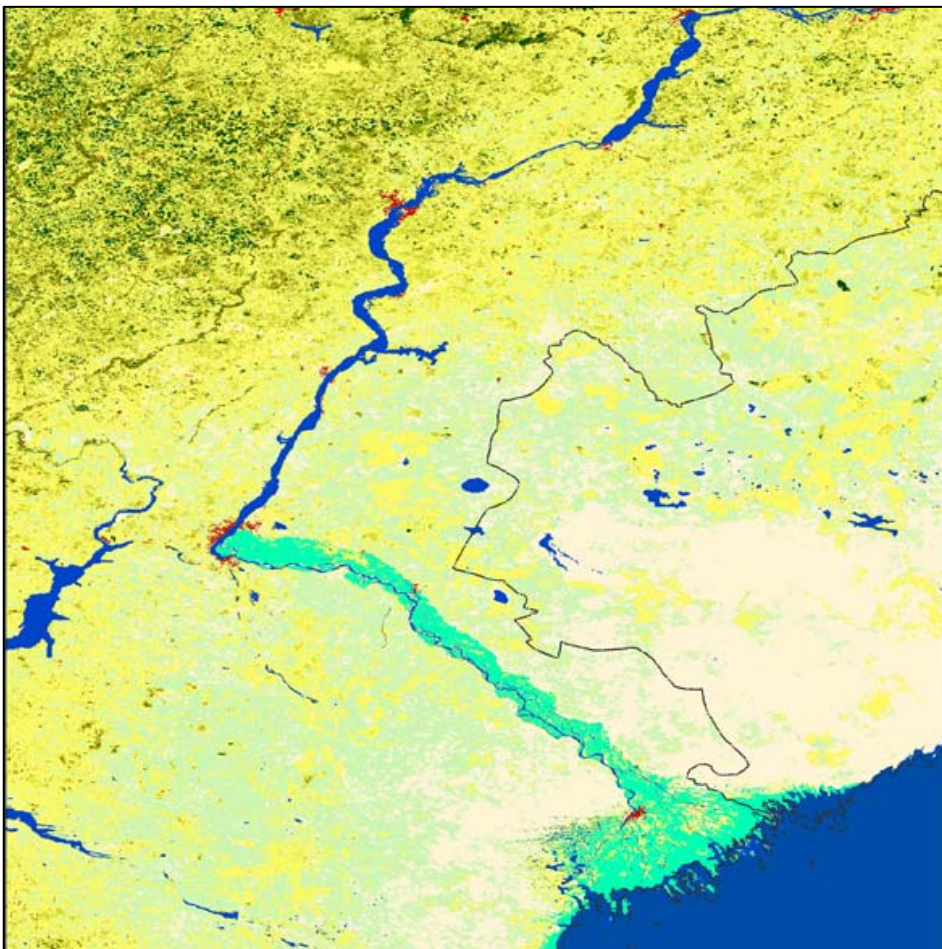
Globcover 2005-2006 (300 m)

Corine LC 2000 (300 m)

First End Users Meeting Ispra,
June 20, 2007

First classification results

Russia - Kazakhstan



Globcover 2005-2006 (300 m)

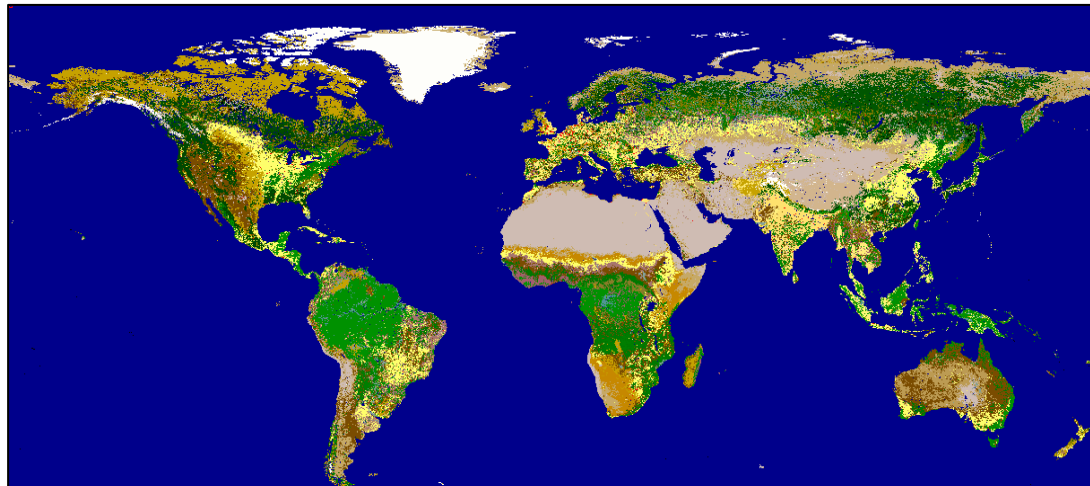


GLC2000 (1km)

