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. 58139

Implementing Framework service contract No

EEA/IDM/R0/16/009/Finland

Finland

Pekka Härmä, Suvi Hatunen, Kaisu Harju, Minna Kallio, Markus Törmä

Finnish Environment Institute, Data and Information Services

Final Report

Background

Finnish Environment Institute (SYKE) signed a Framework Service Contract for the Copernicus Land Monitoring services with EEA 16th of February 2017. SYKE as the main-contractor and Natural Resources Institute Finland (LUKE) as a sub-contractor has provided EEA a series of mapping, verification, enrichment etc. activities related to pan-European and local component land monitoring products.

Totally 3 specific contracts within this Framework contract were agreed and completed between 2018-2021. Specific contacts 1 and 2 included the following tasks:

- Verification of 2012 reference year local component products, including
  - Urban Atlas and street tree layer
  - Riparian zones
  - Green linear elements
  - Natura2000
- Update of Corine Land Cover including
  - Land cover map for the 2018 reference year
  - Detection of Corine land cover changes between 2012-2018
- Post-production verification of the High Resolution Layers (HRL's) for the 2015 reference year was completed together with LUKE (Natural resources Institute Finland), who was responsible for the verification of forest layers. This task included verification of
  - Imperviousness
  - Forest layers
    - Tree cover density
    - Dominant leaf type
  - Grassland
  - Permanent water
- Enrichment of Urban Atlas
- Inventory of Land Use (LU) information and other landscape characteristics (CH)
- Dissemination

This document is the final report of the 3rd specific contract, which originally included following tasks:

- Post-production verification of the High Resolution Layers for the 2018 reference year
- Support and testing of future CLC+ with methodological improvements, based on CLC2018 products
  - Verification of CLC+ Backbone
  - Refinement of land use (LU) inventory, Establishment of Ingestion Rulesets and Establishment of Extraction Rulesets

There was a significant delay in the data production of CLC+ Backbone data and only part of the service i.e. a raster version of the data was available for national validation during the last month of the project, in early December 2021. It was not possible to complete this task in its entirety within one month and thus this task was not completed in Finland.

Additionally, establishment of ingestion and extraction rulesets were not completed. With the agreement of EEA these tasks were replaced and resources reallocated into provision of wall-2-wall data on LULUCF LUA layer based on national, high resolution data.

Post-production verification of the HRLs for the 2018 reference year and refinement of land use (LU) inventory were completed as planned.

The completion of these tasks in Finland is described below.

Task 1.1: Post-production verification of the High Resolution Layers (HRLs) for the 2018 reference year

Verification of the High Resolution Layers (HRLs) for the 2018 reference year consisted of the verification of following HRLs:

- Degree of Imperviousness status 2018 (IMD)
    - General overview, look-and-feel verification and quantitative verification were performed.
    - Link to report: https://www.syke.fi/download/noname/%7B9DA3B29C-DDC2-41F8-954B-37FE6A053272%7D/172400
- Degree of Imperviousness change 2015-2018 (IMCC)
    - General overview, look-and-feel verification and quantitative verification were performed.
    - Link to report: https://www.syke.fi/download/noname/%7B2CB9EB07-6096-4BB4-8390-C672BA5EC948%7D/172399
- Tree Cover Density status 2018 (TCD)
    - General overview, look-and-feel verification and quantitative verification were performed.
    - Link to report: https://www.syke.fi/download/noname/%7B38A4547F-EAA1-4A86-B1CC-7EF8895463D5%7D/172402
- Dominant Leaf Type status 2018 (DLT)
    - General overview, look-and-feel verification and quantitative verification were performed.
    - Link to report: https://www.syke.fi/download/noname/%7B2FC7B826-771B-4844-A1FA-AE5741E16536%7D/172394
- Dominant Leaf Type Change 2015-2018 (DLTC)
    - General overview, look-and-feel verification and quantitative verification were performed.
    - Link to report: https://www.syke.fi/download/noname/%7B6448D370-F3A6-437E-B4B5-601AC8D01F64%7D/172395
- Grassland status 2018 (GRA)
    - General overview, look-and-feel verification and quantitative verification were performed.
    - Link to report: https://www.syke.fi/download/noname/%7B9FC82506-23AF-4A77-A5A6-B588EF794DD0%7D/172396
- Grassland change 2015-2018 (GRAC)

