Climate Change Initiative Extension (CCI+) Phase 2
New Essential Climate Variables (NEW ECVS)
High Resolution Land Cover ECV (HR_LandCover_cci)

Product Validation Plan

(PVP)

Prepared by:

Università degli Studi di Trento
Fondazione Bruno Kessler
Università degli Studi di Pavia
Università degli Studi di Genova
Université Catholique de Louvain
Politecnico di Milano
LSCE
CREAF
University of Exeter
e-GEOS s.p.a.
Planetek Italia

Changelog

Issue	Changes	Date
1.0	First version	03/07/24
1.3	Use a more recent Mapbiomas dataset (collection 9 instead of 8), Therefore, all the relevant parts were updated: Table 11, Table 12 and Table 22 were adjusted. Modifications in Table 11 related to the temporal resolution of dataset GL30, FNF and Tree canopy cover and legend for FNF new version is also added, now in Table 32. Added sentence to explain the removal of CCI Prototype Africa dataset in Section 4.2 to ensure consistency with PSD. The legend of this dataset and legend link in Table 40 were also removed. Adjust the typo of legend description in Table 19 and Table 24.	12/09/2024

Detailed Change Record

RID	Description of discrepancy	Sections	Change
1.0	Hyperlink was broken **Error! Reference source not found**. shows the newly collected HR LCs as well as the HR LCs collected during Phase 1, which are indicated in the "Phase 1 Dataset" column.	4	Updated the broken hyperlink

Contents

Introduction

1.1 Purpose and scope

The objective of this Product Validation Plan (PVP) is to describe the strategies selected for the validation of the European Space Agency (ESA) Climate Change Initiative (CCI) High Resolution (HR) Land Cover (LC) products: 10-m static LC maps for the year 2019 at the sub-continental level and 30-m regional historical (1990-2024) LC maps over reduced areas within this level [AD3]. The results regarding the qualitative and quantitative assessments of the CCI HR LC products are included in the Product Validation and Intercomparison Report (PVIR) [AD2].

1.2 Applicable documents

[AD1] CCI_HR LC _Ph1-D2.3_E3UB, v1.0, 03/07/2019
[AD2] CCI_HR LC _Ph1-D4.1_PVIR, the latest version
[AD3] CCI_HR LC _Ph2-D1.2_PSD, the latest version.

1.3 Acronyms and abbreviations

ALOS	Advanced Land Observing Satellite
AUE	Atlas of Urban Expansion
C3S	Copernicus Climate Change Service
CCI	Climate Change Initiative
CEOS	Committee on Earth Observation Satellites
CI	Confidence Interval
DLR	German Aerospace Center
DUE	Data User Element
E3UB	End-to-End ECV Uncertainty Budget
ECV	Essential Climate Variables
EO	Earth Observation
ESA	European Space Agency
ETM	Enhanced Thematic Mapper
FAO	Food and Agriculture Organization
FNF	Forest/Non-Forest
FROM-GLC	Finer Resolution Observation and Monitoring of Global Land Cover
GCOS	Global Climate Observing System
GEE	Google Earth Engine
GEO	Group for Earth Observation
GFC	Global Forest Cover
GHS BU LDS	Global Human Settlement BuiltUp derived from Landsat image collections
GHS BU S1	Global Human Settlement BuiltUp derived from Sentinel1 image collections
GISD30	Global 30 m Impervious-Surface Dynamic dataset
GL30	GlobeLand30
GLC	Global Land Cover
GLC_FCS30	Global Land-Cover product with Fine Classification System at 30 m
GLC_FCS30D	Global 30 m Land-Cover Dynamics monitoring product with a Fine Classification System
GLS	Global Land Survey
GLanCE	Global Land Cover Mapping and Estimation
GOFC-GOLD	Global Observation of Forest Cover – Global Observation of Land Dynamics

GRASS-GIS	Geographic Resources Analysis Support System - Geographic Information System
GSW	Global Surface Water
GUF	Global Urban Footprint
GWL_FCS30	Global 30 m Wetland Map with a Fine Classification System
GWL_FCS30D	Global annual Wetland dataset at 30 m with a Fine Classification System
HR	High Resolution
IFOV	Instantaneous Field of View
INPE	National Institute for Space Research
JAXA-EORC	Japan Aerospace Exploration Agency - Earth Observation Research Center
JRC	Joint Research Centre
LC	Land Cover
LCC	Land Cover Change
LCCS	Land Cover Classification System
LPVS	Land Product Validation Subgroup
LULC	Land Use Land Cover
MMU	Minimum Mapping Unit
MR	Medium Resolution
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetative Index
NGCC	National Geomatics Center of China
NYU	New York University
OA	Overall Accuracy
PA	Producer Accuracy
PCC	Priority of Change
PSD	Product Specification Document
PVIR	Product Validation and Intercomparison Report
PVP	Product Validation Plan
S2	Sentinel-2
SEEG	Greenhouse Gas Emissions and Removals Estimates
TM	Thematic Mapper
UA	User Accuracy
UCLouvain	Université catholique de Louvain
UMD	University of Maryland
USGS	United States Geological Survey
VHR	Very High Resolution
WGCV	CEOS Working Group on Calibration & Validation
WSF	World Settlement Footprint

2 The CCI HR LC products to be validated thematically

2.1.1 Three types of CCI HR LC outputs

Three types of land cover products will be generated and validated thematically within the ESA CCI HR LC project:

- **10-m static sub-continental LC maps for the year 2019**, evaluated per class.
- **Historical 30-m regional dynamic LC maps** in the **1990-2024** period generated every five years for the period on reduced areas inside the sub-continental static LC maps. Yet, land cover change was detected on an annual basis among these five-year intervals. We evaluated the **Best Class layer.**
- **Historical 30-m change detection maps** for **1990-2024** obtained from Landsat 5, 7 and 8 products at a spatial resolution of 30m covering the same extents as the historical 30-m regional LC maps.

  Each historical change detection map comes with an indication of the priority of change (PCC) where PCC equals 1 for the low priority of changes and 2 for the high priority of change.
  For changes with a high priority of change (PCC = 2), additional information is provided:
  - *Year*: the value is the year in which the pixel has changed and 0 for no change.
  - *Probability*: the probability of change from 0 to 100
  - *Temporal reliability*: temporal distance between years in which the change has been calculated

Each type of product was delivered in the form of tiles and mosaics. The mosaic format was assessed in priority, although the tiles were also checked to some extent for comparison.

The products will be generated over three areas selected through key users' consultation, with varying extents (Figure 1). The Phase 1 static LC maps, in green, cover the extended regions in the Amazon (including Mato Grosso), Sahel and Siberia. The historical LC maps are restricted to the orange areas.

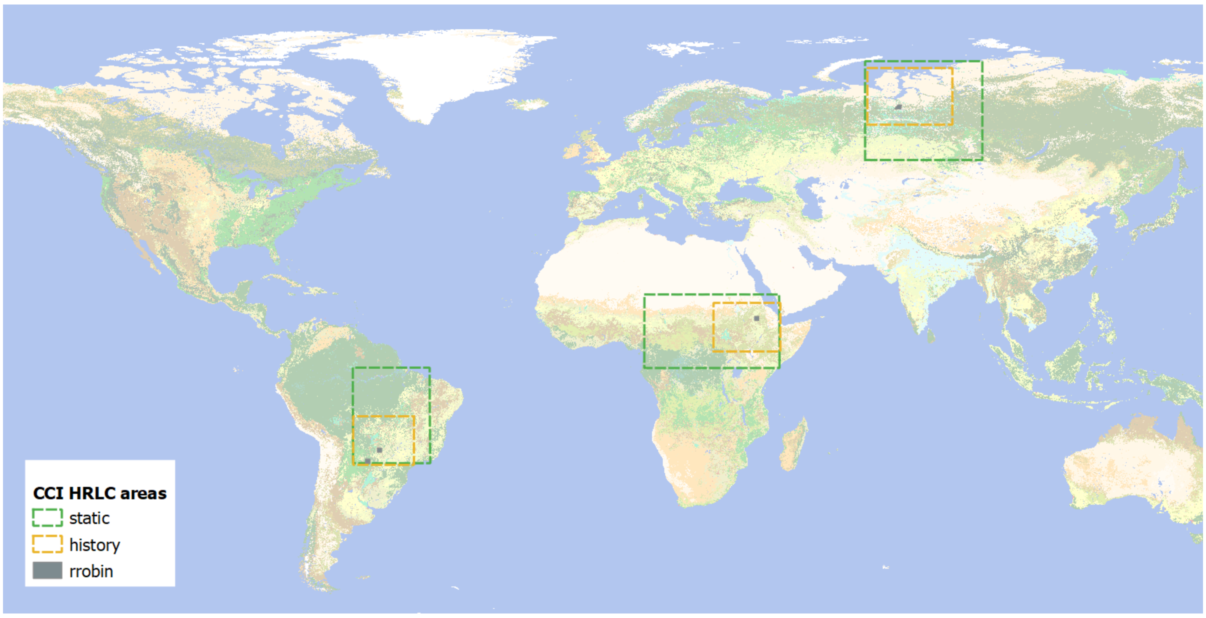

Figure 1: Distribution of the Phase 1 study sites per type of CCI HR LC product.

The Phase 2 Amazon area is expanded to the East and North-West side for HRLC10, HRLC30 and HRLCC30 considering Figure 2. The current extension is being fine-tuned in line with user requirements and technical aspects.

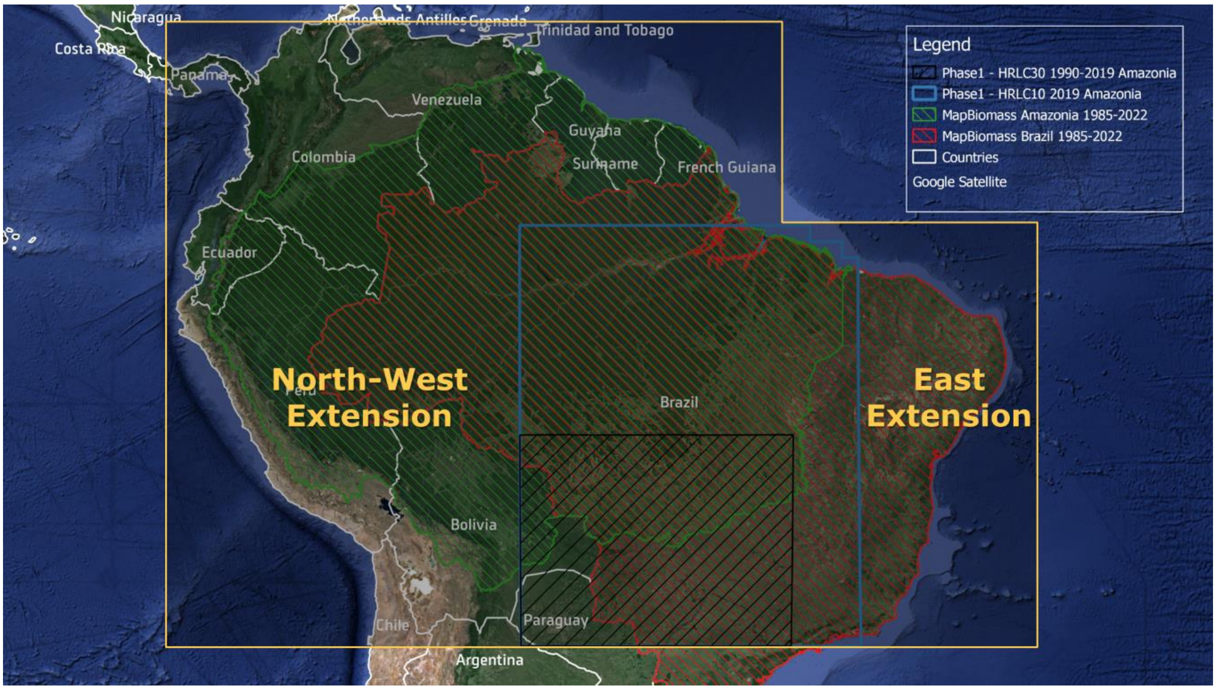

Figure 2: Amazon area expansion to South America.

Table 1 summarizes the extent and areas of each region according to the land cover classification activity.

Table 1. Spatial extents of the different HR land cover products.

Region	CCI HR LC outputs	Extent	Surface [km²]	Lat min	Lon min	Lat max	Lon max
Amazon	LC map 2019	Full extent	5370097	-24	-62	0	-43
	Historical LC maps	Reduced extent	2230570	-24	-62	-12	-47
Amazon extension	LC map 2019 and Historical LC maps	Full extent	21335500	-24	-82	12	-34
Sahel	LC map 2019	Full extent	7338692	0	10	18	43
	Historical LC maps	Reduced extent	2453620	4	27	16	43.5
Siberia	LC map 2019	Full extent	3895765	51	64.5	75.5	93.5
	Historical LC maps	Reduced extent	3643260	60	65	74	86

2.1.2 The CCI HR LC legend

The Food and Agriculture Organization (FAO) Land Cover Classification System (LCCS) was found pertinent to support the description of the CCI HR LC maps. Based on key user consultations and after adaptation to the FAO LCCS framework, a set of 16 main classes is proposed for the LC mapping at 10 m spatial resolution (Figure 3) with class descriptions in Table 2. The legend was built with four levels of thematic complexity. The main 16 classes correspond to Level 1. Descriptions of Levels 2, 3 and 4 are available in Annex 2.

HRLC CLASSES			
CODE	DESCRIPTION		
0	No data		
10	Tree cover evergreen broadleaf		
20	Tree cover evergreen needleleaf		
30	Tree cover deciduous broadleaf		
40	Tree cover deciduous needleleaf		
50	Shrub cover evergreen		
60	Shrub cover deciduous		
70	Grasslands		
80	Croplands		
90	Woody vegetation aquatic or regularly flooded		
100	Grassland vegetation aquatic or regularly flooded		
110	Lichens and mosses		
120	Bare areas		
130	Built-up		
140	**Open water**	141	Open water seasonal
		142	Open water permanent
150	Permanent snow and/or ice		

Figure 3. The CCI HR LC legend includes 16 main land cover classes.

Table 2. FAO LCCS description of the 1st level of land cover classes selected for the CCI HR LC products.