Labelling procedure based on best previous products

Reference dataset based on existing LC data:

- GLC2000
- Corine LC 2000
- FAO- Africover product (2000)
- US National LC database (USGS) 2005
- National LC map of China (Chinese Academy of Sciences)
- RDCongo LC map (UCL-Geomatics, 2000)
- Australia Land use map (Australian Government)



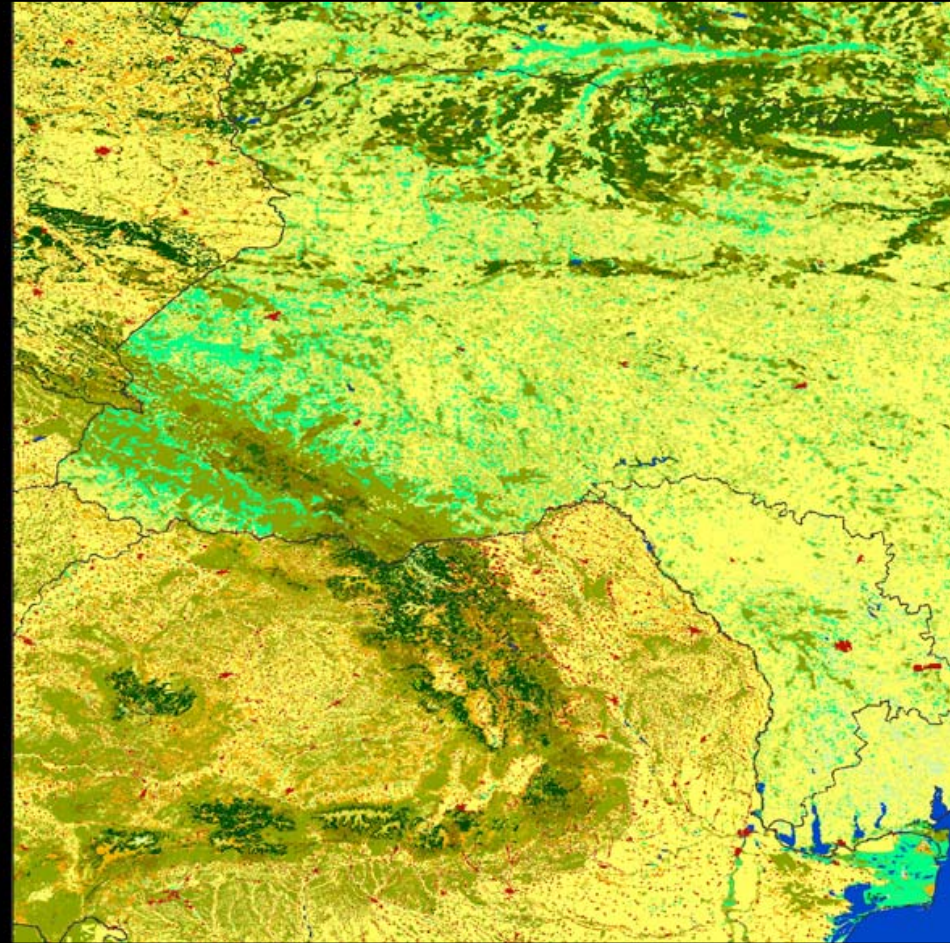
First End Users Meeting Ispra,
June 20, 2007

But best exiting is not always relevant to label GLOBCOVER product

Romania



Globcov 2005-2006 (300 m)



Corine LC & GLC2000 (300 m)

Validation of Globcover product

Its concepts and its methods

Leon Schouten

Contents

- Objectives
- Concept
 - Statistics
 - LCCS
 - Assumptions (time, productivity)
- Tools
- Data sets
- Feedback from experts (Alterra)

Objectives

- The main objective of validation is to allow a potential user to determine the map's "fitness for use" for his or her application.
- Validation - The process of assessing, by independent means, the quality of the data products derived from the system outputs (WGCV).
- Here we look at the thematic information only!

Concepts

- For obvious reasons no field surveys were foreseen to conduct the validation
- In stead we decided to engage “regional experts” in the validation.
- This implies that we accept a certain level of subjectivity.
- Two sets of validation points are required.

