

# GMES Initial Operations / Copernicus Land monitoring services – Validation of products

Validation Services for the geospatial products of the  
Copernicus land Continental and local components  
including in-situ data (lot 1)

**Open Call for Tenders - EEA/MDI/14/010**

**Fourth Specific Contract - N°3436/R0-  
COPERNICUS/EEA.57889**

## RIPARIAN ZONES LC/LU 2018 – VALIDATION REPORT



**Document:** CLMS\_RZ\_LCLU\_Validation\_Report\_SC04

**Issue:** 1.3

**Date:** 17/11/2020

## Consortium Partners

<b>Prime Contractor</b>	SIRS SAS (France)
<b>Main Partners</b>	GAF AG (Germany), Joanneum Research (Austria), e-GEOS (Italy)
<b>Experts</b>	Specto Natura (UK), ISPRA (Italy), IGN France International (France), Terranea (Germany), NLSI (Iceland)

## Document Preparation and Release

	<b>Affiliation</b>	<b>Name(s)</b>	<b>Date</b>	<b>Signature</b>
<b>Author</b>	SIRS	Alexandre Pennec	29/05/2020	
<b>Contributions</b>	SIRS	Alice Lhernould	23/07/2020	
<b>Review</b>	SIRS	Alexandre Pennec, Christophe Sannier	31/07/2020	
<b>Approval</b>	EEA	Hans Dufourmont		

## Document Issue Record

<b>Issue</b>	<b>Date</b>	<b>Author(s)</b>	<b>Description of Change</b>
1.0	29/05/2020	Alexandre Pennec	Creation of the report
1.1	23/07/2020	Alice Lhernould, Alexandre Pennec	Update of the document
1.2	31/07/2020	Christophe Sannier	Review and minor corrections
1.3	17/11/2020	Alexandre Pennec	Integration of feedback from EEA and SPs

## Applicable Documents

ID	Title
AD01	Tender Specifications « Open Call for Tenders n° EEA/MDI/14/010 - Framework service contract for the GMES Initial Operations / Copernicus Land monitoring services – Validation Services for the geospatial products of the Copernicus land Continental and local components including in-situ data (lot 1)»
AD02	Framework service contract No EEA/MDI/14/010
AD03	Second Specific Contract N°3436/R0-COPERNICUS/EEA.56601
AD04	MAES Nomenclature, First, Second & Third Version
AD05	Nomenclature Guideline, First, Second & Third Version
AD06	Technical Note on the use of UA data (3 <sup>rd</sup> Specific Contract Task no. 18)
AD07	Final Delivery Report, Issue 3.0
AD08	Recoding tables UA-MAES and CLC-MAES
AD09	Third Specific Contract N°3436/R0-COPERNICUS/EEA.57056
AD10	Task 5: Revision and maintenance of the nomenclature guideline, Issue 1.3 (06/06/2018)
AD11	Fourth Specific Contract N°3436/R0-COPERNICUS/EEA.57889

## Executive Summary

This report provides the evaluation results of the Riparian Zones (RZ) for the reference year 2018 and change layer 2012-2018 (Strahler 2 to 9), specifically the Land Cover Land Use (LCLU) product. This analysis was performed over 28 Delivery Units (DU) during the fourth Specific Contract of the validation project (11 DUs from the 1<sup>st</sup> delivery of Task2 and 17 DUs in the 2<sup>nd</sup> delivery). The validation process mainly consisted in the implementation of the thematic accuracy assessment.

The results from the LCLU semantic checks performed by ETC/ULS and made available to the validation team revealed 4 DUs (out of 6) were considered as “good” and 2 DUs are “acceptable (redelivered after ETC checks). Both artificial / urban features and natural features were mapped with the same geometrical precision and detail of mapping features was good.

The thematic accuracy assessment was conducted in a two-stage process:

1. An initial blind interpretation in which the validation team did not have knowledge of the product’s thematic classes. However, the product polygon was provided to the validation team together with the point sample unit to consider boundary effects and geometric differences between the validation and production data
2. A plausibility analysis was performed on all sample units in disagreement with the production data to consider the following cases:
  - 1: Uncertain code, both producer and operator codes are plausible. Final validation code used is producer code
  - 2: Error from first validation interpretation. Final validation used is producer code
  - 3: Error from producer. Final validation code used is from first validation interpretation
  - 4: Producer and operator are both wrong. Final Validation code used is a new code from this second interpretation.

The **thematic validation of the LCLU 2018 products** shows **very good results** with an Overall Accuracy that meets the validation requirement in a plausibility analysis (with an **Overall Accuracy of 91.5% ±0.2**), even though the blind interpretation results show slightly but satisfying lower overall accuracies (88.7% ±0.2).

The **Change layer (2012-2018)** also shows good results at Level 4 with an Overall Accuracy of **86.2% ±0.3**.

These results in a plausibility analysis are in line with the assessment of the RZ LCLU product performed previously.

The differences between the blind interpretation and plausibility results highlights the complexity of the nomenclature of this product which is still probably on the edge of what can be extracted from the available imagery for certain classes. Nevertheless, it should be noticed that the results obtained for the blind interpretation for the reference year 2018 (Overall Accuracy of 88.7% ±0.2) are much greater than the Overall Accuracy obtained in the previous Specific Contracts for the reference year 2012 (**Overall Accuracy of 68.8% ±0.4** for the blind interpretation). The differences can be explained because PSUs from 2012 were re-interpreted for the reference year 2018 for which a plausibility had already been applied.

At Biogeographical region level, the results show that all the Overall Accuracies reach the minimum threshold required and at Delivery Unit level, most of the deliveries meet the minimum accuracy assessment (9 DUs under the threshold but with accuracy figures very close to 85%).

The analysis of the validation results at class and disaggregated level should provide insights on where the product could be improved and cases for which some thematic classes could be regrouped to improve their identification if required.

## Table of Contents

1.	Validation Framework .....	9
1.1.	Products to be validated.....	9
1.2.	Validation Criteria .....	12
2.	Validation approach .....	13
2.1.	Completeness .....	13
2.2.	Logical consistency .....	13
2.2.1.	Conceptual consistency.....	13
2.2.2.	Domain consistency .....	13
2.2.3.	Format consistency .....	13
2.2.4.	Topological consistency.....	14
2.2.5.	Additional logical consistency checks .....	14
2.3.	Positional Accuracy.....	14
2.4.	Thematic Accuracy.....	15
2.4.1.	Level of reporting .....	15
2.4.2.	Stratification and sample design.....	17
2.4.3.	Response Design.....	25
2.4.4.	Estimation and analyses procedures.....	26
2.5.	Temporal Quality .....	26
2.6.	Usability .....	26
2.7.	INSPIRE compliant metadata.....	26
3.	Validation check list.....	27
4.	Additional Tables.....	35
4.1.	Attributes.....	35
4.2.	Coordinate reference system description .....	36
4.3.	Nomenclature .....	37
5.	Thematic accuracy.....	39
6.	Conclusions.....	41
6.1.	Main findings .....	41
6.2.	Recommendations.....	41
Annex 1.	RZ LCLU confusion matrix for the blind and plausibility interpretation for the delivered DUs .....	28
		43

## List of Figures

Figure 1: RZ 2018 product over EEA39 for the reference year 2018 .....	10
Figure 2: Level of reporting at pan-European level.....	15
Figure 3: Level of reporting accordingly to the Delivery Units .....	16
Figure 4: Level of reporting accordingly to the biogeographical regions .....	16
Figure 5: Simple random (left) and random systematic (right) sampling designs.....	17
Figure 6: Replicates and sub-replicates used on LUCAS grid .....	18
Figure 7: RZ LCLU 2012 confusion matrix for the blind interpretation for the 28 delivered DUs .....	43
Figure 8: RZ LCLU 2012 confusion matrix for the plausibility interpretation for the 28 delivered DUs .....	43
Figure 9: RZ LCLU 2018 confusion matrix for the blind interpretation for the 28 delivered DUs .....	44
Figure 10: RZ LCLU 2018 confusion matrix for the plausibility interpretation for the 28 delivered DUs .....	44
Figure 11: RZ LCLU Change 2012-18 confusion matrix at MAES Level 1 for the blind interpretation for the 28 delivered DUs .....	45
Figure 12: RZ LCLU Change 2012-18 confusion matrix at MAES Level 1 for the plausibility interpretation for the 28 delivered DUs .....	45

## List of Tables

Table 1: Detailed specification of the Riparian Zones LCLU product .....	11
Table 2: RZ LCLU Delivery Units (DUs).....	19
Table 3: Distribution and surface area for each Delivery Unit on each Specific Contracts .....	21
Table 4: RZ LCLU distribution of sample units for each stratum et per Specific Contract.....	23
Table 5: RZ LCLU distribution of weight factors for each stratum and per Specific Contract .	24
Table 6: Class grouping for the LCLU accuracy assessment with the thematic rules and the old/new codes .....	25
Table 7: RZ LCLU attribute names and types .....	35
Table 8: Description of the RZ coordinate reference system .....	36
Table 9: RZ LCLU Nomenclature v5.0 (20/03/2020) .....	37
Table 10: Thematic Overall Accuracy (OA) for LCLU 2012, 2018 products and Change layer product: Blind and Plausibility interpretation, per delivery unit, per lot and bio-geographical region. For LCLU products: n is the number of observations, green color corresponds to OA greater than 85%, brown within 10% and red below 10% considering 95% CI. For Change product: n is the number of observations, green color corresponds to OA greater than 80%, brown within 10% and red below 10% considering 95% CI.....	40

## List of Abbreviations

ALP	Alpine region
ATL	Atlantic region
BLS	Black Sea region
BOR	Boreal region
CLC	CORINE Land Cover
CI	Confidence Interval
CON	Continental region
DU	Delivery Unit
EEA	European Environment Agency
ETC ULS	European Topic Centre Urban land and soil systems
EU-DEM	Digital Elevation Model over Europe
ESA	European Spatial Agency
FAO	Food and Agriculture Organization of the United Nations
FTY	Forest Type
GIO	GMES Initial Operations
GMES	Global Monitoring for Environment and Security
HRL	High Resolution Layer
IMD	Imperviousness Degree
JRC	Joint Research Centre
LAEA	Lambert Azimuthal Equal-Area
LCLU	Land Cover Land Use
LUCAS	Land Use/Cover Area frame Survey
MAES	Mapping of Assessment of Ecosystems and their Services
MED	Mediterranean region
MMU	Minimum Mapping Unit
MMW	Minimum Mapping Width
PAN	Pannonian region
PSIL	Parent Scene Identification Layer
PSU	Primary Sample Unit
RZ	Riparian Zones
SP	Service Provider



SC	Specific Contract
STE	Steppic region
TCD	Tree Cover Density
UA	Urban Atlas

# 1. Validation Framework

The validation framework is defined by a comprehensive analysis of the product specifications to determine the criteria to be used for the validation exercise.

## 1.1. Products to be validated

The Riparian Zones products represent transitional areas occurring between land and freshwater ecosystems, characterized by distinctive hydrology, soil and biotic conditions and strongly influenced by the stream water. They provide a wide range of riparian functions (e.g. chemical filtration, flood control, bank stabilization, aquatic life and riparian wildlife support, etc.) and ecosystem services.

Riparian Zones products cover the whole EEA39 countries (without French DOMS), plus Andorra and the Vatican City.