Code	Label	Description
10	Tree cover evergreen broadleaf	Primarily vegetated areas with a tree canopy cover of more than 50 % at the time of fullest development. Snow and/or ice, open water or built-up areas cover less than 50% of the area. A tree is a woody, perennial plant with a simple and well-defined stem, bearing a more or less defined crown [1] and a minimum height of 5 m. Tree canopy cover is composed of trees that are never entirely without green foliage [1]. Trees are broadleaved and come from the Angiospermae group.
20	Tree cover evergreen needleleaf	Primarily vegetated areas with a tree canopy cover of more than 50 % at the time of fullest development. A tree is a woody, perennial plant with a simple and well-defined stem, bearing a more or less defined crown [1] and a minimum height of 5 m. Tree canopy cover is composed of trees that are never entirely without green foliage [1]. Trees carry typical needle-shaped leaves and come from the Gymnospermae group.
30	Tree cover deciduous broadleaf	Primarily vegetated areas with a tree canopy cover of more than 50 % at the time of fullest development. Snow and/or ice, open water or built-up areas cover less than 50% of the area. A tree is a woody, perennial plant with a simple and well-defined stem, bearing a more or less defined crown [1] and a minimum height of 5 m. Tree canopy cover composed of trees that are leafless for a certain period during the year [1]. Trees are broadleaved and come from the Angiospermae group.
40	Tree cover deciduous needleleaf	Primarily vegetated areas with a tree canopy cover of more than 50 % at the time of fullest development. Snow and/or ice, open water or built-up areas cover less than 50% of the area. A tree is a woody, perennial plant with a simple and well-defined stem, bearing a more or less defined crown [1] and a minimum height of 5 m. Tree canopy cover is composed of trees that are leafless for a certain period during the year [1]. Trees carry typical needle-shaped leaves and come from the Gymnospermae group.
50	Shrub cover evergreen	Primarily vegetated areas with a shrub canopy cover of more than 50 % at the time of fullest development. Snow and/or ice, open water or built-up areas cover less than 50% of the area. A shrub is a woody perennial plant with persistent woody stems and

		without any defined main stem [1], being less than 5 m tall. Shrub canopy cover composed of shrubs that are never entirely without green foliage [1].
60	Shrub cover deciduous	Primarily vegetated areas with a shrub canopy cover of more than 50 % at the time of fullest development. Snow and/or ice, open water or built-up areas cover less than 50% of the area. A shrub is a woody perennial plant with persistent woody stems and without any defined main stem [1], being less than 5 m tall. Shrub canopy cover is composed of shrubs that are leafless for a certain period during the year [1].
70	Grassland	Primarily vegetated areas with a herbaceous cover of more than 50% at the time of fullest development. Snow and/or ice, open water or built-up areas cover less than 50% of the surface. Herbaceous plants are defined as plants without persistent stems or shoots above ground and lacking definite firm structures [2].
80	Croplands	Primarily vegetated areas with a herbaceous cover of more than 50 % at the time of fullest development. Snow and/or ice, open water or built-up areas cover less than 50%. Croplands are mainly herbaceous plants that are sowed/planted and harvestable at least once within the 12 months after the sowing/planting date. Herbaceous plants are defined as plants without persistent stems or shoots above ground and lacking definite firm structures [2]. Cropland includes rainfed crops, irrigated crops, aquatic crops and annual pastures. It is an adaptation of the Joint Experiment for Crop Assessment and Monitoring (JECAM) cropland definition [3]. Croplands exclude permanent crops like woody plantations that are part of the tree or shrub classes.
90	Woody vegetation aquatic or regularly flooded	Primarily vegetated areas with trees and/or shrubs, grasslands or lichens and mosses covering more than 50 % of the area flooded by water for more than 4 months throughout the year. The water can be saline, fresh or brackish.
100	Grassland vegetation aquatic or regularly flooded	Primarily vegetated areas with grasslands and/or lichens and mosses covering more than 50 % of the area flooded by water for more than 4 months throughout the year. The water can be saline, fresh or brackish.
110	Lichen and mosses	Primarily vegetated areas with a cover of more than 50% at the time of fullest development. Snow and/or ice, open water or built-up areas cover less than 50% of the surface. Mosses are a group of photo-autotrophic land plants without true leaves, stems or roots [4]. Lichens are composite organisms formed from the symbiotic association of fungi and algae [4].
120	Bare areas	Areas where the sum of vegetation cover is less than 50% at the time of fullest development. Snow and/or ice, open water or built-up areas cover less than 50% of the surface. Bare rock areas, sands and deserts are classified as bare areas. Extraction sites (open mines and quarries) and salt flats covered by water for less than 5 months are classified as bare areas.
130	Built-up	Areas where any predominant type of linear and non-linear artificial surface covers at least 50%. Snow and/or ice, and open water cover less than 50% of the surface. Built-up areas include buildings, roads, airports, greenhouses, etc. but may, however, exclude temporary settlements.
141	Open Water seasonal	Areas where open water covers at least 50% of the surface and remains between 5 and 9 months a year, except in special circumstances (particularly dry years, construction of dams, etc.). Snow and/or ice and built-up areas cover less than 50% of the surface. Water bodies can be natural or artificial. Water can be saline, fresh or brackish.
142	Open Water permanent	Areas where open water covers at least 50% of the surface and remains for more than 9 months a year, except in special circumstances (particularly dry years, construction of dams, etc.). Snow and/or ice and built-up areas cover less than 50% of the surface. Water bodies can be natural or artificial. Water can be saline, fresh or brackish.
150	Permanent snow and ice	Areas where snow and/or ice cover at least 50% of the surface for more than 9 months a year. Built-up areas and open water cover less than 50% of the surface.

	and/or Ice

2.1.3 Definition of change

The definition of "change" is equally important for validation as the definition of the LC classes. Here, LC change refers to a permanent modification of the LC – and not of the seasonality of the surface – in comparison with the baseline status. Whether a change is observed or not also depends on the spatial scale of the products being analysed. The spatial resolution of the CCI HR LC historical LC maps is 30 m, corresponding to an area of approximately 0.09 ha per pixel. A pixel is considered to be changed when the change covers more than half a pixel.

Among the historical products, we consider the 'Change Detection' mosaic filtered for PCC = 2, i.e. for a high change priority. The 'Year' layer is used to define the year of change and the 'Dynamic Land Cover Best Class' is used to define the land cover class before and after the change.

For the Minimum Mapping Unit (MMU), erosion is applied to the "Change Detection" layer, retaining only hotspots with an area of 0.81 ha. This decision is supported by the reliability of the change detection algorithm and enhances the reliability of photo-interpretation. Consequently, derived accuracy metrics will be more robust as they will be based on significant changes. These decisions influence the response design.

3 State-of-the-art quantitative accuracy assessment

The validation is an essential step for providing high-quality products, endorsed by the ESA climate modelling and broader user community. The current validation exercise is based on the lessons learned from previous projects like the Global Land Cover (GLC) 2000 [5], Globcover [6], [7], and the CCI LC 1992-2015/C3S 2016-2017 maps [7]. It is also intended to reflect state-of-the-art standard protocols of LC validation such as the CEOS Working Group on Calibration and Validation (LC validation subgroup). In particular, the design and implementation of the validation plan follow the general recommendations of the GOFC-GOLD validation report [8] and other scientific publications from these groups [5], [9], [10].

The overall validation process follows accepted state-of-the-art methodologies (see Section 3) and includes an independent statistical quantitative validation of the three HR LC outputs. The methodology is fine-tuned to the specific challenges in validating each type of product. The validation of the static HR LC map aims at quantifying and reporting the quality of the product from an overall, producer and user accuracy point of view. Validating the historic HR LC maps represents the most significant challenge of this exercise as building a statistically sound validation of change in the context of multiple LC classes is a rather new domain to explore by the community. For all types of products, the validation process is composed of 4 steps (Section 3.2.3): the collection of Very-High-Resolution (VHR) imagery, the sampling design, the response design using VHR imagery and Google Earth Engine Time Series Explorer temporal profiles, and the reporting.

3.1 Definition and standard protocols

For each type of CCI HR LC product and region, complementary qualitative and quantitative assessments are performed.

The qualitative assessment consists of a systematic visual quality control of each product received. It provides an overall appreciation of the completeness and quality of the products which can exhibit some macroscopic artefacts or classification errors affecting specific LC classes (wrong label, missing classes) or the spatial pattern to be delineated (wrong position of the boundary between classes, the disappearance of small patches, etc.).

Each quantitative assessment follows the state-of-the-art reporting content described in Section 3.2.5.

There are several definitions of validation available from various agencies but within the CCI program, the definition from the Committee on Earth Observing Satellites Working Group on Calibration and Validation (CEOS-WGCV) was adopted. It defines validation as: *"The process of assessing, by independent means, the quality of the data products derived from the system outputs".*

It is assumed that the term "data products" in the above definition refers to both the geophysical parameter (i.e. the Level-4 LC classification) and its uncertainties. Information related to the characterization of uncertainties is documented in the End-to-End Uncertainty Budget (E3UB) [AD1].

3.1.1 Validation stages

The Committee on Earth Observation Satellites (CEOS), recognized as the space arm of the Group for Earth Observation (GEO), plays a key role in coordinating the land product validation process that depends on the temporal and spatial coverage of available reference data, thus providing a confidence estimate for each product even if there is little or no in situ data (Table 3). The CEOS-WGCV has defined initially three validation stages. However, in response to the evolving ECV monitoring activities, validation stage 4 was included to define an operational component to ensure that the time series of land products are systematically validated.

We propose to fill the CEOS WGCV stage 3 within this project as no systematic and operational validation updates are planned.

Table 3. Four levels of validation were adopted by the Committee on Earth Observation Satellites Working Group on Calibration and Validation.

Stage 1	Product accuracy is assessed from a small (typically < 30) set of locations and periods by comparison with reference in situ and/or higher resolution airborne or satellite data. Spatial and temporal consistency of the product and consistency with similar products has been evaluated over selected locations and periods.
Stage 2	Product accuracy is estimated over a significant set of locations and periods by comparison with reference in situ and/or higher resolution airborne or satellite data. Spatial and temporal consistency of the product and consistency with similar products has been evaluated over globally representative locations and periods. Results are published in the peer-reviewed literature.
Stage 3	Uncertainties in the product and its associated structure are well-quantified from comparison with reference in situ and higher resolution airborne and satellite data. Uncertainties are characterized in a statistically robust way over multiple locations and periods representing global conditions. Spatial and temporal consistency of the product and consistency with similar products has been evaluated over globally representative locations and periods. Results are published in the peer-reviewed literature.
Stage 4	Validation results for stage 4 are systematically and operationally updated by independent actors for comparative assessment of existing products when new products are released and as the time series expands.

3.1.2 Validation requirements

The validation procedure of the CCI HR LC maps is also driven by the main GCOS requirements [11], summarized in Table 4.

Table 4. Maps of high-resolution land cover terrestrial ECV product requirements from GCOS.

Coverage and sampling	
Geographic coverage	Regional
Temporal sampling	Yearly and every 5 years (breakthrough and threshold requirements)
Temporal extent (time span)	10 to 30 years (breakthrough requirement)

Resolution	
Geometrical Resolution	10 - 30 m (breakthrough requirement)
Error/Uncertainty	
Accuracy-Uncertainty	5, 20 and 35 % for the goal, breakthrough and target requirements, respectively. This is the maximal percentage for accuracy and errors of omission and commission and hectares for area estimates incl. 95 % confidence intervals. The geolocation accuracy should be better than 1/3 IFOV with target IFOV 10–30 m
Stability	As above, per decade.

With these requirements in mind, the validation procedure will highlight the following aspects:

- A **per-class accuracy** analysed in the light of the expected rate of omission and commission error
- The need for **stable accuracy** should be reflected in implementing an accuracy assessment of LC change, at least per decade.

3.2 Good practices of accuracy assessment

3.2.1 Independence of the validation process

The static HR LC and the historical LC maps were validated using a transparent traceable procedure relying on statistical quantities, independent from the production process to be considered a scientifically credible input for climate assessments and modelling. The validation procedure follows defined protocols approved by the CEOS Land Product Validation Subgroup (LPVS) (http://lpvs.gsfc.nasa.gov/) [8], [10].

The independence of the validation process adopted in this PVP is two-fold:

1. In situ, other suitable reference datasets and auxiliary datasets used for validation should not have been used during the production of the products to be validated. As [12] states, the accuracy assessment is conducted independently of classifier training.
2. The validation is carried out by staff not involved in the generation of the LC products.

3.2.2 Reference data specification

The collection of ground information (i.e. through field surveys) is considered the best option to support the validation of remote sensing products. The cost of manpower and logistics to organize field visits to remote areas with difficult or impossible access if historical LC maps need to be validated makes the collection of ground truth data not feasible for a large number of plots distributed over large areas.

Reference data should be of an equal or finer level of detail than the data used to create the map [10]. Existing "reference data sources" like VHR imagery interpreted by experts are good surrogates to "ground truth".

3.2.3 Sampling frame requirements

To satisfy the requirements of design-based inference, the sampling design should be a probability sampling design, and the estimators should be constructed following the principle of a consistent estimation [13].

The sampling scheme will be designed with the following general requirements:

- to be statistically valid for the accuracy assessment of the CCI HR LC products;
- to be reusable for future products of a similar type;
- to be designed before (i.e. independently) the CCI HR LC product;
- to use the most recent picture of global LC distribution (as the best proxy of the current LC distribution);
- to be scale-independent to allow the evaluation of scaling issues between the CCI HR LC and MRLC products;
- to consider the availability of reference VHR resolution imagery for the recent and historical years.

3.2.3.1 Sampling designs for LC change assessment

While large-scale validation standards are well recognised in the international community, the **validation of LC change** remains very open. Validating broad-scale change products is often challenging because it is subject to a twofold constraint. First, change is a rare event [14]. The commission rate is easy to quantify by examining objects identified as having changed but it is much more complex to estimate the omission rate among large numbers of objects identified as unchanged [8]. Second, the availability (and quality) of reference data decreases when going back in time and the poor match between observation dates, i.e. validation versus detection, is a source of uncertainty.

3.2.4 Response design

The reference data sources are then intended to be interpreted over each sample by LC experts in a standardized manner. LC experts should meet the following criteria:

- Recognized expertise in LC over large areas;
- Familiarity with the interpretation of remote sensing imagery;
- Good understanding of the LC legend of the products;
- Good understanding of the definition of LC change.