    - General overview, look-and-feel verification and quantitative verification were performed.
    - Link to report: https://www.syke.fi/download/noname/%7BB170D82D-002E-46B0-BFC1-57926534C263%7D/172397
- Water & Wetness status 2018 (WAW)
    - General overview and look-and-feel verification were performed.
    - Link to report: https://www.syke.fi/download/noname/%7B35BB612D-6A53-4EA3-A97F-70A7659E2FA8%7D/172398
- Small Woody Features 2015 (SWF)
    - General overview and look-and-feel verification were performed.
    - Link to report: https://www.syke.fi/download/noname/%7B9DC8F795-176B-48B4-BB10-BFE962CE68FC%7D/172401

All proposed verification methods were applied including general overview, look-and-feel and statistical verification. The verification task was completed together with Natural Resource Institute Finland (LUKE) who was subcontractor for SYKE. LUKE was responsible for verifying the HR forest layers and SYKE for rest of the HR layers.

General overview was completed by overlaying each HRL with national HR land cover data using GIS techniques. This comparison between HRL and national data shows possible differences and indicates typical and significant omission and commission errors. The following national data sets were applied:

- National High Resolution Corine Land Cover 2018 data (20x20m)
- Finnish Multisource National Forest Inventory 2017 of Natural Resources Finland
- Topographic Database of National Land Survey
- Land Parcel Identification System
- National High Resolution Corine Land Cover changes 2012 – 2018
- National Forestry Statistics

Also, overall areal statistics of HRL and country overview map were provided.

The look-and-feel verification of HRLs were completed as proposed in the guidelines, meaning that the strata of potential omission and commission errors identified in the guidelines were visually studied using aerial and satellite imagery and national spatial data. There were typically over 10 samples per stratum, the number of samples varying from 3 to 50. There were some strata that were not studied because they either do not exist in Finland (e.g. Mediterranean shrublands, Sclerophyllous or Dehesa/Montado forests) or their area is very small in Finland (e.g. forest damage, solar panel parks). In the case of HRL change products, the classes to be verified were proposed in the guidelines. The raster data of selected classes were vectorized, polygons sorted from the largest to the smallest and 100 largest polygons were studied visually.

The quantitative verification of HRLs were completed as proposed in the guidelines, with the exception of forests, where field sample plots measured in the National Forest Inventory were used as ground truth. The verification was performed to all HRL-products, except Water & Wetness status

2018 (WAW) and Small Woody Features status 2015 (SWF). Statistical sampling was usually performed by making systematic sampling with 100 meter interval in East- and North-directions and then performing random sampling (Matlab, rand-function) from this set of locations; resulting 300 points per evaluated class. Non-HRL area was stratified using national GIS data in order to concentrate on areas where omission errors were more likely. In other words, areas where national in-situ GIS data indicates to be part of HRL but is non-HRL in the HRL 2018 data. Some sample points were excluded due to positional shifts in the data. The final set of sample points were checked against relevant in-situ data sets (e.g. aerial or satellite imagery, Corine Land Cover 2018/2012 classifications, LPIS, Topographic Database) and assessed as correct or incorrect. In case of incorrect, the correct class was given. There were some exceptions concerning sampling:

- Degree of Imperviousness change 2015-2018 (IMCC): Stratification was based on the area of IMCC-classes and availability of 2015 aerial images. Systematic sampling distance for IMCC class 2 was 40 meter.
- Tree Cover Density status 2018 (TCD): 13496 field sample plots of National Forest Inventory measured years 2017-2019, except 2012-2013 in Northern Finland.
- Dominant Leaf Type status 2018 (DLT): 13496 field sample plots of National Forest Inventory measured years 2017-2019, except 2012-2013 in Northern Finland.
- Dominant Leaf Type Change 2015-2018 (DLTC): The field sample consists of permanent field plot pairs 2012-2014 and 2017-2019 measured in National Forest Inventory; except for sampling region of Ahvenanmaa island (plots visited 2013 and 2018) and excluding Northern Lapland. The data set contained 7777 pairs of permanent field sample plots.
- Grassland change 2015-2018 (GRAC): Stratification was based on the area of GRAC-classes and availability of year 2015 and 2016 aerial images.

The final statistical accuracy values were computed using the Map Accuracy Tool.

Table 1. Main findings and the results of quantitative verification of HRLs 2018.

HR layer Accuracy values	Main findings
Degree of Imperviousness status 2018 (IMD) Overall accuracy: 60.8% Classwise accuracies (Producer's / User's): • Non-impervious: 64.7% / 80.6% • Low imperviousness: 41.9% / 28.2% • High imperviousness: 65.0% / 56.3%	Overall: Acceptable Comments / Problems: • IMD has typically identified built-up area as impervious quite accurately. No major omission errors in urban and suburban housing areas, industrial and commercial areas nor traffic areas (airports, parking lots, harbours) • Major mistakes are found in unvegetated areas where the ground reflectance is high. Commission errors in unvegetated agricultural fields, natural sand areas and sand pits, bare rocks and dump sites which are erroneously classified as impervious. • Small features are often missing from the data.

Degree of Imperviousness change 2015-2018 (IMCC) Sample-based probability of correct classification: • IMCC class 1 - New impervious cover: 26.0% • IMCC class 2 - Loss of impervious cover: 1.7%	Overall: Insufficient Comments / Problems: • New impervious surface: the largest mistakes consist of industrial or storage areas with no change or non-impervious areas like clear-cuts in forest • Loss of impervious surface: errors are due to misclassified impervious surfaces year 2015 (these kinds of changes are very rare in Finland)
Tree Cover Density status 2018 (TCD) Overall accuracy: 88.6% Classwise accuracies (Producer's / User's): • 0-29%: 71.9% / 84.1% • 30-100%: 94.9% / 90.0%	Overall: Acceptable Comments / Problems: • TCD has good spatial consistency in the detailed forest stand structure and tree cover was situated on plausible land use/land cover classes. • HRL overestimated the area of no tree cover. • Above-zero tree cover densities are overestimated, particularly in Northern Finland.
Dominant Leaf Type status 2018 (DLT) Overall accuracy: 74.6% Classwise accuracies (Producer's / User's): • Non-forest: 93.8% / 20.9% • Broadleaved: 52.7% / 59.6% • Coniferous: 77.5% / 96.6%	Overall: Acceptable Comments / Problems: • DLT has good spatial consistency in the detailed forest stand structure, and leaf type dominance of the tree cover was in line with what was expected for different land cover types. • Detected dominance of coniferous leaf types better than dominance of the broadleaved. • Classification accuracy improved as the forest matured, especially for coniferous-dominated areas.
Dominant Leaf Type Change 2015-2018 (DLTC) Overall accuracy: 4.5% Classwise accuracies (Producer's / User's): • Unchanged areas with no tree cover: 0.0% / 0.0% • New coniferous cover: 0.0% / 0.0% • Loss of Broadleaved cover: 100.0% / 10.0% • Loss of coniferous cover: 100.0% / 3.6% • Unchanged areas with tree cover: 0.0% / 0.0% • Potential change among dom. leaf: 0.0% / 0.0%	Overall: Insufficient Comments / Problems: • DLTC underestimates area of New tree cover, thus implying decline in forest area (Total area for loss of tree cover >> Total area for New tree cover according to DLTC). This conflicts with national data. • There are many detected changes in DLTC which are unchanged according to field measurements. • Very low tree cover density threshold used by the HRL (1%) is very difficult to verify.
Grassland status 2018 (GRA) Overall accuracy: 38.9% Classwise accuracies (Producer's / User's): • Non-grassland: 10.0% / 22.0% • Grassland: 66.3% / 43.7%	Overall: Insufficient Comments / Problems: • Natural grasslands are very rare in Finland, most are agricultural grasslands • Croplands, peatbogs and clear-cut areas frequently and incorrectly classified as grassland • Some managed and semi-managed grasslands were classified as non-grassland due possible errors in the ploughing indicator

