Climate Change Initiative Extension (CCI+) Phase 1

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
Université de Versailles Saint Quentin
CREAF
e-GEOS s.p.a.
Planetek Italia
GeoVille

Changelog

Issue	Changes	Date
1.0	First version	03/07/2019
1.1	Internal second version (not submitted)	16/10/2019
2.0	Second version	06/01/2020

Detailed Change Record

Issue	RID	Description of discrepancy	Sections	Change

Contents

1 Introduction

1.1 Purpose and scope

The objective of this Product Validation Plan (PVP) is to describe the strategies and selected for the validation, benchmarking and scaling analysis of the European Space Agency (ESA) Climate Change Initiative (CCI) High Resolution (HR) Land Cover (LC) products: 10-m Round Robin (RR) local LC prototypes for the year 2018, a 10-m static sub-continental LC maps for the year 2018 and 30-m regional historical (1990-2015) LC maps.

1.2 Applicable documents

[AD1] CCI_HRLC_Ph1-D2.3_E3UB, v1.0, 03/07/2019

1.3 Acronyms and abbreviations

ALOS	Advanced Land Observing Satellite
ASAP	Anomaly Hotspots of Agricultural Production
C3S	Copernicus Climate Change Service
CCI	Climate Change Initiative
CEOS	Committee on Earth Observation Satellites
CMC	Climate Modelling Community
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
G	GeoEye
GCOS	Global Climate Observing System
GEO	Group for Earth Observation
GLC	Global Land Cover
GLS	Global Land Survey
GOFC-GOLD	Global Observation of Forest Cover – Global Observation of Land Dynamics
HR	High Resolution
I	IKONOS
IFOV	Instantaneous Field of View
IKONOS	Commercial Earth Observation Satellite
JRC	Joint Research Centre
LC	Land Cover
LCC	Land Cover Change
LCCS	Land Cover Classification System
LCML	Land Cover Meta-Language
LPVS	Land Product Validation Subgroup
MMU	Minimum Mapping Unit
MR	Medium Resolution
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetative Index
OA	Overall Accuracy
PA	Producer Accuracy

PSU	Primary Sampling Unit
PVP	Product Validation Plan
ROC	Relative Operating Characteristics
RR	Round Robin
S2	Sentinel-2
SPOT	Satellite Pour l'Observation de la Terre
SSU	Secondary Sampling Unit
TM	Thematic Mapper
TPM	Third Party Mission
TREES	Tropical Ecosystem Environment Observations by Satellite
UA	User Accuracy
UCLouvain	Université catholique de Louvain
USGS	United States Geological Survey
VHR	Very High Resolution
WGCV	CEOS Working Group on Calibration & Validation
WV	WorldView

2 CCI HRLC products to be validated thematically

2.1.1 Three types of CCI HRLC outputs

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

- **10-m Round Robin local LC prototypes for the year 2018**, produced through a Round Robin (RR) exercise during which optical and SAR classifications will be benchmarked.
- **10-m static sub-continental LC maps for the year 2018**. The algorithms from the RR exercise with the best accuracy figures will be selected and applied at the sub-continental scale.
- **30-m regional historical (1990-2017) LC maps**, generated every five years, since 1990 on reduced areas. Change will first be detected on an annual basis on Landsat time series. It will then be used to backdate, on a 5-year basis back to 1990, the detailed spatio-temporal 10-m static LC map.

All three types of land cover products will be generated over three areas selected through key users' consultation, with varying extents (Figure 1). The RR sites, in black, cover 3 Sentinel-2 (S2) tiles located in the Amazonian region (21KUQ, 21KXT) and in Siberia (42WXS). One additional tile has been finally been included in Sahel. The static LC maps, in red, will cover the regions of Amazon (including Mato Grosso), Sahel and Siberia. The historical LC maps are restricted to the blue areas.

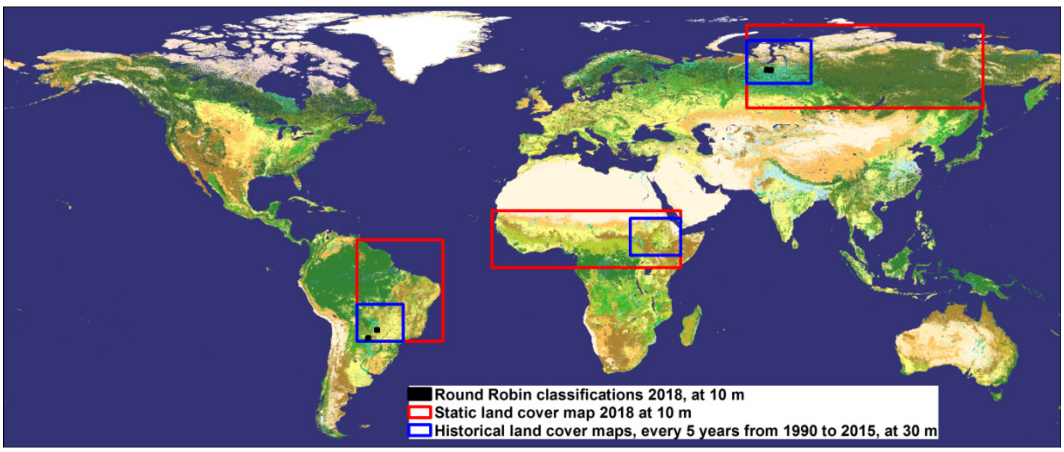


Figure 1: Distribution of the study sites per type of CCI HRLC product.

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

Table 1. Spatial extents to which the various HRLC products will be generated.

Region	CCI HRLC outputs	Extent	Surface [km²]	Lat min	Lon min	Lat max	Lon max
Amazon	Round robin site 1	S2 tile 21KUQ	10000	-23	-59	-22	-58
	Round robin site 2	S2 tile 21KXT	10000	-21	-56	-20	-55
	LC map 2018	Full extent	11450200	-24	-62	9	-34
	Historical LC maps	Reduced extent	2230570	-24	-62	-12	43.5
Sahel	Round robin site	S2 tile 37PCP	10000	12	37	13	38
	LC map 2018	Full extent	14099000	0	-18	18	43.5
	Historical LC maps	Reduced extent	2453620	16	44	4	27
Siberia	Round robin site	S2 tile 42WXS	10000	64	71	72	65
	LC map 2018	Full extent	25763000	52	65	79	142
	Historical LC maps	Reduced extent	3643260	74	86	60	65

2.1.2 HRLC legend

The Food and Agriculture Organization (FAO) Land Cover Classification System (LCCS) was found pertinent to support the description of the CCI HRLC maps. Based on key user consultations and after adaptation to the FAO LCCS framework, a set of 15 main classes is proposed for the LC mapping at 10 m spatial resolution (Table 2).

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

Region	CCI HRLC outputs
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 composed of trees that are never entirely without green foliage [1]. Trees are broadleaved and come from the Angiospermae group.
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 composed of trees that are never entirely without green foliage [1]. Trees carry typical needle-shaped leaves and come from the Gymnospermae group.
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.
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 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.

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].
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 composed of shrubs that are leafless for a certain period during the year [1].
Grassland	Primarily vegetated areas with an 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 stem or shoots above ground and lacking definite firm structure [2].
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 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 stem or shoots above ground and lacking definite firm structure [2]. Cropland includes rain fed 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.
Vegetation aquatic or regularly flooded	Primarily vegetated areas with trees, 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.
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].
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.
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.
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 year, 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.

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 year, 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.
Snow and/or 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.

2.1.3 Definition of change

Definition of what is meant by "change" is equally important as the definition of the LC classes for the validation. The list of expected change transitions between the different HRLC classes .

3 Overall validation process

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], the CCI LC 1992-2015/C3S 2016-2017 maps [7]. It is also intended to reflect state-of-the-art standard protocols of land cover validation such as the CEOS Working Group on Calibration and Validation (LC validation subgroup). In particular, the design and implementation of the validation plan and the creation of the reports follows 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 4) and includes an independent statistical quantitative validation of the three HRLC outputs and a benchmarking of the static and historic HRLC maps to other existing products (Figure 2). The methodology is fine-tuned to the specific challenges in validating each type of product. The aim of the RR exercise is to select the pre-processing and classification algorithms providing the best results and that will eventually be upscaled to the sub-continental extent. Accuracy figures need to clearly discriminate the quality of the various RR prototypes and a sampling design biased towards error-prone areas and highlighting differences between maps is selected [11]. The validation of the static HRLC map aims at quantifying and reporting the quality of the product from an overall, producer and user accuracy point of view. Validating the historic HRLC maps represents the most significant challenges of this exercise as, to the knowledge of the consortium, building a statistically sound validation of change in the context of multiple LC classes has not been tackled yet by the community. For all types of products, the validation process is composed of 4 steps (section 4.2.3): the collection of Very-High-Resolution (VHR) imagery, the sampling design, the response design and the reporting.

The activity of benchmarking (i.e. comparing) between the HRLC static and historic maps and other LC products will be performed in the second year of this project so that the PVP will be updated accordingly in due time.

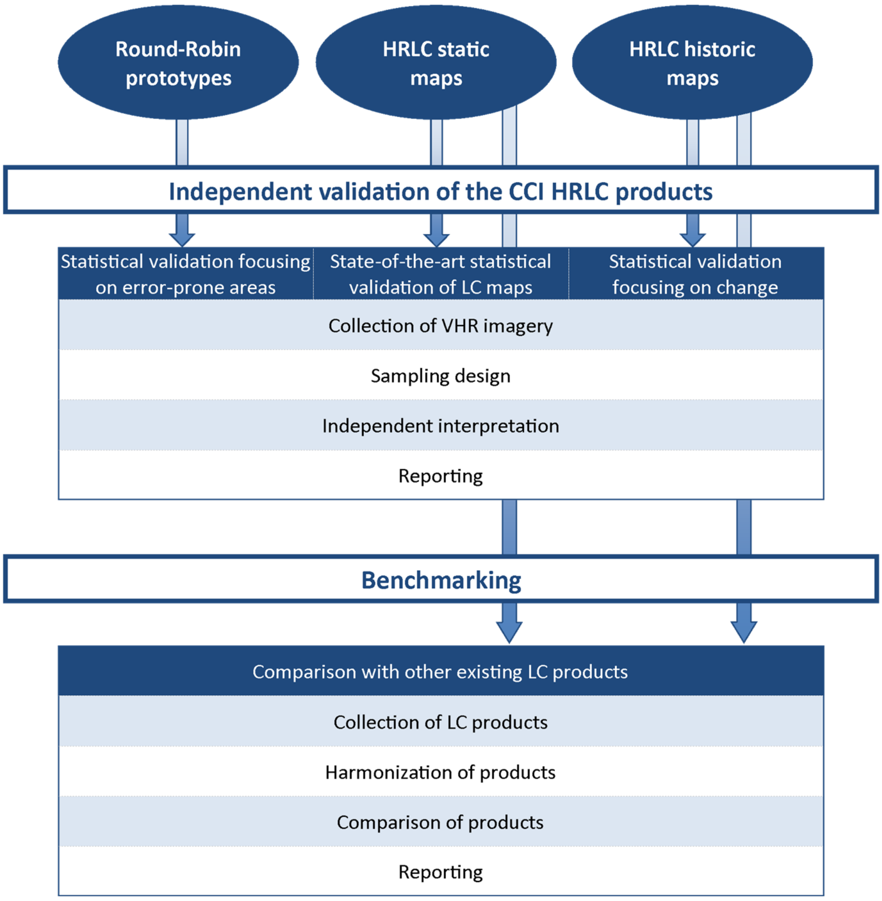


Figure 2. Detailed description of the different validation components of the CCI HRLC project.

4 State-of-the art of accuracy assessment methodology

4.1 Definition and standard protocols

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 ECV Uncertainty Budget (E3UB) [AD1].

4.1.1 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 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 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 time 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 time periods.
Stage 2	Product accuracy is estimated over a significant set of locations and time 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 time 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 time 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.

4.1.2 Validation requirements

The validation procedure of the HRLC maps is also driven by the main GCOS requirements [12], 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	5 years
TEMPORAL EXTENT	1990 - present
RESOLUTION	
GEOMETRICAL RESOLUTION	10/30 m
ERROR/UNCERTAINTY	
ACCURACY-UNCERTAINTY	5% (max. error of omission and commission in mapping individual classes), location accuracy 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 a **stable accuracy** should be reflected in implementing an accuracy assessment of LC change, at least per decade.

4.2 Good practices of accuracy assessment

4.2.1 Independence of the validation process

The RR prototypes, the static HRLC map and the historical LC, will be validated using a transparent traceable validation 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 dataset used for validation should not have been used during the production of the products to be validated. As [13] state it, 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.

4.2.2 Reference data specification

The collection of ground information (i.e. through field surveys) is considered as 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".

4.2.3 Sampling frame requirements

To satisfy 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 [14].

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

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

4.2.3.1 Number of sample plots

The sample size required for a given confidence level and a given acceptable error in the sample can be calculated from the binomial distribution [15]:

$$n = \frac{p * q}{\left(\frac{E}{z_\alpha}\right)^2} \qquad (1)$$

where n is the number of sample plots, E is the allowable error in the sample, z_α is drawn from the normal distribution for the given level of confidence, p is the required accuracy and q is 1-p.

The allowable error stands for the error that is made when validating the product using a sampling strategy (instead of exhaustively assessing the product in any locations). For example, if the allowable error has been set at 5% and the validation process (based on a given sample) yields to a result of 77% successes, then it is safe to claim that an accuracy value between 72% and 82% would have been obtained if the whole accuracy had been validated.

The z-value (z_α) is directly related to the statistical normal distribution and a certain level of confidence. This level of confidence expresses the percentage that indicates how often a validation performed on the basis of a given sample dataset would yield a result that lies within the confidence interval. Considering the previous example (allowable error of 5% and accuracy of 77%) and adding the concept of this level of confidence (set at 95% for instance), we could say that we are 95% sure that the quality of the product falls between 72% and 82%.

The number of samples is also constrained 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 of VHR resolution imagery for validation is the third parameter constraining the sample size.

4.2.3.2 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 [16]. 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.

4.2.4 Response design

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

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

Based on the GlobCover [17], [18] and CCI MRLC validation experience, the process of samples 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. An image interpretation protocol ensures each sample unit is interpreted – i.e. labelled – by the expert in a systematic and consistent way:

- If the expert cannot derive the land cover because of the poor quality of reference imagery, the sample has to be skipped. The expert must specify that no land cover 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 land cover class, the expert needs to indicate the classes from which the expert cannot choose with certainty. It is clear that more attributes than the dominant land cover classes are relevant, especially for the analysis of observed discrepancies between classification and expert's labelling.

4.2.5 Reporting

4.2.5.1 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, Kappa statistics and area statistics. The confusion matrix is recognised to efficiently organize and summarize the agreement between the maps and reference classification [13](Figure 3).

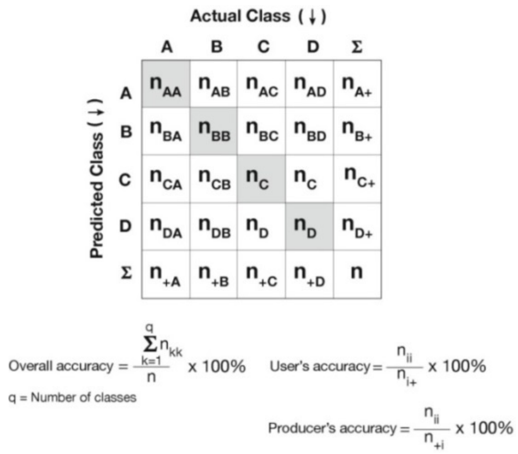


Figure 3. Layout of a typical confusion or error matrix, showing the computation of user's and producer's accuracies (from [8]). The fields in grey mark the correspondence between classification and labelling by the expert.

A shortcoming of the overall accuracy is that it does not account for chance agreement. A complete random classification would also generate a certain level of accuracy. Cohen's Kappa is an index used to make this comparison. It expresses the proportionate reduction in error generated by a classification process, compared with the error of a completely random classification. The closer Kappa gets to 1.0, the higher the accuracy of the data. A value of 1.0 would imply that the classification process was avoiding all of the errors that a completely random classification would generate. This Kappa index is frequently written as follows:

$$C_{Kappa} = \frac{Acc_{Overall} - Prob_{chance}}{1 - Prob_{chance}}$$

where $Acc_{Overall}$ = Overall accuracy and $Prob_{chance}$ = Probability that the agreement is due to chance.

Finally, the F-score will also be reported. It represents for a class k the harmonic mean of the user and producer accuracies and ranges between 0 and 1.

$$F - Score_k = 2 * \frac{UA_k * PA_k}{UA_k + PA_k}$$

4.2.5.2 Reporting accuracy figures and area

To calculate the overall accuracy of the product when the sampling design is not equal probability [13], each class should be weighted by the area it represents in the map.

5 Accuracy assessment tailored to each CCI HRLC products

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

5.1 Reference data sources collection

Several types of high and VHR geolocalized imagery with spatial resolutions below 10 m have been identified for the purpose of validation. Their specifications are detailed in Table 5. Planet data are ignored in this study due to current uncertainties in the geolocation accuracy and consistency of the images.

Table 5. Specifications of the potential very high-resolution imagery suitable for validation.

Product	Spatial Coverage	Temporal Coverage	Resolution (m)
WorldView-1/-2/-3	Global	2007 - present	0.5, 046, 0.31
Pleiades	Global	2011 - present	0.5

Product	Spatial Coverage	Temporal Coverage	Resolution (m)
GeoEye-1	Global	2008 - present	0.41, 1.65
IKONOS	65 N, 9 S, 8 W, 75 E	1999 - 2008	0.82
RapidEye	Global	2008 - present	5
SPOT 1 - 5	Global	1986 - 2011	20
SPOT 6 - 7	Global	2012 - present	1.5

5.1.1 ESA Third Party Mission collections

Table 6 summarizes the VHR imagery availability from the various ESA TPM collections. No matter the sensor, the quantity, spatial and temporal distributions of the images do not allow properly validating any of the CCI HRLC products. Annex 1 provides the full details of investigation.

Table 6. Evaluation of the availability of images from ESA TPM data archives per type of CCI HRLC products.

HRLC product	Year	Region	Pleiades	SPOT1-5	SPOT6-7	Rapideye	Deimos	IKONOS2	TropForest
Static LC map 2018	2018 or more recent	Amazon	3 (2013+2x2016)	7 (2000, 2001, 3x2006, 2x2007)	3 (2x2015, 2016)	24	0	0	0
		Sahel	0	248 (1980-2011)	1 (2014)	17	0	0	0
		Siberia	0	0	0	0	0	0	0
Historical LC	2015,-5,1990	Amazon	2 (2013,2016)	0	1 (2016)	1 mosaic Bolivia and Paraguay (2013-2015) + 1 time series	0	0	Brazil, Bolivia and Paraguay samples on 2009-2010
		Sahel	0	170 (1980-2011)	0	4 time series	0	0	0
		Siberia	0	0	0	0	0	0	0
Round Robin	2018	S2 tile 21KUQ (Amazon)	0	0	0	0	0	0	0
		S2 tile 21KXT (Amazon)	0	0	0	0	0	0	0
		S2 tile 42WXS (Siberia)	0	0	0	0	0	0	0

5.1.2 Availability of Airbus data archives investigated for validation

Given the insufficient amount of the VHR imagery from the ESA TPM collection, the SPOT and Pleiades archives of the Airbus Geostore have been investigated first with ideal criteria of cloud cover and incidence angles set to 0 and minus 5°, respectively. Then, criteria were relaxed to cloud cover in a 0-10% range and incidence angles in a 0-15° range according to the amount of results. Figure 4, Figure 5 and Figure 6 localize the various types of VHR imagery with spatial resolution in the 0-10m range for the regions of Amazon, Africa and Siberia, respectively. Table 7 to Table 10 include the number and surface covered by the VHR imagery, per type of CCI HR LC product.

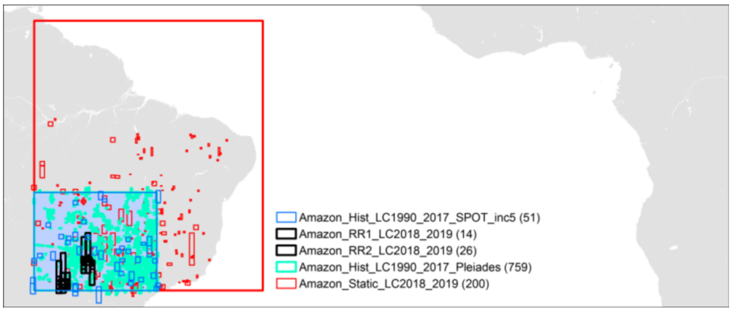


Figure 4. Footprints of the images available from the Airbus Geostore for the validation of the RR, static and historic maps of Amazon.

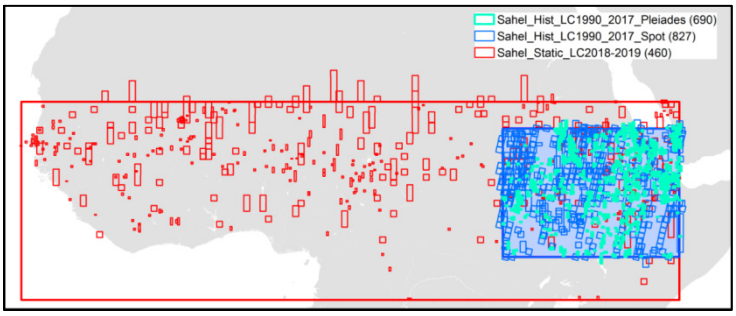


Figure 5. Footprints of the images available from the Airbus Geostore for the validation of the static and historic maps of Sahel.

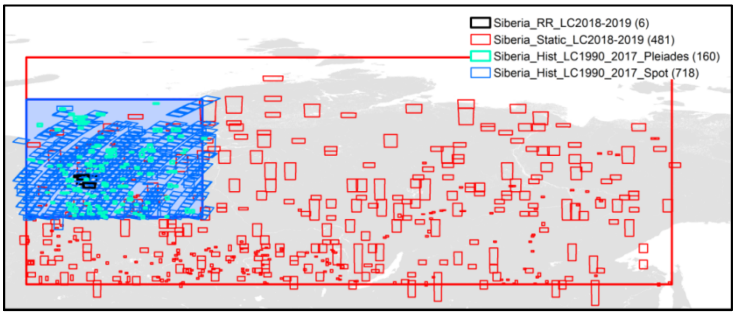


Figure 6. Footprints of the images available from the Airbus Geostore for the validation of the RR, static and historic maps of Siberia.

Table 7. Historic HRLC maps of Siberia. Number and spatial distribution of the Airbus Geostore VHR imagery with spatial resolution in the 0-20 m for the validation of the most critical periods: 1990_2005 and 2006-2012.

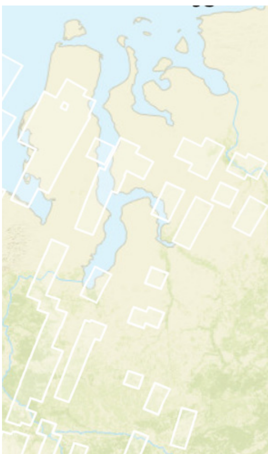
1990-2005 (110 images)	2006-2012 (746 images)
	

Table 8. Historic HRLC maps of the Sahelian study area. Number and spatial distribution of the Airbus Geostore VHR imagery with spatial resolution in the 0-20 m for the validation of the most critical periods: 1990_2005 and 2006-2012.

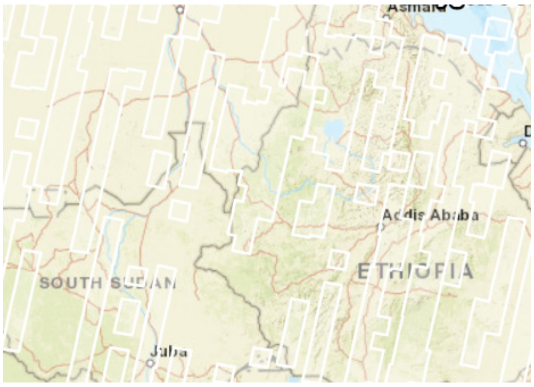
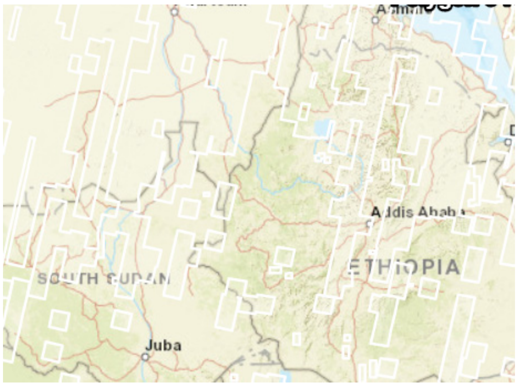
1990-2005 (7893)	2006-2012 (7772)
	

Table 9. Historic HRLC maps of Amazon. Number and spatial distribution of the Airbus Geostore VHR imagery with spatial resolution in the 0-20 m for the validation of the most critical periods: 1990_2005 and 2006-2012.

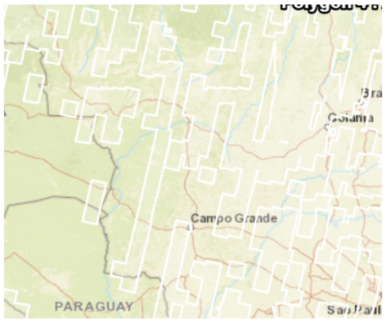
1990-2005 (1178)	2006-2012 (3841)
	

Table 10. Number of VHR imagery and surface of non-overlapping images from the Airbus Geostore suitable for the validation of the RR and Static HRLC products.

HRLC product	Year	Region	Total images	% of study area covered
Round Robin	2018	S2 tile 21KUQ (Amazon)	14	100
		S2 tile 21KXT (Amazon)	26	100
		S2 tile 42WXS (Siberia)	6	35
Static LC map 2018	2018 or more recent	Amazon	200	3
		Sahel	460	8
		Siberia	481	13

5.1.3 Auxiliary information

To complement the VHR images possibly made available to the project and according to the quantity and location of these, other types of information will be used.

First, the image interpretation protocol will rely on very high spatial resolution data available from **Google Earth**. However, it has to be noted that their use could be limited because the level of details can vary from site to site.

In addition, the possibility to consult multi-temporal annual profiles of spectral indices will also be recommended in order to characterize the seasonal variations of the various LC. Indeed, **S2 vegetation indices**

profiles can be extracted on a yearly basis with Google Earth Engine or other web interfaces like the Joint Research Centre (JRC) Anomaly Hotspots of Agricultural Production (ASAP) High-Resolution viewer.

Finally, the **Global Land Survey** (GLS) was derived from orthorectified and geodetically accurate global land dataset of Landsat TM (30 m × 30 m) and Enhanced Thematic Mapper (ETM+) (30 m × 30 m) satellite images with global coverage. It was created from the epochs circa 1990, circa 2000 and 2005 by NASA at the global scale. Although the spatial resolution of 30 m cannot be used to validate the CCI historical HRLC maps at 30 m due to geolocation issues, it could be exploited to determine if a change has occurred between the 5-year epochs. The accuracy of the land cover label should be derived from imagery with higher spatial accuracy than the HRLC product itself. The GLS datasets present the advantages of wall-to-wall standardized imagery but should be used with care to evaluate the change as the GLS time indications stand, roughly, for the period of 1986 – 1993 for GLS-1990 and of 1999 – 2001 for GLS-2000.

5.2 Sampling designs

More specifically to the type of CCI HRLC product, the sampling scheme needs:

- to highlight accuracies that allow comparing RR prototypes and select the best algorithms;
- to address the issue of rare classes with a strong impact on the climate system (urban areas, wetlands, etc.) in the CCI HRLC static maps;
- to address the issue of rare events of change in LC.

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.

5.2.1 For the Round Robin prototypes

The assessment of 2018 RR prototypes will derive the overall accuracy and per-class accuracy with the aim of selecting the best classification method to generate the static LC maps at the regional scale in a second step.

The area validated is constrained by the availability of VHR imagery, price and minimum order size area of 100 km² from the Airbus Geostore (Table 11). For Amazon, where each S2 tile is fully covered by VHR imagery, it is proposed to select nine footprints of 100 km² spread as illustrated in Figure 7. Fore Siberia, where only 35% of the tile is covered, the 9 footprints of 100 km² will be selected randomly but ensuring there is no overlap.

Table 11. Surface of non-overlapping images from the Airbus Geostore suitable for the validation of the RR HRLC products.

HRLC product	Year	Region	% / km² of study area covered by VHR images	Surface [km²] selected for validation
Round Robin	2018	S2 tile 21KUQ (Amazon)	100 / ~12000	900
		S2 tile 21KXT (Amazon)	100 / ~12000	900
		S2 tile 42WXS (Siberia)	35/ 4200	900

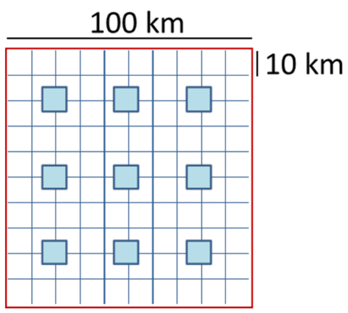


Figure 7. Distribution of the 9 footprints of 100 km² (blue) distributed within each S2 tiles of the Amazon RR.

Within the validation area, a **stratified random sampling** is tailored for a quality assessment when reference data is scarce and the objective is to highlight differences between products.

It is expected that most prototypes should accurately describe most LC classes, especially in homogeneous areas such as stretches of intact broadleaf evergreen forest. With marginal differences between prototypes, using random samplings will bring similar high accuracy figures for each prototype, thus making the selection of the best RR prototype impossible. Two strata are defined, based on the stratification from [11], where one class was marginally distributed. Stratum 1 is biased towards areas of discrepancies between RR prototypes and stratum 2 focuses on areas where prototypes agree. Within stratum 1, the sample random selection is further stratified per LC classes to ensure that even the rare LC classes are represented. This second-stage stratification is defined by an external LC product, the Finer Resolution Observation and Monitoring of Global Land Cover (FROM-GLC) 2017 [19], which includes similar LC classes as the CCI HRLC legend.

Samples allocation is not proportional to the area of each stratum as we expect higher accuracy in areas of agreement and therefore fewer samples are necessary to maintain narrow confidence intervals of 5%. The **number of samples** per stratum is defined according to equation 1 (Table 3). The total number per sample is then evenly distributed among the 10 LC classes under assessment.

Table 12. Number of samples per stratum for the evaluation of each RR prototype deduced from the choice of parameters from equation 1.

Stratum	p	q	E	Zα	n	n/class
1	0.65	0.35	0.03	1.96	971	~100
2	0.85	0.15	0.03	1.96	544	~55

5.2.2 For the HRLC static maps

The sampling scheme dedicated to the assessment of the HRLC static maps takes advantage of the systematic sampling of the Tropical Ecosystem Environment Observations by Satellite (TREES) dataset built on a **two-stage stratified clustered sampling**. The 2600 Primary Sampling Units (PSUs) were optimally selected based on latitude and landscape fragmentation and composition [5]. 5 Secondary Sampling Units (SSUs) were defined within a square 20-km × 20-km, one SSU at the centre and the others spread towards the box corners separated by a distance of 4-km from the centre. A total of 596, 811 and 1273 SSUs are therefore available as potential samples for the validation of the Amazon, Sahel and Siberia HRLC maps, respectively (Figure 8).

The **number of SSUs** per region will be defined according to eq. 1 using the overall accuracy of the selected RR algorithm per region and a confidence interval of 5%. SSUs will be randomly selected until an equal class distribution is reached.

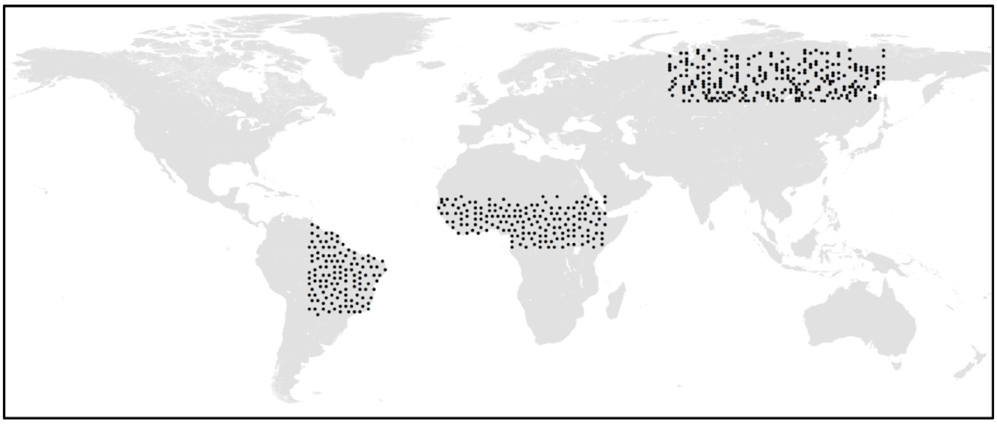


Figure 8. Potential sampling units derived from the TREES dataset from which actual samples will be selected to meet an equal class distribution.

5.2.3 For the HRLC historical maps

Considering the specificities of change validation (see section 4.2.3.2), the sampling should be stratified in space in order to ensure a significant representation of the areas known to be experiencing high rates of change [20] (in [8]) and in time, to account for reference data availability.

A **two-stage stratified sampling scheme** is therefore proposed, with 2005 considered as a sensible landmark year for data availability. Before 2005, scarce very high spatial resolution reference data imply defining sampling scheme driven by the availability of SPOT and Pleiades. After 2005, it becomes feasible to rely on a theoretical sampling that prioritizes the likely places for change to occur.

Based on the definition of change that will prevail in the CCI HRLC product, stratification layers will be generated based:

- a-priori knowledge of the location of hot-spots of change;
- distance to change (change is more likely to occur close to an already changed area);
- distance to roads;
- resistance to change (natural factors like altitude or slope lower the probability of change occurrence);
- etc.

Table 13 presents the number of samples per stratum, following eq. 1.

Table 13. Number of samples per stratum for the evaluation of the CCI HRLC historic LC maps derived from the choice of parameters from equation 1.

1st stage stratum	2nd stage stratum	p	q	E	Zα	n
1: 1990 - 2004	1.1: change	0.65	0.35	0.05	1.96	350
	1.2: no change	0.85	0.15	0.05	1.96	196
2: 2005- 2018	2.1: change	0.65	0.35	0.05	1.96	350
	2.2: no change	0.85	0.15	0.05	1.96	196

5.3 Response design

5.3.1 Choice of the sample spatial unit

The spatial resolution of the CCI HRLC RR and Static products is a pixel of 10 m × 10 m, with change detected at 30 m × 30 m spatial resolution for historical LC maps.

However, based on the lessons learnt from the GLC200, GlobCover and CCI MRLC experiences, it seems unwise to match the sample spatial unit of interpretation for the validation to the size of the pixel for the following reasons:

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

For the validation of the GLC2000 product, GlobCover and CCI MRLC products, blocks of 3 × 3 pixels, blocks of 5×5 pixels and blocks of 3×3 pixels were analyzed, respectively [5]. For the present exercise, it is envisaged to interpret plots of 3×3 pixels, of size of 10 m x 10 m for the validation of the CCI HRLC RR prototypes and static maps and of 30 m x 30 m for the validation of the CCI HRLC historical maps.

5.3.2 Labelling protocol

For each sampling unit, a set of attributes are recorded in a consistent manner. Table 14 presents an optimum attribute table designed for this validation exercise.

Table 14. Information included in the validation database for each SSU.

Field name	Details
For each sampling unit	
SSU ID	Unique identifier
PSU ID	Identifier of the associated PSU (if applicable)
Lat / Long	Centre coordinates of the observational unit to interpret
Level of certainty	Level of certainty (certain, reasonable, doubtful) associated with the interpretation of the expert
Land cover change	Presence/absence of LC change within the sample (if applicable)
Year of change	Year when LC change has occurrence. If not possible due to sparse temporal sampling of VHR imagery, an indication of the 5-year epoch of change (if applicable)
Comments	Comments given by the expert to explain/detail its interpretation (e.g. for indication why the labelling was not successful, or to give the local name used for the concerned LC type)
For each object in the sample and year/epoch	
Object ID	Unique identifier
SSU ID	Identifier of the associated SSU
PSU ID	Identifier of the associated PSU (if applicable)
Object geometry	Area and perimeter
Land cover class	The class ID of the LC legend (Table 4) associated with each object, for each year/epoch

Although error-free validation database do not exist, a rigorous validation protocol is a prerequisite to build it as closed as possible to “ground truth”. First, the legend of the LC and LCC classes should be described exhaustively and be well understood by the interpreter. The ground, as visible on very-high resolution imagery, can be complex spatially and include a mixture of classes within the sampling spatial unit or evolve temporarily in ways that requires auxiliary information to make accurate LC interpretations. A graphical validation interface is a valuable tool to gather evidences that effectively help the interpreter converge towards the best guess.

5.3.3 Graphical interface for image interpretation

The scale-independent validation tool developed for hosting this interpretation process, based on the experience gained during the previous validation exercises has been successfully used in the CCI MRLC validation and for the recent SIGMA validation experiment. The validation tool provides an online interface available to the expert on the reception of the URL. Figure 9 presents this interface, highlighting different functionalities. Note that time series of vegetation profiles can be generated with S2 at the sample centroid for recent years.

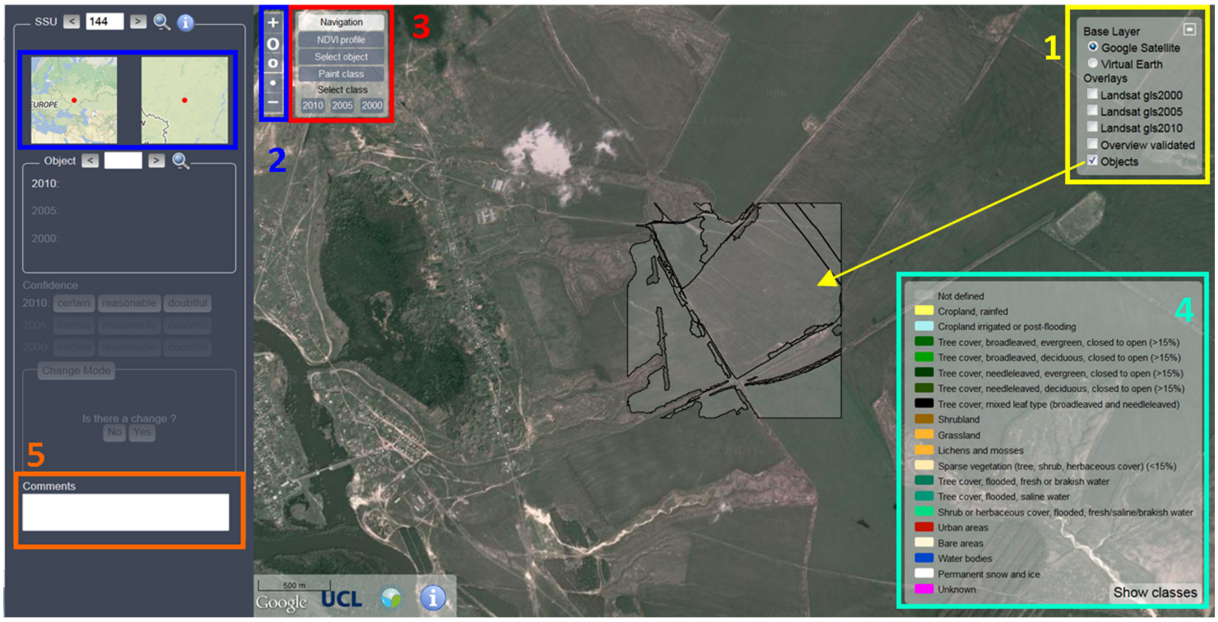


Figure 9. Main page of the validation tool, with the following functionalities: 1) Layer box to display different layouts; 2) Zooming functionalities; 3) Tools box to activate navigation, display NDVI profile, select objects or assign a LC class; 4) Legend description; 5) Comments box to include free text that should help understanding the labelling choices.

Figure 10 and Figure 11 illustrate the interfaces developed for the two main steps of the interpretation process: the LC classes' interpretation based on very high spatial resolution data available from Google Satellite/Virtual Earth and the LC change evaluation using the 3 Landsat TM or ETM+ images obtained from the GLS datasets.

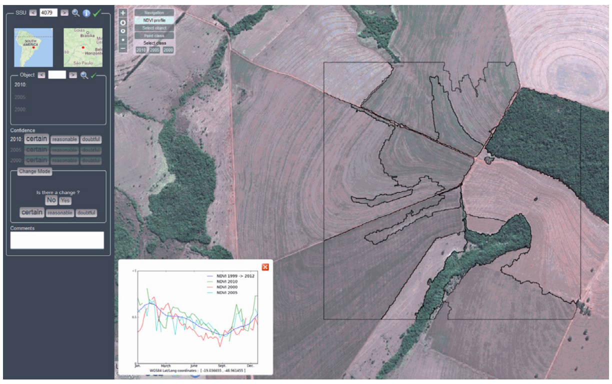


Figure 10. Example of segmented SSU over Brazil.

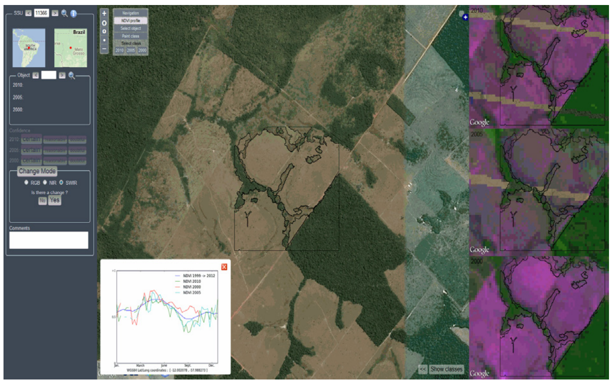


Figure 11. Example of segmented SSU over Brazil to be interpreted for 3 periods (the right panel providing the 3 Landsat images to validate the historical HRLC maps every 5 years.

6 Benchmarking with other existing products

The comparison with other HRLC products (based on international requirements) has the objective of building confidence in the CCI HRLC products, thus increasing their use (i.e. for non-climate model applications) and integrating the results into other land cover monitoring efforts. It is planned to perform benchmarking (inter-comparison) of static map and historical maps. However, the feasibility of this procedure highly depends on the availability of the other high-resolution land cover products with similar spatial and temporal resolution.

6.1 Existing high-resolution land cover maps

Collection of existing HRLC data is currently in its initial phase. The data collected so far are mostly global HRLC, because these datasets are of global interest, and usually available online or upon a request. On the opposite, local datasets are rarely available online, and sometimes subject to legal restrictions. Therefore, collection of local datasets is more challenging with respect to global datasets. Nevertheless, few local HRLC have been collected as a result of contacting national mapping authorities.

Besides the difficulties related to the collection of local HRLC, another critical point is availability of HRLC for benchmarking of historical maps, especially ones derived from satellite imagery with acquisition date older than year 2000.

Collected HRLC are shown in the Table 15.

Table 15. Existing HRLC for benchmarking

Name of LC map	Resolution	Year	Spatial coverage
GlobeLand30 (GL30) **	30 m	2000, 2010,2015	Global
FROM-GLC *	30 m	2010, 2015	Global
Global Urban Footprint ***	12 m	2011	Global
Global Human Settlement Layer (GHS BUILT-UP GRID S1) *	20 m	2016	Global
Global Surface Water *	30 m	1984 - 2015	Global
Forest / Non-Forest map *	25 m	2007 – 2010 2015 - 2016	Global
Tree canopy cover *	30 m	2000	Global
Global forest cover gain *	30 m	2000-2012	Global
Global forest cover loss *	30 m	2000-2015	Global
TerraClass Dataset*	30 m	2004, 2008, 2010, 2012, 2014	Brazilian Amazon
National LC of Suriname*	30 m	2015 and 2017	Country
National LC of Senegal**	30 m	2010	Country
National LC of Senegal**	10 m	2016	Country
Atlas of Urban Expansion*	30 m	1984-2015	Cities in Nigeria, Brazil, Uganda, Sudan, Ghana, Mali, Russia, Ethiopia

*Freely available

**Available upon agreement

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

6.2 Methodology

Benchmarking or inter-comparison will be pixel-by-pixel comparison of a project output and an existing HRLC. It means that agreement/disagreement between the two maps will be computed based on all pixels available in the region of interest in both the maps.

To compare two maps pixel-by-pixel it is necessary for maps to have same spatial resolution. Since it will not always be the case, the map with the lower resolution will be resampled to the resolution of the other map.

Moreover, for inter-comparison it is required that the two maps have the same legend. In the absence of a standard that regulates land cover legends, existing HRLC have different legends with respect to the legend that is planned for project outputs. Therefore, it will be needed to carefully match different legends. This may imply merging several classes of one map into one class in order to match legend of another map.

Inter-comparison will analyse both, similarities and dissimilarities between project outputs (static and historical maps) and existing HRLC.

Similarity between project output and existing HRLC will be computed based on accuracy indexes. More in detail, Overall Accuracy, Producer's Accuracy and User's Accuracy, described in the Section 3.4, will be computed. Although the names of the indexes are referring to the accuracy, in case of inter-comparison they will be expressing agreement between the products. Confusion matrix will serve as a base for computing mentioned indexes.

Dissimilarities or disagreement between project output and existing land cover map will be analysed in terms of disagreement pattern. Analyses will be done on difference image i.e. map of all disagreements, as well as on each combination of class disagreement separately. The aim is to understand if the disagreement has a meaningful pattern which can help in understanding a cause of disagreement. Furthermore, analysis of spatial pattern might be useful to point out if some areas are more prone to disagreement than the others.

Taking into consideration that the area of interest is rather large, and that inter-comparison is pixel-by-pixel, assessment of disagreement pattern will be computationally intensive. Therefore, in case that multiple datasets are available for the same region, disagreement pattern will be analysed for one dataset with the highest accuracy indexes (similarity).

7 Investigating scaling issues between the HRLC and MRLC products

The role of the spatial-temporal resolution on the consistency of the land cover classification will be investigated as a basis for understanding the variability in classifications as resolution spatially degrades. This research question will be addressed along with three different components.

First, a conceptual approach will deal with the land cover typology evolving across spatial scales to propose the most appropriate HRLC typology compatible with the already existing 300 m CCI MRLC time series. Secondly, an empirical approach will proceed based on the respective HRLC and MRLC maps to analyse the conditions of their consistency, to explain their discrepancies and to document their complementarities.

Thirdly, the resolution impact on the accuracy assessment metrics will be investigated thanks to two different methods specifically designed to quantify the scaling issues. On one hand, the Pareto boundary approach [21] will be used to disentangle the error related to the spatial resolution and the error related to the classifier. On the other hand, the ROC (Relative Operating Characteristics) Curve analytical framework [22] will support the validation and the scale impact assessment on the LC change detection across scales. Both analytical frameworks, i.e. the Pareto boundaries and the ROC curve, will be respectively applied on three subsets of 10 km x 10 km representative of the main landscape patterns of the 3 selected regions of interest. This empirical analysis should be supported by the VHR imagery described in section 3.1 which will be downscaled progressively to coarser resolution (e.g. from 1 m to 300 m) taking into account typical point spread function distribution.

7.1 LCML-LCCS3 nomenclature to address the semantic interoperability across scales

The graded and fuzzy nature of common LC categories derived from traditional classifications/legends has been recognized for a long time by the remote sensing community, however limited efforts to sort out semantic uncertainty in land cover studies has been proposed. As described in [23], the LCML-LCCS3 framework demonstrates that alternative solutions of LC data representation and functional management of their semantic Interoperability exists, this is particularly important with the increasing need to capture as much as possible the diversity of information of "real world" features. LCCS works by creating a set of standard attributes (called classifiers) to create or describe different LC classes. The classifiers act as standardized building blocks and can be combined to describe the more complex semantics of each LC class. Each LC class is no longer described by a name and a text but by a set of clearly quantifiable attributes. The CCI MRLC typology is built on the LCCS2 which prescribed a fixed threshold for each classifier. The evolution of LCCS2 to the LCML-LCCS3 takes advantage of the database flexibility. The meta-language operates by representing LC features in terms of the subtypes of LCML "basic elements" (or a set of them) with associated attributes and characteristics. The LCML "basic Elements" (LCElements) form the basic building blocks of any Land feature representation. This integrated system for LC observation provides worldwide consistency and links local and global levels of observation. Both have been recognized as international standards developed to address classification systems in general (ISO 19144-1 Classification Systems) and to address LC (ISO 19144-2 Land Cover Meta Language). The theoretical concept of the latter assumes that there is an observational continuity across all observation scales and products are consistent and compatible, using common approaches to characterize, describe and compare LC information (standardization, harmonization, and validation) and to facilitate the joint use of mapping products [24].

Both "building blocks" approaches allow translating an existing typology into another one, in spite of some ambiguities. For instance, the Mosaic classes, much reduced in the last version of the CCI MRLC dataset, cannot be allocated to one or another class in a rigorous manner and could rely partly on its spatial context to reduce the uncertainty. Based on this theoretical reflection, the Climate Users Requirements and the CCI MRLC typology, this first step has contributed to the definition of the HRLC typology relevant for the 10 to 30 m resolution products. The 10 m resolution is close to the typical size of common landscape features like a tree, a road, or a house introducing probably a difference with the 30m classically related to a forest cover, an impervious surface or an urban fabric. This should allow defining a first target HRLC typology shortly after the Climate Users Requirements deliverable and useable for the round robin activities.

The two most recent global reference databases, respectively developed by UCLouvain in the context of the CCI MRLC validation and by the FP7-SIGMA project, are scale-independent and rely on object-based delineated on VHR imagery. These databases will be exploited to assess the impact the different HRLC typologies at the different spatial resolution and their compatibility with the MRLC typology.

In a second step, the compatibility of these typologies will be tested directly from classified VHR images, progressively downgrade to 300 m resolution. In a third step, the HRLC reference database to be developed for the product validation will be used along the classified VHR images to quantify the proportion of the different landscape features in order to fine-tune the HRLC typology, to characterize quantitatively the MRLC typology, and to update the cross-walking table translating the LC classes into Plant Functional Types distribution [25]. The current absence of objective quantification has been recently identified as a key source of uncertainty in the climate model simulation.

7.2 Pareto boundary analysis to assess quantitatively the spatial resolution impact

Confusion matrices do not consider contextual influence of mixed pixels on the product accuracy [21]. Besides, when validating a coarse resolution product with an HR reference map, the assumption of equal spatial resolution between the reference and the product is violated. The Pareto boundary method is an alternative to deal with these shortcomings. The number of low-resolution pixels covering multiple classes is closely linked to the ground features (reference data) and is a function of their shape, size and spatial patterns [26], [27]. The difference in spatial resolution between high and low-resolution data is referred to as the low-resolution bias [21]. The resolution bias sets down the omission and commission errors (OE and CE) as conflicting objectives. Effectively, residual error after classification cannot be avoided. Any attempt to reduce the commission errors will inevitably lead to an increase of the omission errors and conversely.

Therefore, in the OE/CE bi-dimensional space, a region of unreachable accuracy limited by the Pareto boundary separates the errors due to the spatial resolution and the method. The Pareto boundary determines the maximum user and producer's accuracy values that could he attained jointly and represents such a lower limit as a boundary. To generate the Pareto boundary, the HR binary reference map is degraded to the low-resolution pixel size (Figure 12). Each new pixel value corresponds to the percentage of HR pixels of the class of interest. A set of low-resolution products is obtained by thresholding the low-resolution reference map. For each threshold defining the percentage for which a pixel is considered as vegetation, the pair of efficient error rates OE/CE is computed. The line joining all these points defines the Pareto boundary of a specific HR reference to a defined low-resolution pixel size. The distance between the product and the boundary indicates the performance of the method. The area under the efficient solution curve indicates the accuracy of the detection algorithm.

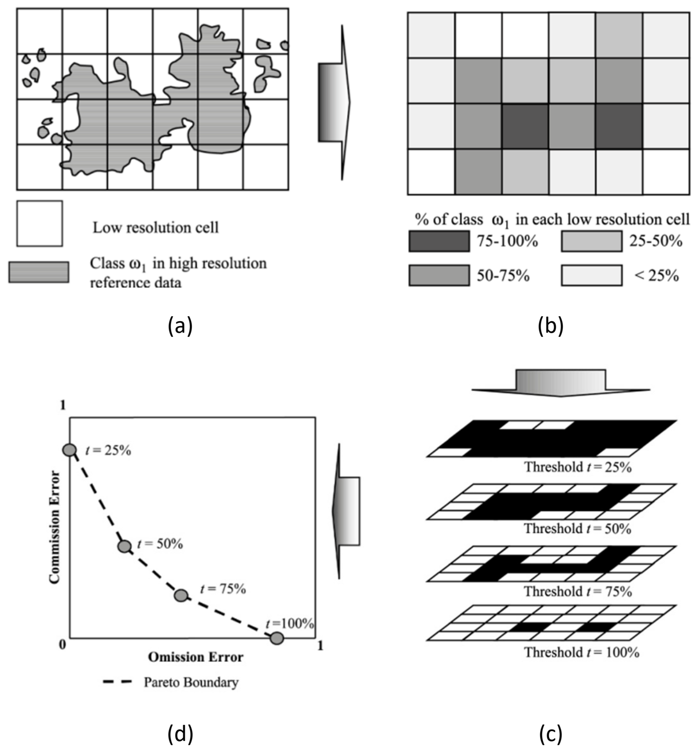


Figure 12. The procedure for generating a discrete set of points belonging to the Pareto Boundary, starting from the HR map, for a desired low spatial resolution. (a) A low-resolution grid is overlaid on the HR reference map. (b) The percentage of class ω_1 is computed for each cell. (c) A set of low-resolution products is generated by thresholding the percentage of class ω_1; threshold t varies in the interval (0, 1). These maps are efficient *solutions* according to Pareto's criterion. The map generated with *t* = 100% will have no commission (no mixed pixels included) but large omission, and the map with *t* = 1% will have no omission (all the mixed pixels are included) but large commission; in general, the higher *t*, the lower the commission and the higher the omission. (d) The confusion matrix is produced for each one of these maps. Omission error and commission error are derived and plotted in the omission error/commission error space [21].

As proposed by [28], the resolution-dependent error can be quantified based on the Pareto boundary approach of [21]. Its computation includes two steps. First, a Very High Resolution (VHR) map is aggregated to a coarser resolution so that the low-resolution pixels reflect the sub-pixel proportion of the class of interest. Second, the omission and commission errors that occur at different sub-pixel proportion thresholds are calculated. Once reported in the OE/CE bi-dimensional space, they delineate the Pareto boundary. Hypothetical Pareto boundaries for a higher and a coarser spatial resolution are presented in red and blue in Figure 13. The resolution error (*E)* is derived at the intersection of the 1:1 line with the Pareto boundary, i.e., omission and commission errors were given equal weights. The increase in resolution-dependent error (ΔE) when moving from a fine scale to a coarser one is represented by the length of the green line in Figure 13. It should be noted that the methodology does not require a threshold on the sub-pixel class proportion to define whether a pixel corresponds to one class or the other.

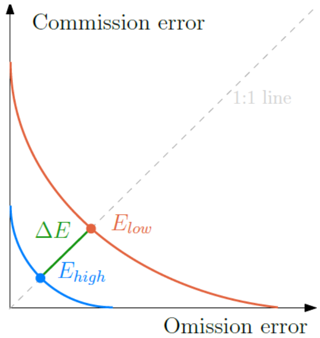
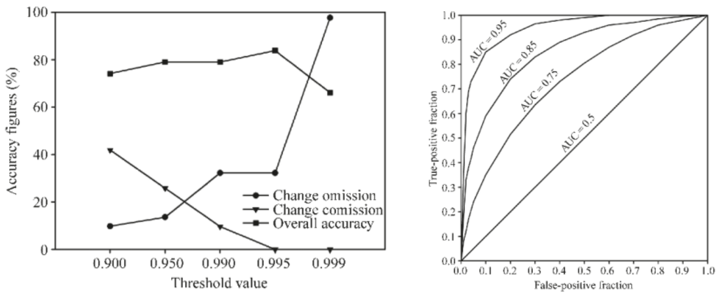


Figure 13. Pareto boundaries for a high and a low spatial resolution in red and blue, respectively. The area below the boundary is a region of unreachable accuracy because of resolution-dependent errors. The resolution error *E* corresponds to the intersection between the Pareto boundary and the 1:1 line. The distance between the resolution error for the low and the high resolution is the resolution-dependent error ΔE (represented in green) [28]**.**

7.3 ROC curves analysis to quantify the change detection performance at different scales

[22] introduced the concept of "accuracy assessment curve", which plots the accuracy figures derived from a standard confusion matrix (e.g., overall accuracy) against the threshold values used to classify the changed areas (Figure 14a). The different accuracy figures are continuously represented across the range of tested threshold values but there is no explicit link between them. In this respect, the ROC curve should receive considerable attention [29]. The ROC curve plots the false-positive fraction (i.e., the unchanged data erroneously classified as changed) against the true-positive fraction (i.e., the correctly classified changed data) for all possible threshold values (Figure 14b).

The optimal threshold value is not directly provided by the ROC curve but may be easily expressed in terms of ROC curve parameters. The curve shape illustrates the relationship between omission and commission errors and the area under the curve (AUC) measures the method's performance independently of the selected threshold value. The AUC provides a single integrated measure of overall accuracy that is not dependent upon a particular threshold value and that therefore gives information on the method's performance [30]].



(a) (b)

Figure 14. (a) Accuracy assessment curves, which show the relationship between the different accuracy figures and the threshold value (1- α) used to detect changed areas. (b) ROC curves, in which the true-positive fraction (on the y-axis) is

plotted against the false-positive fraction (on the x-axis) for all possible settings of the decision criterion. The area under the curve (AUC) informs about the method's performance, independently of the threshold value.

8 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.

[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.

[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.

[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.

[11] C. Lamarche *et al.*, "Compilation and validation of sar and optical data products for a complete and global map of inland/ocean water tailored to the climate modeling community," *Remote Sensing*, vol. 9, no. 1, 2017.

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

[13] 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.

[14] 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.

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

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

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

[18] 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.

[19] 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, 2013.

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

[21] L. Boschetti, S. P. Flasse, and P. A. Brivio, "Analysis of the conflict between omission and commission in low spatial resolution dichotomic thematic products: The Pareto Boundary," *Remote Sensing of*

Environment, vol. 91, no. 3–4, pp. 280–292, Jun. 2004.

[22] J. T. Morisette and S. Khorram, "Accuracy assessment curves for satellite-based change detection," *Photogrammetric Engineering and Remote Sensing*, vol. 66, no. 7, pp. 875–880, 2000.

[23] A. Di Gregorio *et al.*, "Land Cover Classification System: Classification Concepts. Software version 3," Rome (Italy) FAO, Rome, 2016.

[24] M. Herold, P. Mayaux, C. E. Woodcock, A. Baccini, and C. Schmullius, "Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1??km datasets," *Remote Sensing of Environment*, vol. 112, no. 5, pp. 2538–2556, May 2008.

[25] C. Ottlé, J. Lescure, F. Maignan, B. Poulter, T. Wang, and N. Delbart, "Use of various remote sensing land cover products for plant functional type mapping over Siberia," *Earth System Science Data*, vol. 5, no. 2, pp. 331–348, Nov. 2013.

[26] H. Eva and E. F. Lambin, "Remote sensing of biomass burning in tropical regions: Sampling issues and multisensor approach," *Remote Sensing of Environment*, vol. 64, no. 3, pp. 292–315, 1998.

[27] P. Mayaux and E. F. Lambin, "Estimation of tropical forest area from coarse spatial resolution data: a two-step correction function for proportional errors due to spatial aggregation," *Remote sensing of environment*, vol. 53, no. 1, pp. 1–15, 1995.

[28] F. Waldner and P. Defourny, "Where can pixel counting area estimates meet user-defined accuracy requirements?," *International journal of applied earth observation and geoinformation*, vol. 60, pp. 1–10, 2017.

[29] A. H. Fielding and J. F. Bell, "A review of methods for the assessment of prediction errors in conservation presence / absence models," *Environmental Conservation*, vol. 24, no. 1, pp. 38–49, 1997.

[30] J. M. DeLeo, "Receiver operating characteristic laboratory (ROCLAB): Software for developing decision strategies that account for uncertainty," in *2nd International Symposium on Uncertainty Modeling and Analysis*, 1993, no. 2, pp. 318–325.

[31] J. Delrue, L. Bydekerke, H. Eerens, S. Gilliams, I. Piccard, and E. Swinnen, "Crop mapping in countries with small-scale farming: A case study for West Shewa, Ethiopia," *International journal of remote sensing*, vol. 34, no. 7, pp. 2566–2582, 2013.

[32] C. Vancutsem, E. Marinho, F. Kayitakire, L. See, and S. Fritz, "Harmonizing and combining existing land cover/land use datasets for cropland area monitoring at the African continental scale," *Remote Sensing*, vol. 5, no. 1, pp. 19–41, Dec. 2013.

[33] M. J. Lambert, F. Waldner, and P. Defourny, "Cropland mapping over Sahelian and Sudanian agrosystems: A Knowledge-based approach using PROBA-V time series at 100-m," *Remote Sensing*, vol. 8, no. 3, 2016.

[34] G. L. Velthof *et al.*, "Grassland areas, production and use," *Lot 2. Methodological studies in the field of Agro-Environmental Indicators*, 2014.

9 Annex 1 - Investigating reference data source requirements

9.1 Ideal reference data requirements

- **Type of very-high-resolution imagery**

For each CCI HRLC product and regions (Amazon, Sahel and Siberia), Table 6 specifies the ideal requirements in terms of data quantity and time period concerned. Given the high quality and spectral resolution of the WorldView (WV) data, it is preferred for the validation of the RR prototypes of the year 2018, the static LC map 2018 as well as the validation of the scattered landscape of the Sahel back to 2010. For 2000 and 2005, and for 1995 and 1990, only IKONOS (I) and SPOT (S) data are available, respectively. Among the potential very high-resolution imagery, no preference is shown among WV, Pleiades (P) or GeoEye (G) data for the validation of the

Amazon and Siberia areas in 2015 and 2010. This would leave some flexibility for the repartition of TPM quota of data.

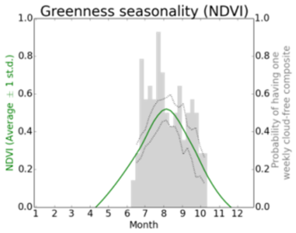
- **Number of VHR image footprints per region**

To validate each CCI HRLC product, various combinations of footprint quantities and number of images along the year have been defined, depending on the region and the vegetation seasonality. The static LC map 2018 corresponds to the largest extent and 50 footprints per region are suggested. Historical LC maps will be processed on a reduced extent and therefore 25 footprints would be sufficient, corresponding to less than 5% of the total area. The choice of the classification algorithms for the whole project will be selected on the accuracy of the RR prototypes. In this case, the suggested number of footprints remains 25 but covers 68% of the RR sites.

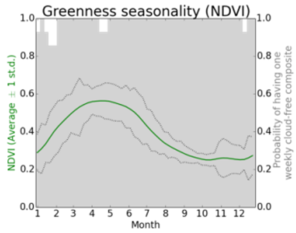
- **Number of images along the year**

Over these footprints, 1, 2 or 3 images distributed along the year would ideally be required for Siberia, Amazon, and the Sahel, respectively. These numbers would allow to visually interpreting distinct LC classes (e.g. the distinction between grassland and cropland) according to the vegetation seasonality and complexity of the landscapes. The number of images per year and the preferable months of acquisition are derived from the preliminary analysis of aggregated Normalized Difference Vegetative Index (NDVI) profiles, as illustrated in Figure 15. Each figure represents, in green, the seasonality of the vegetation present in Siberia, Amazon and Sahel and, in grey, the probability of having cloud-free observations along the year. In Siberia, clear observations are expected to be retrieved from June to October due to solar illumination constraints. One image acquisition along the year, centred in August, would be ideal. In Amazon, the seasonal pattern of vegetation along the year is more visible. Two images, acquired around September and around March would allow discriminating the land cover classes, even taking into account the high latitudinal gradient from North to South of the selected area (Figure 15), 2. (a) compared to 2. (c)). Sahelian landscapes show a high fragmentation, spatial heterogeneity of LC classes, high diversity of the cropping systems, mosaics of cropland, fallow and natural grassland [31]–[33] so that a set of 3 images along the year would be preferred. Vegetation profiles from Figure 15, 3. (a to c), show that one image around March, one around June and one around September could offer good discrimination of land covers. Table 6 summarizes the ideal data requirement for VHR imagery for the validation and scaling issue investigation.

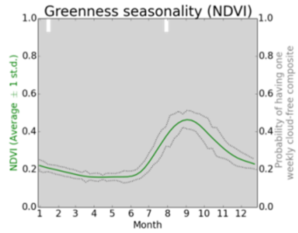
1. Siberia	2. Amazon	3. Sahel



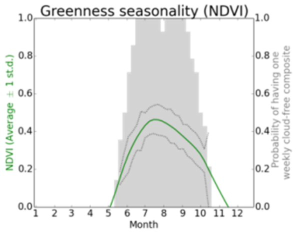
(a) LON/LAT=71.45/71.86



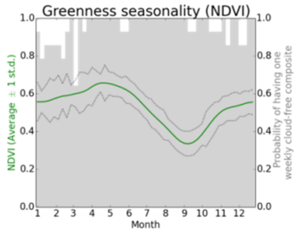
(a) LON/LAT=-38.77/-6.12



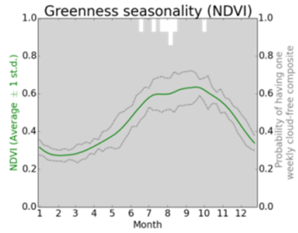
(a) LON/LAT= -7.57/ 13.88



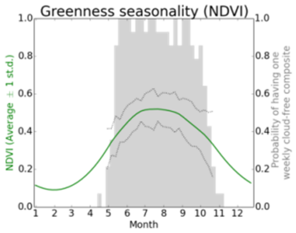
(b) LON/LAT= 75.89/66.25



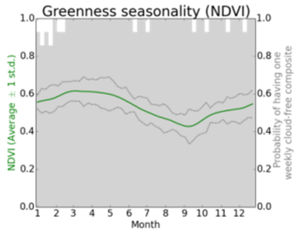
(b) LON/LAT= -58.85/-15.57



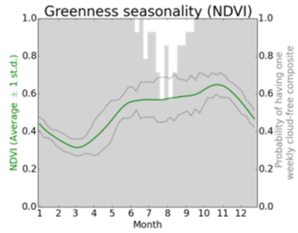
(b) LON/LAT= 21.26/9.13



(c) LON/LAT= 81.87/62.10



(c) LON/LAT= -53.27/-23.31



(c) LON/LAT= 36.90/8.25

Figure 15. Representative vegetation profiles from the three areas of LC map production defined in the ESA CCI HRLC project.

Table 16. Summary characteristics of the TPM data that could be used for the CCI+HRLC product evaluation. WorldView (WV), Pleiades (P), GeoEye (G), Ikonos (I) and SPOT (S). Multiple sources of VHR data are indicated by order of preference, starting from the ideal sensor. # stands for "number of".

LC map	Year	Region	Region area [km²]	Sensor	Im. footprint [km²]	# Footprints	Total Footprint [km²]	Proportion of region area	# im. per per year	Months [1-12]	TOTAL # images
Static LC map 2018	2018 or more recent	Amazon	11450200	WV	268.96	50	13448	0.001	2	3, 9	100
		Sahel	14099000	WV	268.96	50	13448	0.001	3	3, 6, 9	150
		Siberia	25763000	WV	268.96	50	13448	0.001	1	8	50
Historical LC	2015	Amazon	2230570	WV or P or G	268.96	25	6724	0.003	2	3, 9	50
		Sahel	2453620	WV	268.96	25	6724	0.003	3	3, 6, 9	75
		Siberia	3643260	WV or P or G	268.96	25	6724	0.002	1	8	25
	2010	Amazon	2230570	WV or P or G	268.96	25	6724	0.003	2	3, 9	50
		Sahel	2453620	WV	268.96	25	6724	0.003	3	3, 6, 9	75
		Siberia	3643260	WV or P or G	268.96	25	6724	0.002	1	8	25
	2005	Amazon	2230570	I	121	25	3025	0.001	2	3, 9	50
		Sahel	2453620	I	121	25	3025	0.001	3	3, 6, 9	75
		Siberia	3643260	I	121	25	3025	0.001	1	8	25
	2000	Amazon	2230570	I	121	25	3025	0.001	2	3, 9	50
		Sahel	2453620	I	121	25	3025	0.001	3	3, 6, 9	75
		Siberia	3643260	I	121	25	3025	0.001	1	8	25
	1995	Amazon	2230570	S	3600	25	90000	0.040	2	3, 9	50
		Sahel	2453620	S	3600	25	90000	0.037	3	3, 6, 9	75
		Siberia	3643260	S	3600	25	90000	0.025	1	8	25
	1990	Amazon	2230570	S	3600	25	90000	0.040	2	3, 9	50
		Sahel	2453620	S	3600	25	90000	0.037	3	3, 6, 9	75
		Siberia	3643260	S	3600	25	90000	0.025	1	8	25
Round Robin	2018	S2 tile 21KUQ	10000	WV	268.96	25	6724	0.672	2	3, 9	50
		S2 tile 21KXT	10000	WV	268.96	25	6724	0.672	2	3, 9	50
		S2 tile 42WXS	10000	WV	268.96	25	6724	0.672	1	8	25

9.2 Availability of ESA Third Party Mission open archives investigated for validation

ESA has a collection of VHR imagery archives that is available through the L-OADS online dissemination service at https://tpm-ds.eo.esa.int/oads/access/collection/. The ESA collection archive of Pleiades, SPOT1-5, SPOT 6-7, Rapideye, Deimos, IKONOS2 and the Tropforest 2010 dataset have been browsed in search of suitable images for the validation of the CCI HRLC products.

In total, the **SPOT1-5 ESA collection** includes 248 products covering the Sahelian extent of the static LC map and among those, 170 images are available over the reduced extent dedicated to the historic LC maps. Unfortunately, the Image distribution is clustered in space (Figure 16) and in time (Figure 17). SPOT images are evenly spread over the CCI HRLC extent: along the Atlantic coastline of Africa, at the border between Niger and Nigeria and in East Sudan. The most recent images available date back to 2011 and the temporal distribution of images does not allow validating the historical LC maps before 2003. As for the **SPOT 6-7 ESA collection**, three images are available over the Amazon static LC map extent, including 1 image falling in the historic LC map extent. Over Africa, only one image was found across the Mbam River in Cameroun.

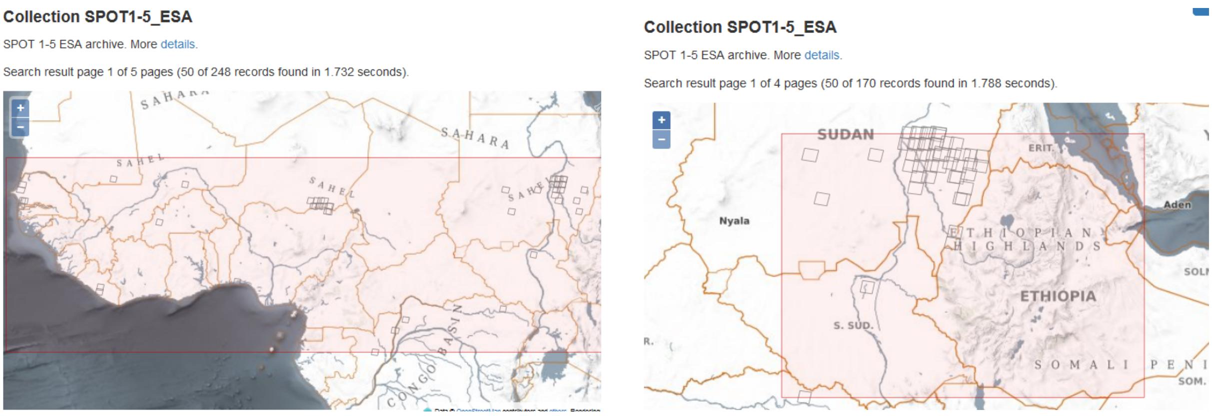


Figure 16. Spatial distribution of the images from the ESA SPOT1-5 archives over the Sahelian static LC map extent (left) and historic LC map extent (right).

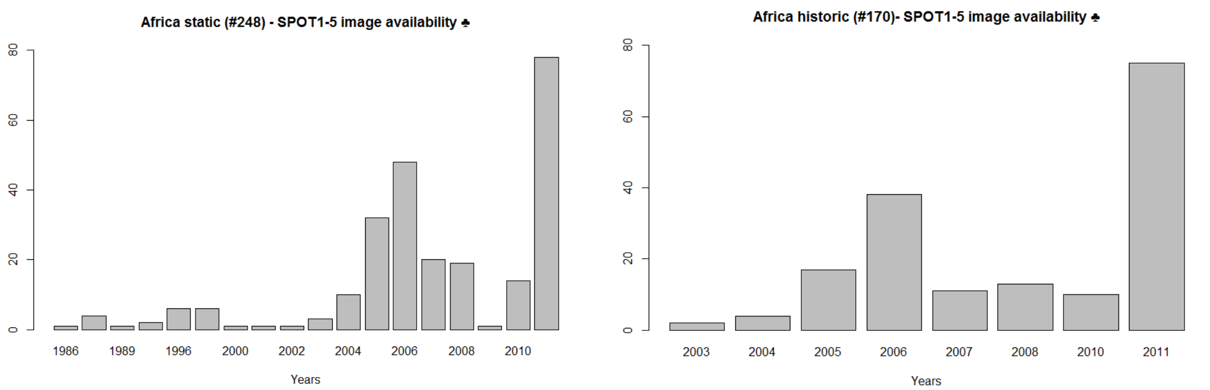


Figure 17. Temporal distribution of the images from the ESA SPOT1-5 archives over the Sahelian static LC map extent (left) and historic LC map extent (right).

The **ESA RapidEye collection** comes in a series of 4 datasets: the Planet - PlanetScope and SkySat data familiarisation phase, the RapidEye ESA archive, the RapidEye South America and the RapidEye time series for S2. No image from the *ESA Planet collection* can be used in the context of the CCI HRLC project because acquisitions are limited to three European demonstration sites. The *RapidEye ESA archive* includes 107 VHR imagery of which 83 are located in the Africa HRLC static extent and 24 in the Amazon HRLC static extent. Among these images, 17 and 8 images can be used to validate the historic LC maps of Africa and Amazon, respectively. The *RapidEye time series for Sentinel-2* consists in 5-day time series multispectral L3A orthorectified imagery at 5 m spatial resolution, acquired in 2013 and in 2015 and located over 24 sites around the world (Figure 18). Four times series acquired for sites in Mauritania, Ethiopia, Gabon and Congo can be used for the validation of the HRLC maps of Africa and one time series over Bolivia. As illustrated in Figure 19, images from the *ESA RapidEye South America* collection available from 2013 to 2015 overlap partially the extent of the Amazon HRLC maps, including Bolivia and Paraguay.

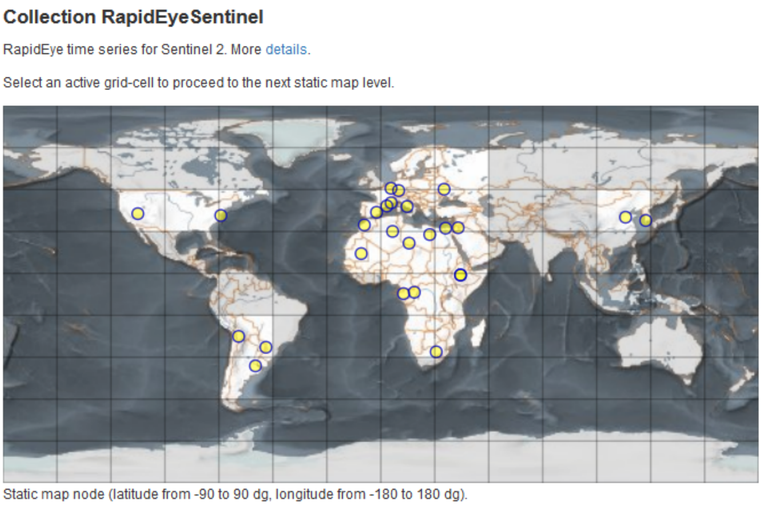


Figure 18. Spatial distribution of the images from the ESA RapidEye time series for Sentinel-2.

Figure 19. Spatial distribution of the images from the ESA RapidEye collection of South America.

Pleiades ESA collection offers three images centred on Kuru that could be used to validate the Amazon HRLC maps. The **Deimos 1-2** and **IKONOS 2** ESA collections were also investigated but none of the images covered our regions of interest (see Figure 20).

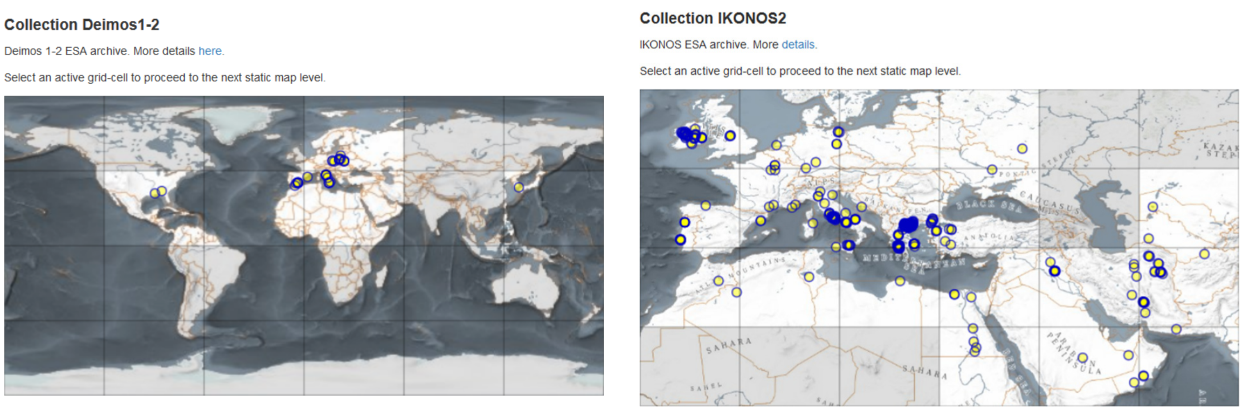


Figure 20. Spatial distribution of the images from the ESA Deimos 1-2 (left) and IKONOS 2 (right) collections.

The ESA category-1 project entitled "***TropForest 2010***" allowed the acquisition of high spatial resolution satellite imagery for South America and South East Asia for the 2010 period. For the 1˚ × 1˚

confluence points, satellite imagery from the Advanced Land Observing Satellite (ALOS) Advanced Visible and Near-Infrared Radiometer (AVNIR) 2 sensor (at 10 m × 10 m resolution) or from DEIMOS-1 sensor (at 22 m × 22 m resolution) have been be acquired (circa 85% covered by AVNIR-2 and 15% by DEIMOS-2). For the 2˚ × 2˚ confluence points, satellite at 4 m × 4 m resolution have been acquired (presently 60% of 2˚ × 2˚ confluence points are covered). This dataset is very much suited to validate the HRLC historic map over Amazon for the year 2010.

Figure 21. Location of the high-resolution satellite imagery available through the ESA CAT-1 "TropForest 2010" project (https://earth.esa.int/web/guest/-/tropforest).

10 Annex 2 - HRLC legend - level 2, level 3 and level 4.

Table 17 presents the 2nd level of LC classes of the CCI HRLC legend.

Table 17. Description of the level-2 categories of the HRLC legend.

LC class	Description
Shrub cover evergreen broadleaf	The evergreen shrubs are broadleaved and come from the Angiospermae group.
Shrub cover evergreen needleleaf	The evergreen shrub cover is composed of shrubs carrying typical needle-shaped leaves (Gymnospermae group) that are never entirely without green foliage [1]
Shrub cover deciduous broadleaf	The deciduous shrub cover is composed of broadleaved shrubs coming from the Angiospermae group.
Shrub cover deciduous needleleaf	The deciduous shrub cover is composed of shrubs carrying typical needle-shaped leaves (Gymnospermae group) that are never entirely without green foliage [1]
Natural grassland	Herbaceous areas not planted by humans but that can be influenced by human actions to some extent. Generally, low yields and high biodiversity value natural or semi-natural herbaceous covers can be grazed extensively [34].
Managed grassland (pastures)	Herbaceous areas sowed/planted by humans and that are sowed and planted at least twice a year
Winter crops	Sowed/planted herbaceous areas present at spring time
Summer crops	Sowed/planted herbaceous cover not present before spring time
Multicropping	Several crop cycles a year.
Unconsolidated bare areas	Bare areas with an unconsolidated aspect (e.g. bare soil, sands)

Consolidated bare areas	Bare areas with a consolidated aspect (e.g. made of rocks, hard pans)
Buildings	Built-up areas where buildings cover at least 50% of the surface
Artificial roads	Built-up areas where roads cover at least 50% of the surface
Snow	Permanent snow
Glaciers	A glacier is defined as a perennial mass of ice, and possibly firn and snow, originating on the land surface from the recrystallization of snow or other forms of solid precipitation and showing evidence of past or present flow

At third level, we add the rainfed and irrigated classes for the 3 crop categories (Table 18).

Table 18. Description of the level-3 categories of the HRLC legend.

LC class	Description
Rainfed	Sowed/planted herbaceous areas that are harvestable at least once within the 12 months after the sowing/planting date that depends on rainfall to grow
Irrigated	Sowed/planted herbaceous areas that are harvestable at least once within the 12 months after the sowing/planting date that depends on artificial water supply to grow

And at fourth level, the type of irrigation could be specified between sparkling and flooding categories (Table 19).

Table 19. Description of the level-4 categories of the HRLC legend.

LC class	Description
Sparkling	In sprinkler or overhead irrigation, water is piped to one or more central locations within the field and distributed by overhead high-pressure sprinklers or guns.
Flooding	In surface (flood or level basin) irrigation systems, water moves across the surface of an agricultural lands, in order to wet it and infiltrate into the soil.

11 Annex 3 - List of land cover transitions required by the climate modellers.

The main land cover changes on the Amazonian, African and Siberian regions are indicated in Table 20, Table 21 and Table 22, respectively.

Table 20. Transition classes expected on the Amazonian historical region (SW part of Amazonia). Double crosses indicate the transitions which should require more attention from the climate modellers point of view.

Year N	EBT 10	ENT 20	DBT 30	DNT 40	ShrE 50	ShrD 60	Grass 70	Crops 80	Flood 90	Li&Mo 100	Bare 110	Built 120	OpWs 130	OpWp 140	Sn&Ic 150
Year(N+1)															
10					x							x			
20															

30						X						X			
40															
50	X		X				X	X	X		X X	X			
60							X	X	X		X X	X			
70	X		X		X	X		X	X		X X	X			
80	X X		X X		X X	X X	X X		X		X X	X			
90	X		X		X	X	X	X							
100															
110	X		X		X	X	X	X				X			
120	X X		X X		X X	X X	X X	X X			X X				
130												X		X	
140													X		
150															

Table 21. Transition classes expected on the African historical region (Ethiopia). Double Crosses indicate the transitions which should require more attention from the climate modellers point of view.

Year N	EBT 10	ENT 20	DBT 30	DNT 40	ShrE 50	ShrD 60	Grass 70	Crops 80	Flood 90	Li&Mo 100	Bare 110	Built 120	OpWs 130	OpWp 140	Sn&Ic 150
YearN+1															
10					X		X	X				X			
20															
30						X	X	X				X			
40															
50	X		X								X X	X			
60	X		X								X X	X			
70	X		X		X	X		X			X X	X			
80	X X		X X		X X	X X	X X				X X	X			
90													X		
100															
110	X X		X X		X X	X X	X X	X X							
120	X X		X X		X X	X X	X	X			X X				
130									X					X	
140									X				X		
150															

Table 22. Transition classes expected on the Siberian historical region (Western Siberia). Double Crosses indicate the transitions which should require more attention from the climate modellers point of view.

Year N	EBT	ENT	DBT	DNT	ShrE	ShrD	Grass	Crops	Flood	Li&Mo	Bare	Built	OpWs	OpWp	Sn&Ic

	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150
YearN+1															
10															
20					XX										
30															
40						XX									
50		X	X	X			XX	X		X					
60		X	X	X			XX	X		X					
70		X	X	X	X	X		XX		XX	X	X			
80		X	X	X	X	X	X								
90															
100															
110		X	X	X	X	X	X	X				X			
120		X	X	X	X	X	X	X							
130					XX	XX	XX	XX			XX			XX	
140													XX		
150															