Concepts -statistics

● Binomial distribution, according to equation 1:

$$n = \frac{pq}{\left(\frac{E}{z_{\alpha}}\right)^2}$$

- N = the original sample size,
- E = the allowable error; set at 0.05
- z_{α} = drawn from the normal curve for the given level of confidence,
- p = the required accuracy of the data, $q = (1-p)$; set at 0.70.

Concepts - statistics

- Contingency matrix will be made, but the main aim is to establish the validity of the assumed accuracy of 0.70
- Depending on E and z_{α} we will be able to accept this, even when the overall accuracy is less than 0.70.

Concepts - statistics

🌐 Sampling:

- 🌐 Stratified random sampling
- 🌐 Minimum of 5 points per class
- 🌐 Sampling on a Globcover product that is in an equal area projection (Lambert azimuthal equal area)

Concepts - LCCS

- The Globcover legend is based on the LCC System
- This legend is based on several land cover classifiers and one so-called environmental attribute.
- Rather than validating the legend we seek to validate these classifiers and the attribute. (This approach is an adjustment on the GVP!)

Concepts - assumptions

- Number of samples need to be known → what can we expect from an expert?
 - 10 minutes per point
 - 7 hours per day
 - 3 working days
- This would amount to 126 points and this can be used as input in the equation 1.

Tools

- **Two dedicated tools are made/foreseen:**
 1. A dedicated environment to fill in the data (in MS Access)
 2. Tailor made information that allows quick panning to the concerned validation point

Main menu of
the (current)
Access
environment:

Sample points

ID:	1	clas2:	O
x:		clas2-1:	O3
y:		clas2-2:	o-1-2
person_id:	Leon Schouten	clas2-3:	o-2-2
timestamp:	08/06/2007 18:55:24	clas2-4:	o-3-2
source_id:	landsat	clas2-5:	o-4-1
level_certainty_id:	certain	clas3:	O
clas1:	N	clas3-1:	O2
clas1-1:	N1	clas3-2:	
clas1-2:		clas3-3:	
clas1-3:		clas3-4:	
clas1-4:		clas3-5:	
clas1-5:			

Class 1 Class 2 Class3

Tools

● Filling in the form:

- Up to three land cover types per point.
- One land cover type is always the most dominant one.
- Based on an observational unit of $\pm 5 \times 5$ pixels.

Tools

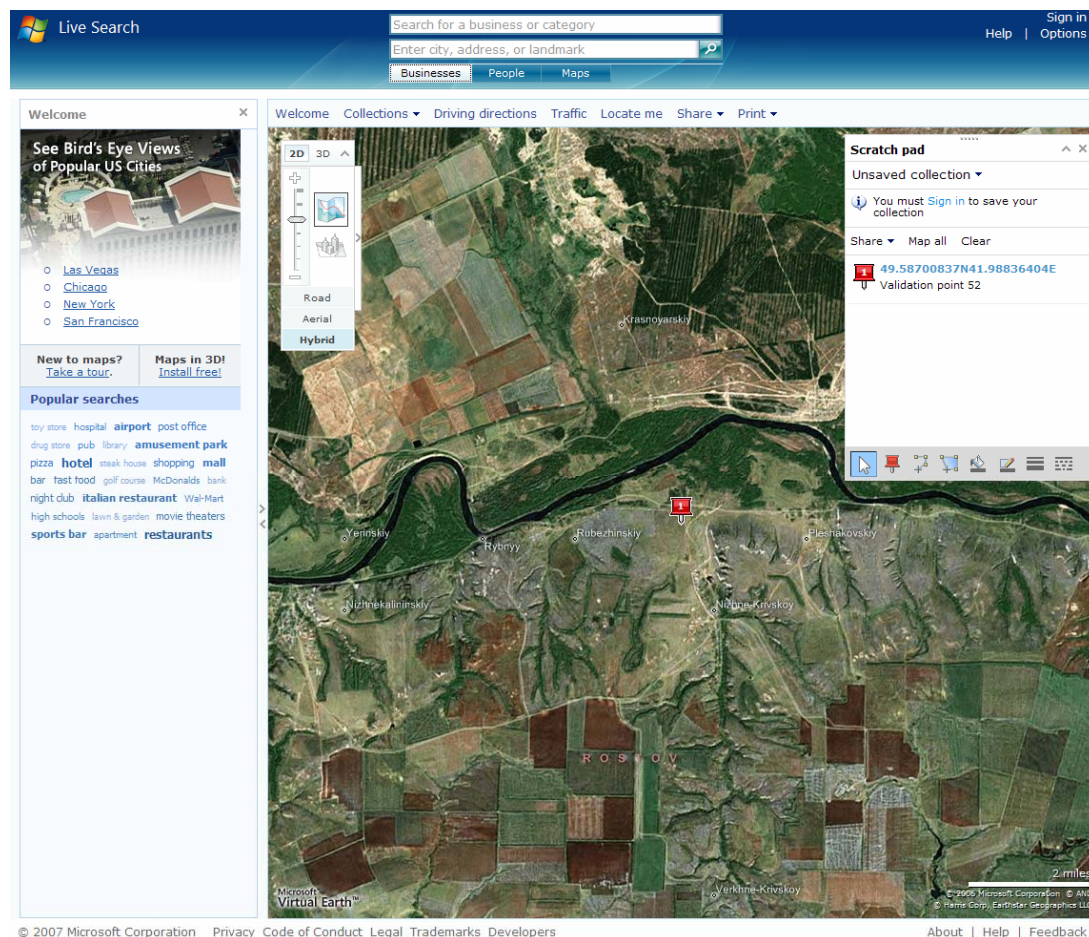
● Second tool:

A clickable url that will open Virtual Earth.

Example:

http://maps.live.com/default.aspx?v=2&cp=49.58700837~41.98836404&style=h&lvl=12&sp=an.49.58700837_41.98836404_49.58700837N41.98836404E_Validation%20point%2052

Tools



The screenshot shows the Microsoft Live Search Maps interface. At the top, there's a search bar with the text "Search for a business or category" and "Enter city, address, or landmark." Below the search bar are tabs for "Businesses", "People", and "Maps". The main map area displays a satellite view of a rural landscape with a river and fields. A red pin is placed on the map, and a "Scratch pad" window is open on the right, showing the coordinates "49.58700837N 41.98836404E" and "Validation point S2". The left sidebar contains a "Welcome" message, a "See Bird's Eye Views of Popular US Cities" section with links to Las Vegas, Chicago, New York, and San Francisco, and a "Popular searches" section with various categories like airport, post office, amusement park, etc. The bottom of the page includes copyright information for Microsoft Corporation and links for About, Help, and Feedback.

Feedback

- Integrate the two tools
- Give clear definitions on LCCS classifiers (hardpans?)
- Avoid errors due to misinterpretations rather than misclassifications
- Suggestions to improve the user interface of the Access form

Feedback

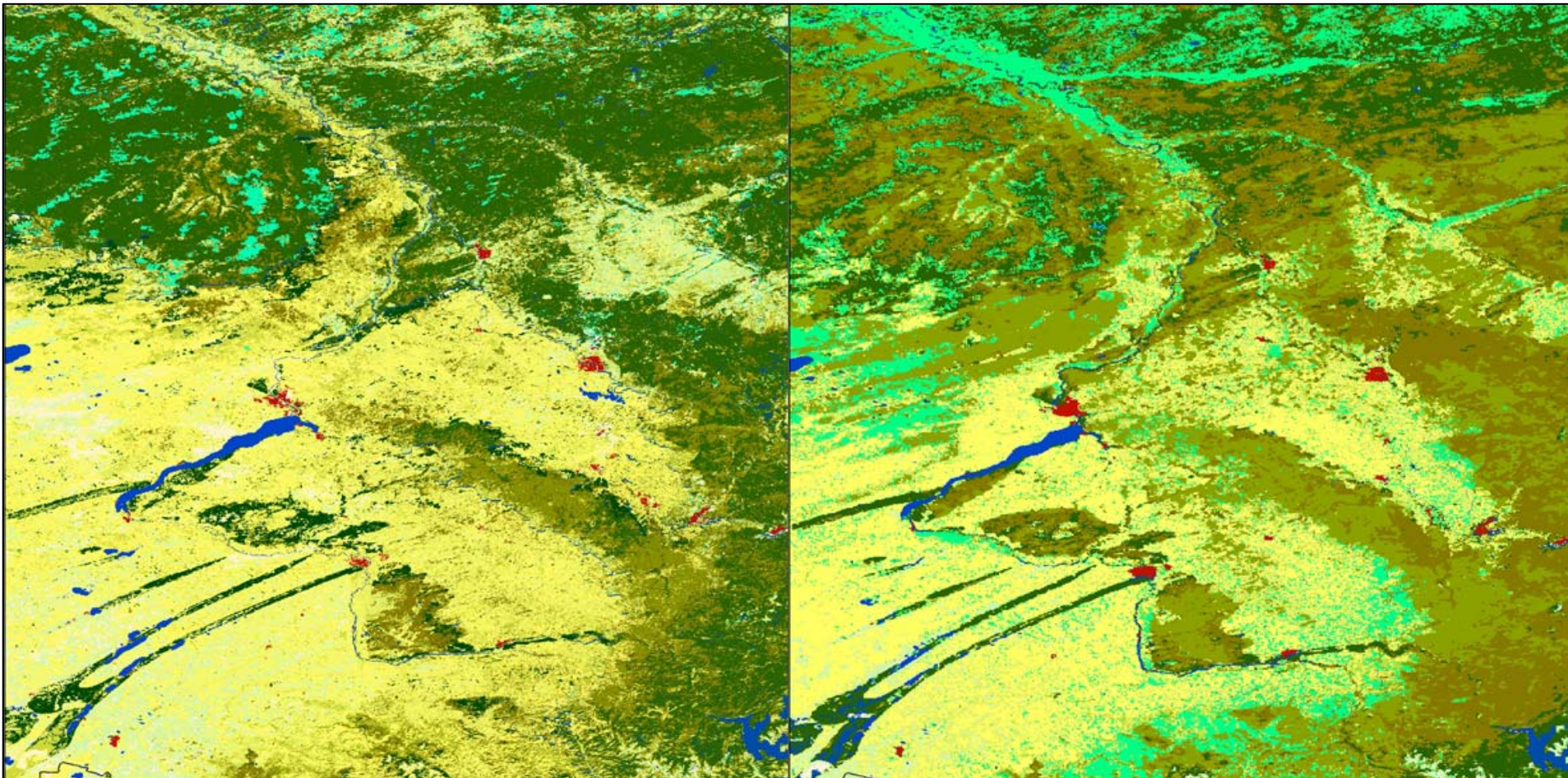
- Experiences up to now by Alterra staff: take 5 minutes per hour, but more breaks
- 7 hours per day may be too optimistic
- Assumption: $6.5 \text{ hours} * 12 \text{ (points/hour)} * 3 \text{ days} = 234$
- But no Meris data layers available in Virtual Earth, yet.