Based on the GlobCover [15], [16] and CCI MRLC validation experiences, the process of sample interpretation can be ambiguous for the following reasons: (i) inadequate quality of the reference imagery; (ii) heterogeneity of the landscape, (iii) limited knowledge by the expert and (iv) ambiguity of the LC legend. Therefore, clear description of each class of the CCI HR LC legend (Table 2) as well as a detailed image interpretation protocol (Sections 3.3.3 and 3.4.3) ensure each sample unit is interpreted – i.e. labelled – by the expert systematically and consistently:

- If the expert cannot derive the LC because of the poor quality of reference imagery, the sample has to be skipped. The expert must specify that no LC class have been assigned to the sample because of the insufficient quality of the data.
- If the landscape is heterogeneous, the expert has to explicitly specify that the landscape is complex. The segmentation procedure tackles this heterogeneity issue and will generate many small polygons in heterogeneous landscapes.
- If the expert is not sure how to interpret the sample, he/she can indicate a lower level of certainty. When there is serious doubt about the exact LC class, the expert needs to indicate the classes from which the expert cannot choose with certainty. More attributes than the dominant LC classes are relevant, especially for the analysis of observed discrepancies between classification and expert labelling.

3.2.5 Reporting

3.2.5.1 The error matrix

This validation report will analyse in detail the various parameters describing the accuracy of the map: contingency matrix, user's and producer's accuracy. The confusion matrix is recognised to efficiently organize and summarize the agreement between the maps and reference classification [17] (Figure 4).

		Reference				
		Class 1	Class 2	Class 3	Class 4	Total
Map	Class 1	p_{11}	p_{12}	p_{13}	p_{14}	$p_{1\cdot}$
	Class 2	p_{21}	p_{22}	p_{23}	p_{24}	$p_{2\cdot}$
	Class 3	p_{31}	p_{32}	p_{32}	p_{34}	$p_{3\cdot}$
	Class 4	p_{41}	p_{42}	p_{43}	p_{44}	$p_{4\cdot}$
	Total	$p_{\cdot 1}$	$p_{\cdot 2}$	$p_{\cdot 3}$	$p_{\cdot 4}$	1

Figure 4. The layout of a typical confusion or error matrix is as in [10].

The overall accuracy (O), user accuracy (U_i) and producer accuracy (P_j) will be calculated according to Eq. 1, Eq. 2, and Eq. 3, respectively.

$$O = \sum_{j=1}^{q} p_{jj}$$

Eq. 1

$$U_i = \frac{p_{ii}}{p_{i\cdot}}$$

Eq. 2

$$P_j = \frac{p_{jj}}{p_{\cdot j}}$$

Eq. 3

The F-score could also be reported. It represents for a class *i* the harmonic mean of the user and producer accuracies and ranges between 0 and 1, as defined in Eq. 4.

$$F - Score_i = 2 * \frac{UA_i * PA_i}{UA_i + PA_i}$$

Eq. 4

3.2.5.2 Reporting area-based accuracy figures

To calculate the overall accuracy of the product when the sampling design is not of equal probability [10], each class (or stratum in the case of a stratified sampling) should be weighted by the area it represents in the map. A confusion matrix is computed as unbiased by considering the weight of the class/stratum, which is inversely proportional to the sampling effort. The weights are computed for each stratum based on Eq. 5 (Eq. 4 in [10]).

$$\hat{p}_{ij} = W_i * \frac{n_{ij}}{n_{i\cdot}}$$

Eq. 5

where W_i is the proportion of area mapped as class i.

Variances of the overall, user and producer accuracy estimates are derived according to Eq. 6, Eq. 7, and Eq. 8.

$$\hat{V}(\hat{O}) = \sum_{i=1}^{q} W_i^2 \hat{U}_i (1 - \hat{U}_i)/(n_{i\cdot} - 1)$$

Eq. 6

$$\hat{V}(\hat{U}_i) = \hat{U}_i (1 - \hat{U}_i)/(n_{i\cdot} - 1)$$

Eq. 7

$$\hat{V}(\hat{P}_j) = \frac{1}{\hat{N}_{.j}^2}\left[\frac{\hat{N}_{j.}^2\left(1-\hat{P}_j\right)^2\hat{U}_j(1-\hat{U}_j)}{n_{j.}-1} + \hat{p}_{j.}^2\sum_{i\neq j}^{q} N_{i.}^2\frac{n_{ij}}{n_{i.}}(1-\frac{n_{ij}}{n_{i.}})/(n_{i.}-1)\right]$$

Eq. 8

where $\hat{N}_{.j} = \sum_{i=1}^{q}\frac{N_{i.}}{n_{i.}}n_{ij}$ is the estimated marginal total number of pixels of reference class j, $N_{j.}$ is the marginal total of map class *j* and $n_{j.}$ is the total number of sample units in map class j [10].

3.3 Static LC maps quantitative accuracy assessment

This Section presents how the state-of-the-art of accuracy assessment methodology will be applied in the context of this project.

3.3.1 Reference data sources

The use of the GEE Time Series Explorer allows easy access to a wide range of high and VHR reference images.

The image interpretation protocol is mainly based on the very high spatial resolution data available on Google Earth Pro. However, their use may be limited as the level of detail and availability of images may vary from site to site. Furthermore, the possibility of consulting annual multi-temporal spectral index profiles is recommended to characterise the seasonal variations of the different LC. Indeed, Sentinel-2 vegetation index profiles can be extracted with GEE for the year of the static LC map. The tool offers false colour date-specific images.

The use of Planet monthly mosaics with Catalog Planet Lab also provides easy access to these collections with a spatial resolution of 3 metres from 2014 to the present. These very high-resolution images will help to validate the static map.

Note that several types of high and VHR geolocalized imagery with spatial resolutions below 10 m were also identified as potential sources of reference images for validation at the very beginning of the project. The trade-off between the time allocated for validation; the time needed for VHR data ordering and the already extensive access to the reference data in GEE Time Series Explorer explains why VHR geolocalized imagery was ultimately not used.

3.3.2 Sampling designs

The sampling scheme, tailored to each type of CCI HR LC static product, needs to address the issue of rare classes with a strong impact on the climate system (urban areas, wetlands, etc.) in the CCI HR LC static maps.

Following these requirements, three aspects of the sampling design will be addressed: the number of sample plots, their size and the way they are selected from the total population.

Given the time needed for sample photo-interpretation, the construction of the validation database started before the product generation. The number of samples is also constrained by the time for interpreting the data. The GlobCover/CCI MRLC validation exercise showed that experts can interpret between 30 and 50 sample plots per day. Finally, the availability and quality of VHR resolution imagery, the certainty of the LC photo-interpreter and the heterogeneity of the landscape constrain the sample size.

To evaluate the static LC maps, the Phase 1 HR static LC maps were used to estimate the number of samples using the sample size formula for stratified random samplings, using the static LC map classes as strata [Eq. 13 in 10]:

$$n \approx \left(\frac{\sum W_i S_i}{S(\hat{O})}\right)^2$$

Eq. 9

where $S(\hat{O})$ is the standard error of the estimated overall accuracy that we would like to achieve, W_i is the mapped proportion of the area of class *i*, and S_i is the standard deviation of stratum *i*, $S_i = \sqrt{U_i(1 - U_i)}$, with U_i being the user's accuracy of class *i*. In our case, values for $S(\hat{O})$ were set to 0.01 and U_i were taken from the U_i derived from the assessment of the final products of the Phase 1 [AD2].

In order to find an optimal sample distribution that seeks a balance between the number of samples that can be photo-interpolated and precise confidence intervals per class, we preferred an intermediate distribution between proportional and equal. This optimal allocation allows us to set a minimum number of samples for the minority classes, i.e. 40, and to redistribute the remaining points proportionally among the remaining classes. A large number of points are thus retained in the majority classes, maintaining accurate confidence intervals (CIs) while significantly improving the CIs of the minority classes. Considering the Eq2 parameter values as indicated in Table 5, Table 6, and Table 7, a set of 1827, 1878, and 1540 samples will be collected for Siberia, Amazonia, and Africa.

Table 5. Sample allocation for the validation of the Siberia static map. Total samples = 1827.

Siberia area: HRLC10 product sampling allocation with the native legend													Min Alloc.	40
Label	**Code**	**Area (m²)**	**Ui (PVIR)**	**S(Ô)**	**z (95% CI)**	**Si**	**Wi**	**n**	**Equal**	**CI**	**Prop**	**CI**	**Optimal Allocation**	**CI**
Tree cover evergreen needleleaf	20	5.99E+11	69%	0.01	1.96	0.46	17.87%	1827	152	7%	327	5%	314	5%
Tree cover deciduous broadleaf	30	4.80E+11	54%			0.50	14.31%		152	8%	262	6%	251	6%
Tree cover deciduous needleleaf	40	1.18E+11	43%			0.49	3.51%		152	8%	64	12%	62	12%
Shrub cover	50	9.33E+10	59%			0.49	2.78%		152	8%	51	14%	49	14%
Grassland	70	1.00E+12	56%			0.50	29.90%		152	8%	546	4%	525	4%
Croplands	80	2.90E+11	81%			0.39	8.65%		152	6%	158	6%	152	6%
Woody vegetation aquatic or regularly flooded	90	1.96E+11	3%			0.17	5.85%		152	3%	107	3%	103	3%
Grassland vegetation aquatic or regularly flooded	100	1.71E+11	15%			0.36	5.09%		152	6%	93	7%	90	7%
Lichens and mosses	110	4.10E+10	49%			0.50	1.22%		152	8%	22	21%	40	16%
Bare areas	120	4.69E+10	34%			0.47	1.40%		152	8%	26	18%	40	15%
Built-up	130	6451053109	86%			0.35	0.19%		152	6%	4	37%	40	11%
Open Water	140	3.09E+11	96%			0.20	9.23%		152	3%	169	3%	162	3%

Table 6. Sample allocation for the validation of the Amazonia static map. Total samples = 1878.

Amazon area: HRLC10 product sampling allocation with the native legend													Min Alloc.	40
Label	**Code**	**Area (m²)**	**Ui (PVIR)**	**S(Ô)**	**z (95% CI)**	**Si**	**Wi**	**n**	**Equal**	**CI**	**Prop**	**CI**	**Optimal Allocation**	**CI**
Tree cover evergreen broadleaf	10	2.15E+12	78%			0.42	40.38%		156	7%	738	3%	686	3%
Tree cover deciduous broadleaf	30	4.05E+11	51%			0.50	7.61%		156	8%	139	8%	129	9%
Shrub cover evergreen	50	1.07E+11	27%			0.44	2.01%		156	7%	37	14%	40	14%
Shrub cover deciduous	60	1.34E+11	7%			0.25	2.52%		156	4%	46	7%	43	8%
Grassland	70	1.77E+12	61%			0.49	33.24%		156	8%	607	4%	565	4%
Croplands	80	5.26E+11	81%			0.39	9.88%		156	6%	181	6%	168	6%
Woody vegetation aquatic or regularly flooded	90	6950000000	42%	0.01	1.96	0.49	0.13%	1878	156	8%	2	63%	40	16%
Grassland vegetation aquatic or regularly flooded	100	2.58E+10	28%			0.45	0.48%		156	7%	9	30%	40	14%
Bare areas	120	2.64E+10	67%			0.47	0.50%		156	7%	9	31%	40	15%
Built-up	130	1.86E+10	86%			0.35	0.35%		156	5%	6	27%	40	11%
Open Water seasonal	141	6990000000	8%			0.28	0.13%		156	4%	2	35%	40	9%
Open Water permanent	142	1.48E+11	97%			0.17	2.78%		156	3%	51	5%	47	5%

Table 7. Sample allocation for the validation of the Africa static map. Total samples = 1540.

Africa area: HRLC10 product sampling allocation with the native legend													Min Alloc.	40
Label	**Code**	**Area (m²)**	**Ui (PVIR)**	**S(Ô)**	**z (95% CI)**	**Si**	**Wi**	**n**	**Equal**	**CI**	**Prop**	**CI**	**Optimal Allocation**	**CI**
Tree cover evergreen broadleaf	10	1.16E+12	96%	0.01	1.96	0.19	15.79%	1540	140	3%	289	2%	219	3%
Tree cover deciduous broadleaf	30	5.22E+11	43%			0.50	7.11%		140	8%	130	9%	99	10%
Shrub cover evergreen	50	646000000	57%			0.50	0.01%		140	8%	0	242%	40	16%
Shrub cover deciduous	60	8.20E+11	56%			0.50	11.16%		140	8%	204	7%	155	8%
Grassland	70	2.39E+12	45%			0.50	32.53%		140	8%	595	4%	451	5%
Croplands	80	9.00E+11	64%			0.48	12.25%		140	8%	224	6%	170	7%
Grassland vegetation aquatic or regularly flooded	100	2.46E+10	59%			0.49	0.33%		140	8%	6	39%	40	15%
Bare areas	120	1.31E+12	95%			0.22	17.83%		140	4%	326	2%	247	3%
Built-up	130	1.19E+10	89%			0.31	0.16%		140	5%	3	36%	40	10%
Open Water seasonal	141	4260000000	10%			0.30	0.06%		140	5%	1	57%	40	9%
Open Water permanent	142	2.03E+11	87%			0.34	2.76%		140	6%	50	9%	38	11%

3.3.3 Response design

3.3.3.1 Choice of the sample spatial unit

It is unwise to match the size of the validation sample unit to the spatial resolution of the product to validate for the following reasons:

- Geo-location accuracy of the information. The absolute positional accuracy of the LC product is targeted to a 1/3-pixel dimension;
- S2 time series may result in radiometric information coming from a few adjacent pixels;

Therefore, sample plots of 3×3 pixels were interpreted for the validation of each CCI HR LC product. This results in 30 m x 30 m sample plots for the validation of the CCI HR LC static maps.

3.3.3.2 Labelling protocol

Although error-free validation databases do not exist, a rigorous validation protocol is a prerequisite to building it as close as possible to "ground truth". For each sampling unit, a set of attributes is recorded systematically. Table 8 present an optimum attribute table designed for the validation of the static CCI HR LC. The information is collected in order to test two response designs.

Table 8. The information included in the static validation databases for each validation sample plot.

Field name	Details
Sample ID	Unique identifier
Lat / Long	Centre coordinates of the observational unit to interpret. Centroid of the central pixel and of the 3 x 3 window.
Central_Class	The class ID of the LC legend (Table 2) of the central pixel of the sample plot.
Central_Class_Homogeneity	How many pixels in the window are of a similar label as the pixel central
Window_Class	The class ID of the LC legend (Table 2) of the mode/majority of the 3x3 plot
Window_Class_Homogeneity	How many pixels form the mode/majority
Level of certainty	Level of certainty (certain, reasonable, doubtful) associated with the interpretation of the expert
Comments	Comments given by the expert to explain/detail its interpretation (e.g. for indication of why the labelling was not successful, or to give the local name used for the concerned LC type)

The ground, as visible on the VHR imagery, can be spatially complex and include a mixture of classes within the sampling spatial unit. A graphical validation interface is a valuable tool to gather evidence that effectively helps the interpreter converge towards the best guess (Section 3.5).