	• Grasslands in urban areas seem to be classified with an acceptable accuracy • There are large areas in Lapland, that are classified as grassland and contain mixtures of mountain birch, brush, heath etc...
Grassland change 2015-2018 (GRAC) Sample-based probability of correct classification: • GRAC class 1 - Grassland gain: 16.2% • GRAC class 2 - Grassland loss: 7.6% • GRAC class 11 - Unverified grassland gain: 4.0% • GRAC class 22 - Unverified grassland loss: 3.7%	Overall: Insufficient Comments / Problems: • Natural grasslands are very rare in Finland, most are agricultural grasslands • Grassland gain: the largest errors are due to that area is not grassland 2018 or is grassland at both times (i.e., no change at all), or is forest clear-cut, open wetland or peat production. • Grassland loss: area is no grassland 2015 or grassland or no grassland at both times (i.e., no change at all), increase of grassland or forest clear-cut. • Unverified grassland gain: these are typically very large polygons at Lapland (mixtures of mountain birch, brush, heath etc...), wetlands, peat production sites or forest clear-cuts • Unverified grassland loss: very large agricultural areas with minor part of grassland, very large polygons at Lapland or forest clear-cuts
Water & Wetness status 2018 (WAW) No quantitative verification	Overall: Permanent water acceptable, others insufficient Comments / Problems: • Permanent Water: areas seem to be correctly classified but under-estimate the total water area • Narrow water areas and near shores or areas with emergent vegetation are missing, which are water or (floating) vegetation • There are erroneous dry narrow areas beside the seashores, which are water or wet areas according to national data • Many dry bogs are classified as Permanently Wet • Many high-altitude areas in Lapland are misclassified as Temporary Wet
Small Woody Features status 2015 (SWF) No quantitative verification	Overall: Insufficient Comments / problems: • In Finland the idea of SWF data may be feasible in the landscapes with intensive agriculture, where small-size forest patches are part of the landscape. Inside forest dominated landscapes (70% in Finland) the idea of SWF is questionable. • Agricultural areas: SWF data succeeds to catch the small woody parts inside the field areas • Natural grasslands, fell areas & Wetlands: SWF data is mostly incorrect in these areas

Comparison of classification accuracies of several HRL classes of HRL 2015 and 2018 is presented in table 2. Both producer's and user's accuracies are higher in HRL 2018 than in HRL 2015 for classes TCD: Tree cover, DLT: Coniferous and DLT: Broadleaf. Otherwise, when one accuracy measure of class increases, the other decreases. The largest drop of classification accuracy happens with class DLT: No tree cover, in which case the user's accuracy decreases from 71.1% to 20.9%.

Table 2. Comparison of producer's and user's accuracies of several HRL classes of HRL 2015 and 2018.

HRL: Class	2018 producer's accuracy	2015 producer's accuracy	2018 user's accuracy	2015 user´s accuracy
IMP: built up	**65.0%**	56.6%	56.3%	**74.9%**
GRA: Grassland	66.3%	**69.3%**	**43.7%**	23.2%
DLT: No tree cover	**93.8%**	93.3%	20.9%	**71.1%**
DLT: Broadleaf	**52.7%**	50.5%	**59.6%**	44.8%
DLT: Coniferous	**77.5%**	69.2%	**96.6%**	93.8%
TCD: Tree cover	**94.9%**	93.9%	**90.0%**	73.8%

Producer's accuracy is the proportion of samples of certain class of the reference data which are classified to belong to that class.