Future plans

- ① Adjust the tools conform the suggestions made by Alterra
- ① Conduct the validation for Europe
- ① Invite the experts in the autumn of 2007
- ① Assess the consequences for the Globcover product

First classification results

Russia



Globcover 2005-2006 (300 m)

Corine LC 2000 (300 m)

First End Users Meeting Ispra,
June 20, 2007

First classification results

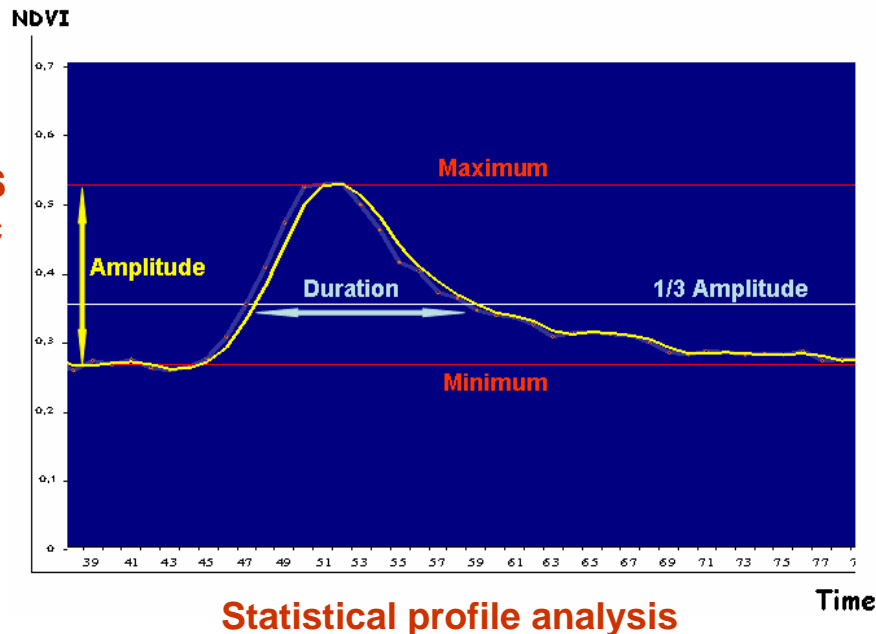


Globcover 2005-2006 (300 m)

Corine LC 2000 (300 m)

Per-object classification based on spectro-temporal characterization

- Robust temporal metrics computed per-object from bi-weekly multispectral composite and associated indices



Object attributes :

Min., Max., Amplitude,
Multispectral reflectance

Spectro-temporal
classes