Following the reporting guidelines of Section 3.2.5, we will use confusion matrices to communicate both overall and per-class accuracy figures. Two confusion matrices are planned:

- The first matrix will compare the reference label at the centre of the 3 x 3 sample window (referred to as "Central_Class" in our validation database) with the map class label at the corresponding "Lat/Long" coordinates.
- The second matrix will consider the reference centre LC label and the mode of the reference LC label from the 3 x 3 sample window referred to as "Window_Class" in our validation database). This means that the reference LC is represented by both the "Central_Class" and an alternative label in the validation dataset. Both labels will be compared with the map class at the corresponding "Lat/Long" coordinates.

3.4 Land cover historic and change maps quantitative accuracy assessment

3.4.1 Reference data sources

The use of the GEE Time Series Explorer allows easy access to a wide range of high and VHR reference images. Landsat vegetation index profiles can be extracted with GEE for the period 2000-present. The tool offers false colour date-specific images. To compensate for the lack of data between 1990 and 2000, the USGS Global Visualization Viewer (GloVis) can be used to validate change by accessing Landsat images and the Global Land Survey (GLS). The GLS was derived from a geodetically accurate, orthorectified global land dataset of Landsat TM (30 m × 30 m) and Enhanced Thematic Mapper (ETM+) (30 m × 30 m) satellite imagery with global coverage. The GLS datasets have the advantage of standardised wall-to-wall imagery available annually. Although it is recommended to rely on a higher spatial resolution than the product to be validated to avoid geolocation problems as much as possible, GLS annual mosaics can be used complementarily with date-specific images to determine whether a change has occurred with annual accuracy.

3.4.2 Sampling design

Considering the specificities of change validation (see Section 3.2.3.1), the sampling should be random and stratified in space to ensure a significant representation of the areas known to experience high rates of change [18] (in [8]). Systematic access to Landsat data from 1992 to the present through the Google Earth, Collect Earth and Glovis interfaces allows us not to stratify over time.

Two criteria guided the generation of the “change” and “no change” stratification layers:

- A priori knowledge of change locations, determined by the PCC values from CCI HR LC historic maps, is used to define stratification and expected user accuracies for each stratum. This approach aims to translate the increasing reliability in change detection, progressing from low-priority changes to high-priority changes and then to stable areas.
- To reduce the uncertainty related to the geolocation issues between the validation footprints and the CCI HR LC products, we applied an erosion buffer of 1 pixel on the union of each change observed in the full period. This means that only hot spots of changes of at least 0.81 ha were attributed to the “change” stratum.

The number of samples selected per stratum is chosen according to Eq. 1 and the evaluation of the time available for the validation exercise. Table 9 provides an example of the values selected per stratum based on the hypothesis that the stratum remaining stable in time is more accurately mapped than the strata (PCC= 1, PCC = 2) experiencing change. Equal allocation is chosen, allowing 306 samples to be distributed in each stratum, for a total of 918 samples per study area. Confidence intervals (CI) for this allocation remain below 5%, and the margin of error $S(\hat{O})$ for overall accuracy is 1%. The current W_i values are assumptions and will be updated once the extent of the historical CCI HR LC maps has been validated.

Table 9. The number of samples per stratum for the evaluation of the CCI HR LC historic LC maps, based on W_i set as an example. U_i is the user accuracy estimated a priori, $S(\hat{O})$ is the standard error on overall accuracy, z is equal to 1. 96 for a 95% confidence interval, S_i is the standard deviation of stratum i, $S_i = \sqrt{U_i(1 - U_i)}$, W_i is the proportion of the class, 'n' is the total number of validation samples, the 'Equal' and 'Prop' columns correspond to equal or proportional allocations to classes, and the associated confidence intervals (CI).

Label	U_i	$S(\hat{O})$	z (95% CI)	S_i	W_i	n	Equal	CI	Prop	CI
PCC=1	75%			0.43	2%		306	5%	18	20%
PCC=2	85%	0.01	1.96	0.36	1%	918	306	4%	9	23%
No Change	90%			0.30	97%		306	3%	892	2%

3.4.3 Response design

3.4.3.1 Choice of the sample spatial unit

It is unwise to match the size of the validation sample unit to the spatial resolution of the product to validate for the following reasons:

- Geo-location accuracy of the information. The absolute positional accuracy of the LC product is targeted to a 1/3 pixel dimension;
- Landsat time series may result in radiometric information coming from a few adjacent pixels;

Therefore, sample plots of 3×3 pixels were interpreted for the validation of each CCI HR LC product. This results in 90 m x 90 m for the validation of the CCI HR LC historical and LCC maps.

3.4.3.2 Labelling protocol

For each sampling unit, a set of attributes is systematically recorded. Table 10 shows an optimal attribute table designed for the validation of CCI HR LC historic and LCC maps. The information is collected to test two response design, i.e. either the central pixel is well classified, or the majority of the 3x3 window is well classified.

Table 10. The information included in the historical validation database for each validation sample plot.

Field name	Details
Sample ID	Unique identifier
Lat / Long	Centre coordinates of the observational unit to interpret
Central_LC_Start	The LC legend class ID (Table 2) of the sample plot center pixel at the start of the period
Central_LC_Start_Homogeneity	How many pixels in the window are of a similar label as the pixel central at the start of the period
Central_LC_End	The LC legend class ID (Table 2) of the sample plot center pixel at the end of the period.
Central_LC_End_Homogeneity	How many pixels in the window are of a similar label as the pixel central at the end of the period
Central_LC_Change	Presence/absence of LC change observed at the central pixel of the sample plot (if applicable). Change is defined as a LC transition from one year to the other, confirmed for at least one year. Seasonal variations of the LC are not considered a change. Multiple LC changes in the period can be recorded in the field “comments”.
Central_Year_Change	Year when LC change was observed (if the field Central LC change was filled). This field is completed for each LC change observed in the previous field.
Level of certainty	Level of certainty (certain, reasonable, doubtful) associated with the interpretation of the expert
Comments	Comments given by the expert to explain/detail its interpretation (e.g. for indication of why the labelling was not successful, or to give the potential change phenomenon observed on the VHR imagery)

3.5 Graphical interface for image interpretation

GEE Time Series Explorer allows connecting a sampling scheme to Google Earth and Google Earth Engine to access various HR and VHR imagery and temporal NDVI profiles derived from Landsat and S2 (Figure 5). The interpretation tool displays an interface in QGIS so that experts can indicate the LC class identified and their confidence level in a form. A similar form is created for the set of change validation footprints where one can indicate the change, certainty, homogeneity, year of change, and add comments (Figure 6). Note that time series of vegetation profiles are generated for Sentinel 2 and Landsat for each year of the period.

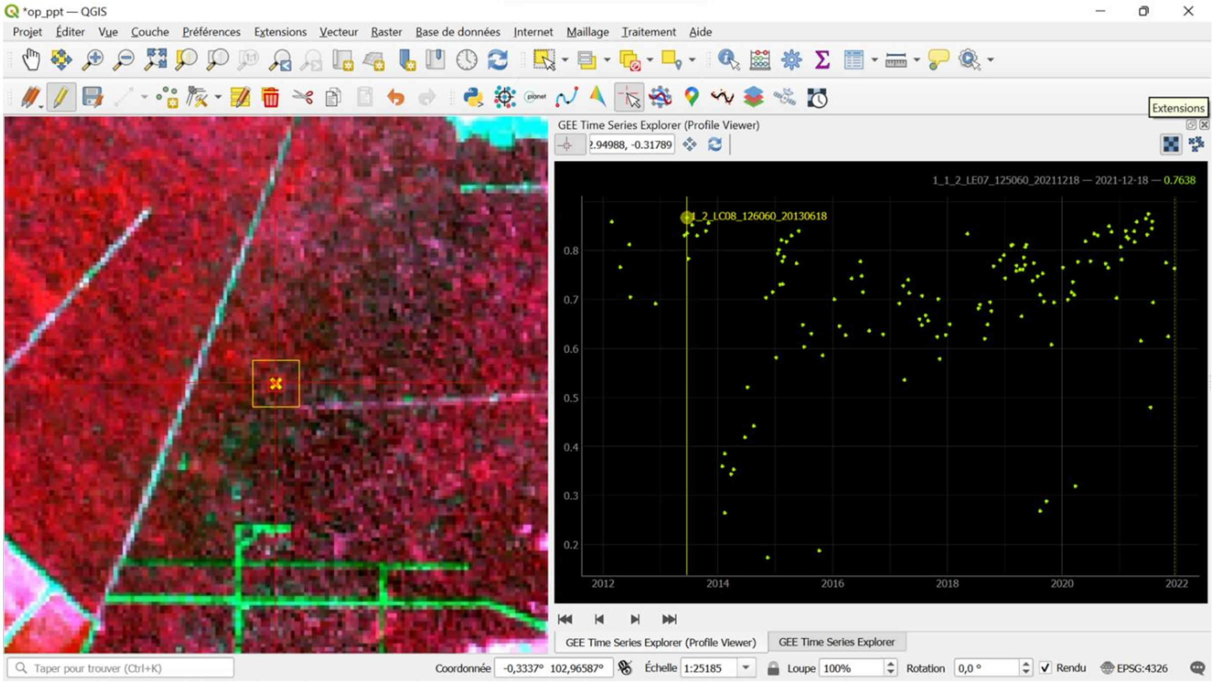

Figure 5. Interface of the Google Earth Engine Time Series Explorer, which displays all available Landsat and Sentinel satellite images. Example of a tropical rainforest in 2013.

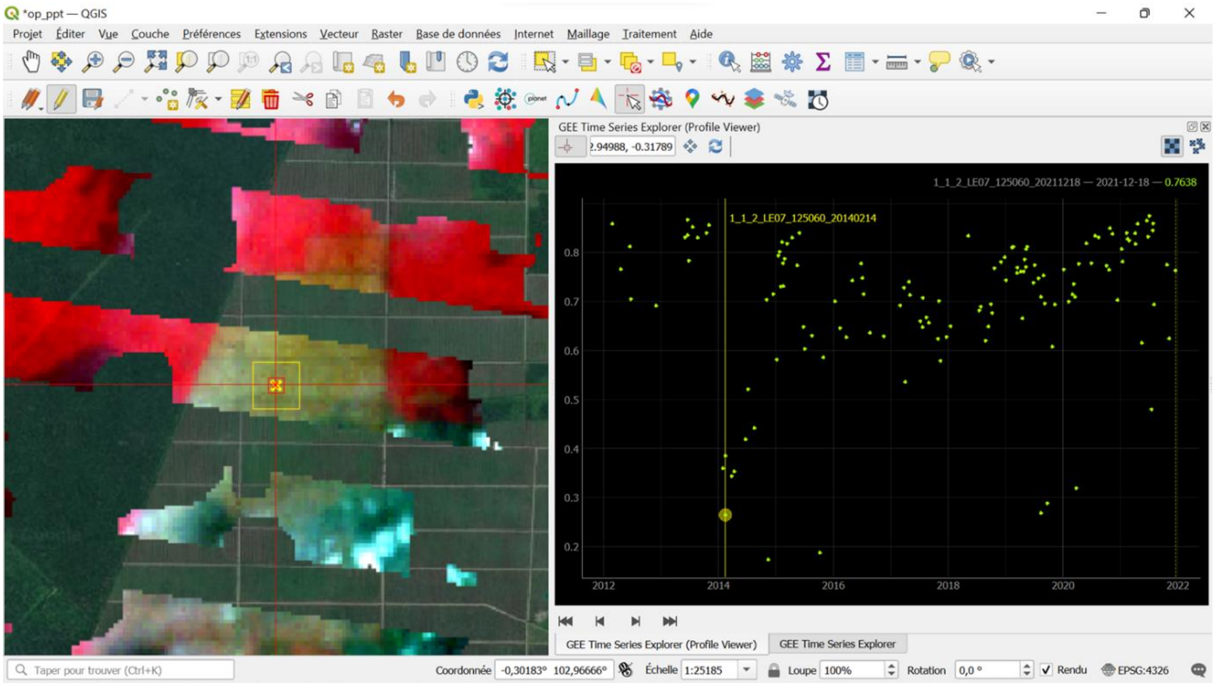

Figure 6. Interface of the Google Earth Engine Time Series Explorer, which displays all available Landsat and Sentinel satellite images. Example of deforestation of the tropical rainforest in early 2014, to lead to a future oil palm plantation.

Once the interpretation of all samples is first achieved by an interpreter, a confidence-building procedure takes place, with a review of all samples by a second interpreter. Samples receiving different labels from the interpreters are discussed. Some may be relabelled or discarded if interpreters agree or not. A special focus is put on samples with “doubtful” and “totally uncertain” confidence levels, for which each data source is reviewed to potentially correct the LC class and/or the confidence level.

4 Benchmarking with other existing products

The data collected during Phase 1 were mostly global and regional HR LC because these datasets are of global interest, and usually available online or upon request with few local HR LC that have been collected upon contacting national mapping authorities. Currently in Phase 2 (as for June 2024), a similar trend is observed where mainly new global datasets were released in the recent few years. Notably, the objectives of the scholars and/or organizations producing HR LC are currently more focused on time-series and/or LC change products. This fact is mainly due to the longer time spans of the ongoing NASA's Landsat and ESA's Sentinel mission, as well as the availability of powerful computing cloud services (e.g., Google Earth Engine, Microsoft Planetary Computer, etc.).

All of the newly collected and recently released HR LC products have global coverage, where the only current exception is GLanCE V001 as the data owners released only partial products covering Europe, North and South America with the intention of releasing the rest of the continents (at the moment of writing June 2024). Table 11 shows the newly collected HR LCs as well as the HR LCs collected during Phase 1, which are indicated in the "Phase 1 Dataset" column. In the cases where the dataset series used in Phase 1 have new releases, then new versions are marked in red colour font.

For the dynamic datasets the following notation for the year of the products is used below in the table:

- When maps are produced for distinct years, the years are separated with comma. For example, "1985,1990" would mean that there is a single map for the 1985 and one for 1990.
- When maps are produced for a period of years, the contained period will be noted with dash between the start and end period. For example, "2000-2005" would mean that there is a single map produced for the period between year 2000 and 2005.
- When maps are produced for a period of years at a certain time-step, the notation will be again using a dash sign between the start and end period, but in brackets will be clarified the time-step period. For example, "2000-2005 (for every year)" would mean that there are maps produced for each year in the period between the year 2000 and 2005.