User's accuracy is the proportion of samples of certain class of the classification result which are classified to belong to that class.

Task 1.3 – Land Use related tasks

Subtask 1 – Refinement of LU inventory

In the previous Specific Contract #2 information about available national data sources were produced by screening existing national/EU level spatial data portals and national repositories. The previous excel table contained several information sources for the same class, and the datasets were prioritized according to their geometric resolution and suitability for the class. Many of these national data sources listed in the table were also used in the production of HR Corine Land Cover (CLC) 2018 national raster, and the national CLC 4th level classes were also listed. The input data sources and their restrictions were described in the associated documentation.

In this Specific Contract LU inventory tables were updated and national LU data sources were separated into three groups as follows:

- High Resolution Corine Land Cover 2018 that was used in the production of simplified LULUCF LUA layer (described below)
- LU datasets that are used as input in the HR Corine Land Cover data
- Other LU datasets that could be used in creating LU and LC dataset from the scratch

The key national datasets that are used in creating the national HR CLC data and are listed in table "Source description". The methods used in combining national GIS data together with EO can be found in the CLC 2018 final report. Table "HR CLC 2018 classes" include the nomenclature on the 4th level of HR CLC data.

SYKE has previously conducted EU funded projects where provision of land use data has been developed and tested at SYKE.

In the project "*Provision of harmonized land use information: LUCAS and national systems*" (funded by EUROSTAT) the aim was to solve how statistical data on land use and land cover could be produced using national spatial data sources with the level of precision and nomenclature required by EUROSTAT. The exercise revealed data gaps in Finland especially in producing spatial data about secondary production and services sector.

Recently (2020-2021) SYKE and LUKE participated into Copernicus User Update project "*Developing support for monitoring and reporting of GHG emissions and removals from land use, land use change and forestry"* together with several organizations responsible for LULUCF reporting in various Member States (FPCUP DG Clima project). The project was coordinated by Finnish Meteorological Institute. The work included recoding of Finnish HR CLC classes into the IPCC land use categories. This recoding was taken into account also in this work. Additionally net GHG emissions and removals were calculated using time series of LU statistics based on HR CLC data. These were compared with official GHGI reports, where LULUCF information is based on intensive network of field sample plots measured in National Forest Inventory (NFI) by LUKE. The study came into conclusion, that it is possible to calculate this information using HR CLC, if additional information on carbon stocks and emission factors are provided to complete the calculations. Also, some additional information on soils in croplands are needed. Land use statistics based on HR CLC data seems to slightly overestimate forest land when compared to NFI based method and slightly underestimate all other classes.

Subtask 2 - Creation of simplified LULUCF LUA layer

In creating simplified LULUCF LUA layer the Finnish HR CLC 2018 was used as a source, because there was no comprehensive LU data available from Finland and there was not enough time to make new LU data from scratch. Additionally recent FPCUP DG Clima project shows that this approach is feasible. HR CLC 2018 data is a 20m raster data that works as a base data for generalized 25ha MMU CLC 2018. HR CLC 2018 is produced by combining EO data with existing spatial data on land cover and land use together to create wall-to-wall data. HR CLC 2018 has 49 classes in the 4th level giving more detailed information on some of the EU CLC 3rd level classes.

LUA layer was created according to guidelines given in LU inventory refinement MS instructions vs 3.5 chapter 4 following steps 1, 2 and 3. First the original HR CLC was reclassified into LU inter-classes according to LU inventory refinement excel. After that it was projected into ETRS1989-LAEA projection using EEA LAEA reference grid to match the grid geometry and finally the pixel size was resampled from 20m to 10m (Figure 1).

In previous projects making a comprehensive LU data has been piloted, but the data was too old and covered only parts of Finland to be usable in this exercise. Based on these previous projects possible data gaps and other problems in producing comprehensive LU with currently available datasets have been found. Thus, it can be stated, that it is possible to produce LU data but there are classes with insufficient data or with no available spatial data at all, so the resulting LU data will not have all the classes HILUCS or the previous EAGLE classification require.

