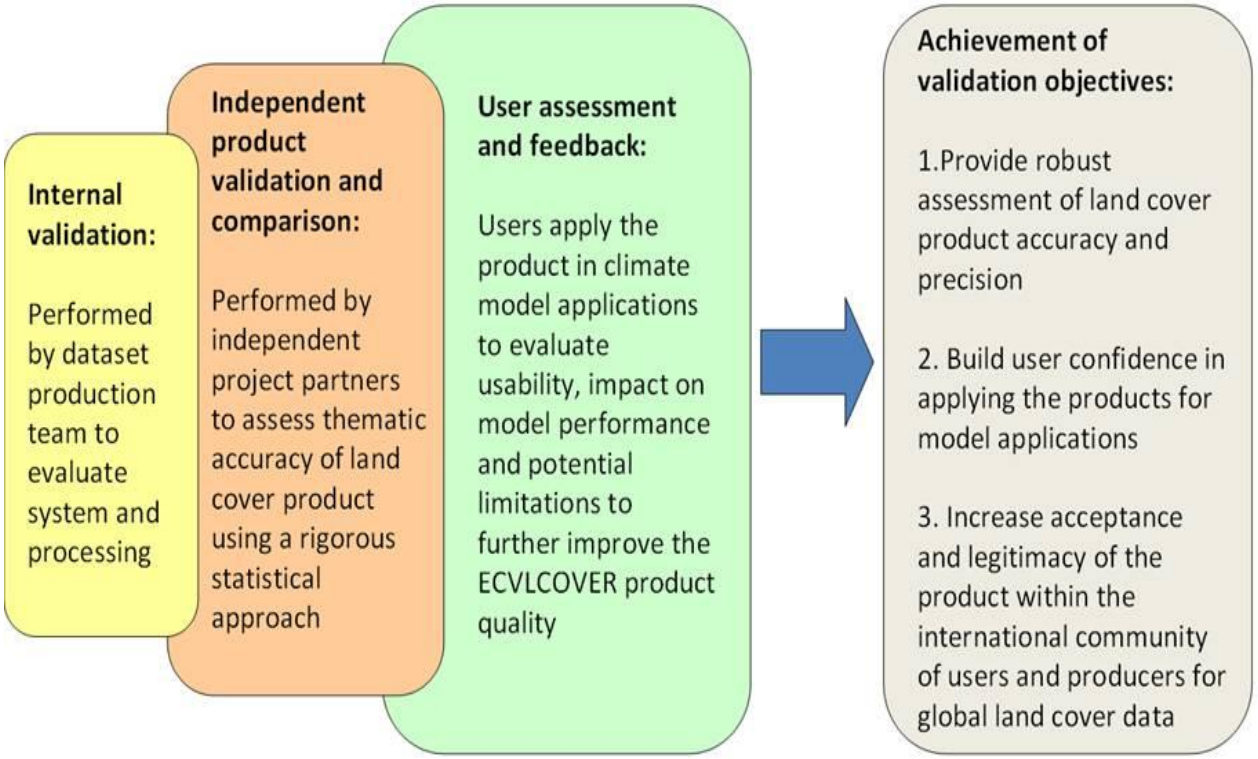
Global land cover maps validation: current strategy for CCI Land Cover

Defourny P. (1), Brockmann C. (3), Bontemps S. (1), Achard F. (2), Boettcher M. (3), De Maet T. (1), Gamba P. (4), Hagemann S. (5), Hartley A. (6) , Hoffman L., (12), Georgievski G. (5), Herold M. (10), Kirches G. (3), Lamarche C. (1), Lisini G. (4), MacBean N. (7), Moreau I. (1), Pathe C. (8), Peylin P. (7), Radoux J. (1), Ramoino F. (11), Riedel T. (8), Santoro M. (9), Schmullius C. (8), Smollich S. (4), Van Bogaert E. (1), Vittek M.(1), Wegmüller U. (9), Zuehlke M. (3), Arino O. (11)

(1) UCLouvain- Earth and Life Institute *Geomatics*, Belgium; (2) JRC, Italy; (3) Brockmann Consult GmbH, Germany;
(4) University of Pavia, Italy; (5) Max Plank Institute for Meteorology, Germany;
(6) MET Office, United Kingdom; (7) LSCE, France; (8) University of Jena, Germany;
(9) GAMMA Remote Sensing, Switzerland; (10) Wageningen Univ., (11) ESA, Italy, (12) LIST, Lux.

Validation standards

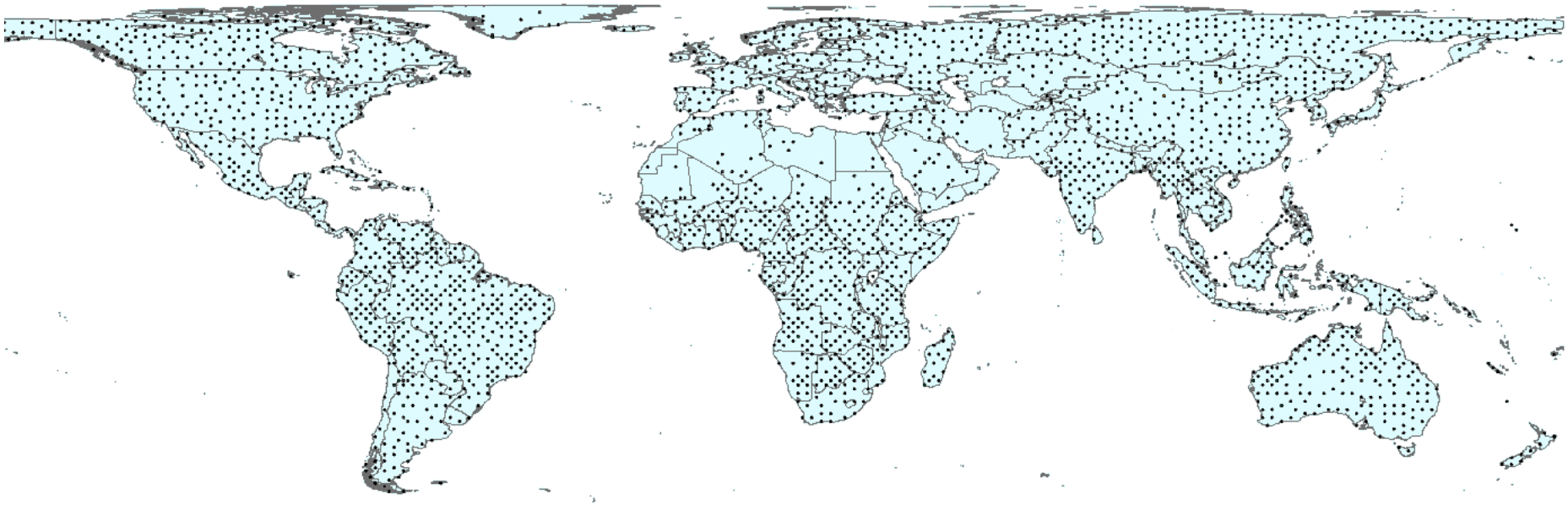
- Quantitative and independent validation