Table 11. Existing HR LC for benchmarking

Name of LC map	Producer	Resolution	Year	Spatial coverage	Phase 1 Dataset
GFC* [19]	JRC	10 m	2020	Global	No
GLC_FCS30/D*[20],[21]	Aerospace Information Research Institute, Chinese Academy of Sciences	30 m	1985,1990,1995, 2000-2022 (map for every year)	Global	No
GISD30* [22]	Aerospace Information Research Institute, Chinese Academy of Sciences	30 m	1985–2020	Global	No
GWL_FCS30/D**[23], [24]	Aerospace Information Research Institute, Chinese Academy of Sciences	30 m	2000-2022 (map for every year)	Global	No
Dynamic World* [25]	World Resources Institute Google	10 m	From 27/06/2015	Global	No
LULC Annual v2* [26]	Impact Observatory, Microsoft, and Esri	10 m	2017-2023 (map for every year)	Global	No
WorldCover*[27] [28]	ESA	10 m	2020,2021	Global	No
GHS-BUILT-S R2023A* [29] [30]	JRC	10 m	2018	Global	No
GLanCE V001* [31],[32]	BU/EE/NASA ES/USGS EROS	30 m	2001-2019 (map for every year)	Global (currently	No

Name of LC map	Producer	Resolution	Year	Spatial coverage	Phase 1 Dataset
				available Europe, North & South America)	
MapBiomas Collection 10m Beta*	SEEG and Climate Observatory	10m	2016-2022 (map for every year)	Regional Brazilian Amazon	No
GL30 **[33],[34],[35]	NGCC	30 m	2000, 2010,2019	Global	Yes
FROM-GLC[36],[37],[38]	Tsinghua University	30 m	2010, 2015,2017	Global	Yes
GUF ***	DLR	12 m	2011	Global	Yes
GHS BU S1 *[39],[40]	JRC	20 m	2016	Global	Yes
GHS BU LDS [39],[40]	JRC	38 m	1990, 2000, 2014	Global	Yes
GSW Yearly History* [41]	JRC	30 m	1984-2021 (map for every year)	Global	Yes
FNF * [42]	JAXA-EORC	25 m	2007 - 2010 (map for every year) 2015 - 2020 (map for every year)	Global	Yes
Tree canopy cover * [43]	Hansen/UMD/Google/USGS/NASA	30 m	2000	Global	Yes
WSF*[44][82], [45]	DLR	10 m	2015, 2019	Global	Yes
TerraClass * [46]	INPE and EMBRAPA	30 m	2004, 2008, 2010, 2012, 2014	Regional Brazilian Amazon	Yes
MapBiomas Collection 9*	SEEG and Climate Observatory	30m	1985 - 2023 (map for every year)	Regional Brazilian Amazon	Yes
ESA-DUE-GlobPermafrost* [47]	ESA-DUE-GlobPermafrost	20 m	2016-2018 (map for period)	Regional Siberia	Yes
CCI Prototype Africa* [48]	CCI Land Cover team	20 m	2016	Regional Africa	Yes
AUE*	NYU Urban Expansion Program	30 m	1984-2015 some years	Cities in Nigeria, Brazil, Uganda, Sudan, Ghana, Mali,	Yes

*Freely available

**Available upon agreement

*** Freely available upon request for non-commercial and scientific purpose

Most of the HR LC products gathered in Phase 2 are multitemporal, providing maps mainly on a yearly basis, except the Dynamic World, which provides LC classification for each available Sentinel-2 scene with a cloud cover ≤35% starting from 27/06/2015. Another exception is the GISD30 dataset, which provides a map with temporal coverage from 1985 to 2020, using different pixel values to represent different year periods.

Some HR LCs from Phase 1 (including GSW Yearly History, FNF, and MapBiomas) have been updated to cover a larger spatial extent. Additionally, MapBiomas has introduced a new collection with a 10-meter resolution, called MapBiomas Collection 10m Beta.

4.1 Accuracy of the existing HR LC

For all of the Phase 2 products are reported accuracy metrics, mainly in terms of Overall Accuracy (OA), even if in several cases more metrics were reported by the data producers. In the table below it was decided to report only OA for consistency; where and when the evaluation is done on a specific epoch or location it was additionally noted. Note that "Overall" in the "year" column indicates that the accuracy is reported by the authors without specifying the year.

Table 12. Accuracy of the existing HR LC

Map	Year	Accuracy
GFC*	2020	OA=76.6% [19]
GLC_FCS30D*	1985,1990, 1995, 2000-2022 (map for every year)	OA_{2015}=82.5% [20], OA_{2020}=80.88% [21], $OA_{(2006,2009,2012,2015,2018)}$=82.11%, 81.99%, 81.97%, 81.82%, 81.64% [22]
GISD30*	1985, 1985–1990, 1990–1995, 1995–2000, 2000–2005, 2005–2010, 2010–2015 and 2015–2020	OA=90.1% [23]
GWL_FCS30*	2020	OA=87.7% [24], OA=86.95%[24]
Dynamic World*	Overall	OA=73.8% [25]
LULC Annual V2*	Overall	OA=85% [26]
WorldCover*	2020, 2021	OA_{2020}=74.4%, OA_{2021}=76.7%
GHS-BUILT-S R2023A*	2018	OA=90.5%
GLanCE V001*	2010	$OA_{NorthAmerica}$=77.0%±2.0%
GL30	2017	[-]
GL30	2010	OA= 80% [33]
GL30	2000	OA=78.6% [33]
FROM-GLC	2017	OA=72.76% [36]
FROM-GLC	2015	OA=70.17% [37]
FROM-GLC	2010	OA=67.08% [49]
GUF	2011	OA=90.23% [50]
GHS BU S1	2016	Kappa=0.31-0.48 median kappa [39]
GHS BU LDS	2014	OA=89% - 96% Balanced accuracy= 67%-77%
GHS BU LDS	2000	[-]
GHS BU LDS	1990	[-]
GSW_Yearly_History	2018	UA=99%; PA=96% [41]
GSW_Yearly_History	2015	UA=99%; PA=96% [41]
GSW_Yearly_History	2010	UA=99%; PA=96% [41]
GSW_Yearly_History	2005	UA=99%; PA=96% [41]
GSW_Yearly_History	2000	UA=99%; PA=96% [41]
GSW_Yearly_History	1995	UA=99%; PA=96% [41]
GSW_Yearly_History	1990	UA=99%; PA=96% [41]
FNF	2017	[-]
FNF	2015, 2016	[-]
FNF	2010	OA=87.14%-91.13% [42]
FNF	2007	OA=85.19%-91.49% [42]
Tree canopy cover	2000	[-]
WSF	2015	Balanced accuracy =83.27 [44]

WSF	2019	[-]
TerraClass	2014	OA=86% [51]
TerraClass	2010, 2012	[-]
TerraClass	2004, 2008	OA=76,64% [46]
MapBiomas Collection 9+	Overall	Mean Global Accuracy (Level 1) = 93.1% Mean Global Accuracy (Level 2) = 89.8% Mean Global Accuracy (Level 3) = 89.8%
MapBiomas Collection 10m Beta*	2016-2022	[-]
ESA-DUE-GlobPermafrost	2016-2018	OA=83% [47]
CCI Prototype Africa	2016	OA_{Kenya}=56%; OA_{Gabon}=91%; $OA_{IvoryCoast}$=47%; $OA_{SouthAfrica}$=44% [48]
Atlas of Urban Expansion	All years	[-]

*New compared to Phase 1

+Updated compared to Phase 1

The CCI Prototype Africa dataset used in Phase 1 will not be used in Phase 2 due to its limited validation across countries and its low overall accuracy in three out of the four countries where it was tested.

4.2 Legend of existing land cover

The legends of the LC map products were reported by their authors in all Phase 2 gathered products (Table 13, Table 14, Table 15, Table 16, Table 17, Table 18, Table 19, Table 20, Table 21). In several cases the authors are communicating when there is more than one level of legend, except the cases of binary maps (e.g., GHS-BUILT-S R2023A).

Most of the multitemporal map products have the same class legend among the epochs. In most cases the different products have utilized different LC classes, where some authors provided homogenization tables (e.g., GLC_FCS30D with GlobeLand30).

The legends of the LC map products included in Phase 1 are also listed (Table 24, Table 25, Table 26, Table 27, Table 28, Table 29, Table 30, Table 31, Table 33, Table 34, Table 35, Table 36, **Error! Reference source not found.**, Table 37). In addition, MapBiomas, which was utilized in the Phase 1, has been updated, resulting in corresponding updates to the legend (Table 22 and Table 23). FNF, which was also utilized in Phase 1, has been updated, resulting in corresponding updates to the legend listed in Table 32.

Table 13. GFC legend and description of all products

Class code	Class	Description
1	Forest	Land spanning more than 0.5 hectares with trees higher than 5 meters and a canopy cover of more than 10 %, or trees able to reach those thresholds in situ, excluding land that is predominantly under agricultural or urban land use [19].

Table 14. GLC_FCS30D legend and description of all products

Class code	Class	Description
10	Rainfed cropland	[-]
11	Herbaceous cover cropland	[-]
12	Tree or shrub cover (Orchard) cropland	[-]
20	Irrigated cropland	[-]
51	Open evergreen broadleaved forest	[-]
52	Closed evergreen broadleaved forest	
61	Open deciduous broadleaved forest (0.15<fc<0.4)	[-]
62	Closed deciduous broadleaved forest (fc>0.4)	[-]
71	Open evergreen needle-leaved forest (0.15< fc <0.4)	[-]
72	Closed evergreen needle-leaved forest (fc >0.4)	[-]
81	Open deciduous needle-leaved forest (0.15< fc <0.4)	[-]

82	Closed deciduous needle-leaved forest (fc >0.4)	[-]
91	Open mixed leaf forest (broadleaved and needle-leaved)	[-]
92	Closed mixed leaf forest (broadleaved and needle-leaved)	
120	Shrubland	[-]
121	Evergreen shrubland	[-]
122	Deciduous shrubland	[-]
130	Grassland	[-]
140	Lichens and mosses	[-]
150	Sparse vegetation (fc<0.15)	[-]
152	Sparse shrubland (fc<0.15)	[-]
153	Sparse herbaceous (fc<0.15)	[-]
181	Swamp	[-]
182	Marsh	[-]
183	Flooded flat	[-]
184	Saline	[-]
185	Mangrove	[-]
186	Salt marsh	[-]
187	Tidal flat	[-]
190	Impervious surfaces	[-]
200	Bare areas	[-]
201	Consolidated bare areas	[-]
202	Unconsolidated bare areas	[-]
210	Water body	[-]
220	Permanent ice and snow	[-]

Table 15. GISD30 legend and description of all products

Class code	Class	Description
0	Pervious	Pervious surface before 1985
1	Impervious	Impervious surface before 1985
2	Impervious	Expanded impervious surfaces in the period 1985–1990
3	Impervious	Expanded impervious surfaces in the period 1990–1995
4	Impervious	Expanded impervious surfaces in the period 1995–2000
5	Impervious	Expanded impervious surfaces in the period 2000–2005
6	Impervious	Expanded impervious surfaces in the period 2005–2010
7	Impervious	Expanded impervious surfaces in the period 2010–2015
8	Impervious	Expanded impervious surfaces in the period 2015–2020

Table 16. GWL_FCS30D legend and description of all products

Class code	Class	Description
0	Non-wetland	[-]
181	Swamp	The forest or shrubs which grow in the inland freshwater
182	Marsh	Herbaceous vegetation (grasses, herbs and low shrubs) grows in the freshwater
183	Flooded flat	The non-vegetated flooded areas along the rivers and lakes
184	Saline	Characterized by saline soils and halophytic (salt tolerant) plant species along saline lakes
185	Mangrove forest	The forest or shrubs which grow in the coastal blackish or saline water
186	Salt marsh	Herbaceous vegetation (grasses, herbs and low shrubs) in the upper coastal intertidal zone
187	Tidal flat	The tidal flooded zones between the coastal high and low tide levels including mudflats and sandflats

Table 17. Dynamic World legend and description of all products

Class code	Class	Description
0	Water	Water is present in the image. Contains little-to-no sparse vegetation, no rock outcrop, and no built-up features like docks. Does not include land that can or has

		previously been covered by water.
1	Trees	Any significant clustering of dense vegetation, typically with a closed or dense canopy. Taller and darker than surrounding vegetation (if surrounded by other vegetation).
2	Grass	Open areas covered in homogenous grasses with little to no taller vegetation. Other homogenous areas of grass-like vegetation (blade-type leaves) that appear different from trees and shrubland. Wild cereals and grasses with no obvious human plotting (i.e. not a structured field).
3	Flooded vegetation	Areas of any type of vegetation with obvious intermixing of water. Do not assume an area is flooded if flooding is observed in another image. Seasonally flooded areas that are a mix of grass/shrub/trees/bare ground.
4	Crops	Human planted/plotted cereals, grasses, and crops.
5	Shrub & Scrub	Mix of small clusters of plants or individual plants dispersed on a landscape that shows exposed soil and rock. Scrub-filled clearings within dense forests that are clearly not taller than trees. Appear grayer/browner due to less dense leaf cover.
6	Built area	Clusters of human-made structures or individual very large human-made structures. Contained industrial, commercial, and private building, and the associated parking lots. A mixture of residential buildings, streets, lawns, trees, isolated residential structures or buildings surrounded by vegetative land covers. Major road and rail networks outside of the predominant residential areas. Large homogeneous impervious surfaces, including parking structures, large office buildings, and residential housing developments containing clusters of cul-de-sacs.
7	Bare ground	Areas of rock or soil containing very sparse to no vegetation. Large areas of sand and deserts with no to little vegetation. Large individual or dense networks of dirt roads.
8	Snow & Ice	Large homogenous areas of thick snow or ice, typically only in mountain areas or highest latitudes. Large homogenous areas of snowfall.

Table 18. LULC Annual V2 legend and description of all products

Class code	Class	Description
1	Water	Areas where water was predominantly present throughout the year; may not cover areas with sporadic or ephemeral water; contains little to no sparse vegetation, no rock outcrop nor built up features like docks; examples: rivers, ponds, lakes, oceans, flooded salt plains.
2	Trees	Any significant clustering of tall (~5 m or higher) dense vegetation, typically with a closed or dense canopy; examples: wooded vegetation, clusters of dense tall vegetation within savannas, plantations, swamp or mangroves (dense/tall vegetation with ephemeral water or canopy too thick to detect water underneath).
4	Flooded vegetation	Areas of any type of vegetation with obvious intermixing of water throughout a majority of the year; seasonally flooded area that is a mix of grass / shrub / trees / bare ground; examples: flooded mangroves, emergent vegetation, rice paddies and other heavily irrigated and inundated agriculture.
5	Crops	Human planted / plotted cereals, grasses, and crops not at tree height; examples: corn, wheat, soy, fallow plots of structured land.
7	Built area	Human made structures; major road and rail networks; large homogenous impervious surfaces including parking structures, office buildings and residential housing; examples: houses, dense villages / towns / cities, paved roads, asphalt.
8	Bare ground	Areas of rock or soil with very sparse to no vegetation for the entire year; large areas of sand and deserts with no to little vegetation; examples: exposed rock or soil, desert and sand dunes, dry salt flats / pans, dried lake beds, mines.
9	Snow/ice	Large homogenous areas of permanent snow or ice, typically only in mountain areas or highest latitudes; examples: glaciers, permanent snowpack, snow fields.
10	Clouds	No land cover information due to persistent cloud cover.
11	Rangeland	Any area of low, non-flooded vegetation with very little-to-no taller (~15m or higher) vegetation, homogeneous or heterogeneous, containing any degree of the following: wild cereals and grasses with no obvious human plotting (i.e. not a plotted field); mix of small clusters of plants or single plants dispersed on a landscape that shows exposed soil or rock; clearings of homogeneous grasses; scrub-filled clearings within dense forests that are clearly not taller than trees; examples: moderate to sparse cover of bushes, shrubs and tufts of grass, savannas with very sparse grasses, trees or other plants.

Table 19. WorldCover legend and description of all products

Class code	Class	Description
10	Tree cover	This class includes any geographic area dominated by trees with a cover of 10% or more. Other land cover classes (shrubs and/or herbs in the understorey, built-up, permanent water bodies, …) can be present below the canopy, even with a density higher than trees. Areas planted with trees for afforestation purposes and plantations (e.g. oil palm, olive trees) are included in this class. This class also includes tree covered areas seasonally or permanently flooded with fresh water except for mangroves.
20	Shrubland	This class includes any geographic area dominated by natural shrubs having a cover of 10% or more. Shrubs are defined as woody perennial plants with persistent and woody stems and without any defined main stem being less than 5 m tall. Trees can be present in scattered form if their cover is less than 10%. Herbaceous plants can also be present at any density. The shrub foliage can be either evergreen or deciduous
30	Grassland	This class includes any geographic area dominated by natural herbaceous plants (Plants without persistent stem or shoots above ground and lacking definite firm structure): (grasslands, prairies, steppes, savannahs, pastures) with a cover of 10% or more, irrespective of different human and/or animal activities, such as: grazing, selective fire management etc. Woody plants (trees and/or shrubs) can be present assuming their cover is less than 10%. It may also contain uncultivated cropland areas (without harvest/ bare soil period) in the reference year
40	Cropland	Land covered with annual cropland that is sowed/planted and harvestable at least once within the 12 months after the sowing/planting date. The annual cropland produces an herbaceous cover and is sometimes combined with some tree or woody vegetation. Note that perennial woody crops will be classified as the appropriate tree cover or shrub land cover type. Greenhouses are considered as built-up.
50	Built-up	Land covered by buildings, roads and other man-made structures such as railroads. Buildings include both residential and industrial building. Urban green (parks, sport facilities) is not included in this class. Waste dump deposits and extraction sites are considered as bare.
60	Bare/sparse vegetation	Lands with exposed soil, sand, or rocks and never has more than 10 % vegetated cover during any time of the year
70	Snow and Ice	This class includes any geographic area covered by snow or glaciers persistently
80	Permanent water bodies	This class includes any geographic area covered for most of the year (more than 9 months) by water bodies: lakes, reservoirs, and rivers. Can be either fresh or salt-water bodies. In some cases the water can be frozen for part of the year (less than 9 months).
90	Herbaceous wetland	Land dominated by natural herbaceous vegetation (cover of 10% or more) that is permanently or regularly flooded by fresh, brackish or salt water. It excludes unvegetated sediment (see 60), swamp forests (classified as tree cover) and mangroves see 95)
95	Mangrove	Taxonomically diverse, salt-tolerant tree and other plant species which thrive in intertidal zones of sheltered tropical shores, "overwash" islands, and estuaries.
100	Moss and lichen	Land covered with lichens and/or mosses. Lichens are composite organisms formed from the symbiotic association of fungi and algae. Mosses contain photo-autotrophic land plants without true leaves, stems, roots but with leaf-and stemlike organs.

Table 20. GHS-BUILT-S R2023A legend and description of all products

Class value	Class	Description
[0-100]	Range built-up	Values are expressed as 8bit integers (uint8) and represent the amount of square metres of built-up surface in the cell.
255	NoData	[-]

Table 21. GLanCE Level 1 legend and description of all products

Class value	Class	Description
1	Water	Areas covered with water throughout the year: streams, canals, lakes, reservoirs, and oceans.

2	Ice/Snow	Land areas where snow and ice cover is greater than 50% throughout the year.
3	Developed	Areas of intensive use; land covered with structures, including any land functionally related to developed/built-up activity.
4	Barren/Sparsely Vegetated	Land consists of natural occurrences of soils, sand, or rocks where less than 10% of the area is vegetated.
5	Tree Cover	Land where the tree cover is greater than 30%. Note that cleared trees (i.e., clear-cuts) are mapped according to current cover (e.g., barren/sparsely vegetated, shrubs, or grasses).
6	Shrublands	Land with less than 30% tree cover, where total vegetation cover exceeds 10% and shrub cover is greater than 10%.
7	Herbaceous	Land covered by herbaceous cover. Total vegetation cover exceeds 10%, tree cover is less than 30%, and shrubs comprise less than 10% of the area.

Table 22. MapBiomas legend and description for Collection 9

Class value	Class	Description	Level
1	Forest	[-]	1
3	Forest Formation	[-]	2
4	Savanna Formation	[-]	2
5	Mangrove	[-]	2
6	Floodable Forest	[-]	2
49	Wooded Sandbank Vegetation	[-]	2
10	Herbaceous and Shrubby Vegetation	[-]	1
11	Wetland	[-]	2
12	Grassland	[-]	2
32	Hypersaline Tidal Flat	[-]	2
29	Rocky Outcrop	[-]	2
50	Herbaceous Sandbank Vegetation	[-]	2
14	Farming	[-]	1
15	Pasture	[-]	2
18	Agriculture	[-]	2
19	Temporary Crop	[-]	3
39	Soybean	[-]	4
20	Sugar cane	[-]	4
40	Rice	[-]	4
62	Cotton (beta)	[-]	4
41	Other Temporary Crops	[-]	4
36	Perennial Crop	[-]	3

Class value	Class	Description	Level
46	Coffee	[-]	4
47	Citrus	[-]	4
35	Palm Oil	[-]	4
48	Other Perennial Crops	[-]	4
9	Forest Plantation	[-]	2
21	Mosaic of Uses	[-]	2
22	Non vegetated area	[-]	1
23	Beach, Dune and Sand Spot	[-]	2
24	Urban Area	[-]	2
30	Mining	[-]	2
25	Other non Vegetated Areas	[-]	2
26	Water	[-]	1
33	River, Lake and Ocean	[-]	2
31	Aquaculture	[-]	2
27	Not Observed	[-]	1

Table 23. MapBiomas legend and description for Collection 10M Beta

Class value	Class	Description	Level
1	Forest	[-]	1
3	Forest Formation	[-]	2
4	Savanna Formation	[-]	2
5	Mangrove	[-]	2
49	Wooded Sandbank Vegetation	[-]	2
10	Non Forest Natural Formation	[-]	1
11	Wetland	[-]	2
12	Grassland	[-]	2
32	Salt Flat	[-]	2
29	Rocky Outcrop	[-]	2
50	Herbaceous Sandbank Vegetation	[-]	2
13	Other non Forest	[-]	2

Class value	Class	Description	Level
	Formations		
14	Farming	[-]	1
15	Pasture	[-]	2
18	Agriculture	[-]	2
19	Temporary Crop	[-]	3
39	Soybean	[-]	4
20	Sugar cane	[-]	4
40	Rice	[-]	4
62	Cotton (beta)	[-]	4
41	Other Temporary Crops	[-]	4
36	Perennial Crop	[-]	3
46	Coffee	[-]	4
47	Citrus	[-]	4
48	Other Perennial Crops	[-]	4
9	Forest Plantation	[-]	2
21	Mosaic of Uses	[-]	2
22	Non vegetated area	[-]	1
23	Beach, Dune and Sand Spot	[-]	2
24	Urban Area	[-]	2
30	Mining	[-]	2
25	Other non Vegetated Areas	[-]	2
26	Water	[-]	1
33	River, Lake and Ocean	[-]	2
31	Aquaculture	[-]	2
27	Not Observed	[-]	1

Table 24. GL30 legend and description for all products

Class value	Class	Description
10	Cultivated land	Lands used for agriculture, horticulture, and gardens, including paddy fields, irrigated and dry farmland, vegetation and fruit gardens, etc.
20	Forest	Lands covered with trees, with vegetation cover over 30%, including deciduous and coniferous forests, and sparse woodland with cover 10 - 30%, etc.

Class value	Class	Description
30	Grassland	Lands covered by natural grass with cover over 10%, etc.
40	Shrubland	Lands covered with shrubs with cover over 30%, including deciduous and evergreen shrubs, and desert steppe with cover over 10%, etc.
50	Wetland	Lands covered with wetland plants and water bodies, including inland marsh, lake marsh, river floodplain wetland, forest/shrub wetland, peat bogs, mangrove, and salt marsh, etc.
60	Water bodies	Water bodies in the land area, including river, lake, reservoir, fish pond, etc.
70	Tundra	Lands covered by lichen, moss, hardy perennial herb and shrubs in the polar regions, including shrub tundra, herbaceous tundra, wet tundra and barren tundra, etc.
80	Artificial surfaces	Lands modified by human activities, including all kinds of habitation, industrial and mining area, transportation facilities, and interior urban green zones and water bodies, etc.
90	Bareland	Lands with vegetation cover lower than 10%, including desert, sandy fields, Gobi, bare rocks, saline and alkaline lands, etc.
100	Permanent snow and ice	Lands covered by permanent snow, glacier and ice cap.

Table 25. FROM-GLC legend and description for products for 2010

Class value	Class	Description
10	Croplands	This type of land has clear traits of intensive human activity. It varies a lot from bare field, seeding, crop growing to harvesting. It can be easily identified if edges or textures are visible with sufficiently large land parcels. Fruit trees are classified into forests. Bare field is classified into bare land. Pasture could be transitional from croplands to natural grasslands.
11	Rice fields	Land for rice cultivation.
12	Greenhouse farming	Land with plastic foam or grass roof protection with distinguishing spectral properties.
13	Other croplands	This category includes arable and tillage land.
20	Forest	Trees observable in the landscape from the images. Forest has a distinct canopy texture on TM images.
21	Broadleaf forests	Usually higher reflectivity than conifer species in the near infrared (NIR) spectral band. Shaded and sunlit sides less contrast.
22	Needleleaf forests	Lower reflectivity than broadleaf trees in the NIR band.
23	Mixed forests	Neither coniferous nor broadleaf trees dominate in a mixed forest stand.
24	Orchards	Parcels planted with fruit trees or shrubs: single or mixed fruit species, fruit trees associated with permanently grassed surfaces.
30	Grasslands	-
31	Pastures	Grasslands for grazing.
32	Other grasslands	Natural grasslands identifiable.
40	Shurblands	Shrub cover identifiable in the image. Has a texture finer than tree

Class value	Class	Description
		canopies but coarser than grasslands.
50	Wetlands	Although wetland is defined in the RAMSAR convention to maximize wetland areas, we intend to include only marshland with distinctively high reflectivity in the NIR band. Low relief areas with perched bogs, playas, and patholes may also be included depending on the season of image acquisition time. Forested wetland is not included here as it cannot be well identified from TM images.
51	Marshland	Aquatic and hydrophytic herbaceous plants observable from the image as non-water cover.
52	Mudflats	Generally unvegetated expanses of mud, sand or rock lying between high and low water lines.
60	Waterbodies	All inland waterbodies with >3 pixels in width or 8 pixel—8 pixel (6 ha) in area. Patches of fish ponds are included in this category. Spectral characteristics vary widely and the waterbody change in area with season.
61	Lake	Natural waterbodies.
62	Reservoir/ Pond	Dammed waterbodies.
63	River	Natural or artificial water-courses serving as water drainage channels. Minimum width for inclusion 3 pixels.
64	Ocean	Salinity water.
70	Tundra	Located at high mountains above tree line and high latitude regions with low height vegetation. The growing season is between 1 and 2 months.
71	Shrub and Brush Tundra	Dominated by low shrubs with grasses, lichens, and mosses at the background.
72	Herbaceous Tundra	Dominated by various sedges, grasses, forbs, lichens, and mosses, all of which lack woody stems.
80	Impervious	Primarily based on artificial cover such as asphalts, concrete, sand and stone, bricks, glasses, and other cover materials.
81	Impervious high albedo	Impervious road cover with high albedo materials (e.g. concrete, cement).
82	Impervious low albedo	Impervious roof tops covered by low albedo materials (e.g. asphalts, black shingles).
90	Barren Land	Vegetation is hardly observable but dominated by exposed soil, sand, gravel, and rock backgrounds.
91	Dry salt flats	Dry salt flats occurring on the flat floored bottoms of interior desert basins.
92	Sandy areas	Sandy areas are composed primarily of dunes accumulations of sand transported by wind.
93	Bare exposed rock	Gravel land and bare rocks.
94	Bare herbaceous croplands	Just harvested, fallow land and all other types of land not covered by vegetation such as lake bottoms in dry season.
95	Dry lake/river bottoms	Other types of land not covered by vegetation such as lake/river bottoms in dry season.
96	Other barren lands	All other types of land not covered by vegetation.

Class value	Class	Description
100	Snow and ice	Distributed in the polar areas and high mountains.
110	Snow	Lands under perennial or non-perennial snow over.
120	Ice	Lands under perennial or non-perennial ice over.

Table 26. FROM-GLC legend and description for products for 2015

Class value	Class	Description
10	Cropland	[-]
11	Rice paddy	[-]
12	Greenhouse	[-]
12	Cloud	[-]
13	Other	[-]
14	Orchard	[-]
15	Bare farmland	[-]
20	Forest	[-]
21	Broadleaf, leaf-on	[-]
22	Broadleaf, leaf-off	[-]
23	Needleleaf, leaf-on	[-]
24	Needleleaf, leaf-off	[-]
25	Mixed leaf, leaf-on	[-]
26	Mixed leaf, leaf-off	[-]
30	Grassland	[-]
31	Pasture	[-]
32	Natural grassland	[-]
33	Grassland, leaf-off	[-]
40	Shrubland	[-]
41	Shrubland, leaf-on	[-]
42	Shrubland, leaf-off	[-]
50	Wetland	[-]
51	Marshland	[-]
52	Mudflat	[-]

Class value	Class	Description
53	Marshland, leaf-off	[-]
60	Water	[-]
70	Tundra	[-]
71	Shrub and brush tundra	[-]
72	Herbaceous tundra	[-]
80	Impervious surface	[-]
90	Bareland	[-]
92	Sand	[-]
100	Snow/Ice	[-]
110	Snow	[-]
120	Ice	[-]

Table 27. FROM-GLC legend and description for products for 2017

Class value	Class	Description
10	Cropland	[-]
20	Forest	[-]
30	Grassland	[-]
40	Shrubland	[-]
50	Wetland	[-]
60	Water	[-]
70	Tundra	[-]
80	Impervious surface	[-]
90	Bareland	[-]
100	Snow/Ice	[-]

Table 28. GUF legend and description for all products

Class value	Class	Description
0	Non-built-up areas	Other
255	Built-up areas	A region featuring man-made building structures with a vertical component
188	NoData	[-]

Table 29. GHS BU legend and description for all products

Class value	Class	Description
0	Non-built-up areas	[-]
1	Built-up areas	BU areas are the spatial generalization of the notion of building defined as: 'areas (spatial units) where buildings can be found'. The working definition of BU structure (building)used in this experiment setting is as follows: 'buildings are enclosed constructions above ground which are intended or used for the shelter of humans, animals, things or for the production of economic goods and that refer to any structure constructed or erected on its site. The GHSL notion of BU structure is more inclusive, accepting to describe also structures belonging to temporary human settlements as refugee or internal displaced people (IDP) camp. The GHSL repository includes also BU areas falling in the 'slum' or informal settlement concept: the area of a city characterized by sub-stan-dard housing and squalor and lacking in tenure security, also called 'shanty town', 'squatter settlement' and similar.

Table 30. GSW Yearly History legend and description for all products

Class value	Class	Description
0	No observations	[-]
1	Not water	Other
2	Seasonal water	Seasonal water surface is underwater for less than 12 months of the year
3	Permanent water	Permanent water surface is underwater throughout the year

Table 31. FNF 3-class legend and description for all products

Class value	Class	Description
1	Forest	"Forest" is defined as the tree covered land with the area larger than .5 ha and canopy cover over 1 %
2	Non-forest	Other
3	Water	[-]

Table 32. FNF 4-class legend and description for all products

Class value	Class	Description
1	Dense Forest	Forest (>90% crown cover)
2	Non-dense Forest	Forest (10-90% crown cover)
3	Non-forest	
4	Water	[-]

Table 33. Tree Cover legend and description for all products

Class value	Class	Description
0-100 integer	Tree Cover	Defined as canopy closure for all vegetation taller than 5m in height. Encoded as a percentage per output grid cell, in the range –1

Table 34. WSF legend and description for all products

Class value	Class	Description
0	Non-settlements	[-]
255	Settlements	[-]

Table 35. TerraClass legend and description for all products

Class value	Class	Description
1	Annual Crops	Extensive areas with predominance of annual crops, specially grains, highly technological such as certified seeds, enriched soil, chemicals, fertilizers, mechanization among other resources.
2	Non-observed area	Areas not possible to be interpreted due to clouds or cloud shade at the moment of the satellite overpass or recently burned areas.
3	Urban area	Population concentration forming small inhabited places, villages and cities that present differentiated infrastructure from the rural areas with street design and higher density of dwellings such as houses, buildings and other public spaces
4	Deforestation 28	Areas recently deforested covered by soil, shrubs, herbage and felled trees with no defined land use at this stage, defined as areas that were mapped by PRODES project as deforested in 28
5	Forest	-
6	Water	-
7	Mining	Areas of mineral extraction with the presence of bare soil and deforestation in the proximity of water bodies.
8	Mosaic of Uses	Characterized by land cover units that, due to the spatial resolution of the satellite images, cannot be broken down further into specific components. For example, this classification might include family agriculture practiced in conjunction with the traditional cattle raising.
9	Not Forest	-
10	Others	Areas not encompassed by other categories such as rocky or mountain outcrops, river shores and sand banks, among others.
11	Pasture with exposed soil	Pasture areas, exhibiting signs of severe degradation, containing at least 5% bare soil.
12	Herbaceous Pasture	Pasture in productive process with predominance of herbage and coverage between 9 and 1% by different species of grass.
13	Shrubby Pasture	Areas of pasture in productive process with predominance of herbage and coverage by species of grass between 5% and 8% associated to the presence of shrubby vegetation with coverage between 2% and 5%.
14	Regeneration with Pasture	Areas that were clear-cut, later developed as pasture and are at the beginning of a regenerative process containing shrubs and early successional vegetation.
15	Secondary Vegetation	Areas that were clear-cut and are at an advanced stage of regeneration with trees and shrubs. Includes areas that were used for forestry (silviculture) or permanent agriculture with use of native or exotic species.

Table 36. ESA-DUE-GlobPermafrost legend and description for all products

Class value	Class	Description
1	Sparse vegetation	Sparse vegetation (without shrubs), mostly sandy soil, flood plains, recent landslides, also within fire scars
2	Sparse vegetation	Dry cryptogamic-crust or sparse vegetation
3	Shrub tundra	Graminoid, prostrate dwarf shrub, patterned ground, partially bare
4	Shrub tundra	Dry to moist prostrate to erect dwarf shrub tundra
5	Shrub tundra	Moist to wet graminoid prostrate to erect dwarf shrub tundra
6	Shrub tundra	Wet to waterlogged graminoid prostrate to low shrub tundra
7	Shrub tundra	Moist low dense shrubs
8	Forest	Tall shrubs, deciduous forest
9	Forest	Mixed forest
10	Forest	Coniferous (partially mixed) forest
11	Grassland	Meadows, grass and herb-dominated
12	Floodplain	Wet ecotops, especially in floodplains
13	Disturbed	Disturbed, including forest fire scars, seasonally inundated areas and landslide scars
14	Floodplain	Floodplain, mostly fluvial
15	Floodplain	Floodplain, mostly lacustrine
16	Floodplain	Seasonally inundated
17	Barren	Barren, rare vegetation (petrophytes and psammophytes)
18	Barren	Barren, including articficial surfaces
19	Water	Water (shallow or high sediment yield)
20	Water	Water (medium depth or medium sediment yield)
21	Water	Water (low sediment yield)

Table 37. AUE legend and description for all products

Class value	Class	Description
1	Urban built up area	Urban pixels are the majority of built-up pixels (50% or more) in the radius of 584m
2	Suburban built-up area	Suburban pixels are 25–50% built-up pixels in in the radius of 584m
3	Rural built-up area	Rural pixels are less than 25% built-up pixels in the radius of 584m
4	Fringe open space	Open space pixels within 100 meters of urban or suburban pixels
5	Captured open space	Open space clusters that are fully surrounded by urban and suburban built-up pixels and the fringe open space pixels around them, and that are less

Class value	Class	Description
		than 200 hectares in area
6	Rural open space	All open spaces that are not fringe or captured open spaces.
7	Water	Water

4.3 Linking legend of existing HRCL with HR LC legend

The link between the legends utilized in Phase 1 is contained in Table 38, Table 39 and Table 40 for link with global HR LC with multiple class, with global binary HR LC and with local and regional HR LC with multiple class respectively. The link is based on the class code. Links to new datasets are still under discussion and determination, so they have not been listed yet. Besides, Links to MapBiomas is excluded in Table 40 considering the updates. And links to CCI Prototype Africa is excluded in Table 40 considering the removal.

Table 38. Link between HR LC legend and legend of existing global HR LC with multiple class

Class code	Label	GL30 class code (Table 24)	FROM-GLC 2010 class code (Table 25)	FROM-GLC 2015 class code (Table 26)	FROM-GLC 2017 class code (Table 27)
10	Tree cover evergreen broadleaf	20	21	21, 22, 23, 24, 25, 26, 14	20
20	Tree cover evergreen needleleaf	20	22	21, 22, 23, 24, 25, 26, 14	20
30	Tree cover deciduous broadleaf	20	21	21, 22, 23, 24, 25, 26, 14	20
40	Tree cover deciduous needleleaf	20	22	21, 22, 23, 24, 25, 26, 14	20
50	Shrub cover evergreen	40	40	40, 41, 42, 71	40
60	Shrub cover deciduous	40	40	40, 41, 42, 71	40
70	Grasslands	30	31, 32, 39, 72	31, 32, 33	[-]
80	Croplands	10	13, 19, 94	11,13,15	10
90	Vegetation aquatic or regularly flooded	50	[-]	50, 51, 53	50
100	Grassland vegetation aquatic or regularly flooded	50	51	50, 51, 53	50
110	Lichen and Mosses	70 (only Siberia)	70, 71, 72	72	70 (only Siberia)
120	Bare areas	90	52, 91, 92, 93, 95, 96, 99	52, 90	120
130	Built-up	80	80, 81, 82, 83	80, 12	80
141	Open Water seasonal	60	60	60, 61, 62, 63, 64	60
142	Open Water permanent	60	60	60, 61, 62, 63, 64	60

160	Snow and/or Ice	100	101, 102	100, 101, 102	100

Table 39. Link between HR LC legend and legend of existing global binary HR LC

Class code	Label	GUF class code (Table 28)	GHS BU class code (Table 29)	GSW class code (Table 30)	FNF class code (Table 31)	Tree cover class code (Table 33)	WSF class code (Table 34)	AUE (Table 37)
10	Tree cover evergreen broadleaf	0	0	1	1	50-100	0	4, 5, 6
20	Tree cover evergreen needleleaf	0	0	1	1	50-100	0	4, 5, 6
30	Tree cover deciduous broadleaf	0	0	1	1	50-100	0	4, 5, 6
40	Tree cover deciduous needleleaf	0	0	1	1	50-100	0	4, 5, 6
50	Shrub cover evergreen	0	0	1	2	[-]	0	4, 5, 6
60	Shrub cover deciduous	0	0	1	2	[-]	0	4, 5, 6
70	Grasslands	0	0	1	2	[-]	0	4, 5, 6
80	Croplands	0	0	1	2	[-]	0	4, 5, 6
90	Vegetation aquatic or regularly flooded	0	0	1	2	[-]	0	4, 5, 6
100	Grassland vegetation aquatic or regularly flooded	0	0	1	2	[-]	0	4, 5, 6
110	Lichen and Mosses	0	0	1	2	[-]	0	4, 5, 6
120	Bare areas	0	0	1	2	[-]	0	4, 5, 6
130	Built-up	255	1	1	2	[-]	255	1, 2, 3
141	Open Water seasonal	0	0	2	3	[-]	0	4, 5, 6
142	Open Water permanent	0	0	3	3	[-]	0	4, 5, 6
160	Snow and/or Ice	0	0	1	2	[-]	0	4, 5, 6

Table 40. Link between HR LC legend and legend of existing regional and local HR LC with multiple class

Class code	Label	TerraClass class code (Table 35)	ESA-DUE-GlobPermafrost class code (Table 36)
10	Tree cover evergreen broadleaf	5	[-]
20	Tree cover evergreen needleleaf	5	10
30	Tree cover deciduous broadleaf	5	8
40	Tree cover deciduous needleleaf	5	[-]
50	Shrub cover evergreen	15	3,4,5,6,7
60	Shrub cover deciduous	15	[-]
70	Grasslands	12, 13, 14	11
80	Croplands	1	[-]
90	Vegetation aquatic or regularly flooded	[-]	14, 15
100	Grassland vegetation aquatic or regularly flooded	[-]	14, 15
110	Lichen and Mosses	[-]	2
120	Bare areas	7, 11	1, 17
130	Built-up	3	[-]
141	Open Water seasonal	6	16
142	Open Water permanent	6	19, 20, 21
160	Snow and/or Ice	[-]	[-]

4.4 Methodology

The process of benchmarking or inter-comparison involves a pixel-by-pixel comparison between the output of a project and an existing High-Resolution Land Cover (HR LC). This means that the concurrence or discrepancy between the two maps will be determined based on all available pixels in the region of interest on both maps.

For a pixel-by-pixel comparison, it's essential that the maps have the same spatial resolution. If this is not the case, the map with the lower resolution will be adjusted to match the resolution of the other map.

In addition, inter-comparison requires that the two maps use the same class code for corresponding classes. To address the challenge of legend matching across different products, the team will be working in conjunction with climatology experts. Their insights will be invaluable in ensuring the legends across different products are harmonized effectively.

Given that different existing HR LCs use different coordinate reference systems, WGS84 has been chosen for inter-comparison due to its widespread use.

The inter-comparison will analyse both similarities and differences between project outputs (including static and historical dynamic maps) and existing HR LCs. The similarity between project outputs and existing HR LCs will be calculated using accuracy indexes, specifically Overall Accuracy (OA), Producer's Accuracy (PA), and User's Accuracy (UA). While these indexes are typically associated with accuracy, in the context of inter-comparison, they express the agreement between the products. A confusion matrix will serve as the foundation for calculating these indexes.

Due to the large volume of data to be processed, the inter-comparison will be conducted on the CINECA High-Performance Computing (HPC) system - GALILEO, using automated Python scripts with GRASS-GIS.

4.4.1 Map of Land Cover Agreement

A unique methodology has been introduced during Phase 1, aiming to maximize the efficiency of existing High-Resolution Land Cover (HR LC) datasets and rationalise the interpretation of results. This methodology involves the intersection of all accessible HR LCs within a specific region to extract uniform information. The product of this intersection is a map, designated as the Map of Land Cover Agreement (MOLCA).

The fundamental assumption of this approach is that each land cover map aims to represent the Earth's surface materials with maximum accuracy. When a comparison is made among multiple existing datasets, the areas exhibiting consistent information across all datasets are likely to be the most accurate. As achieving correct classification is an objective of the classification process, rather than a random event, pixels that are correctly classified are expected to appear in the same location across different datasets.

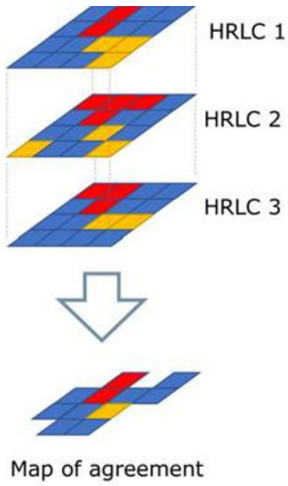

Figure 7. Illustration of extraction of the map of agreement

On the other hand, errors in land cover classification may be attributed to various factors such as the type of image, pre-processing, training data, and classification algorithms. Considering that different land cover maps are produced using diverse procedures and input data, it is reasonable to assume that errors across different datasets are random and uncorrelated. Therefore, the intersection of multiple land cover maps can aid in identifying areas with shared information, which can subsequently be used to extract accurate training and validation samples for the creation of a new land cover map.

The computation of the MOLCA, as illustrated in Figure 7. Illustration of extraction of the map of agreement, is based on an example of intersecting three HR LCs. Generally, a pixel is included in the Map of Agreement if all land cover maps provide identical information for that pixel. If any intersected map displays a different class for a pixel than the other maps, that pixel is designated as null.

The number of maps utilized to construct the MOLCA can vary, depending on the availability and extent of HR LCs within a region. Furthermore, for certain classes such as Forest, Water, and Built-up, the number of intersected HR LCs is larger due to the existence of binary maps specifically focused on these classes.

It will be exploited the possibility of applying the MOLCA methodology to both historical and dynamic land cover products. MOLCA's ability to extract consistent information from multiple high-resolution land cover datasets would potentially make it a powerful tool for analysing historical land cover changes. However, the application of MOLCA to historical and dynamic land cover products would necessitate careful consideration of several additional factors. For example, changes in satellite sensor technology and data processing methods over time could lead to inconsistencies between datasets from different periods. Similarly, the rapid land cover changes captured in dynamic products may not be accurately reflected in Maps of Agreement if the intersected datasets do not all cover the same time-period. Therefore, it will be evaluated and assessed the applicability of introducing a scoring system to determine the certain level of dis/agreement of MOLCA outputs when applied to multitemporal LC maps.

Prior to intersecting the existing HR LCs, it is imperative to harmonize their resolutions, legends, and coordinate reference systems. In the static case, a resolution of 10m was selected as it was the highest resolution among the available HR LCs. The chosen coordinate reference system was WGS84 (EPSG:4326), which is the most prevalent among existing HR LCs. Finally, the legends were harmonized into the following classes: Forest, Cropland, Grassland, Shrubland, Bareland, Built-up, Water, Wetland, and Permanent ice and snow. Now, considering also the new products, these parts need to be updated considering their characteristics.

The legend of the Map of Agreement could be less detailed than the legend of CCI HR LC. Certain types of Water, Forest, Shrubland, and Wetland are either not included in the legends of the existing HR LCs, or they are included only in one dataset in a region, which is inadequate for intersection. References

5 References

[1] F. C. Ford-Robertson, *Terminology of forest science, technology practice and products: English-language version*. Washington, DC (USA) SAF, 1971.

[2] H. J. Scoggan, *The flora of Canada.*, no. 581.9 SOC. CIMMYT. distributeur Ottawa, 1978.

[3] JECAM, "JECAM Guidelines for cropland and crop type definition and field data collection," pp. 1–7, 2014.

[4] P. Gray, *The Encyclopedia of the Biological Sciences*. New York, 1970.

[5] P. Mayaux *et al.*, "Validation of the global land cover 2000 map," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 7, pp. 1728–1737, 2006, doi: 10.1109/TGRS.2006.864370.

[6] O. Arino, P. Bicheron, F. Achard, J. Latham, R. Witt, and J. L. Weber, "GlobCover: The most detailed portrait of Earth," *European Space Agency Bulletin*, vol. 2008, no. 136, pp. 24–31, 2008, doi: WOS:000261325400004.

[7] P. Defourny *et al.*, "The first 300 m global land cover map for 2005 using ENVISAT MERIS time series: A product of the GlobCover system," in *The 33rd International Symposium on Remote Sensing of Environment*, 2009.

[8] A. H. Strahler *et al.*, "Global Land Cover Validation: Recommendations for Evaluation and Accuracy Assessment of Global Land Cover Maps," *Scientific and Technical Research series*, vol. EUR 22156, p. 48, 2006, doi: 10.1080/01431160512331326521.

[9] M. Herold *et al.*, "Assessment of the status of the development of the standards for the Terrestrial Essential Climate Variables: T9 Land Cover," Rome, 2009.

[10] P. Olofsson, G. M. Foody, M. Herold, S. V. Stehman, C. E. Woodcock, and M. A. Wulder, "Good practices for estimating area and assessing accuracy of land change," *Remote Sensing of Environment*, vol. 148, pp. 42–57, 2014, doi: 10.1016/j.rse.2014.02.015.

[11] A. Belward, "The global observing system for climate: Implementation needs," *Reference Number GCOS-200*, p. 315, 2016.

[12] S. V. Stehman and G. M. Foody, "Key issues in rigorous accuracy assessment of land cover products," *Remote Sensing of Environment*, vol. 231, no. May, p. 111199, 2019, doi: 10.1016/j.rse.2019.05.018.

[13] A. H. Strahler *et al.*, "Global Land Cover Validation: Recommendations for Evaluation and Accuracy Assessment of Global Land Cover Maps," *Scientific and Technical Research series*, vol. EUR 22156, no. 25, p. 48, 2006, doi: 10.1080/01431160512331326521.

[14] R. G. Congalton and K. Green, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Second edi. Press, CRC, 1999.

[15] P. Bicheron, C. Henry, S. Bontemps, and G. Partners, "Globcover Products Description Manual," 2008.

[16] S. Bontemps, P. Defourny, E. Van Bogaert, V. Kalogirou, and J. R. Perez, "GLOBCOVER 2009 Products Description and Validation Report," *ESA Bulletin*, vol. 136, p. 53, 2011, doi: 10013/epic.39884.d016.

[17] W. J. Stadelman, "Contaminants of liquid egg products," in *Microbiology of the avian egg*, Springer, 1994, pp. 139–151.

[18] S. Khorram, "Accuracy Assessment of Remote Sensing-Derived Change Detection," *Bethesda, Md., American Society for Photogrammetry and Remote Sensing Monograph Series*, p. 65, 1999.

[19] European Commission. Joint Research Centre., *Mapping global forest cover of the year 2020 to support the EU regulation on deforestation-free supply chains.* LU: Publications Office, 2024. Accessed: Aug. 01, 2024. [Online]. Available: https://data.europa.eu/doi/10.2760/262532

[20] X. Zhang, L. Liu, X. Chen, Y. Gao, S. Xie, and J. Mi, "GLC_FCS30: global land-cover product with fine classification system at 30 m using time-series Landsat imagery," *Earth System Science Data*, vol. 13, no. 6, pp. 2753–2776, Jun. 2021, doi: 10.5194/essd-13-2753-2021.

[21] X. Zhang *et al.*, "GLC_FCS30D: the first global 30 m land-cover dynamics monitoring product with a fine classification system for the period from 1985 to 2022 generated using dense-time-series Landsat imagery and the continuous change-detection method," *Earth System Science Data*, vol. 16, no. 3, pp. 1353–1381, Mar. 2024, doi: 10.5194/essd-16-1353-2024.

[22] X. Zhang, L. Liu, T. Zhao, Y. Gao, X. Chen, and J. Mi, "GISD30: global 30 m impervious-surface dynamic dataset from 1985 to 2020 using time-series Landsat imagery on the Google Earth Engine platform," *Earth System Science Data*, vol. 14, no. 4, pp. 1831–1856, Apr. 2022, doi: 10.5194/essd-14-1831-2022.

[23] X. Zhang *et al.*, "GWL_FCS30: global 30 m wetland map with fine classification system using multi-sourced and time-series remote sensing imagery in 2020," *Earth System Science Data*, vol. 15, no. 1, pp. 265–293, Jan. 2023, doi: 10.5194/essd-15-265-2023.

[24] X. Zhang, L. Liu, T. Zhao, J. Wang, W. Liu, and X. Chen, "Global annual wetland dataset at 30 m with a fine classification system from 2000 to 2022," *Scientific Data*, vol. 11, no. 1, p. 310, Mar. 2024, doi: 10.1038/s41597-024-03143-0.

[25] C. F. Brown *et al.*, "Dynamic World, Near real-time global 10 m land use land cover mapping," *Sci Data*, vol. 9, no. 1, p. 251, Jun. 2022, doi: 10.1038/s41597-022-01307-4.

[26] K. Karra, C. Kontgis, Z. Statman-Weil, J. C. Mazzariello, M. Mathis, and S. P. Brumby, "Global land use / land cover with Sentinel 2 and deep learning," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, Brussels, Belgium: IEEE, Jul. 2021, pp. 4704–4707. doi: 10.1109/IGARSS47720.2021.9553499.

[27] D. Zanaga *et al.*, "ESA WorldCover 10 m 2020 v100." Zenodo, Oct. 20, 2021. doi: 10.5281/zenodo.5571936.

[28] D. Zanaga *et al.*, "ESA WorldCover 10 m 2021 v200." Zenodo, Oct. 28, 2022. doi: 10.5281/zenodo.7254221.

[29] European Commission. Joint Research Centre., *GHSL data package 2023*. LU: Publications Office, 2023. Accessed: Apr. 29, 2024. [Online]. Available: https://data.europa.eu/doi/10.2760/098587

[30] M. Pesaresi and P. Politis, "GHS-BUILT-S R2023A - GHS built-up surface grid, derived from Sentinel2 composite and Landsat, multitemporal (1975-2030)." European Commission, Joint Research Centre (JRC), May 08, 2023. doi: 10.2905/9F06F36F-4B11-47EC-ABB0-4F8B7B1D72EA.

[31] P. Arevalo *et al.*, "Global Land Cover Mapping and Estimation Yearly 30 m V001." NASA EOSDIS Land Processes Distributed Active Archive Center, 2022. doi: 10.5067/MEASURES/GLANCE/GLANCE30.001.

[32] M. A. Friedl *et al.*, "Medium Spatial Resolution Mapping of Global Land Cover and Land Cover Change Across Multiple Decades From Landsat," *Front. Remote Sens.*, vol. 3, Jun. 2022, doi: 10.3389/frsen.2022.894571.

[33] J. Chen *et al.*, "Global land cover mapping at 30 m resolution: A POK-based operational approach," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 103, pp. 7–27, May 2015, doi: 10.1016/j.isprsjprs.2014.09.002.

[34] C. Jun, Y. Ban, and S. Li, "Open access to Earth land-cover map," *Nature*, vol. 514, no. 7523, pp. 434–434, Oct. 2014, doi: 10.1038/514434c.

[35] J. Chen, X. Cao, S. Peng, and H. Ren, "Analysis and Applications of GlobeLand30: A Review," *International Journal of Geo-Information*, vol. 6, no. 8, p. 230, Jul. 2017, doi: 10.3390/ijgi6080230.

[36] P. Gong *et al.*, "Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data," *International Journal of Remote Sensing*, vol. 34, no. 7, pp. 2607–2654, Apr. 2013, doi: 10.1080/01431161.2012.748992.

[37] C. Li *et al.*, "The first all-season sample set for mapping global land cover with Landsat-8 data," *Science Bulletin*, vol. 62, no. 7, pp. 508–515, Apr. 2017, doi: 10.1016/j.scib.2017.03.011.

[38] P. Gong *et al.*, “Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017,” *Science Bulletin*, vol. 64, no. 6, pp. 370–373, Mar. 2019, doi: 10.1016/j.scib.2019.03.002.

[39] C. Corbane *et al.*, “Big earth data analytics on Sentinel-1 and Landsat imagery in support to global human settlements mapping,” *Big Earth Data*, vol. 1, no. 1–2, pp. 118–144, Dec. 2017, doi: 10.1080/20964471.2017.1397899.

[40] Joint Research Centre (European Commission) *et al.*, *GHSL data package 2019: public release GHS P2019*. Publications Office of the European Union, 2019. Accessed: Apr. 22, 2024. [Online]. Available: https://data.europa.eu/doi/10.2760/290498

[41] J.-F. Pekel, A. Cottam, N. Gorelick, and A. S. Belward, “High-resolution mapping of global surface water and its long-term changes,” *Nature*, vol. 540, no. 7633, pp. 418–422, Dec. 2016, doi: 10.1038/nature20584.

[42] M. Shimada *et al.*, “New global forest/non-forest maps from ALOS PALSAR data (2007–2010),” *Remote Sensing of Environment*, vol. 155, pp. 13–31, Dec. 2014, doi: 10.1016/j.rse.2014.04.014.

[43] M. C. Hansen *et al.*, “High-Resolution Global Maps of 21st-Century Forest Cover Change,” *Science*, vol. 342, no. 6160, pp. 850–853, Nov. 2013, doi: 10.1126/science.1244693.

[44] M. Marconcini *et al.*, “Outlining where humans live, the World Settlement Footprint 2015,” *Sci Data*, vol. 7, no. 1, p. 242, Jul. 2020, doi: 10.1038/s41597-020-00580-5.

[45] M. Marconcini, A. Metz-Marconcini, T. Esch, and N. Gorelick, “Understanding Current Trends in Global Urbanisation - The World Settlement Footprint Suite,” *giForum*, vol. 1, pp. 33–38, 2021, doi: 10.1553/giscience2021_01_s33.

[46] C. A. de Almeida *et al.*, “High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data,” *Acta Amaz.*, vol. 46, pp. 291–302, Sep. 2016, doi: 10.1590/1809-4392201505504.

[47] A. Bartsch, B. Widhalm, G. Pointner, K. A. Ermokhina, M. Leibman, and B. Heim, “Landcover derived from Sentinel-1 and Sentinel-2 satellite data (2015-2018) for subarctic and arctic environments.” 2019. doi: 10.1594/PANGAEA.897916.

[48] M. Lesiv *et al.*, “Accuracy Assessment of the ESA CCI 20M Land Cover Map: Kenya, Gabon, Ivory Coast and South Africa.” Accessed: Apr. 22, 2024. [Online]. Available: https://iiasa.dev.local/

[49] L. Yu, J. Wang, and P. Gong, “Improving 30 m global land-cover map FROM-GLC with time series MODIS and auxiliary data sets: a segmentation-based approach,” *International Journal of Remote Sensing*, vol. 34, no. 16, pp. 5851–5867, Aug. 2013, doi: 10.1080/01431161.2013.798055.

[50] T. Esch *et al.*, “Where We Live—A Summary of the Achievements and Planned Evolution of the Global Urban Footprint,” *Remote Sensing*, vol. 10, no. 6, Art. no. 6, Jun. 2018, doi: 10.3390/rs10060895.

[51] Barroes *et al.*, “Validação dos dados TerraClass para o município de Paragominas.” 2017.

[52] “MapBiomas Brasil.” Accessed: Jun. 06, 2024. [Online]. Available: https://brasil.mapbiomas.org/en/estatistica-de-acuracia/