HR CLC 2018 reclassification into LU inter-classes was completed using EU CLC translation tables given by EEA and reclassification work (HR CLC to Finnish LULUCF classification) done in above mentioned FPCUP DG Clima project. Based on the latter three 4th level classes have been classified differently from the 3rd level EU CLC translation table: 1421 (Summer cottages) has been classified into residential instead of Services, 2441 (Agro-forestry) into Grasslands because in Finland this class consists of pastures with trees and 3246 (Transitional woodland/shrub under power lines) into Utilities instead of Forest.

The composition of LU inter-classes based on HR CLC 2018 is described in LU inter-classes table column “National data source class name field”. This was possible at Level-2 as requested by EEA. Some shortcomings in the correspondence are listed in the comments. These include:

- Agroforestry Use and Greenhouse Use within Agriculture Use
- Aquaculture Use from the Water use
- Forest areas include all forested areas regardless of their usage; protected areas and areas managed mainly for recreation (these uses also overlap) are included in the Forestry Use.
- Transport & Logistic Use: There are no waterways included. In general, all the logistics areas are included in the associated land and cannot be separated from other classes.
- Utilities: In addition to Dump sites and Construction sites Utilities includes Power lines but no other power distribution services or water and sewage infrastructure.

Most of HR CLC 2018 classes fit to EAGLE LUA classes mentioned in the excel, but there are a couple of exceptions, which are described above (refinement of LU Inventory).