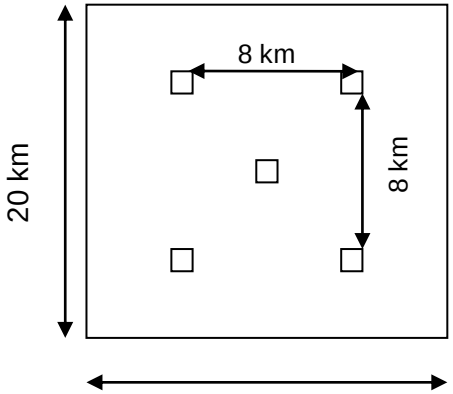


- Assessment by climate modellers => climate assessment report

Stratified random sampling design

2600 Primary Sampling Units (10*10 km)



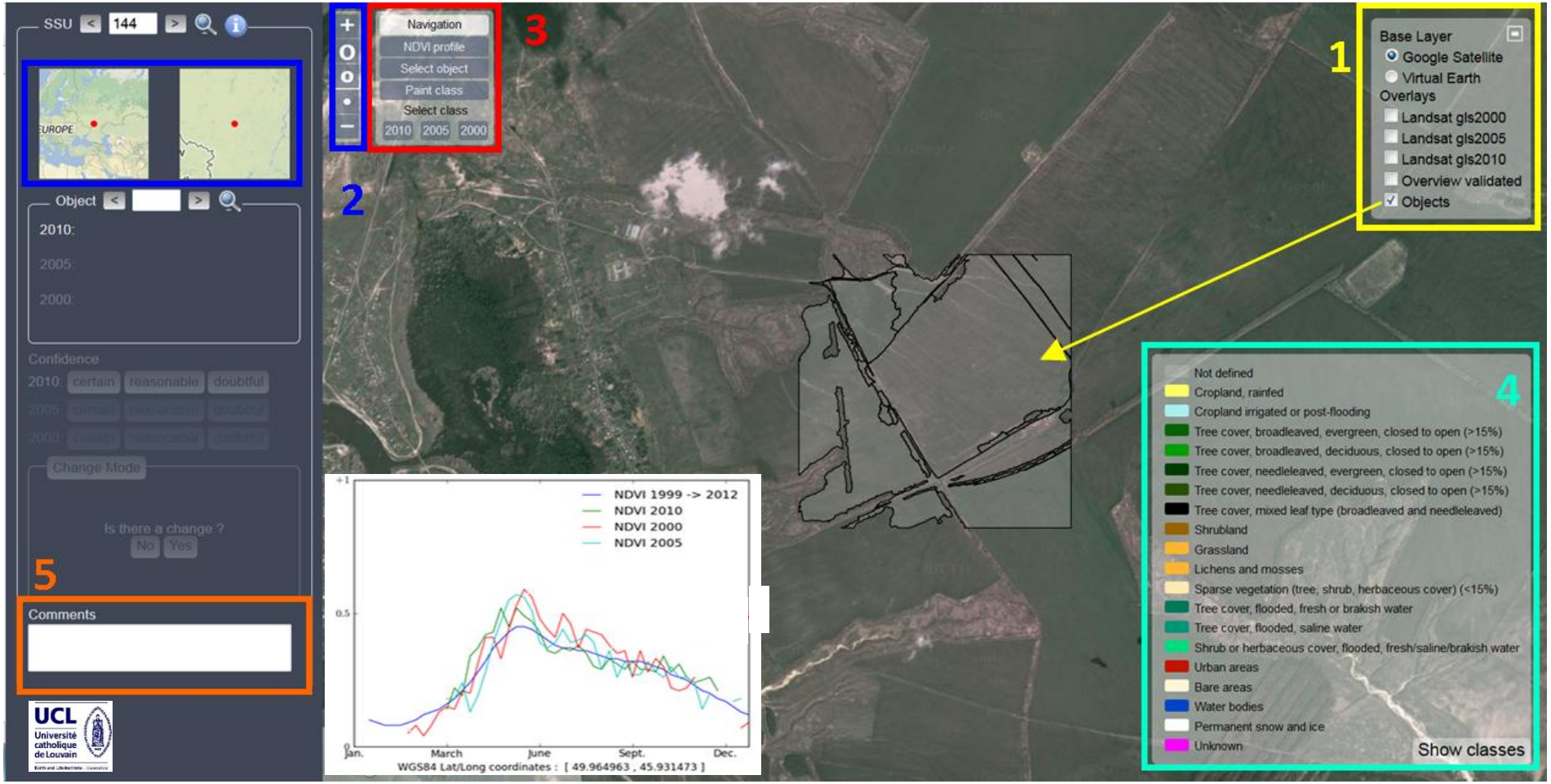


5 Secondary Sampling Units (SSU) per PSU

Size of each SSU = 900 m x 900 m

SSU = validation units to be interpreted
=> 13.000 potential units

CCI LC Validation : very efficient interface

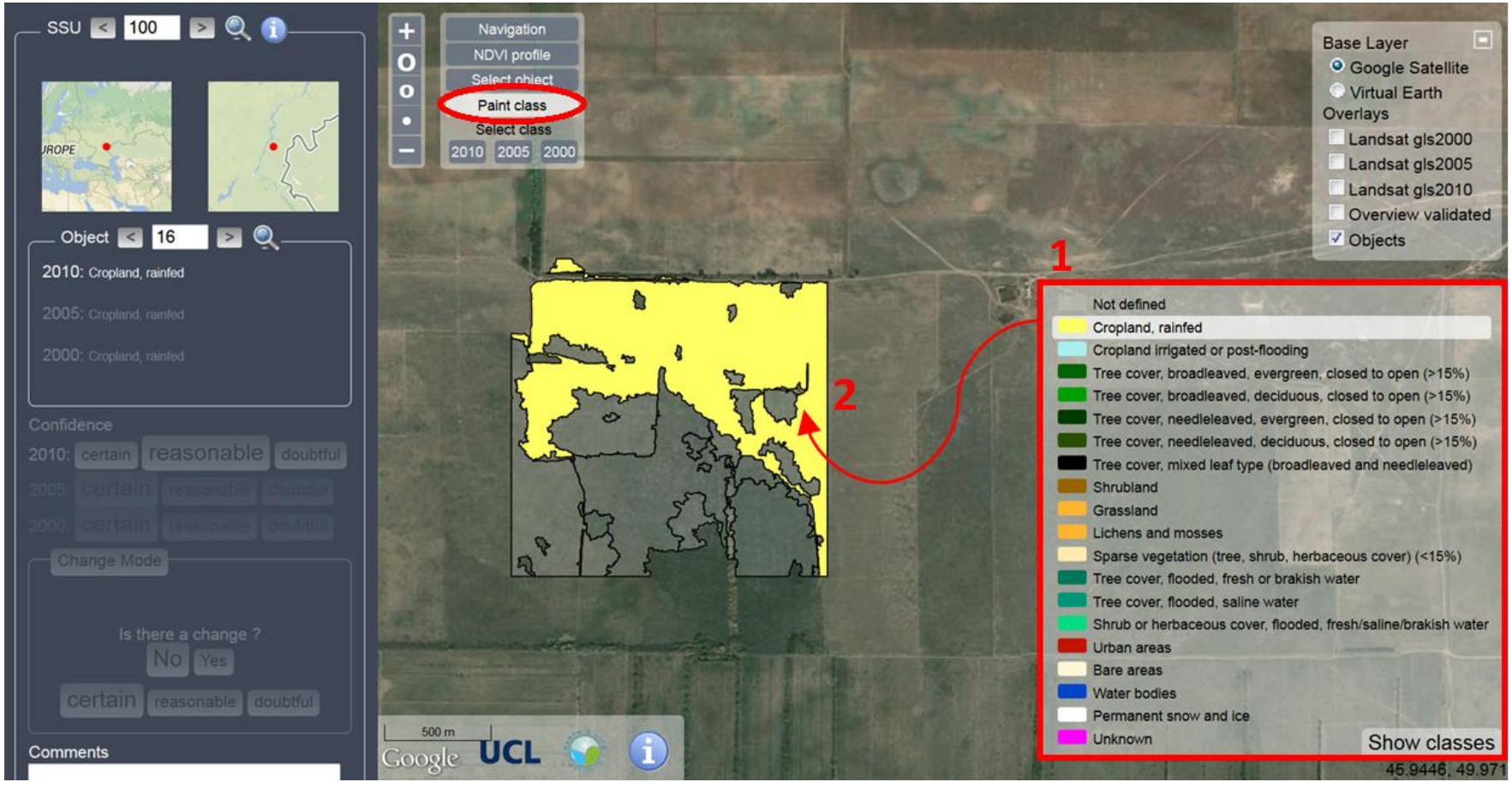


1. Layer box
2. Zooms
3. Tools (navigation, NDVI, select object, paint class)
4. Legend
5. Comment

SSU Interpretation

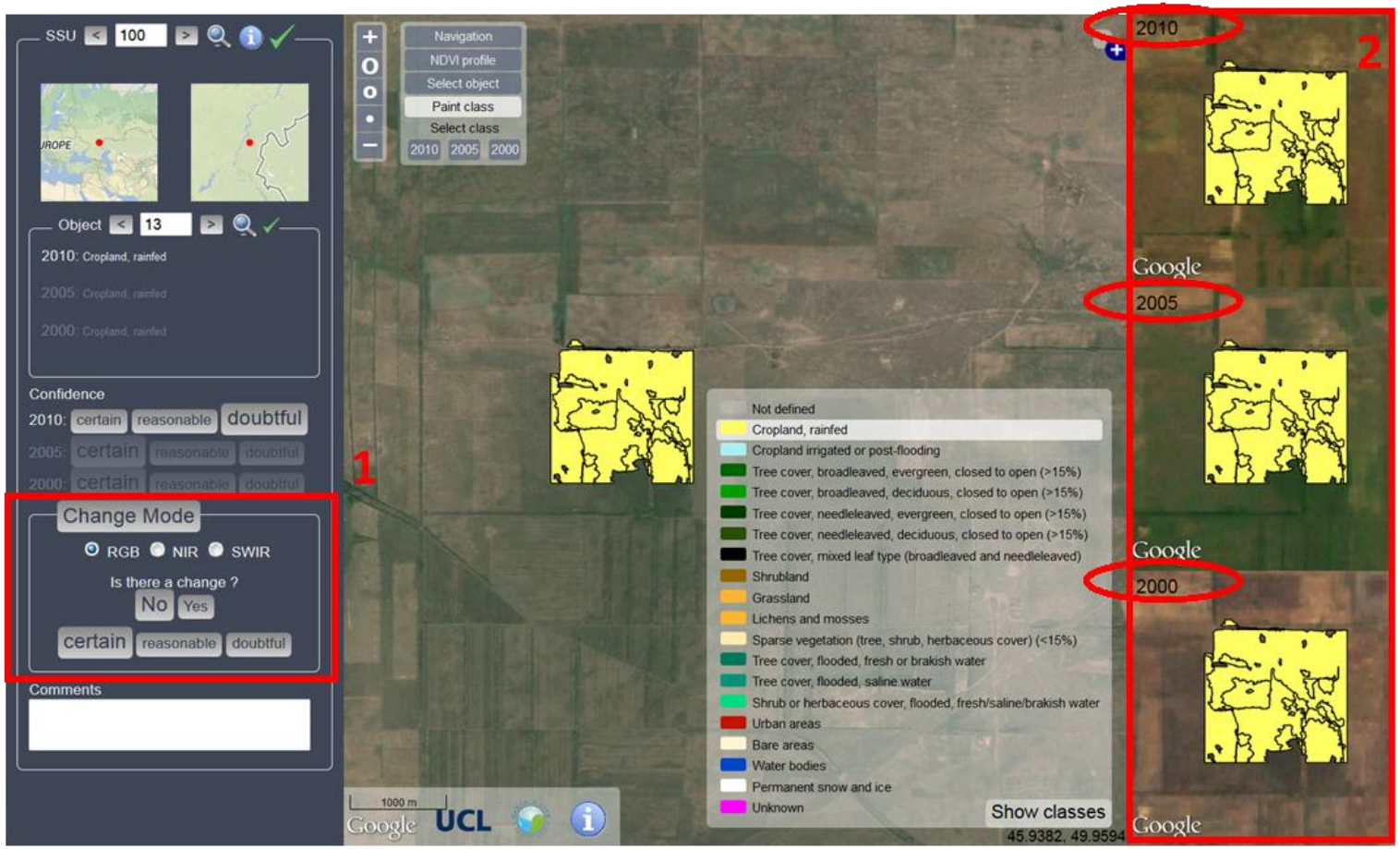
2. Assigning the 2010 Land Cover Classes

- *Using the "Paint class" button*



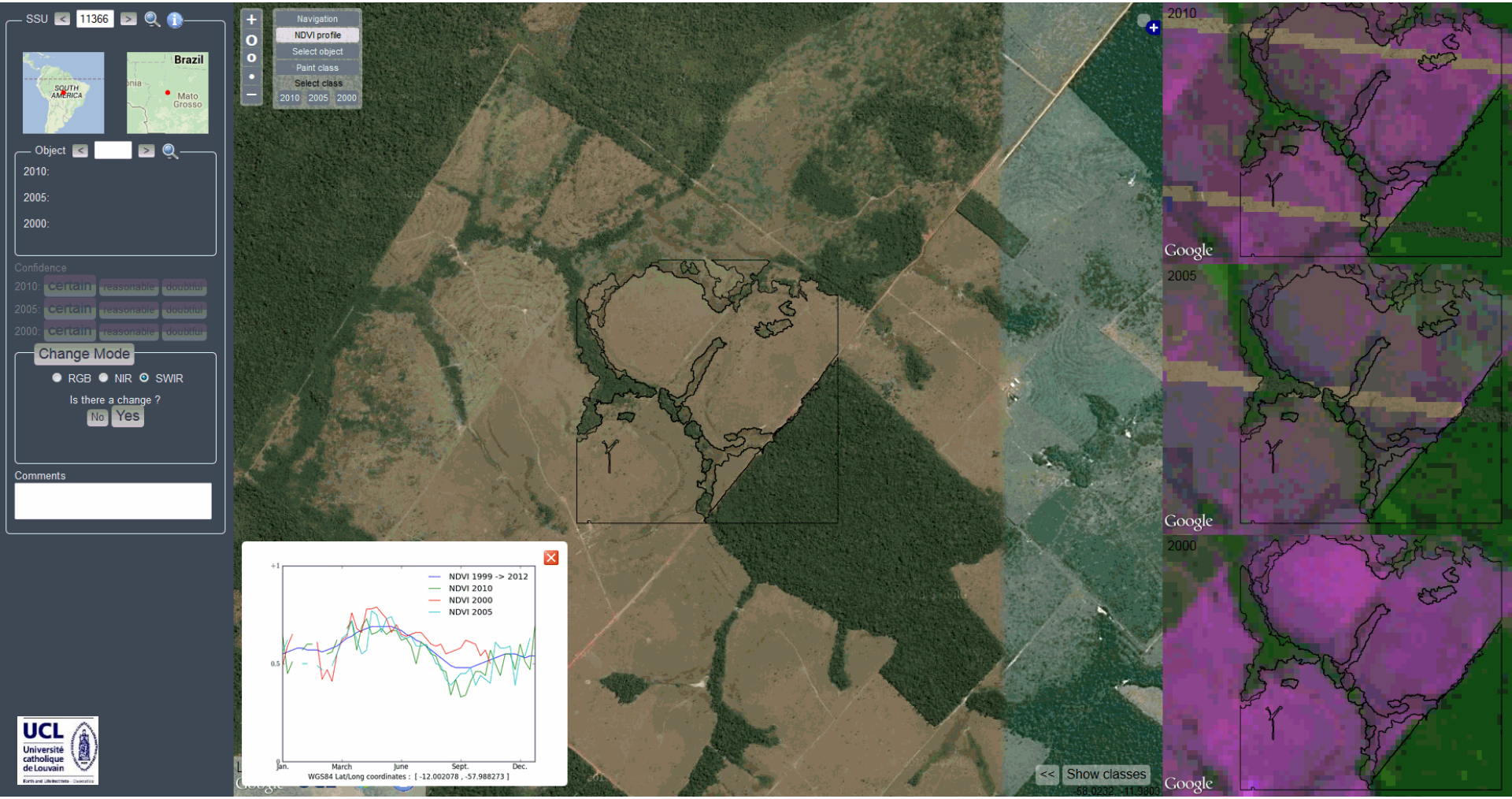
SSU Interpretation

3. Evaluating the change between the 3 epochs

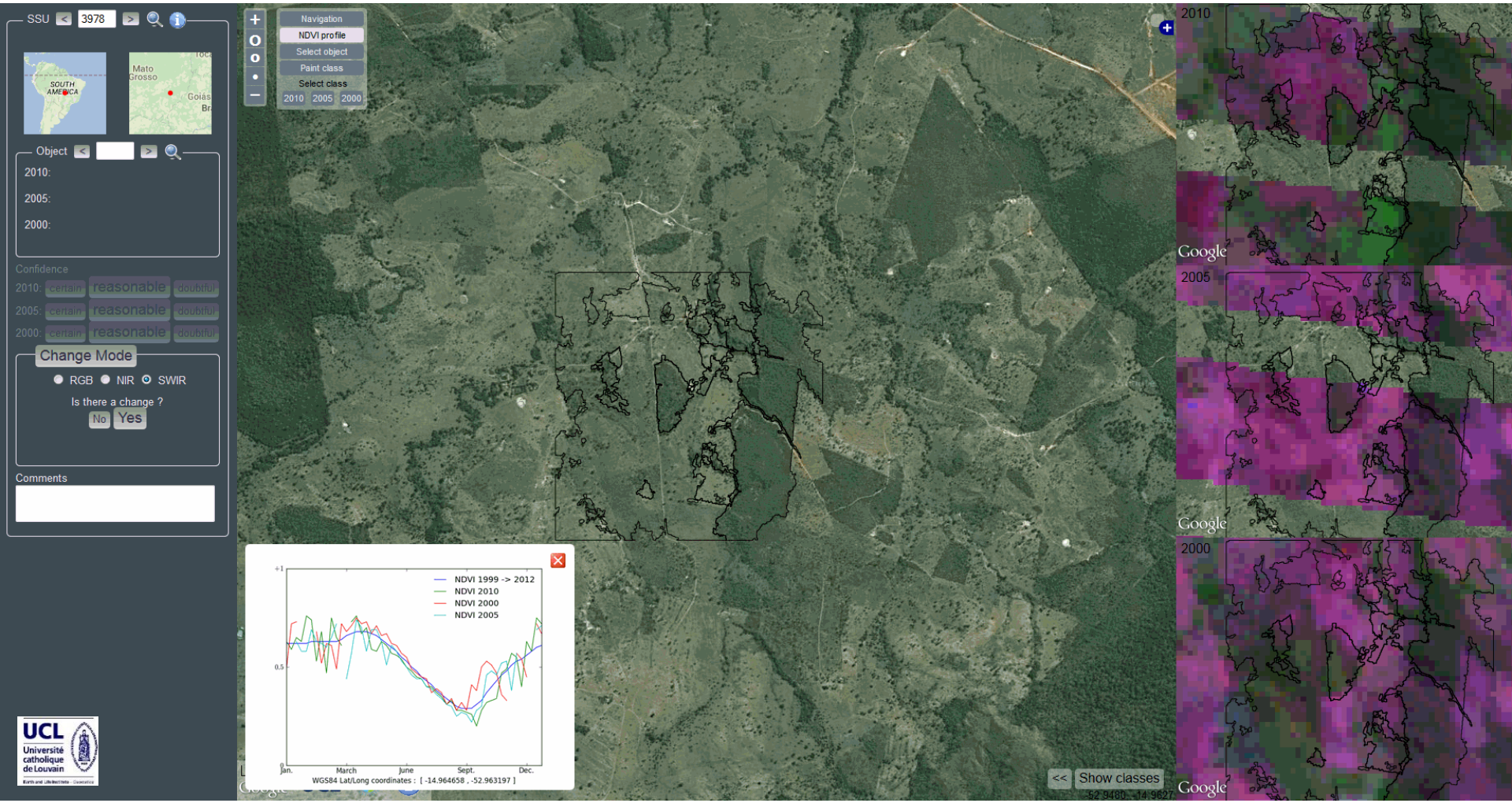


Example of segmented SSU over Brazil

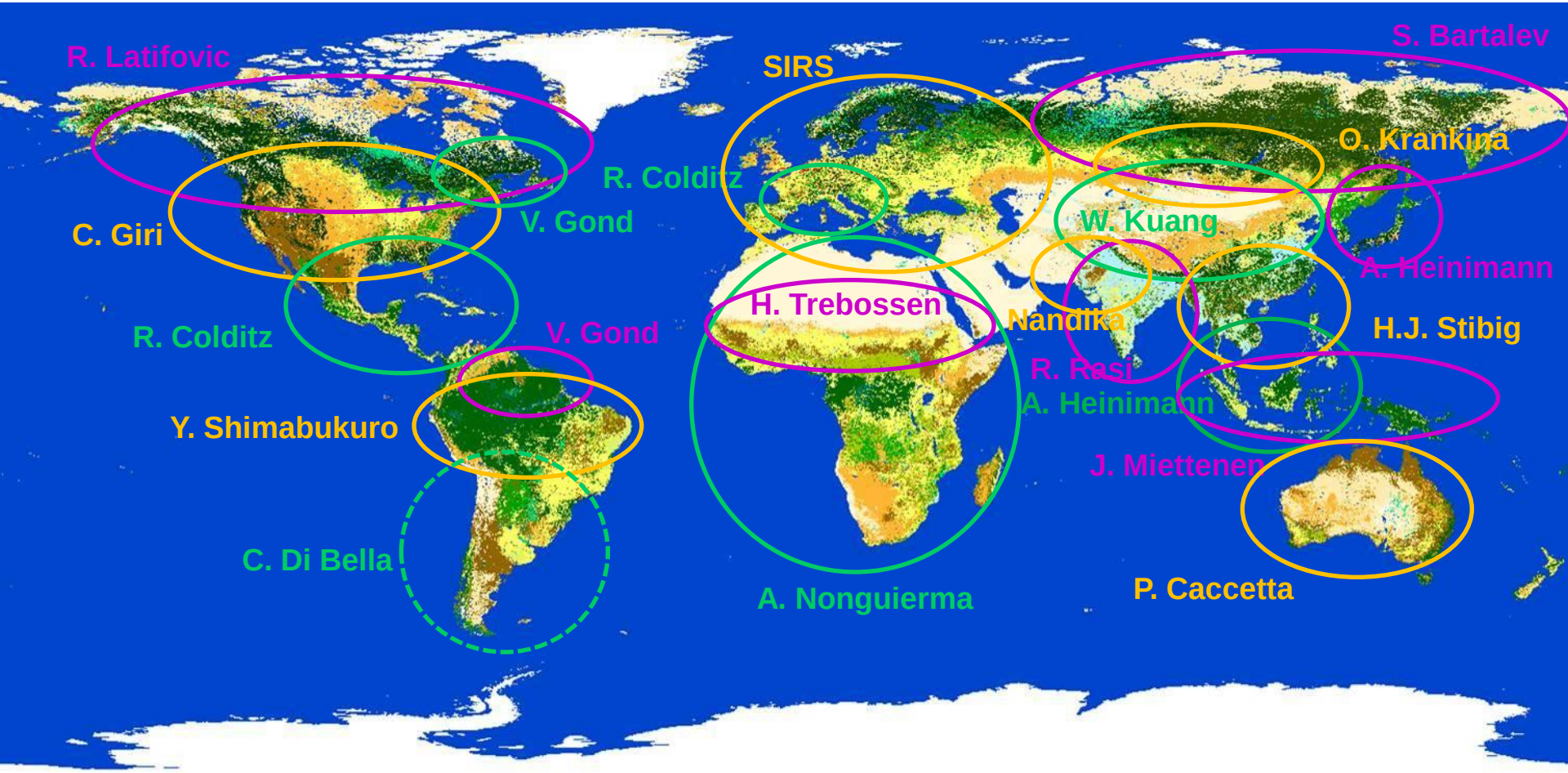
Good segmentation and availability of images



Example over Brazil:
Changes can not be assessed due to bad imagery



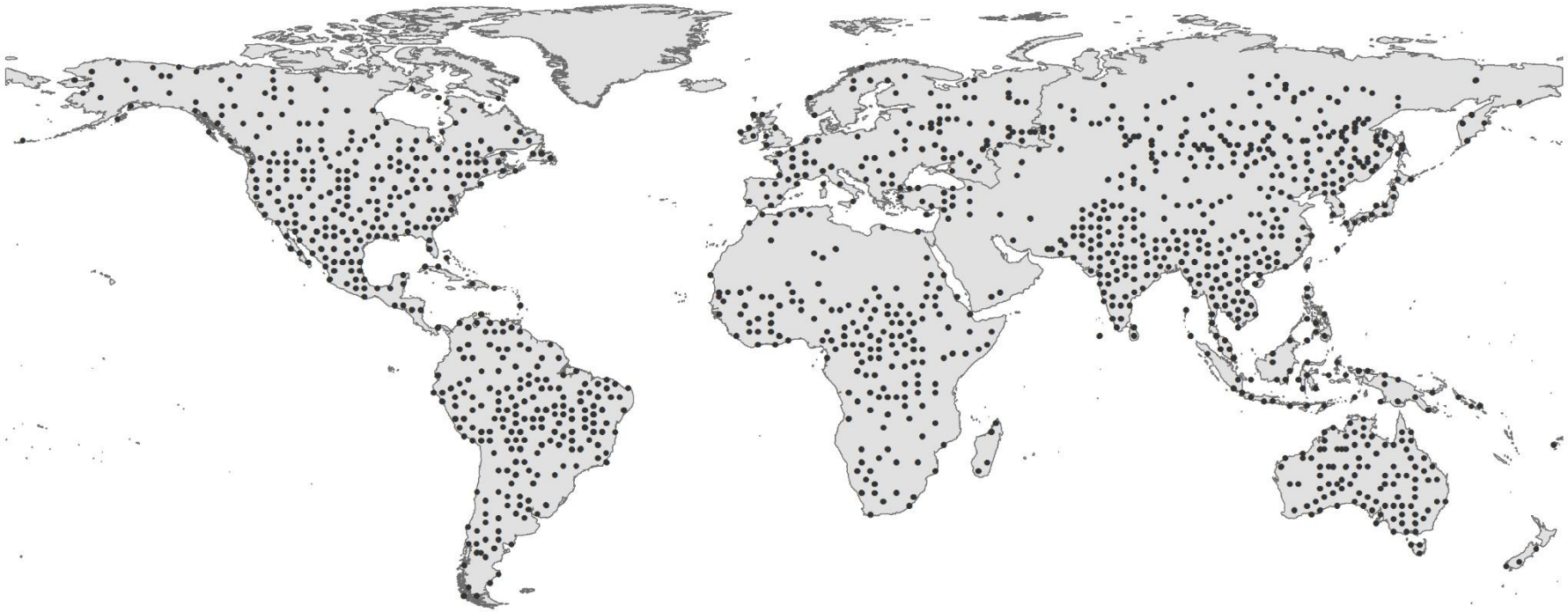
Network of experts



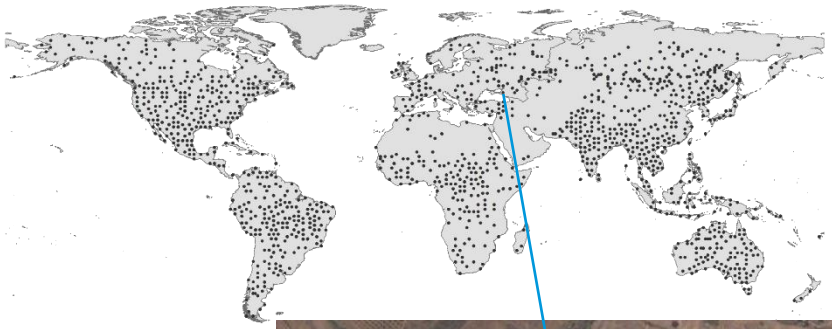
Database completed: 19 experts (150-250 points/expert)

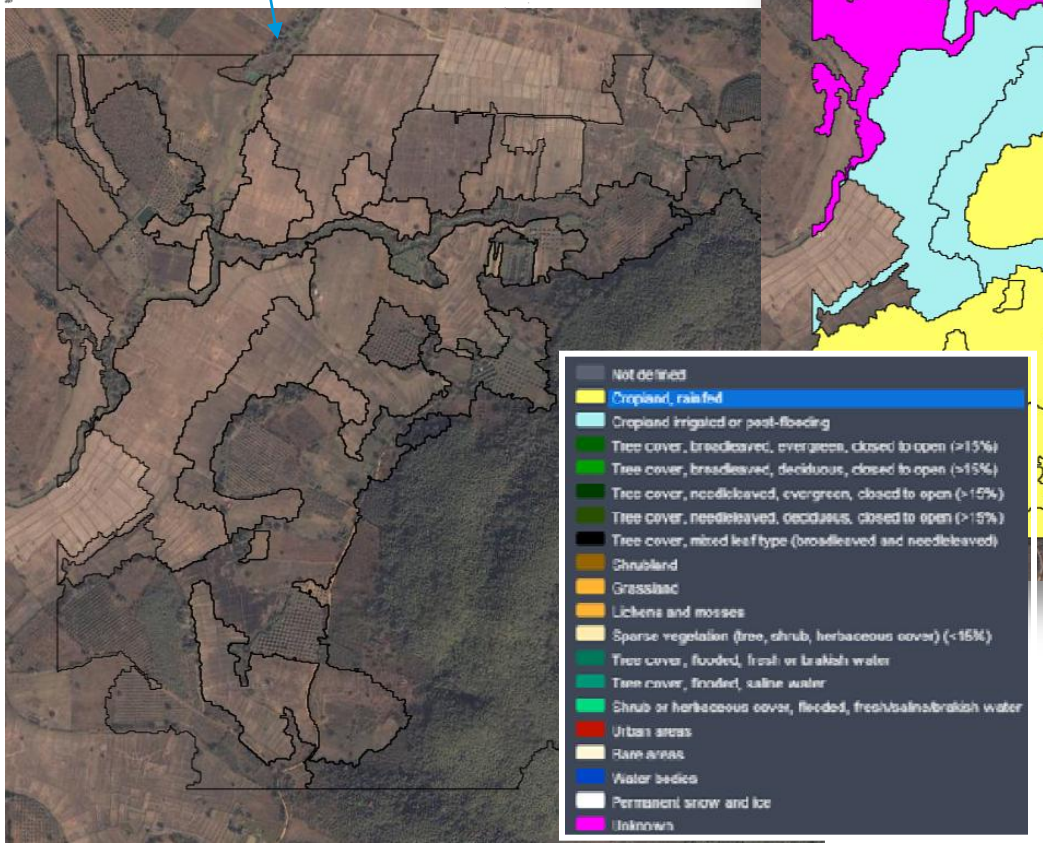
Completion of the multi-epoch CCI LC Validation database

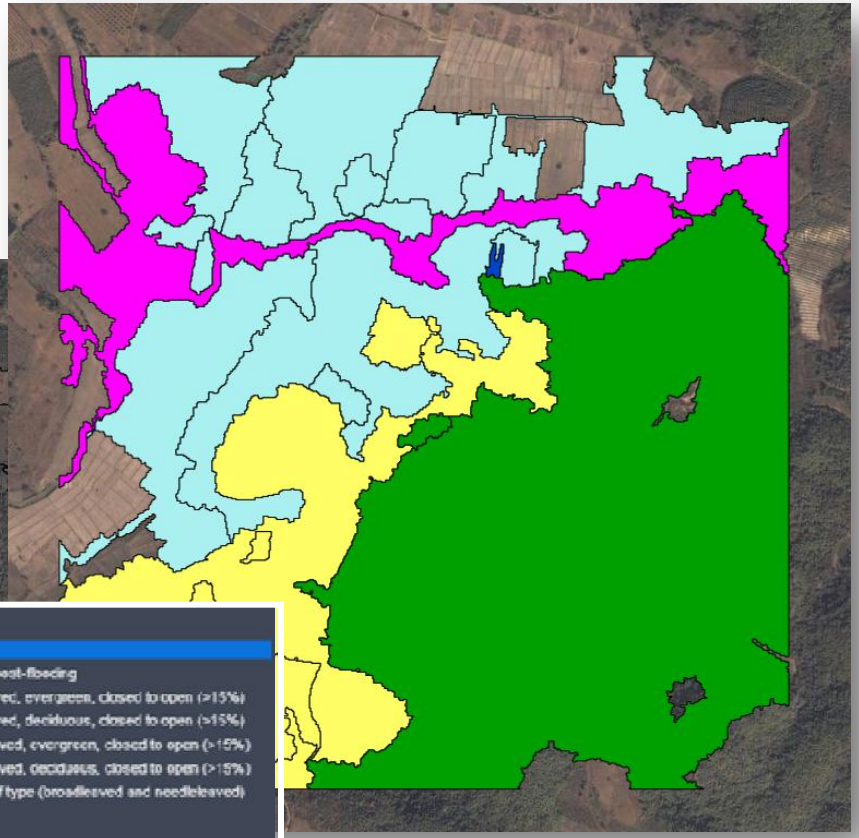
- 1450 PSUs proposed & 1352 PSUs interpreted (90%)
- Min 1 SSU by PSU and 2.4 SSU by PSU in average
- 2591 SSU of 900 x 900 m interpreted for each epoch (2000/2005/2010)



A 900 x 900 m validation sample (SSU) and its interpretation for 2010

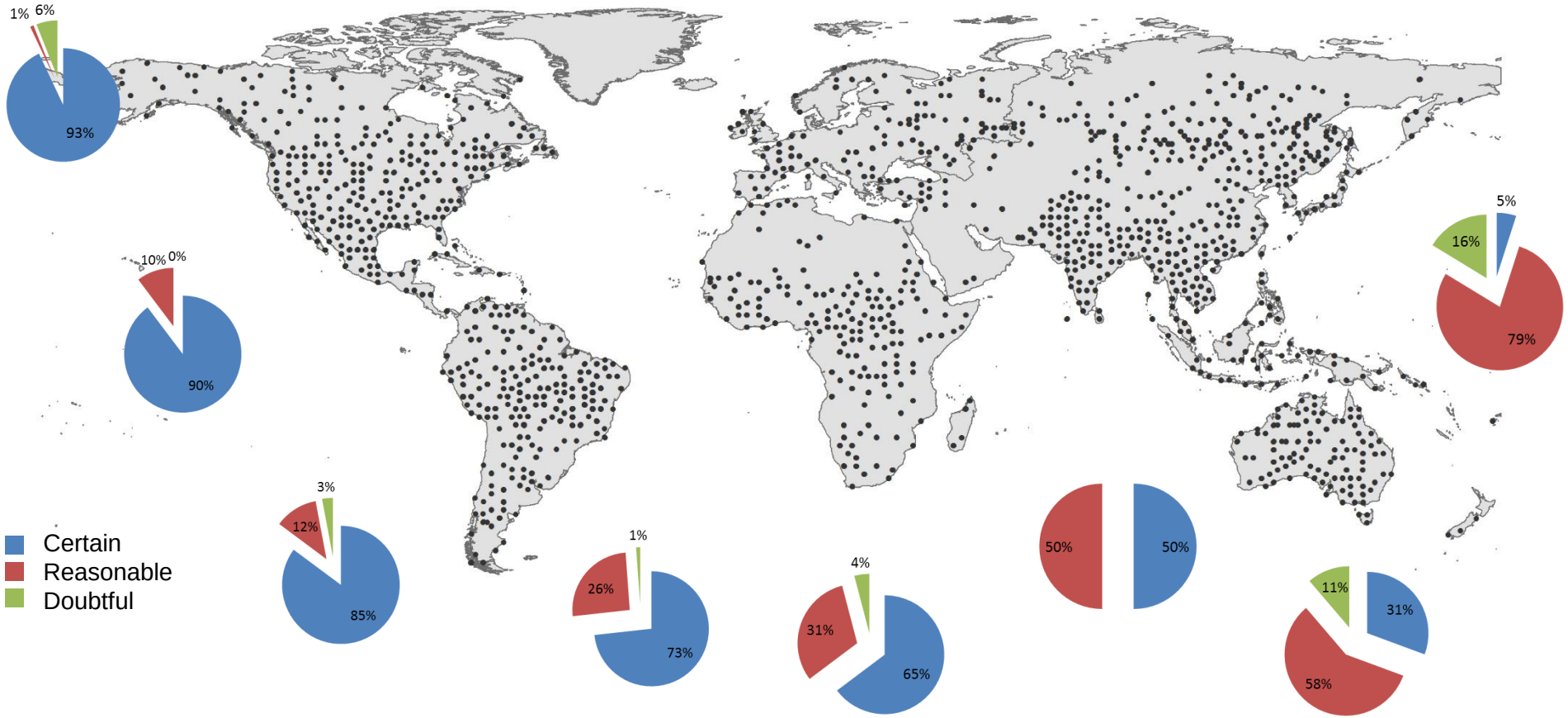






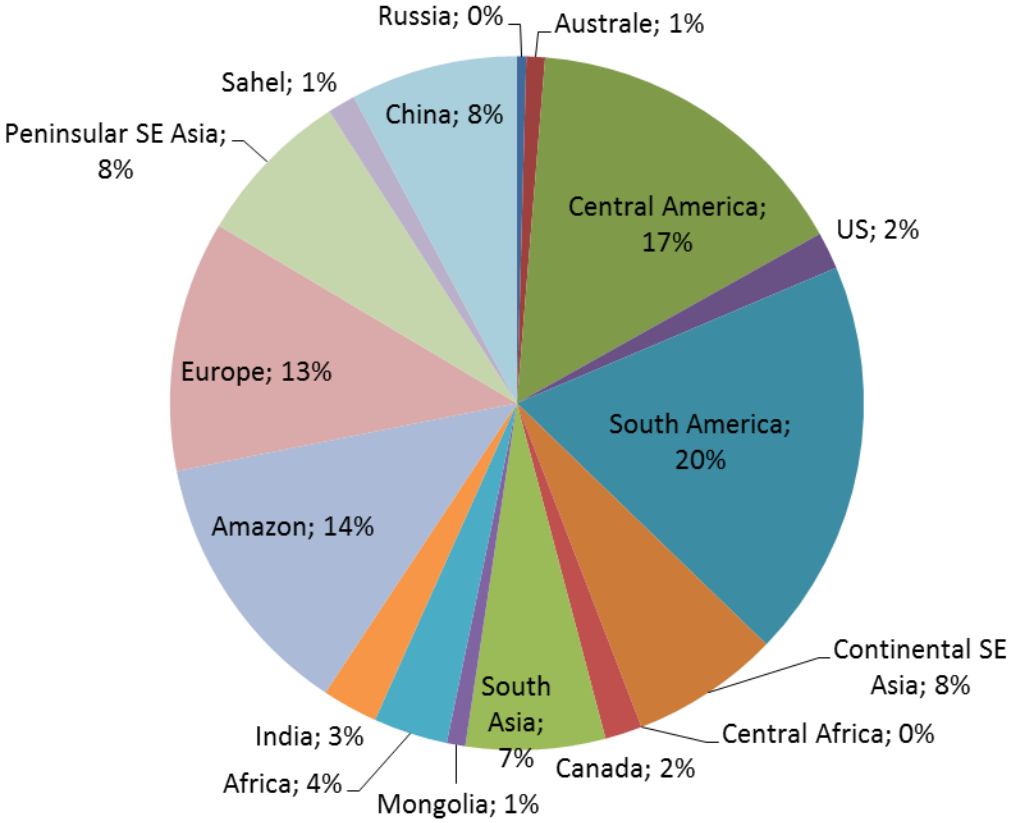
Level of certainty as assessed by the expert

- 65% of SSUs interpreted as certain; 31% as reasonable; 4% as doubtful
- Evaluating his subjectivity is subjective

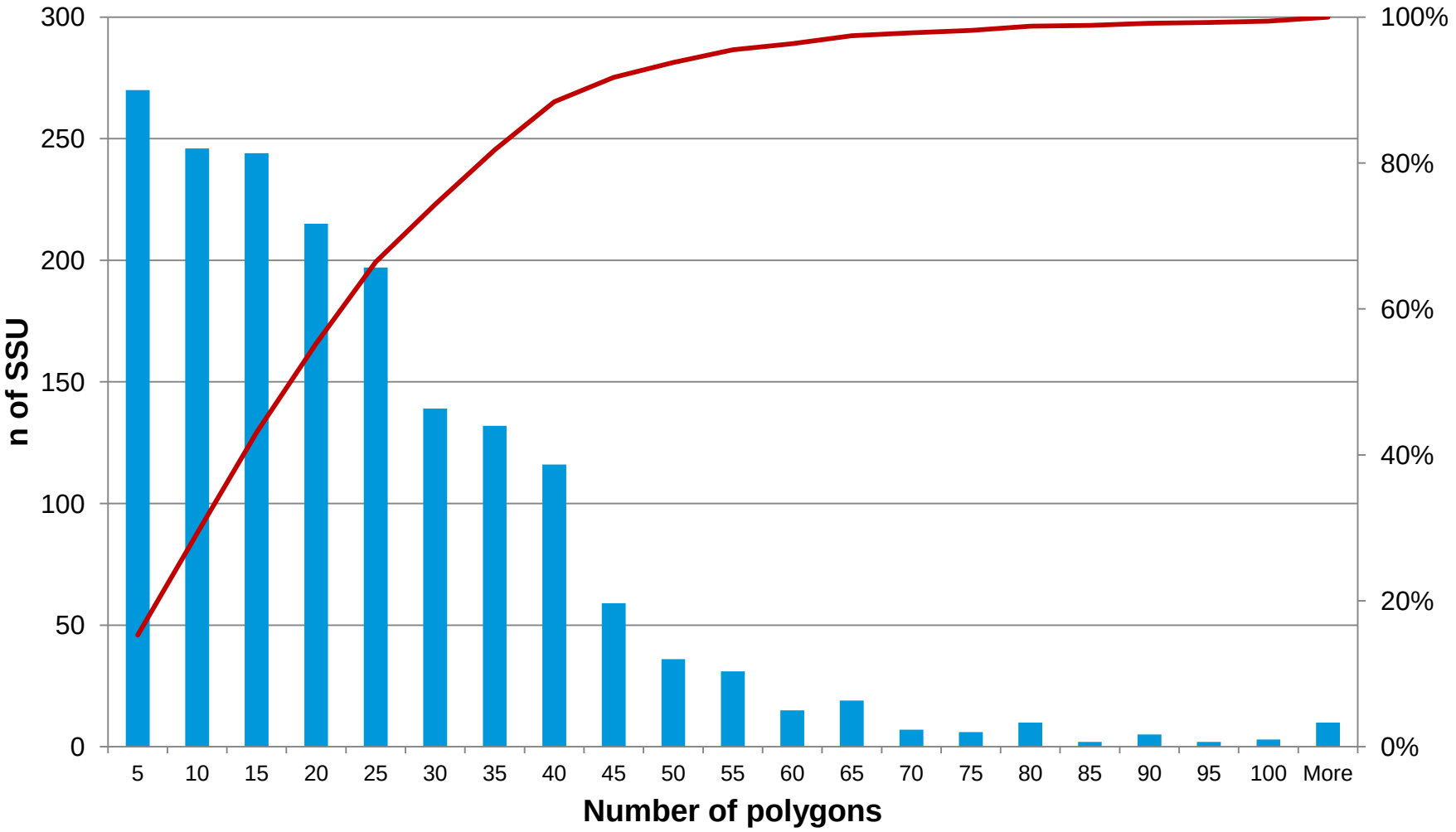


Land Cover Change : 8 % of validation SSU

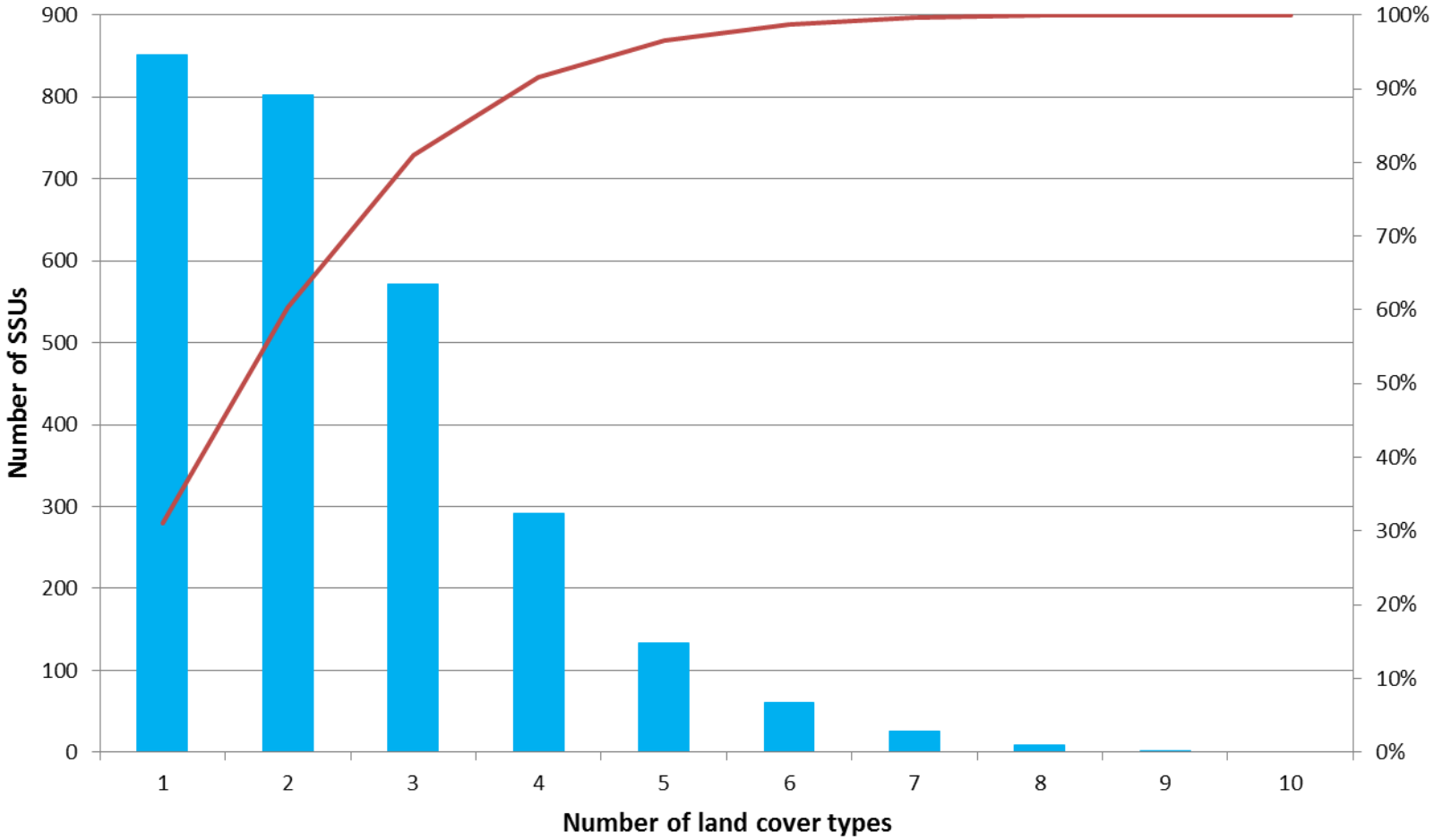
- 8% of SSUs with at least 1 change between the 3 epochs
- Large spatial variability for change detection : main LC changes in Central America, Amazon, South and Southeast Asia, China and Europe



Validation dataset : distribution of number of objects per SSU



Validation dataset : distribution of number of LC type per SSU



SSUs interpretation for 1 epoch

- SSUs are described by 1 (min) to 9 (max) LC classes and in average by 2.3 LC classes

ssu	object ID	f_area	lccode2010
4648	1	288370,440907999	130
4648	2	927238,267814000	130
4648	3	54224,168645100	130
4648	4	24939,539963700	130
4648	5	15234,173347800	130
4648	6	3908,392787790	130
4648	7	27758,147317100	130
4648	8	3315,001765940	130
4648	9	1289,484335800	130
4648	10	2436,326599120	130
4648	11	4444,726980490	130
4648	12	2898,486914070	130
4648	13	1243,838872560	130

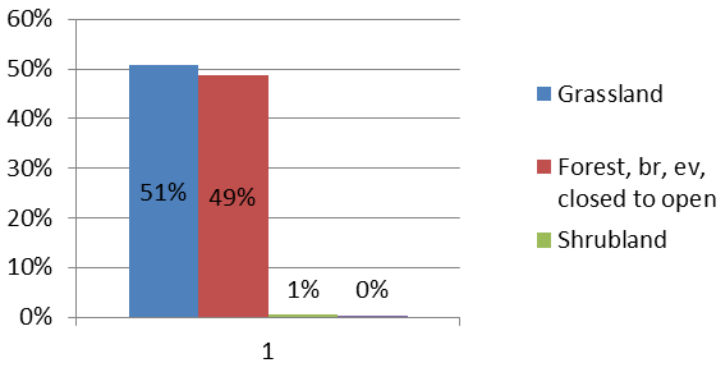
ssu	object ID	f_area	lccode2010
3800	1	1056421,05	130
3800	2	281371,587	50
3800	3	1009,90587	150
3800	4	732233,562	50
3800	5	13807,7526	120

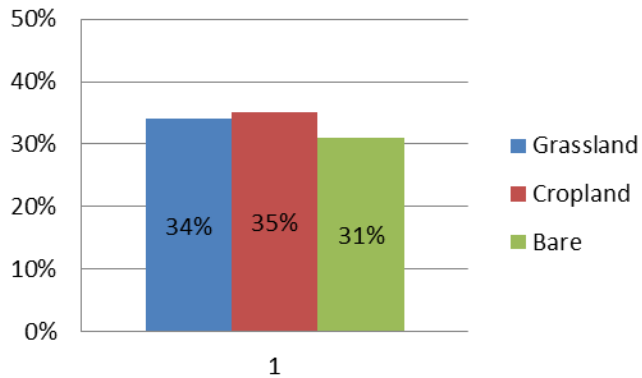
ssu	object ID	f_area	lccode2010
3802	1	22183,69512220000	120
3802	2	4108,09168939000	50
3802	3	11485,53968150000	50
3802	4	29566,84879720000	150
3802	5	22891,19980220000	120
3802	6	10932,08843980000	130
3802	7	2276,56747779000	120
3802	8	1169,66499498000	120
3802	9	1118,31384864000	120
3802	10	3029,71762092000	120
3802	11	839873,41629199900	130
3802	12	26936,52897920000	150
3802	13	4130,91442090000	120
3802	14	25618,51622940000	130
3802	15	4695,77702817000	120
3802	16	40878,92053640000	120
3802	17	169063,21614000000	130
3802	18	21647,36092940000	120
3802	19	137298,26310499900	130
3802	20	141009,33624700000	120
3802	21	149512,54692699900	130
3802	22	27062,05400300000	150

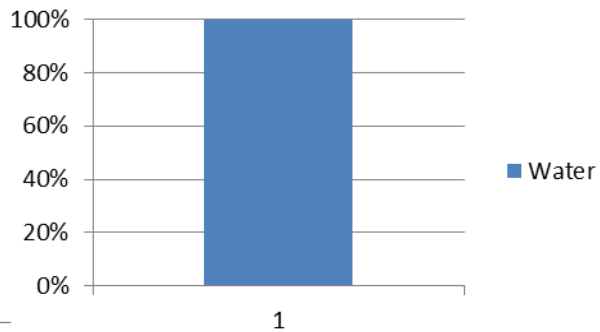
- 3 modes of interpretation:
  - 1 LC class by SSU
  - several LC classes by SSU
  - mosaic class by SSU

1 LC type by SSU

- Identify the « ***dominant »*** LC type
- What is the threshold to be « dominant »?
  - Majoritary (independently from the area covered)
  - > 50%
  - > arbitrary threshold (e.g. 60, 70, 80, 90%)
  - 100%



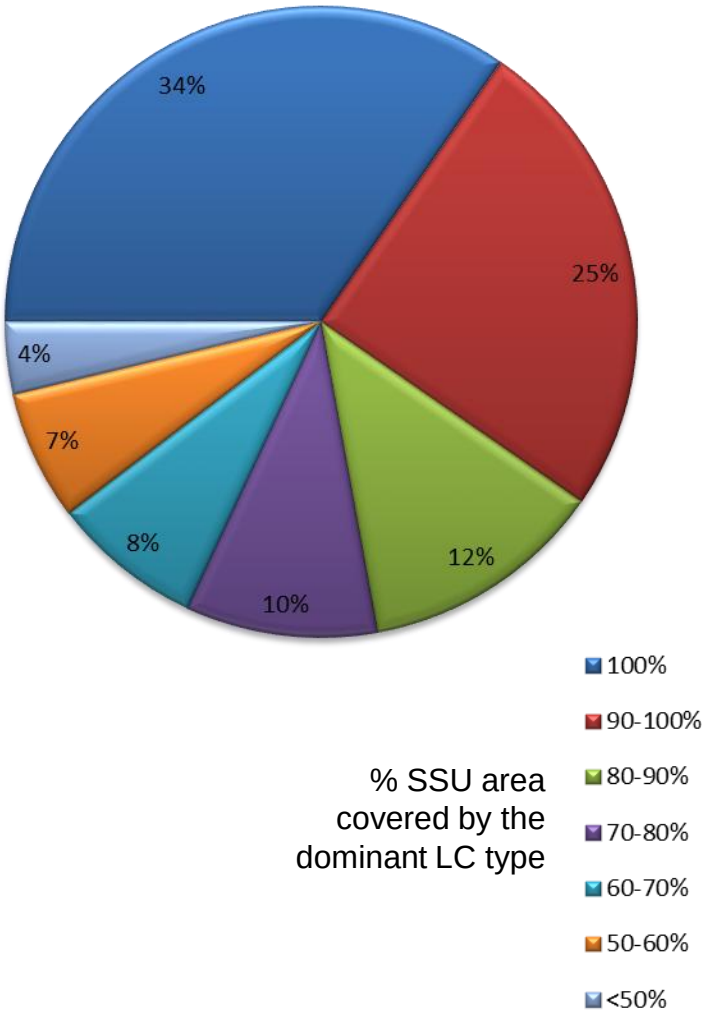




Area covered by the dominant LC class

- The dominant LC class covers more than 50% of the SSU area for 96% of the SSU

- 80% of the SSUs are covered by a dominant LC class representing more than 70% of the box



% SSU area covered by the dominant LC type

Validation activities in process

- **CCI Land Cover validation** of the Global land cover time series (CCI LC v.2)

- **2-tier cropland validation experiment** with IIASA and JRC (SIGMA):

  Expert-based validation database including expert reliability assessment

  (all experts belonging to 20+ SIGMA partners)

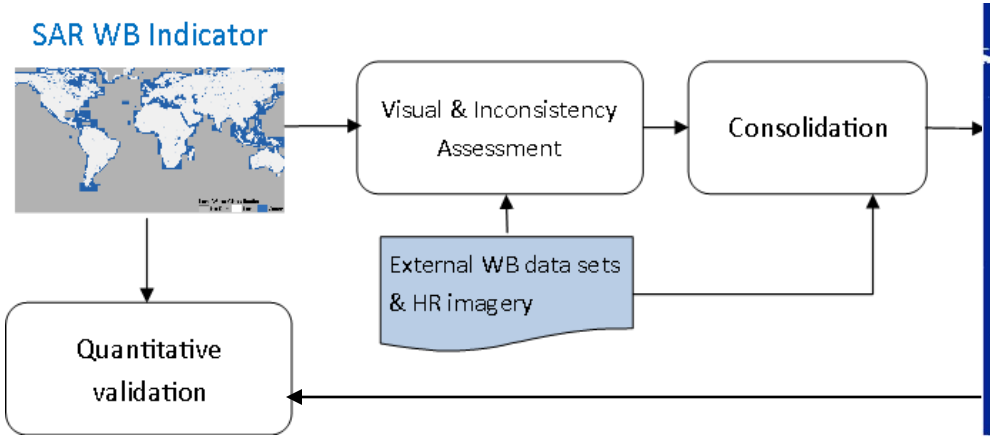
  versus

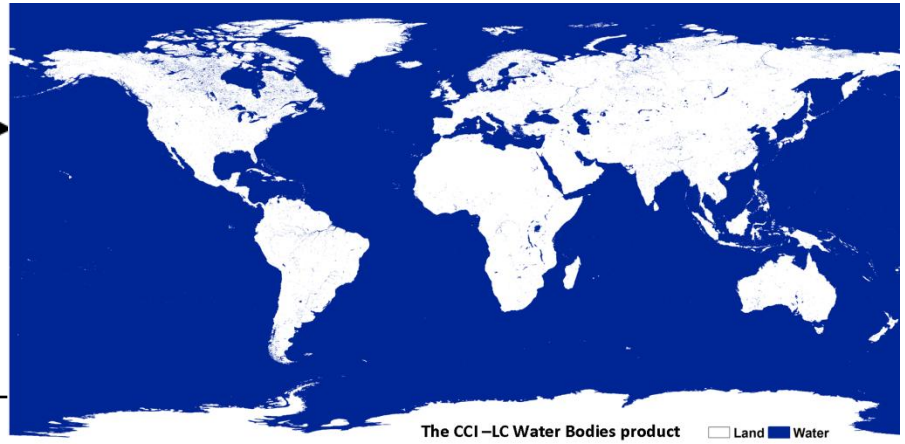
  Crowd sourcing based validation database for the same points

- **CEOS Global Land Cover Validation recommendations** not so efficient to inter-compare **global land cover products**

Inter-comparison products for 150 m Global Water Body and Coastline product

	SAR-WBI	GFC- datamask	GIW v1.0
Sensor/mission	ASAR Envisat	Landsat	Landsat
Data source	Wide Swath Mode (WSM), Image Mode Medium-resolution (IMM), Global Monitoring Image Mode (GM1)	Landsat time series	Global Land Survey collection of Landsat images
Spatial resolution	75 m, 150 m, 1,000 m for WSM, IMM and GM1, respectively	30 m	30 m
Time interval	2005 – 2012	2000 - 2012	2000
Legend	3 classes: potential WB, land, NaN	3 classes: no data, mapped land surface, permanent WB	6 classes: no data, land, water, snow/ice, cloud shadow and cloud.
Spatial extent	84° N - 60° S	80° N - 57° S	90° N - 60° S
Reference	[12]	[23]	[24]



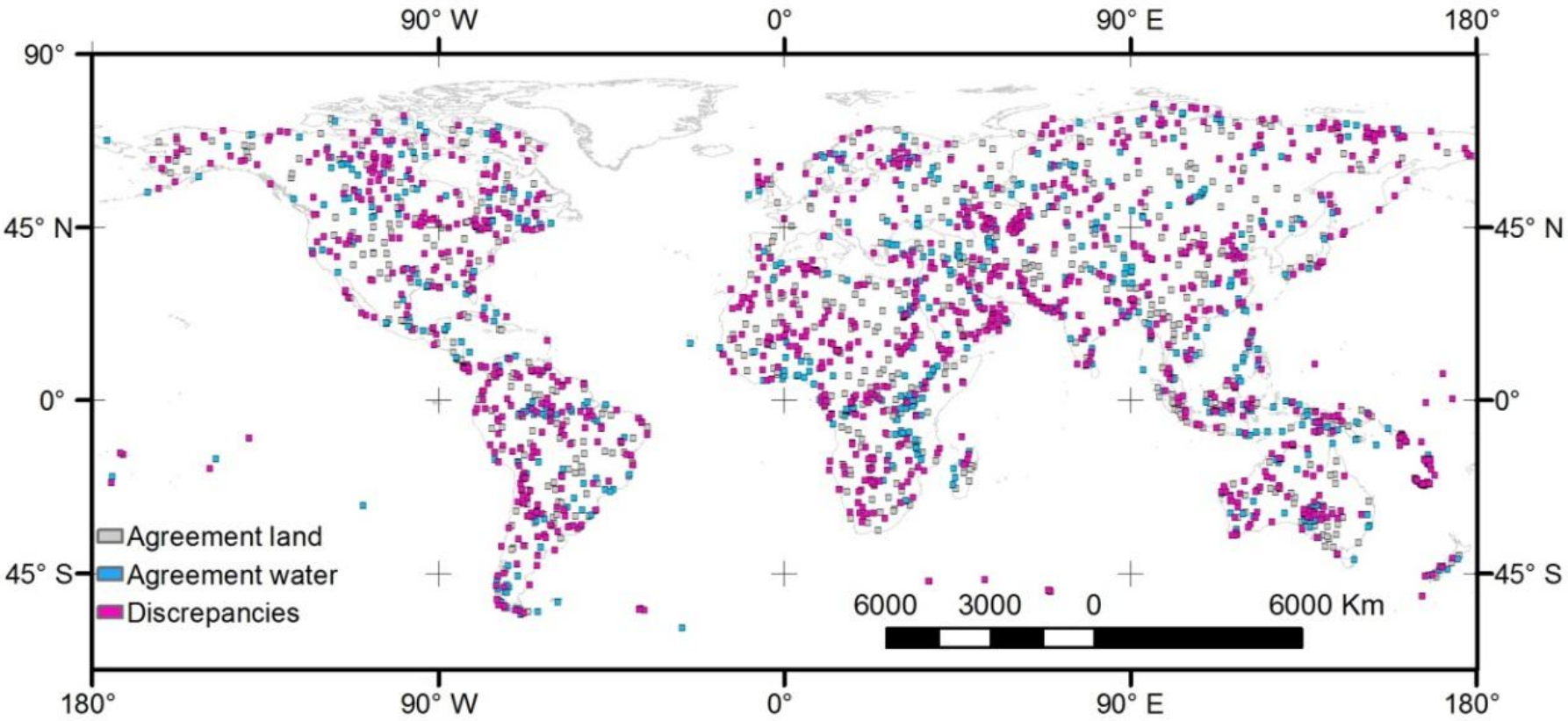


Overall accuracy of 4 global water / land mask : not really sensitive

	Spatial coverage		Non-Water			Water		
	(LAT)	OA (%)	PA (%)	UA (%)	F-Score (%)	PA (%)	UA (%)	F-Score (%)
SAR-WBI*	84° N - 60° S	98	99	99	99	62	84	71
GIW v1.0*	90° N - 60° S	98	99	99	99	78	77	77
GFC-datamask*	80° N - 57° S	99	100	99	100	79	97	87
global map of open WB	90° N - 90° S	99	99	100	100	92	86	89

* Area weighted overall accuracy

Sampling biased towards error-prone areas (50 % of samples in discrepancy areas)



Overall accuracy based on sampling biased towards error-prone areas

		Non-Water			Water		
	OA (%)	**PA (%)**	**UA (%)**	**F-Score (%)**	**PA (%)**	**UA (%)**	**F-Score (%)**
SAR-WBI	74	90	79	84	23	43	30
GIW v1.0	86	92	90	91	67	72	70
GFC-datamask	89	98	89	93	61	89	73
CCI global map of open WB	89	93	92	93	75	78	77

Lamarche et al., 2016 in revision

Some items for discussion

- Overall accuracy assessment standards to be updated:

  Sthraler et al., 2006 applied several times at global scale

  => lessont learnt to update the CEOS Cal/Val LC standards

- Inter-comparison of products to advise on products and recommended uses should be a separate quality assessment process different from the validation

- New validation data collection strategy proved to be very productive for training data but it is urgent to assess their respective efficiency for accuracy assessment

- How to consolidate a validation dataset as a community effort and to keep it as common validation set for standard assessment ?

- What could be the common strategy to validate the land cover change precisely ?