

Large scale automatic land cover map production with Sentinel-2 & 1 image time series: current status and outlooks

J. Inglada^{1,2}, M. Arias¹, A. Vincent¹, B. Tardy^{1,2}, D. Morin¹, J. Michel², O.Hagolle^{1,2}



[2016-10-11 Tue]



Outline

① Fully automatic land cover map production

② Examples

③ How we do it

④ Ongoing Improvements

⑤ Conclusion



Fully automatic land cover map production

- ▶ Up-to-date land cover maps are a crucial asset for many research works and operational applications
- ▶ Need for regular and frequent updates
- ▶ Need for a shorter delay between image acquisition and map availability:
 - ▶ CLC 2012 was delivered in may 2015;
 - ▶ Copernicus HRL 2012 "no yet validated" and 2015 reference year AO opened in 2016.

Fully automatic land cover map production

- ▶ Up-to-date land cover maps are a crucial asset for many research works and operational applications
- ▶ Need for regular and frequent updates
- ▶ Need for a shorter delay between image acquisition and map availability:
 - ▶ CLC 2012 was delivered in may 2015;
 - ▶ Copernicus HRL 2012 "no yet validated" and 2015 reference year AO opened in 2016.
- ▶ For high quality maps and short production time we need:
 - ▶ automatic, robust and reliable methods,
 - ▶ able to efficiently exploit all available data.

Fully automatic land cover map production

- ▶ Up-to-date land cover maps are a crucial asset for many research works and operational applications
- ▶ Need for regular and frequent updates
- ▶ Need for a shorter delay between image acquisition and map availability:
 - ▶ CLC 2012 was delivered in may 2015;
 - ▶ Copernicus HRL 2012 "no yet validated" and 2015 reference year AO opened in 2016.
- ▶ For high quality maps and short production time we need:
 - ▶ automatic, robust and reliable methods,
 - ▶ able to efficiently exploit all available data.
- ▶ Which data ?
 - ▶ Need of large coverage and high resolution images
 - ▶ Vegetation phenology conveys a lot of information
 - ▶ ⇒ use of Sentinel-2 image time series.
 - ▶ Work going on to also add Sentinel-1

The Theia Land Cover product

- ▶ In the frame of the French Theia Land Data Centre
 - ▶ <http://www.theia-land.fr/>
- ▶ **Fully automatic** production of LC maps at the national scale:
 - ▶ 20 class nomenclature;
 - ▶ 20 m resolution;
 - ▶ yearly updated;
 - ▶ delivered at the latest 3 months after the reference period.

The Theia Land Cover product

- ▶ In the frame of the French Theia Land Data Centre
 - ▶ <http://www.theia-land.fr/>
- ▶ **Fully automatic** production of LC maps at the national scale:
 - ▶ 20 class nomenclature;
 - ▶ 20 m resolution;
 - ▶ yearly updated;
 - ▶ delivered at the latest 3 months after the reference period.
- ▶ Input data:
 - ▶ high resolution optical image time series (mainly Sentinel-2, but also Landsat8);
 - ▶ later on Sentinel-1, SPOT6, Pleiades
 - ▶ reference data for **automatic** model calibration and **automatic** product validation.

The Theia Land Cover product

- ▶ In the frame of the French Theia Land Data Centre
 - ▶ <http://www.theia-land.fr/>
- ▶ **Fully automatic** production of LC maps at the national scale:
 - ▶ 20 class nomenclature;
 - ▶ 20 m resolution;
 - ▶ yearly updated;
 - ▶ delivered at the latest 3 months after the reference period.
- ▶ Input data:
 - ▶ high resolution optical image time series (mainly Sentinel-2, but also Landsat8);
 - ▶ later on Sentinel-1, SPOT6, Pleiades
 - ▶ reference data for **automatic** model calibration and **automatic** product validation.
- ▶ Compatible with and complementary to existing products:
 - ▶ CLC is less *fresh*, has lower temporal and spatial resolution, but has more classes (44) and probably better accuracy;
 - ▶ Urban Atlas is very detailed but limited in spatial extent and not frequently updated;

The map or the process?

- ▶ Non goal
 - ▶ Produce a hand-crafted

The map or the process?

- ▶ Non goal
 - ▶ Produce a hand-crafted, semantically rich

The map or the process?

- ▶ Non goal
 - ▶ Produce a hand-crafted, semantically rich, 99% accurate land cover map

The map or the process?

- ▶ Non goal
 - ▶ Produce a hand-crafted, semantically rich, 99% accurate land cover map, which is available too late.

The map or the process?

- ▶ Non goal
 - ▶ Produce a hand-crafted, semantically rich, 99% accurate land cover map, which is available too late.
- ▶ Constraints
 - ▶ Full automation: for efficiency and reproducibility.
 - ▶ No hand-made or class/landscape/region specific rules: open to nomenclature changes.
 - ▶ Computationally efficient processing: short delivery delay, reprocessing when specs change.

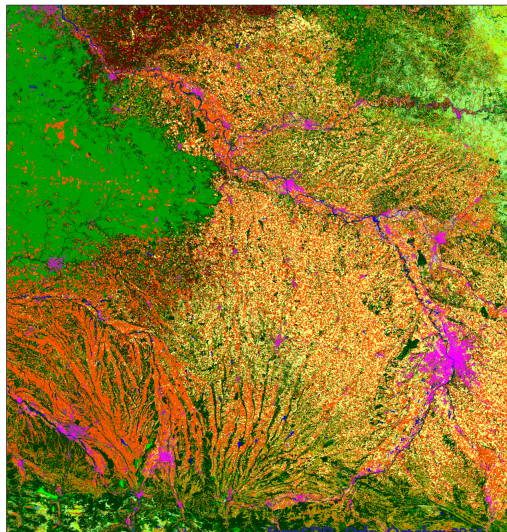
Outline

- ① Fully automatic land cover map production
- ② **Examples**
- ③ How we do it
- ④ Ongoing Improvements
- ⑤ Conclusion

Product prototypes using Landsat8



Product prototypes using Landsat8



Légende

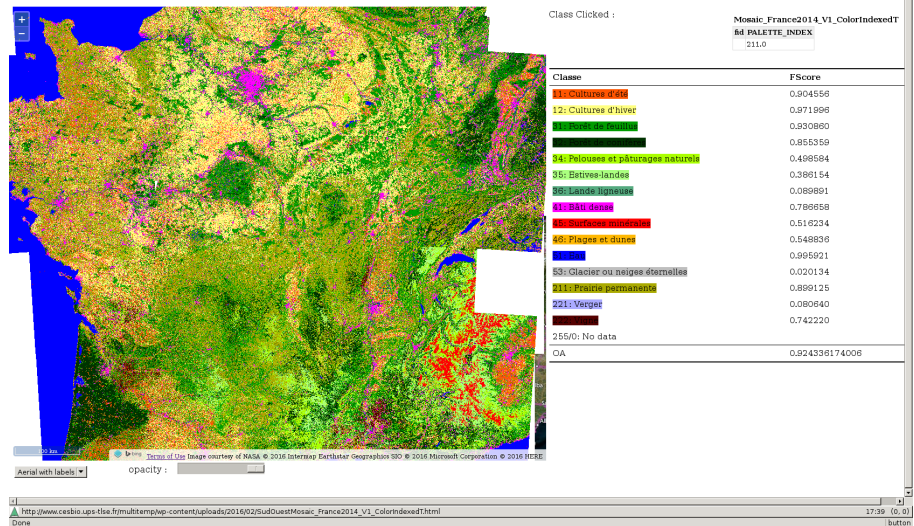
OccupationSol_2013_16classes

- Cultures d'été
- Cultures d'hiver
- Forêt Feuilles Caduques
- Forêt Feuilles Persistantes
- Forêts mélangées
- Pelouses
- Estives-landes
- Lande ligneuse
- Bâti
- Surfaces minérales
- Eau
- Mer et océans
- Glacier ou neiges éternelles
- Prairie permanente
- Verger
- Vigne



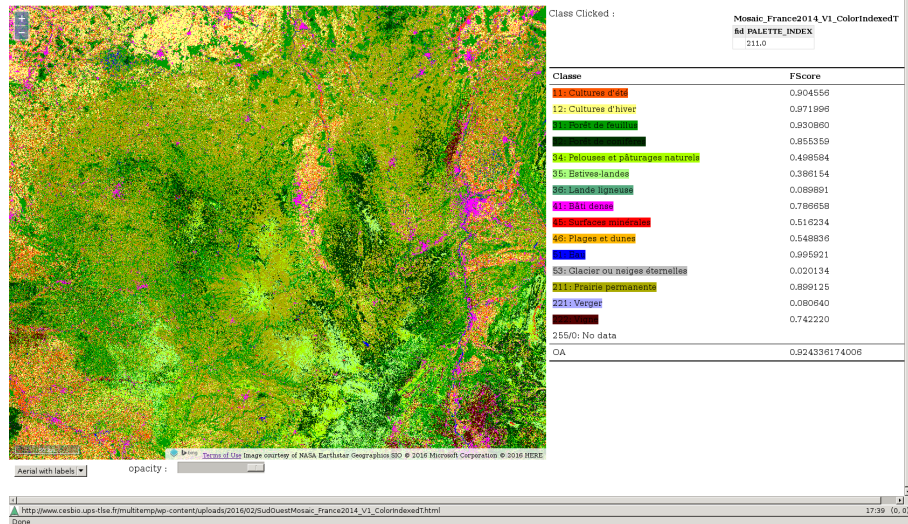
10 0 10 20 30 40 Km.

France Landsat8 2014 v1



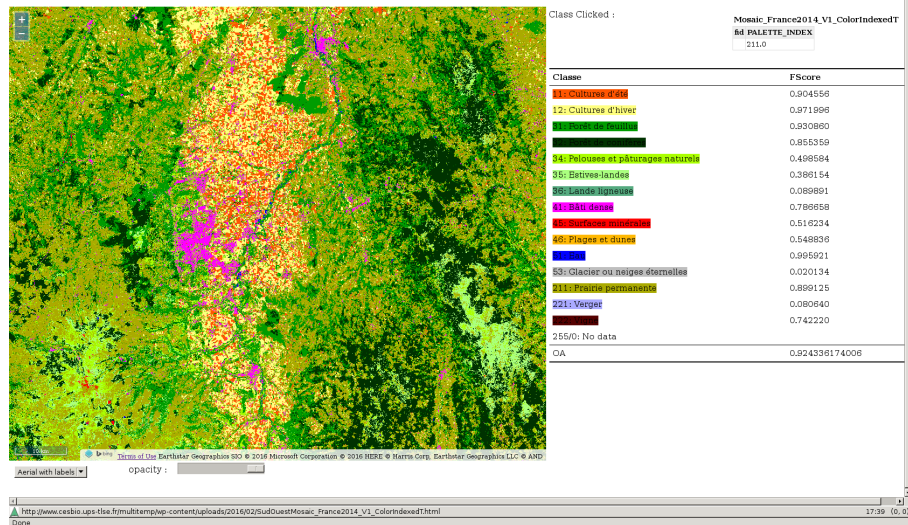
<https://frama.link/ocs2014ba>

France Landsat8 2014 v1



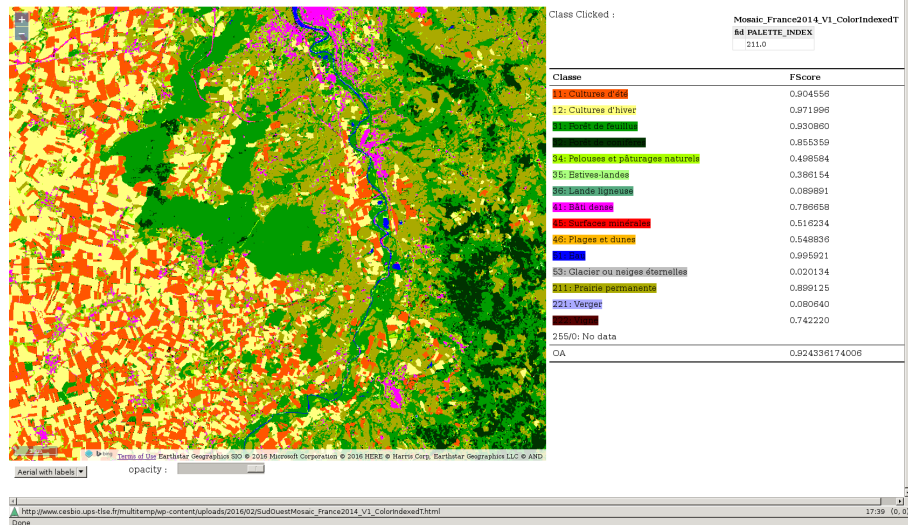
<https://frama.link/ocs2014ba>

France Landsat8 2014 v1



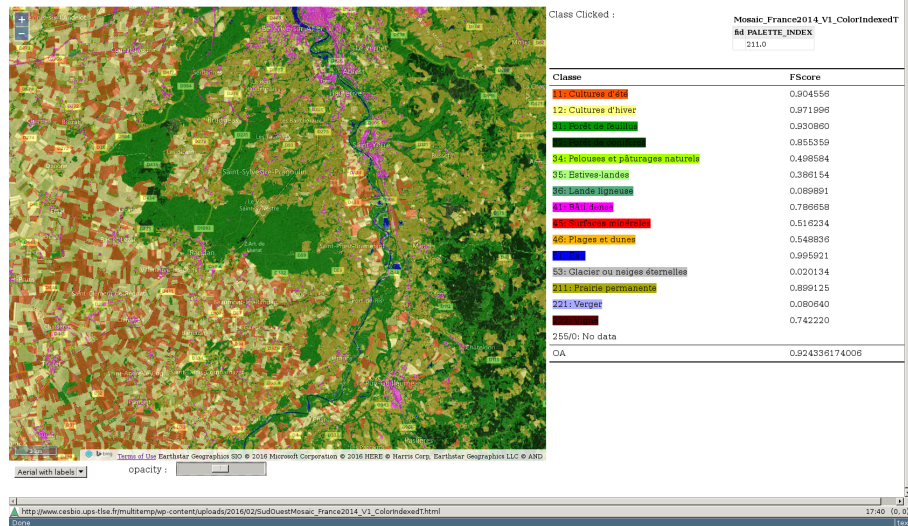
<https://frama.link/ocs2014ba>

France Landsat8 2014 v1

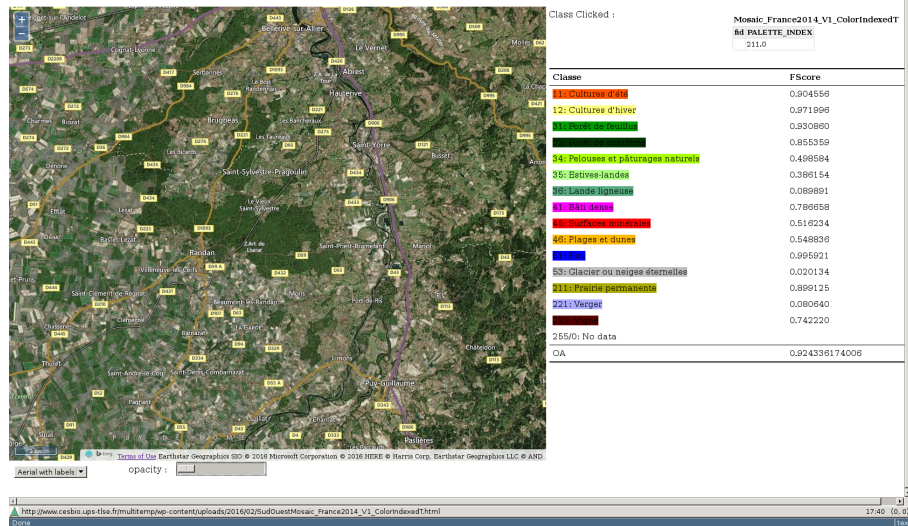


<https://frama.link/ocs2014ba>

France Landsat8 2014 v1

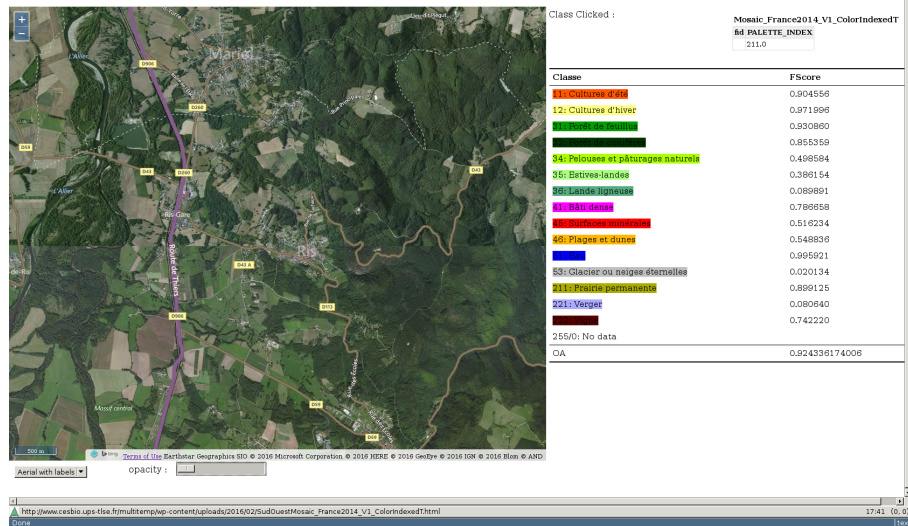


France Landsat8 2014 v1



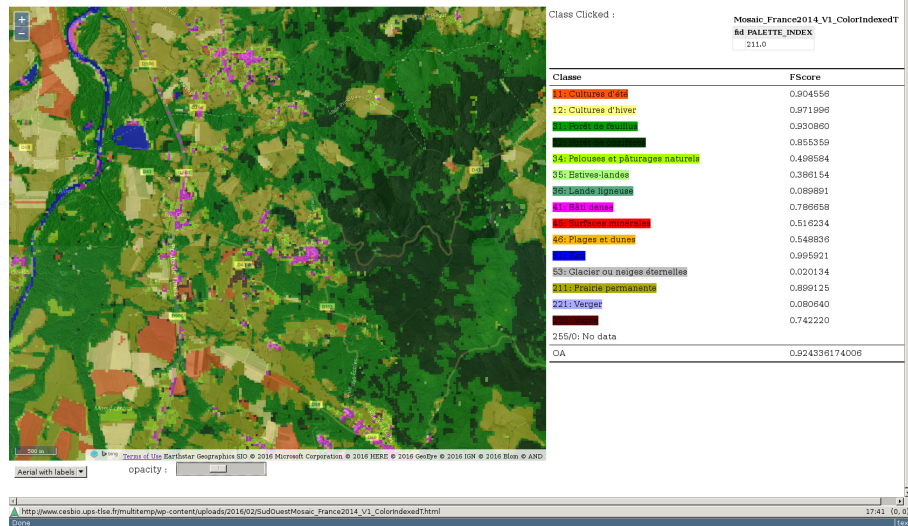
<https://frama.link/ocs2014ba>

France Landsat8 2014 v1



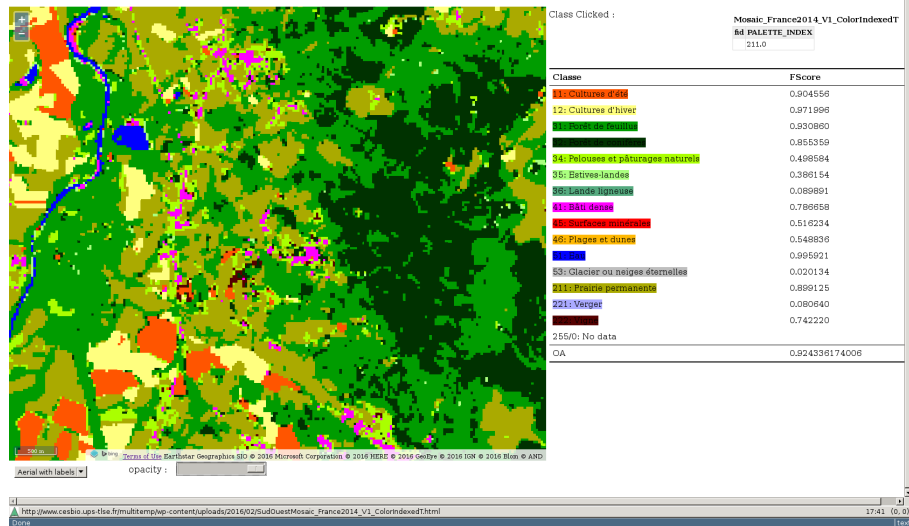
<https://frama.link/ocs2014ba>

France Landsat8 2014 v1



<https://frama.link/ocs2014ba>

France Landsat8 2014 v1



<https://frama.link/ocs2014ba>

Outline

- ① Fully automatic land cover map production
- ② Examples
- ③ How we do it
- ④ Ongoing Improvements
- ⑤ Conclusion

The iota2 processing chain

- ▶ Free software implementation available for download and contributions
 - ▶ 2
 - ▶ <https://frama.link/iota2>
- ▶ HPC cluster or classical workstation

¹Glue code for GDAL/OGR and Orfeo Toolbox Applications

²-based custom modules

scheduling

J. Inglada et al.

The iota2 processing chain

- ▶ Free software implementation available for download and contributions
 - ▶ 
 - ▶ <https://frama.link/iota2>
- ▶ HPC cluster or classical workstation
- ▶ Total Physical Source Lines of Code (SLOC)=11,382

Language	SLOC	
python ¹	8476	(74.47%)
c++ ²	2044	(17.96%)
unix shell ³	858	(7.54%)

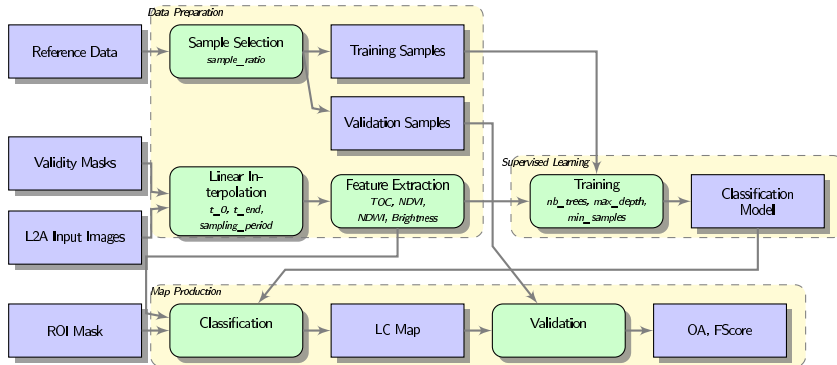
¹Glue code for GDAL/OGR and Orfeo Toolbox Applications

²-based custom modules

³scheduling



Kernel of the processing chain



Particularities of the approach

- ▶ Fully automatic, no human operator
 - ▶ Supervised classification.

Particularities of the approach

- ▶ Fully automatic, no human operator
 - ▶ Supervised classification.
- ▶ All available images are used
 - ▶ no date selection;
 - ▶ no scene elimination because of cloud cover: 80% of clouds is 20% of clear pixels.

Particularities of the approach

- ▶ Fully automatic, no human operator
 - ▶ Supervised classification.
- ▶ All available images are used
 - ▶ no date selection;
 - ▶ no scene elimination because of cloud cover: 80% of clouds is 20% of clear pixels.
- ▶ Efficient processing
 - ▶ 1 year of Landsat8 data over France processed (from L2A to LC map) in a day

Particularities of the approach

- ▶ Fully automatic, no human operator
 - ▶ Supervised classification.
- ▶ All available images are used
 - ▶ no date selection;
 - ▶ no scene elimination because of cloud cover: 80% of clouds is 20% of clear pixels.
- ▶ Efficient processing
 - ▶ 1 year of Landsat8 data over France processed (from L2A to LC map) in a day
 - ▶ OK, a week if you don't have a 100 node HPC cluster!

Particularities of the approach

- ▶ Fully automatic, no human operator
 - ▶ Supervised classification.
- ▶ All available images are used
 - ▶ no date selection;
 - ▶ no scene elimination because of cloud cover: 80% of clouds is 20% of clear pixels.
- ▶ Efficient processing
 - ▶ 1 year of Landsat8 data over France processed (from L2A to LC map) in a day
 - ▶ OK, a week if you don't have a 100 node HPC cluster!
 - ▶ and some optimisations are on the way;
 - ▶ allows for easy benchmarking and evaluation.

From tile to country

- ▶ Homogeneous acquisition dates
 - ▶ several orbits needed for a large territory;
 - ▶ gaps due to cloud cover;
 - ▶ time series used as classification features.

From tile to country

- ▶ Homogeneous acquisition dates
 - ▶ several orbits needed for a large territory;
 - ▶ gaps due to cloud cover;
 - ▶ time series used as classification features.
- ▶ Spatial stratification, tiling?
 - ▶ data volume;
 - ▶ intra-class variability, landscape variability.

From tile to country

- ▶ Homogeneous acquisition dates
 - ▶ several orbits needed for a large territory;
 - ▶ gaps due to cloud cover;
 - ▶ time series used as classification features.
- ▶ Spatial stratification, tiling?
 - ▶ data volume;
 - ▶ intra-class variability, landscape variability.
- ▶ Reference data availability
 - ▶ spatial distribution;
 - ▶ outdated samples;
 - ▶ quality of the reference data (label noise).

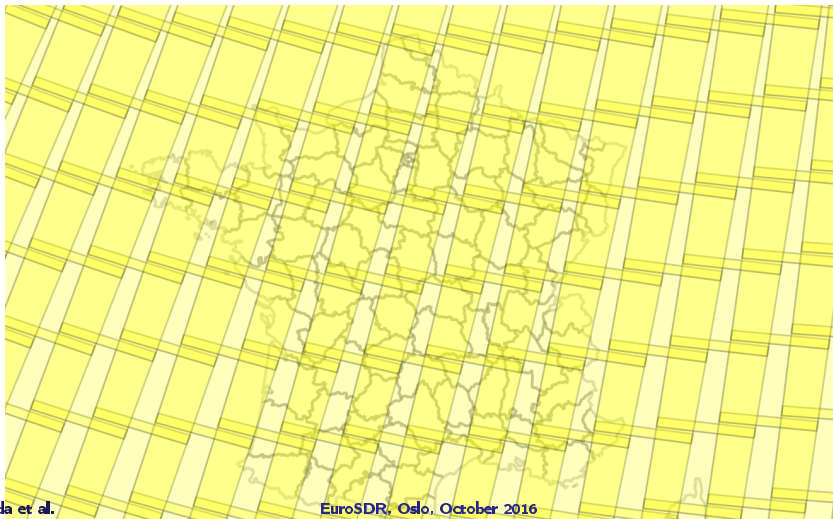
The problem of homogeneous cover

- The area to map



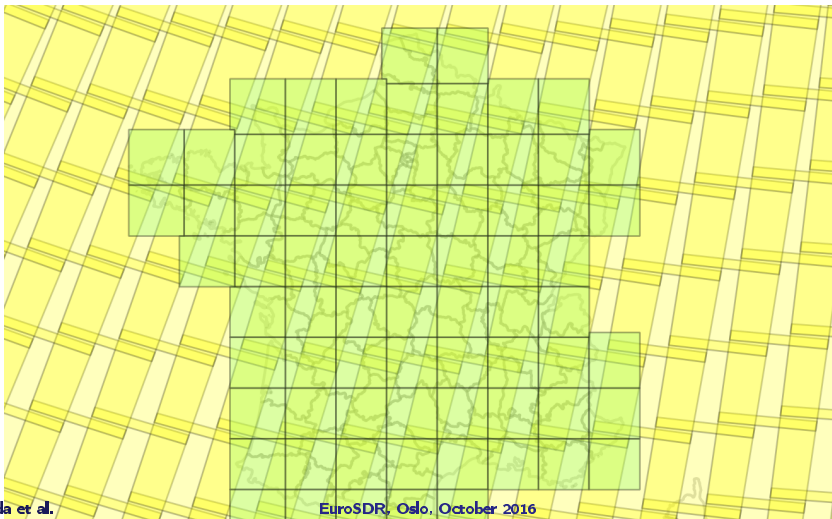
The problem of homogeneous cover

- Image acquisitions: each orbit has different dates



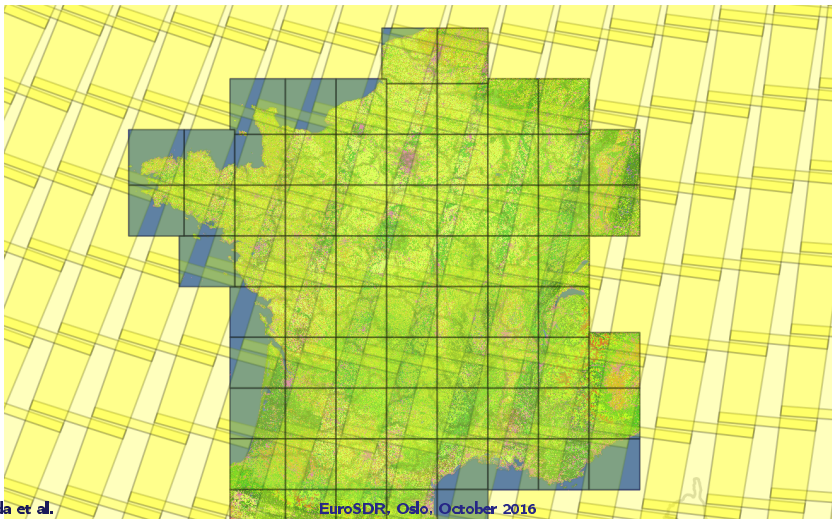
The problem of homogeneous cover

- Data storage: tiles for convenience



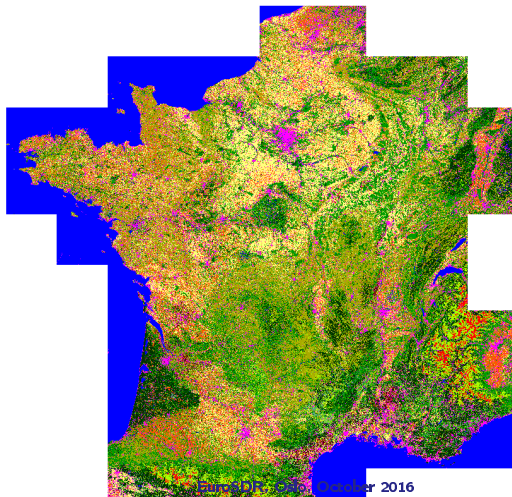
The problem of homogeneous cover

- The map: continuity and homogeneity



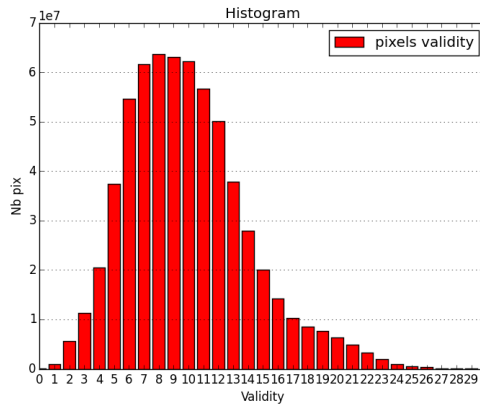
The problem of homogeneous cover

- The map: continuity and homogeneity



The problem of homogeneous cover

- Count of clear pixels (validity)



The problem of homogeneous cover

- ▶ Solution: resample all image data to a common temporal grid
 - ▶ Interpolation is already used in the gap-filling step
 - ▶ Use a set of common virtual dates (twice a month)
- ▶ Does this degrade accuracy with respect to the case where a single orbit is used?

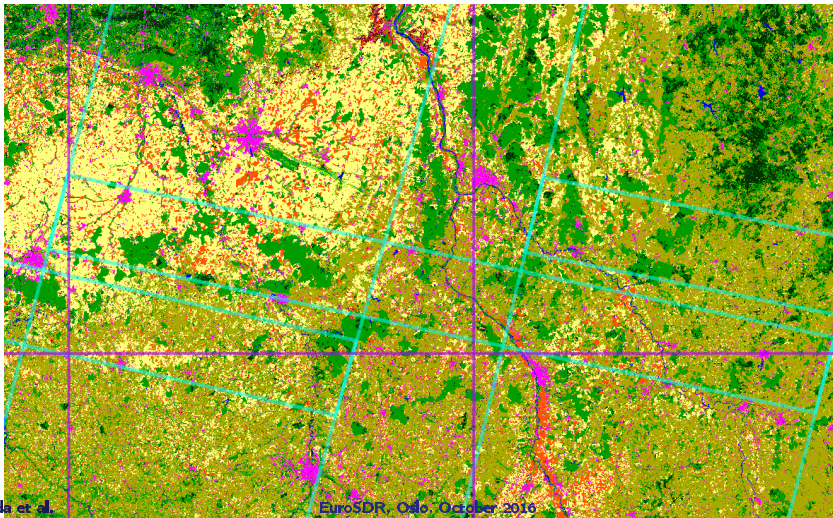
The problem of homogeneous cover

- ▶ Solution: resample all image data to a common temporal grid
 - ▶ Interpolation is already used in the gap-filling step
 - ▶ Use a set of common virtual dates (twice a month)
- ▶ Does this degrade accuracy with respect to the case where a single orbit is used?
 - ▶ yes but only **very slightly**, and enhances **a lot** multi-orbit classification

Site	OA original	OA resampled
Argentina	0.904 ± 0.026	0.893 ± 0.028
Belgium	0.816 ± 0.001	0.816 ± 0.001
Burkina Faso	0.522 ± 0.018	0.503 ± 0.018
China	0.927 ± 0.010	0.911 ± 0.031
France	0.904 ± 0.004	0.905 ± 0.004
Madagascar	0.501 ± 0.058	0.498 ± 0.070
Morocco	0.876 ± 0.004	0.875 ± 0.009
Pakistan	0.727 ± 0.035	0.723 ± 0.036
Russia	0.665 ± 0.019	0.665 ± 0.019
South Africa	0.907 ± 0.011	0.894 ± 0.015
Ukraine	0.740 ± 0.025	0.724 ± 0.024
US-Maricopa	0.911 ± 0.006	0.911 ± 0.006

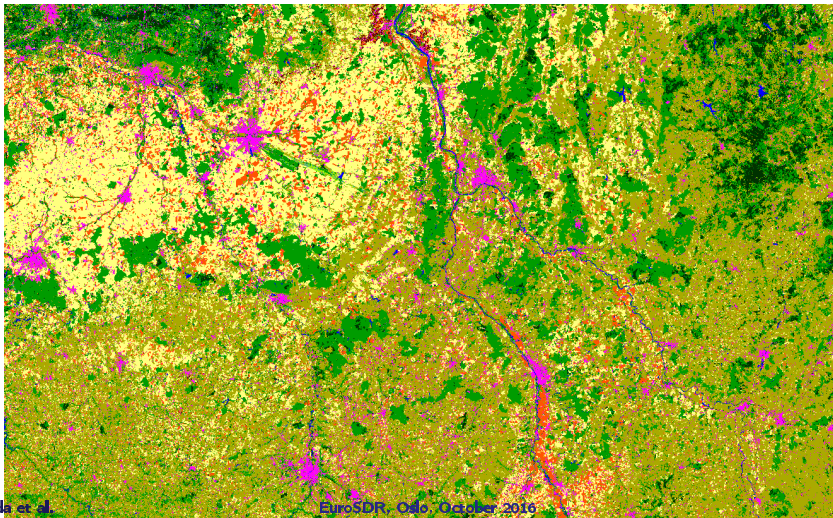
The problem of homogeneous cover

- The map: continuity and homogeneity

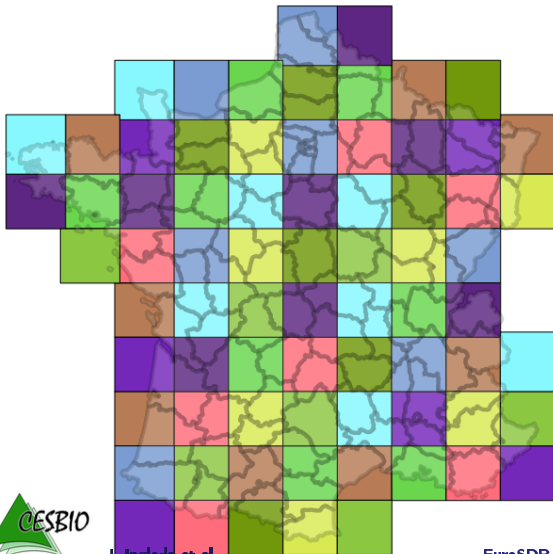


The problem of homogeneous cover

- The map: continuity and homogeneity



Speeding up the classification



- ▶ A different classifier per group of tiles
 - ▶ allows for parallel learning.