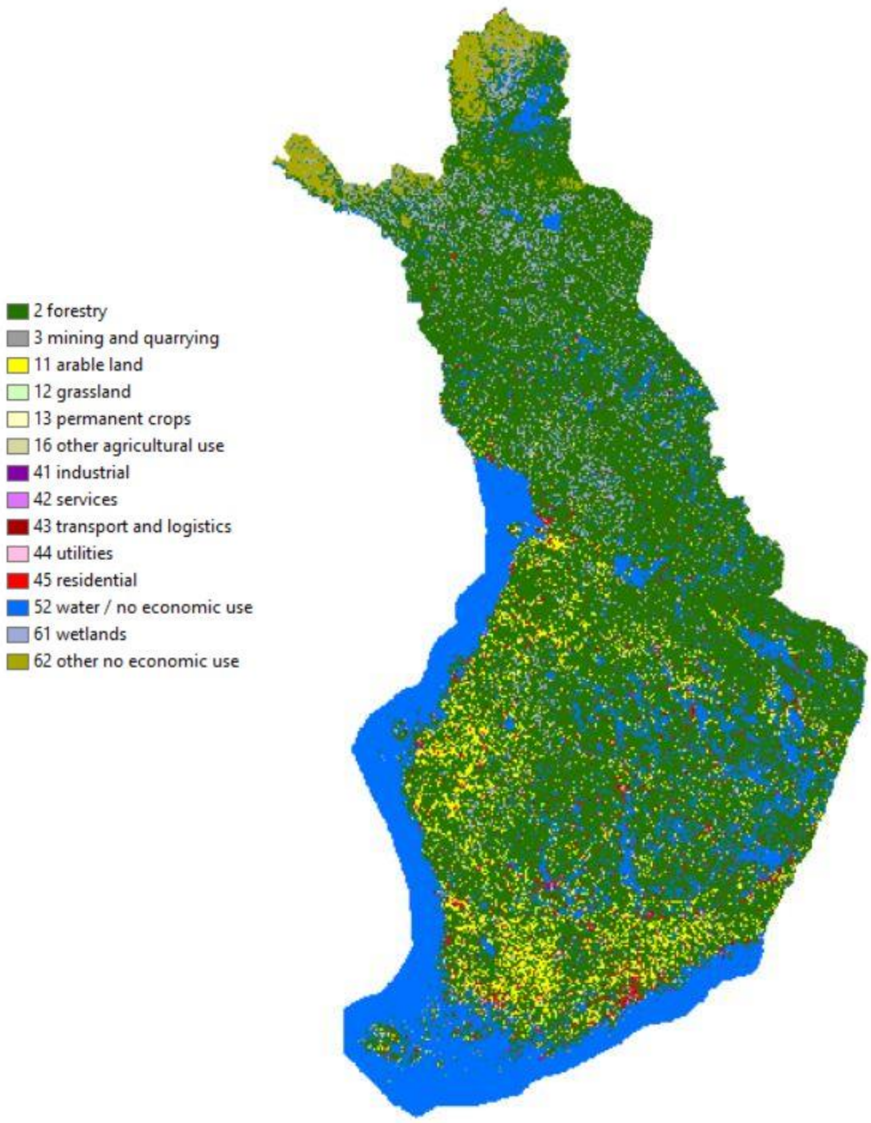

Figure 1. Simplified LULUCF LUA layer from Finland

Presently SYKE is coordinating a large national cooperation project funded by the Ministry of Agriculture and Forestry, where the objective is to support Finnish LULUCF sector in the reduction of GHG emissions by providing relevant LU information on regular bases (https://mmm.fi/en/climate-plan-for-the-land-use-sector). The outputs of the project are not yet available, but these data will possibly be useful in the ingestion of land use data for CLC+ Core in the future.

Conclusions

This Framework contract included the last update of Corine Land Cover data completed by Member states. In Finland monitoring of Corine land cover started in year 2000, and 2018 update was the fourth CLC dataset in the timeseries of land cover information. Participation into this joint European effort included useful cooperation within Europe (Eionet) and Nordic countries. Additionally, this provided significant resources and data, which were beneficial also for national data production and fostered also national cooperation in land monitoring. In Finland, in each CLC update also high resolution (25-20 m resolution raster) data were produced as input for EU version of CLC and also for national use. According to significant amounts of data downloads these HR land cover data are heavily used. Thus, participation into data production of Copernicus services has been very useful also from national perspective in Finland. In 2020, SYKE conducted an end user survey about the usage of High Resolution CLC land cover data. These datasets were regarded very useful and future updates of these data were requested. At the moment continuation of this service remains open.

In the future European land monitoring will be completed centrally by service providers (CLC+). The national usefulness of these data is unknown, since for example CLC+ backbone data were not validated in this project and thus we do not have yet any experience on the data. Data provision organized using top-down approach may provide homogenous data services for pan-european use at EU. However national and local usage of these data remains questionable, which is also our present experience learned in the previous post-production verifications of HRLs and local components. This is especially the case in Finland, where high quality national data sources are available as open data.

As an example of nationally useful centralized Copernicus data services for land applications are Mosaic Hub (https://apps.sentinel-hub.com/mosaic-hub/#/) and HR vegetation phenology and productivity. This is due to the fact that these services are highly data intensive and require heavy processing of temporal, raw satellite data. Additionally, production of these services does not necessarily require local in-situ data and expertise. Centralized provision of these services provides high benefits and saves local resources. Hopefully these and corresponding future services are developed in Copernicus 2.0. Sentinel 2 reflectance mosaics provided by Mosaic-Hub and seasonal phenological parameters of HR-VPP are already used/tested in Finland for example in habitat mappings as input for machine learning algorithms. These Copernicus data create possibilities for new national applications.

A significant part of originally planned tasks, like verification of CLC+ Backbone and Establishment of Rulesets for ingestion and extraction, were not completed within this specific contract. Thus, the feasibility of CLC+ Core remains open.

In order to involve national expertise, data and get hands-on experience on the quality and usability of CLC+ data services, national participation is needed also in Copernicus 2.0. This is required to guarantee high quality of these services and put these services into use.

Done at Helsinki, Finland

Date: 27 January 2022

Name: Pekka Härmä

Signature: