

## Final Report

**Copernicus Land Monitoring 2014 – 2020  
in the framework of Regulation (EU) No 377/2014 of the European  
Parliament and of the Council of 3 April 2014**

**Specific Contract No 3436/R0-COPERNICUS/EEA. 56936  
Implementing Framework service contract No  
EEA/IDM/R0/16/009/Finland  
Finland**

### Tasks:

1. <b>Verification</b> of 2012 reference year local component products and enrichment of Urban Atlas (if applicable)	<b>Y Not yet</b>
2. Production of <b>CLC</b> for the 2018 reference year	<b>Y</b>
1. Post-production <b>verification</b> of the High Resolution Layers (HRL's) for the 2015 reference year	<b>Y</b>
2. <b>Dissemination</b>	<b>Y</b>

### Authors:

Pekka Härmä, Iida Autio, Riitta Teiniranta, Suvi Hatunen, Markus Törmä, Minna Kallio, Minna Kaartinen

## Table of content

- Background
- Project management and participating experts
- Verification of 2012 reference year local component products
  - The usability and potential use of the Copernicus Land Local Component Products in Finland
- Corine Land Cover 2018 update
  - Organisation of work at national level
    - Interpretation of EO data
    - Processing of national spatial data
    - Mapping of land cover changes
    - Production of CORINE Land Cover 2018 map
  - Internal quality control
    - QC during data production
    - Comparison of HR Corine land cover 2018 with national statistics
    - LUCAS 2015 micro data and HR Corine land Cover 2018
  - External quality control
  - Main difficulties and solutions
  - Ancillary data used
- Post-production verification of the High Resolution Layers for the 2015 reference year
- Dissemination
  - Spatial Data Infrastructure of SYKE facilitates efficient data dissemination
  - INSPIRE compatible map services
- Conclusions
- References

## Background

Finland joined the Corine land cover programme year 2000. Thereafter CLC2000, CLC2006, CLC2012 and CLC2018 updates have been produced. Finnish Environment Institute (SYKE) has been responsible for the data production.

Additionally

- During GIO Land 2011-2012 SYKE participated also in the verification and enhancement of the High Resolution Layers (together with former Forest Research Institute, now Natural Resources Institute).
- SYKE has participated in the Geoland2, GSE-land and HELM projects and has been part of the EAGLE group.
- SYKE is a service provider in Copernicus Global Component Cryosphere.
- SYKE is also an active member of the national Copernicus User Forum.
- The Finnish Meteorological Institute's (FMI) hosts the National Satellite Data Centre (NSDC) in northern Finland (Sodankylä) focusing on fast delivery remote sensing product generation. This includes the hosting of a Copernicus Sentinel Collaborative Ground Station for Sentinel satellite data distribution and archiving, cloud services and maintaining external processing chains in virtual environments.
- SYKE has parallel processing facilities (Calvalus developed by Brockmann Consult GmbH) at NSDC. Selected EO data (Sentinel 2 and 3, Landsat 8, Envisat, Terra, Aqua) are stored online covering Finland and surrounding areas. This enables development of new algorithms and user defined re-processing (L1 -> L3) of EO data producing time-series of spatio-temporally gridded products.
- Presently processing chains of selected intermediate products based on Sentinel 1, Sentinel 2 and Sentinel 3 are under development. This includes production of Level 3 data on regular basis, like time series of calibrated backscattering mosaics, cloud-free reflectance mosaics and spectral indices. These data will be put available for different users in Finland using standard interfaces (WMS, WCS, ...). The objectives are to make basic, common pre-processing steps only once and make deployment of Sentinel data easy for various applications.

In the current contract SYKE has completed following tasks

- Verification of 2012 reference year local component products
- Production of CLC (Corine Land Cover) and CLCC for the 2018 reference year
- Post-production verification of the High Resolution Layers (HRL's) for the 2015 reference year (with Natural Resource Institute Finland as a subcontractor)
- Dissemination

Enrichment of Urban Atlas was not included in this 1<sup>st</sup> specific contract, since the amount of work was already large with tight time-table. This enrichment task will be proposed for the 2<sup>nd</sup> contract.

## Project management and participating experts

Project Steering group met when needed and consisted of following persons:

- Pekka Härmä (Technical project manager), Pekka Salminen (Administrative manager), Riitta Teiniranta, Jarkko Miettinen (Administrative support)

Technical team met regularly every other week during the project and consisted of following persons:

- Pekka Härmä, Suvi Hatunen, Iida Autio, Markus Törmä, Minna Kallio, Minna Kaartinen, Riitta Teiniranta, Mikko Kervinen, Sofia Juntila

## Verification of 2012 reference year local component products

The task was completed according to the guidelines. However, Street Tree Layer (STL) was verified using airborne Lidar data by comparing the whole coverage of the data with DSM (digital surface model) derived using Lidar data. Thus no sampling was needed. STL data covered only city of Lahti in Finland. Also in the verification of Green Linear Elements (GLE) Lidar data was used in a subset of GLE ta in order to evaluate omission errors. Summary of the characteristics of the local component products are presented in following the table 1.

Table 1. Summary of the statistical verification of the local component products

Characteristics of the local component products in Finland	Urban Atlas	Street tree layer	Riparian zones	Green linear element	Natura2000
Area covered within Finland	10,5 %	0,01 %	13 %	0,03 %	0,06 %
Number of valid classes	25	1	65	4	20
Number of samples selected	172	100 % of data verified	573	100	154
Overall accuracy	89,0 %	82,8%	52,7 %	98,5 %	94,6 %
omission error	-	52,4%	-	98,1 %	
commission error	-	11,2%	-	0,0 %	
Detail of delineation	84,8 %	-	74,7 %	95,0 %	90,9 %
Correctness of delineated area	28,0 %	-	15,2 %	35,0 %	61,0 %
Positional accuracy	59,2 %	-	94,6 %	43,0 %	96,7 %

The classification of the Urban Atlas dataset is generally accurate especially in the urban classes. Improvements are required in delineation of polygons, especially roads as well as classification of rural land use, e.g. clear cuts and arable land which are often mixed.

The characteristics and quality of the Street Tree Layer (STL) for Finland (city of Lahti) are unacceptable for national and local applications. A significant share of tree covered areas are omitted from the data, especially small scale and narrow features.

The accuracy of the Riparian Zones status layer varies a lot between different LC/LU themes. Urban and agricultural classes are fairly well recognized but e.g. forest, grassland, and wetlands need major improvements in classification and delineation. In general the classification of the feature layer is too ambitious as the features of different classes cannot be distinguished in satellite images or even in more precise national in-situ data.

The quality of the Green linear elements (GLE) dataset is not acceptable. Large areas of trees and hedgerows are missing from the dataset and the delineation of the GLE-polygons doesn't make sense as they're often located in a forest. The purpose and usability of this dataset is questionable in Finland and is probably more relevant in e.g. central Europe where forests are less abundant.

Delineation of the N2K is fairly accurate and data is mainly not shifted. Delineation of fields is produced with road polygons which sometimes causes errors. More attention should be paid to tree patterns in semi-natural grasslands, wooded pastures and set-asides. It seems like the Finnish national datasets have been directly used to produce some classes e.g. peatlands.

## The usability and potential use of the Copernicus Land Local Component Products in Finland

The user survey was part of the verification of Copernicus local component products. Following products were evaluated:

- Urban Atlas (UA) 2012 status layer
- Riparian Zones LCLU 2012 status layer (RZ) and Green Linear Elements 2012 status layer (GLE)
- Natura 2000 - Grassland LCLU 2012 status layer (N2K)

Significant efforts were made in order to get as many as possible responses from different users about the present use, usability and potential use of the local components. The aim was also to inform users and stakeholders about the availability and information content of the Copernicus services (local components and HR layers). Following methods were applied:

- A WebRopol online survey tool was used to collect responses to the questionnaire. The survey was sent to potential users, who were selected in stakeholder analysis. Following forums were utilized in the analysis: national Copernicus user forum, experts at environmental administration responsible for reporting obligations, visitors of SYKE's Open Data portal, partners in land monitoring of Finland etc.
- SYKE's open data service ([www.syke.fi/openinformation](http://www.syke.fi/openinformation)) is widely used and thus offers a good forum to reach people interested in Land monitoring data and activities. In order to promote Local components and the verification task (and links questionnaire) dedicated [web pages](#) (in Finnish) were published. The Local component products and potential use cases set by Copernicus programme were presented.

- A web-based application for viewing the Copernicus data with relevant background data was developed in-house. The application utilized directly the local component web map services of EEA.
- The local component products, including UA, RZ, and N2K, were made available using tailored ArcGIS projects for internal users of SYKE. Several layers of national data were available and tools to select more GIS layers on screen.
- Information on the local component products were presented to 25 participants in a national seminar September 2017. The programme concentrated on presentations on different Copernicus land monitoring services and their usability.
- The launch of the national verification and usability assessment of the local components was promoted in various media including links to end user survey and Web map applications. The survey and verification process was announced in the SYKE's open data service. Intranet news was repeated several times along the project progress. Notifications in social media were published in Twitter, Facebook and Yammer accounts. Corresponding information were e-mailed to selected user groups, individual experts and participants of the seminar
- User meetings and personal face-2-face interviews were organized. The questionnaire was browsed together with the users and the topics in the questionnaire were discussed and filled together if needed.

Despite of above mentioned efforts only 14 answers were received to the questionnaire. It was difficult to motivate users to allocate their working time and answer the query after they found out that the data to be evaluated could not be used directly in their own work. The preliminary expectations of the usability of EO based data were low, since experts are used to work with national data based on field observations and aerial orthophotos. It was learned that filling the questionnaire and expressing an opinion of the usability of the data would have required more resources (time) than was available. Some users gave their feedback be emails or during interviews, which was more feasible in terms of time needed for the task.

#### Main findings:

- The local component products themselves did not raise spontaneous interest of the users; even though the data sets were promoted and disseminated in different and easily accessible ways and advertised repeatedly.
- We did not find any operative user cases of the local components in Finland. Some users had viewed the data before this study. This may be partly due to the availability of accurate and detailed national data sets in Finland as open data.

#### Potential usability was found and reported, like:

- Urban Atlas data is evaluated to be sufficient in land use monitoring, but it could cover more cities geographically both in Finland and in Nordic countries. Time series is of major benefit. More information on urban green areas was requested.
- The data content and quality of Riparian Zones LCLU data was found unacceptable. However the need for land cover information on riparian zones was recognized. Following factors

hamper the usability of data: It is based on elevation model with low accuracy, unclear descriptions of data production methods and too ambitious nomenclature. Also the delineation of the riparian zone itself received criticism from the users, although it was not included in the questions posed. More simple data that covers also smaller streams/ rivers would serve better. Riparian zone data have potential and should be further developed.

- Natura 2000 data covers only two Natura 2000 sites in Finland. Grassland data is welcomed in the monitoring of the habitat pressures. More habitats should be covered in the future. The habitats themselves cannot be recognized in the data, but the general trends in the area can be monitored. Additionally the data are usable in evaluating the status and change in the surroundings of grassland habitats.

#### Suggestions

- The information content of the local components should be tailored prior to the production/ updating phase according to real user cases together with selected users in different scales (local, national, regional), which user cases are relevant also nationally (for example reporting obligations based on EU directives). These user cases should be reported in a way where usability and benefits are clearly visible.
- The VHR satellite data sets, which are the input for the local components, are valuable and useful output of the Copernicus program. Presently deployment of VHR data is laborious (scene-wise FTP download, clouds, ...). These data should be delivered in a way which is more easily accessible for authorized users. Also cloud-free intermediate products could be produced and delivered, like for example NDVI. These could be a new service in Copernicus land monitoring. There are numerous user organizations in the member states, which have knowledge and facilities to take benefit of this service. This would enable usage of VHR data together with national data sets like digital surface models (DSM) based on airborne Lidar or photos. The homogeneity of the VHR EO coverage is of vital importance in terms of number of instruments used, receiving window and pre-processing

## Corine Land Cover 2018 update

The approach for the CLC2018 update was the same which was applied already in CLC2000, CLC2006 and CLC2012 projects i.e. instead of visual interpretation of satellite images the production of CLC is based on automated processing of existing national GIS data together with satellite images.

National data are first pre-processed into raster format (20 meter) and merged into single raster coverage describing the land cover of whole of Finland with CLC compatible nomenclature. CLC land cover categories which cannot be retrieved from national data sources are mapped using automated satellite image interpretation or digitized manually. The European version of CLC data is produced from high resolution (HR) national land cover with automated generalization procedures. Change mapping is based on the analysis of multi-temporal (2000, 2006, 2012 and 2018) HR national land cover data sets together with image-to-image change detection.

## Organisation of work at national level

The whole of Finland was processed as one working unit. The production included following steps:

### Interpretation of EO data (IMAGE2017)

- Pre-processing of Sentinel 2 data
  - Sentinel 2 A&B data were retrieved and processed at Sodankylä National Satellite Data Centre NSDC. About 60 Sentinel 2 image tiles cover the area of Finland (see Annex 1).
  - In the processing facilities at Sodankylä NSDC Sentinel 2 data are stored and pre-processed online. Sentinel 2 A&B image tiles received spring, summer and autumn seasons 2016 and 2017, totally 3400 individual tile-wise scenes, were processed into Level 2a using SEN2COR.
  - Intermediate cloud masks were generated using Idepix. The most cloud-free scenes (1-8) for each tile, received in given receiving windows, were selected (totally 480). Intermediate cloud masks of these selected scenes were visually checked and manually corrected.
  - Image tiles (both Level 1c and 2a) were mosaicked together into cloud free nationwide data sets in the national projection (EUREF\_FIN\_TM35FIN). Multi-temporal TOA and BOA mosaics were produced: spring and mid-summer (see Annex 2.).
- Production of intermediate themes
  - Normalized Difference Vegetation Index (NDVI)
  - Infrared index (IR)
  - Image segments
- Image-2-image change detection between 2012 and 2017
  - Difference between red and near infrared bands of IMAGE2012 and IMAGE2017 mosaics
  - Classification of changes in terms of type (loss or increase of biomass) and degree of change

### Processing of national spatial data

- Acquisition of relevant national data sets (representing reference year 2016-2017). These include:
  - Agricultural Land Parcel Identification System
  - Topographic database
  - Building and Dwelling Register
  - Road database
  - Multisource National forest inventory
  - Registers on dump and mineral extraction sites and peat production areas
  - Lidar based DSM
- Manual digitizing of specific land cover categories not included in national dbs.
  - Golf courses, dump sites, harbours, open cast mines and large construction sites
- Data integration into HR land cover



- Rasterizing with spatial resolution of 20 meters
- Reclassification according to CLC nomenclature
- Combination with intermediate themes based on Sentinel 2 data and airborne Lidar
- Integration into single coverage of land cover

## Mapping of land cover changes

- Analysis of multi-temporal (2000, 2006, 2012, 2018) land cover data for each CLC 1st level category separately
  - Evaluation of changes between multi-temporal land cover data sets
  - Combination with image-2-image changes
  - Post processing of changes including filtering out small patches and changes due to technical incompatibility of multi-temporal data
- Integration into single coverage of land cover changes
- Generalization
- Conversion into vector format
- Subdivision and labeling of changes using 2012 and 2018 data

## Production of CORINE Land Cover 2018 map

- Combination of HR land cover and land cover changes
- Generalization with MMU of 25 ha including
  - Simplification of the input data
  - Elimination of narrow linear features (roads, rivers) and single pixels
  - Aggregation of the parcels to reach 25 ha inside each level 1 CORINE class
  - Amalgamation of remaining small parcels to the most appropriate neighbouring area according to the priority list
  - Smoothing the boundaries of the parcels
- Vectorizing the raster database

## Internal quality control of results

### QC during data production

During the process the internal quality control was completed for each CLC level 1 categories separately. The present status of land cover and subset of detected changes were checked visually using ortophotos, IMAGE2012 and IMAGE2017, LPIS and topographic maps. A change layer based on multitemporal satellite images was produced and potential land cover changes (based on various input data) were evaluated and verified together with image based change layer.

## Comparison of HR Corine land cover 2018 with national statistics

Forestry causes the majority of land cover changes in Finland (over 90%). National statistics of annual forest management practises (regeneration of forests) were compared with areas of forest changes (3.1.x – 3.2.4) in CLC2018 update. Analysis was completed with total areas of change in whole of Finland using the national HR version of land cover change data (MMU 1 ha). The area of detected forest clear cuttings was found to be 4-19 % smaller than reported in forest statistics. The difference depends on the time interval taken into analysis i.e. 2012-2016 or 2013-2016. Comparison of these data is problematic:

- Applied MMU in national land cover change data (1 ha) causes systematic underestimate into area estimated based on the data.
- The time interval between multi-temporal satellite data sets used in CLC change detection is different in different parts of Finland. Satellite images for IMAGE2012 were received 2012-2013 and images for IMAGE2017 were received 2016-2017.

## LUCAS 2015 micro data and national HR Corine land Cover 2018

The Finnish National high resolution Corine Land Cover 2018 data (HR CLC18) was compared with the micro data measured in the LUCAS 2015 survey (European Land Use/Cover Area Frame Statistical Survey). Both datasets provide information on land cover and land use. The objective was to compare the correspondence of these data at LUCAS field sample plots (see details of the method in Annex 3). A method similar to Maucha et al. 2006 was used in the analysis.

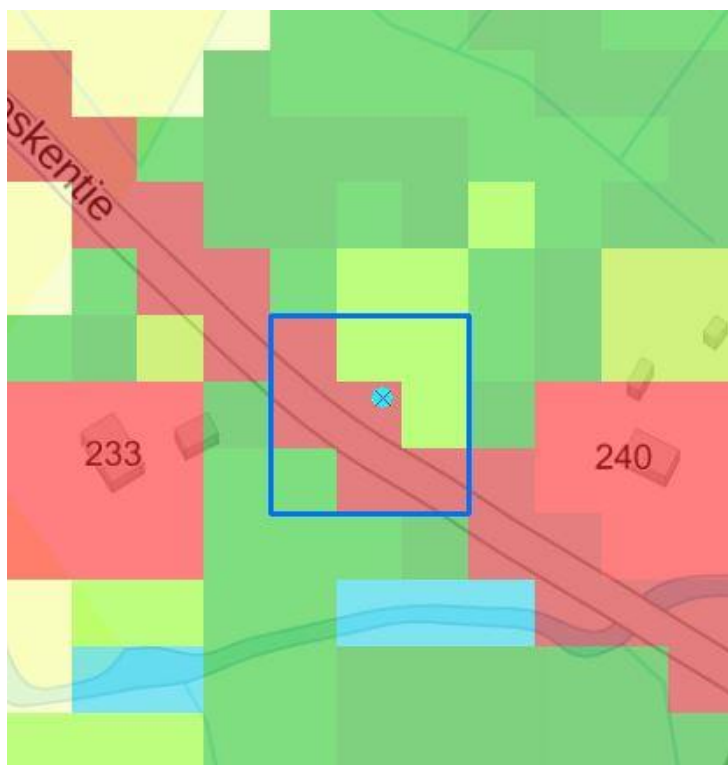
Comparison of LUCAS field survey and HR Corine Land Cover map is problematic mainly due to different nomenclatures, which causes poor correspondence of the data sets. This varies between classes. Challenges with LUCAS nomenclature and data collection at Nordic conditions are reported by Finnish Environment Institute and Natural Resource Institute Finland, who have studied possibilities of providing harmonized land cover / land use information for EUROSTAT using national systems (LUCAS GRANTS 2014 & 2015). Several problems were found in the applicability and practical adoption of LUCAS nomenclature in the field work of LUCAS 2012 in Finland.

Totally 11338 LUCAS 2015 field sample plots, which were measured in the field at the maximum distance of 100 meters from the plot centre were selected to the analysis. Due to the different spatial resolution of LUCAS survey (point observation) and HR CLC2018 (20 meter raster) comparison of these data were made using two approaches. The information content of HR CLC2018 for each LUCAS plot was selected using:

1. the closest pixel or
2. 9 pixels in 3 by 3 window

A positive correspondence between LUCAS and HR CLC2018 was assumed if the closest pixel or any of the 9 pixels matches according to the correspondence table (Annex 3 table 1).

Figure 1. The LUCAS sample point (blue) presented on top of HR CLC2018 map. The point is located in a forest and is assessed as such in the LUCAS survey. Because of the resolution and production method of the HR CLC2018 map (20x20m), the closest pixel of HR CLC2018 is labelled as class 122 i.e. 'Road and rail network and associated land'. A 3 by 3 pixel window is delineated with dark blue line.



When using approach 1 i.e. the match was decided according to the closest pixel the overall percentage of total agreement (PTA) of HR CLC18 data compared to the LUCAS 15 survey was as follows:

- 67 %, if land cover only was a match
- 93 %, if land use only was a match and
- 35 %, if both land cover and use must match

Corresponding figures were calculated also for single CLC classes in 3rd and 1st levels and these results are presented in following table 2. In the 1st level comparison, wetlands and water areas seem to have a better PTA than the rest, when both land cover and land use must match. Especially poor is the classification success for artificial surfaces (52 %). Many CLC classes have only a few LUCAS observations, which must be taken into account when analysing the results. Those CLC classes where  $n > 100$ , the lowest PTA figures are in CLC forests. Highest figures are for "Non-irrigated arable land" (83,9 %), "Transitional woodland and scrub" (88,2 %), "Peatbogs" (88,1 %) and "Lakes" (92,1 %).

Table 2. The results of Comparison using the single pixel approach

CLC Code	Code description	Number of LUCAS plots (n)	LC OK %	LU OK %	Both LC and LC OK i.e. Percentage of total agreement PTA	PTA for the 1st level class
111	Continuous urban fabric	9	66,7 %	88,9 %	66,7 %	
112	Discontinuous urban fabric	190	88,4 %	76,3 %	71,6 %	
121	Industrial or commercial units	76	59,2 %	46,1 %	39,5 %	
122	Road and rail networks and associated land	144	57,6 %	56,9 %	54,2 %	
123	Port areas	0				
124	Airports	2	50,0 %	50,0 %	0,0 %	
131	Mineral extraction sites	21	85,7 %	42,9 %	38,1 %	
132	Dump sites	2	50,0 %	0,0 %	0,0 %	
133	Construction sites	2	0,0 %	0,0 %	0,0 %	
141	Green urban areas	3	100,0 %	33,3 %	33,3 %	
142	Sport and leisure facilities	85	84,7 %	23,5 %	21,2 %	
1	Artificial surfaces	534				51,9 %
211	Non-irrigated arable land	1265	85,4 %	94,5 %	83,9 %	
222	Fruit trees and berry plantations	2	50,0 %	100,0 %	50,0 %	
231	Pastures	5	40,0 %	60,0 %	40,0 %	
243	Agricultural land no longer in use	135	40,7 %	21,5 %	18,5 %	
244	Agro-forestry areas	3	33,3 %	66,7 %	0,0 %	
2	Agricultural areas	1410				77,2 %
311	Broad-leaved forest	332	51,8 %	95,2 %	50,6 %	
312	Coniferous forest	5420	56,3 %	98,3 %	56,0 %	
313	Mixed forest	1812	64,7 %	96,6 %	64,4 %	
321	Natural grassland	0				
322	Moors and heathland	5	40,0 %	60,0 %	20,0 %	
324	Transitional woodland/shrub	1195	93,6 %	89,1 %	88,2 %	
331	Beaches, dunes, and sand plains	1	100,0 %	100,0 %	100,0 %	
332	Bare rock	19	0,0 %	10,5 %	0,0 %	
333	Sparsely vegetated areas	0				
3	Forests and semi-natural areas	8784				61,8 %
411	Inland marshes	55	65,5 %	83,6 %	63,6 %	
412	Peatbogs	219	89,5 %	98,2 %	88,1 %	
421	Salt marshes	8	37,5 %	87,5 %	37,5 %	
4	Wetlands	282				81,9 %
511	Rivers	30	73,3 %	86,7 %	66,7 %	
512	Lakes	277	92,4 %	96,8 %	92,1 %	
523	Sea and ocean	0				
5	Water bodies	307				89,6 %
All		<b>11317</b>	<b>66,9 %</b>	<b>93,4 %</b>	<b>35,5 %</b>	

According to the second approach, where 9 HR CLC2018 pixels were included in the evaluation of the match, the correspondence was as follows:

- 87 %, if land cover only was a match
- 97 %, if land use only was a match
- 84 %, if both land cover and use must match

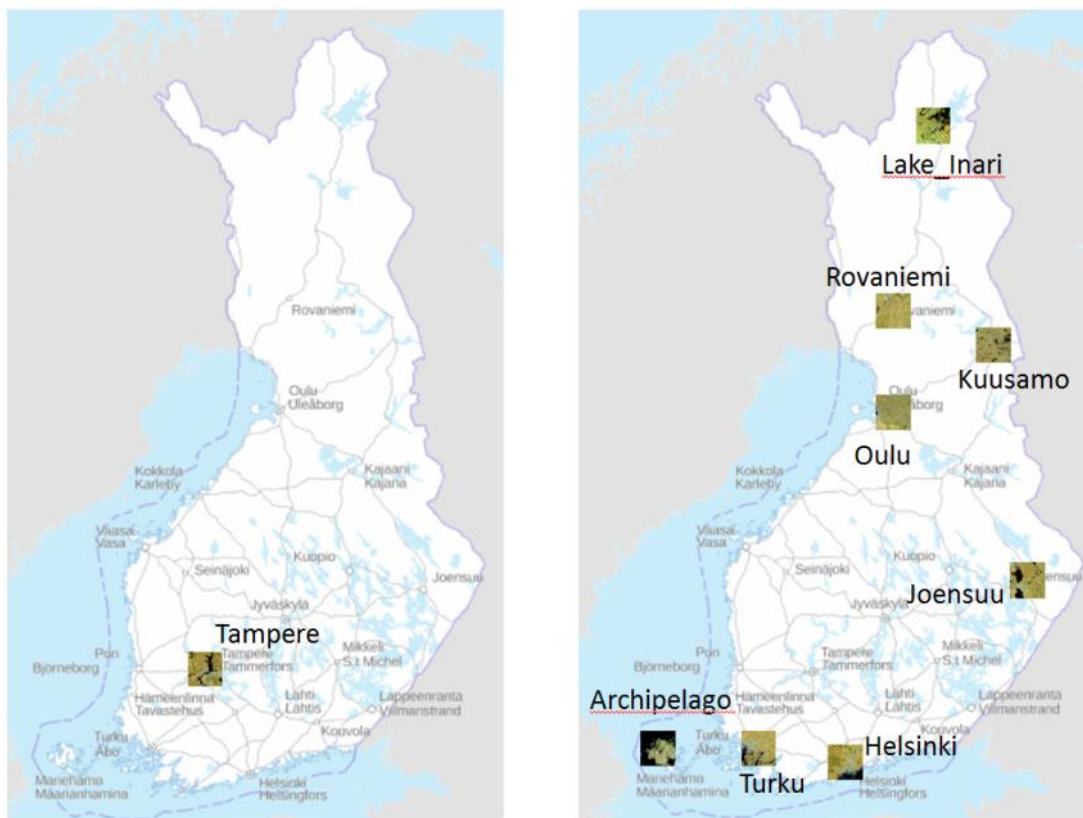
The results of the two approaches are very different (PTA of 35% vs. 84% when both LC and LU must match). The second approach takes into account the different spatial resolutions of the two datasets and is more likely to give reliable results of the overall PTA. This figure (84 %) indicates a fairly positive signal of the accuracy of the HR CLC18 dataset, which is the input data in the automated generalization process for the European version of CLC2018 in Finland.

## External quality control

The external verification by the CLC Technical team from ETC/ULS (technical team) was completed remotely and included following phases:

1. The 1<sup>st</sup> verification was completed in one sample site Tampere (see figure 1.) and included:
  - Evaluation of CLC changes 2012-2018 together with multi-temporal satellite data.
  - These data were Accepted/Conditionally 6.4.2018. For example following improvements were proposed:
    - better mapping of clear cuts (more synchronised with satellite imagery)
    - Improved mapping of changes related to mineral extraction sites and dumpsites
    - Changes between agriculture and forest (211/231-31x)
  - An enhanced version of CLC changes and CLC map 2018 were send to the technical team 27.4.2018.
  - This sample site was accepted by the technical team 22.5.2018
2. For the 2<sup>nd</sup> verification eight (8) sample sites (see figure 2.) were selected by the technical team. Verification was completed 23.7.2018 having proposals for improvements for example in
  - Changes related to artificial surfaces (especially internal changes of artificial surfaces)
  - The area of 243 class is proposed to be reduced by using pure CLC classes, wherever possible
  - Mapping of coastal areas with sparse vegetation.
3. Final steps:
  - Recommendations specified in the 2nd verification report were integrated into the data by the national team.
  - Technical control of the data was done using the online CLC QA/QC tool
  - Datasets were upload CDR (Central Data Repository) by the end of August 2018

Figure 2. Sample sites in the external validation by technical team in phase 1 (left panel) and phase 2 (right panel).



## Main difficulties and solutions

During preparation of CLC2018 update it was assumed that Sen2cor tool will detect clouds in Sentinel 2 data with acceptable accuracy and no or only minor manual correction will be needed for cloud masks. However, the performance of the preprocessing tools (2016 versions) were poor and clouds and their shadows were not detected properly.

**Solution:** cloud-masks were checked and corrected (or totally reinterpreted) manually; which work took lot of labour resources in a very short time frame. Also extra labour resources had to be acquired.

Atmospheric correction implemented in Sen2cor caused anomalies (especially in the water bodies) to the image data.

**Solution:** Image mosaics without atmospheric correction (Level 1C) were also produced. Intermediate data products (spectral indices) were computed using TOA mosaic.

High spectral and spatial resolution of Sentinel satellite data caused challenges in automated change detection based on multi-temporal image data (i.e. between IMAGE2012 and IMAGE2017 mosaics). The aim of change detection was mapping of clear cuts only in forests. However also thinning cuttings cause significant change in red band of Sentinel 2 and thus it was difficult to separate thinning and clear cuts with difference of red bands in forests.

Solution: Sentinel 2 data was resampled into resolution of IMAGE2012 i.e. 20 meters prior to change detection. In the separation of thinning and clear cuts also historical land cover data and reflectance values of Sentinel 2 mosaic were utilized. However some intense thinning cuttings were left into the change data.

## Ancillary data used

In the production and quality control of CLC 2018 update following ancillary data sources were used:

- Agricultural Land Parcel database (Finnish Food Authority)
- National high resolution Corine Land Cover 2000, 2006, 2012 and Corine Land Cover 2012 (Finnish Environment Institute)
- Topographic database (National Board of Survey)
- False colour orthophoto (National Land Survey)
- Multisource National Forest Inventory: Crown coverage-% by main tree species, tree height, site type (Natural Resource Institute)
- Road database (Finnish Transport Infrastructure Agency)
- Airborne lidar (National Board of Survey)
- Building and Dwelling Register (Population Register Centre)
- LUCAS data 2015 (Eurostat)
- Statistics on forest management practices (Natural Resource Institute)
- Environmental registers (Finnish Environment Institute)
- Forest Vegetation zones
- Heat summation

## Post-production verification of the High Resolution Layers for the 2015 reference year

This task was completed together with Natural Resource Institute Finland (LUKE) who was subcontractor for SYKE. LUKE verified the HR forest layers and SYKE HR imperviousness, HR Grassland and HR Water and Wetness.

All proposed verification methods were applied including general overview, look-and-feel and statistical verification. General overview was completed by overlaying each HRL with national HR land cover data using GIS techniques. For example following national data sets were applied:

- National HR CLC2006, 2012 and 2018 land cover data (25 m and 20 m raster)
- Thematic maps of Finnish multisource-NFI (16 m raster)
- Agricultural Land Parcel database (LPIS)
- Topographic Database

Look-and-feel and statistical verification were completed as proposed in the guidelines with the exception of forest layers where field sample plots (n=13729) measured in the national forest inventory were used. Statistical verification was performed also to HRL grasslands, which was not required in the guidelines.

Following ancillary data were also used:

- False colour Aerial Photographs
- IMAGE 2006, 2012 and 2017
- Contour lines
- Biotope Maps of Metsähallitus
- Database of water bodies and riverbeds
- Google map

In the following table 3 the main results of statistical verification are described.

Table 3. Statistical verification and main findings of HRLs 2015 (Corresponding figures in parentheses of HRLs 2012). Accuracy statistics are not fully compatible, since different formulas have been proposed in the guidelines and used in different updates and products.

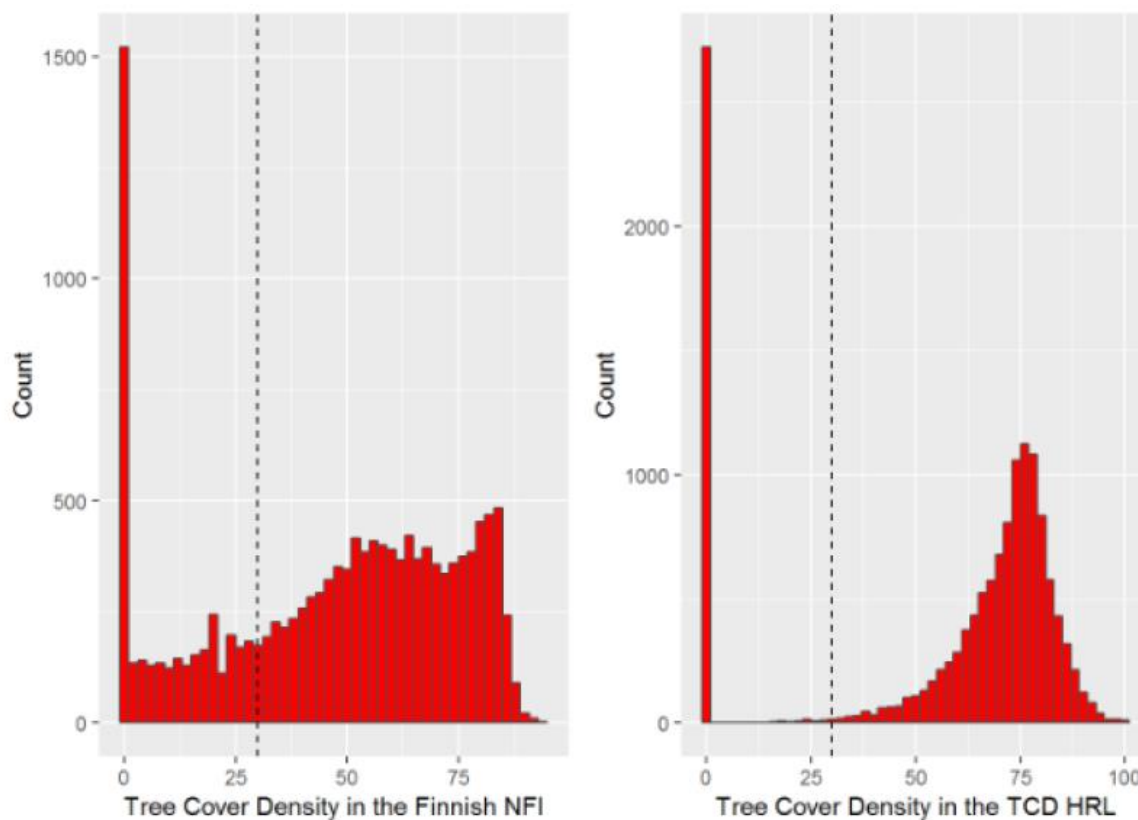
HR layer	Omission error	Commission error	Main findings
Imperviousness	43,4 % (187 %)	25,1 % (6,9 %)	Overall evaluation: good/acceptable; -Discontinuity in roads; -Sport facilities covered by sand or grass included
Tree cover density	6,1 % (10,5 %)	26,2 % (9,5 %)	Overall evaluation: acceptable -Distribution of tree cover density values was not realistic to Finland (overly condensed around value 75 %) -High commission error (26 %) was present on all development stages of forestry land -Areal distribution of tree cover density values showed patchiness that was not explained by our in situ (ground truth) data, and



			systematic over-estimation error that increased northward
Dominant leaf type	No tree cover 6,7 % (11 %) broadleaved 49,5 % (56 %) coniferous: 30,8 % (10 %)	No tree cover 28,9 % (7 %) broadleaved 55,2 % (56 %) coniferous: 6,2 % (17 %)	Overall evaluation: acceptable -Especially in young mixed forests, forests dominated by coniferous species were often erroneously classified as dominated by broadleaved species, and vice versa. Classification accuracy improved as trees matured -Areas with young forest or low tree cover were often erroneously classified as not having any trees, and as a consequence, the HRL layer did not contain information of their leaf type dominance.
Grassland	30,7 % (14 %)	76,8 % (87 %)	Overall evaluation: insufficient -Arable land is often included in the HRL Grassland layer -Clear cut areas and coniferous forest erroneously classified as grassland in HRL Grassland -Pastures are often omitted
Permanent water	22,6 % (1 %)	0 % (8 %)	Overall evaluation of permanent water: acceptable Wet classes in WAW: insufficient -large areas of forests and transitional woodlands erroneously included in permanently of temporary wet areas -large areas of open fells erroneously included in permanent or temporary wet areas

Accuracy statistics presented in the table above do not describe properly the characteristics of Tree Cover Density values in TCD HRL. Distribution of tree cover density values in the high resolution layer does not agree with the canopy cover estimated for the plots in the national forest inventory (see histograms in Figure 3). Especially the left tail of the distribution (values > 0 AND < 25 %) is almost non-existent and values are overly condensed around values 0 % and 75 % in the high resolution layer.

Figure 3. Histogram of the tree cover density value on forestry land, based on estimates for NFI field plots (left) and in the HRL layer (right). Dashed line denotes 30 % threshold used in statistical verification.



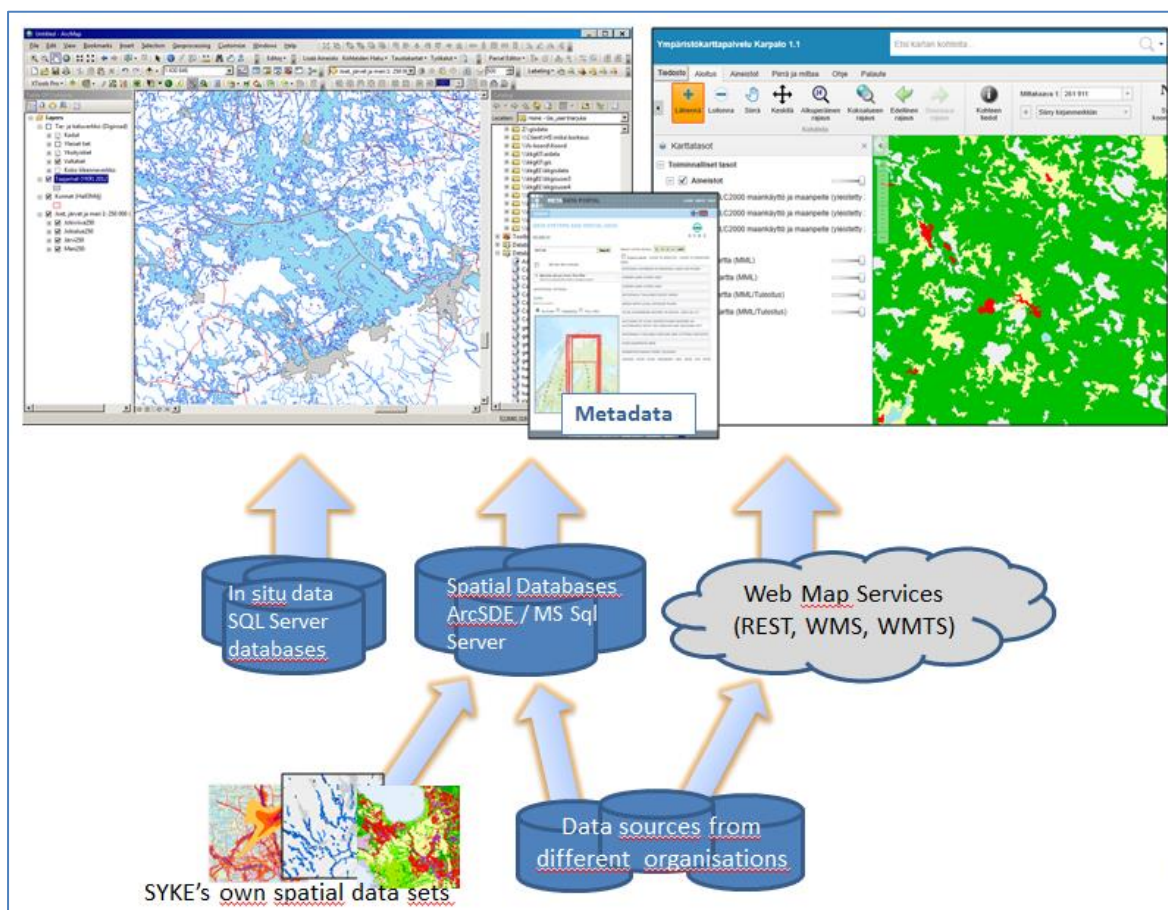
## Dissemination

### Spatial Data Infrastructure of SYKE facilitates efficient data dissemination

SYKE maintains the Environmental Spatial Data Infrastructure (ESDI) with the aim of efficient data dissemination. The ESDI consists of components for data, metadata, technology, human resources and coordination structures. The ESDI is connected with national and European INSPIRE infrastructures. The ESDI has tools and processes for data delivery and compatibility with the INSPIRE requirements can be reached when applicable.

SYKE has adapted open data policy and data sets are put available in the SYKE's open data portal ([www.syke.fi/opendata](http://www.syke.fi/opendata)). Environmental data are accessible by utilizing web services, spatial datasets and satellite observations, as well as data stored in environmental information systems..

Figure 4. Centralized SDI and data delivery system at SYKE



### INSPIRE compatible map services

CORINE Land Cover 2018 products are published according to regulations in INSPIRE directive. Also the INSPIRE compatible metadata for data and services are available. All dataset and services are openly available in SYKE's open data portal ([www.syke.fi/opendata](http://www.syke.fi/opendata)). More detailed information of the services are recorded in the figure 5.

INSPIRE View Services allow users and computer programs to view spatial datasets. The recommended approach to implement INSPIRE view services is the Web Map Service (WMS) 1.3.0. SYKE's WMS on Land Cover data is published using ArcGIS for Server.

There are two types of INSPIRE Download Services that are implemented: pre-defined dataset download service (Atom implementation) and direct access download service (Web Feature Service WFS). SYKE has transformed CORINE Land Cover 2018 vector –dataset into the INSPIRE Land Cover vector application schema. FME Workbench was used for creating the ETL process. The GML file according to INSPIRE application scheme is downloadable in Atom feed. INSPIRE Download services

are also implemented with Web Feature Service WFS using technology provided by HALE and GeoServer.

Figure 5. INSPIRE compatible data, metadata and services

<p>INSPIRE compatible WMS-service for the Corine Land Cover 2018 : <a href="http://paikkatieto.ymparisto.fi/ArcGIS/services/INSPIRE/SYKE_Maanpeite/MapServer/WMSServer">http://paikkatieto.ymparisto.fi/ArcGIS/services/INSPIRE/SYKE_Maanpeite/MapServer/WMSServer</a></p> <p>INSPIRE compatible data download provided by ATOM-feed and GML –file for Corine Land Cover 2018: <a href="http://www.d3.ymparisto.fi/d3/INSPIREAtom/LC_Corine2018_25ha.xml">http://www.d3.ymparisto.fi/d3/INSPIREAtom/LC_Corine2018_25ha.xml</a></p> <p>INSPIRE compatible download service WFS for Corine Land Cover 2018: <a href="http://geoserver.ymparisto.fi/geoserver/wfs?service=wfs&amp;version=2.0.0&amp;request=GetCapabilities">http://geoserver.ymparisto.fi/geoserver/wfs?service=wfs&amp;version=2.0.0&amp;request=GetCapabilities</a></p> <p>INSPIRE compatible metadata (in English) in XML format of Corine Land Cover 2018 dataset : <a href="http://metatieto.ymparisto.fi:8080/geoportal/rest/document?id=%7B26EEEEBBB-FB5C-4045-B6DF-439F9B7D5C46%7D">http://metatieto.ymparisto.fi:8080/geoportal/rest/document?id=%7B26EEEEBBB-FB5C-4045-B6DF-439F9B7D5C46%7D</a></p> <p>The metadata of INSPIRE compatible ATOM feed (in Finnish) : <a href="http://www.paikkatietohakemisto.fi/geonetwork/srv/eng/catalog.search;jsessionid=1o0kvmsoxqip61fh6cx8ui2q3e#/metadata/%7B615B1B36-3140-4A6A-BFAF-1BF94F8F8D6B%7D">http://www.paikkatietohakemisto.fi/geonetwork/srv/eng/catalog.search;jsessionid=1o0kvmsoxqip61fh6cx8ui2q3e#/metadata/%7B615B1B36-3140-4A6A-BFAF-1BF94F8F8D6B%7D</a></p> <p>All versions of CORINE Land Cover products from years 2000, 2006, 2012 and 2018 are also available for download “as is” -version using ATOM feed and EsriShape/geotiff: <a href="http://www.d3.ymparisto.fi/d3/atom/inspireatom.xml">http://www.d3.ymparisto.fi/d3/atom/inspireatom.xml</a></p>
---

## Conclusions

According to experiences gained in the verification of HRL and the Local component products usability and present use of these Copernicus services are low in Finland. When comparing HRLs year 2012 and 2015 it is visible that the quality of these data is improving. This may be the case especially in coming 2018 updates, when time series of Sentinel data will be available as input.

There are significant overlaps between Copernicus services and national monitoring programmes. For example, EO based thematic forest maps are produced operationally as open data in Finnish National Forest Inventory, whose information content is about the same as Copernicus HRLs Tree Cover Density (TCD) and Dominant Leaf Type (DLT). In order to improve the relation between data quality and production costs these efforts should be joined.

In Finland, nationally feasible data with high spatial resolution (20 m pixel) are produced parallel to European Corine data (MMU 25 ha) and Finnish High Resolution Corine Land Cover classification has been widely used nationally. Best available national data and local expertise have been available for this work. This supports usage of bottom-up approach when possible in production of Copernicus services.

In the near future there will be available wall-to-wall information based on airborne lidar, which data will be regularly updated. This will cause fundamental improvement in the information content, thematic accuracy and spatial resolution of land cover information; especially when combined with time-series of EO data. This emphasises the importance of easy access to satellite data provided in the Copernicus programme, including VHR data.

## References

Büttner, G. & Maucha, G. 2006: The thematic accuracy of Corine land cover 2000 – Assessment using LUCAS (land use/cover area frame statistical survey). EEA Technical report, No. 7/2006.

Härmä Pekka, Suvi Hatunen, Markus Törmä, Elise Järvenpää, Minna Kallio, Riitta Teiniranta, Tiia Kiiski, Jaakko Suikkanen. GIO Land Monitoring 2011 – 2013 in the framework of regulation (EU) No 911/2010; Pan-EU Component; Grant Agreement 3541/B2012/R0-GIO/EEA.55037; Final Report; Finland;  
<https://www.syke.fi/download/noname/%7B2D87A112-C054-4A79-BEC2-41F00974EBC0%7D/107966>

Härmä, P.; Teiniranta, R.; Törmä, M.; Repo, R.; Järvenpää, E.; Kallio, M. CLC2000 Finland - Final Report. Finnish Environment Institute (SYKE), Geoinformatics and Land Use Division (GEO), May 2005.  
<http://www.syke.fi/download/noname/%7BB4231601-439E-4903-B6FC-2F2FE718B465%7D/39571>

Härmä, P., Teiniranta, R., Törmä, M., Repo, R., Järvenpää, E., Kallio, M. The production of Finnish CORINE land cover 2000 classification. Geo-Imagery Bridging Continents. XXth ISPRS Congress, 12-23 July 2004 Istanbul, Turkey, In: Geo-Imagery Bridging Continents. XXth ISPRS Congress, 12-23 July 2004 Istanbul, Turkey Commission IV. Vol. XXXV, part B4. Istanbul, International Society for Photogrammetry and Remote Sensing and Spatial Information Sciences (ISPRS). 6 p. 2004.  
<http://www.isprs.org/proceedings/XXXV/congress/comm4/papers/548.pdf>

Härmä P., Järvenpää E., Kallio M., Teiniranta R., Törmä M., Hatunen S. CLC2006 Finland - Final technical report. Finnish Environment Institute, November 2009.  
<http://www.syke.fi/download/noname/%7BC7C849EB-3F4D-42AE-9A94-5B8069FFDFFB%7D/37641>

LUCAS GRANT 2014: Provision of harmonized land cover / land use information. LUCAS and national systems. EUROSTAT/Contract No.: 08441.2014.003-2014.756  
[https://circabc.europa.eu/sd/a/9fde551b-239a-4e66-a718-27d277a3e2e2/FI\\_LUCAS\\_Grant2014\\_MethodologicalReport\\_OBJ1\\_Final.pdf](https://circabc.europa.eu/sd/a/9fde551b-239a-4e66-a718-27d277a3e2e2/FI_LUCAS_Grant2014_MethodologicalReport_OBJ1_Final.pdf)

[https://circabc.europa.eu/sd/a/629cb5b5-10dc-4733-9aa1-1cf309282fdb/FI\\_LUCAS\\_Grant2014\\_MethodologicalReport\\_OBJ2.pdf](https://circabc.europa.eu/sd/a/629cb5b5-10dc-4733-9aa1-1cf309282fdb/FI_LUCAS_Grant2014_MethodologicalReport_OBJ2.pdf)

LUCAS GRANT 2015: Provision of harmonized land cover / land use information. LUCAS and national systems. EUROSTAT/Contract No.: 08441.2015.002-2015.723

[https://circabc.europa.eu/sd/a/713ac40a-b56c-48c1-a374-5cae848e3f38/FI\\_LUCAS\\_Grant\\_2015\\_VII\\_a\\_Final\\_MethodologicalReport.pdf](https://circabc.europa.eu/sd/a/713ac40a-b56c-48c1-a374-5cae848e3f38/FI_LUCAS_Grant_2015_VII_a_Final_MethodologicalReport.pdf)

Törmä, M., Härmä, P., Hatunen, S., Teiniranta, R., Kallio, M., Järvenpää, E. Change detection for Finnish CORINE land cover classification. In Earth Resources and Environmental Remote Sensing/GIS Applications II, edited by Ulrich Michel, Daniel L. Civco, Proceedings of SPIE Vol. 8181 (SPIE, Bellingham, WA 2011) 81810Q.

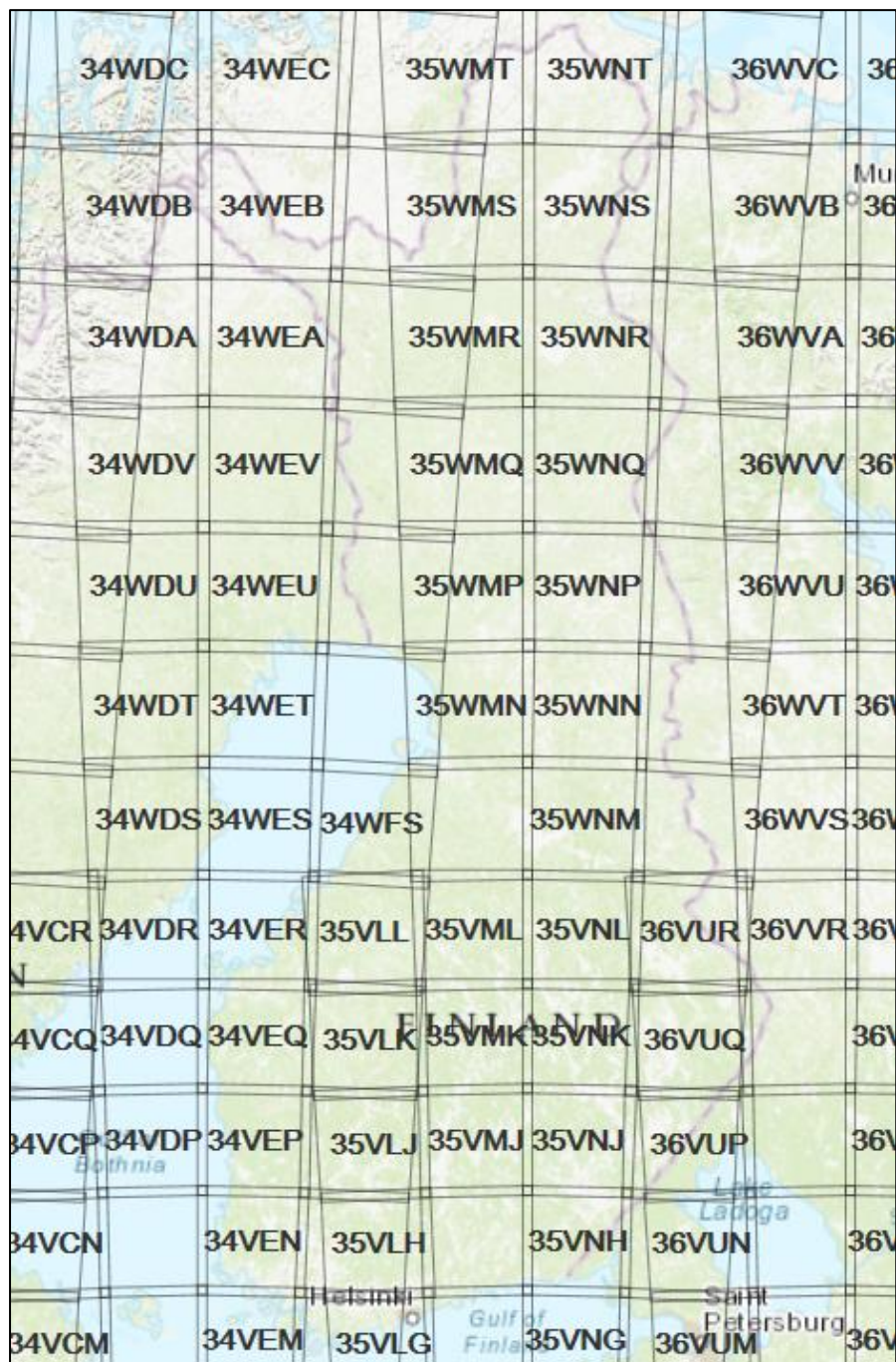
Done at Finnish Environment Institute, Helsinki

Date: 22.2.2019

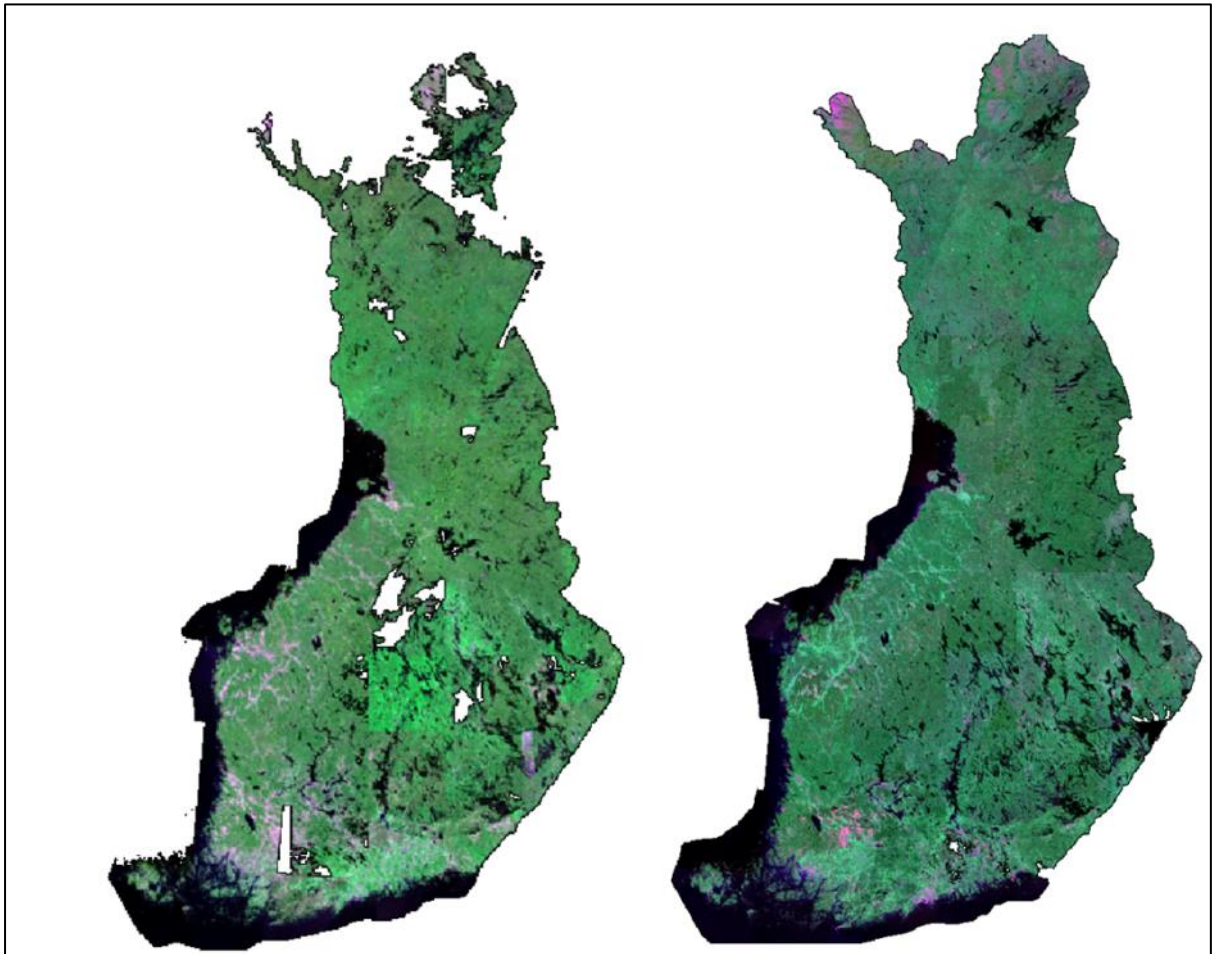
Name: Pekka Härmä

Signature:

Annex 1. Sentinel 2 tiling grid covering Finland



Annex 2. Sentinel 2 a+b mosaics (bands 03, 04 and 08), spring on the left and summer on the right.





### Annex 3. Comparison of LUCAS micro data and HR Corine Land cover

The LUCAS survey is based on sample plots located in a 2 km grid. In the LUCAS 2015 survey, observations were collected from 16107 sample points in Finland. Plots which could not be accessed in-situ were estimated from a distance or by photo-interpretation. Only the LUCAS plots (totally 11338) were selected to this analysis, where the surveyor had visited at the maximum distance of 100 meters.

A Correspondence table (see Table 1 - Annex 3) was produced in order to determine which LUCAS2015 LC/LU classes correspond (match) to each HR CLC18 class on 3rd level. Some adaptations to Finnish particularities of CLC LC/LU classes were necessary. For example, in the European CLC dataset, class 243 refers to “a mosaic of agricultural and natural areas” but in the national dataset this is “agricultural land no longer in use”. The translation of CLC classes into LUCAS class combinations is not straightforward. E.g. the definition of forest in the LUCAS nomenclature is a woodland with tree crown cover of >10% at maturity. In CLC, forest is defined as having apparent crown cover of >30%. Transitional woodlands (young forest /clear cuts) are missing in LUCAS. Some classes, such as Discontinuous urban fabric (1.1.2) in CLC is a mix of land cover and use and can include several LUCAS LC and LU classes.

GIS overlay was performed to combine each LUCAS sample point with the HR CLC18 value at the same location. The data was validated based on the correspondence table that gives each HR CLC18 class a set of acceptable LUCAS LU and LC classes. A sample point was assigned as a match (value 1) if HR CLC18 code corresponded to one or more LUCAS LU and LC codes in the correspondence table. This way a percentage of total agreement (PTA) could be obtained (Maucha et al 2006). The validation was automated with a python script.

The resolution differences between the two datasets causes challenges to this approach. If a LUCAS sample point is located at the edge of two LC/LU classes, the HR CLC18 class can be validated as mismatch if the neighbouring class extends to the location of the sample point. This problem is especially evident with narrow LC/LU classes such as roads. To tackle this issue, another approach was also implemented to the validation. Here a buffer area of 8 HR CLC18 pixels was also checked around the pixel where the LUCAS-sample point landed. HR CLC18 was validated as a match if any of the HR CLC18 classes within the buffer area corresponds with the LC/LU class of the LUCAS15 sample point. An example of such situation is demonstrated on Figure 1.

Annex 3 table 1. Correspondence table between HR CLC18 and LUCAS2015 classes.

CLC code	CLC category	LUCAS land cover	LUCAS land use
111	Continuous urban fabric	A12 A11 A21 A22 A30 E10 E20	U370 U350 U361 U341 U342 U311 U312 U317
112	Discontinuous urban fabric	A11 A12 A21 A22 A30 B43 B44 C10 C21 C22 C23 C31 C32 C33 E10 E20	U350 U370 U341 U342 U311 U312 U317 U113 U361 U362
121	Industrial or commercial units	A12 A11 A21 A22 A30 E30 E20	U210 U221 U222 U223 U224 U225 U226 U227 U341 U342 U350 U311 U312 U317 U319 U321
122	Road and rail network and associated land	A22 A21 A30 E10 E20 F40	U311 U312
123	Port areas	A21 A22 G11 G21 G30 A11 A12 F40	U313 U361 U362
124	Airports	A21 A22 E10 E20 A11 A12 F40	U314
131	Mineral extraction sites	A22 G11 F20 F10 F40 C10 C21 C22 C23 C31 C32 C33	U140
132	Dump sites	A21 A22 A30	U322
133	Construction sites	A11 A21 A22 F10 F20 F30 F40 E20	U330
141	Green urban areas	A22 C10 C21 C22 C23 C31 C32 C33 E10 E20 G11	U361
142	Sport and leisure facilities	A11 A21 A22 C10 C21 C22 C23 C31 C32 C33 E10 E20 G11	U362 U361 U350
211	Non-irrigated arable land	A13 A21 A22 B11 B12 B13 B14 B15 B18 B21 B22 B23 B31 B32 B35 B37 B41 B43 B44 B45 B51 B53 B54 B55 B83 E30 F40	U111 U112
222	Fruit trees and berry plantations	B71 B75 B84 A22 E20	U111
231	Pastures	B51 B52 B53 B54 B55 E10 E20 E30	U111 U410
243	Abandoned agricultural land	E30 E20 E10	U410 U420 U112
244	Agro-forestry areas	A22 A21 B11 B12 B13 B14 B15 B18 B51 B53 B54 C10 C21 C22 C23 C31 C32 C33 E10 E20	U111
311	Broad-leaved forest	A22 C10	U120 U410 U420 U312
312	Coniferous forest	A22 C21 C22 C23	U120 U410 U420 U312
313	Mixed forest	A22 C31 C32 C33	U120 U410 U420 U312
321	Natural grasslands	E10 E20 E30	U420 U410
322	Moors and heathland	D10 D20 F30	U420 U410
324	Transitional woodland shrub	B83 C10 C21 C22 C23 C31 C32 C33 D10 D20 E10 E20 E30 H12 F40	U120 U410 U420
331	Beaches, dunes and sand plains	F20	U420 U410
332	Bare rocks	F10	U420 U410
333	Sparsely vegetated areas	F40 F30 F10 D10 D20 E10 E20	U420 U410
411	Inland marshes	H11 E20 D20 G11	U420 U410 U321 U111 U361
412	Peat bogs	H12 C10 C21 C22 C23 C31 C32 C33	U140 U410 U420 U120
421	Salt marshes	H21 E20 D20 G30	U410 U420
511	Water courses	G21 G30	U313 U410 U420 U130 U361
512	Water bodies	G11	U313 U410 U420 U130 U361