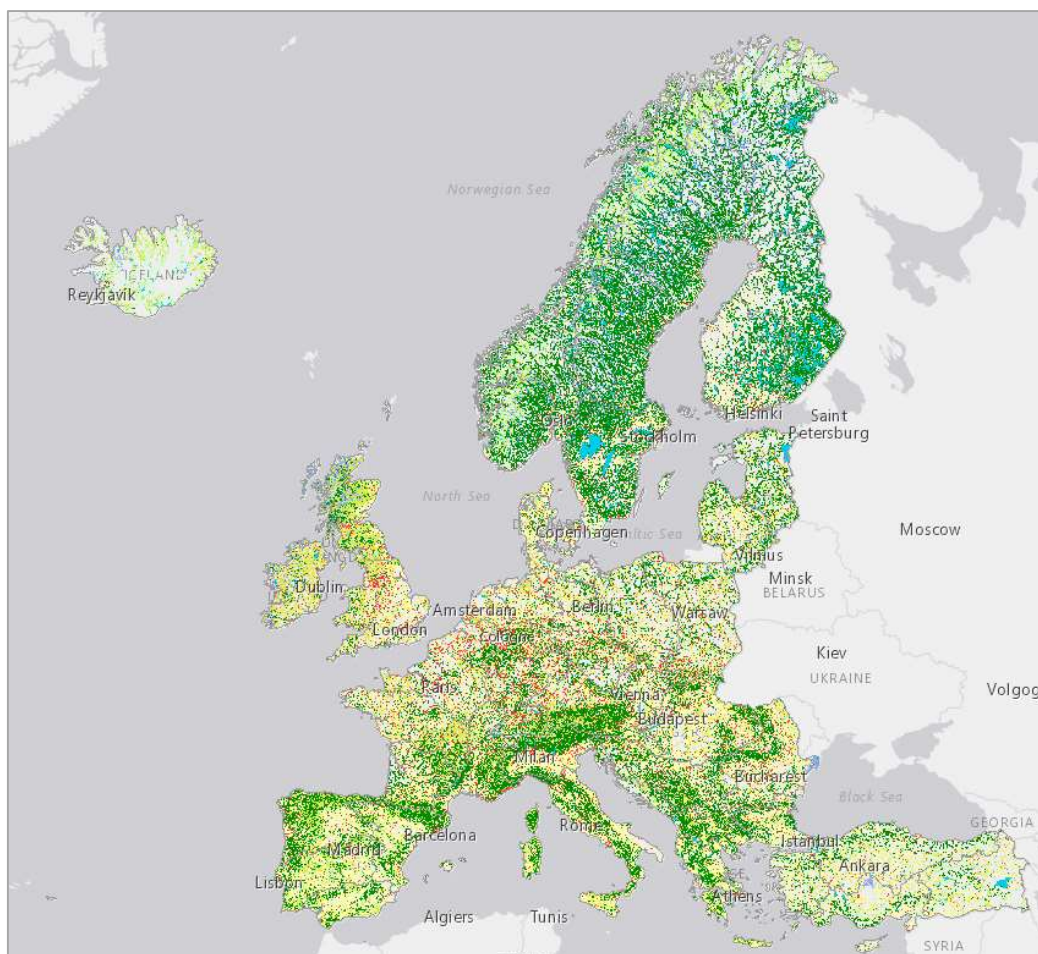
The main objective of Land Cover Land Use (LCLU) product is to support the Mapping and Assessment of Ecosystems and their Services (MAES) as part of the EU Biodiversity Strategy to 2020, but also the Habitats and Birds Directives and the Water Framework Directive. Its aim is to provide detailed LC/LU mapping dataset for areas along a buffer zone of selected rivers covering the EEA39:

- LC/LU layer for the reference years 2012 and 2018 on areas that are comprised of a merge of selected rivers with Strahler level 3 to 9 with different buffer sizes derived from the EU-HYDRO dataset and the area of the Pan-EU Flood Hazard Map produced by JRC (approx. 550,000 km<sup>2</sup>).
- LC/LU layer for the reference year 2012 and 2018 for Strahler level 2 (approx. 255,000 km<sup>2</sup>). By adding Strahler level 2 river segments for the reference years 2012 and 2018, the area covered by the Riparian zone mapping has been enlarged by about 35% to cover more than 805,000 km<sup>2</sup>.
- LC/LU change between 2012 and 2018 for Strahler level 2-9.

The 3 first Specific Contracts concentrate the validation on Strahler level 2 to 9 for the reference year 2012 and now the fourth Specific Contract emphasizes the exercise on the reference year 2018.

The territories mapped by Riparian Zones products are divided onto 43 Delivery Units (DU) based on aggregated sub-basins of EEA's European Catchments and Rivers Network Systems (ECRINS v1.1) and based on the GISCO boundaries and the shoreline from EU-Hydro database.

The LCLU classification is based on the MAES nomenclature which was revised multiple time during the four Specific contracts. The last revision of the nomenclature came up with 4 levels and up to 56 thematic classes with a Minimum Mapping Unit of 0.5 ha and 10 m of Minimum Mapping Width. The aim of the last revision is to harmonize the Local Component products (mainly Riparian Zones, Natura 2000, Urban Atlas and future Coastal Zones products)



**Figure 1: RZ 2018 product over EEA39 for the reference year 2018**

For the reference year 2012, the production of the LCLU product was mainly based on imagery dataset:

- DWH\_MG2b\_CORE\_03

Additional datasets were used to fill the gaps within the vegetation season of the relevant ESA DHW CORE\_03 VHR datasets such as:

- D2\_MG2b\_LOLA\_011b
- D2\_MG2b\_NARA\_011b
- DAP\_MG2b\_01
- DWH\_MG2b\_GEMS\_ADD\_003b
- VHR\_IMAGE\_2015 (only data acquired in 2014 for the Reference year 2012 since only data within the range 2010 - 2014 are eligible)

But other sources of multi-temporal imagery and geospatial products were used, such as Urban Atlas (UA) 2006/2012, Corine Land Cover (CLC) 2006/2012, the High-Resolution Layers 2012 (HRL) Imperviousness (IMD) and Tree Cover Density (TCD).

For the reference year 2018, a cloud free VHR dataset (VHR\_IMAGE\_2018) was available and used during the production. Additionally, VHR\_IMAGE\_2018\_ENHANCED dataset and Sentinel-2 were also used for gap filling.

The detailed specifications of the Riparian Zones LCLU product are shown below:

**Table 1: Detailed specification of the Riparian Zones LCLU product**

Product Title / Content	Product Short Name
Riparian Zones: Land Cover and Land Use Classification within buffer zone of selected rivers for reference years 2018 and 2012 including change mapping.	LCLU
<b>Product Definition</b>  <i>The Riparian LC/LU product provides a detailed LC/LU dataset for areas along a buffer zone of selected rivers of Strahler levels 2 – 9 covering EEA-39.</i>	
<b>Input Data Sources</b>  1) Riparian Zones (Str. 2-9).  2) Image data:  <b>2012</b>  <div style="display: flex; justify-content: space-between;"> <div> <b>Products:</b> <ul style="list-style-type: none"> <li>D2_MG2b_LOLA_011b</li> <li>D2_MG2b_NARA_011b</li> <li>DAP_MG2b_01</li> <li>DWH_MG2b_CORE_03</li> <li>DWH_MG2b_GEMS_ADD_003b</li> <li>VHR_IMAGE_2015</li> </ul> </div> <div> <b>Missions:</b> <ul style="list-style-type: none"> <li>GeoEye1 (2m)</li> <li>Pléiades (2.0m)</li> <li>SPOT-5 HRG (2.5m)</li> <li>SPOT-6 (1.5m)</li> <li>WorldView-2 (1.8m)</li> </ul> </div> </div> <b>2018</b>  <div style="display: flex; justify-content: space-between;"> <div> <b>Products:</b> <ul style="list-style-type: none"> <li>VHR_IMAGE_2018</li> <li>VHR_IMAGE_2018_ENHANCED</li> <li>Sentinel-2</li> </ul> </div> <div> <b>Missions:</b> <ul style="list-style-type: none"> <li>Pléiades 1A/B (2m)</li> <li>SPOT-6 (4 m)</li> <li>SPOT-7 (4 m)</li> <li>Kompsat-03/04 (2 m)</li> <li>SuperView (2 m)</li> <li>Planet DOVE (4 m)</li> <li>TripleSat (4m)</li> <li>Deimos-02 (4 m)</li> </ul> </div> </div> 3) Additional data:  CLC 2006/2012/2018; Urban Atlas 2006/2012/2018 (tbc); HRLs Imperviousness Degree and Tree Cover Density; DWH_MG2_CORE_01 Coverage 1 (IRS 20m) & 2 (RapidEye, 5m); Landsat-8, Sentinel 2, National orthophoto WMS, Google Earth Pro, Bing Maps; Numerous reference data sources.	
<b>Methodology</b>  Computer assisted visual refinement of the status 2018 data based on the LC/LU product for reference year 2012. The main dataset for visual image interpretation of reference year 2012 was the DWH_MG2b_CORE_03 dataset (Optical VHR2 coverage over EU 2011-2013 and Riparian zones), mainly consisting of 1.5m VHR2 SPOT-6 and 2.5m VHR2 SPOT-5 HRG and 2m Pléiades satellite data. Image interpretation for reference year 2018 used the VHR_IMAGE_2018 dataset (Optical VHR coverage of EEA-39 2017-2019), consisting of 2m Pléiades, Kompsat 3/3A, SuperView-1 and 4m SPOT-6/7, TripleSat,	
<b>Geographic Coverage)</b>  EEA-39 (without French DOMs) plus Andorra and Vatican City: ca. 5,819,516.50 km <sup>2</sup>	

Product Title / Content	Product Short Name
Riparian Zones: Land Cover and Land Use Classification within buffer zone of selected rivers for reference years 2018 and 2012 including change mapping.	LCLU
<b>Temporal Reference</b>  <i>Reference year 2012: 2010 – 2014; 2018: 2017 – 2019</i>	
<b>Geometric Resolution / Equivalent Scale</b>  <i>1:10.000</i>	
<b>Nomenclature</b>  <i>56 thematic classes</i>	
<b>Minimum Mapping Unit</b>  <i>0.5 ha</i>	
<b>Minimum Mapping Length</b>  <i>N/A</i>	
<b>Minimum Mapping Width</b>  <i>10m</i>	
<b>Thematic/Positional Product Accuracy</b>  <i>Overall thematic accuracy demanded is &gt; 85% taking into account the relative occurrence of the LC/LU classes for status layers of reference years 2012 and 2018 and &gt; 80 % for change layer 2012-2018.</i>	

## 1.2. Validation Criteria

The validation exercise focuses on thematic accuracy. The RZ LCLU overall accuracy is expected to be greater than 85% taking into account the relative occurrence of the LC/LU classes for status layers of reference years 2012 and 2018 and greater than 80 % for change layer 2012-2018.

## 2. Validation approach

The validation approach provides guidance on how the products were validated by defining suitable indicators or metrics. Detailed completeness and logical consistency checks are already performed as part of the semantic checks undertaken by the Service Provider since the QC Tool was not available during upload of the products. Therefore, the aim of this validation exercise is not to repeat these, but to review the existing documentation and perform additional checks if deemed necessary. We recommend however to make the QC Tool reports available openly as additional quality information to the users.

Thematic accuracy represents the bulk of the work undertaken as part of the validation.

The quality assessment is performed according to INSPIRE Data Specifications. The data quality elements considered are: (i) Completeness, (ii) Logical Consistency, (iii) Positional Accuracy, (iv) Thematic Accuracy, (v) Temporal quality and (vi) Usability. Each of them forms a section in the Validation Check list.

### 2.1. Completeness

Completeness should only be performed when the product is complete.

**Description:** For non-thematic raster products (Image mosaics & EU-DEM), completeness provides an indication for missing data or omission within the intended area. For Land Cover and Land Use products (both raster & vector), the notion of completeness in INSPIRE provides an indication of omission and commission errors.

**Indicators:** the rate of excess items is used for commission errors and the rate of missing items is used to verify omission errors.

### 2.2. Logical consistency

Logical consistency evaluates the degree of adherence to logical rules of data structure, attribution and relationships. In INSPIRE Data Specifications, Logical Consistency comprises four sub-elements described hereafter: conceptual consistency, domain consistency, format consistency and topological consistency.

#### 2.2.1. Conceptual consistency

**Description:** indicates that the data structure follows the data specifications in terms of data model and relationships.

**Indicators:**

- Type of feature used
- Minimum Mapping Unit (MMU)
- Coordinate Reference System
- The presence of a unique identifier for each feature
- Nomenclature used

#### 2.2.2. Domain consistency

**Description:** involves the detection of attribute values that are outside the pre-defined range of values. For vector data each attribute has a pre-defined set of range of values. For raster data, the correct encoding of data is checked.

**Indicator:** Value domain non-conformance: number of items not in conformance with their expected value domain.

#### 2.2.3. Format consistency

**Description:** includes detection of file format, file or attribute names or attribute types which do not correspond to the specifications. In addition, for raster data the pixel depth is also considered here.

#### Indicators:

- File format conformance
- File name conformance
- Attribute names conformance
- Attribute types conformance

### 2.2.4. Topological consistency

**Description:** topological consistency is applicable to vector data and describes the degree of correctness of the topological characteristics described in the product specification of the dataset.

#### Indicators:

- Number of overlaps
- Number of gaps
- Number of multipart features
- Number of neighboring features
- Number of self-intersections
- Number of null geometries
- Number of unclosed rings
- Number of duplicate vertex
- Number of pseudo nodes
- Number of non-matching nodes

### 2.2.5. Additional logical consistency checks

- **Map projection:** the conformity of the map projection parameters is also checked.

## 2.3. Positional Accuracy

Detailed positional accuracy as described below is only required for the validation of image mosaics. Positional accuracy of UA2012 and RZ is directly related to the positional accuracy of the underlying VHR imagery. Positional accuracy of the HRLs is directly related to the underlying HR imagery.

Visual checks were undertaken in relation to imagery used for validation and during production.

## 2.4. Thematic Accuracy

### 2.4.1. Level of reporting

The level of reporting for the validation results is at pan-European level. However, results are also provided at different levels of aggregation. The analysis at disaggregated levels will contribute to assess regional differences, if any, and the nature of these differences.

The internal validation needs to find a compromise between the number of sample units and representativeness of the results at sub-European level. Therefore, the envisaged levels of reporting are:

- Pan-European



Figure 2: Level of reporting at pan-European level



- Major River Basin related Delivery Units (DU)

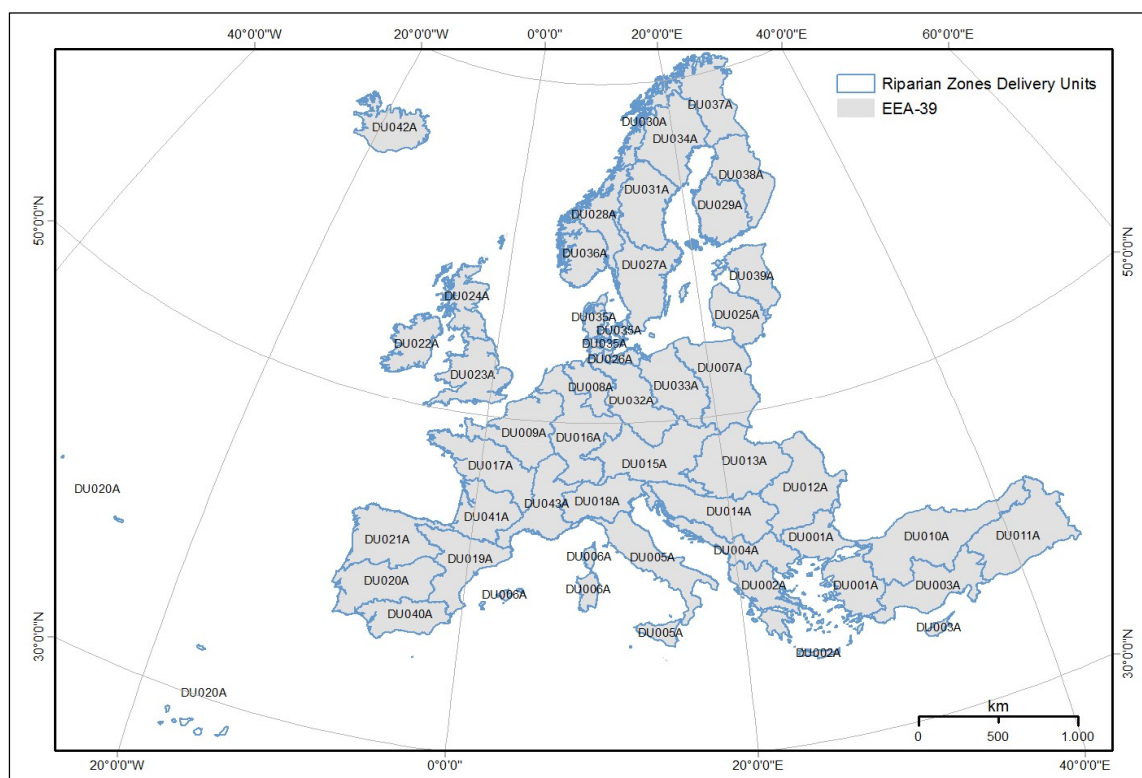


Figure 3: Level of reporting according to the Delivery Units

- Biogeographical regions 2016

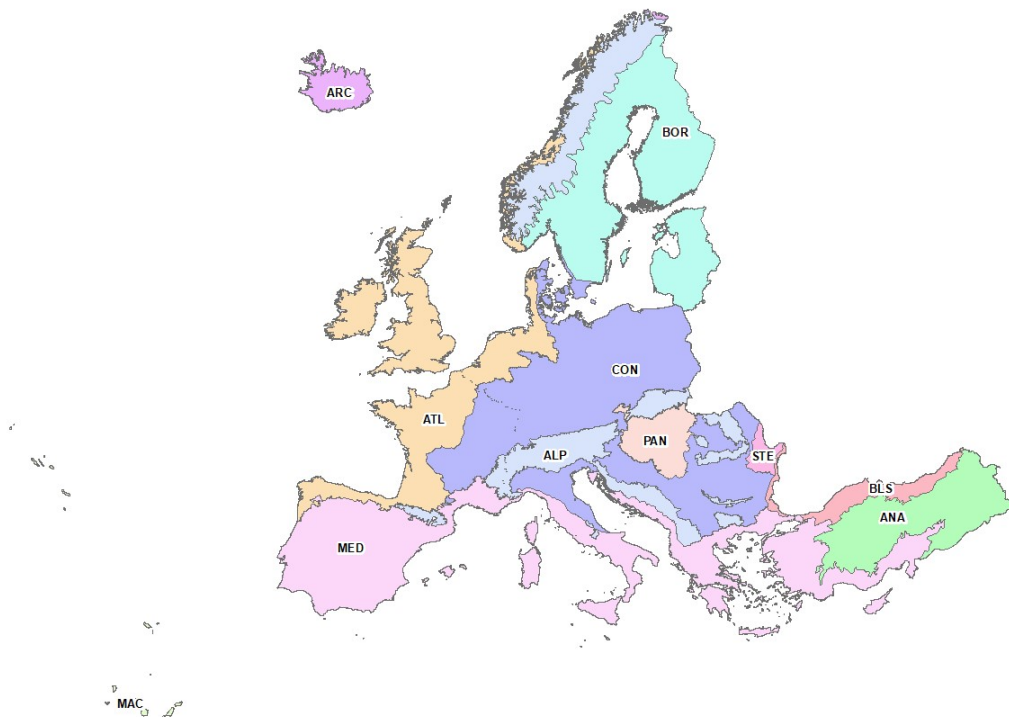


Figure 4: Level of reporting according to the biogeographical regions

## 2.4.2. Stratification and sample design

### 2.4.2.1. Overview

The stratification and the sampling design primarily consist in selecting an appropriate sampling frame and sampling unit. The sampling units can either be “defined on a cartographic representation of the surveyed territory” (Gallego, 2004<sup>1</sup>), in which case it is an area frame, or on a list of the features. According to (Gallego, 2004), area frames give a better representation of the population as the spatial dimension is kept.

In an area frame, sample units can be points, lines (often referred to as transects) or areas (often referred to as segments, described by Gallego, 1995<sup>2</sup>). The first step is to define the geographical area for which the accuracy assessment is to be reported and the type of sample units. For the majority of cases, point samples will be used, but areas or segments may be used in specific cases such as when not only thematic accuracy needs to be reported, but also the geometry of mapped objects. Points are considered as the most appropriate unit for our purpose. Polygons have also the drawback of being specific to a single map. In case of changes, the sample may not be adapted anymore.

Sampling design refers to the protocol whereby the samples are selected. A probability sampling design is preferred for its objectivity. “Simple random, stratified random, clustered random and systematics designs are all examples of probability sampling designs” (Stehman et al., 1998<sup>3</sup>). Even though a simple random design is easy to implement, its main drawback is that some portions of the population may not be adequately sampled. Cluster sampling is often used to reduce the costs of the collection of reference data but does not resolve geographic distribution problems. A systematic approach would solve this problem, yet it is not appropriate if the map contains cyclic patterns. A stratified approach consists in allocating a pre-defined number of samples per land-cover class. As explained in (Stehman et al., 1998), stratification ensures that each class is represented.

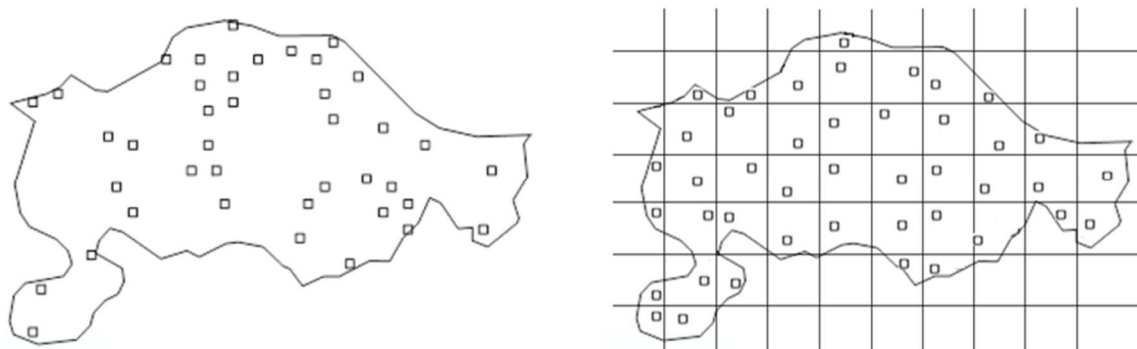


Figure 5: Simple random (left) and random systematic (right) sampling designs

A stratified systematic sampling approach based on LUCAS (Land Use/Cover Area frame statistical Survey) is used for all thematic layers adapting the number of replicates to each stratum. The choice of this approach was made for the following reasons:

<sup>1</sup> Gallego, J. and JRC-IES, I.I., 2004. Area Frames for Land Cover Estimation: Improving the European LUCAS Survey. Presented at the Proceedings of the 3rd World Conference on Agricultural and Environmental Statistical Application, Cancun, Mexico, 2–4.

<sup>2</sup> Gallego, F.J., 1995. Sampling frames of square segments. Office for Official Publ. of the European Communities.

<sup>3</sup> Stehman, S.V. and Czaplewski, R.L., 1998. Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sensing of Environment*, 64 (3), 331–344.

- A systematic approach ensures full traceability of the results compared to a random approach in which the random selection of samples is difficult to trace
- Some of the sample units will be shared amongst different data layers fostering economies of scale for the validation of Copernicus Land pan-European and local geospatial products
- Some of the sample units will also be surveyed on the ground as part of LUCAS

LUCAS corresponds to a grid of approximately 1,100,000 points throughout the European Union. The LUCAS sampling is densified for small strata.

A set of 81 points located on an 18x18 km square constitutes a group (red points shown in Figure 6) in which every point is associated with a number comprised between 1 and 81 (the numbers do not follow each other spatially). The same pattern with the same numbers allocation is repeated all over the grid. A replicate refers to the points with the same number selected on the whole LUCAS grid.

At first, the number of samples to allocate to each stratum (or MAES4 land cover class) was calculated as a function of their area. In this manner the sampling design is not only systematic but also stratified. A minimum number of sample units per stratum was defined to ensure that even small strata are represented in the sample.

The number of replicates to be selected for a stratum depends on its area and the number of LUCAS points intersecting the stratum.

For land cover classes covering a large proportion of the study area, 1 replicate may already exceed the defined number of samples for this class. To solve this problem, replicates are split into four sub-replicates, as illustrated by the blue numbers in Figure 6.

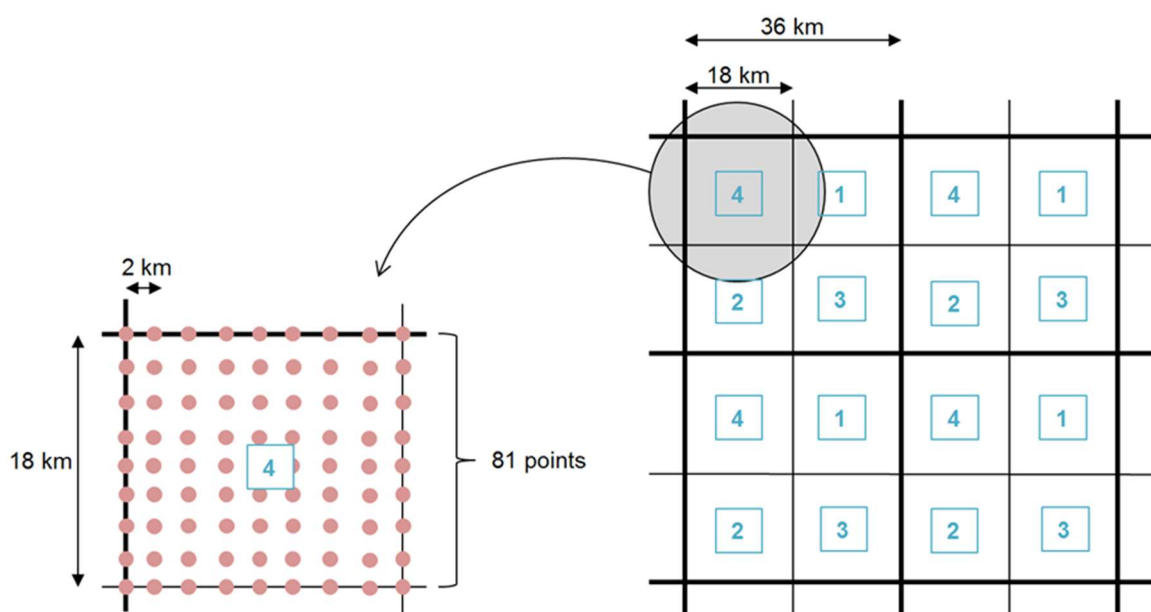


Figure 6: Replicates and sub-replicates used on LUCAS grid

The opposite problem is encountered for land cover classes covering a small proportion of the study area: even by selecting 81 replicates (the maximum number), the intersecting area between the stratum and LUCAS points is too small to reach the required number of samples. Therefore, LUCAS grid was densified by creating one point every 200 m.

## 2.4.2.2. Stratification approach

For the RZ LCLU product, the stratification is applied based on the LCLU classes and the map products.

The number of sample units per stratum should be such to ensure a sufficient level of precision at reporting level. The minimum number of sample units per stratum should be set at 5 if possible. The stratification and sample selection cover the whole study area in order to be valid.

For the RZ LCLU products the thematic validation is not performed on the thematic classes including density levels derived from HRL IMD as this layer is validated separately.

The Fourth Specific Contract includes 28 DU of the initial 43 DUs of the RZ LCLU product, covering around 60% of the total areas. Delivery Units covered by the Fourth Specific Contract are listed in Table 2.

**Table 2: RZ LCLU Delivery Units (DUs)**

Riparian Zones Delivery Units			
No.	DU ID	Catchment Name(s) per Unit	RZ Aol Str. 2-9 [km <sup>2</sup> ]
1	DU001A	Aegean Islands, Black Sea Basin District, Turkey_West	17.003,02
2	DU002A	Attica, Crete, Eastern Peloponnese, Eastern Sterea Ellada, Seman, Thessalia, Western Macedonia, Western Sterea Ellada	14.627,61
3	DU003A	Turkey_South, Cyprus	12.048,10
4	DU004A	Adriatic Sea coastal catchments and small basins, Central Macedonia, Crni Drim / Drin, Drini i Zi, Neretva, North Adriatic, West Aegean Region Basin District	15.487,18
5	DU005A	Middle Appenines, Serchio, Sicily, Southern Appenines	25.454,12
6	DU006A	Balearic Islands, Corsica, Sardinia	5.330,56
7	DU007A	Dniestr, Pregolya, Vistula	22.657,55
10	DU010A	Turkey_North	17.425,68
13	DU013A	Danube_North	46.978,77
19	DU019A	Ebro, Internal Basins of Catalonia, Jucar	19.386,97
22	DU022A	Eastern, IE South Eastern, IE South Western, Neagh Bann, North Eastern, North Western, Shannon	9.537,20
23	DU023A	Anglian, Dee, Humber, North West, Thames, UK South East, UK South West, Western Wales, Isle of Man	14.400,28
24	DU024A	Northumbria, Scotland, Solway Tweed	13.839,72
25	DU025A	Lielupe, Nemunas, Venta	9.725,29
26	DU026A	Schlei/Trave, Vidaa-Krusaa, Warnow/Peene	2.821,89
27	DU027A	North Baltic, Skagerrak and Kattegat, South Baltic	36.208,84
28	DU028A	Glomma, Moere and Romsdal, Troendelag	12.664,46
29	DU029A	Kokemoenjoki-Archipelago Sea-Bothnian Sea, Kymijoki-Gulf of Finland	19.171,74

Riparian Zones Delivery Units			
No.	DU ID	Catchment Name(s) per Unit	RZ AoI Str. 2-9 [km <sup>2</sup> ]
31	DU031A	Bothnian Sea	26.224,37
34	DU034A	Tornionjoki (Finnish part)	27.538,23
35	DU035A	Jutland and Funen, Zealand	3.965,58
37	DU037A	Finnmark, Kemijoki, Teno-, Nootom- and Paatsjoki (Finnish part)	17.137,80
38	DU038A	Oulujoki-Iijoki, Vuoksi	27.767,52
39	DU039A	Daugava, East Estonia, Gauja, West Estonia	10.152,22
40	DU040A	Andalusia Atlantic Basins, Andalusia Mediterranean Basins	13.589,11
41	DU041A	Adour, Garonne, Dordogne, Charente and coastal waters of Aquitania	15.198,04
42	DU042A	Iceland	7.527,65
43	DU043A	Rhone and Coastal Mediterranean	21.910,16

As shown in Table 3, most of the Delivery Units were already sampled during the previous Specific contracts. In details, all the DUs from Strahler Level 2-8 were covered during the two first Specific contracts and 16 out of 48 DUs were covered for Strahler Level 9 during the Third Specific contract. The actual sampling scheme take advantage of stratification and the sampling design already performed, and the sample units interpreted for the reference year 2012 can be mostly reused for the reference year 2018. The Table 3 shows the DUs that still need to be sampled for Strahler Level 9, covering 81,312 km<sup>2</sup> out of 485,780 km<sup>2</sup> of the total areas covered by the Fourth Specific Contract.

Table 3: Distribution and surface area for each Delivery Unit on each Specific Contracts

Riparian Zones Delivery Units									
No.	DU ID	Catchment Name(s) per Unit	RZ Aol Str. 2-9 [km²]	RZ Aol Str. 2-8 [km²]	RZ Aol Str. 9 [km²]	Incl. SC01-02	Incl. SC03	RZ Aol Str. 2-8 sampled [km²]	RZ Aol Str. 9 sampled [km²]
1	DU001A	Aegean Islands, Black Sea Basin District, Turkey_West	17 003,02	10 366,28	6 636,74	yes	no	10 366,28	-
2	DU002A	Attica, Crete, Eastern Peloponnese, Eastern Sterea Ellada, Seman, Thessalia, Western Macedonia, Western Sterea Ellada	14 627,61	9 194,39	5 433,22	yes	yes	9 194,39	5 433,22
3	DU003A	Turkey_South, Cyprus	12 048,10	6 950,23	5 097,87	yes	no	6 950,23	-
4	DU004A	Adriatic Sea coastal catchments and small basins, Central Macedonia, Crni Drim / Drin, Drini i Zi, Neretva, North Adriatic, West Aegean Region Basin District	15 487,18	10 445,75	5 041,43	yes	no	10 445,75	-
5	DU005A	Middle Appenines, Serchio, Sicily, Southern Appenines	25 454,12	15 915,40	9 538,72	yes	yes	15 915,40	9 538,72
6	DU006A	Balearic Islands, Corsica, Sardinia	5 330,56	2 793,79	2 536,77	yes	yes	2 793,79	2 536,77
7	DU007A	Dniestr, Pregolya, Vistula	22 657,55	17 318,66	5 338,89	yes	no	17 318,66	-
10	DU010A	Turkey_North	17 425,68	10 662,95	6 762,73	yes	no	10 662,95	-
13	DU013A	Danube_North	46 978,77	40 960,98	6 017,79	yes	yes	40 960,98	6 017,79
19	DU019A	Ebro, Internal Basins of Catalonia, Jucar	19 386,97	12 801,42	6 585,55	yes	yes	12 801,42	6 585,55
22	DU022A	Eastern, IE South Eastern, IE South Western, Neagh Bann, North Eastern, North Western, Shannon	9 537,20	5 942,54	3 594,66	yes	no	5 942,54	-
23	DU023A	Anglian, Dee, Humber, North West, Thames, UK South East, UK South West, Western Wales, Isle of Man	14 400,28	8 410,13	5 990,15	yes	yes	8 410,13	5 990,15
24	DU024A	Northumbria, Scotland, Solway Tweed	13 839,72	7 585,44	6 254,28	yes	yes	7 585,44	6 254,28
25	DU025A	Lielupe, Nemunas, Venta	9 725,29	5 586,03	4 139,26	yes	yes	5 586,03	4 139,26
26	DU026A	Schlei/Trave, Vidaa-Krusaa, Warnow/Peene	2 821,89	1 800,37	1 021,52	yes	yes	1 800,37	1 021,52
27	DU027A	North Baltic, Skagerrak and Kattegat, South Baltic	36 208,84	25 475,64	10 733,20	yes	no	25 475,64	-
28	DU028A	Glomma, Moere and Romsdal, Troendelag	12 664,46	7 529,90	5 134,56	yes	no	7 529,90	-
29	DU029A	Kokemoenjoki-Archipelago Sea-Bothnian Sea, Kymijoki-Gulf of Finland	19 171,74	13 968,23	5 203,51	yes	yes	13 968,23	5 203,51
31	DU031A	Bothnian Sea	26 224,37	14 988,76	11 235,61	yes	no	14 988,76	-
34	DU034A	Tornionjoki (Finnish part)	27 538,23	17 331,65	10 206,58	yes	no	17 331,65	-
35	DU035A	Jutland and Funen, Zealand	3 965,58	725,45	3 240,13	yes	yes	725,45	3 240,13
37	DU037A	Finnmark, Kemijoki, Teno-, Nootom- and Paatsjoki (Finnish part)	17 137,80	9 948,95	7 188,85	yes	no	9 948,95	-
38	DU038A	Oulujoki-Iijoki, Vuoksi	27 767,52	20 701,62	7 065,90	yes	yes	20 701,62	7 065,90
39	DU039A	Daugava, East Estonia, Gauja, West Estonia	10 152,22	6 566,04	3 586,18	yes	yes	6 566,04	3 586,18
40	DU040A	Andalusia Atlantic Basins, Andalusia Mediterranean Basins	13 589,11	8 901,54	4 687,57	yes	yes	8 901,54	4 687,57
41	DU041A	Adour, Garonne, Dordogne, Charente and coastal waters of Aquitania	15 198,04	10 253,98	4 944,06	yes	yes	10 253,98	4 944,06
42	DU042A	Iceland	7 527,65	3 366,38	4 161,27	yes	no	3 366,38	-
43	DU043A	Rhone and Coastal Mediterranean	21 910,16	16 186,62	5 723,54	yes	yes	16 186,62	5 723,54
Total SC04 60% Delivery [km²]			485 779,66	322 679,12	163 100,54			322 679,12	81 968,15
Total already sampled SC01-02-03 [km²]			404 647,27						
Total to be sampled SC04[km²]			81 132,39						

As shown in the Table 4, there was a of 21,000 sample units for the 28 selected DU, covering 485,780 km<sup>2</sup>. The level of stratification is based on the MAES Level 4 classes. The overall share of sample units per main strata and per Specific Contract are presented in Table 4. In the frame of the Fourth Specific Contract, 3,507 additional sample units were sampled since not covered previously.

In order to ensure that unequal sampling intensities are accounted for in the construction of the error matrix, weights are applied to each stratum and Specific Contract (Table 5) following the formula described in section 2.4.4.

The sample units were provided to the interpretation team as one shapefile along with the polygons from the original dataset in which all the information on strata or thematic classes was removed to ensure the independence of the interpretation (section 2.4.3).



Table 4: RZ LCLU distribution of sample units for each stratum et per Specific Contract

		MAES Level 4 New Riparian Zones	Number of Sample Units SC01-02	Number of Sample Units SC03	Number of Sample Units SC04	TOTAL
Strata	1110	Urban Fabric	954	89	64	1,107
	1120	Industrial, Commercial and Military Units	406	36	27	469
	1210	Road Networks and Associated Land	296	24	25	345
	1220	Railways and associated land	123	16	12	151
	1230	Port Areas and Associated Land	6	3	11	20
	1240	Airports and Associated Land	40	6	11	57
	1310	Mineral Extraction, Dump and Construction Sites	152	33	18	203
	1320	Land without Current Use	18	5	11	34
	1400	Green Urban, Sports and Leisure Facilities	269	33	14	316
	2110	Arable Irrigated and Non-Irrigated Land	1,416	268	385	2,069
	2120	Greenhouses	55	18	12	85
	2210	Vineyards, Fruit Trees and Berry Plantations	383	58	38	479
	2220	Olive Groves	139	63	17	219
	2310	Annual Crops Associated with Permanent Crops	18	8		26
	2320	Complex Cultivation Patterns	171	26	24	221
	2330	Land Principally occupied by Agriculture with Significant Areas of Natural Vegetation	147	23	16	186
	2340	Agro-Forestry	75	12		87
	3110	Natural & Semi-Natural Broadleaved Forest	1,671	254	329	2,254
	3120	Highly Artificial Broadleaved Plantations	37		11	48
	3210	Natural & Semi-Natural Coniferous Forest	1,321	194	614	2,129
	3220	Highly Artificial Coniferous Plantations	29	9	1	39
	3310	Natural & Semi-Natural Mixed Forest	1,043	89	114	1,246
	3320	Highly artificial mixed plantations	6			6
	3410	Transitional Woodland and Scrub	749	84	167	1,000
	3420	Lines of Trees and Scrub	43	15	11	69
	3500	Damaged Forest			11	11
	4100	Managed Grassland	786	143	157	1,086
	4210	Semi-Natural Grassland	892	128	182	1,202
	4220	Alpine and Sub-Alpine Natural Grassland	20	26	14	60
	5110	Heathland and Moorland	207	47	154	408
	5120	Other Scrub Land	64		11	75
	5200	Sclerophyllous Vegetation	186	79	38	303
	6100	Sparsely Vegetated Areas	296	24	97	416
	6210	Beaches and Dunes	68	31	12	111
	6220	River Banks	221	26	29	276
	6310	Bare Rocks and Rock Debris	206	27	39	272
	6320	Burnt Areas (except Burnt Forest)	3	1	11	15
	6330	Glaciers and Perpetual Snow	1	3	10	14
	7100	Inland Marshes	378	35	42	455
	7210	Exploited Peat Bog	33	10	13	56
	7220	Unexploited Peat Bog	308	53	223	584
	8110	Coastal salt marshes	104	22	12	138
	8120	Salines	45	11	11	67
	8130	Intertidal Flats	15	9	11	35
	8210	Coastal Lagoons	53	12	15	80
	8220	Estuaries	10	7	13	30
	9110	Interconnected Water Courses	633	38	40	711
	9120	Highly Modified Natural Water Courses and Canals	84	17	13	114
	9130	Separated Water Bodies belonging to the River System	38	3	10	51
	9210	Natural Water Bodies	870	120	366	1,356
	9220	Artificial Standing Water Bodies	44	19	10	73
	9230	Intensively Managed Fishponds	13	17	11	41
	9240	Standing Water Bodies of Extractive Industrial Sites	17	4	10	31
	10000	Sea and Ocean	28	25	10	63
		<b>TOTAL</b>	<b>15,190</b>	<b>2,303</b>	<b>3,507</b>	<b>21,000</b>



Table 5: RZ LCLU distribution of weight factors for each stratum and per Specific Contract

		MAES Level 4 New Riparian Zones	Weight of the strata SC01-02	Weight of the strata SC03	Weight of the strata SC04
Strata	1110	Urban Fabric	0.434753821	1.969292682	0.938491416
	1120	Industrial, Commercial and Military Units	0.365068863	1.316057549	0.693184877
	1210	Road Networks and Associated Land	0.26743531	1.785576373	0.665005372
	1220	Railways and associated land	0.14582727	0.312372322	0.118485687
	1230	Port Areas and Associated Land	0.20352036	0.17745375	0.010757649
	1240	Airports and Associated Land	0.170573332	0.092400722	0.038657079
	1310	Mineral Extraction, Dump and Construction Sites	0.2143412	0.253199279	0.453058581
	1320	Land without Current Use	0.335486168	0.19913638	0.028976276
	1400	Green Urban, Sports and Leisure Facilities	0.180704175	0.433047646	0.309789982
	2110	Arable Irrigated and Non-Irrigated Land	2.276246701	3.814659186	1.084635071
	2120	Greenhouses	0.191981664	0.145452	0.153914528
	2210	Vineyards, Fruit Trees and Berry Plantations	0.49890956	1.55593762	0.80154266
	2220	Olive Groves	0.529081326	1.42295726	0.401919176
	2310	Annual Crops Associated with Permanent Crops	0.210810927	0.037951534	
	2320	Complex Cultivation Patterns	0.710611277	1.278658781	0.641276865
	2330	Land Principally occupied by Agriculture with Significant Areas of Natural Vegetation	0.362897906	1.228698446	0.362977459
	2340	Agro-Forestry	0.287373128	5.358825986	
	3110	Natural & Semi-Natural Broadleaved Forest	0.63356202	3.857945614	1.077750263
	3120	Highly Artificial Broadleaved Plantations	0.164315147		0.024234301
	3210	Natural & Semi-Natural Coniferous Forest	0.801584338	3.016718756	1.096557992
	3220	Highly Artificial Coniferous Plantations	0.285245948	0.019455216	0.002968622
	3310	Natural & Semi-Natural Mixed Forest	0.486278343	2.497976191	1.01482826
	3320	Highly artificial mixed plantations	0.054929272		
	3410	Transitional Woodland and Scrub	0.629872541	1.931443455	1.045475245
	3420	Lines of Trees and Scrub	0.11704583	0.029058052	0.008644239
	3500	Damaged Forest			0.031034049
	4100	Managed Grassland	1.18348577	2.809412601	1.037077538
	4210	Semi-Natural Grassland	0.579782794	3.109436155	1.048022251
	4220	Alpine and Sub-Alpine Natural Grassland	0.177415874	1.107017067	0.256273108
	5110	Heathland and Moorland	0.508620134	2.043086734	1.041638212
	5120	Other Scrub Land	0.268252985		0.006705587
	5200	Sclerophyllous Vegetation	0.687507098	2.244514424	0.810169508
	6100	Sparsely Vegetated Areas	0.3130241	1.100109275	0.999357274
	6210	Beaches and Dunes	0.0945863	0.037565814	0.130251311
	6220	River Banks	0.274691886	0.536600305	0.709985376
	6310	Bare Rocks and Rock Debris	0.247950751	0.493589455	0.806503989
	6320	Burnt Areas (except Burnt Forest)	0.190156982	0.139328921	0.004773862
	6330	Glaciers and Perpetual Snow	0.266933417	0.014418117	0.014156465
	7100	Inland Marshes	0.434574139	1.322525567	0.832476532
	7210	Exploited Peat Bog	0.278921349	0.050283184	0.205062592
	7220	Unexploited Peat Bog	0.655015681	1.179387339	1.062162763
	8110	Coastal salt marshes	0.464741955	0.171410492	0.191168468
	8120	Salines	0.239891284	0.169387671	0.138742871
	8130	Intertidal Flats	0.08487185	0.042099054	0.124369082
	8210	Coastal Lagoons	0.808802365	0.150974323	0.407988071
	8220	Estuaries	0.179541777	0.101184287	0.274007223
	9110	Interconnected Water Courses	0.581396884	0.420274253	0.825687039
	9120	Highly Modified Natural Water Courses and Canals	0.272322043	0.255978666	0.251791569
	9130	Separated Water Bodies belonging to the River System	0.1470933	0.017745375	0.003864334
	9210	Natural Water Bodies	2.5035468	1.997488037	1.085157684
	9220	Artificial Standing Water Bodies	0.11347213	0.046100033	0.014449123
	9230	Intensively Managed Fish Ponds	0.358645397	0.060771385	0.053738984
	9240	Standing Water Bodies of Extractive Industrial Sites	0.147208616	0.08432519	0.019144293
	10000	Sea and Ocean	0.104141528	0.120796095	0.037633256

### 2.4.3. Response Design

LUCAS points are re-interpreted based on available in situ data. LUCAS thematic information is not used directly.

Response design for most data set are based on the interpretation of thematic class at the point level taking into account product specifications (MMU, MMW, class definitions ...) based on combination of available in situ data. Virtual globes and imagery used in production are the main source of in situ reference data.

A blind approach guarantees complete independence from the map products and may underestimate their accuracy for complicated and difficult classes which cannot be unambiguously detected due to differences in the input data between validation and production or due to class definitions when sometimes several LCLU classes are possible. This is resolved by the plausibility approach for which the interpreter checks the map value to assess whether it can be considered plausible, within the frame of accepted product specifications. However, the plausibility analysis should be combined with the blind approach to ensure full traceability and transparency of the validation process.

A blind approach was first applied. This consists in constructing the validation data set without any knowledge about the corresponding map layer information, i.e. the validation team did not have knowledge of the product's thematic classes for the selected sample units. However, the product polygon was provided to the validation team together with the point sample unit to consider boundary effects and geometric differences between the validation and production data.

Used reference data which covers the DU area at a wall-to-wall basis:

- VHR Image 2012 and 2015 NIR with spatial resolution of 2.5m (imagery used for the production and provided by GAF).
- VHR Image 2018 NIR with spatial resolution of 2-4m (imagery used for the production and provided by GAF).

As complementary data:

- Google Earth Images
- LUCAS field survey photos
- National and regional web map services (RGB and/or CIR imagery with varying spatial resolution)

For the initial blind interpretation, operators interpreting each sample unit by selecting code from LCLU Nomenclature. Operators can complete Uncertain (Yes or No) and Comments fields if problems were found (no data, clouds, doubt about 2 or 3 codes, etc).

The LCLU interpretations were done according to the latest MAES nomenclature (LC/LU PRODUCT SPECIFICATIONS, Issue 5.0, 20/03/2020). Some of the classes were regrouped to account for the fact that some of them were not be interpreted by the validation team for the following reason:

- Use of HRL IMD as part of the class definitions

The regrouping of classes was more prevalent for the first specific contract of the validation project, but still applied during the fourth specific contract in relation to classes defined according HRL IMD thresholds. Class regrouping of producer and validation codes is shown in the table below. In addition, sample units extracted from UA and CLC were not considered in these results.

**Table 6: Class grouping for the LCLU accuracy assessment with the thematic rules and the old/new codes**

THEMATIC RULES	RECODING RULES		
	PRODUCER CODE	QC CODE	NEW CODE
MERGING DENSITY IMD	1111	1111	1110
	1112	1112	
	1113	1113	

A second interpretation, carried out by a second interpreter as part of the plausibility analysis, was done with only the interpreted sample units in disagreement with product codes. The operator gives a QC code and corrects validation code if necessary:

- 1: Uncertain code, both producer and operator codes are plausible. Final validation code used is producer code
- 2: Error from first validation interpretation. Final validation used is producer code
- 3: Error from producer. Final validation code used is from first validation interpretation
- 4: Producer and operator are both wrong. Final Validation code used is a new code from this second interpretation.

The Plausibility Analysis results in revised accuracy metrics based on the results from this second interpretation.

#### 2.4.4. Estimation and analyses procedures

Thematic accuracy should be presented in the form of an error matrix. Unequal sampling intensity resulting from the stratified systematic sampling approach should be accounted for by applying a weight factor ( $p$ ) to each sample unit based on the ration between the number of samples and the size of the stratum considered:

$$\hat{p}_{ij} = \left(\frac{1}{N}\right) \sum_{x \in (i,j)} \frac{1}{\pi_{uh}^*}$$

Where  $i$  and  $j$  are the columns and rows in the matrix,  $N$  is the total number of possible units (population) and  $\pi$  is the sampling intensity for a given stratum.

Overall accuracy and User and producer accuracy should be computed for all thematic classes and 95% confidence intervals should be calculated for each accuracy.

The standard error of the error rate can be calculated as follows:  $\sigma_h = \sqrt{\frac{p_h(1-p_h)}{n_h}}$  where  $n_h$  is the sample size for stratum  $h$  and  $p_h$  is the expected error rate. The standard error is calculated for each stratum and an overall standard error is calculated based on the following formula:

$$\sigma = \sqrt{\sum w_h^2 \cdot \sigma_h^2}$$

In which is the proportion of the total area covered by each stratum. The 95% confidence interval is +/- 1.96.

### 2.5. Temporal Quality

Temporal quality is evaluated by providing an indication of the closeness of the acquired image data to the reference year, e.g. the percentage area covered outside the accepted reference period as defined in the tender/product specification i.e. 2018 +/- 1-2 year(s).

### 2.6. Usability

Usability relates to the appropriateness of the metadata description and accompanying documentation to describe the processes and workflows involved in the production of the data. Although it is difficult to describe usability in quantitative terms, it provides a clear evaluation based on objective criteria of any limitation in the intended use of the data.

### 2.7. INSPIRE compliant metadata

Presence of INSPIRE compliant metadata should be verified.

### 3. Validation check list

The final validation exercise for LCLU has been performed over 6 DU by ETC/ULS and 6 Look & Feel Checks reports have been analyzed for DU001B, 004B, 005B, 0027B, 035B and 040B. 4 DUs has been declared "good" (meaning that operator is confident that accuracy of the product is at least 85 %; only sporadic errors are encountered in the verified areas) and 2 DUs has been declared "acceptable" (meaning that accuracy of the product is estimated to reach 85 % in most of the verified areas, minor errors can be detected in the verified areas).

PRODUCT: RZ LCLU (6 DU)						
VALIDATION LEVEL: DU						
SERVICE PROVIDER: GAF, Geoville, Indra and E-Geos				SERVICE USER: EEA	ISSUE/REVISION: 1.0	
VALIDATION DATE: 11/10/2019				REVIEW DATE: 20/07/2020		
CONDUCTED BY: Gisat				REVIEWED BY: Alice LHERNOULD	APPROVED BY:	
No.	DATA QUALITY SUB-	DATA QUALITY MEASURE	DATA QUALITY RESULT	COMMENTS BY AUDIT TEAM	DRAFT AUDIT CONCLUSION	FINAL AUDIT CONCLUSION
1	COMPLETENESS					
1.1	Commission	Rate of excess items		No information provided by QC reports.	-	
1.2	Omission	Rate of missing items		No information provided by QC reports.	-	
2	LOGICAL CONSISTENCY					
2.1	Format consistency	File format	100%	All File formats conforming to ESRI shapefile	Accepted	
2.2		File name	100%	All File Names conforming	Accepted	

PRODUCT: RZ LCLU (6 DU)						
VALIDATION LEVEL: DU						
SERVICE PROVIDER: GAF, Geoville, Indra and E-Geos				SERVICE USER: EEA	ISSUE/REVISION: 1.0	
VALIDATION DATE: 11/10/2019				REVIEW DATE: 20/07/2020		
CONDUCTED BY: Gisat				REVIEWED BY: Alice LHERNOULD	APPROVED BY:	
No.	DATA QUALITY SUB-	DATA QUALITY MEASURE	DATA QUALITY RESULT	COMMENTS BY AUDIT TEAM	DRAFT AUDIT CONCLUSION	FINAL AUDIT CONCLUSION
2.3		Attributes names	100%	All attribute names in file conforming	Accepted	
2.4		Attributes types	100%	All attribute types conforming according	Accepted	
2.5	Conceptual consistency	Feature type		No information provided by QC reports	-	

PRODUCT: RZ LCLU (6 DU)						
VALIDATION LEVEL: DU						
SERVICE PROVIDER: GAF, Geoville, Indra and E-Geos				SERVICE USER: EEA	ISSUE/REVISION: 1.0	
VALIDATION DATE: 11/10/2019				REVIEW DATE: 20/07/2020		
CONDUCTED BY: Gisat				REVIEWED BY: Alice LHERNOULD	APPROVED BY:	
No.	DATA QUALITY SUB-	DATA QUALITY MEASURE	DATA QUALITY RESULT	COMMENTS BY AUDIT TEAM	DRAFT AUDIT CONCLUSION	FINAL AUDIT CONCLUSION
2.6		MMU		N/A	-	
		MMW		N/A	-	
2.7		Coordinate reference system	100%	File format conforms to specifications. File format consistency is checked by EEA QC tool during product upload. QC reports are not open access.	Accepted	
2.8		Unique identifier		OK	Accepted	
2.9		Nomenclature	100%	Nomenclature conforms to specifications. Nomenclature consistency is checked by EEA QC tool during product upload. QC reports are not open access.	Accepted	

PRODUCT: RZ LCLU (6 DU)						
VALIDATION LEVEL: DU						
SERVICE PROVIDER: GAF, Geoville, Indra and E-Geos				SERVICE USER: EEA	ISSUE/REVISION: 1.0	
VALIDATION DATE: 11/10/2019				REVIEW DATE: 20/07/2020		
CONDUCTED BY: Gisat				REVIEWED BY: Alice LHERNOULD	APPROVED BY:	
No.	DATA QUALITY SUB-	DATA QUALITY MEASURE	DATA QUALITY RESULT	COMMENTS BY AUDIT TEAM	DRAFT AUDIT CONCLUSION	FINAL AUDIT CONCLUSION
2.10	Domaine consistency	Value domain non-conformance	100%	Domaine consistency conforms to specifications	Accepted	
2.11	Topological consistency	Overlaps		N/A	-	
2.12		Gaps		N/A	-	

PRODUCT: RZ LCLU (6 DU)						
VALIDATION LEVEL: DU						
SERVICE PROVIDER: GAF, Geoville, Indra and E-Geos				SERVICE USER: EEA	ISSUE/REVISION: 1.0	
VALIDATION DATE: 11/10/2019				REVIEW DATE: 20/07/2020		
CONDUCTED BY: Gisat				REVIEWED BY: Alice LHERNOULD	APPROVED BY:	
No.	DATA QUALITY SUB-	DATA QUALITY MEASURE	DATA QUALITY RESULT	COMMENTS BY AUDIT TEAM	DRAFT AUDIT CONCLUSION	FINAL AUDIT CONCLUSION
2.13		Multipart features		N/A	-	
2.14		Neighboring features		N/A	-	
2.15		Self-intersections		N/A	-	
2.16		Null geometry		N/A	-	
2.17		Unclosed rings		N/A	-	



PRODUCT: RZ LCLU (6 DU)						
VALIDATION LEVEL: DU						
SERVICE PROVIDER: GAF, Geoville, Indra and E-Geos				SERVICE USER: EEA	ISSUE/REVISION: 1.0	
VALIDATION DATE: 11/10/2019				REVIEW DATE: 20/07/2020		
CONDUCTED BY: Gisat				REVIEWED BY: Alice LHERNOULD	APPROVED BY:	
No.	DATA QUALITY SUB-	DATA QUALITY MEASURE	DATA QUALITY RESULT	COMMENTS BY AUDIT TEAM	DRAFT AUDIT CONCLUSION	FINAL AUDIT CONCLUSION
2.18		Duplicate vertex		N/A	-	
2.19		Pseudo nodes		N/A	-	
2.20		Non-matching nodes		N/A	-	
3	POSITIONAL ACCURACY					
3.1		Positional accuracy of polygon	100%	Correct or shifted positional accuracy of polygons	Accepted	
4	THEMATIC ACCURACY					
4.1	Classification correctness	Overall accuracy	From 85% (DU 005B) to 97.7% (DU 004B)	Semi-quantitative assessment based on 180 samples on MAES 4 level LCLU class code (plausibility approach)	Accepted	

PRODUCT: RZ LCLU (6 DU)						
VALIDATION LEVEL: DU						
SERVICE PROVIDER: GAF, Geoville, Indra and E-Geos				SERVICE USER: EEA	ISSUE/REVISION: 1.0	
VALIDATION DATE: 11/10/2019				REVIEW DATE: 20/07/2020		
CONDUCTED BY: Gisat				REVIEWED BY: Alice LHERNOULD	APPROVED BY:	
No.	DATA QUALITY SUB-	DATA QUALITY MEASURE	DATA QUALITY RESULT	COMMENTS BY AUDIT TEAM	DRAFT AUDIT CONCLUSION	FINAL AUDIT CONCLUSION
4.2		Correctness of object delineation (i.e. omissions / commissions)	From 77.4% (DU 040B) to 98.8% (DU001B)		Conditionally accepted	
4.3		Polygon delineation detail	From 66.6% (DU 005B) to 100% (DU004B)		Conditionally accepted	
5	TEMPORAL QUALITY					
5.1	Temporal quality	Closeness of the acquired image data to the	N/A	Dependant on the VHR 2018 data set accuracy. No accuracy report available for VHR 2018 mosaic.	-	
6	USABILITY					
6.1	Usability					
7	METADATA					
7.1	INSPIRE compliant metadata	File format		N/A	Accepted	

PRODUCT: RZ LCLU (6 DU)						
VALIDATION LEVEL: DU						
SERVICE PROVIDER: GAF, Geoville, Indra and E-Geos				SERVICE USER: EEA	ISSUE/REVISION: 1.0	
VALIDATION DATE: 11/10/2019				REVIEW DATE: 20/07/2020		
CONDUCTED BY: Gisat				REVIEWED BY: Alice LHERNOULD	APPROVED BY:	
No.	DATA QUALITY SUB-	DATA QUALITY MEASURE	DATA QUALITY RESULT	COMMENTS BY AUDIT TEAM	DRAFT AUDIT CONCLUSION	FINAL AUDIT CONCLUSION
7.2		File content		N/A	Accepted	
7.3		INSPIRE compliant completeness		INSPIRE compliant	Accepted	

## 4. Additional Tables

### 4.1. Attributes

For information's purpose, Table 7 below recaps the attributes of the LCLU product.

**Table 7: RZ LCLU attribute names and types**

ATTRIBUTE NAME	DESCRIPTION	ATTRIBUTE TYPE
OBJECTID	Dynamic Feature Identifier	Object ID
Shape	Polygon (default by ESRI)	Geometry
ID	Unique Identifier of the feature geometry	Long, Precision 6
DU_ID	Unique Identifier of the Riparian Zones delivery unit	String, Length 10
MAES_1_18	Class code of MAES Level 1 for LC/LU 2018	Long, Precision 9
MAES_2_18	Class code of MAES Level 2 for LC/LU 2018	Long, Precision 9
MAES_3_18	Class code of MAES Level 3 for LC/LU 2018	Long, Precision 9
MAES_4_18	Class code of MAES Level 4 for LC/LU 2018	Long, Precision 9
UA_18	Urban Atlas 2018 flag	String, Length 10
NODATA_18	Unclassifiable area due to clouds, shadows, snow, haze or missing data (no satellite data available)	Long, Precision 9
COMMENT_18	Comment field for additional information for LC/LU 2018	String, Length 254
MAES_1_12	Class code of MAES Level 1 for LC/LU 2012	Long, Precision 9
MAES_2_12	Class code of MAES Level 2 for LC/LU 2012	Long, Precision 9
MAES_3_12	Class code of MAES Level 3 for LC/LU 2012	Long, Precision 9
MAES_4_12	Class code of MAES Level 4 for LC/LU 2012	Long, Precision 9
UA_12	Urban Atlas 2012 flag	String, Length 10
NODATA_12	Unclassifiable area due to clouds, shadows, snow, haze or missing data (no satellite data available)	Long, Precision 9
COMMENT_12	Comment field for additional information for LC/LU 2012	String, Length 254
CHANGECODE	Change Class code of MAES Level 4 for LC/LU 2018 & 2012	String, Length 11
AREA_HA	Area in hectare [ha]	Double

## 4.2. Coordinate reference system description

Table 8 lists the projection parameters of the RZ product.

**Table 8: Description of the RZ coordinate reference system**

<b>NAME OF COORDINATE REFERENCE SYSTEM</b>	ETRS_1989_LAEA
WKID	3035
PROJECTION	Lambert_Azimuthal_Equal_Area
FALSE EASTING	4321000.0
FALSE NORTHING	3210000.0
CENTRAL MERIDIAN	10.0
LATITUDE OF ORIGIN	52.0
LINEAR UNIT	Meter
<b>GEOGRAPHIC COORDINATE SYSTEM</b>	GCS_ETRS_1989
ANGULAR UNIT	Degree (0,0174532925199433)
PRIME MERIDIAN	Greenwich (0,0)
DATUM	D_ETRS_1989
SPHEROID	GRS_1980
SEMIMAJOR AXIS	6378137,0
SEMIMINOR AXIS	6356752,314140356
INVERSE FLATTENING	298,257222101

### 4.3. Nomenclature

For information purpose, Table 9 below describes the latest LCLU nomenclature used in the semantic checks (LC/LU PRODUCT SPECIFICATIONS, Issue 5.0, 20/03/2020).

Table 9: RZ LCLU Nomenclature v5.0 (20/03/2020)

CODE	CLASS NAME
1111	Continuous Urban Fabric (IM.D $\geq$ 80%)
1112	Dense Urban Fabric (IM.D $\geq$ 30-80%)
1113	Low Density Fabric (IM.D < 30%)
1120	Industrial, commercial and military units
1210	Road networks and associated land
1220	Railways and associated land
1230	Port areas and associated land
1240	Airports and associated land
1310	Mineral extraction, dump and construction sites
1320	Land without current use
1400	Green urban, sports and leisure facilities
2110	Arable land
2120	Greenhouses
2210	Vineyards, fruit trees and berry plantations
2220	Olive groves
2310	Annual crops associated with permanent crops
2320	Complex cultivation patterns
2330	Land principally occupied by agriculture with significant areas of natural vegetation
2340	Agro-forestry
3110	Natural and semi-natural Broadleaved forest
3120	Highly artificial broadleaved plantations
3210	Natural and semi-natural coniferous forest
3220	Highly artificial coniferous plantations
3310	Natural and semi-natural mixed forest
3320	Highly artificial mixed plantations
3410	Transitional woodland and scrub
3420	Lines of trees and scrub
3500	Damaged forest
4100	Managed grasslands
4210	Semi-natural grassland
4220	Alpine and sub-alpine natural grassland
5110	Heathland and Moorland
5120	Other scrub land
5200	Sclerophyllous vegetation
6100	Sparsely vegetated areas
6210	Beaches and dunes
6220	River banks
6310	Bare rocks and rock debris
6320	Burnt areas (except burnt forest)
6330	Glaciers and perpetual snow
7100	Inland marshes
7210	Exploited peat bog
7220	Unexploited peat bog
8111	Coastal salt marshes

CODE	CLASS NAME
8120	Salines
8130	Intertidal flats
8210	Coastal lagoons
8220	Estuaries
9110	interconnected water courses
9120	Highly modified water courses and canals
9130	Separated water bodies belonging to the river system
9210	Natural water bodies
9220	Artificial standing water bodies
9230	Intensively managed fish ponds
9240	Standing water bodies of extractive industrial sites
10000	Sea and ocean

## 5. Thematic accuracy

Table 10 summarizes overall accuracies obtained at pan-European level, by Delivery Units (DU) and by biogeographical region, for the blind and plausibility interpretations. Plausibility analysis results for 2018 are almost all greater than the 85% threshold (91.51% at pan-European level). Therefore, the RZ LCLU 2018 results meet and even exceed the overall accuracy requirements. The Change layer (2012-2018) accuracy is also in line with an overall result for EEA39 at 86.15%. It should be noticed that the results obtained for the blind interpretation for the reference year 2018 are much greater than the Overall Accuracy obtained in the previous Specific Contracts for the reference year 2012 (Overall Accuracy of 68.8% for the blind interpretation). The differences can be explained because PSUs from 2012 were re-interpreted for the reference year 2018 for which a plausibility had already been applied.

At a disaggregated level, almost all the plausibility results by biogeographical regions are greater than the 85% minimum requirement, the weakest result is obtained for the Boreal region (77.35%). Overall accuracies by DU are also satisfactory with a large majority of them at a level greater or equal to 85%. A few DUs below 85% are located in the Boreal region (DU042A and DU37A) and characterized with very complex landscape (e.g. mosaic landscape of forests and wetlands). The results on the Change layer between 2012 and 2018 also show a discrepancy on Boreal regions with an overall accuracy of 77.16%.

The complete confusion matrices (blind and plausibility interpretation) for all the DUs are available in RZ LCLU confusion matrix for the blind and plausibility interpretation for the 28 delivered DUs (source Excel files are provided for more detailed analysis) providing more insight on thematic classes that are more difficult to characterize. Some of the MAES level 4 thematic classes exhibit user and producer accuracies that are substantially lower than the overall accuracy threshold of 85 %, but these classes tend to cover small areas or suffer from confusion with similar thematic classes. Overall, blind interpretation results are substantially lower, suggesting the following comments:

- The use of the imagery alone is not sufficient to characterize some of the RZ LCLU classes due to the complexity of the nomenclature and the reliance on ancillary data to resolve some ambiguity especially for the forestry (3xxx) and grassland (4xxx) thematic classes.
- Thematic classes including notions densities such as field or tree coverings have been problematic in blind interpretation and increase confusion between classes. This is particularly relevant for agricultural, forest and grassland thematic classes because vegetation density rules are equivalent.

It should also be noted that the stratification and sampling strategy provided a high level of precision since the 95% Confidence intervals were around 0.2% at EEA39 level and around 0.1% at DU level.



Table 10: Thematic Overall Accuracy (OA) for LCLU 2012, 2018 products and Change layer product: Blind and Plausibility interpretation, per delivery unit, per lot and bio-geographical region. For LCLU products: n is the number of observations, green color corresponds to OA greater than 85%, brown within 10% and red below 10% considering 95% CI. For Change product: n is the number of observations, green color corresponds to OA greater than 80%, brown within 10% and red below 10% considering 95% CI.

		Blind								Plausibility					
		2012		2018		2012-2018				2012		2018		2012-2018	
		n	Overall	95% CI	Overall	95% CI	Overall			95% CI	Overall	95% CI	Overall	95% CI	
EEA39	21,000	68.82%	0.40%	88.75%	0.26%	67.41%	0.40%	EEA39	21,000	87.86%	0.23%	91.51%	0.22%	86.15%	0.26%
ALP	1,940	64.28%	0.13%	83.03%	0.09%	62.92%	0.13%	ALP	1,940	84.80%	0.09%	88.11%	0.08%	83.25%	0.09%
ATL	808	76.92%	0.05%	86.23%	0.03%	75.28%	0.05%	ATL	808	87.19%	0.03%	88.62%	0.03%	85.52%	0.04%
BLS	863	71.34%	0.17%	73.05%	0.13%	64.81%	0.17%	BLS	863	84.72%	0.10%	77.35%	0.11%	77.16%	0.11%
BOR	2,901	76.50%	0.02%	92.31%	0.01%	75.35%	0.02%	BOR	2,901	91.35%	0.01%	93.92%	0.01%	90.11%	0.01%
CON	383	64.85%	0.24%	79.44%	0.15%	62.86%	0.24%	CON	383	85.40%	0.13%	84.05%	0.13%	81.28%	0.15%
MAC	5,848	72.77%	0.13%	89.53%	0.07%	71.25%	0.14%	MAC	5,848	89.89%	0.08%	92.55%	0.06%	88.60%	0.08%
MED	2,454	71.09%	0.04%	89.83%	0.02%	69.91%	0.04%	MED	2,454	89.41%	0.02%	92.38%	0.02%	87.78%	0.02%
PAN	5,424	57.62%	0.18%	87.53%	0.09%	56.29%	0.18%	PAN	5,424	82.97%	0.12%	90.19%	0.08%	80.82%	0.12%
STE	379	78.19%	0.05%	97.07%	0.02%	77.39%	0.05%	STE	379	94.10%	0.01%	97.48%	0.01%	92.89%	0.02%
DU001A	860	63.73%	0.08%	82.17%	0.05%	62.72%	0.08%	DU001A	860	77.17%	0.07%	85.07%	0.05%	75.51%	0.07%
DU002A	757	59.35%	0.07%	90.16%	0.03%	58.75%	0.07%	DU002A	757	86.83%	0.04%	92.22%	0.03%	84.96%	0.04%
DU003A	831	69.93%	0.05%	82.21%	0.03%	68.44%	0.05%	DU003A	831	77.30%	0.04%	84.97%	0.03%	75.62%	0.05%
DU004A	810	56.22%	0.08%	77.11%	0.06%	55.58%	0.08%	DU004A	810	70.70%	0.07%	82.70%	0.05%	70.23%	0.07%
DU005A	910	61.79%	0.13%	90.74%	0.04%	61.11%	0.13%	DU005A	910	89.03%	0.05%	93.71%	0.03%	87.33%	0.05%
DU006A	651	46.56%	0.04%	88.54%	0.01%	45.13%	0.04%	DU006A	651	84.51%	0.02%	93.38%	0.01%	83.78%	0.02%
DU007A	800	65.28%	0.06%	86.45%	0.04%	64.78%	0.06%	DU007A	800	87.89%	0.03%	90.28%	0.03%	86.16%	0.04%
DU010A	872	74.02%	0.06%	84.14%	0.05%	71.97%	0.06%	DU010A	872	89.80%	0.04%	87.45%	0.04%	86.62%	0.04%
DU013A	660	76.06%	0.08%	95.44%	0.03%	75.23%	0.08%	DU013A	660	92.43%	0.04%	96.60%	0.02%	91.07%	0.04%
DU019A	759	55.00%	0.10%	90.07%	0.05%	54.35%	0.10%	DU019A	759	79.67%	0.07%	92.15%	0.04%	78.09%	0.07%
DU022A	640	77.44%	0.04%	83.79%	0.04%	76.62%	0.05%	DU022A	640	86.46%	0.03%	88.59%	0.03%	85.72%	0.03%
DU023A	515	82.28%	0.07%	95.92%	0.02%	81.17%	0.07%	DU023A	515	93.15%	0.03%	96.83%	0.01%	92.00%	0.03%
DU024A	768	67.23%	0.07%	92.99%	0.03%	66.15%	0.08%	DU024A	768	89.87%	0.04%	94.27%	0.03%	88.79%	0.04%
DU025A	587	66.78%	0.06%	94.36%	0.02%	66.66%	0.06%	DU025A	587	89.52%	0.02%	95.30%	0.02%	89.33%	0.02%
DU026A	471	78.72%	0.02%	93.63%	0.00%	75.83%	0.02%	DU026A	471	94.84%	0.00%	94.04%	0.00%	91.69%	0.01%
DU027A	997	81.18%	0.08%	90.53%	0.05%	79.99%	0.08%	DU027A	997	93.66%	0.04%	93.31%	0.04%	92.66%	0.04%
DU028A	661	75.35%	0.04%	89.37%	0.02%	74.17%	0.04%	DU028A	661	92.17%	0.02%	91.75%	0.02%	91.30%	0.02%
DU029A	746	82.15%	0.04%	95.13%	0.01%	80.63%	0.04%	DU029A	746	94.75%	0.02%	95.43%	0.01%	93.15%	0.02%
DU031A	1,038	71.96%	0.11%	83.57%	0.08%	70.35%	0.12%	DU031A	1,038	86.36%	0.08%	89.57%	0.06%	85.95%	0.08%
DU034A	993	60.05%	0.12%	84.96%	0.07%	59.44%	0.12%	DU034A	993	84.28%	0.07%	91.36%	0.05%	83.84%	0.07%
DU035A	376	71.95%	0.05%	87.26%	0.02%	67.64%	0.05%	DU035A	376	91.48%	0.02%	90.00%	0.02%	89.64%	0.02%
DU037A	887	68.10%	0.06%	79.84%	0.05%	67.57%	0.06%	DU037A	887	82.69%	0.04%	84.59%	0.04%	82.19%	0.04%
DU038A	734	66.92%	0.08%	90.88%	0.03%	64.52%	0.08%	DU038A	734	91.37%	0.04%	92.74%	0.02%	88.93%	0.04%
DU039A	681	69.53%	0.04%	85.56%	0.02%	65.07%	0.04%	DU039A	681	87.67%	0.02%	90.91%	0.02%	83.52%	0.03%
DU040A	766	55.40%	0.05%	89.73%	0.03%	52.98%	0.05%	DU040A	766	84.58%	0.03%	90.92%	0.03%	81.43%	0.03%
DU041A	718	80.24%	0.06%	95.99%	0.03%	78.92%	0.06%	DU041A	718	96.48%	0.03%	96.71%	0.03%	94.79%	0.03%
DU042A	851	71.74%	0.03%	73.39%	0.03%	64.98%	0.04%	DU042A	851	84.83%	0.02%	77.19%	0.03%	77.00%	0.03%
DU043A	661	64.20%	0.07%	89.00%	0.04%	60.45%	0.07%	DU043A	661	89.83%	0.04%	89.75%	0.04%	84.89%	0.05%

## 6. Conclusions

### 6.1. Main findings

The LCLU 2018 product shows **high overall thematic accuracy** for most of the 28 DUs, namely higher than the 90% for the plausibility validation accuracy. This dataset also contains a high disparity in accuracy levels, from 77.2% of overall accuracy for DU042A (Iceland) to 96.8% for DU023A (Anglian, Dee, Humber, North West, Thames, UK South East, UK South West, Western Wales, Isle of Man). There is an obvious link between landscape complexity and weaker overall accuracy. At biogeographical level, the Boreal DUs present the lowest accuracy values (77.3%) but most of the other regions are close to or well above 90%.

On the whole, blind accuracy assessment results are lower, but it is important to highlight the complexity of this product's nomenclature for which is at the very edge of what can be extracted from the available imagery. This is further illustrated when scrutinizing confusion matrices at MAES level 4 which exhibit accuracies that can sometimes be lower than the overall accuracy threshold of 85%. This is because some of these classes often cover very small areas or can be confused with similar classes. Nevertheless, it should be noticed that the results obtained for the blind interpretation for the reference year 2018 are much greater than the Overall Accuracy obtained in the previous Specific Contracts for the reference year 2012 (Overall Accuracy of 68.8% for the blind interpretation). The differences can be explained because PSUs from 2012 were re-interpreted for the reference year 2018 for which a plausibility had already been applied.

The **stratification and sampling design** proved efficient ensuring a good representation of all LCLU classes whilst providing precise overall thematic accuracy results at DU and EEA39 level.

The results from the LCLU **Look & feel checks reports** performed by ETC/ULS have revealed that 4 DUs has been declared "good" and 2 "acceptable". Both artificial / urban features and natural features were mapped with the same geometrical precision and detail of mapping features was good.

### 6.2. Recommendations

These results are in line with the previous assessment of the RZ LCLU product performed previously. Therefore, the present validation was implemented based on the full product without any further class grouping. However, the differences between the **blind interpretation and plausibility results** highlights the **complexity of the nomenclature** of this product which is probably beyond what can be extracted from the available imagery for certain MAES level 4 classes and suggests that for a given land unit several interpretations are possible despite the detailed documentation available. In fact, some classes remain ambiguous based on the VHR satellite imagery alone. This can only be resolved if:

- Multi-temporal HR satellite imagery is available (Landsat or Sentinel-2)
- Auxiliary data sources are available (used as auxiliary data).

It appears that **auxiliary data sources** were sometimes used in the production, but this was not always entirely clear to the validation team and these could be detailed in an "Auxiliary Source Identification Layer" similar to the PSIL used for the identification of source VHR imagery. In addition, despite the comprehensive documentation available for the product, its usability could be further enhanced if the complexity of the nomenclature could be summarized in a decision tree. This could be included as part of a user handbook which would summarize the comprehensive documentation available any limitations in the use of the data.

The analysis of the **validation results at class and disaggregated geographical level** should provide insights on where the product could be improved thematically focusing on weaker classes and geographical regions. and cases for which some thematic classes could be regrouped to improve their identification if required.

The validation should be implemented when the production is finalized based on a final set of products and documentation to avoid issues related to product definition and nomenclature which has yet to be finalized. This should also contribute on further improving the efficiency of the **stratification and sampling design** by avoiding

a first stratification at DU level which is only necessary when the dataset is not yet complete. This first level would also make sense if there were large discrepancies in the size of the DU which was not the case.

[illegible]

**Figure 7: RZ LCLU 2012 confusion matrix for the blind interpretation for the 28 delivered DUs**

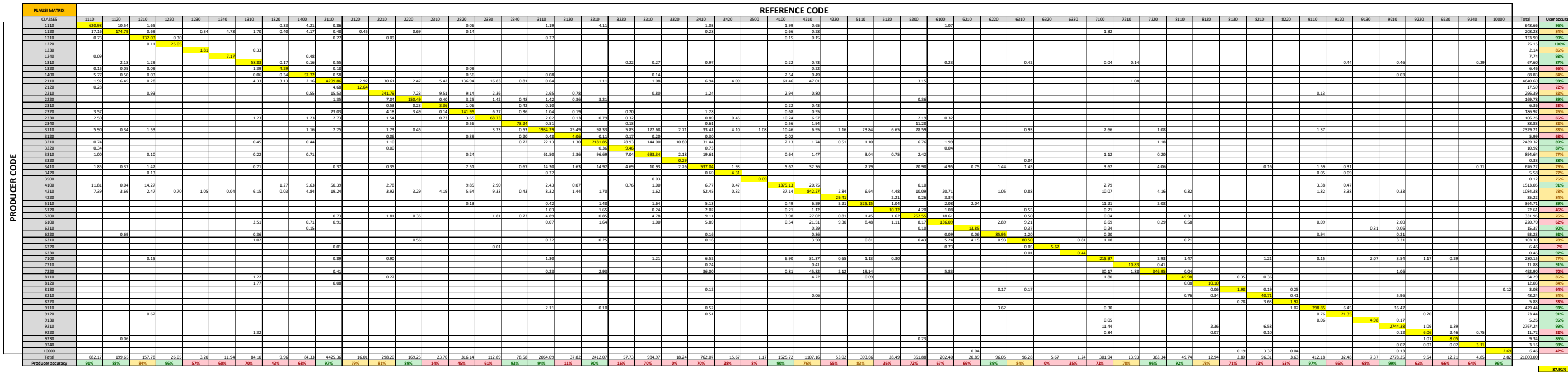


Figure 8: RZ LCLU 2012 confusion matrix for the plausibility interpretation for the 28 delivered DUs



CHANGES - BLIND	Class 10 - Sea and Ocean	Class 1 - Urban	Class 2 - Cropland	Class 3 - Woodland and Forest	Class 4 - Grassland	Class 5 - Heathland and Scrub	Class 6 - Sparsely Vegetated Land	Class 7 - Wetland	Class 8 - Lagoons, Coastal Wetlands and Estuaries	Class 9 - Rivers and Lakes	Total	User accuracy
Class 10 - Sea and Ocean	1.53						0.19		4.46	0.28	6.46	24%
Class 1 - Urban		1120.13	11.61	15.23	13.05		3.54	2.55	0.28	2.46	1168.85	96%
Class 2 - Cropland		32.45	4781.76	73.14	593.97	26.73	2.75	1.08		0.93	5512.83	87%
Class 3 - Woodland and Forest		32.65	40.17	5842.71	154.88	223.39	32.40	25.31	1.64	9.19	6362.33	92%
Class 4 - Grassland		84.03	332.12	131.71	1910.04	58.66	59.15	43.69	1.32	11.92	2632.65	73%
Class 5 - Heathland and Scrub			4.11	63.27	147.17	428.83	49.65	23.36	1.76	1.11	719.26	60%
Class 6 - Sparsely Vegetated Land		6.13	0.93	19.74	51.68	25.03	292.77	13.03	4.68	25.62	439.60	67%
Class 7 - Wetland		0.44	3.57	89.93	150.19	51.41	10.76	459.48	4.15	14.99	784.93	59%
Class 8 - Lagoons, Coastal Wetlands and Estuaries	0.54	1.68	0.47	0.12	6.73	1.09	0.64	17.16	82.24	12.79	123.47	67%
Class 9 - Rivers and Lakes		5.96		32.63		0.23	6.33	21.54	19.52	3163.39	3249.61	97%
Total	2.08	1283.49	5174.73	6268.49	3027.70	815.38	458.18	607.19	120.07	3242.70	21000.00	
Producer accuracy	74%	87%	92%	93%	63%	53%	64%	76%	68%	98%		86.11%

Figure 11: RZ LCLU Change 2012-18 confusion matrix at MAES Level 1 for the blind interpretation for the 28 delivered DUs

CHANGES - PLAUSI	Class 10 - Sea and Ocean	Class 1 - Urban	Class 2 - Cropland	Class 3 - Woodland and Forest	Class 4 - Grassland	Class 5 - Heathland and Scrub	Class 6 - Sparsely Vegetated Land	Class 7 - Wetland	Class 8 - Lagoons, Coastal Wetlands and Estuaries	Class 9 - Rivers and Lakes	Total	User accuracy
Class 10 - Sea and Ocean	2.69						0.04		3.60	0.13	6.46	42%
Class 1 - Urban		1142.79	5.01	8.54	8.07		1.73	1.49		1.22	1168.85	98%
Class 2 - Cropland		28.55	5299.36	33.00	133.40	16.99	0.32	1.08		0.13	5512.83	96%
Class 3 - Woodland and Forest		16.89	14.30	6140.38	64.05	96.94	11.58	13.91	0.16	4.11	6362.33	97%
Class 4 - Grassland		59.35	111.94	77.35	2307.55	23.78	25.97	17.02	0.32	9.37	2632.65	88%
Class 5 - Heathland and Scrub			5.56	33.24	45.44	596.33	24.86	13.53	0.31		719.26	83%
Class 6 - Sparsely Vegetated Land		6.46	1.48	9.50	35.52	19.11	348.24	8.61	0.78	9.92	439.60	79%
Class 7 - Wetland		0.15	2.20	48.44	87.59	20.56	5.83	609.13	2.72	8.30	784.93	78%
Class 8 - Lagoons, Coastal Wetlands and Estuaries	0.12	2.99	0.35	0.12	4.28	0.09	0.34	1.80	107.40	5.96	123.47	87%
Class 9 - Rivers and Lakes		2.00		3.24		0.23	3.62	12.63	10.13	3217.76	3249.61	99%
Total	2.82	1259.18	5440.19	6353.81	2685.90	774.03	422.53	679.20	125.42	3256.90	21000.00	
Producer accuracy	96%	91%	97%	97%	86%	77%	82%	90%	86%	99%		94.15%

Figure 12: RZ LCLU Change 2012-18 confusion matrix at MAES Level 1 for the plausibility interpretation for the 28 delivered DUs