

**Copernicus Land Monitoring Service – CLC+ Backbone Production, including
Raster and Vector Products based on Satellite Input Data from 2017/2018/2019**

**CLC+ BACKBONE
PRODUCT SPECIFICATION
AND USER MANUAL**

Issue 3.0



European Environment Agency

Framework Service Contract No. EEA/DIS/R0/19/012

Issue: 3.0
Date: 30/09/2022

TABLE OF CONTENTS

1	Executive Summary	1
2	Background of the Document	2
2.1	Scope of the Document.....	2
2.2	Content and Structure	2
3	Product Description.....	3
3.1	Overview of the product portfolio	3
3.2	Raster Product.....	4
3.2.1	Product Specifications	4
3.2.2	Production Methodology and Workflow.....	17
3.3	Vector Product.....	26
3.3.1	Product Specifications	26
3.3.2	Production Methodology and Workflow.....	35
3.4	Additional Quality Layers and Expert Products	55
3.4.1	Data Score Layer (DSL).....	55
3.4.2	Raster Confidence Layer (CL) Specifications.....	57
3.4.3	Raster Post-processing Layer.....	59
3.4.4	Segmentation Confidence Layer.....	61
4	Product Quality.....	63
4.1	Raster Product Thematic Accuracy.....	63
4.1.1	Methodological approach	63
4.1.2	Results	67
4.2	Vector Product Thematic Accuracy	69
4.2.1	Methodological approach	69
4.2.2	Results	71
4.3	Vector Product Geometric Accuracy	75
4.3.1	Methodological approach	75
4.3.2	Results	79
5	Terms of Use and Product Technical Support	83
5.1	Terms of Use.....	83
5.2	Citation	83
5.3	Product Technical Support	83
6	References	84
	ANNEXES.....	85
	ANNEX 1: NAMING CONVENTIONS	85
	ANNEX 2: COLOUR PALETTES	86
	ANNEX 3: PROJECTION PARAMETERS	90
	ANNEX 4: THEMATIC ACCURACIES FOR THE RASTER UNITS OF EEA-38 + UK	91
	ANNEX 5: THEMATIC ACCURACIES FOR THE VECTOR UNITS OF EEA-38 + UK.....	99
	ANNEX 6: GEOMETRIC VALIDATION RESULTS FOR THE VECTOR PRODUCT PER BIOGEOGRAPHIC REGION	111

LIST OF FIGURES

Figure 1: High-level overview of the CLC+ Backbone product portfolio.	3
Figure 2: Exemplary 10x10m plot illustration of a situation where no single land cover class reaches 50%.	5
Figure 3: Levels 1-4 of the EAGLE Land Cover Components (LCC) v3.1. The component hierarchy shall be used to resolve the class assignment in cases where no single land cover class dominates.	6
Figure 4: Decision tree for the usage of area thresholds for pixels with (pure or) spatially mixed land cover in the Raster Product.....	10
Figure 5: Illustration of the CLC+ Backbone raster product decision tree approach to evaluate area thresholds for mixed land cover at olive groves, Castile-La Mancha, Spain. Upper left: 100x100m plot overlaid on aerial imagery. Upper right: Outlines of surfaces which are covered by Sealed, Permanent herbaceous, Broadleaved evergreen trees and Non- and sparsely-vegetated. 10m pixel grid overlaid on top. Lower left: Resulting class attribution of the 10m pixel when applying the threshold of Biotic >30% to decide between Non- and sparsely-vegetated vs. one of the vegetation classes. Lower right: ©Google Street View of the area, showing the four dominant land cover classes.....	12
Figure 6: CLC+ Backbone Raster LC Product (left) and ESRI World Imagery (right) for a rural area in France illustrating the interruption of narrow linear elements which do not reach the majority per 10 m pixels (black grid lines) and typically show mixed spectral signals in the Sentinel-2 time series. ...	13
Figure 7: CLC+ Backbone Raster LC Product superimposed on ESRI World Imagery (left) and VHR IMAGE 2018 (right) for an area with sparse tree cover in the Alps. Sparse tree canopies dominated by herbaceous understory are mapped largely correct as class 6.....	13
Figure 8: Example of the mapping of ephemeral classes in the CLC+ Backbone Raster Product. Water is mapped according to its spatio-temporal coverage during the reference year 2018. In the given example, areas which are more than 50% of the observations covered by water (i.e. NDVI typically below 0), are correctly mapped as Water. Areas with shorter water coverage show an NDVI slightly above 0 for most of the time in this example and are correctly mapped as Sparsely- and Non-vegetated.....	14
Figure 9: Example of the mapping of ephemeral classes in the CLC+ Backbone Raster Product. Permanent Snow and Ice is mapped according to its spatio-temporal coverage during the reference year 2018. In the given example, areas which are more than 90% of the observations covered by snow or ice (i.e. NDSI typically above 0.41), are correctly mapped as Snow and Ice. Adjacent areas with shorter snow or ice coverage are correctly mapped as Sparsely- and Non-vegetated.	14
Figure 10: Example of the mapping of temporarily barren agricultural fields in CLC+ Backbone Raster Product in Spain. Parcels that are cultivated during the main growing season (Mediterranean winter rain season) show a strong NDVI peak in early 2018. In contrast, parcels which are not cultivated during the vegetation season 2018 show a very faint NDVI signal until October of the reference year (i.e. actually the beginning of the cultivation season 2019) and are hence correctly mapped as Non- and Sparsely Vegetated.....	15
Figure 11: Example of the mapping of dry grasslands as Permanent Herbaceous in CLC+ Backbone Raster Product in Turkey. While in Central European and Northern Europe, permanently herbaceous areas are typically characterized by a strong NDVI peak during the vegetation season, dry grasslands in Southern Europe and Turkey are characterized by far fainter vegetation signals.	16

Figure 12: Example of peat extraction areas mapped as Sparsely Vegetated. CLC+ Backbone raster product (left) and false colour VHR Image 2018 (right).....	17
Figure 13: Peat production area in the process of renaturation, already showing dense vegetation cover in large parts.....	17
Figure 14: Production Units (PUs) subdividing the EEA-38 + UK area into homogenous biogeographical sub-strata for the Raster Product.....	19
Figure 15: Border of tree covered area (cyan dots) extending into a field, resulting at least partially from geometric shifts in S-2 time-series. Red dots mark pixels where this leads to commission errors in tree cover (grid: 10m pixels in LAEA projection)	22
Figure 16: Mapping of a complex mix of wild olive trees, shrub, coniferous and permanent herbaceous vegetation. Left: CLC+ Backbone Raster product. Right: VHR Image 2018 false colour infrared. ...	22
Figure 17: Example of the uncertainty in the mapping of Low-growing woody vegetation. This example from Northern Sweden illustrates that a sharp distinction of shrubs and trees is often difficult and sometimes remains fuzzy even at the level of the same species (here different specimens of Betula with sometime more tree-like, sometime more shrub-like habitus). In addition, the canopy is rather sparse and the understory comprises a mix of herbaceous species, dwarf shrubs as well as mosses and lichens.....	23
Figure 18: Example of the uncertainty in the mapping of Lichens and Mosses in the CLC+ Backbone Raster Product. These two examples from Southern Norway illustrate the typical similarity between surfaces covered by Lichens and Mosses versus Non- and Sparsely Vegetated areas with some fractions of sparse herbaceous vegetation. Both land cover compositions occur under similar biogeographic conditions and are nearly indistinguishable in VHR imagery or NDVI time-series. .	24
Figure 19: Greenhouse partially classified as sparsely vegetated class. Greenhouses should be classified as Sealed, but in some cases, unfavourable reflection angles may cause confusion with the Sparsely Vegetated class.....	24
Figure 20: Example of the mapping of intensively used grassland (e.g. fodder crops). The distinction between the classes Permanent Herbaceous and Periodically Herbaceous depends on whether the bare soil is exposed (e.g. ploughed parcel on the right, mapped as Periodically Herbaceous) within the reference year or not. Especially under temporarily very dry conditions, the distinction of mowing events (i.e. no ploughing on the left parcel, mapped as Permanent Herbaceous) from actual exposure of bare soil can be very difficult.	25
Figure 21: Decision tree for the usage of area thresholds for polygons with (pure or) spatially mixed land cover in the Vector Product.	33
Figure 22: Example of Vector Product in part of Luxembourg.....	35
Figure 23: Example of Vector Product in part of Hungary	35
Figure 24: Secondary Production Units (SPU) subdividing the EEA-38 + UK area into homogenous biogeographical sub-strata for production of Vector Product.....	36
Figure 25: MMW tolerance for the connectivity.....	38
Figure 26: Illustration of cell centre extraction method	40
Figure 27: FAD classes and respective value ranges	48
Figure 28: Example MSPA raster showing all possible MSPA classes.....	48
Figure 29: Example of apparent regional segmentation density difference in Croatia (SPU 129_1 & 129_2) at small and large scales	52

Figure 30: Example of regional segmentation density difference in Spain (SPU 74, 70 and 71_1) at small and large scales	53
Figure 31: Artificial cut (yellow) in river and major road polygons at the SPU border (black).....	54
Figure 32: Artificial cut (yellow) in ocean polygons at the EEA 10km grid (red)	54
Figure 33: CLC+ Backbone Data Score Layer for Sentinel-2 observations of the reference year 2018 (± 6 months) over the area of the EEA-38 + UK.....	56
Figure 34: CLC+ Backbone Raster Confidence Layer 2018	58
Figure 35: Distribution of CLC+ Backbone Raster Confidence Layer values across the EEA-38 + UK.....	59
Figure 36: CLC+ Backbone Raster Post-processing Layer 2018	60
Figure 37: Segmentation Confidence Layer 1 (CONF_TEX, above) and Confidence Layer 2 (CONF_SPEC, below)	62
Figure 38: Distribution of the sampling and reporting units for the EU27 Raster Product.....	64
Figure 39: QGIS validation tool to support the interpretation of large numbers of sampling units for Raster Product (above) and Vector Product (below)	66
Figure 40: QGIS validation tool supporting the interpretation of large numbers of sampling units	69
Figure 41: NDVI – Example of spectral profile units.....	70
Figure 42: Overview of geometric criteria for CLC+ Backbone vector product validation.....	75
Figure 43: Blind interpretation process.....	76
Figure 44: Positional accuracy analysis	77
Figure 45: Potential over-segmentation accepted as correct: (a) same land cover, but statistics are likely to be different enough to be mapped separately, (b) slight variation of single feature not applicable for all layers, (c) small objects below MMU, that could be detected as separate class	77
Figure 46: Graphic illustration of segment evaluation for Jaccard index.....	78
Figure 47: Geometric plausibility check via visual interpretation	78

LIST OF TABLES

Table 1: Overview of usage of auxiliary datasets during the post-processing routine.....	20
Table 2: Examples for interpretation of Vector Product nomenclature and decision tree for different land cover compositions	34
Table 3: Overview of potential input datasets for Hardbone production (region specific selection applied)...	37
Table 4: Selected OSM code and category for the application of a buffer.	38
Table 5: Overview of Segmentation Confidence Layers.....	61
Table 6: Overview of the validated areas and distribution of sampling units in each CLC+ Backbone raster class.....	64
Table 7: Area fractions and validation sample distributions for the 11 LC classes of the EEA-38+UK Raster Product.....	65
Table 8: Overview of computed accuracy metrics.	67
Table 9: Colour coding used for the presentation of the different accuracy levels.....	67
Table 10: Area-weighted Overall Accuracy for CLC+ Backbone Raster Product in percent for the reporting units covering the EEA-38 + UK area.....	68
Table 11: Area-weighted Producer's and User's Accuracies in the CLC+ Backbone Raster Product in percent for the reporting units covering the EEA-38 + UK area.....	68
Table 12: Overview of computed accuracy metrics.	70
Table 13: General requirement: Specified Accuracy thresholds for CLC+ Backbone Vector Product	71
Table 14: Overall Accuracies for blind and plausibility analyses with Confidence Intervals (95%) for the CLC+ Backbone Vector Product, per country or group of countries for EEA-38+UK area.....	72
Table 15: Producer's and User's Accuracies of CLC+ Backbone Vector product, given by land cover class– Blind and plausibility analysis for EEA-38+UK area (colour coding for PA: see Table 13).....	73
Table 16: Producer's and User's Accuracies of CLC+ Backbone Vector Product, given by land cover class and by country (group) – Plausibility analysis for EEA-38+UK	74
Table 17: Geometric accuracy metrics	76
Table 18: Results of geometric validation of the CLC+ Backbone Vector Product (summary, excluding TRF, DOM).....	79
Table 19: CLC+ Backbone Vector Product geometric validation results per biogeographic region (relative percentage values).....	80
Table 20: Results of plausibility analysis of the 100m buffer region for EEA-38+UK	81
Table 21: Overview of buffer zone check of 100m buffer of approx. 20% of the samples.....	82
Table 19: Colour palette for the Raster Product	86
Table 20: Colour palette for the Vector Product.....	87
Table 21: Colour palette for the Data Score Layer	88
Table 22: Adapted colour palette for the Data Score Layer of the DOMs	88
Table 23: Colour palette for the Raster Confidence Layer	88
Table 24: Colour palette for the Raster Post-processing layer.....	89
Table 25: Colour palettes for the Segmentation Confidence Layer 1 (above) and 2 (below)	89

ACRONYMS AND ABBREVIATIONS

AD	Applicable Document
AoI	Area of Interest
ARVI	Atmospherically Resistant Vegetation Index
BAI	Burned Area Index
CL	Confidence Layer
CLC, CLC+	CORINE Land Cover, CORINE Land Cover +
CLMS	Copernicus Land Monitoring Service
CZ	Coastal Zones
DEM	Digital Elevation Model
DG	Directorate General
DOMs	French Overseas Departments
DSL	Data Score Layer
EAGLE	EIONET Action Group on Land monitoring in Europe
EEA	European Environment Agency
EEA-38 + UK	The 32 member and 6 cooperating countries of the EEA plus the United Kingdom
EFFIS	European Forest Fire Information System
EIONET	European Environment Information and Observation Network
EO	Earth Observation
EPSG	European Petroleum Survey Group
ETRS89	European Terrestrial Reference System 1989
EU	European Union
EU28	The 28 member states of the European Union
EU-Hydro	European Hydrography Layer
EUROSTAT	European Statistical Office
FMask	Function of mask
GAFSEG	software developed by GAF AG
GIS	Geographic Information System
HDF5	Hierarchical Data Format 5
HR	High resolution
IMD	Imperviousness Density
INSPIRE	INfrastructure for SPatial InfoRmation in Europe
JRC	Joint Research Centre
L2A	Level 2A
LAEA	Lambert Azimuthal Equal Area
LC	Land Cover
LC/LU	Land Cover/Land Use
LCC	Land Cover Component
LiDAR	Light detection and ranging
LPIS	Land parcel identification system
LUCAS	Land Use/Cover Area frame Survey
LZW	Lempel–Ziv–Welch data compression algorithm
MMU	Minimum Mapping Unit
MMW	Minimum Mapping Width
NBR	Normalized Burn Ratio
NDVI	Normalized Difference Vegetation Index
NDWI	NDWI Normalized Difference Water Index
OSM	Open Street Map
PU	Production Unit
QA	Quality Assurance
QC	Quality Control

RZ	Riparian Zones
S-2	Sentinel-2
SPU	Secondary Production Unit
TCD	Tree Cover Density
TF	Time Feature
TIFF	Tagged image File Format
UA	Urban Atlas
UK	United Kingdom
UTM	Universal Transverse Mercator
VHR	Very High Resolution
WAW	Water and Wetness
WGS84	World Geodetic System 1984
WISE	Water Information System for Europe
WKT	Well-known-Text
XML	Extensible Markup Language

1 Executive Summary

Copernicus is the European Union's Earth Observation Programme. It offers information services based on satellite Earth observation and in situ (non-space) data. These information services are freely and openly accessible to its users through six thematic Copernicus services (Atmosphere Monitoring, Marine Environment Monitoring, Land Monitoring, Climate Change, Emergency Management and Security).

The **Copernicus Land Monitoring Service (CLMS)** provides geographical information on land cover and its changes, land use, vegetation state, water cycle and earth surface energy variables to a broad range of users in Europe and across the world in the field of environmental terrestrial applications. The CLMS is jointly implemented by the European Environment Agency (EEA) and the European Commission's DG Joint Research Centre (JRC).

The **CLC+ Backbone** constitutes the first component of the CLMS's new 'CLC+ Product Suite', which represents a true paradigm change in European land cover/land use (LC/LU) monitoring, building on the rich legacy of the European CORINE Land Cover (CLC) flagship product. The CLC+ Backbone is an object-oriented, large scale, wall-to-wall (EEA-38 + UK), high-resolution (HR) inventory of European LC in a vector format accompanied by a raster product layer, providing a consistent pan-European geometric backbone of Landscape Objects with limited, but robust thematic detail, on which many other applications can be built.

The CLC+ Backbone products are provided in the European Terrestrial Reference System 1989 (ETRS89) in Lambert Azimuthal Equal Area (LAEA) projection, for all of the EEA-38 + UK area. Country-wise products in national projections are made available by the EEA as well. CLC+ Backbone products of the reference year 2018 consist of two main products, based primarily on Copernicus Sentinel satellite imagery from 2017, 2018 and 2019:

- 1) a pixel-based, multi-temporal Sentinel-2 time-series based **Raster Product** with 10m spatial resolution and 11 basic LC classes;
- 2) an object-based **Vector Product** with 0.5 ha MMU, derived from a combination of linear traffic and hydrological networks (Hardbone) and image segmentation (Softbone), and 18 LC classes attributed with aggregated statistics from the Raster Product as well as various **additional characteristics**.

This document constitutes the third and final Issue of the Product Specifications and User Manual. It provides detailed specifications of all CLC+ Backbone products and attributes, as well as a documentation of the applied production methodologies, together with discussions of the strengths and limitations of the products. Furthermore, the document contains summaries of the extensive internal validations conducted for the Raster and Vector products.

The CLC+ Backbone Product Specification and User Manual is intended to provide all product information to users that may be required for successful further-reaching analyses and applications.

2 Background of the Document

2.1 Scope of the Document

This Product Specification and User Manual is the primary document that users are recommended to read before using the product. It provides a description of the product characteristics, production methodologies and workflows, and information about the product quality. Furthermore, it gives information on the terms of use and product technical support. It constitutes the consolidated final Issue 3.0 of the document.

2.2 Content and Structure

In more detail, the document is structured as follows:

- Chapter 3 presents the detailed product descriptions, including details on the applied production methodology and workflows;
- Chapter 4 summarizes the information on the product quality, including details on the applied assessment procedure;
- Chapter 5 provides information about product access and use conditions as well as on the technical product support;
- Chapter 6 lists references to the cited literature; and
- the Annexes provide technical details with respect to product file naming, colour palettes and projection parameters.

3 Product Description

This chapter provides a comprehensive overview of the CLC+ Backbone products' specifications, putting them in the context of the overall CLC+ product suite portfolio (section 3.1), presenting the characteristics of the Raster Product (section 3.2), the Vector Product (section 3.3) and the various accompanying Additional Quality Layers and Expert Products (section 3.4).

3.1 Overview of the product portfolio

The CLC+ Backbone comprises two main status layers for the reference year 2018 (Figure 1): The Raster Product represents an 11-class land cover classification at 10m spatial resolution; the Vector Product delineates landscape objects with an MMU of 0.5ha and assigns an 18-class classification to each landscape object polygon according to the land cover composition of the Raster Product within the respective object. The Raster Product is accompanied by three expert products being a Confidence Layer (the confidence of the initial classification) a Data Score Layer (the number of used valid Sentinel-2 observations) and a Post-Processing Layer (marking pixels whose initial classification class code was adjusted during post-processing correction steps. The land cover class-assignment in the Vector Product is complemented by series of > 50 attributes derived from the CLC+ Backbone Raster product, existing CLMS products, and products from the European Forest Fire Information System (EFFIS), the Copernicus DEM and from the Data Score Layer. Furthermore, it comprises a spatially explicit Segmentation Confidence Layer expert Product (as attribution).

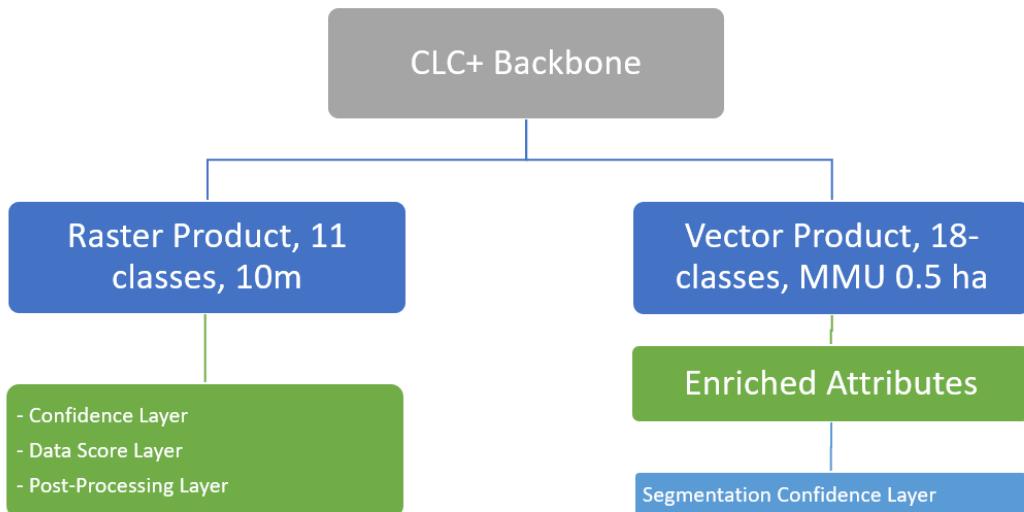


Figure 1: High-level overview of the CLC+ Backbone product portfolio.

3.2 Raster Product

The CLC+ Backbone Raster product is a 10m pixel-based land cover map based on Sentinel data for the reference year 2018. Each pixel shows the dominant land cover among the 11 basic land cover classes. The following sections provide information on the product specifications (section 3.2.1) including the nomenclature concept and class definitions (section 3.2.1.1) as well as the related decision tree approach (section 3.2.2). Additionally, section 3.2.1.3 provides some illustrated examples of typical cases of the Raster Product nomenclature and decision tree application. Details on the production methodology (section 3.2.2), including an overview of the input data, the time series classification approach and post-processing steps are given in sections 3.2.2.1 to 3.2.2.3. Additionally, section 3.2.2.4 provides a comprehensive overview of the strengths and limitations of the applied approach.

3.2.1 Product Specifications

The product specifications for the CLC+ Backbone Raster product are summarized in the table below. Further product details and information on the methodology, nomenclature etc. can be found in the following chapters.

CLC+ Backbone Raster Product	Acronym	Product family
	RASTER	CLMS_CLCplus
Summary		
CLC+ Backbone is a spatially detailed, large scale, EO-based land cover inventory. The CLC+ Backbone Raster Product is a 10m pixel-based land cover map based on Sentinel time series from July 2017 to June 2019 and auxiliary features. For each pixel it shows the dominant land cover among the 11 basic land cover classes.		
The Raster product is clipped at each national border and re-projected to respective national projections.		
Reference year		
2018		
Geometric resolution		
Pixel resolution 10m x 10m, fully conform with the EEA reference grid		
Coordinate Reference System		
European ETRS89 LAEA projection / for French DOMs WGS84 and the respective UTM zone		
National products re-projected to respective national projections.		
Coverage		
6,002,168 km ² (covering the full EEA-38 + UK)		
Geometric accuracy (positioning scale)		
equals the Sentinel-2 positional accuracy in 2018 (~11m at 95.5% confidence)		
Thematic accuracy		
90 % overall accuracy, not more than 15 % omission errors and 15 % commission errors per class (the amount of omission and commission errors for particular difficult classes such as Low-growing woody plants and Lichens and Moses might regionally exceed those thresholds)		
Data type		
8bit unsigned raster with LZW compression		

Minimum Mapping Unit (MMU)
Pixel-based (no MMU)
Necessary attributes
Raster value, class name, pixel count
Raster coding (thematic pixel values)
1: Sealed 2: Woody – needle leaved trees 3: Woody – Broadleaved deciduous trees 4: Woody – Broadleaved evergreen trees 5: Low-growing woody plants (bushes, shrubs) 6: Permanent herbaceous 7: Periodically herbaceous 8: Lichens and mosses 9: Non- and sparsely-vegetated 10: Water 11: Snow and ice 254: outside area 255: NoData
Metadata
XML metadata files according to INSPIRE metadata standards
Delivery format
GeoTIFF incl. pyramids (*.ovr, level: 11, resampling: Nearest Neighbour), attribute table (*.dbf), statistics (*.aux.xml), integrated as well as external colour tables (*.clr) and INSPIRE-compliant metadata in XML format

By definition, the above land cover nomenclature does not provide a separation between inland water and coastal/sea water along a country's sea border and in the respective 250m coastal buffer zone. For potential use cases intending to separate both, a clipping with a national/European coast line dataset is recommended.

3.2.1.1 Nomenclature concept and class definitions

Most of the class definitions for CLC+ Backbone Raster Product comprise a 50% area threshold to express that the dominant land cover should be assigned. While this is a plausible approach there are many situations where the land cover within a single 10m pixel comprises a spatial mix of different classes and no single class reaches an absolute majority of 50%. A simple example is provided in Figure 2.

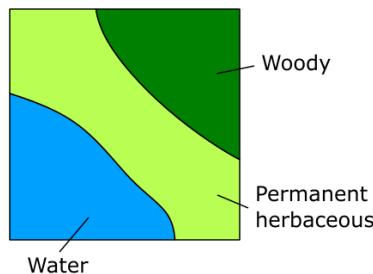


Figure 2: Exemplary 10x10m plot illustration of a situation where no single land cover class reaches 50%.

To address this issue a relative majority is used in most case for the class assignment. This detailed logic is implemented through a decision tree approach which does not only define the area threshold but further clarifies and limits for each class the Land Cover Component (Figure 3) considered when evaluation the area threshold. The approach is detailed in section 3.2.1.2. In general, the area fractions of different land cover classes are to be understood as what can be reasonable be evaluated in HR and VHR remote sensing imagery at nadir.

Land Cover Components (LCC)																			
ABIOTIC / NON-VEGETATED SURFACES AND OBJECTS							BIOTIC / VEGETATION												
Artificial Surfaces and Constructions				Natural Material Surfaces			Woody Vegetation		Herbaceous Vegetation (grass-like, forbs, ferns)			WATER							
Sealed Artificial Surfaces and Constructions	Artificial Surfaces and Constructions	Non-Sealed and Artificial Surfaces	Consolidated Surfaces	Natural Material Surfaces			trees	Bushes, Shrubs	Graminoids (grass-like)	Succulents and cacti	Lichens, Mosses and Algae	Liquid waters	Solid waters						
Buildings	specific structures and facilities	open sealed surfaces	waste materials	non-sealed and semi-sealed artificial surfaces	bare rock	hard pan	Mineral Fragments	bare soils	Natural Deposits	regular bushes	dwarf shrubs, grasses, sedges, rushes, cereals (low ground), reeds, bamboo, canes (high)	Graminoids (grass-like)	macro algae	micro algae (In plankton)	water courses	Inland Waters	Marine Waters	Snow	Ice, Glaciers

Figure 3: Levels 1-4 of the EAGLE Land Cover Components (LCC) v3.1. The component hierarchy shall be used to resolve the class assignment in cases where no single land cover class dominates.

Remark on the temporal dimensions of the land cover classes: In cases of land cover changes during the reference year, and unless stated otherwise in the class definitions (see below i.e. Permanent herbaceous vs. Periodically herbaceous, Snow and Ice) the dominant land cover (i.e. present > 6 month/ year) is mapped.

Textual descriptions of the main land cover components included in each of the 11 classes are provided in the following paragraphs. The area thresholds are excluded here and detailed in section 3.2.1.2.

1. Sealed:

Sealed Artificial Surfaces include all impervious and sealed surfaces that are covered mainly by features with a specific height above ground (buildings and artificial constructions) or features without a specific height above ground (flat impervious surfaces). Flat surfaces covered by any type of impervious material that is used for artificial surface pavements (e.g. asphalt, concrete, tarmacadam).

- Includes: All sealed artificial surfaces and constructions including Buildings, Specific structures and facilities, and open sealed surfaces (EAGLE land cover components). Also vegetated rooftops are to be mapped under this class. Railway tracks are also considered as part of this class since they typically comprise impervious structural elements and a highly compacted subsoil.
- Excludes: Waste materials (e.g. communal / industrial waste), non-sealed and semi-sealed artificial surfaces (e.g. nat. mat. displaced from original place, artificially consolidated, e.g. logistic and storage areas, festive squares, non-vegetated sport fields, grass pavers, permeable paving (de: "Rasengittersteine"). Such areas are to be mapped as Non- and sparsely-vegetated since Biotic LC components do typically not exceed 30%.

2. Woody - trees

Perennial woody plants with single, self-supporting main stem or trunk, containing woody tissue and branching into smaller branches and shoots.

Excluded	Destination class
Pinus mugo and Alnus viridis	Low-growing woody plants
Ephedra	Low-growing woody plants
Shrub forms of Taxus, Juniperus and Betula	Low-growing woody plants
Musa	Permanent herbaceous

2. 1. Woody – needle leaved trees

Needle leaved trees: referring to trees of the botanical group Gymnospermae (Ford-Robertson, 1971) carrying typical needle-shaped leaves. An exception is Ginkgo biloba which belongs to the Gymnospermae but is considered here as Broadleaved deciduous tree.

2. 2. Broadleaved trees: referring to trees of the botanical group Angiospermae, with the exception of ginkgo (Ginkgo biloba), which belongs to the Gymnospermae taxonomically.

2. 2. 1 Woody – Broadleaved deciduous trees: broadleaved trees which are leafless for a certain period during the year

2. 2. 2 Woody – Broadleaved evergreen trees: trees that are never entirely without green foliage (includes palm-leaved species)

5. Low-growing woody plants (bushes, shrubs)

Perennial woody plants with shrub growth form i.e. multiple stems arising at or near the base, height usually less than 5 metres. Leaf type can be needle leaf, broadleaf or palm leaf, phenology either evergreen or deciduous, leaf surface type can be regular or sclerophyllous.

- Includes: regular bushes and dwarf shrubs, species such as Pinus mugo, Alnus viridis, shrub forms from the genus Ephedra, Taxus, Juniperus and Betula, subshrubs such as Thymus spec., Salvia, Rosmarinus, Calluna vulgaris, Erica spec. . The class also includes Vitis spec. and Humulus spec., which are typically permanent crops. Individual small trees in shrub-dominated areas are allowed in this class.
- Excludes: Low-growing fruit trees (e.g. apple plantations), tree cover regrowth (e.g. after clear cuts) with sufficient density and trees in nurseries, which should be classified according to the definitions for Woody-trees.

Remark: Due to the difficulty of differentiating shrubs / bushes from trees and herbaceous vegetation the accuracies could be regionally below the defined target accuracies.

6. Permanent herbaceous

- Permanent herbaceous areas are characterized by a continuous vegetation cover throughout a year. No bare soil occurs within a year. These areas are either unmanaged or extensively managed natural grasslands or permanently managed grasslands, or arable areas with a permanent vegetation cover

(e.g. fodder crops) or even set-aside land in agriculture. For managed grasslands, the biomass varies over the year, depending on the number of mowing (grassland cuts) or grazing events.

- According to IACS/LPIS definition a permanent and managed grassland may be ploughed every 3-5 years for amelioration purposes followed by an artificial seeding phase and a renewal of vegetation cover, thus potentially showing a phase of bare soil within a time frame of 5-6 years. Given that the observation period for the CLC+ Backbone Raster Product is of 1 year with a focus on land cover, such longer-term land use patterns cannot be considered and therefore grasslands which underwent ploughing within the reference year are typically mapped under the class 7. Periodically Herbaceous.

The class includes regular graminaceous (grasses), reeds and forbs, notably also natural dry grassland in Southern, South-Eastern Europe and Turkey, as well as banana plantations (*Musa spec.*).

7. Periodically herbaceous

- Periodically herbaceous areas are characterized by at least one land cover change (in the sense of EAGLE land cover components) between bare soil and herbaceous vegetation within one year. Depending on the management intensities these areas can also have up to several changes between these two EAGLE land cover components within a year. Normally these areas are managed as arable areas.

8. Lichens and mosses

- Any type of lichens – composite organisms formed by a symbiotic relationship of a fungus and a photosynthetic partner (usually green algae or cyanobacteria);
- Mosses: Non-vascular plants in the land plant division Bryophyta. They are small (a few centimetres tall) herbaceous (non-woody) plants that absorb water and nutrients mainly through their leaves and but also photosynthesize;
- Typical vegetation class of northern European Tundra vegetation;

Remark: The mapping of this class is focused on larger areas in northern Europe where sufficient in situ data is available. Due to the difficulty to distinguish Lichens and Mosses from herbaceous vegetation and dwarf shrubs, Producer's and User's Accuracies below 80% and a corresponding impact on the Overall Accuracy are to be expected.

9. Non- and sparsely-vegetated (i.e. rock, screes, sand, permanent bare soils)

Contains consolidated and unconsolidated materials as well as permanent bare soils, where non-vegetated areas cover >= 70 % of the land surface;

- Consolidated surfaces (rocks):
 - The rock surface is continuous except perhaps for a few cracks in the material. Some areas may be covered by shallow layers of soil or there could be isolated pockets of soil or a mixture of both;
 - Examples: solid (closed) rock formations, fresh lava flows, quarries, mineral extraction sites, open pit mines.
- Unconsolidated surfaces (screes, sand, permanent bare soil)

- Mineral Fragments come to be through mainly physical disintegration of geological formations and are the result of becoming smaller and smaller along time. They are accumulated on site due to sedimentary processes or human activity;
- Includes variable particle sizes: boulders, stones, pebble, gravel, sand and clay
- Examples: mountain slope debris, gravel river banks, open pit pebble mining of fossil river banks or fluvial sediments, volcanic lapilli fields, sand dunes, sand beaches, river sand banks, volcanic ash.
- Permanent bare soil:
 - Mixture of mineral and organic material that is fertile enough and capable of sustaining plant life, but being continuously un-vegetated during the entire observation period.
- Sparsely vegetated areas:
 - Sparsely vegetated on unstable areas (stones, boulders, rubble on steep slopes, or anthropogenic activity), due to harsh environmental conditions or anthropogenic interference. Biotic land cover below <30%.
- Any other non-sealed artificial surfaces and constructions with a vegetation cover <30%.
- Organic and in-organic deposits with a vegetation cover <30%.

10. Water

- Water in liquid state of aggregate regardless of location, shape, salinity and origin (natural or artificial);
- Includes: running water (water courses) and standing water (natural lakes, fishponds, man-made reservoirs, pools, irrigation ponds, etc.);

Remark: Regarding the temporal coverage, the area should be under water at least 50% of the observation period; temporary ice cover of water bodies included.

11. Snow and ice

- Snow: areas covered permanently (> 90 % of observation period) with snow throughout the year;
- Ice: persistent ice cover formed by accumulation of snow (> 100 % of observation period);
- and combinations of both (e.g. in case of glaciers being covered by snow for parts of the year)

3.2.1.2 Decision tree for area thresholds

The CLC+ Backbone Raster Product decision tree (Figure 4) complements the class definitions in section 3.2.1.1 and the EAGLE LCC matrix (Figure 3), to assure a seamless definition of the classes not only for “pure” pixels, but also in cases of mixed land cover at the scale of the 10m resolution of the raster product. At each decision level the tree defines the reference area which should be considered (blue text next to the rhombus) and the area threshold for a specific land cover class in percentages of the reference area (white text in rhombus). Generally, the decision tree targets to define an unambiguous assignment of pixels with (pure or) mixed land cover to the dominant land cover class. In accordance to the EAGLE concept it refers to the dominant land cover independent of the land use. Pixels which are for example dominated by Permanent Herbaceous should be considered as such even though the dominant land use might be an orchard, fruit plantation or Dehesa. Further examples are given in sub-section 3.2.1.3.

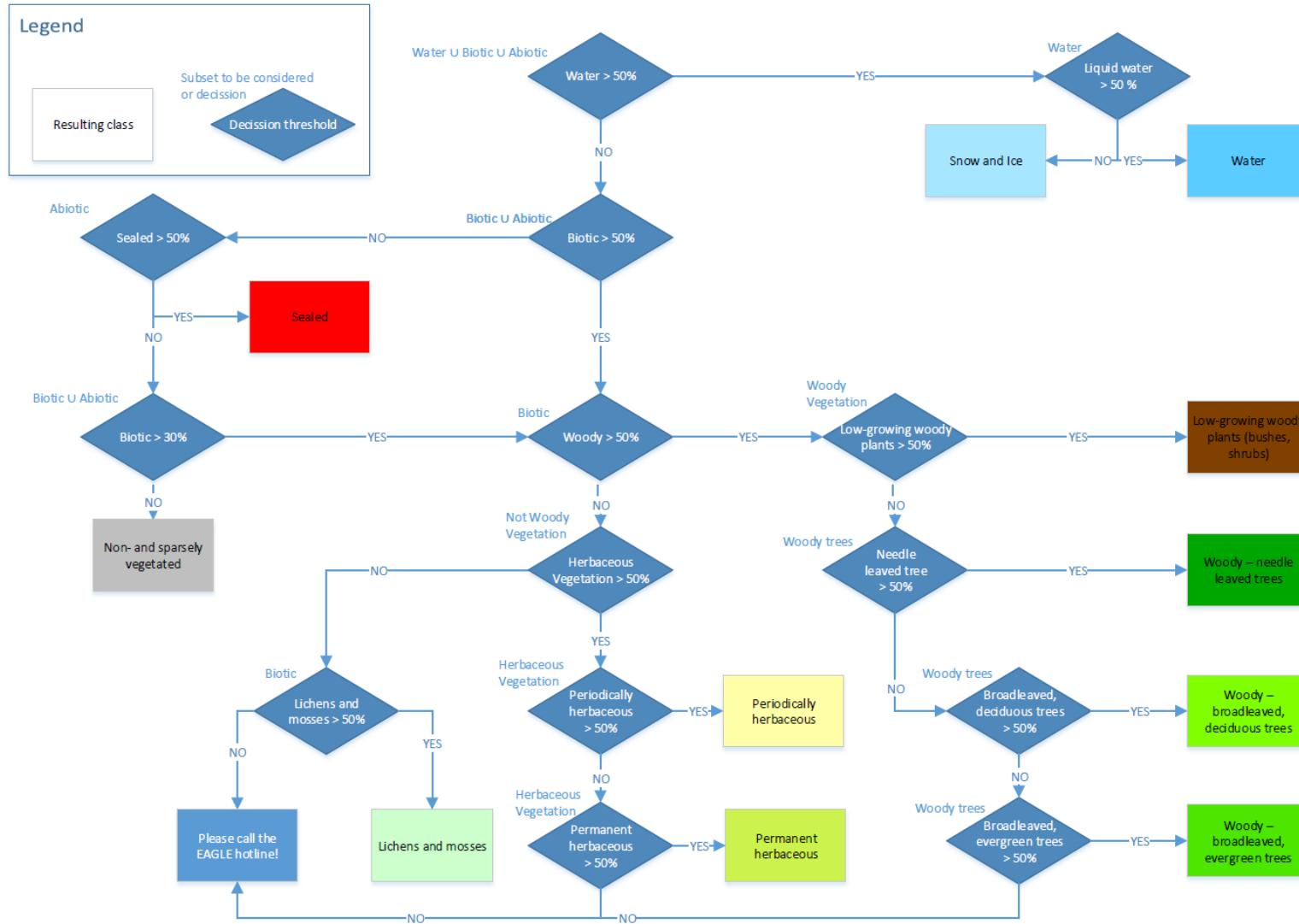
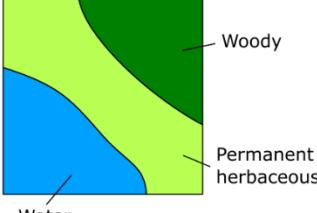


Figure 4: Decision tree for the usage of area thresholds for pixels with (pure or) spatially mixed land cover in the Raster Product.

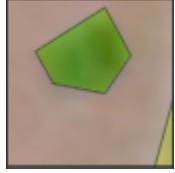
3.2.1.3 Examples

Some concrete examples for the usage of the decision tree are provided in the following:

Example 1 – Hypothetical example - considering the mix which was already presented in Figure 2:

 <p>Water</p> <p>Permanent herbaceous</p> <p>Woody</p>	<ol style="list-style-type: none"> 1. Does Water cover more than 50% (considering the sum of Water U Biotic U Abiotic)? > NO 2. Does Biotic cover more than 50% (considering the sum of Biotic U Abiotic)? > YES 3. Does Woody cover more than 50% (considering the sum of Biotic)? > NO 4. Does Herbaceous vegetation cover more than 50% (considering the sum of Not Woody vegetation)? > YES 5. Does Periodically herbaceous cover more than 50% (considering the sum Herbaceous vegetation)? > NO 6. Does Permanent herbaceous cover more than 50% (considering the sum Herbaceous vegetation)? > YES 7. Permanent herbaceous
---	--

Example 2 – Practical example extracted from Figure 5 (Row 2, Column 5):

 <p>Sealed: 0 m²</p> <p>Broadleaved evergreen trees: 17.54 m²</p> <p>Permanent herbaceous: 2.34 m²</p> <p>Non-vegetated: 80.12 m²</p>	<ol style="list-style-type: none"> 1. Does Water cover more than 50% (considering the sum of Water U Biotic U Abiotic)? > NO 2. Does Biotic cover more than 50% (considering the sum of Biotic U Abiotic)? > NO 3. Does Sealed cover more than 50% (considering the sum of Abiotic)? > NO 4. Does Biotic cover more than 30% (considering the sum of Biotic and Abiotic)? > NO 5. Non- and sparsely-vegetated
--	--

Example 2 illustrates the application of the decision tree over an area with olive groves in Southern Spain. Since the growth of the understory is largely suppressed (tillage and use of herbicides are common) and the tree cover is relatively sparse, the Biotic coverage percentage for most pixels does not exceed the 30% threshold and the pixel should hence be considered as Non- and sparsely-vegetated. In cases where the suppression of the herbaceous understory is not continuous (i.e. alteration between herbaceous cover and bare soil within one year), such areas should be mapped as Periodically Herbaceous instead. As illustrated in Figure 5, the same applies for wide areas where such land cover characteristics are dominant, in line with the key paradigm of the EAGLE concept to disentangle land cover and land use.



Figure 5: Illustration of the CLC+ Backbone raster product decision tree approach to evaluate area thresholds for mixed land cover at olive groves, Castile-La Mancha, Spain. Upper left: 100x100m plot overlaid on aerial imagery. Upper right: Outlines of surfaces which are covered by Sealed, Permanent herbaceous, Broadleaved evergreen trees and Non- and sparsely-vegetated. 10m pixel grid overlaid on top. Lower left: Resulting class attribution of the 10m pixel when applying the threshold of Biotic >30% to decide between Non- and sparsely-vegetated vs. one of the vegetation classes. Lower right: ©Google Street View of the area, showing the four dominant land cover classes.

Example 3 below (Figure 6) shows the map result for a rural area in France, illustrating some interruptions of small linear landscape elements (rural road in this case). The apparent omission of such narrow linear elements (e.g. narrow roads and water courses, narrow tree lines and hedge rows) is a common feature of the CLC+ Backbone Raster Product, since a) such narrow landscape elements do often not occupy the majority of the pixels and b) are registered in the time series with a mixed spectral signal due to the sensor point spread function and imprecisions in the multi-temporal spatial co-registration.

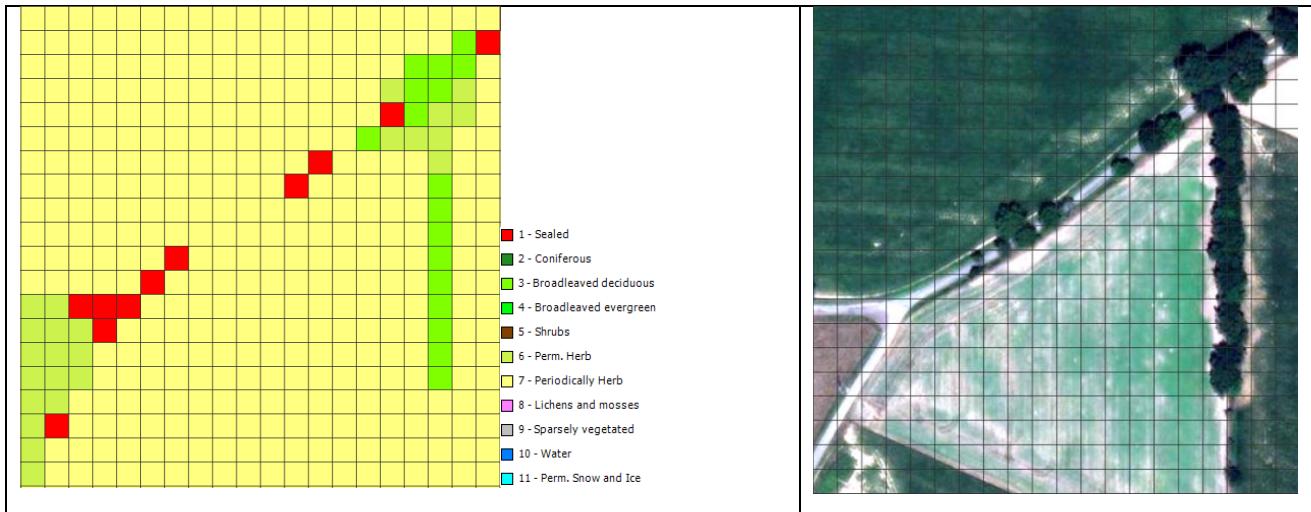


Figure 6: CLC+ Backbone Raster LC Product (left) and ESRI World Imagery (right) for a rural area in France illustrating the interruption of narrow linear elements which do not reach the majority per 10 m pixels (black grid lines) and typically show mixed spectral signals in the Sentinel-2 time series.

Example 4 (Figure 7) illustrates the largely correct mapping of an area with sparse coniferous tree cover and herbaceous understory as class 6 (Permanent Herbaceous). While the mapping of such mixed areas might be unexpected for some users and might not be reflected in national / regional land cover maps, it correctly represents the dominant land cover and its spectral-temporal properties.

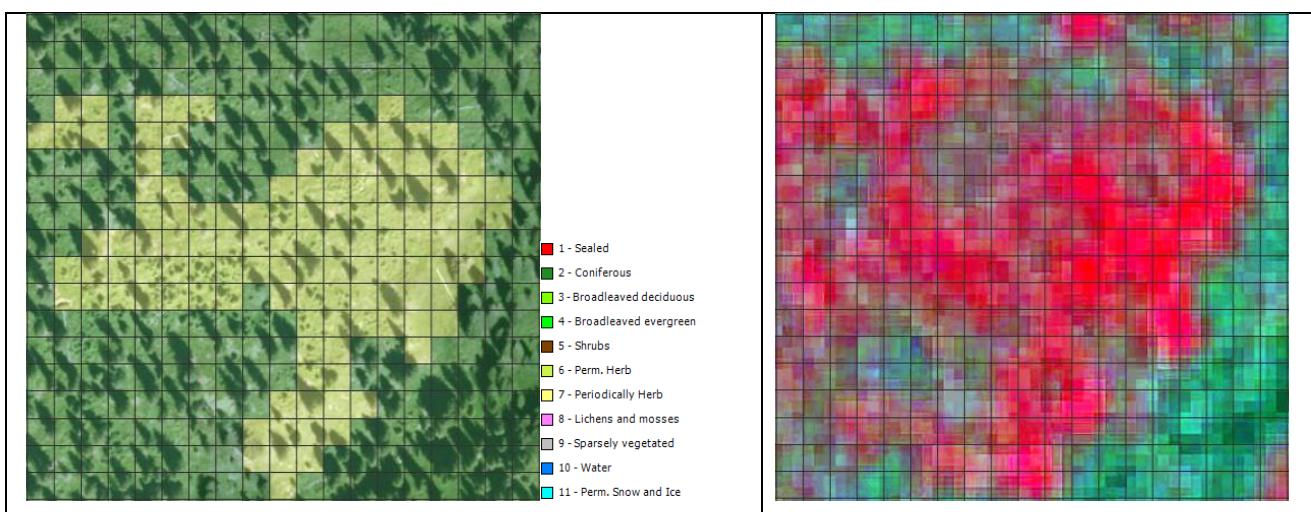


Figure 7: CLC+ Backbone Raster LC Product superimposed on ESRI World Imagery (left) and VHR IMAGE 2018 (right) for an area with sparse tree cover in the Alps. Sparse tree canopies dominated by herbaceous understory are mapped largely correct as class 6.

Example 5 concerns the mapping of Water and Permanent Snow & Ice, which are defined according to their spatio-temporal extent throughout the reference year, with at least 50% and 90% permanence within the period, respectively. As illustrated in Figure 8 and Figure 9, the production has taken these thresholds into account, considering temporal profiles of the NDVI and NDSI during the generation of training data, quality control and internal validation. It is therefore important to note that single

satellite images at deliberate times can suggest different spatial extents for these two classes, depending on the time of acquisition. Time-series should be consulted in case of doubts for assessing the correctness of the extent in the Raster Product.

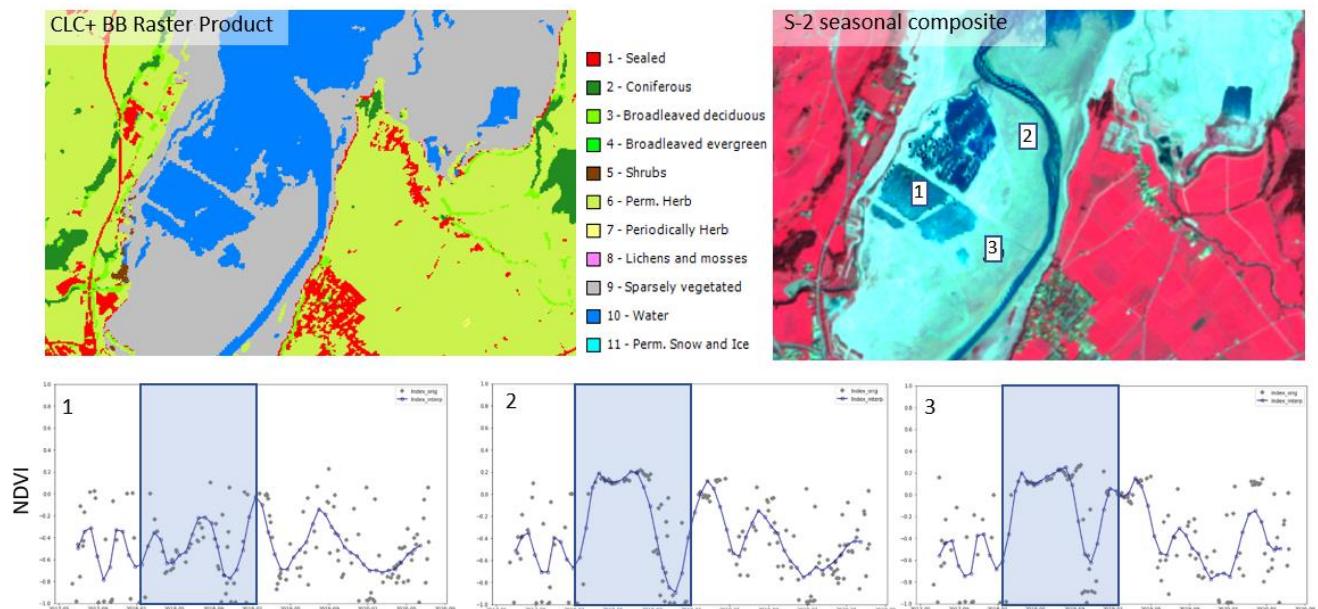


Figure 8: Example of the mapping of ephemeral classes in the CLC+ Backbone Raster Product. Water is mapped according to its spatio-temporal coverage during the reference year 2018. In the given example, areas which are more than 50% of the observations covered by water (i.e. NDVI typically below 0), are correctly mapped as Water. Areas with shorter water coverage show an NDVI slightly above 0 for most of the time in this example and are correctly mapped as Sparsely- and Non-vegetated.

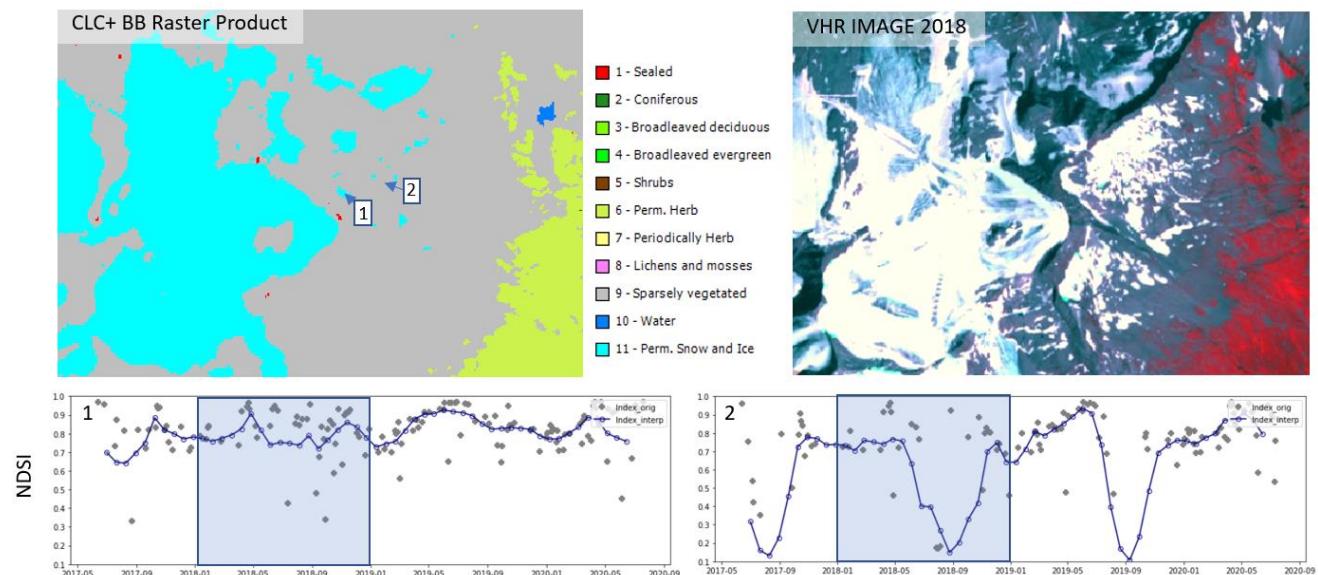


Figure 9: Example of the mapping of ephemeral classes in the CLC+ Backbone Raster Product. Permanent Snow and Ice is mapped according to its spatio-temporal coverage during the reference year 2018. In the given example, areas which are more than 90% of the observations covered by snow or ice (i.e. NDSI typically above 0.41), are correctly mapped as Snow and Ice. Adjacent areas with shorter snow or ice coverage are correctly mapped as Sparsely- and Non-vegetated.

Example 6 (Figure 10) illustrates the mapping of some agricultural parcels in the Mediterranean, correctly as Non- and Sparsely Vegetated, if they have not been cultivated in the main growing season 2018 (i.e. October 2017 to July 2018 in the Mediterranean). The contrast is in particular clear, when contrasting with the temporal trajectory of neighbouring parcels, which were cultivated in the growing season (Figure 10).

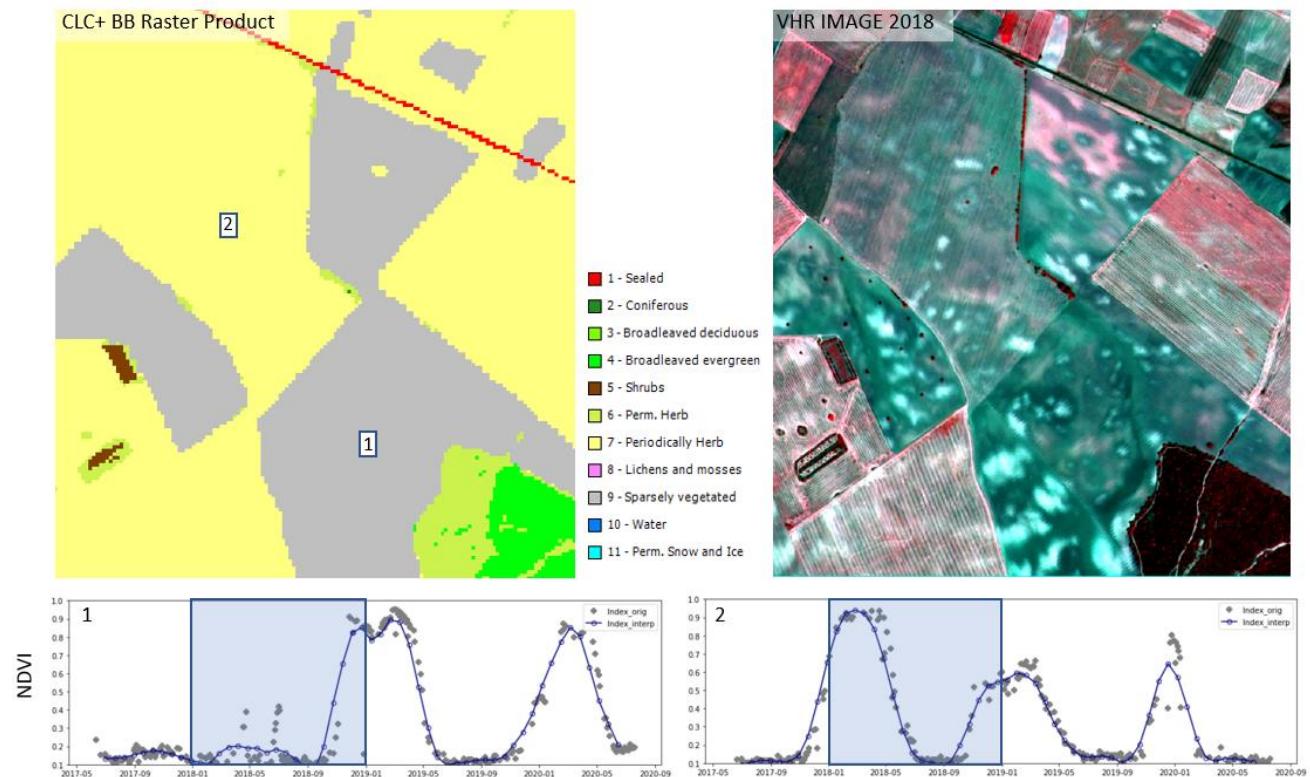


Figure 10: Example of the mapping of temporarily barren agricultural fields in CLC+ Backbone Raster Product in Spain. Parcels that are cultivated during the main growing season (Mediterranean winter rain season) show a strong NDVI peak in early 2018. In contrast, parcels which are not cultivated during the vegetation season 2018 show a very faint NDVI signal until October of the reference year (i.e. actually the beginning of the cultivation season 2019) and are hence correctly mapped as Non- and Sparsely Vegetated.

Example 7 (Figure 11) illustrates the mapping of dry grasslands (widespread in Southern Europe and Turkey) as Permanent Herbaceous. Though the vegetation signal is rather subtle, when compared to Permanent Herbaceous, for example in Central Europe, it is still sufficient to correctly distinguish such areas from Non- and Sparsely vegetated areas in those regions.

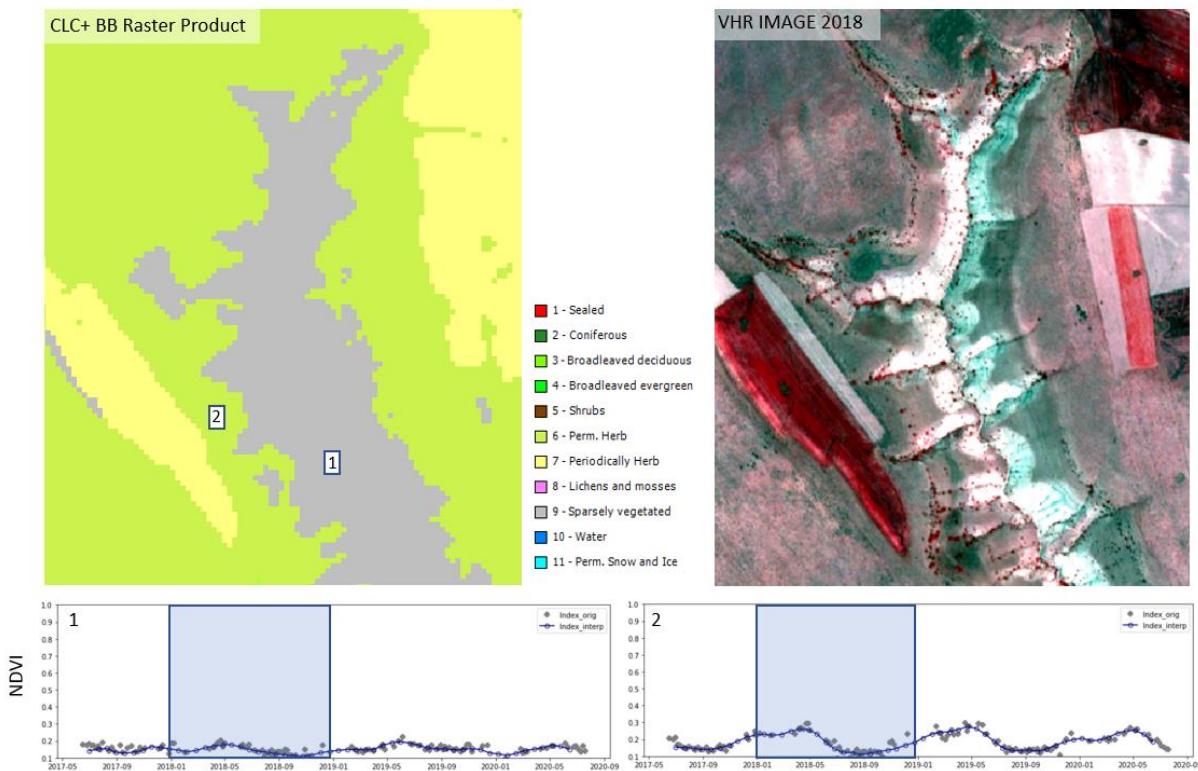


Figure 11: Example of the mapping of dry grasslands as Permanent Herbaceous in CLC+ Backbone Raster Product in Turkey. While in Central European and Northern Europe, permanently herbaceous areas are typically characterized by a strong NDVI peak during the vegetation season, dry grasslands in Southern Europe and Turkey are characterized by far fainter vegetation signals.

Example 8 (Figure 12) illustrates the mapping of peat extraction areas in Ireland as class 9, “Non- and Sparsely Vegetated”. Other than natural unexploited peat bogs, active peat extraction sites do typically not have noticeable vegetation cover and are thus correctly classified as Non- and Sparsely Vegetated. In order to turn such areas from CO₂ sources back into sinks, more and more peat production areas are reinstated and partially covered by vegetation again, as shown in Figure 13. Where vegetation cover (mostly shrub and permanent herbaceous) is dense enough, it is classified accordingly as vegetation cover in the raster product.

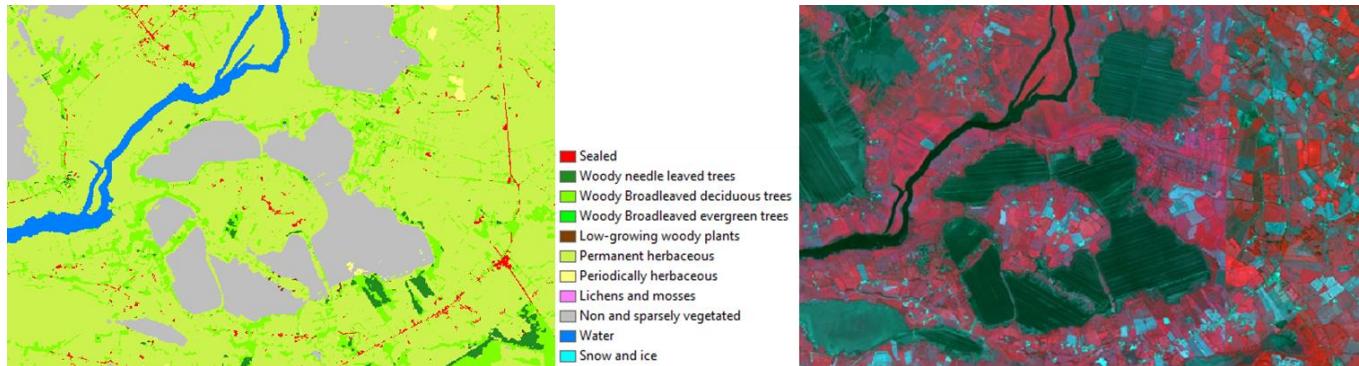


Figure 12: Example of peat extraction areas mapped as Sparsely Vegetated. CLC+ Backbone raster product (left) and false colour VHR Image 2018 (right).



Figure 13: Peat production area in the process of renaturation, already showing dense vegetation cover in large parts.

3.2.2 Production Methodology and Workflow

This section provides a high-level overview of the production methodology applied for the CLC+ Backbone raster product classification, to allow product users better understand certain characteristics, strengths and limitations of the product, beyond the pure technical specifications as outlined in the previous sections.

3.2.2.1 Input data

The main input data sources for the raster classification are Sentinel-2 time-series covering the period 2017-07-01 till 2019-06-31. Initially all scenes at a processing level of L2A and with a cloud coverage lower than 80% are retrieved. Additional cloud-masks are computed using the FMask algorithm version 4.1 (Qiu et al. 2019). The EO data is re-projected to Lambert azimuthal equal-area projection (EPSG 3035) and for each pixel all valid observations are taken into account after masking out clouds and cloud shadows to interpolate an equidistant time-series with a total of 72 time-steps. This includes all S-2 bands except B10 and several spectral indices such as NDVI, NDWI or NBR. Optional input data to support the classification in particularly difficult areas (Sentinel-1 time-series, mono-temporal auxiliary features extracted from DEMs, distance features derived from existing LC/LU maps) were extensively tested at several test sites during the early production phase. These tests demonstrated, that the

inclusion of such additional features did in most cases not lead to any significant accuracy gains. Considering also the overhead in terms of implementation, I/O and computational footprints, such additional data sources were finally not integrated. In particularly difficult areas, additional sampling and post-processing strategies were found more efficient, to address remaining issues in the classification results.

The training / validation and test data required for the model calibration are compiled from various sources, such as from adjusted and filtered LUCAS (EUROSTAT 2018) data of 2018, from stratified automated LC class annotations based on existing land use/land cover maps, as well as from additional visual sample point photo-interpretation from VHR imagery, NDVI time series and auxiliary datasets. The latter comprise, amongst others, national LC datasets, aerial imagery, or LiDAR data collected by a European network of involved regional experts. CLMS HR-S&I was also considered to improve the sampling and classification for the class Permanent Snow & Ice, but was finally not integrated due to deviating product specifications, that would have rather degraded the classification results (e.g. the HR-S&I Permanent Snow Area does not consider ice, and sets a threshold of 95% permanence instead of 90% defined for the CLC+ Backbone Raster Product).

3.2.2.2 Time series classification

Given the heterogeneity of the addressed European landscapes, all classifier training, testing and, finally, LC classification, is performed along substrata based on biogeographical regions (Metzger et al. 2013) and existing LC layers. The AOI is subdivided in 232 of these substrata (i.e. Production Units, Figure 14). The regional calibration is performed iteratively using initially all readily collected samples for a specific Production Unit to train and test a first model (i.e. 20% of the samples are withheld from training for testing). Map previews are generated for a subset of 15-25 10x10km tiles which are distributed representatively across Production Unit. The previews are checked manually and additional samples are allocated in areas where the classification quality is not yet satisfactory. Subsequently the model is trained and tested again; this process is repeated until both the accuracy metrics on the test set as well as map previews indicate that the required target accuracies are reached.

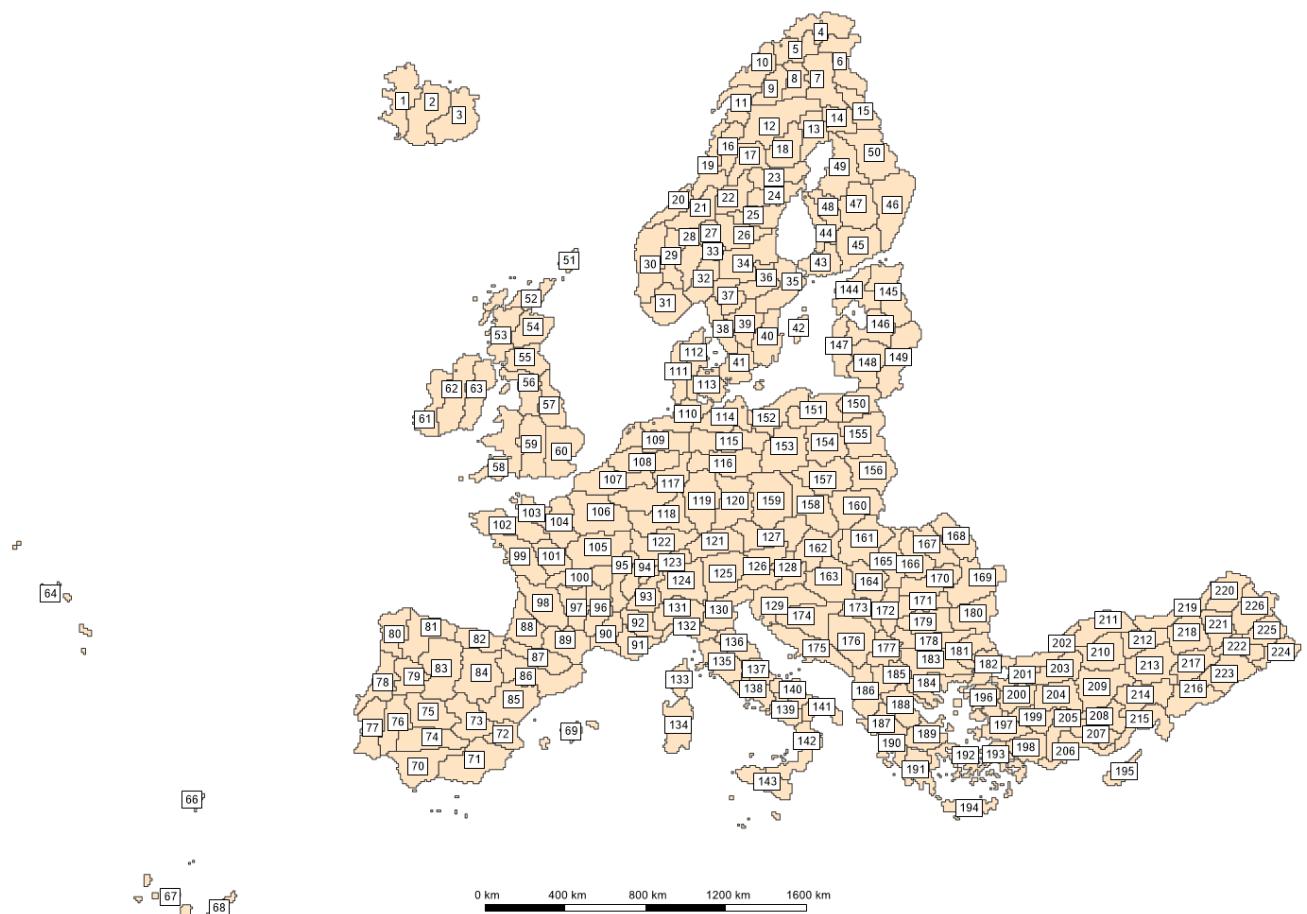


Figure 14: Production Units (PUs) subdividing the EEA-38 + UK area into homogenous biogeographical sub-strata for the Raster Product.

3.2.2.3 Post-processing

Post-processing steps comprise 1) bilateral filtering of the class probabilities, 2) blending of the probabilities along Production Unit borders and 3) adjustments of decision thresholds based on auxiliary LC information.

A bilateral filter is applied to reduce the salt and pepper noise in the classification while preserving edges. This type of filter tends to provide a good compromise between computational complexity and accuracy improvements (Schindler 2012). A small window size and low standard deviation was used to parametrize the filter in a way that favours the preservation of small details while still reduce label noise significantly.

The training data used for model calibration of neighbouring Production Units typically has some overlap to assure consistency in the classification of neighbouring units. Nevertheless, differences in the model calibration could still lead to undesired edge effects at the borders of Production Units. To circumvent this issue, each Production Units is initially produced with a buffer of 5km, which assures an overlap of 10km among neighbouring units. Within this overlap, a distance weighted averaging of the classification probabilities is performed, so that probabilities from both models have an equal

weight at the centre of the overlap area and increasing / decreasing weights respectively as a function of the distance to the centre of the overlap.

Finally, the probabilities and decision thresholds can be adjusted in areas where appropriate auxiliary datasets are available, to further reduce potential omission and commission errors among classes, which are spectrally and temporally very similar, or omission of small structures below or close the geometric precision of the multi-temporal co-registration of the Sentinel-2 time-series (e.g. small roads, trees in urban contexts). An overview of the usage of different auxiliary datasets is provided in Table 1. The order in the table presents the hierarchy of the applied rules; in cases of conflicts among the auxiliary input layers, the more reliable input (i.e. matching better the CLC+ Backbone product definitions and/or large spatial scale), was given priority. A number of further rules have been implemented to conservatively address some local phenomena (e.g. reduction of omission of urban trees based on Urban Atlas Street Tree Layer, reduction of commission of Water on artificial sports fields in urban areas). All rules remain connected to the original input probabilities, in order to avoid discontinuities in the final map product. The following Table 1 provides an overview of the used input data and the post-processing purpose. The complete and detailed ruleset is documented and implemented as Python code and included in the final delivery of the CLC+ Backbone System. All pixels affected by changes during the post-processing are documented in the Raster Post-processing Layer (section 3.4.3).

Table 1: Overview of usage of auxiliary datasets during the post-processing routine.

Auxiliary dataset	Purpose
OSM: Open Street Map roads, parking lots, runways and building footprints IMD: Imperviousness Density High Resolution Layer 2018 CopLocal: Sealed areas in Copernicus Local Component maps (UA 2018, CZ 2018, RZ 2018, N2K 2018)	<ul style="list-style-type: none"> Reduction of omission of Sealed areas in particular for small features and in mixed areas Reduction of commission of Non- and sparsely vegetated in urban areas Reduction of commission of Sealed in areas with little or no vegetation cover Reduction of omission of Sealed for small roads
DLT: Dominant Leaf Type High Resolution Layer 2018 TCD: Tree Cover Density High Resolution Layer 2018 CopLocal: Forests in Copernicus Local Component maps (UA 2018, CZ 2018, RZ 2018, N2K 2018) CLC: CORINE Land Cover 2018 (class 223: Olive Groves)	<ul style="list-style-type: none"> Reduce omission and commission of all three tree cover classes where auxiliary layers indicate / do not indicate tree cover Same factors for all three tree cover classes to leave distinction among them to the classification algorithm No reduction of the probability for pixel predicted as class 4 since auxiliary data is sparser there Extra increase of the probability for Olive Groves (class 4) if indicated in auxiliary layers to reduce omission of less dense canopies of Broadleaved Evergreen
LPIS: National datasets (where available) of the Land Parcel Information System and complemented by National Land Cover maps (mainly vineyards and other shrubby crops) CLC: CORINE Land Cover 2018 (Vineyards, 221)	<ul style="list-style-type: none"> Reduce omission of Low-growing woody plants and commission of Perm. and Period. Herbaceous in vineyards If indicated by high-quality auxiliary data reduce omission of Low-growing woody plants and commission of Perm. and Period. Herbaceous even further
GRA: Grassland High Resolution Layer 2018	<ul style="list-style-type: none"> Reduce omission of Perm. Herb. and Period. Herb. where auxiliary layers indicate their presence Reduce commission of Perm. Herb. and Period. Herb. on narrow roads

<p>CopLocal: Grassland, crop land, and sparsely vegetated classes in Copernicus Local Component maps (UA 2018, CZ 2018, RZ 2018, N2K 2018)</p> <p>CLC: Grassland, crop land, and sparsely vegetated classes in CORINE Land Cover 2018 (211, 212, 213, 231, 321, 331, 332, 333, 334)</p> <p>OSM: Open Street Map roads, parking lots, runways and building footprints</p>	<ul style="list-style-type: none"> Reduce commission of Perm. Herb. and Period. Herb. where auxiliary layers indicate sparsely vegetated areas Apply the same factors for both classes to leave distinction among them to the classification algorithm
<p>CopLocal: Sparsely vegetated classes in Copernicus Local Component maps (UA 2018, CZ 2018, RZ 2018, N2K 2018)</p> <p>CLC: Sparsely vegetated classes in CORINE Land Cover 2018 (331, 332, 333, 334)</p> <p>OSM: Open Street Map roads, parking lots, runways and building footprints</p> <p>IMD: Imperviousness Density High Resolution Layer 2018</p> <p>CopLocal: Sealed areas in Copernicus Local Component maps (UA 2018, CZ 2018, RZ 2018, N2K 2018)</p>	<ul style="list-style-type: none"> Reduce omission of Non-and Sparsely Vegetated where the auxiliary suggests sparse vegetation Reduce commission Non-and Sparsely Vegetated where auxiliary data suggests roads
<p>WAW: High Resolution Layer 2018 Water and Wetness</p> <p>CopLocal: Water in Copernicus Local Component maps (UA 2018, CZ 2018, RZ 2018, N2K 2018)</p>	<ul style="list-style-type: none"> Reduce omission and commission Permanent Water where auxiliary layers indicate / do not indicate water
<p>NDVI90 : 90th percentile of the NDVI derived from the Sentinel-2 for the vegetation season 2018</p>	<ul style="list-style-type: none"> Limit the application of all rulesets above to areas where they corrections seems consistent with the 90th percentile derived from the Sentinel-2 time-series

3.2.2.4 Strengths and Limitations of the Applied Methodology

The strongest point of the applied methodology is a general high accuracy and robustness for most land cover classes and biogeographical regions. This is the results of several elements, including:

- The ingestion of a full time series of 2 years of Sentinel-2 (i.e. 2018 ± 0.5 years) data, which comprises very rich information on the spectral-temporal dynamics of different land cover classes, generally enables a very good separability of the 11 target classes and partially even compensates some shortcomings in the Sentinel-2 L2A input data, such as topographic over-normalization on north- to west-facing slopes.
- The usage of a state-of-the-art Deep Learning architecture, that enables to fully leverage the full time-series, without the need for feature engineering / selection, that typically leads to some loss of information.
- Enhanced cloud mask based and FMask version 4, that addressed some of the shortcomings of the default Scene Classification Layer (SCL) provided by ESA.
- A rich sample database, which have been allocated and quality-checked based on various sources.
- An iterative approach to sampling and regional calibration, which allows to address regional differences and particularly difficult cases in the classification.
- Post-processing comprising the reduction of label noise (bilateral filtering), assurance of wall-to-wall consistency (blending) and rule-based reduction of omission and commission errors.

Despite excellent quality achieved with the applied methodology, some limitations remain. It is worth mentioning, that **areas of heterogeneous land cover and small landscape features typically** have a relatively higher uncertainty, due to mixed spectral and temporal signals. The inherent uncertainty of mixed pixels at the borders between different land cover types is further amplified by remaining shortcomings of the multi-temporal co-registration of the Sentinel-2 archive (i.e. before processing Baseline 03.00 30/03/2021). As exemplified in Figure 15, the geometric uncertainties of the input data propagate into the classification with a typical uncertainty of min. 1 pixel of the exact class boundary.

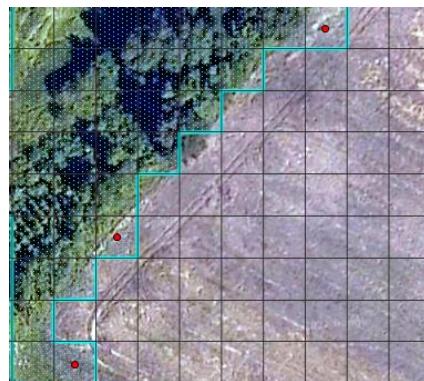


Figure 15: Border of tree covered area (cyan dots) extending into a field, resulting at least partially from geometric shifts in S-2 time-series. Red dots mark pixels where this leads to commission errors in tree cover (grid: 10m pixels in LAEA projection)

Similarly, in **landscapes with a generally strong mixture of different land cover types at the pixel level**, it remains sometimes difficult to capture the dominant type correctly for all pixels. Figure 16 illustrates the issue for an area in Turkey, where the land cover is a mix of dry herbaceous vegetation, wild olive trees, shrubs and coniferous trees, with sometimes only sparse vegetation cover. While the general mixture is also depicted in the CLC+ Backbone Raster Product, the fractions of the respective classes are subject to over- and underestimations, which are challenging to overcome, both in terms of mapping and validation. The spectral signature mix can partially lead to overestimation of shrubby vegetation, hence the olive trees are partially mapped as shrub in less dense areas, for example. Furthermore, the signal from the dry herbaceous vegetation understory is very dominant, leading to a mapping of permanent herbaceous, where the tree cover actually seems to be dense enough, at the first glance, to be captured.

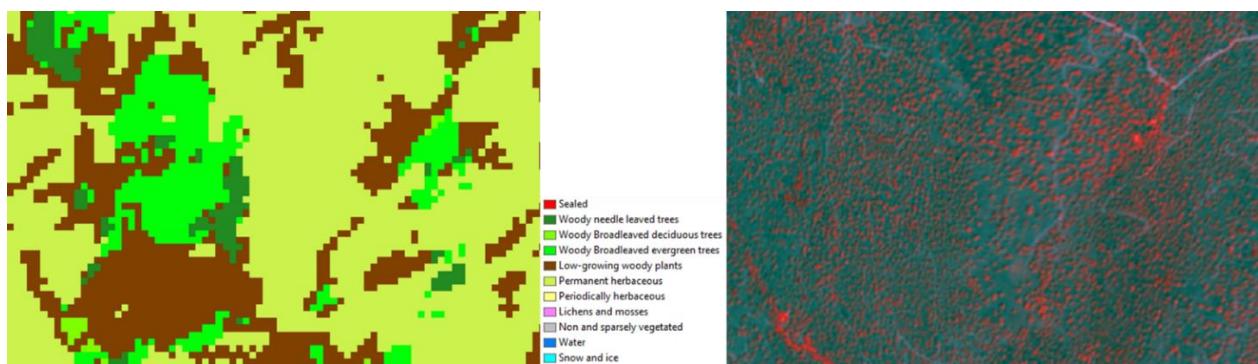


Figure 16: Mapping of a complex mix of wild olive trees, shrub, coniferous and permanent herbaceous vegetation. Left: CLC+ Backbone Raster product. Right: VHR Image 2018 false colour infrared.

Also particularly the classes **Low-growing woody vegetation** (Figure 17) and **Lichens and Mosses** (Figure 18) have some inherent uncertainty due to i) some fuzziness in the class definition, ii) limited spectral-temporal separability and/or iii) limited reference data, in particular in remote areas. All three factors cannot be easily resolved on a methodological level. Figure 17 illustrates uncertainties in the mapping of the class Low-growing woody plants (bushes, shrubs) versus Woody trees. The main defining characteristics for the class Low-growing woody plants comprise the habitus (i.e. multiple stems emerging from the ground) and the typical height (i.e. typically below 5m), both of which are criteria which are difficult to evaluate from Sentinel time series alone. While the spectral-temporal signature of shrubs allows a distinction from other land cover types to some degree (e.g. vineyards), there are many cases, where this is less obvious. These comprise in particular sparse woody canopies in Nordic countries with mixtures of shrubs and trees even at the species level, Mediterranean maquis mingled with trees or heathlands, with mixtures of dwarf shrubs and herbaceous species.

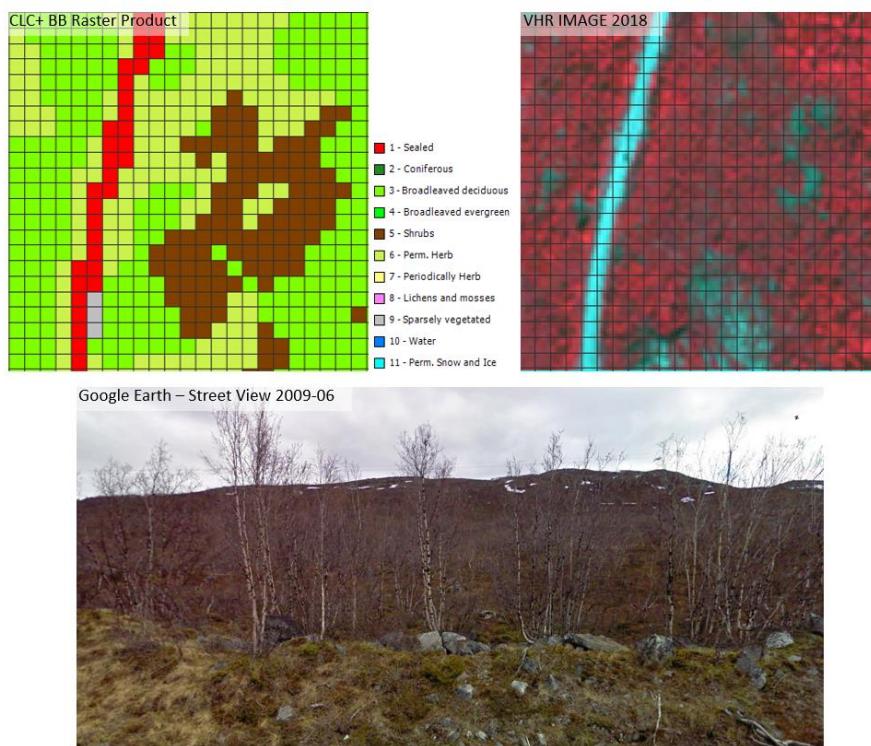


Figure 17: Example of the uncertainty in the mapping of Low-growing woody vegetation. This example from Northern Sweden illustrates that a sharp distinction of shrubs and trees is often difficult and sometimes remains fuzzy even at the level of the same species (here different specimens of Betula with sometime more tree-like, sometime more shrub-like habitus). In addition, the canopy is rather sparse and the understory comprises a mix of herbaceous species, dwarf shrubs as well as mosses and lichens.

Figure 18 illustrates the typical similarity between surfaces covered by Lichens and Mosses versus Non- and Sparsely Vegetated areas with some fractions of sparse herbaceous vegetation. Both land cover compositions occur in the same biogeographic regions and are nearly indistinguishable in VHR imagery or NDVI time-series. In areas, where street level imagery and auxiliary data sources are available, it is possible to map and verify larger occurrences of this class, however, large uncertainties remain, in particular in remote areas with no or little adequate reference data.

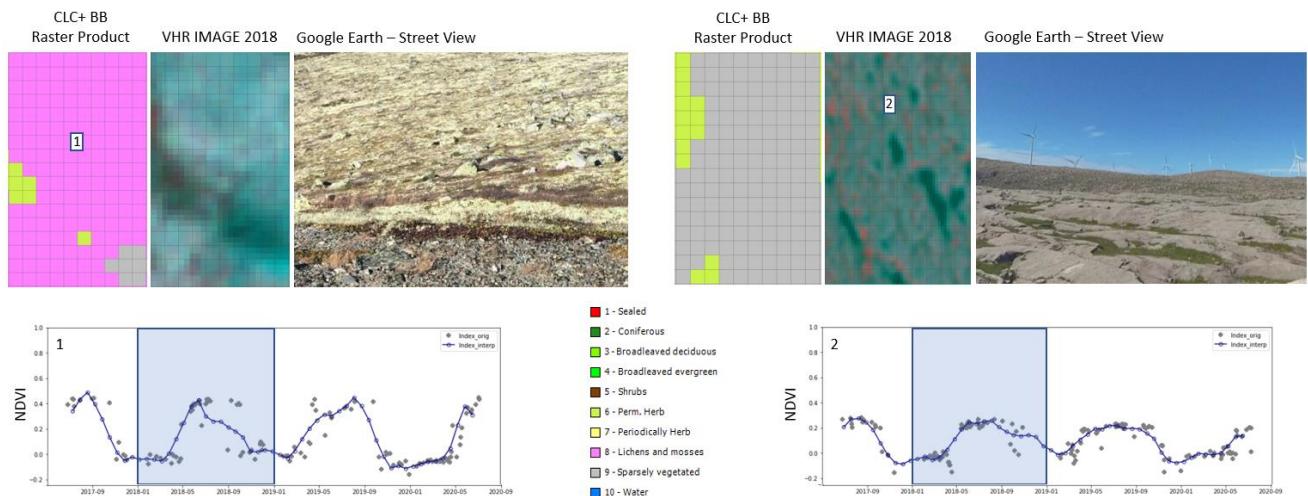


Figure 18: Example of the uncertainty in the mapping of Lichens and Mosses in the CLC+ Backbone Raster Product. These two examples from Southern Norway illustrate the typical similarity between surfaces covered by Lichens and Mosses versus Non- and Sparsely Vegetated areas with some fractions of sparse herbaceous vegetation. Both land cover compositions occur under similar biogeographic conditions and are nearly indistinguishable in VHR imagery or NDVI time-series.

Another further issue worth mentioning is the potential **confusion between the Sealed and Sparsely vegetated areas**, in particular in areas with very high reflectance (i.e. bright surfaces), where the spectral signature of both classes is very similar. As an example, Figure 19 presents an example of a greenhouse which was partially classified as a sparsely vegetated area. Dedicated post-processing steps were implemented to address such issues, where adequate reference datasets covered such areas (see Table 1) but some confusion still remains in the final product.

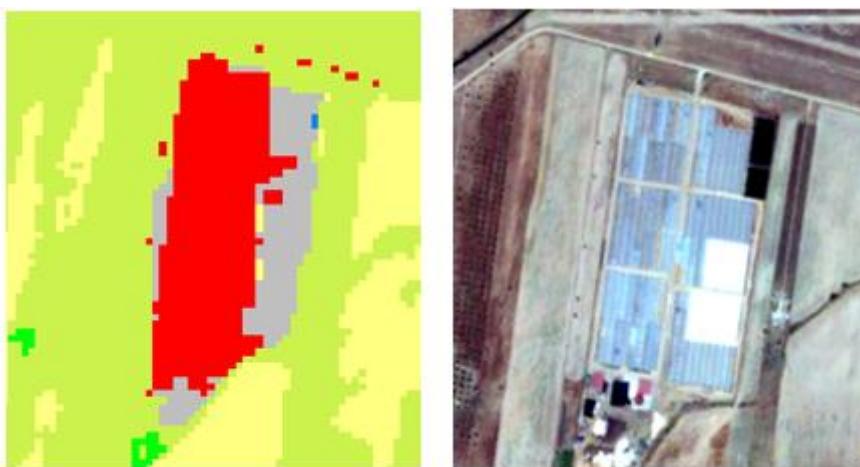


Figure 19: Greenhouse partially classified as sparsely vegetated class. Greenhouses should be classified as Sealed, but in some cases, unfavourable reflection angles may cause confusion with the Sparsely Vegetated class

Figure 20 illustrates uncertainties in the mapping of **intensively managed grassland** (e.g. fodder crops) as either Permanent Herbaceous or Periodically Herbaceous. While the exposure of bare soil within the reference years can be clearly detected in most ploughed parcels (i.e. to be mapped as Periodically Herbaceous), there can be border cases where drought events and mowing events under rather dry conditions are not fully distinguishable from a bare soil exposure.

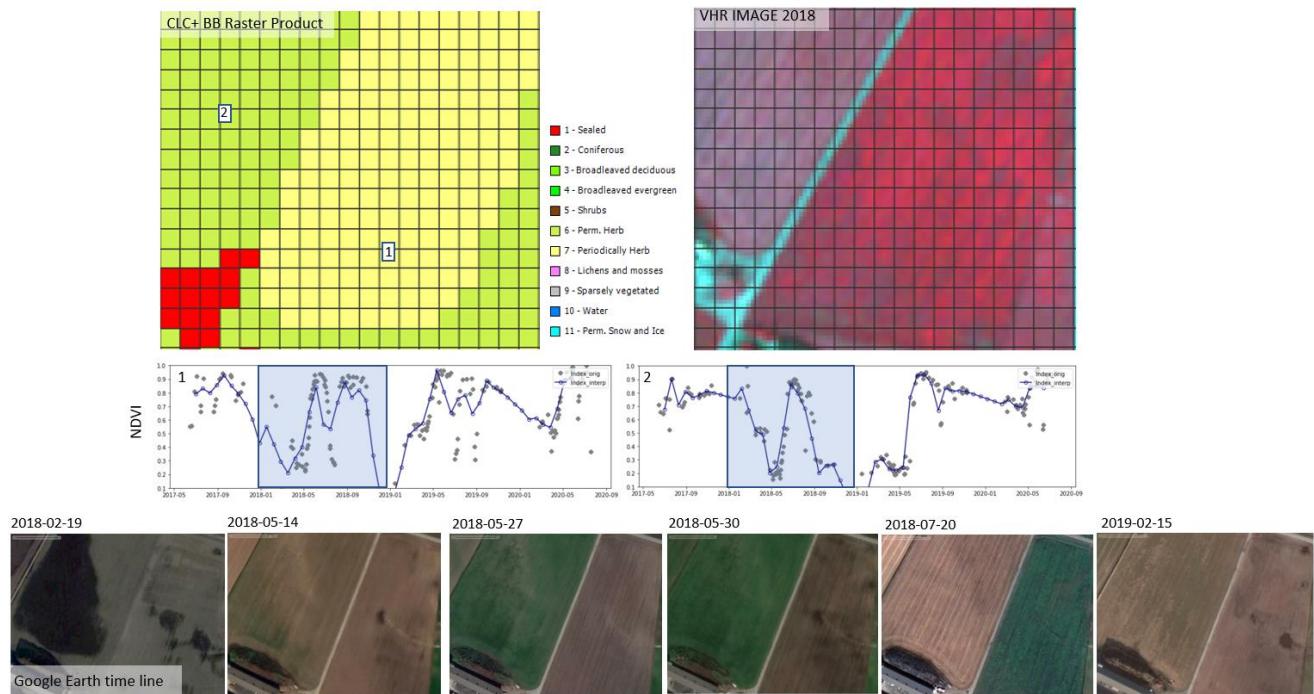


Figure 20: Example of the mapping of intensively used grassland (e.g. fodder crops). The distinction between the classes Permanent Herbaceous and Periodically Herbaceous depends on whether the bare soil is exposed (e.g. ploughed parcel on the right, mapped as Periodically Herbaceous) within the reference year or not. Especially under temporarily very dry conditions, the distinction of mowing events (i.e. no ploughing on the left parcel, mapped as Permanent Herbaceous) from actual exposure of bare soil can be very difficult.

By definition, the CLC+ Backbone land cover nomenclature does not provide a separation between **inland water vs. sea water** along a country's coastline. For potential use cases intending to separate both, a clipping with a national/European coast line dataset is recommended.

3.3 Vector Product

The CLC+ Backbone Vector Product combines the geometries of the Landscape Objects with the thematic land cover information of the raster classification for the reference year 2018. Each polygon is attributed with the dominant LC class (18 classes) and has some further attributes attached. The following sections provide information on the product specifications (section 3.3.1), the nomenclature and class definitions (section 3.3.1.1) as well as the methodology and production workflows (section 3.3.2). Additionally, section 0 provides some illustrated examples of typical cases of the Vector Product nomenclature and decision tree application. Further details on the attribution of the vector product with EO, LC and CLMS attributes can be found in section 3.3.2.6 .

3.3.1 Product Specifications

The product specifications for the CLC+ Backbone Vector product are summarized in the table below. Further product details and information on the methodology and workflows, nomenclature and attribution etc. can be found in the following chapters.

CLC+ Backbone Vector Product	Acronym	Product family
	VECTOR	CLMS_CLCplus
Summary		
CLC+ Backbone is a spatially detailed, large scale, EO-based land cover inventory. The CLC+ Backbone Vector Product combines the Landscape Object geometries and the thematic land cover information provided by the raster classification. Each polygon is attributed with its dominant Land Cover class (majority rule based 18-class LC classification), the percentage of all underlying raster LC classes in relation to the total area of the Landscape Object and an ordered list of LC class percentages of the three most dominant raster LC classes. In addition, the polygons are further enriched by Earth Observation data related attributes and CLMS product related attributes.		
Reference year		
2018		
Geometric resolution		
Approximate mapping scale of 1:10 000		
Coordinate Reference System		
European ETRS89 LAEA projection / for French DOMs WGS84 and the respective UTM zone		
Coverage		
6,002,168 km ² covering the full EEA-38 + UK		
Geometric accuracy (positioning scale)		
equals the Sentinel-2 positional accuracy in 2018 (~11m at 95.5% confidence)		
Thematic accuracy		
95 % overall accuracy, not more than 15 % omission errors and 15 % commission errors per class		
Data type		
GIS compatible Vector Format		
Minimum Mapping Unit (MMU)		
0.5 ha		

Minimum Mapping Width (MMW)				
20m				
Vector classes				
11: Very High Sealing Degree				
12: High Sealing Degree				
21: Pure needle leaved				
22: Dominantly needle leaved				
31: Pure broadleaved deciduous				
32: Pure broadleaved evergreen				
33: Dominantly broad leave				
40: Shrubland				
51: Permanent herbaceous without trees				
52: Permanent herbaceous with few trees				
53: Permanent herbaceous with many trees				
60: Periodically herbaceous				
70: Lichens and mosses				
81: Partly vegetated land - Low vegetation cover				
82: Partly vegetated land - Intermediate vegetation cover				
90: Non-vegetated land				
100: Water				
110: Snow and ice				
Attributes				
Field	Description	Type	Value(s)	NoData v
1. LC based Attributes				
Shape	Polygon	Geometry	Polygon	
ID	Landscape Object ID	Integer	1 to 1,8E308	
Site_id	Sub-Production Unit ID	Integer	4-19200	
Shape_Area	Area	Double	0,001 to 1,8E308	
Shape_Length	Length	Double	0,001 to 1,8E308	
LC_code18	Final Vector LC Class	Integer	11 to 110, see Nomenclature	254
Drcl_1	1st Dominant Raster LC class	Integer	1-11	
Drcl_2	2nd Dominant Raster LC class	Integer	1-11	
Drcl_3	3rd Dominant Raster LC class	Integer	1-11	
Drcl_1pc	Percent of 1st Dominant Raster LC class	Double	0-1	
Drcl_2pc	Percent of 2nd Dominant Raster LC class	Double	0-1	
Drcl_3pc	Percent of 3rd Dominant Raster LC class	Double	0-1	
Rcl_01pc	Cover Percent of Raster LC class 1	Double	0-1	
Rcl_02pc	Cover Percent of Raster LC class 2	Double	0-1	
Rcl_03pc	Cover Percent of Raster LC class 3	Double	0-1	
Rcl_04pc	Cover Percent of Raster LC class 4	Double	0-1	
Rcl_05pc	Cover Percent of Raster LC class 5	Double	0-1	
Rcl_06pc	Cover Percent of Raster LC class 6	Double	0-1	
Rcl_07pc	Cover Percent of Raster LC class 7	Double	0-1	

Rcl_08pc	Cover Percent of Raster LC class 8	Double	0-1		
Rcl_09pc	Cover Percent of Raster LC class 9	Double	0-1		
Rcl_10pc	Cover Percent of Raster LC class 10	Double	0-1		
Rcl_11pc	Cover Percent of Raster LC class 11	Double	0-1		
Ref_year	Reference Year	Integer	2018		
2. EO based Attributes					
S2_quality	Sentinel-2 Valid Cloud Free Observations	Integer	0-705	-32768	
3. CLMS based Attributes					
PRZ_mean	Mean percentage likelihood of membership (LOM) for Potential Riparian Zones (PRZ)	Double	0-1	255	
PRZ_perc	Area percentage of vector geometry covered by PRZ	Double	0-1	255	
RZ_1400_perc	Percentage of object belonging to Riparian Zones (RZ) class 1400	Double	0-1	-32768	
RZ_8110_perc	Percentage of object belonging to RZ class 8110	Double	0-1	-32768	
RZ_8120_perc	Percentage of object belonging to RZ class 8120	Double	0-1	-32768	
RZ_1120_perc	Percentage of object belonging to RZ class 1120	Double	0-1	-32768	
RZ_1230_perc	Percentage of object belonging to RZ class 1230	Double	0-1	-32768	
RZ_1240_perc	Percentage of object belonging to RZ class 1240	Double	0-1	-32768	
CZ_perc	Area percentage of vector geometry covered by Coastal Zones (CZ)	Double	0-1	-32768	
CZ_11200_perc	Percentage of object belonging to CZ class 11200	Double	0-1	-32768	
CZ_12300_perc	Percentage of object belonging to CZ class 12300	Double	0-1	-32768	
CZ_12400_perc	Percentage of object belonging to CZ class 12400	Double	0-1	-32768	
CZ_22100_perc	Percentage of object belonging to CZ class 22100	Double	0-1	-32768	
CZ_22200_perc	Percentage of object belonging to CZ class 22200	Double	0-1	-32768	
CZ_72100_perc	Percentage of object belonging to CZ class 72100	Double	0-1	-32768	
CZ_72200_perc	Percentage of object belonging to CZ class 72200	Double	0-1	-32768	
CZ_72300_perc	Percentage of object belonging to CZ class 72300	Double	0-1	-32768	
TCD_mean	Mean tree cover density derived from HRL2018 TCD layer [0-100%]	Double	0-1	255	
TCD_std	Standard deviation in tree cover density derived from HRL2018 TCD	Double	0 - 0,5	255	

IMD_mean	Mean imperviousness degree derived from HRL2018 IMD layer [0-100%]	Double	0-1	255	
IMD_std	Standard deviation in imperviousness degree derived from HRL2018 IMD layer	Double	0 - 0,5	255	
IMC1518_mean	Mean sealing density change between 2015 and 2018	Double	-1 - 1	255	
WAW_dry_perc	Area percentage covered by Water and Wetness (WAW) class dry	Double	0-1	255	
WAW_perm_water_perc	Area percentage covered by WAW class permanent water	Double	0-1	255	
WAW_temp_water_perc	Area percentage covered by WAW class temporary water	Double	0-1	255	
WAW_perm_wet_perc	Area percentage covered by WAW class permanent wet	Double	0-1	255	
WAW_temp_wet_perc	Area percentage covered by WAW class temporary wet	Double	0-1	255	
We_occurrence_mean	Occurrence of wetness according to WAW masks	Double	0-1	255	
Wa_occurrence_mean	Occurrence of water according to WAW masks	Double	0-1	255	
BAI_perc	Percentage of object covered by EFFIS Burnt Areas	Double	0-1	255	
BAI_Occ	EFFIS Burnt Areas - Month of Fire Occurrence	Integer	0-12	255	
UA_perc	Area percentage of object covered by Urban Atlas (UA)	Double	0-1	-32768	
UA_core_perc	Area percentage of object covered by UA core	Double	0-1	-32768	
UA_12100_perc	Percentage of object belonging to UA class 12100	Double	0-1	-32768	
UA_12300_perc	Percentage of object belonging to UA class 12300	Double	0-1	-32768	
UA_12400_perc	Percentage of object belonging to UA class 12400	Double	0-1	-32768	
UA_14100_perc	Percentage of object belonging to UA class 14100	Double	0-1	-32768	
UA_14200_perc	Percentage of object belonging to UA class 14200	Double	0-1	-32768	
Building_perc	Area percentage of buildings per landscape object	Double	0-1	255	
Building_height	Building height –90th percentile in metres for capitals core urban areas	Double	0-225	-32768	
SLP_dg_med	Slope – median in degree	Double	0-90	-9999	
EXP_c_maj	Exposition (4 Cardinal directions: North, East, South, West + Flat)	Integer	0-4	255	
MASL_min	Metres above Sea level of the land cover object minimum	Integer	-299 - 5120	-32767	
MASL_mean	Metres above Sea level of the land cover object mean	Double	-299 - 5120	-32767	

MASL_max	Metres above Sea level of the land cover object maximum	Integer	-299 - 5120	-32767
FFI_name	Forest fragmentation index (multiscale Foreground area density): category	Integer	1-6	255
FFI_num	Forest fragmentation index (multiscale foreground area density): average density per object	Double	0-1	255
MSPA_core	Area percentage of object covered by Morphological Spatial Pattern Analysis (MSPA) class core	Double	0-1	129
MSPA_isle	Area percentage of object covered by MSPA class islet	Double	0-1	129
MSPA_loop	Area percentage of object covered by MSPA class loop	Double	0-1	129
MSPA_brid	Area percentage of object covered by MSPA class bridge	Double	0-1	129
MSPA_perf	Area percentage of object covered by MSPA class perforation	Double	0-1	129
MSPA_edge	Area percentage of object covered by MSPA class edge	Double	0-1	129
MSPA_bran	Area percentage of object covered by MSPA class branch	Double	0-1	129
SOS_mean	Start of Season from HRL Phenology	Double	17086-18365	-32768
LOS_mean	Length of Season from HRL Phenology	Double	18-383	-32768
LI_mean	Large Integral from HRL Phenology	Double	1-11438	-32768

Metadata

XML metadata files according to INSPIRE metadata standards

Delivery format

Geopackage (intermediate deliveries), HDF5 (final delivery), external colour tables in *.lyr and *.lyrx format, and INSPIRE-compliant metadata in XML format

Remark: By definition, the above land cover nomenclature does not provide a separation between inland water and coastal/sea water along a country's sea border and in the respective 250m coastal buffer zone. For potential use cases intending to separate both, a clipping with a national/European coast line dataset is recommended.

3.3.1.1 Nomenclature concept and class definitions

The 18 classes of the Vector Product are assigned based on the composition of the land cover classes in the Raster Product according to the following definitions and area thresholds. Please note that the area-percentage thresholds typically refer to a subset of the land cover (e.g. > 50% of the Woody vegetation should be Shrubs to qualify as Shrubland) to resolve ambiguities in areas where no single class is dominant. Please refer to section 3.3.1.2 for a more detailed elaboration that fully covers the hierarchy of the class definition.

- 10. Built-up land
 - 11: very high sealing degree (sealed surfaces > 80 %¹)
 - 12: high sealing degree (sealed surfaces 50 - 80 %)
- 20. Woodland – needle leaved trees ²
 - 21: pure needle leaved >75 %
 - 22: dominantly needle leaved 50 – 75 %
- 30. Woodland – broadleaved trees
 - pure broadleaved > 75 %
 - 31: pure deciduous > 50 %
 - 32: pure evergreen > 50 %
 - 33: dominantly broadleaved 50 - 75 %
- 40. Shrubland > 50%
- 50. Permanent herbaceous land (i.e. grasslands)
 - 51: without trees (woody trees <= 10 %)
 - 52: with few trees (woody trees 10 - 30 %)
 - 53 with many trees (woody trees 30 - 50 %)
- 60. Periodically herbaceous land (i.e. arable land) > 50%
- 70. Lichens and mosses land > 50%
- 80. Partly vegetated land (Non-vegetated land 50 – 90%)
 - 81: low vegetation cover 10 – 30 %
 - 82: intermediate vegetation cover 30 - 50 %
- 90. Non-vegetated land (i.e. rock, screes, sand, lichen, permanent bare soil) >= 90 %
- 100. Water > 50%
- 110. Snow and ice > 50%

¹ Calculated as ratio of sealed surfaces and total polygon area (standard calculation for all threshold except forest mixture classes)

² Calculated as ratio of needle leaved and total forest area within polygon (calculation method only for forest mixture classes)

3.3.1.2 Decision tree for area thresholds

In areas where no single class reaches absolute dominance, a class assignment based on the relative dominance of the respective classes is used. This is implemented through a hierarchical decision tree which is represented in Figure 21. For example, the classes 11 and 12 are assigned only if Water is not dominant, if within the biotic and abiotic land cover components the abiotic once have a relative dominance and if within the abiotic land cover components Sealed has a relative dominance.

Furthermore, there can be cases where several classes can reach exactly equal area percentages within the polygon. To resolve such ties the following priority listing for Raster Product classes is implemented:

1. Snow and Ice
2. Water
3. Sealed
4. Woody Broadleaved Evergreen
5. Woody Broadleaved Deciduous
6. Woody Needle leaved
7. Low-growing woody plants (bushes, shrubs)
8. Permanent Herbaceous
9. Periodically Herbaceous
10. Lichens and mosses
11. Non- and sparsely vegetated

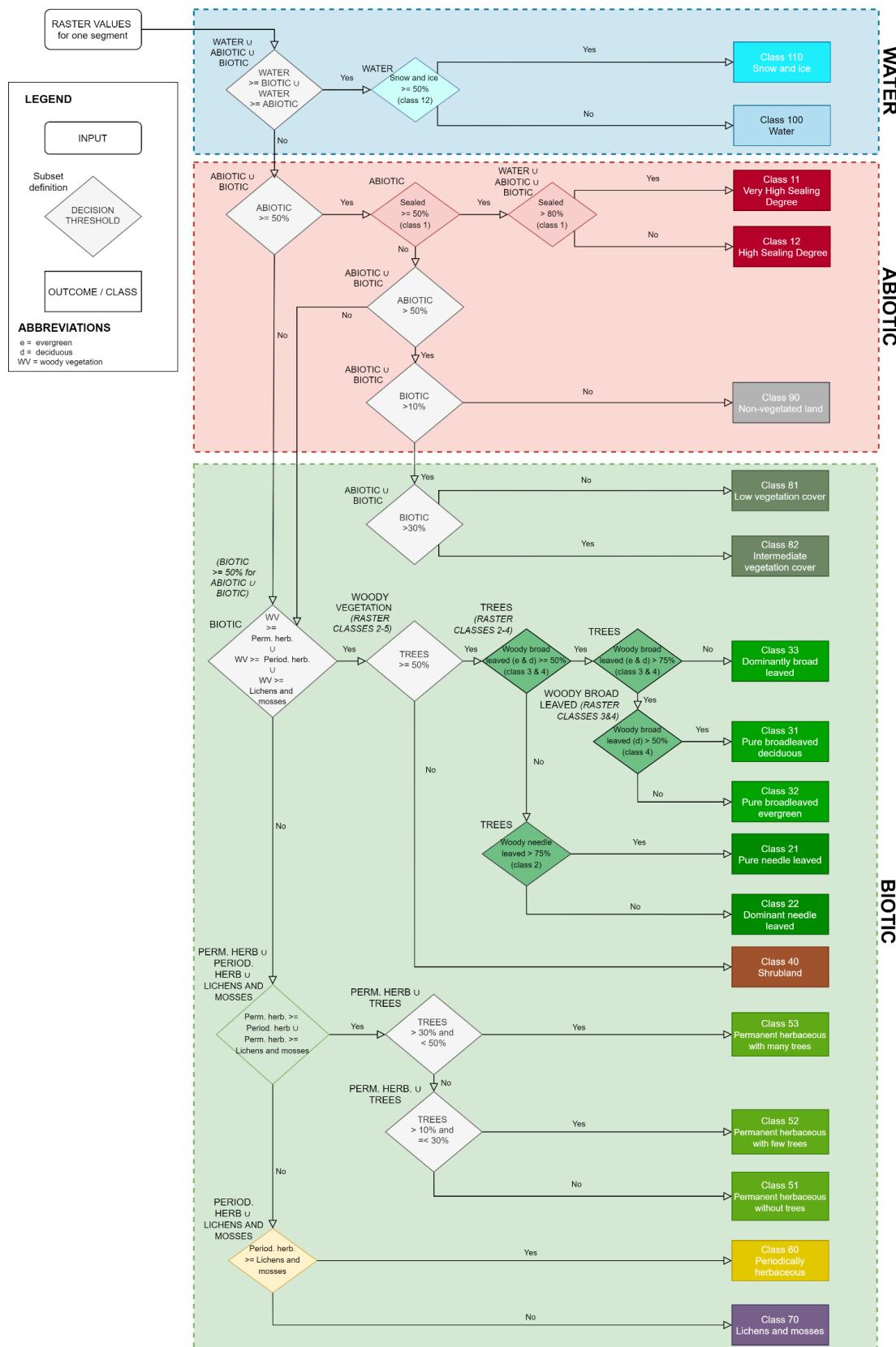


Figure 21: Decision tree for the usage of area thresholds for polygons with (pure or) spatially mixed land cover in the Vector Product.

3.3.1.3 Examples

Table 2 provides a series of examples for the application of the Vector Product nomenclature and decision tree for different land cover compositions.

Table 2: Examples for interpretation of Vector Product nomenclature and decision tree for different land cover compositions

Example (Raster Product class percentages per Landscape Object)	Resulting Vector Product class / Explanation
<ul style="list-style-type: none"> • Example 1: 33 % Permanent Herbaceous + 33 % Periodically Herbaceous + 33% Low-growing woody plants 	<ul style="list-style-type: none"> • <u>Shrubland</u> ➢ Higher priority for Low-growing woody plants
<ul style="list-style-type: none"> • Example 2: 50% Woody Needle leaved + 50% Woody Broadleaved Deciduous 	<ul style="list-style-type: none"> • Dominantly Broadleaved ➢ Higher priority for broadleaved
<ul style="list-style-type: none"> • Example 3: 40 % Permanent Herbaceous + 30 % Water + 30% Sealed 	<ul style="list-style-type: none"> • Permanent Herbaceous land ➢ Biotic has relative dominance (30% Water, 30% Abiotic, 40% Biotic) and Permanent Herbaceous is dominant in Biotic
<ul style="list-style-type: none"> • Example 4: 40 % Herbaceous Permanent + 30 % Non- and sparsely vegetated + 30% Sealed 	<ul style="list-style-type: none"> • High Sealing Degree ➢ Abiotic has relative dominance (60% vs. 40%) → Only consider abiotic classes: Sealed vs. Non- and sparsely vegetated → Sealed has a higher priority than Non- and sparsely vegetated
<ul style="list-style-type: none"> • Example 5: 40% Permanent Herbaceous + 30% Woody Broadleaved Deciduous + 30 Non- and sparsely vegetated 	<ul style="list-style-type: none"> • Permanent Herbaceous with many trees ➢ Biotic has relative dominance → Inside Biotic there are 57% Permanent Herbaceous & 43% Woody Broadleaved Deciduous
<ul style="list-style-type: none"> • Example 6: 40 % Non- and sparsely vegetated + 30% Sealed + 30% Permanent Herbaceous 	<ul style="list-style-type: none"> • Partly vegetated land - Intermediate vegetation cover
<ul style="list-style-type: none"> • Example 7: 40 % Non- and sparsely vegetated + 30% Sealed + 30% Woody Needle leaved 	<ul style="list-style-type: none"> • Partly vegetated land - Intermediate vegetation cover
<ul style="list-style-type: none"> • Example 8: 48% Sealed + 29% Woody Needle leaved + 23% Woody Broadleaved Deciduous 	<ul style="list-style-type: none"> • Dominantly needle leaved ➢ Biotic is dominant and Woody Needle Leaved fraction is higher than Woody Broadleaved Deciduous
<ul style="list-style-type: none"> • Example 9: 80% Non- and sparsely vegetated + 5% Herbaceous Permanent + 15% Woody Needle leaved 	<ul style="list-style-type: none"> • Partly vegetated land - Low vegetation cover ➢ Biotic part is only 20%
<ul style="list-style-type: none"> • Example 10: 30% Sealed + 30% Water + 40% Non- and sparsely vegetated 	<ul style="list-style-type: none"> • Non-vegetated land ➢ Abiotic dominant (70% Abiotic and 30% Water)
<ul style="list-style-type: none"> • Example 11: 29% Woody Needle leaved + 31% Permanent herbaceous + 40% Non- and sparsely vegetated 	<ul style="list-style-type: none"> • Permanent herbaceous with many trees ➢ Biotic dominant → Permanent herbaceous dominant inside Biotic

The following examples (Figure 22, Figure 23) illustrate typical visual appearances of the CLC+ Backbone Vector Product in different biogeographic regions.

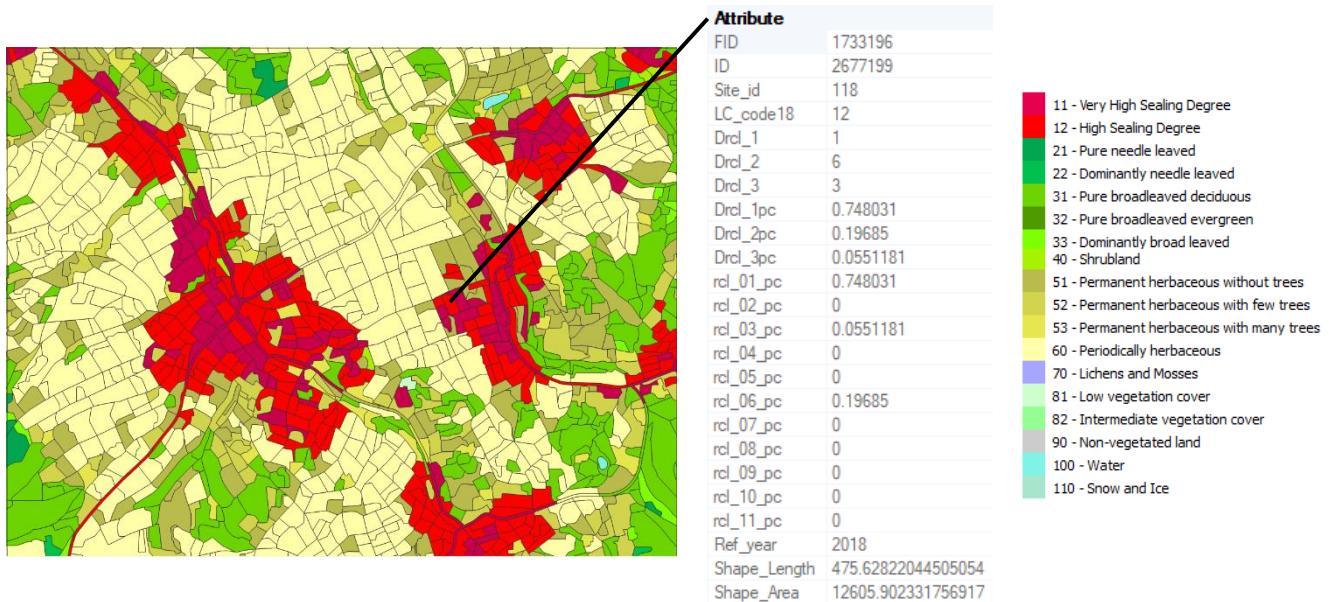


Figure 22: Example of Vector Product in part of Luxembourg

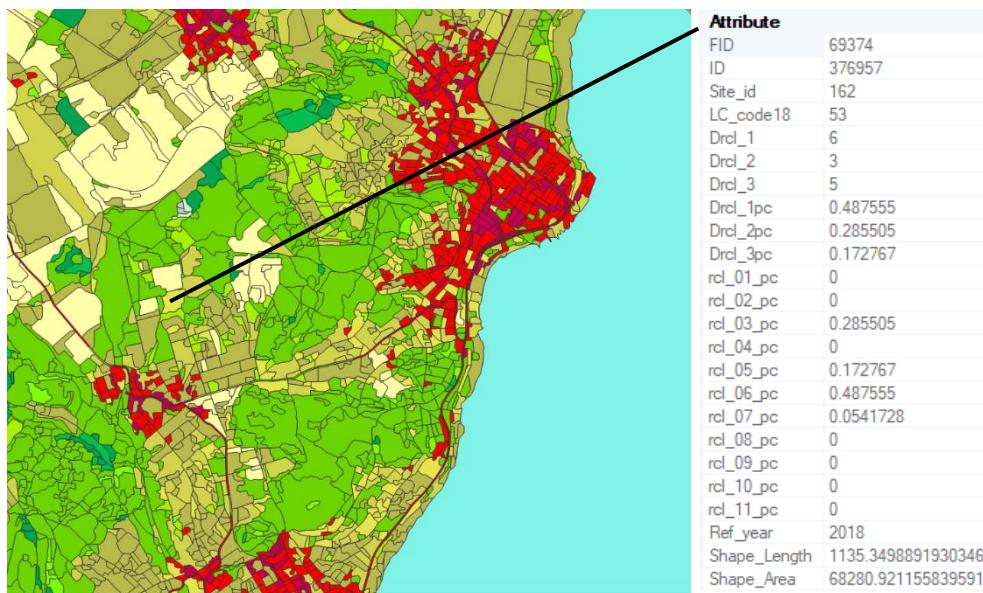


Figure 23: Example of Vector Product in part of Hungary

3.3.2 Production Methodology and Workflow

This section provides a high-level overview of the production methodology applied for the CLC+ Backbone Vector Product creation, to allow product users better understand certain characteristics, strengths and limitations of the product, beyond the pure technical specifications as outlined in the previous sections.

3.3.2.1 Definition of Landscape Objects

Landscape objects are supposedly stable and unique segments which represent spectrally and/or thematically homogeneous features throughout a year, having minimum size (MMU; 0.5 ha) and width (MMW; 20m) bounds. For example, single agricultural parcels, different type of tree stands or building blocks are each landscape object since each partition can be delineated by its spectral characteristics. Landscape objects are generated by combining two levels of objects - level 1 and level 2 objects - so called "Hardbone" and "Softbone", respectively. Hardbone represents a geometric skeleton of persistent linear landscape objects such as transportation network (i.e. roads, railways) and hydrological waterways (i.e. rivers, lakes, canals), which comprises of various line vector data. Softbone further delineates landscape objects in a controlled automatic manner based on the image segmentation technique, which generates landscape object polygons by their spectral response and variation throughout a year within the frame of Hardbone.

Given the heterogeneity of the European landscapes, all production steps are performed along substrata based on biogeographical regions (Metzger et al. 2013). The area is subdivided into 267 substrata (i.e. Secondary Production Units), as displayed in Figure 24. The SPUs used for the vector production are sub-units of the bigger PUs (as used for the raster production), to enable more performant vector data processing. All units are based on the same EEA reference grid.

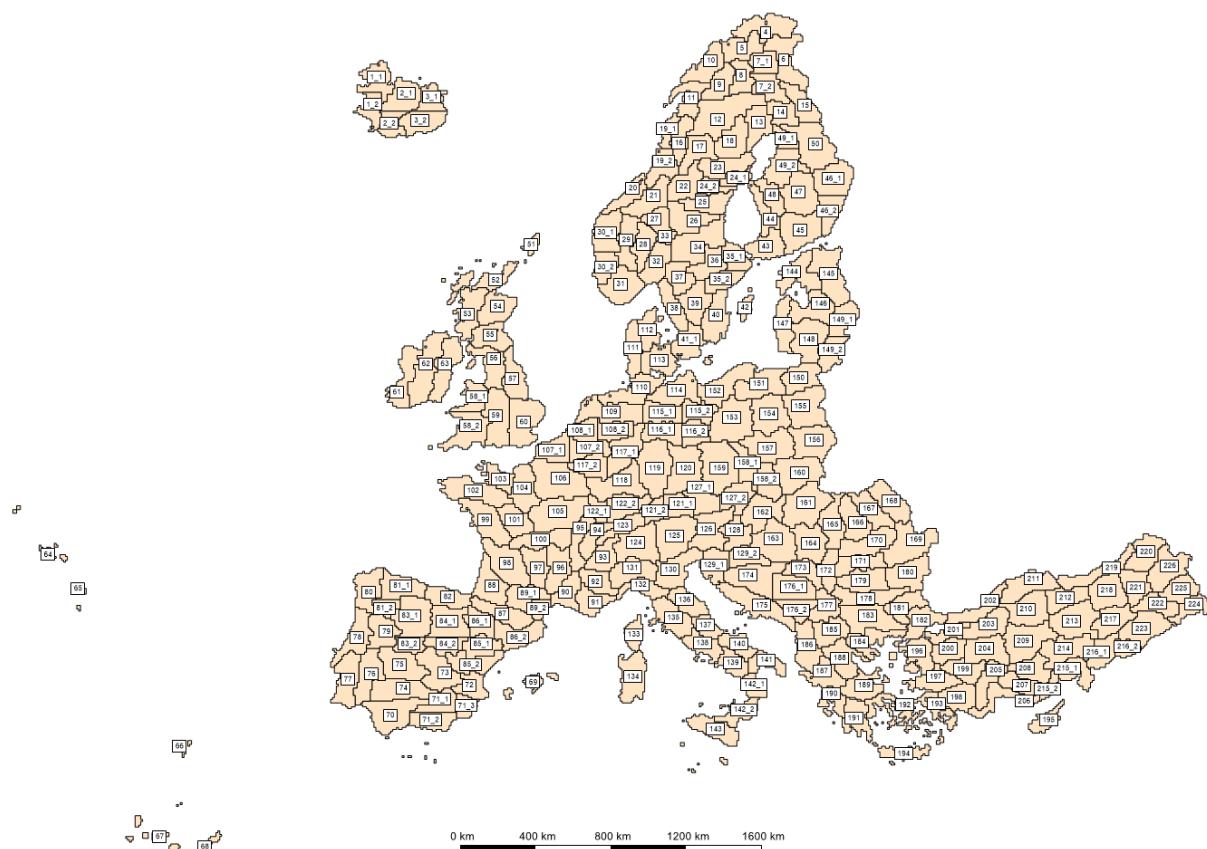


Figure 24: Secondary Production Units (SPU) subdividing the EEA-38 + UK area into homogenous biogeographical sub-strata for production of Vector Product

3.3.2.2 Hardbone generation

3.3.2.2.1 Input data

Table 3 shows the overview of the input data for Hardbone generation, being used in different combinations per Production Units, dependent on the respective data availability and density. The freely available in-situ data are downloaded from respective sources and only line vector layers are considered as an input to compose Hardbone. The primary source of the transportation network are OSM data, which contain categorized road and railway features. To complement the transportation network, road data derived by Deep Learning techniques from the VHR_2018 dataset as well as national/regional road data (and ©Facebook road data) are used. For the hydrological network, mainly EU-Hydro and/or WISE data are used, supplemented by national/regional hydro data.

Table 3: Overview of potential input datasets for Hardbone production (region specific selection applied).

Transportation Network	Hydrological Network
<ul style="list-style-type: none"> • Open Street Map (OSM) roads and railways • National/regional road data • Deep-learning-derived road data • Facebook AI derived road data 	<ul style="list-style-type: none"> • EU-Hydro river and canal network (2019) • Water Information System for Europe (WISE) – surface waterbody lines • National/regional hydro data

3.3.2.2.2 Workflow

Hardbone generation starts with the pre-evaluation of the available input datasets. The evaluation is performed for each Secondary Production Unit (SPU) to gain the insight of the data completeness, geometric accuracy and the usability. During the assessment, input data layers and a hierarchy of the datasets are determined for a single SPU, and the categories of the OSM roads are analysed and selected for a buffer that is applied for major roads (i.e. highways, motorways, see Table 4).

Each input data is re-projected to LAEA projection (EPSG 3035), pre-processed (i.e. dangle line and tunnel removal, railway cleaning, MMW and MMU cleaning, short segments cleaning, buffering) and combined together, resulting in a cleaned Hardbone line data. The hierarchy of the data is considered when assembling different input data, especially for the cleaning process of MMW and MMU. For MMW, a tolerance of 200m is applied for the connectivity of transportation and hydrological network (Figure 25). Water mask derived from raster product is used to remove the centrelines along the large waterbodies (i.e. large lakes or wide rivers) to avoid having artificial segments inside of homogeneous landscape object. The last step is the final cleaning of MMU, MMW and short segments and feature snapping for the seamless connection to the neighbouring Production Units. Final Hardbone lines after the quality check are used as an input for subsequent processes.

Table 4: Selected OSM code and category for the application of a buffer.

OSM code	OSM category	Applied buffer [m]
5111	Motorway	15
5112	Trunk	10
5113	Primary	10
5114	Secondary	10

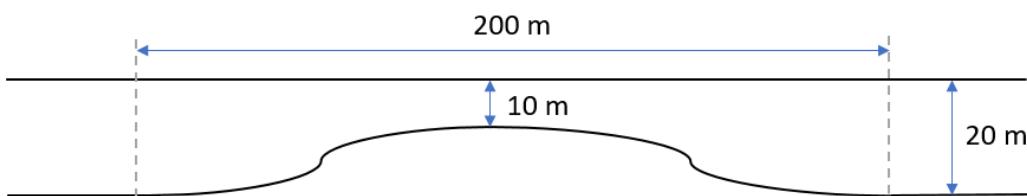


Figure 25: MMW tolerance for the connectivity.

3.3.2.3 Softbone segmentation

3.3.2.3.1 Input data

The two main input data for the Softbone segmentation are i) Hardbone and ii) Sentinel-2 temporal statistics (hereinafter referred to as “Time Features”).

i) Hardbone

The Hardbone line data produced in the previous step are vectorised and subsequently rasterized for the Softbone segmentation. Dangle lines that do not form a closed polygon are incorporated in the Hardbone raster as well to provide additional “guidance” for the segmentation. The rasterized Hardbone is a mandatory layer, which acts as a spatial constraint in the Softbone segmentation process, in that the Softbone segments cannot be built across the Hardbone geometries.

ii) Sentinel-2 Time Features (TFs)

Sentinel-2 Time Features (TFs) are derived from Sentinel-2 L2A data covering the period from 2018-04-01 to 2018-10-31. Initially all scenes at a processing level of L2A with a cloud coverage lower than 80% are retrieved. Additional cloud masks are computed using the FMask algorithm version 4.1. (Qiu et al. 2019) to improve the insufficient cloud masks available from Sentinel-2 L2A data as a standard. These EO data are then re-projected to ETRS89_LAEA projection (EPSG 3035) and for each pixel, all valid observations after masking out clouds and cloud shadows are considered for the TF computation. The parameters for TF calculation, including time window and cloud coverage, can be adapted according to the varying conditions in different bio-geographical regions across the production area (EEA-38 + UK). Possible TFs are standard deviation and percentiles (i.e. 10th, 25th, 50th, 75th and 90th) computed from all S-2 bands except B10 and spectral indices such as NDVI, NDWI, NBR or ARVI. Additionally, probability layers coming from the Raster Product are used if required to support final aggregation steps in the segmentation.

3.3.2.3.2 Workflow

The Softbone Segmentation includes all processing steps relevant to delineate stable Landscape Objects using the above described TFs (reference year 2018) and the Hardbone per SPU (based on the environmental zones according to Metzger et al. (2013). The main steps comprise:

- Selection of best suitable EO data acquisition window and parameters per bio-geographical region
- Identification of suitable Time Features for the respective bio-geographical regions, i.e. SPU
- Iterative semi-automatic Softbone Segmentation based on TFs and Hardbone
- Automatic border harmonization (between 10km processing cells and SPUs)
- Retransformation (smoothing, reintroduction of HB geometries where needed, snapping, MMU cleaning, gaps and overlaps check)
- Quality Assurance

The Softbone segmentation has been performed along a modern graph-based image partitioning approach (Bosilj et al. 2018). Together with the Hardbone, the most suitable 3 to 7 features are selected among the available S-2 TFs and fed to the algorithm to obtain the initial segments based on the spectral information. Initial segments are a combined result of a pixel based Hierarchical Tree segmentation algorithm (Alpha-tree) (Havel et al. 2019) and a relative position of segments which defines a scale space of the tree “Omega”. Several pre-processing options (i.e. bilateral filter, gamma filter and histogram equalization) are optionally applied to the input data to reduce noise (e.g. “salt and pepper” effect) coming from the EO data. The calibration of parameter sets, so called “deck building”, is performed firstly in a subset (5~20 patches of EEA 10km x 10km grid) which represents the landscape objects across the SPU (Figure 24). Sometimes one SPU is subdivided to address highly dynamic regional characteristics. The user-defined ruleset-based aggregations and divisions are configured to obtain spectrally unique stable landscape objects. Additional TFs as well as probability layers from the raster product are used for more effective aggregation and/or division. If the Softbone segments (stable landscape objects) failed to meet the “look-and-feel” internal quality control, deck (parameter set) is adjusted to enhance the quality of Softbone segments. Once the quality criteria are met, the segments are passed to the subsequent process, “Retransformation”.

3.3.2.4 Retransformation

Retransformation is the very last step of the landscape object generation where Softbone geometries are vertex-reduced and smoothed and Hardbone geometries are re-imposed. This includes smoothing Softbone geometries, reintroduction of Hardbone lines, snapping vertices, cleaning the remaining MMU errors, and gaps and overlaps check. This is performed with an internally developed software programmed with open source libraries. The final landscape objects are subjected to an internal quality control concerning delineation, MMU, MMW and topological consistency.

3.3.2.5 Vector Product class assignment

The vector class assignment involves the integration and enrichment of the Landscape Object geometries with the thematic land cover information derived by the raster classification. To avoid counting raster cells on the border of segments twice or not at all, the cell centre extraction method is used, where only those raster cells are considered for a particular object, whose cell centres lie within the object geometry (Figure 26).

The attribution and classification of the Landscape Objects is based on LC classes mapped in the Raster Product and calculated as class percentages related to the total area of a Landscape Object. The following attributes are derived for each polygon:

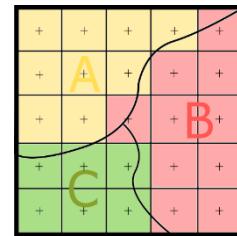


Figure 26: Illustration
of cell centre extraction
method

- Dominant LC class (majority rule), resulting in an 18-class land cover classification expressed by the dominant land cover class per polygon
- Percentage of all underlying LC classes in relation to the total area of the Landscape Object
- Ordered list of LC class percentages with three most dominant LC classes.

The subdivision into 18 classes is generated from the 11 LC classes of the Raster Product according to a predefined set of rules (Figure 4). In general, a class is expressed by the dominant LC class from the 11-class raster product within a Landscape Object ($\geq 50\%$). Exceptions from the 50% rule are ‘Non-vegetated land’ ($\geq 90\%$) and ‘Low vegetation cover’ (10-30%). Subclasses for ‘built-up land’, ‘woodland’ and ‘permanent herbaceous land’ are determined based on the LC statistics.

3.3.2.6 Vector Product attribute enrichment

In addition to the attribution with these 18 classes, the vector product is attributed with 58 additional attributes, derived from EO data, the CLC+ Backbone Raster classification and other CLMS products (from the pan-European, Hotspots and in-situ components). These additional attributes are described in more detail in the following:

S2_quality

The input for this attribute is the Data Score layer (DSL, described in section 3.4.1), which shows the valid cloud free Sentinel-2 observations per pixel. The attribute shows the mean number of observations per landscape object, calculated from all pixels whose centroids are covered by the object.

PRZ_mean

The input for this attribute is a continuous raster product from the Riparian Zones project, which shows the likelihood of membership (LOM) to potential riparian zones (PRZ) (0-100%) per pixel. Each object of the vector product is assigned with the mean LOM, calculated from all pixels whose centroids are covered by the object. Since the Riparian Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas of continental Europe, NoData is set for polygons not covered by the product.

PRZ_perc

The input for this attribute is a binary raster layer, derived from the PRZ LOM raster, showing the extent of the PRZ. The attribute shows how much area of a landscape object is covered by PRZ. Since the Riparian Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas of continental Europe, NoData is set for polygons not covered by the product.

RZ_1400_perc

The input for this attribute is a raster layer, derived from the Riparian Zones Dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by "Green urban, sports and leisure facilities". Since the Riparian Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas of continental Europe, NoData is set for polygons not covered by the product.

RZ_8110_perc

The input for this attribute is a raster layer, derived from the Riparian Zones Dataset, showing the different Landcover Classes. The attribute shows how much area of a landscape object is covered by "Coastal salt marshes". Since the Riparian Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas of continental Europe, NoData is set for polygons not covered by the product.

RZ_8120_perc

The input for this attribute is a raster layer, derived from the Riparian Zones Dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by "Salines". Since the Riparian Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas of continental Europe, NoData is set for polygons not covered by the product.

RZ_1120_perc

The input for this attribute is a raster layer, derived from the Riparian Zones Dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by "Industrial, commercial and military units". Since the Riparian Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas of continental Europe, NoData is set for polygons not covered by the product.

RZ_1230_perc

The input for this attribute is a raster layer, derived from the Riparian Zones Dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by "Port areas and associated land". Since the Riparian Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas of continental Europe, NoData is set for polygons not covered by the product.

RZ_1240_perc

The input for this attribute is a raster layer, derived from the Riparian Zones Dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by

“Airports and associated lands”. Since the Riparian Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas of continental Europe, NoData is set for polygons not covered by the product.

CZ_perc

The input for this attribute is a binary raster layer, derived from the Coastal Zones Dataset, showing the extent of the Coastal Zone Classes. The attribute shows how much area of a landscape object is covered by CZ. Since the Coastal Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

CZ_11200_perc

The input for this attribute is a raster layer, derived from the Coastal Zones Dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by “Industrial, commercial, public and military units”. Since the Coastal Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

CZ_12300_perc

The input for this attribute is a raster layer, derived from the Coastal Zones Dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by “Port areas and associated land”. Since the Coastal Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

CZ_12400_perc

The input for this attribute is a raster layer, derived from the Coastal Zones Dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by “Airports and associated land”. Since the Coastal Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

CZ_22100_perc

The input for this attribute is a raster layer, derived from the Coastal Zones Dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by “Vineyards, fruit trees and berry plantations”. Since the Coastal Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

CZ_22200_perc

The input for this attribute is a raster layer, derived from the Coastal Zones Dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by “Olive groves”. Since the Coastal Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

CZ_72100_perc

The input for this attribute is a raster layer, derived from the Coastal Zones Dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by "Salt marshes". Since the Coastal Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

CZ_72200_perc

The input for this attribute is a raster layer, derived from the Coastal Zones Dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by "Salines". Since the Coastal Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

CZ_72300_perc

The input for this attribute is a raster layer, derived from the Coastal Zones Dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by "Intertidal flats". Since the Coastal Zones input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

TCD_mean

The input for this attribute is the HRL 2018 Tree Cover Density layer (TCD) with 10m spatial resolution, showing the tree cover density (0-100%) per pixel. The attribute describes the mean tree cover density per landscape object, calculated from all TCD pixels whose centroids are covered by the object.

TCD_std

The input for this attribute is the HRL 2018 TCD layer with 10m spatial resolution. The attribute shows the tree cover density standard deviation per landscape object, calculated from all TCD pixels whose centroids are covered by the object.

IMD_mean

The input for this attribute is the HRL 2018 Imperviousness Density layer (IMD) with 10m spatial resolution, showing the sealing density (0-100%) per pixel. The attribute describes the mean imperviousness density per landscape object, calculated from all IMD pixels whose centroids are covered by the object.

IMD_std

The input for this attribute is the HRL 2018 IMD layer with 10m spatial resolution. The attribute shows the imperviousness density standard deviation per landscape object, calculated from all IMD pixels whose centroids are covered by the object.

IMC1518_mean

The input for this attribute is the HRL 2018 Imperviousness Density Change layer (IMC) with 10m spatial resolution, showing the increase and decrease of imperviousness density (-100 - 100%) per pixel. The

attribute describes the mean change of imperviousness density per landscape object, calculated from all IMC pixels whose centroids are covered by the object.

WAW_dry_perc

The input for this attribute is the HRL 2018 Water and Wetness layer (WAW) layer. The attribute shows how much area of a landscape object is covered by class “Dry”.

WAW_perm_water_perc

The input for this attribute is the HRL 2018 WAW layer. The attribute shows how much area of a landscape object is covered by class “Permanent Water”.

WAW_temp_water_perc

The input for this attribute is the HRL 2018 WAW layer. The attribute shows how much area of a landscape object is covered by class “Temporary Water”.

WAW_perm_wet_perc

The input for this attribute is the HRL 2018 WAW layer. The attribute shows how much area of a landscape object is covered by class “Permanent Wetness”.

WAW_temp_wet_perc

The input for this attribute is the HRL 2018 WAW layer. The attribute shows how much area of a landscape object is covered by class “Temporary Wetness”.

We_occurrence_mean

The input for this attribute is a raster layer, derived from the seasonal binary water and wetness masks of the HRL 2018 Water and Wetness layer production. It shows the occurrence of raster masks indicating water or wetness per pixel, covering a time span of 13 seasons (2015-2018). The attribute shows the mean occurrence of wetness masks per landscape object, calculated from all raster pixels whose centroids are covered by the object, normalized to a range of 0-1. In case of NoData gaps in the input dataset, the affected landscape object is assigned NoData.

Wa_occurrence_mean

The input for this attribute is a raster layer, derived from the seasonal binary water and wetness masks of the HRL 2018 Water and Wetness layer production. It shows the occurrence of raster masks indicating water or wetness per pixel, covering a time span of 13 seasons (2015-2018). The attribute shows the mean occurrence of water masks per landscape object, calculated from all raster pixels whose centroids are covered by the object, normalized to a range of 0-1. In case of NoData gaps in the input dataset, the affected landscape object is assigned NoData.

BAI_perc

The input for this attribute is a raster layer, derived from vector data provided by the European Forest Fire Information System (EFFIS), showing burnt areas all over Europe within the reference year 2018. The attribute shows the percentage of a landscape object’s area, affected by a fire event, according to the information in the EFFIS Burnt Areas product (BAI). Since the area percentage value is rounded to

4 decimal places, there can be objects, where the area percentage, covered by BAI, is so low, that the value is rounded to zero. This happens only in very rare cases, when a very large landscape object covers only very few pixel centroids of the input raster. This phenomenon is mentioned here, because it leads to diverging numbers of attributed objects for BAI_perc and BAI_Occ, when appearing. While a landscape object's area percentage might be rounded to zero (and thus indicate the absence of burnt areas), it would still have a value for the month of fire occurrence (BAI_Occ), because the extraction of the BAI_Occ value is independent of the object size, the number of pixel centroids covered and any rounding of values.

BAI_Occ

The input for this attribute is a raster layer, derived from the BAI layer, which provides information about the month in which a recorded fire event happened. The attribute shows the month of the fire occurrence for each object. It has the value 0 in case no fire has been detected in 2018. In case a landscape object covers two or more fire events from different dates with the same area percentage, the event which happened earlier in the year is prioritized. In case of multiple fire events with different extents, the largest one is prioritized (majority vote). Since the BAI area percentage value (BAI_perc; see above) is rounded to 4 decimal places, there can be objects, where the area percentage, covered by BAI, is so low, that the value is rounded to zero. This happens only in very rare cases, when a very large landscape object covers only very few pixel centroids of the input raster. This phenomenon is mentioned here, because it leads to diverging numbers of attributed objects for BAI_perc and BAI_Occ, when appearing. While a landscape object's area percentage might be rounded to zero (and thus indicate the absence of burnt areas), it would still have a value for the month of fire occurrence (BAI_Occ), because the extraction of the BAI_Occ value is independent of the object size, the number of pixel centroids covered and any rounding of values.

UA_perc

The input for this attribute is a binary raster layer, derived from the Urban Atlas Dataset, showing the extent of the Urban Areas. The attribute shows how much area of a landscape object is covered by Urban Area.

UA_core_perc

The input for this attribute is a binary raster layer, derived from the Urban Atlas dataset, showing the extent of the Urban Core Areas. The attribute shows how much area of a landscape object is covered by Urban Core Area. Since the Urban Atlas input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

UA_12100_perc

The input for this attribute is a raster layer, derived from the Urban Atlas dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by "Industrial, commercial, public, military and private units". Since the Urban Atlas input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

UA_12300_perc

The input for this attribute is a raster layer, derived from the Urban Atlas dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by “Port areas”. Since the Urban Atlas input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

UA_12400_perc

The input for this attribute is a raster layer, derived from the Urban Atlas dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by “Airports”. Since the Urban Atlas input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

UA_14100_perc

The input for this attribute is a raster layer, derived from the Urban Atlas dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by “Green urban areas”. Since the Urban Atlas input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

UA_14200_perc

The input for this attribute is a raster layer, derived from the Urban Atlas dataset, showing the different Land Cover Classes. The attribute shows how much area of a landscape object is covered by “Sports and leisure facilities”. Since the Urban Atlas input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

Building_perc

The input for this attribute is a binary raster layer, derived from the Urban Atlas 2012 dataset, showing the extent of the building areas. The attribute shows how much area of a landscape object is covered by built-up area. Since the Urban Atlas input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product. Additionally, all landscape object polygons, which do not cover any building, are assigned NoData.

Building_height

The input for this attribute is a continuous raster layer, derived from the Urban Atlas 2012 dataset, showing the height of all building area pixels. The attribute shows the 90th percentile of building height per landscape object, calculated from all pixels whose centroids are covered by the object. Since the Urban Atlas input dataset is a CLMS Hotspots component product and thus limited to selected areas in continental Europe, NoData is set for polygons not covered by the product.

SLP_dg_med

The input for this attribute is a raster, showing the slope in degree per pixel, which was derived from the Copernicus DEM (Digital Elevation Model) at 10m spatial resolution. The attribute describes the median of the slope per landscape object, calculated from all pixels whose centroids are covered by

the object. The calculation of slope per pixel is dependent on the availability of height information of all neighbouring pixels. There are some rare cases at the EEA-38 + UK outer border where this information has not been available and the slope couldn't be calculated. Thus *slope_dg_med* is set to NoData in these cases.

EXP_c_maj

The input for this attribute is a raster layer, showing the exposition per pixel (Flat, North, East, South and West orientation), derived from the above described slope raster. The attribute describes the exposition of each landscape object, based on the majority value of all pixels whose centroids are covered by the object. As a consequence of this dependency on the slope raster, *EXP_C_maj* shows NoData values in those cases where *slope_dg_med* gives NoData (see explanation above).

MASL_min

The input for this attribute is the Copernicus DEM with 10m spatial resolution (version released Jan. 2021). The attribute shows the minimum height above sea level for each landscape object, based on all pixels whose centroids are covered by the object.

MASL_mean

The input for this attribute is the Copernicus DEM with 10m spatial resolution (version released Jan. 2021). The attribute shows the mean height above sea level for each landscape object, calculated from all pixels whose centroids are covered by the object.

MASL_max

The input for this attribute is the Copernicus DEM with 10m spatial resolution (version released Jan. 2021). The attribute shows the maximum height above sea level for each landscape object, based on all pixels whose centroids are covered by the object.

FFI_name

The input for this attribute is the CLC+ Backbone raster classification. All forest related classes, namely coniferous, broadleaved and broadleaved evergreen (raster classes 2, 3 and 4) were reclassified to a binary tree cover – non tree cover map at 10m spatial resolution. Based on this tree cover map, a forest fragmentation index is calculated, using the Foreground Area Density (FAD) algorithm, provided by GUIDOS Toolbox (https://ies-ows.jrc.ec.europa.eu/gtb/GuidosToolbox_Manual.pdf). The FAD measures forest density as the average forest proportion over several observation scales (moving windows with different sizes). The window side lengths reach from 7 to 243 pixels and resulting values (being average percentages) from 0 (no forest in any window) to 100 (only forest in all windows). So, in difference to the HRL Tree Cover Density, the FAD considers the surrounding of each pixel to derive the Forest density and thus provides a different kind of information. The FAD values range from 0-100% and are categorized in 6 classes (Figure 27). The resulting attribute shows the FAD category for each landscape object, based on the mean FAD value calculated from all pixels whose centroids are covered by the object.

1-Rare		FAD < 10%
2-Patchy		10% ≤ FAD < 40%
3-Transitional		40% ≤ FAD < 60%
4-Dominant		60% ≤ FAD < 90%
5-Interior		90% ≤ FAD < 100%
6-Intact		FAD = 100%

Figure 27: FAD classes and respective value ranges

FFI_num

The input for this attribute is the FAD raster described above (**FFI_name**). It shows the average FAD value per landscape object, calculated from all pixels whose centroids are covered by the object.

MSPA_core

The input for this and all subsequent MSPA (Morphological Spatial Pattern Analysis) attributes is a raster layer, calculated with an algorithm provided by GUIDOS Toolbox like the FAD, as well (https://ies-ows.jrc.ec.europa.eu/gtb/GTB/GuidosToolbox_Manual.pdf). The MSPA raster is calculated from the tree cover map, derived from the tree related classes of the CLC+ Backbone raster classification (also see **FFI_name**). The MSPA results in 7 different classes, providing information about the “system” formed by all tree covered areas (Figure 28). These seven classes are:

- **Core:** Interior forest area excluding perimeter;
- **Islet:** Disjoint forest element and too small to contain Core;
- **Loop:** Connected to the same Core area;
- **Bridge:** Connected to different Core areas;
- **Perforation:** Internal object perimeter, showing the margin of holes within forest core areas;
- **Edge:** external object perimeter, showing the outer margin of forest core areas;
- **Branch:** connected at one end to Edge, Perforation, Bridge or Loop.

All above MSPA features are calculated with respect to the "Foreground" input class "tree cover" (comprising Woody needle leaved trees + Woody Broadleaved deciduous trees + Woody Broadleaved evergreen trees of the initial CLC+ Backbone Raster Product, and are considering horizontal, vertical and diagonal pixel connectivity.

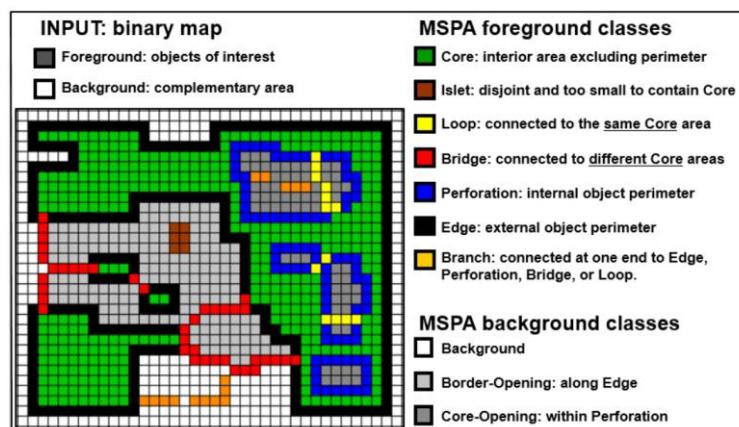


Figure 28: Example MSPA raster showing all possible MSPA classes

The MSPA_core attribute shows the area percentage per landscape object, covered by the MSPA core class. It is calculated, based on all pixels whose centroids are covered by the object.

MSPA_isle

This attribute shows the area percentage per landscape object, covered by the MSPA isle class. It is calculated, based on all pixels whose centroids are covered by the object.

MSPA_loop

This attribute shows the area percentage per landscape object, covered by the MSPA loop class. It is calculated, based on all pixels whose centroids are covered by the object.

MSPA_brid

This attribute shows the area percentage per landscape object, covered by the MSPA bridge class. It is calculated, based on all pixels whose centroids are covered by the object.

MSPA_perf

This attribute shows the area percentage per landscape object, covered by the MSPA perforation class. It is calculated, based on all pixels whose centroids are covered by the object.

MSPA_edge

This attribute shows the area percentage per landscape object, covered by the MSPA edge class. It is calculated, based on all pixels whose centroids are covered by the object.

MSPA_bran

This attribute shows the area percentage per landscape object, covered by the MSPA branch class. It is calculated, based on all pixels whose centroids are covered by the object.

LI_mean

The inputs for this attribute are the two Large Integral Layers (Season 1 and Season 2) from the HRL Vegetation Phenology and Productivity (VPP) for the year 2018. The mean of the Large Integral Layers (LI) is calculated for each landscape object and both seasons, based on all pixels whose centroids are covered by the object. For each polygon, the result of season 1 or season 2 is assigned, depending on which season shows the higher LI mean. This polygon-based information (season 1 or season 2) is further used for the assignment of the other two VPP attributes (SOS_mean and LOS_mean).

SOS_mean

The inputs for this attribute are the two Start of Season Layers (Season 1 and Season 2) from the HRL Vegetation Phenology and Productivity (VPP) for the year 2018. The mean of the Start of Season Layers (SOS) is calculated for each landscape object and both seasons, based on all pixels whose centroids are covered by the object. For each polygon, the result of season 1 or season 2 is assigned, depending on which season shows the higher LI_mean. The values range between 17086 and 18365, with the first two digits indicating the year (2017/2018), and the last three digits referring to the day of the year.

LOS_mean

The inputs for this attribute are the two Length of Seasons Layers (Season 1 and Season 2) from the HRL Vegetation Phenology and Productivity (VPP) for the year 2018. The mean of the Length of Season Layers (LOS) is calculated for each landscape object and both seasons, based on all pixels whose centroids are covered by the object. For each polygon, the result of season 1 or season 2 is assigned, depending on which season shows the higher LI_mean.

3.3.2.7 Strengths and Limitations of the Applied Methodology

The methodology developed and applied for the Vector Product generation has many strengths, but also entails some limitations, as described in the following.

Particular advantages are that landscape objects are well captured according to their spectral, textural and temporal characteristics, independent from the Raster Product classification layers, in an unprecedented pan-European fashion, while paying much attention to regional calibration to biogeographically different vegetation, phenology and management practise patterns. Furthermore, the product provides a unique blend of information for each polygon through numerous attributes, derived from the CLC+ Backbone Land Cover Raster Product and various additional CLMS and other sources. Moreover, the approach allows that landscape objects – although belonging to the same LC class – are still captured separately as individual objects – such as young coniferous tree stands and fully-grown coniferous tree stands, grasslands with different management schemes, or agricultural fields with different crop types or distinctly different phenological development due to soil type/moisture differences, etc. This is the results of several methodological elements, including:

- Incorporation of existing reference data of linear landscape objects (such as traffic and hydrological networks) to form a landscape Hardbone;
- Time Features generated from multi-temporal Sentinel-2 (reference year 2018) L2A data, which comprise essential information on the spectral-temporal dynamics of stable landscape objects;
- enhanced cloud mask based and FMask version 4, that addressed some of the shortcomings of the default Scene Classification Layer (SCL) provided by ESA;
- usage of a state-of-the-art open-source and algorithmically reproducible image segmentation technique;
- regional calibration, which allows proper handling of the complex diversity of biogeographical regions and related land cover characteristics across the EEA-38 + UK.

On the other hand, some technical challenges and limitations have been encountered at the various stages of the workflow. Particularly in terms of input data quality, the S-2 L2A data have proven not ideal in terms of positional accuracy over time, the overall quality of the cloud and cloud shadow masks, and the quality of the L2A topographic normalization. In-situ data such as OSM, EU-Hydro, WISE, and national datasets of traffic and hydrological networks showed certain inconsistency in data quality (e.g. containing some inaccurate data or missing labels), which mostly had to be accepted in the highly automatic processing chain.

It is worth mentioning that **areas of heterogeneous land cover and small landscape features** typically cause less definite outlines during the segmentation, due to mixed spectral-temporal signatures and the small-scale patterns interfering with the product's MMU of 0.5 ha. Heterogeneous land cover often triggers higher densities of segmentation, especially in transition zones such as grassland to forest or barren land to sparsely vegetated areas. This is caused, amongst others, by the S-2 L2A imprecision in positional accuracy over time, which impacts the achievable geometric and thematic accuracies. Furthermore, landscape object boundaries in natural areas of seemingly high homogeneity (such as large sparsely vegetated areas or tropical rainforest in the French DOMs) as grouped by the segmentation algorithm, tend to be sometimes difficult to visually reproduce by looking at mono-temporal images only.

During the final visual checks of the Vector Product, some **micro-gaps and micro-overlaps** were detected mainly at the shared PU borders. Such observed micro-gaps or overlaps vary from 0.01 – 3 mm and have occurred during the final checks in a GIS software due to different innate precisions and the limitation of the topology tools. Since topology tools used for the final topology checks do not guarantee that the supposedly first rank geometries are not changed, it is almost impossible to avoid such micro-gaps and overlaps at the Production Unit border without taking the full geometries (covering EEA 38+UK) into account. These micro-gaps or overlaps are only visible at a very large scale and are not detectable with common commercial GIS software with 6-digit precision like ArcGIS suite.

In some regions of the delivered Vector Product, **different densities of the landscape object polygons** seem to occur along Production Unit boundaries, particularly if these boundaries coincide with the boundaries between biogeographic regions or major landscapes. This is sometimes noticeable when looking at a relatively smaller scale (1:500,000 or smaller), however observed at a larger scale (1:50,000 or larger) this is barely recognizable. Such regional density differences in segmentation are closely coupled with the regional calibration of the segmentation rules which are fitted to capture the main landscape objects in the production region. Mostly, these edges are only visual display effects in the GIS and are typically not related to the quality of the segmentation. For example, in Croatia, a segment density difference between two neighbouring SPUs (129_1 and 129_2) seems identifiable at high zoom-out level (Figure 29). Specifically, SPU 129_2 appears as having smaller and more compact segments, while SPU 129_1 has larger segments. This SPU boundary indeed coincides with the boundary between two geomorphologically and topographically different biogeographical regions. SPU 129_1 extends into a plain, while SPU 129_2 is extending over a more topographically complex area. The segmentation effectively captures the landscape objects' intrinsic particularities in each region. Similarly, in Spain, some density differences seem to appear along the border where three SPUs meet (Figure 30). This is, however, originated from the difference in natural landscapes which coincide with the SPU borders. This is not by chance, but intentional, as the SPUs have been created using the biogeographic region, in order to allow meaningful and consistent local algorithm calibration. As shown in both examples, the apparent visual edge effect is not recognizable once zoomed in to a larger scale in the GIS.

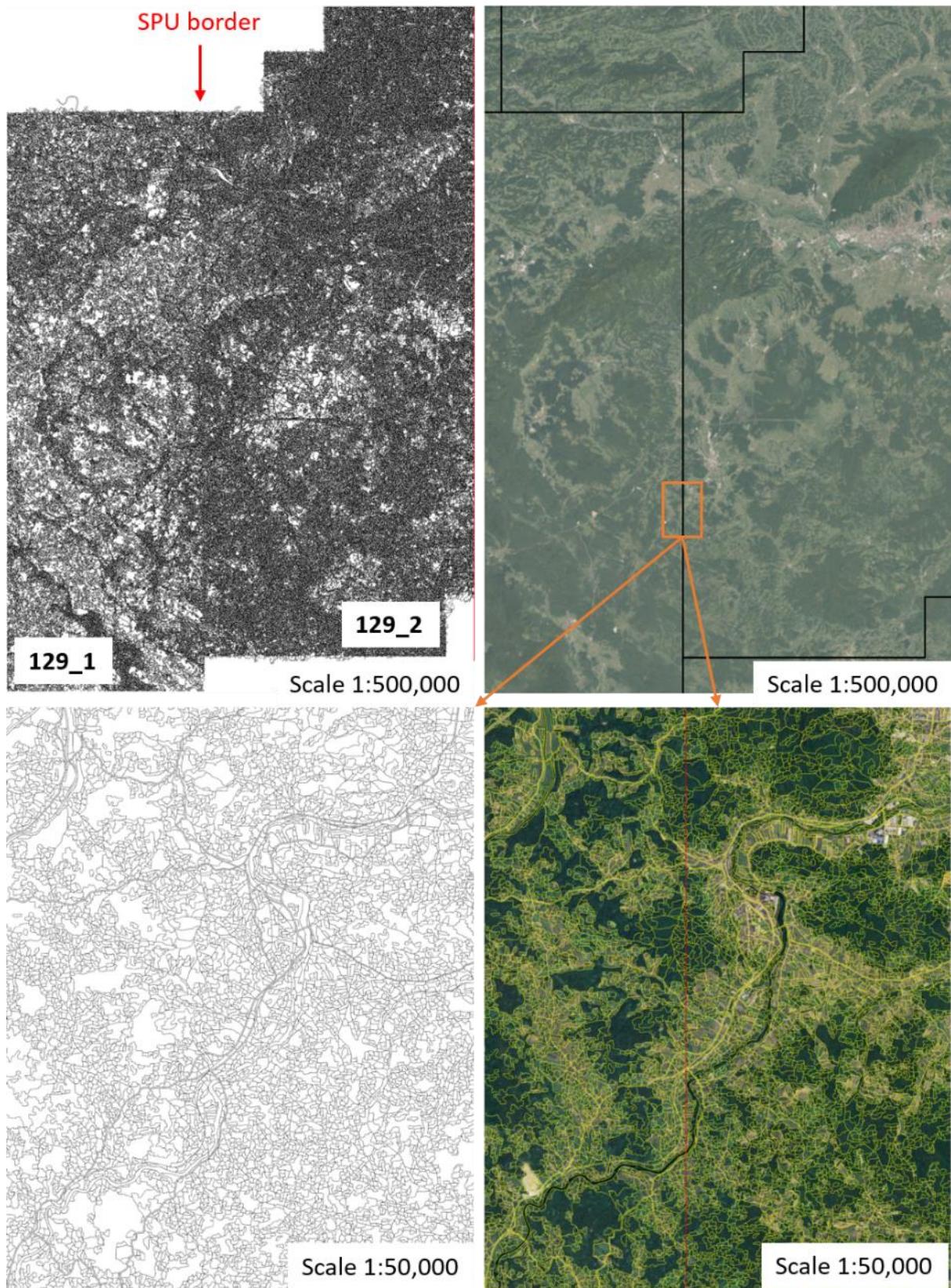


Figure 29: Example of apparent regional segmentation density difference in Croatia (SPU 129_1 & 129_2) at small and large scales

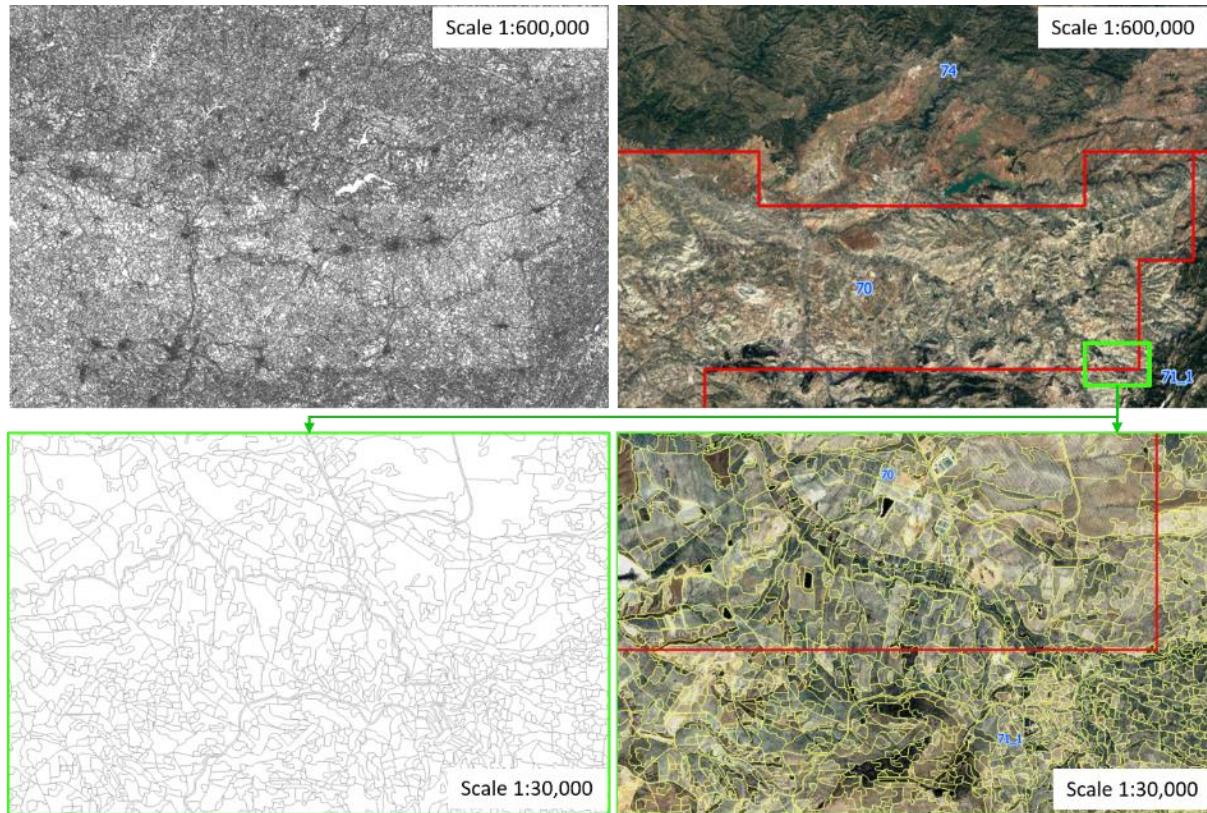


Figure 30: Example of regional segmentation density difference in Spain (SPU 74, 70 and 71_1) at small and large scales

During the Softbone segmentation and border harmonisation, **water bodies – particularly long rivers – and major roads** (OSM code 5111, 5112, 5113 and 5114), had sometimes proven too large and complex to be processed as single polygons in their full length. Considering the size and the geometric complexity of such segments, polygons representing water bodies (such as rivers and lakes) or major roads inevitably had to be technically split at the straight SPU border if the segment was too big (see Figure 31).

As for **sea water** polygons, due to the large size as well as the complexity of the shorelines along the coast, polygons representing open sea water are divided according to the EEA 10km grid (see Figure 32). Such polygons were defined as the intersection of the water mask (calculated from the CLC+ Backbone Raster Product) and the delivery boundary that has a buffer of 250m from the coastline. These measures ensured fluent processing and at the same time improved the usability and stability of the product for users, by reducing the size of the very large and complex segments.

By definition, the CLC+ Backbone land cover nomenclature does not provide a separation between **inland water vs. sea water** along a country's coastline. For potential use cases intending to separate both, a clipping with a national/European coast line dataset is recommended.

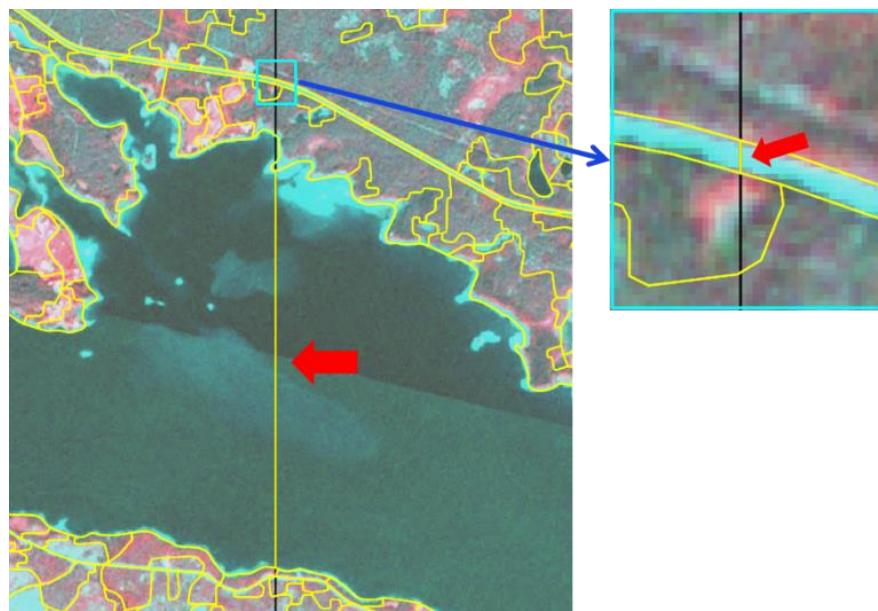


Figure 31: Artificial cut (yellow) in river and major road polygons at the SPU border (black)

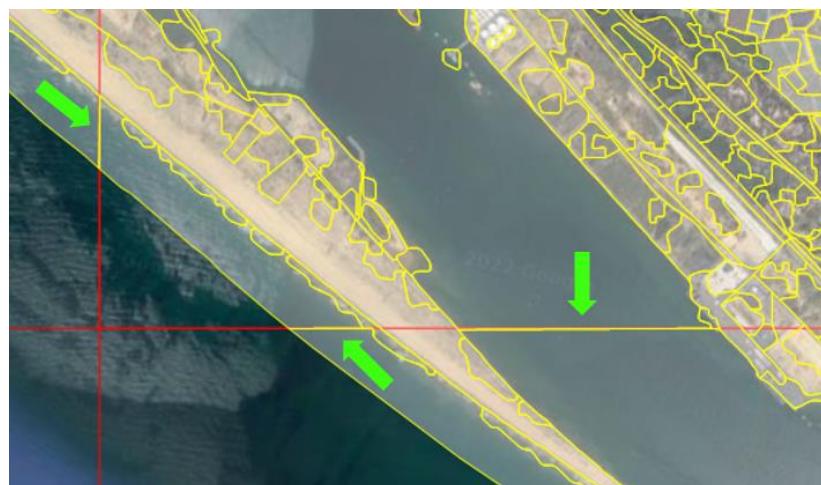


Figure 32: Artificial cut (yellow) in ocean polygons at the EEA 10km grid (red)

3.4 Additional Quality Layers and Expert Products

The CLC+ Backbone Raster Product is accompanied by three additional Quality Layers (section 3.4.1 - 3.4.3) and the Vector Product by one Quality Layer (section 3.4.4), which are described in the following sections.

3.4.1 Data Score Layer (DSL)

The product specifications for the CLC+ Backbone DSL layer are summarized in the following table. Further details and information can be found below the table.

CLC+ Backbone Raster Data Score Layer	Acronym	Product family		
Summary				
The Data Score Layer is a 10m pixel-based Quality Layer for the CLC+ Backbone Raster Product. It is based on Sentinel time series from July 2017 to June 2019 to derive the amount of valid observations during this time frame.				
Reference year				
2018				
Geometric resolution				
Pixel resolution 10m x 10m, fully conform with the EEA reference grid				
Coordinate Reference System				
European ETRS89 LAEA projection / for French DOMs WGS84 and the respective UTM zone				
Coverage				
6,002,168 km ² covering the full EEA-38 + UK				
Geometric accuracy (positioning scale)				
equals the Sentinel-2 positional accuracy in 2018 (~11m at 95.5% confidence)				
Data type				
16bit unsigned raster with LZW compression				
Minimum Mapping Unit (MMU)				
Pixel-based (no MMU)				
Attributes				
Field	Description	Type	Value range	Outside a
Value	Pixel value	Integer		65535
Count	Number of pixels with the corresponding pixel value	Double		
Obs_count	Number of cloud free observations	String	0 – 200 in full EEA-38 + UK / 0 – 85 for French DOMs	"Outside a
Delivery format				
GeoTIFF incl. pyramids (*.ovr, resampling: Nearest Neighbour), attribute table (*.dbf), statistics (*.aux.xml), integrated as well as external colour tables (*.clr).				

The Data Score Layer (DSL) for the reference year 2018 (± 6 months) is shown in Figure 33. For each pixel, it presents the number of valid Sentinel-2 observations available to interpolate the equidistant time series (72 time steps), which is finally ingested into the classifier. Valid observations are defined as i) having a cloud coverage in the Sentinel-2 L2A metadata below 80% and ii) cloud and shadow free according to the FMask cloud masks. The DSL can be considered as indicator for the quality of the available data base for the raster classification. The clearly visible stripe patterns from SW to NE correctly represent the different data quantities between the overlapping areas of the S-2 swaths. Besides this normal general pattern of the DSL, there are some patterns that are worth further explaining. Sentinel-2 GRANULE overlaps are often visible as horizontal and vertical stripes with slightly higher observation counts. This is due to the fact that in the overlap areas, there is a higher chance that one of the GRANULES exceeds the 80% cloud cover threshold.

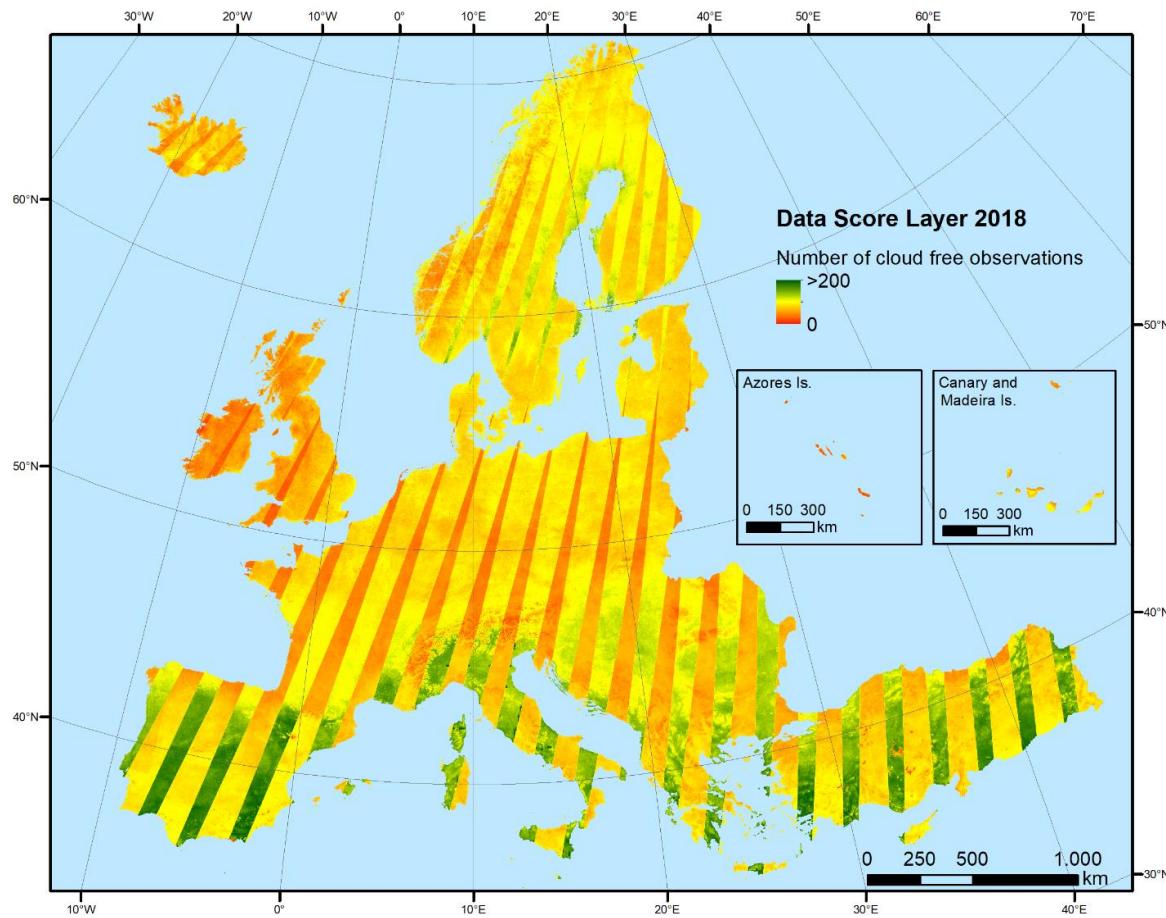


Figure 33: CLC+ Backbone Data Score Layer for Sentinel-2 observations of the reference year 2018 (± 6 months) over the area of the EEA-38 + UK

3.4.2 Raster Confidence Layer (CL) Specifications

The product specifications for the CLC+ Backbone Raster CL layer are summarized in the following table. Further details and information can be found below the table.

CLC+ Backbone Raster Confidence Layer	Acronym	Product family		
Summary				
The Confidence Layer is a 10m pixel-based Quality Layer for the CLC+ Backbone Raster Product. It provides information about the reliability of the land cover class assignment per pixel.				
Reference year				
2018				
Geometric resolution				
Pixel resolution 10m x 10m, fully conform with the EEA reference grid				
Coordinate Reference System				
European ETRS89 LAEA projection / for French DOMs WGS84 and the respective UTM zone				
Coverage				
6,002,168 km ² covering the full EEA-38 + UK				
Geometric accuracy (positioning scale)				
equals the Sentinel-2 positional accuracy in 2018 (~11m at 95.5% confidence)				
Data type				
8bit unsigned raster with LZW compression				
Minimum Mapping Unit (MMU)				
Pixel-based (no MMU)				
Attributes				
Field	Description	Type	Value range	Outside a
Value	Pixel value	Integer		254
Count	Number of pixels with the corresponding pixel value	Double		
Confidence	Confidence value in percentage	String	0 – 100	"Outside a
Delivery format				
GeoTIFF incl. pyramids (*.ovr, resampling: Nearest Neighbour), attribute table (*.dbf), statistics (*.aux.xml), integrated as well as external colour tables (*.clr).				

The Raster Confidence Layer provides information about the reliability of the land cover class assignment per pixel (Figure 34). More specifically, it depicts the difference of the probabilities for the highest ranked class and the second highest ranked class (often referred to as probability margin), whereas high values are an indicator for a higher confidence of the classifier regarding the assigned class. The partially visible edges between adjacent Production Units are the result of different regional

classification model trainings. Areas with lower confidence are typically concentrated in transition areas, with a higher degree of mixed spectral-temporal signatures from mixed land cover types.

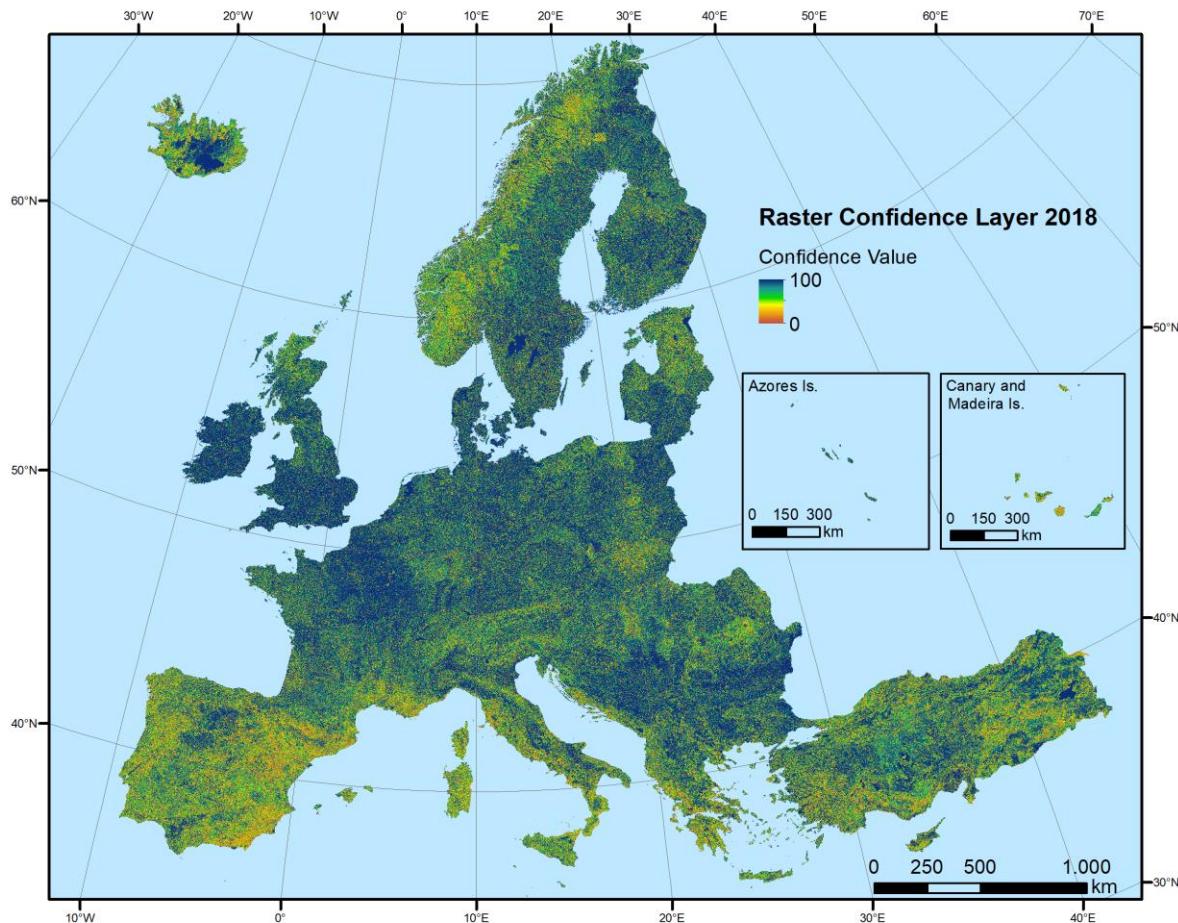


Figure 34: CLC+ Backbone Raster Confidence Layer 2018

The distribution of the raster confidence values is displayed in Figure 35. It shows a significant increase of confidence values from approximately 90% upwards. Overall, 75% of all classified pixels have been assigned to the respective dominant class (out of max. 11 classes) with a confidence of $\geq 50\%$. Considering the classifier's general difficulty to distinguish between up to 11 classes (depending on the present land cover in the respective regions), this shows that the classifier was able to draw on a sufficiently large data and sample basis, to resolve most uncertainties in the classification. While this measure of uncertainty is not always a good proxy for the distribution of errors (i.e. in case of spectral-temporal similarity, the classification can still commit an error with high confidence), it is worth noting that areas with lower confidence generally coincide with areas where land cover classification is generally more difficult (i.e. Southern Europe, Norway and Iceland).

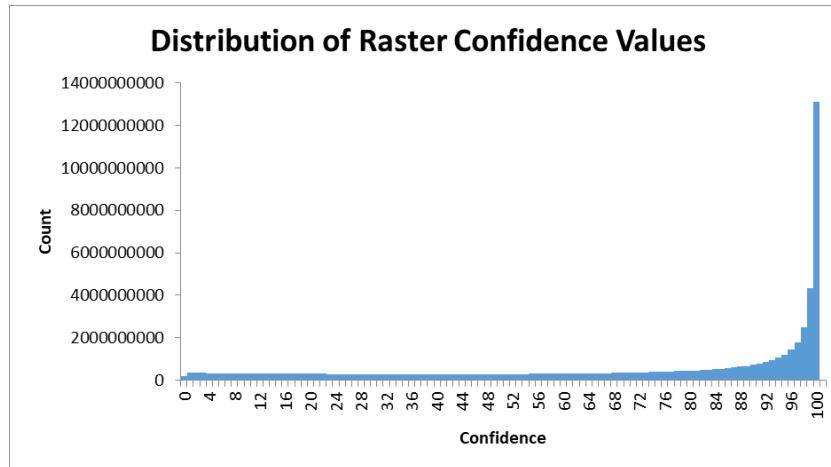


Figure 35: Distribution of CLC+ Backbone Raster Confidence Layer values across the EEA-38 + UK

3.4.3 Raster Post-processing Layer

The product specifications for the CLC+ Backbone Raster Post-processing Layer are summarized in the following table. Further details and information can be found below the table.

CLC+ Backbone Raster Post-processing Layer	Acronym	Product family
	POST	CLMS_CLCplus
Summary		
The Post-processing Layer is a 10m pixel-based Quality Layer for the CLC+ Backbone Raster Product. It provides information of pixels that were re-coded during post-processing of the raster classification.		
Reference year		
2018		
Geometric resolution		
Pixel resolution 10m x 10m, fully conform with the EEA reference grid		
Coordinate Reference System		
European ETRS89 LAEA projection / for French DOMs WGS84 and the respective UTM zone		
Coverage		
6,002,168 km ² covering the full EEA-38 + UK		
Geometric accuracy (positioning scale)		
equals the Sentinel-2 positional accuracy in 2018 (~11m at 95.5% confidence)		
Data type		
8bit unsigned raster with LZW compression		
Minimum Mapping Unit (MMU)		
Pixel-based (no MMU)		
Post-processing coding		
0: No change during post-processing 1: Recoded during post-processing		

Attributes				
Field	Description	Type	Value range	Outside a
Value	Pixel value	Integer	0 – 1	254
Count	Number of pixels with the corresponding pixel value	Double		
Class_name	If the pixel went through change during post-processing or not	String		"Outside a"

Delivery format				
GeoTIFF incl. pyramids (*.ovr, resampling: Nearest Neighbour), attribute table (*.dbf), statistics (*.aux.xml), integrated as well as external colour tables (*.clr).				

The Raster Post-processing Layer for the CLC+ Backbone Raster Product 2018 marks all pixels, which were re-coded during the post-processing of the raster classification (Figure 36). The green areas represent these re-coded and thus improved areas. The rulesets were intentionally tailored to be rather conservative and introduce changes only where auxiliary layers provide very reliable information. This is reflected in the rather small fraction of pixels that underwent automated corrections (5.5%) during the post-processing.

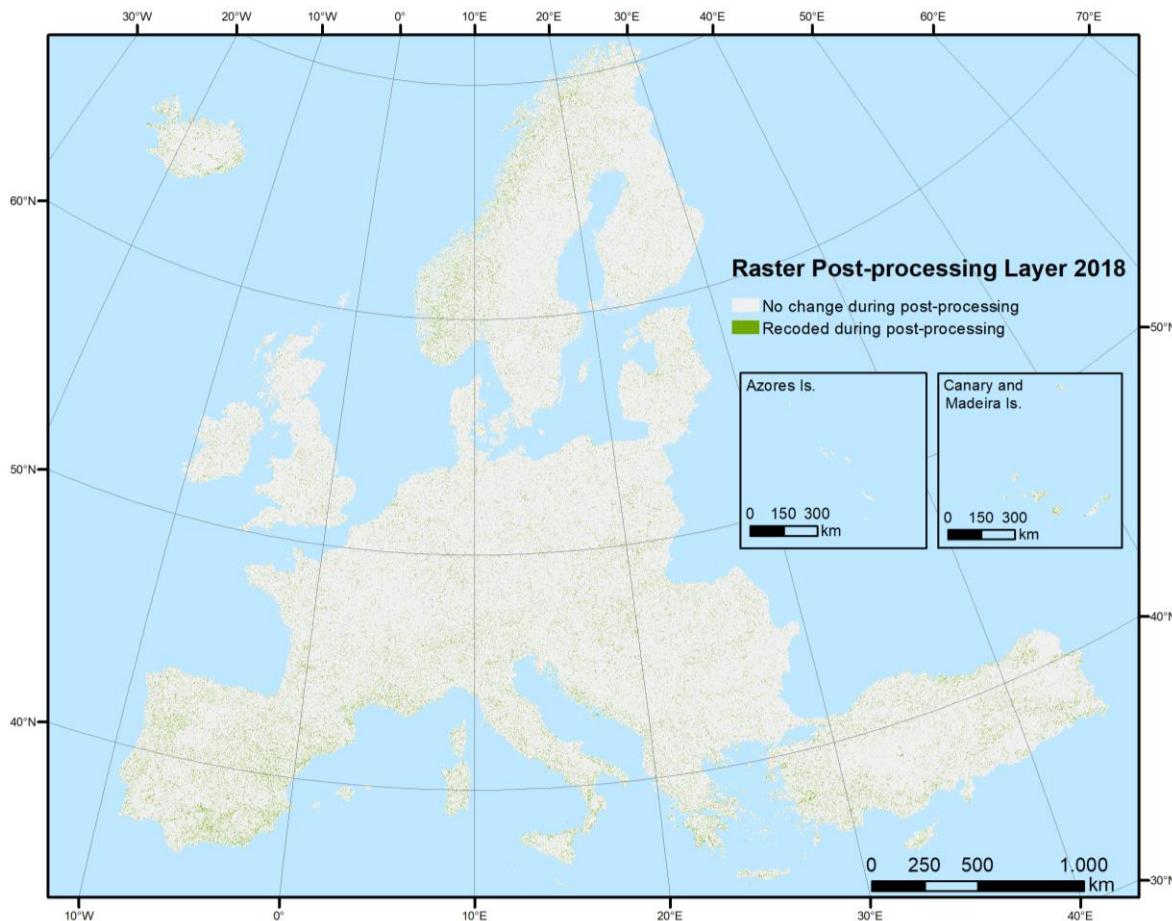


Figure 36: CLC+ Backbone Raster Post-processing Layer 2018

3.4.4 Segmentation Confidence Layer

The Segmentation Confidence Layers (Table 5) provide metrics of the textural and spectral homogeneity of the derived landscape object segments, as assessed versus the independent VHR 2018 dataset. Confidence layer values are calculated per segment and provided as attributes of the final Vector product. Examples are shown in Figure 37 grouped into 7 classes. Since the information content of the Segmentation Confidence Layers is directly triggered by the quality of the VHR 2018 dataset, which, in some cases, contains significant heterogeneity caused by a small-scale mix of different sensors' imagery, it is suggested that the 2018 VHR datasets be consulted additionally in specific cases, to identify possible anomalies.

Table 5: Overview of Segmentation Confidence Layers

Layer name	Data source and transformation	Description	Value range
Confidence layer 1: Textural variance (CONF_TEX)	Source: VHR (Red, Green, Blue, NIR) Transformation: GLCM homogeneity	The mean variance of the homogeneity layer calculated for each landscape object. If a landscape object is continuously homogenous or heterogeneous, the variation of homogeneity is low. (Lower variation in homogeneity = higher confidence)	0 or above NoData: 999
Confidence layer 2: Spectral variance (CONF_SPEC)	Source: VHR (Red, Green, Blue, NIR) Transformation: None	The mean spectral variance calculated for each landscape object. (Lower variation in spectral response = higher confidence)	0 or above NoData: 999

The interpretation of both Segmentation Confidence Layer components is not straightforward in all cases. In general, low variations of both textural and spectral heterogeneity within a polygon might indicate a high confidence of that polygon's adequate segmentation. High variations in both components might be indications of the opposite. In any case, both components have to be looked at synoptically.

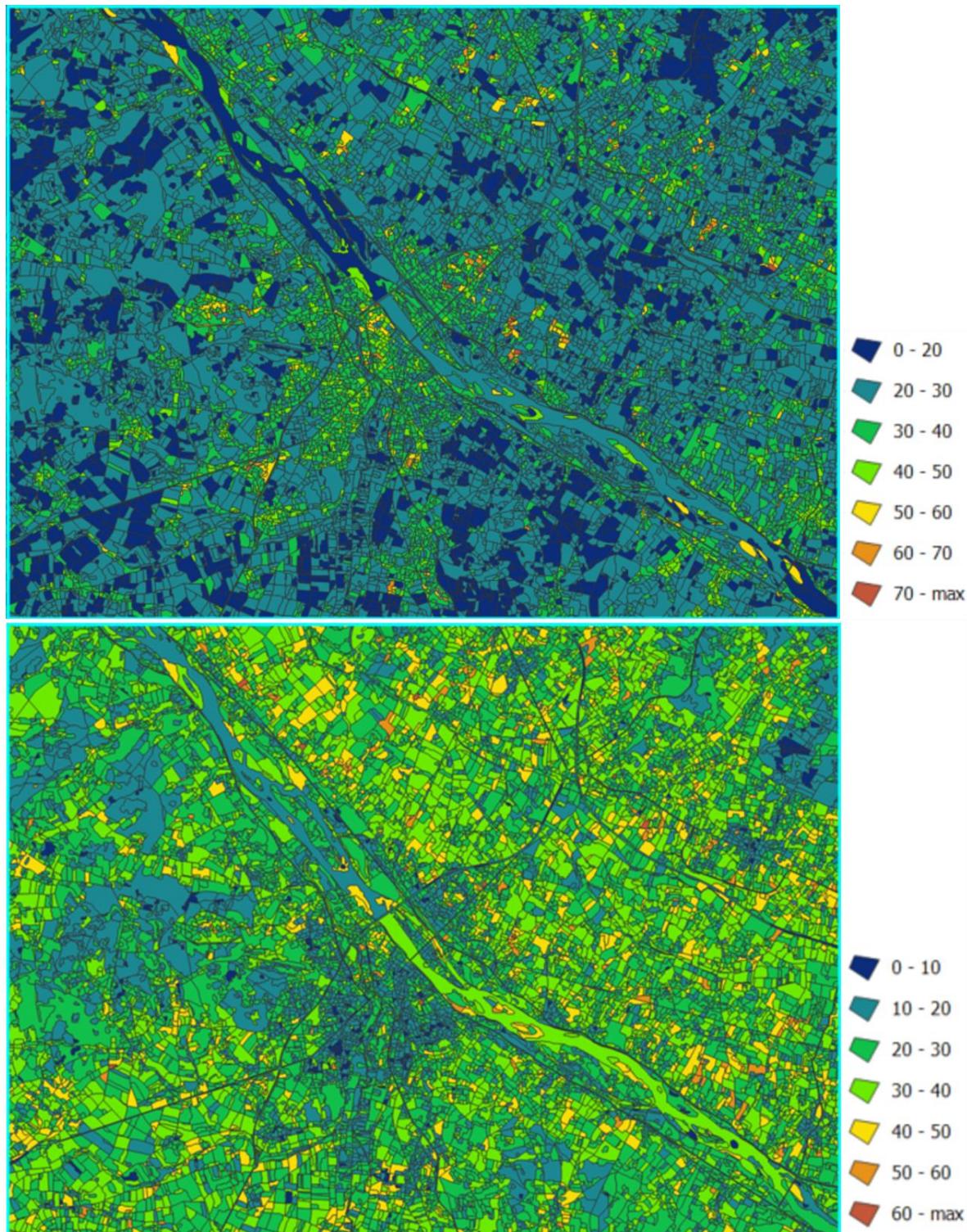


Figure 37: Segmentation Confidence Layer 1 (CONF_TEX, above) and Confidence Layer 2 (CONF_SPEC, below)

4 Product Quality

The following sections provide the final results of internal Validation of the Raster Product's thematic accuracy (section 4.1), the Vector Product's thematic accuracy (section 4.3) and the Vector Product's geometric accuracy, which were performed as internal task of the CLC+ Backbone creation, but completely separated and independent of the production teams.

4.1 Raster Product Thematic Accuracy

This section provides the assessment results of the CLC+ Backbone Raster Product's thematic accuracy.

4.1.1 Methodological approach

The CLC+ Backbone Raster Product's thematic validation approach is designed to be suitable for different reporting levels, i.e. Pan-European (EEA-38 + UK), Biogeographical regions and (groups of) countries (countries < 90,000 km² are grouped into contiguous groups of countries resulting into 23 **reporting units** already used for previous CLMS reporting, see Figure 38). A stratified random sampling approach is implemented for the validation of the CLC+ Backbone raster product. The sampling design is prepared per country (or group of small countries), so that analysis and reporting can be done on both, the aggregate of all reporting units and the country (group) level.

For the Raster Product, the **sampling unit** considered for validation is the pixel, as the purpose is to validate a pixel-based classification. The distribution of samples is designed such that both accuracies – per class and overall – can be assessed properly.

The allocation of a required minimum number of samples according to the **sampling design** is performed by considering a given confidence interval and a given acceptable error in the sample. Similar to the CLC validation and as evoked by Congalton (1991), Plourde (2003) and Olofsson et al. (2013), the methodology considers the binomial distribution and the statistical formula of Cochran (1977) to estimate the sample size (n). The sample size is calculated based on an expected accuracy of each class. Given the accuracy specifications of the CLC+ Backbone LC classes, an average accuracy per class P of 90% is assumed and the target confidence interval is set to 5%. The formula then gives a sample of 140 points per class for both, the Raster and the Vector Product.

An overview of the sampled areas and the quantitative distribution of samples per class for the Raster Product are provided in Table 6 and Table 7.

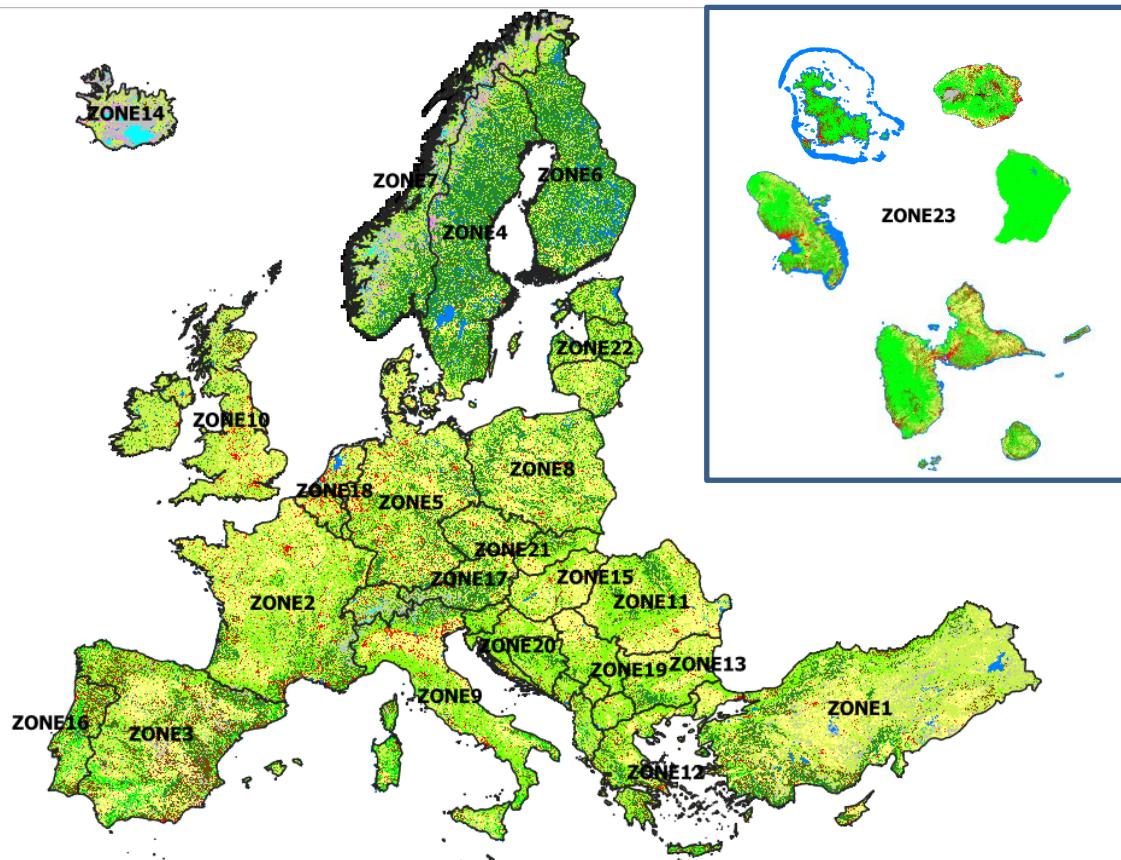


Figure 38: Distribution of the sampling and reporting units for the EU27 Raster Product

Table 6: Overview of the validated areas and distribution of sampling units in each CLC+ Backbone raster class

Zone	Acronym	Country	Coverage (Entirely / Partially / not at all)
Z1	TR	Turkey	Entirely
Z2	FR	France	Entirely
Z3	ES-AD	Spain-Andora	Entirely
Z4	SE	Sweden	Entirely
Z5	DE	Germany	Entirely
Z6	FI	Finland	Entirely
Z7	NO	Norway	Entirely
Z8	PL	Poland	Entirely
Z9	IT-MT	Italy - Malta	Entirely
Z10	UK-IE	United Kingdom - Ireland	Entirely
Z11	RO	Roumania	Entirely
Z12	CY-EL	Cyprus-Greece	Entirely
Z13	BG	Bulgaria	Entirely
Z14	IS	Island	Entirely
Z15	HU	Hungaria	Entirely
Z16	PT	Portugal	Entirely
Z17	AT-CH-LI	Austria-Switzerland -Liechtenstein	Entirely
Z18	BE-DK-LU-NL	Belgium - Danmark - Luxembourg - Netherlands	Entirely
Z19	AL-KX-MK-ME-RS	Albania- Kosovo- Macedonia- Montenegro-Serbia	Entirely
Z20	SI-HR-BA	Slovenia- Croatia- Bosnia & Herzegovina	Entirely
Z21	CZ-SK	Czech Republic - Slovakia	Entirely
Z22	LV-LT-EE	Latvia- Lithuania-Estonia	Entirely
Z23	DOM	French overseas departments	Entirely

Table 7: Area fractions and validation sample distributions for the 11 LC classes of the EEA-38+UK Raster Product

Code	CLC+ Backbone raster product- LC classes	Pixel counts	Area fraction	No. of samples
1	Sealed	1760925020	2,94%	3218
2	Woody – needle leaved trees	10072752914	16,80%	4834
3	Woody – Broadleaved deciduous trees	9386257456	15,65%	5553
4	Woody – Broadleaved evergreen trees	2005711473	3,34%	1883
5	Low-growing woody plants (bushes, shrubs)	2645323662	4,41%	3844
6	Permanent herbaceous	17654726350	29,44%	7267
7	Periodically herbaceous	11332319131	18,90%	5539
8	Lichens and mosses	255896911	0,43%	766
9	Non- and sparsely-vegetated	2454299896	4,09%	4104
10	Water	2231737410	3,72%	3653
11	Snow and ice	165557955	0,28%	1352
	Total	59965508178	100,00%	42013

Regarding the **response design**, a photo-interpretation tool was implemented (customization of QGIS software, Figure 39). Different source data sets are used in the interpretation process: Sentinel-2 data, VHR imagery provided by ESA and also Google and Microsoft (Bing). Within available Sentinel-2 data, whenever possible images of 2 seasons are displayed: e.g. summer and winter, to better distinguish certain land cover classes. NDVI time-series are also available for each control point.

During the **first validation stage**, the interpretation of samples is **performed blindly**, meaning that the interpreter has no access to the information from the map product. Subsequently, all samples which show a difference between the blind interpretation and the map class are re-interpreted in a **second stage**, to take the LC interpretation uncertainty due to the 11-class nomenclature vs. the EO data properly into account. Thereby, the interpreter has also access to the map class for the respective sample, and the **interpreter assesses if the map class can be considered plausible**, even if the blind interpretation would have pointed to a different initial class assignment. Further, potential geometric shifts between Sentinel input imagery and VHR reference data are taken into account at this step, such as e.g. if a sample unit mapped as Water is located close to the border of a mapped lake, it can be considered as mapped plausible, even if the VHR validation data appears to indicate, that the plot would be located 10m outside of the lake, if there are indications for a geometric shift between the Sentinel imagery of 10m and the VHR reference imagery. The interpreter considers the pixel at the point together with its context and also takes the nature of the object into account.

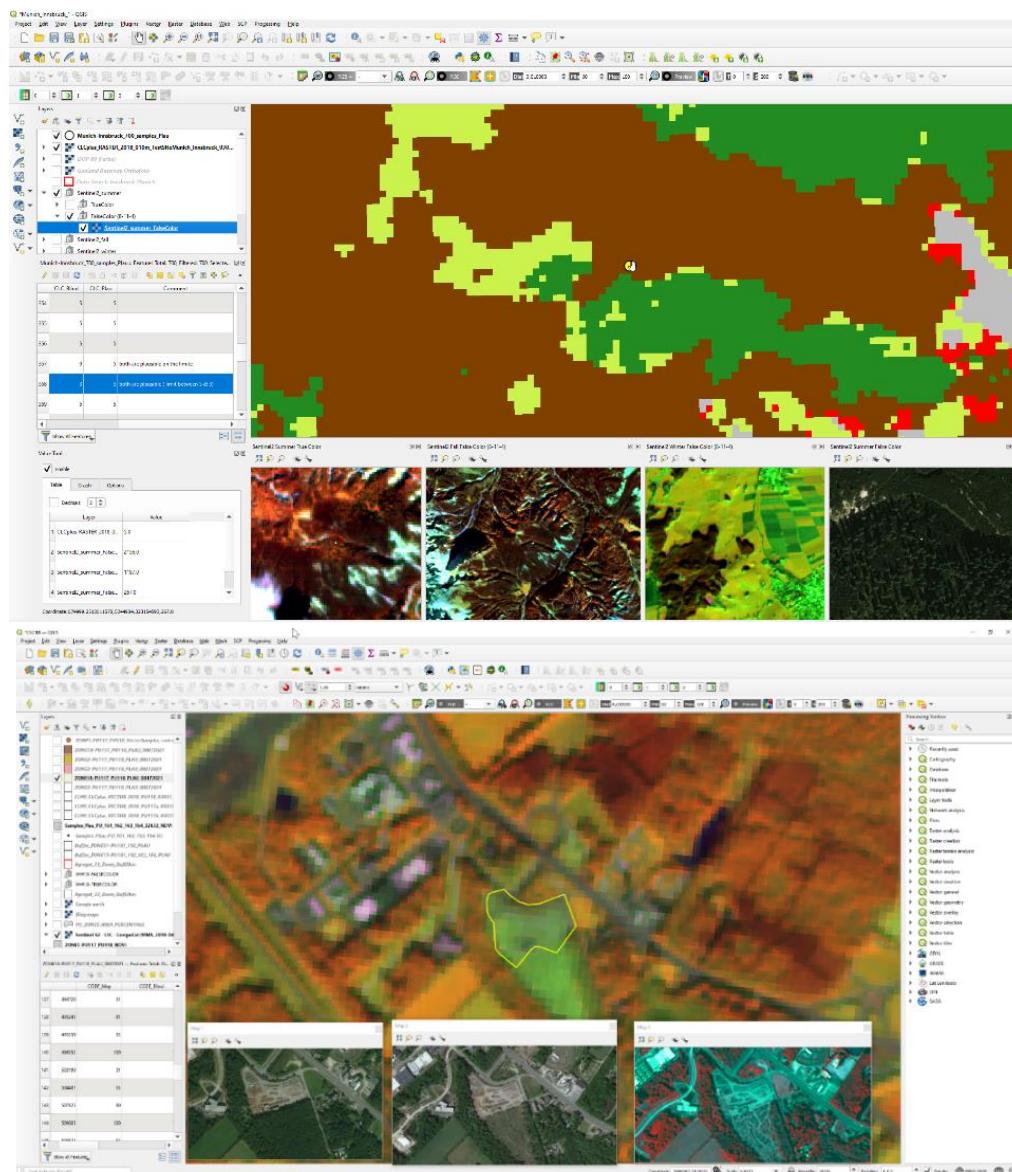


Figure 39: QGIS validation tool to support the interpretation of large numbers of sampling units for Raster Product (above) and Vector Product (below)

The analysis of the interpreted samples is performed using standard procedures for computing the confusion matrices in terms of number of points and respectively estimated area proportions (i.e. area weighted). All subsequently computed metrics for Producer's Accuracy (PA) and the related Omission Error (OE), User's Accuracy (UA) and the related Commission Error (CE), as well as Overall Accuracy (OA) are based on the estimated area proportions. The respective 95% confidence intervals are denoted as ΔPA , ΔUA and ΔOA , respectively. An overview of the computed accuracy metrics is given in Table 8. The resulting accuracy levels are colour coded as summarized in Table 9.

Table 8: Overview of computed accuracy metrics.

Accuracy metric	Abbreviation
Overall Accuracy	OA
Overall Accuracy uncertainty	ΔOA
Omission Error	OE
Producer's Accuracy	PA
Producer's Accuracy uncertainty	ΔPA
Commission Error	CE
User's Accuracy	UA
User's Accuracy uncertainty	ΔUA

Table 9: Colour coding used for the presentation of the different accuracy levels.

Overall Accuracy	> 90 %	85 - 90 %	< 85%
Per Land Cover class	> 85 %	80 - 85 %	< 80 %
Omission and Commission Errors	<15%	15-20%	>20%

4.1.2 Results

This section presents the results of the internal thematic validation of the final EEA-38 + UK Raster Product. The analysis was performed based on 42,013 samples. The Overall accuracy obtained for the entire EEA-38 + UK area is 91.90 % ($\pm 0.30\%$) after plausibility analysis, thus well exceeding the specified 90% on pan-European level.

As shown in Table 10, the area-weighted Overall Accuracies do even exceed 90% also on country (group) level, for all but 3 reporting units after plausibility analysis. For these 3 reporting units (i.e., Italy + Malta, Greece + Cyprus, Portugal), the Overall Accuracies are only marginally below the 90% threshold, reaching 89.9% ($\pm 1.90\%$) in Italy-Malta, 89.9% ($\pm 1.50\%$) in Portugal and 87.9% ($\pm 1.80\%$) in Cyprus-Greece. Considering the Confidence Intervals, validation zone 9 (Italy-Malta) and validation zone 16 (Portugal) actually could be considered to meet 90% Overall Accuracy, while zone 12 (Cyprus-Greece) is still slightly below. Since these areas are known to be especially challenging for land cover classification and validation due to the complex landscape characteristics in the Mediterranean (e.g. mix of evergreen and deciduous forest, mixture of trees and shrubs in the Mediterranean maquis, Eucalyptus plantations typically easily confused with coniferous species), the achieved Overall Accuracies are indeed considered good to excellent also on the country (group) level.

Table 10: Area-weighted Overall Accuracy for CLC+ Backbone Raster Product in percent for the reporting units covering the EEA-38 + UK area

		Overall Accuracy CLC+ Backbone - RASTER PRODUCT			
Zone	Country or group of countries	BLIND ANALYSIS	Confidence interval of 95%	PLAUSIBILITY ANALYSIS	Confidence interval of 95%
Z1	Turkey	74.6%	2.20%	92.20%	1.50%
Z2	France	84.1%	1.9%	94.9%	1.1%
Z3	Spain-Andora	63.80%	2.30%	93.20%	1.20%
Z4	Sweden	84.30%	1.90%	91.90%	1.50%
Z5	Germany	85.40%	1.80%	95.50%	1.00%
Z6	Finland	88.60%	1.60%	94.10%	1.20%
Z7	Norway	78.1%	2.00%	91.60%	1.40%
Z8	Poland	84.30%	2.00%	96.40%	1.00%
Z9	Italy - Malta	69.90%	2.20%	89.90%	1.90%
Z10	United Kingdom - Ireland	81.30%	2.10%	91.30%	1.50%
Z11	Romania	87.50%	1.70%	94.70%	1.10%
Z12	Cyprus-Greece	77.50%	2.30%	87.90%	1.80%
Z13	Bulgaria	84.00%	2.10%	94.10%	1.30%
Z14	Iceland	84.20%	2.00%	91.70%	1.6%
Z15	Hungary	73.30%	2.60%	94.20%	1.20%
Z16	Portugal	51.20%	2.50%	89.90%	1.50%
Z17	Austria - Switzerland - Liechtenstein	80.10%	2.10%	95.90%	1.10%
Z18	Belgium - Denmark - Luxembourg - Netherlands	82.80%	2.00%	91.90%	1.50%
Z19	Albania - Kosovo - N. Macedonia - Montenegro - Serbia	84.10%	1.9%	94.70%	1.20%
Z20	Slovenia - Croatia - Bosnia & Herzegovina	76.90%	2.30%	90.40%	1.6%
Z21	Czech Republic - Slovakia	89.40%	1.50%	97.20%	0.80%
Z22	Latvia - Lithuania - Estonia	82.10%	1.90%	91.60%	1.40%
Z23	French DOMs	93.2%	1.6%	97.60%	0.60%
GLOBAL ACCURACY		77.5%	0.4%	91.9%	0.3%

More differentiated analyses of area-weighted Producer's and User's Accuracies per raster class are presented in Table 11. Considering the plausibility analysis with the respective Confidence Intervals, only two classes feature rates of Omission and Commission Error above the target value of max. 15%.

These exceptions are in particular class 5 (Low-growing woody vegetation), and class 8 (Lichens and Mosses), which both are generally subject to high uncertainties in both, the classification and validation, resulting in the relatively lowest User's and Producer's Accuracies in comparison. These two classes were anticipated from the outset to result in the lowest accuracies. All other classes meet the required 85% User's and Producer's Accuracy, mostly even exceeding 90%.

Table 11: Area-weighted Producer's and User's Accuracies in the CLC+ Backbone Raster Product in percent for the reporting units covering the EEA-38 + UK area.

Code	Legend	Blind analysis					Plausibility analysis						
		Producer (%)	ΔPA (%)	OE (%)	User (%)	ΔUA (%)	CE (%)	Producer (%)	ΔPA (%)	OE (%)	User (%)		
1	Sealed	73,5	2,3	26,5	73,4	1,7	26,6	86,0	1,9	14,0	88,1	1,2	11,9
2	Woody – needle leaved trees	89,4	0,7	10,6	80,9	1,1	19,1	97,0	0,4	3,0	94,4	0,7	5,6
3	Woody – Broadleaved deciduous trees	78,2	1,0	21,8	79,1	1,0	20,9	92,1	0,7	7,9	93,5	0,6	6,5
4	Woody – Broadleaved evergreen trees	67,4	2,4	32,6	58,8	2,5	41,2	86,3	1,8	13,7	89,3	1,4	10,7
5	Low-growing woody plants (bushes, shrubs)	33,4	1,4	66,6	46,8	1,7	53,2	72,1	1,9	27,9	82,9	1,2	17,1
6	Permanent herbaceous	76,9	0,7	23,1	74,9	0,9	25,1	92,3	0,5	7,7	90,1	0,7	9,9
7	Periodically herbaceous	85,7	0,8	14,3	87,0	0,9	13,0	94,3	0,6	5,7	95,0	0,6	5,0
8	Lichens and mosses	31,6	2,9	68,4	60,3	3,4	39,7	67,1	4,8	32,9	76,4	3,1	23,6
9	Non- and sparsely-vegetated	66,1	1,8	33,9	69,6	1,4	30,4	86,9	1,5	13,1	84,2	1,1	15,8
10	Water	94,1	1,1	5,9	96,3	0,6	3,7	96,9	0,9	3,1	98,1	0,4	1,9
11	Snow and ice	95,3	4,5	4,7	58,5	3,4	41,5	99,6	0,8	0,4	85,5	2,0	14,5
Overall Accuracy		77,5 ± 0,4%					91,9 ± 0,3%						

The detailed analyses of the single reporting units for EEA-38 + UK are attached in Annex 4.

4.2 Vector Product Thematic Accuracy

This section provides an overview of the methodology and the assessment results of the CLC+ Backbone Vector Product's thematic accuracy.

4.2.1 Methodological approach

The stratification and sampling design for the CLC+ Backbone Vector Product primarily consists of selecting an appropriate sampling frame and sampling unit. The distribution of samples is planned per **reporting area** (i.e. entire country, or group of countries for the smaller ones). The sampling plan corresponds to that used for the CLC 2018 layer validation.

For the CLC+ Backbone Vector Product, the **stratification** is applied at two levels:

- Stratification according to countries and group of countries greater than 90,000km²
- Stratification based on the CLC+ Backbone Vector Product classes based on the map products

Altogether, 25,632 samples were distributed over the entire EEA-38+UK area.

Regarding the **response design**, the same photo-interpretation tool as in case of the Raster Product (customisation of QGIS software, see Figure 40) was applied. Different source data sets were used in the interpretation process: Sentinel-2 data, VHR imagery provided by ESA and also Google and Microsoft (Bing) as well. Within available Sentinel-2 data, whenever possible, images of 2 seasons were displayed: summer and winter, in order to better distinguish certain land cover classes. For the validation of the map Vector Product, the samples consisted of polygons. NDVI was also available for each polygon, whereby the NDVI profile referred to the average of the pixel values in the polygon over time, see Figure 43.

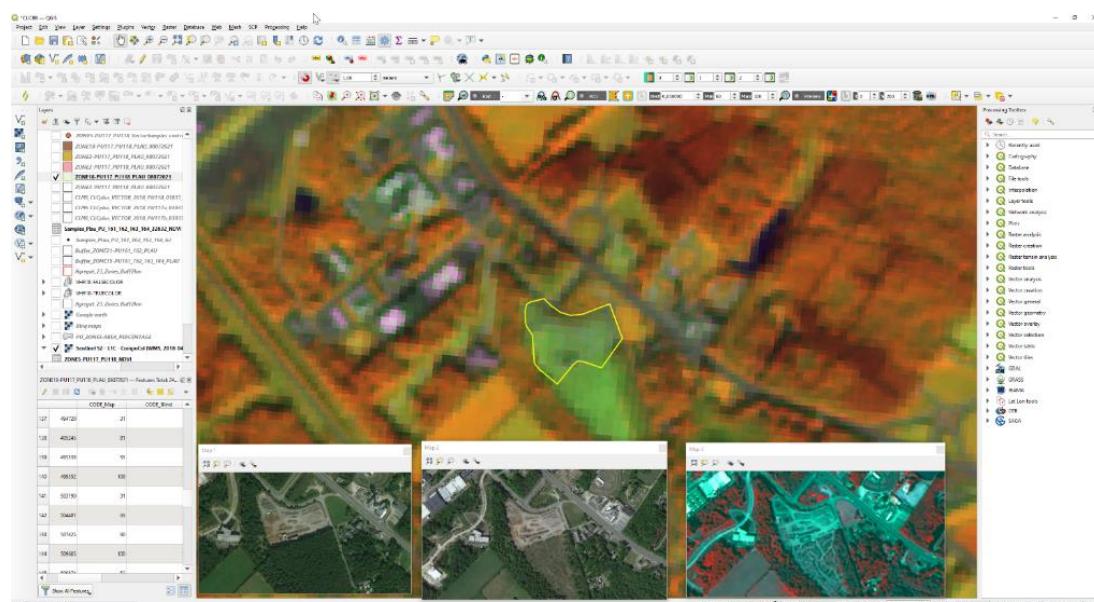


Figure 40: QGIS validation tool supporting the interpretation of large numbers of sampling units

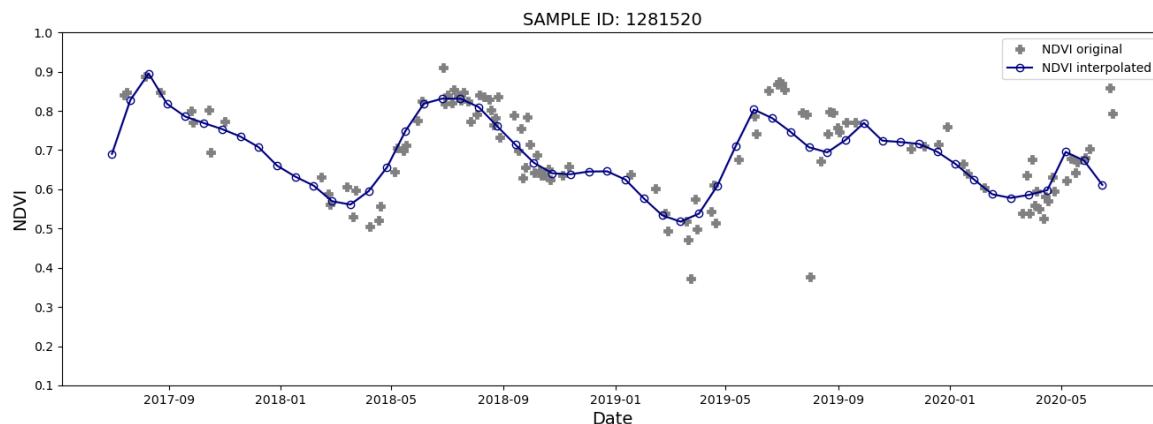


Figure 41: NDVI – Example of spectral profile units

During the **first validation stage**, the interpretation of samples was performed blindly, meaning that the interpreter had no access to the information from the map product. Subsequently, all samples which showed a difference between the blind interpretation and the map class were re-interpreted in a **second stage**. Thereby, the interpreter had also access to the map class for the respective sample, and the interpreter assessed if the map class can be considered plausible, even if the blind interpretation showed a different class assignment. Further, the interpreter considered the whole polygon and the different land cover proportions inside the polygon. The interpreter had the full nomenclature decision tree available together with all technical documentation, to unambiguously apply the defined priority rules to assign the dominant land cover class to each validation sample.

According to the general product specifications, the requested minimum **Overall Accuracy** for the Vector Product is 95%, the minimum individual class accuracies should not be lower than 85% (i.e., Omission Errors and Commission Errors less than 15%).

An overview of the computed accuracy metrics for the Vector Product is given in Table 12. The actually validated accuracy levels are subsequently additionally colour-coded as summarized in Table 13, to allow an intuitive perception of the reached accuracy levels.

Table 12: Overview of computed accuracy metrics.

Accuracy parameter	Abbreviation
Overall accuracy	OA
Overall accuracy uncertainty	ΔOA
Omission error	OE
Producer accuracy	PA
Producer accuracy uncertainty	ΔPA
Commission error	CE
User accuracy	UA
User accuracy uncertainty	ΔUA

Table 13: General requirement: Specified Accuracy thresholds for CLC+ Backbone Vector Product

Overall Accuracy (OA)	> 95 %	90 - 95 %	< 90 %
OA per Land Cover class	> 85 %	80 - 85 %	< 80 %
Omission and Commission Errors	< 15 %	15-20 %	> 20 %

4.2.2 Results

This section provides the internal thematic accuracy assessment results for the entire EEA-38+UK Vector Product, based on 25,632 validation samples.

The pan-European **Overall Accuracy** (OA) obtained was $63.2\% \pm 0.6\%$ for the first validation stage (blind interpretation) and $91.1 \pm 0.4\%$ for the second and decisive validation stage (plausibility analysis). Thus, despite exhibiting a very high Overall Accuracy well above 90%, substantially exceeding all available LC datasets otherwise published in/for Europe, it is a fact that this value does not fully meet the specified 95% OA.

Particularly compared to the results of the thematic Raster Product validation, the present Vector Product validation results show that a split of LC nomenclature from 11 to 18 classes and an aggregation of the pixel-based LC classification into landscape object polygons does not cause the Overall Accuracy to rise from the Raster-based target OA of 90% to a Vector Product-based target OA of 95%. Actually, it rather slightly decreases (cf. section 4.1.2).

In addition to the above pan-European figure, Table 14 summarises Overall Accuracies and associated 95% confidence intervals obtained by zone (i.e., by country or group of countries) for the blind and plausibility interpretations. Whereas in most countries, high Overall Accuracy values of >90% were obtained, there are a few positive and negative outliers: For Finland and Poland, an OA > 95% has been assessed, whereas for Turkey ($86.6\% \pm 1.7\%$), Cyprus-Greece ($81.6\% \pm 2.4\%$) and Iceland ($80.8\% \pm 7.1\%$), the OA values range slightly lower. The former two countries reflect the landscape complexity and heterogeneity as well as the associated higher LC classification confusion potential in the (Eastern) Mediterranean region. Iceland is a special case through its large proportion of complex natural landscapes and sparsely vegetated areas, which are among the most difficult areas in Europe both to map and to validate.

In addition to the Overall Accuracy figures, Figure 15 provides the results of the pan-European detailed assessment of the CLC+ Backbone Vector Product's land cover class-specific Producer's and User's Accuracies, associated Omission and Commission Errors and Confidence Intervals, both for the first validation stage (blind interpretation) and the second validation stage (plausibility analysis).

Table 14: Overall Accuracies for blind and plausibility analyses with Confidence Intervals (95%) for the CLC+ Backbone Vector Product, per country or group of countries for EEA-38+UK area

Zone	Country or group of countries	Overall Accuracy CLC+ Backbone - VECTOR PRODUCT			
		BLIND ANALYSIS	Confidence Interval of 95%	PLAUSIBILITY ANALYSIS	Confidence Interval of 95%
Z1	Turkey	57.3%	2.20%	86.6%	1.7%
Z2	France	57.5%	2.9%	91.3%	1.5%
Z3	Spain-Andorra	56.5%	2.5%	90.4%	1.5%
Z4	Sweden	58.8%	2.7%	94.2%	0.8%
Z5	Germany	69.7%	2.9%	93.6%	1.4%
Z6	Finland	72.1%	2.4%	95.6%	1.00%
Z7	Norway	66.5%	3.00%	94.00%	1.7%
Z8	Poland	66.6%	3.8%	96.1%	2.1%
Z9	Italy - Malta	56.5%	2.7%	92.9%	1.5%
Z10	United Kingdom - Ireland	63.4%	2.8%	91.4%	1.6%
Z11	Romania	66.6%	3.2%	90.7%	2.2%
Z12	Cyprus-Greece	51.4%	2.8%	81.6%	2.4%
Z13	Bulgaria	73.2%	2.8%	93.1%	1.6%
Z14	Iceland	33.00%	7.9%	80.8%	7.1%
Z15	Hungary	69.7%	3.5%	90.8%	2.00%
Z16	Portugal	56.4%	2.6%	90.2%	1.6%
Z17	Austria-Switzerland -Liechtenstein	64.4%	3.1%	91.7%	1.8%
Z18	Belgium - Denmark - Luxembourg - Netherlands	69.1%	3.2%	91.2%	1.5%
Z19	Albania - Kosovo - N. Macedonia - Montenegro - Serbia	71.7%	2.9%	90.6%	2.00%
Z20	Slovenia - Croatia - Bosnia & Herzegovina	78.9%	2.3%	94.5%	1.6%
Z21	Czech Republic - Slovakia	57.7%	3.3%	92.2%	1.9%
Z22	Latvia - Lithuania - Estonia	65.00%	2.6%	91.1%	1.9%
Z23	French Overseas Departments	81.2%	5.00%	91.6%	3.5%
GLOBAL ACCURACY EEA38+UK		63.2%	0.6%	91.1%	0.4%

What the figures show is basically that the 18 LC class-specific **User's Accuracies** are predominantly meeting the 85% criterion. Particularly if taking the Confidence Interval also into account, only the class 53 Permanent Herbaceous with many trees (30-50%) does not entirely reach the requested magnitude ($80.42\% \pm 2.67\%$). This is not surprising, considering the mixed land cover and transitional nature of that LC class, and respective difficulties to unambiguously map and validate it.

The analysis of the validation results for the **Producer's Accuracies** of the Vector Product's LC classes shows that particularly and exclusively the non-pure "mixed" classes by their very definition, i.e. class 12 (Sealing Degree 50-80%), 22+33 (Dominantly Needle-Leaved/Broadleaved 50-75%), 52+53 (Permanent Herbaceous with 10-30% / 30-50% trees), 81+82 (Low/intermediate vegetation cover 10-30% / 30-50%) and 40 Shrubland (which is a mixture by its nature) are those classes, for which the 85% Producer's Accuracy target was not fully met.

Actually, there appears to be a clear Producer's Accuracy "deterioration correlation" with increasing heterogeneity/fuzziness of class mixture definitions, exemplified by the decrease of PA from excellent 92.77% ($\pm 0.77\%$) in case of the rather pure class 51 Permanent Herbaceous (trees $\leq 10\%$), over the mediocre 76.02% ($\pm 2.00\%$) for class 52 Permanent Herbaceous (10-30% trees), to the not satisfying 66.86% ($\pm 3.06\%$) for class 53 Permanent Herbaceous (30-50% trees).

The detailed analyses of the single reporting units for EEA-38 + UK are attached in Annex 5 of this document.

Table 15: Producer's and User's Accuracies of CLC+ Backbone Vector product, given by land cover class – Blind and plausibility analysis for EEA-38+UK area (colour coding for PA: see Table 13).

Code	Legend	Blind analysis						Plausibility analysis					
		Producer (%)	ΔPA (%)	OE (%)	User (%)	ΔUA (%)	CE (%)	Producer (%)	ΔPA (%)	OE (%)	User (%)	ΔUA (%)	CE (%)
11	Very high sealing degree (sealed surfaces > 80 %)	62,62	3,24	37,38	62,76	2,86	37,24	89,77	2,67	10,23	91,81	1,49	8,19
12	High sealing degree (sealed surfaces 50 - 80 %)	41,31	2,95	58,69	51,72	2,95	48,28	73,98	3,59	26,02	87,62	2,03	12,38
21	Woodland – needle leaved trees. Pure needle leaved >75 %	87,35	0,97	12,65	69,44	1,78	30,56	98,41	0,35	1,59	94,78	0,81	5,22
22	Woodland – needle leaved trees. Dominantly needle leaved 50 – 75 %	14,93	1,67	85,07	35,68	3,09	64,32	76,66	3,72	23,34	86,93	2,66	13,07
31	Woodland – broadleaved trees. Pure broadleaved deciduous > 50 %	80,76	1,36	19,24	59,54	1,88	40,46	97,49	0,52	2,51	91,31	0,97	8,69
32	Woodland – broadleaved trees. Pure broadleaved evergreen > 50 %	62,54	3,07	37,46	65,07	3,18	34,93	89,54	2,19	10,46	90,52	1,98	9,48
33	Woodland – broadleaved trees. Dominantly broadleaved 50 - 75 %	12,31	1,45	87,69	32,35	2,81	67,65	77,04	3,64	22,96	88,18	2,51	11,82
40	Shrubland > 50%	38,59	2,01	61,41	53,51	2,38	46,49	76,38	2,24	23,62	87,81	1,59	12,19
51	Permanent herbaceous land (i.e. grasslands) without trees (woody trees <= 10 %)	69,59	1,26	30,41	61,40	1,43	38,60	92,77	0,77	7,23	89,24	0,87	10,76
52	Permanent herbaceous land (i.e. grasslands) with few trees (woody trees 10 - 30 %)	30,13	1,92	69,87	34,82	2,30	65,18	76,02	2,00	23,98	84,05	1,75	15,95
53	Permanent herbaceous land (i.e. grasslands) with many trees (woody trees 30 - 50 %)	17,10	1,91	82,90	27,89	2,75	72,11	66,86	3,06	33,14	80,42	2,67	19,58
60	Periodically herbaceous land (i.e. arable land) > 50%	87,09	0,96	12,91	78,06	1,62	21,94	97,67	0,42	2,33	93,77	0,97	6,23
70	Lichens and mosses land > 50%	25,64	13,59	74,36	18,18	10,69	81,82	89,91	6,67	10,09	86,36	9,92	13,64
81	Low vegetation cover 10 – 30 %	14,13	1,76	85,87	32,86	3,21	67,14	57,70	4,24	42,30	82,25	2,94	17,75
82	Intermediate vegetation cover 30 - 50 %	11,91	2,12	88,09	24,37	3,68	75,63	68,77	5,03	31,23	83,50	3,54	16,50
90	Non-vegetated land (i.e. rock, scree, sand, lichen, permanent bare soil) >= 90 %	55,51	2,78	44,49	65,54	2,96	34,46	88,22	2,26	11,78	90,65	1,90	9,35
100	Water > 50%	94,55	1,40	5,45	97,40	0,79	2,60	98,00	0,85	2,00	99,68	0,28	0,32
110	Snow and ice > 50%	77,56	5,28	22,44	93,67	4,53	6,33	88,41	4,80	11,59	94,94	4,48	5,06
OVERALL ACCURACY		63,20 ± 0,6%						91,10 ± 0,4%					

The following Table 16 provides a comprehensive overview of the Producer's and User's Accuracies for all Vector Product's land cover classes vs. all countries/country groups. Although there have been no specified Producer's and User's Accuracy target values on this (country) disaggregation level, Table 16 has been colour-coded "as if" – in order to increase its intuitive overview value. Basically, it generally confirms the trends already discussed above, i.e. of generally lower accuracies in some Mediterranean countries and Iceland, and generally lower accuracies of LC classes with higher intrinsic "mixing levels". In summary, both at European and at country (group) level, all homogeneous LC classes such as class 11, 21, 31, 32, 51, 60, 90, 100 and 110 have a very good accuracy. In contrast, the heterogeneous classes (containing by definition various land cover elements or components) have lower overall accuracies. This concerns classes 12, 22, 33, 52, 53, 81, 82 and, by definition, also class 40 (Shrubland).



Table 16: Producer's and User's Accuracies of CLC+ Backbone Vector Product, given by land cover class and by country (group) – Plausibility analysis for EEA-38+UK

4.3 Vector Product Geometric Accuracy

This section provides the results of the geometric accuracy assessment of the CLC+ Backbone Vector Products.

4.3.1 Methodological approach

The accuracy of delineation was evaluated applying sampling protocols based on standards of ISO 19157, ISO 2859 and ISO 3951-1:2005. A **probability-based stratified random sampling** was applied considering the size of the segments, assuring a sufficient coverage for all size classes according to their frequency of occurrence. In order to reflect the diversity in European landscapes, the **biogeographical regions** were used as the baseline geographic extent for the sampling design. In addition, the sampling was performed considering the 23 country (group) sub-reporting areas applied also within the thematic validation (sections 4.1, 4.2), as potential subregions, whereas stratification was done based on size criteria of segments within units of unique Production Units, subregions and biogeographical regions.

According to ISO 2859, sample size is defined by population size of the product tested for conformity/non-conformity (with an accuracy quality level of 95%). It defines a stratified sample count of at least 1,250 samples for products / a stratum containing > 0.5 mio units (here: polygons representing landscape elements) as suitable. Proportion of required samples is logarithmic and grouped to the number of units: 800 samples are required for strata containing > 0.15 mio units, 500 samples for strata > 35,000, and 315 samples for strata > 10,000 samples.

Based on the analysis of Europe's Land Cover characteristics as well as previous similar project experience, a stratification of randomly selected landscape object polygons into **three size classes** was regarded appropriate concerning the abovementioned ISO standards. The stratification was implemented by applying classes of equal number, whereas the thresholds were elaborated by size criteria of all segments of each biogeographical region within the individual Production Units.

Figure 42 provides a comprehensive overview of the methodology applied, addressing different aspects of geometric validation considered and described within this section.

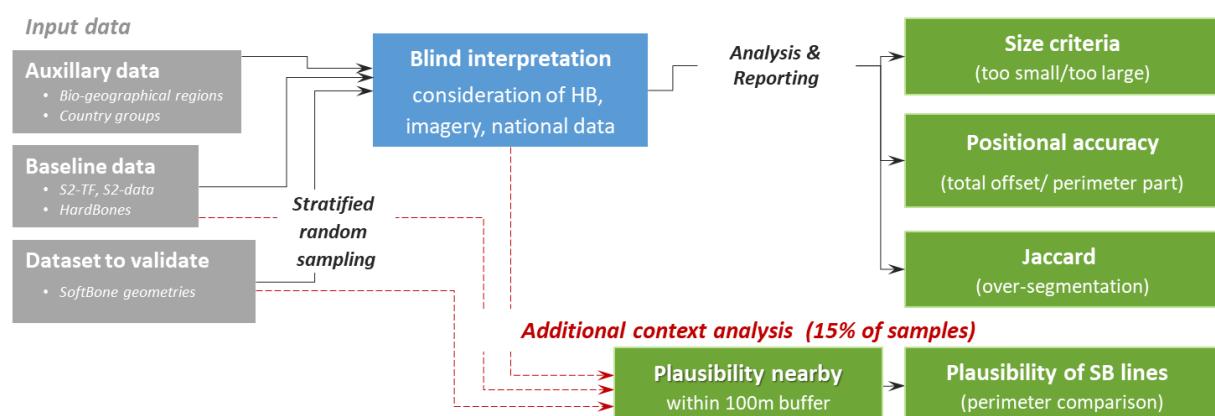


Figure 42: Overview of geometric criteria for CLC+ Backbone vector product validation

Figure 43 shows the overall workflow of blind interpretation. Each sample selected by the described sampling design is checked first by **blind interpretation** and then by **plausibility check**.

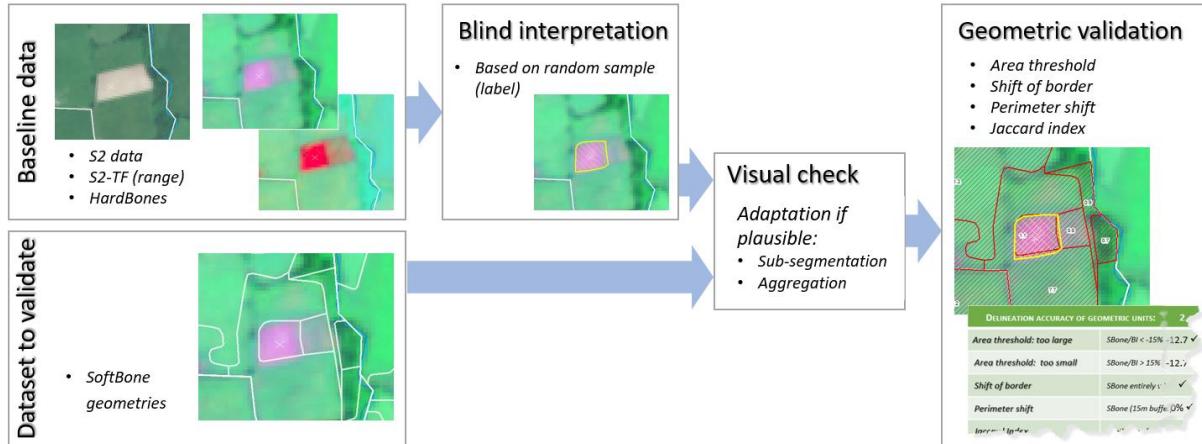


Figure 43: Blind interpretation process

An additional **context analysis** (100m buffer zone) of delineation plausibility along equally distributed and randomly selected samples (~15% of all) was performed. The applied requirements for the geometric validation of the CLC+ Backbone Vector Product are comprehensively shown in Table 17.

Table 17: Geometric accuracy metrics

DELINEATION ACCURACY OF GEOMETRIC UNITS:	
APPROPRIATE SIZE	<ul style="list-style-type: none"> 90% of all segments are not too large 85% of all segments are not too small Acceptable area range defined: 15%
APPROPRIATE DELINEATION (POSITIONAL ACCURACY)	<ul style="list-style-type: none"> Shift of border (modified to 30m) – total segment Y/N Overall perimeter without offset (modified to < 25 m)
JACCARD INDEX	<ul style="list-style-type: none"> Similarity of polygons (1 = highest)
PLAUSIBILITY ANALYSIS	
PLAUSIBLE DELINEATION	<ul style="list-style-type: none"> Overall distribution of elements (HB, SB, area) Estimation of plausible SB 10% of all samples are reviewed (exception LUX= 100%)

As shown in Table 17, in terms of **Delineation Accuracy** of the landscape object polygons, various size, positional accuracy and over-segmentation criteria were assessed:

Size criteria of the Softbone segment and the polygon derived from blind interpretation are compared. The result is evaluated if the size difference lies outside of the acceptable range of 15%, according to overestimation (deviation of ≤10% of all polygons is accepted) and underestimation (deviation of ≤15% of all polygons is accepted). With reference to geometric offset of input data, a range of 1 pixel is considered along the delineation of the provided polygon (delineation ± 5m) and the appropriate area range is considered as suitable reference value.

Appropriate Delineation: positional accuracy is determined by comparing geometric delineation of the selected polygon outlines within a defined buffer range (Figure 44), considering the blind interpretation vector as reference. Estimation of border shift (30m) beyond the given range gets calculated and should not exceed a defined relative proportion (10% of the total segment length).

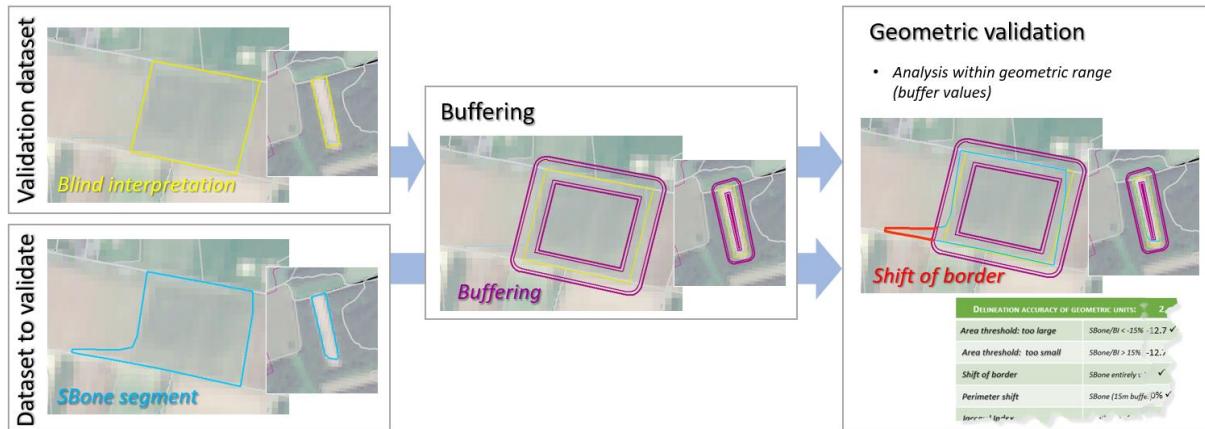


Figure 44: Positional accuracy analysis

Apart from potential geometric errors in the CLC+ Backbone Vector Product, the automatic segmentation process is largely affected by variations significantly related to spatial resolution of the input data and objects. This was considered in an appropriate way within the validation process.

Jaccard Index: Sub-segmentations beyond necessity are not necessarily identified as error (see Figure 45**Error! Reference source not found.**). Such over-segmentation can apply within automatic/process-based image analysis. It often is due to features not necessarily visible to an interpreter within single time step imagery, but relevant within time series classification. They are also happening due to the nature of the segmentation process when small and very large landscape elements intermingle.

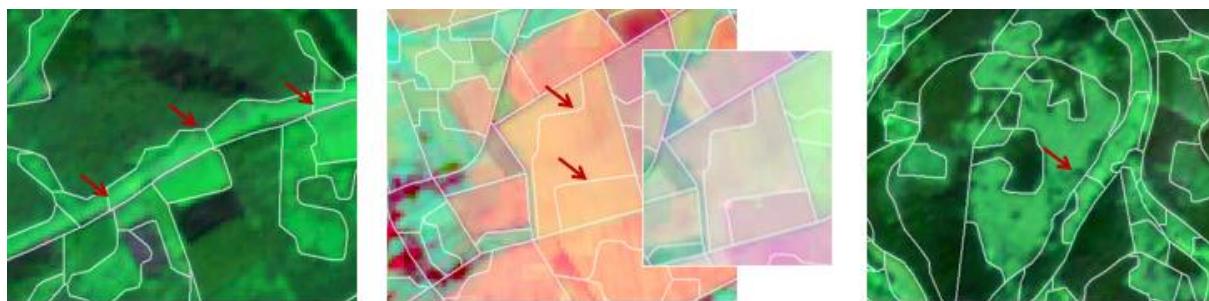
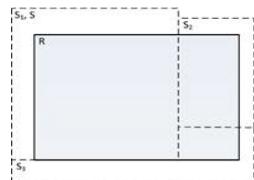


Figure 45: Potential over-segmentation accepted as correct: (a) same land cover, but statistics are likely to be different enough to be mapped separately, (b) slight variation of single feature not applicable for all layers, (c) small objects below MMU, that could be detected as separate class

The Jaccard Index (Jaccard 1901), also known as Intersection over Union, is used as a baseline metric, amended to be invariant to over-segmentation (over-segmentation independently considered via imposing MMU filtering and tuning CC (Coefficient of Community) appropriately). Since automatic CLC+ Backbone Product Specification and User Manual

segmentation approach considers a wide range of numeric values, not always recognizable by visual means, a certain (and plausible) over-segmentation is likely to happen and needs to be considered as effect, not an error. Figure 46 **Error! Reference source not found.** illustrates a reference level 2 Landscape Object (R), with a generated segment having the most overlap (S), or segments with some overlap (1, 2, 3), denoted by dotted lines.



The Jaccard Index, invariant to over-segmentation, is written as:

$$\text{Jaccard}_{\text{Io}} = \frac{|R \cap \sum S_i|}{|R \cup \sum S_i|} \quad (\text{Equation 1})$$

Figure 46: Graphic illustration of segment evaluation for Jaccard index

With $\sum S_i$ in Equation 1 denoting all segments with at least a 50% area overlap with the reference Landscape Object. The range of the index is [0, 1], with 1 indicating a perfect match with the reference Landscape Object geometry. Subjectively, a metric score of over 0.85 may be considered a very good geometric delineation.

In terms of further **Plausibility Analysis**, respective plausibility checks were additionally performed for the close surroundings of the random samples checked via blind interpretation. Plausibility analysis applies exclusively for Softbone geometries that lie within a 100 m distance, excluding the sample itself (Figure 47). It aims to provide a further independent metric on the overall polygon geometric accuracy. For this analysis, around 18% of the initial sample amount were randomly selected, representing all 3 size classes equally (in total 3,232 buffer regions).

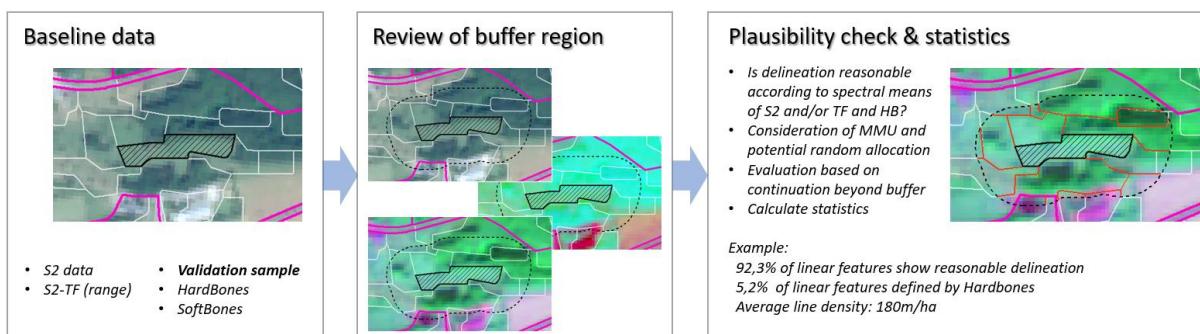


Figure 47: Geometric plausibility check via visual interpretation

Following the listed criteria, line features are classified as plausible and/or non-plausible. For each site the length is reported and aggregated by statistical means (defined by 100m buffer):

- Content and distribution of Hardbones and Softbones,
- Length and average density of line features considered,
- Overall plausibility value of Softbones [%], as well as Summarized plausibility measures for the total reporting region.

4.3.2 Results

The overall results of the internal geometric validation of the CLC+ Backbone Vector Product's delineation accuracy, along the abovementioned criteria and metrics, are presented in Table 18, summarised over all European biogeographical regions (BGR), however excluding the French DOMs (figures for which are given in Table 19 and Annex 6).

Table 18: Results of geometric validation of the CLC+ Backbone Vector Product (summary, excluding TRF, DOM)

Sampling design		
No. of samples (equally distributed)		14,934
No. of samples within size classes		5,185 small / 5,116 medium / 5,186 large
Delineation accuracy of geometric units:		
Appropriate size	Defined threshold of landscape elements that can be 15% larger or smaller	Validation result
too large: (+/- 5m tolerance)	≤ 10% (≤ 1,493 Softbones)	11.1 % (1724) 2.3 % (344)
too small: (+/- 5m tolerance)	≤ 15% (≤ 2,240 Softbones)	20.4 % (3163) 7.5 % (1,131)
Appropriate delineation		
Shift of border within 30m buffer (at least once beyond threshold)		41.0 % (6357)
Perimeter shift beyond 25m buffer (proportion located outside)		23.4 % (3626)
Jaccard Index		
Overall Index (range of single landscape elements)		0.80 %

The values represent an aggregation of the partial results of the geometric validation performed along biogeographical regions. According to these figures, the product essentially meets the specified overall geometric requirements on pan-European level. As expected, a significant range in acceptance level of the requirements applies for absolute size criteria. Regional exceedances exist, especially within the biogeographical regions outside of Europe, see Annex 6. Reviewing all results within the individual biogeographical regions, it can be stated, that:

- Identification and/ or delineation of landscape elements of natural features is in general more difficult. Validation results related to 30m/25m buffer analysis often exceed this criterion, since they are independently reported according to their percentage (shift of border). So, in many of the expressed cases, just small portions of the Softbone vector do not fulfil this requirement.
- Time features are affected by the intrinsic geometric accuracy of the Sentinel-2A and Sentinel-2B time series data. In some cases, especially when steep and fragmented terrain is involved, pre-processing aiming to level illumination effects (i.e. topographic normalisation) as part of the standard Level 2A processing algorithms in turn significantly impacts the resulting spectral information that is the basis for the Softbone generation. Over-corrections (or, to a lesser extent, under-corrections) directly propagate into the Softbone.

- Regions with scarce vegetation (e.g. Anatolian BGR) or with permanent vegetation (e.g. Tropical Rainforest) show lower confidence in Softbone generation. Here, the low temporal variance does not support identification of landscape elements well (cf. Annex 6, TRF)
- A regionally high density of Hardbone elements typically indicates a high human influence on the landscape. In such regions, Softbone elements are often of higher geometric quality. In rare occasions it may happen that Hardbones are not located correctly, resulting in confusion of segmentation along the defined Hardbone skeletons.
- Aggregation of patches below the given MMU of 0.5 ha to larger segments substantially influences the Softbone delineation. Without further knowledge of the small-segment generalization process, independent blind validation may get to different conclusions, hence affecting the validation result.
- For the validation process, those Time Features, which had been considered as most important during the Softbone segmentation, were provided. Within each of the Softbone delineation decisions, they were analysed and weighted accordingly. While performing blind interpretation, band combinations of the latter were helpful additional interpretation support data sources. However, different prioritisation and stretching may also impact the delineation results and thus the overall validation result.

Table 19 below provides an overview of the regional distribution of the above geometric validation of the CLC+ Backbone Vector Product's delineation accuracy, for all biogeographical regions of the EEA-38 + UK. More detailed values per biogeographical region are documented in Annex 6.

Table 19: CLC+ Backbone Vector Product geometric validation results per biogeographic region (relative percentage values)

Code	Defined threshold of landscape elements that can be 15% larger or smaller				Shift of border within 30m buffer (at least once beyond threshold)	Perimeter shift beyond 25m buffer	Mean Jaccard Index			
	too large (< 10% of Softbones)		too small (< 15% of Softbones)							
	value	(± 5m)	value	(± 5m)						
ALP	9.9%	1.7%	20.9%	8.3%	41.1%	24.8%	0.79			
ANA	5.6%	0.9%	21.9%	5.1%	35.1%	15.9%	0.78			
ARC	16.3%	3.9%	18.4%	6.0%	55.7%	35.8%	0.82			
ATL	16.0%	2.8%	19.5%	9.1%	44.1%	29.3%	0.81			
BLS	8.4%	0.6%	17.8%	3.2%	31.3%	12.7%	0.80			
BOR	11.7%	2.1%	36.3%	18.8%	56.0%	38.8%	0.74			
CON	10.6%	2.4%	17.2%	5.6%	34.9%	18.1%	0.82			
MAC	8.8%	3.1%	14.4%	9.1%	32.9%	16.9%	0.85			
MED	8.6%	0.9%	22.7%	7.2%	38.8%	19.1%	0.81			
PAN	5.4%	0.3%	18.6%	4.7%	35.8%	16.6%	0.81			
STE	11.5%	0.6%	16.2%	4.2%	38.1%	19.5%	0.81			
Total cont. Europe	10.3%	1.8%	20.3%	7.4%	40.4%	22.5%	0.80			
DOM	33.9%	19.5%	28.4%	13.0%	77.2%	68.0%	0.68			
TRF	30.0%	17.2%	15.6%	8.2%	53.3%	44.0%	0.68			
Total	13.6%	4.3%	20.6%	7.9%	44.2%	27.7%	0.78			

The assessed geometric accuracy largely meets the desired accuracy levels concerning absolute size criteria. Allowing a $\pm 5\text{m}$ tolerance reduces the number of samples exceeding the required quality criterion significantly. As for the appropriate size check, the results indicate that even without this $\pm 5\text{m}$ tolerance, only 10.3 % of the validated segments were too large and 20.3 % segments were too small. Positional accuracy of shift border shows that the borders of 59.6 % of segments are completely inside a 30m buffer compared to reference borders mapped manually. Among all assessed Softbone segments within Europe, 77.5 % had no shift or less than 25m shift in perimeter when 25 m buffer was applied (whereas 22.5 % showed such shift).

Limitations in positional accuracies are not rare, but with the type of time series analysis necessary to derive all required CLC+ Backbone products, larger impacts were to be expected. They were however found to be of partially different nature within the different biogeographical regions. Main driving aspect is related to the technical specification of the input data used in the automatic segmentation process, namely Sentinel-2 data with typically 10m geometric intrinsic time series imprecision as well as the variation in the spatial resolutions of spectral bands applied. Particularly 20m bands add necessary spectral information, but also bring more geometric fuzziness in terms of stacked multitemporal data (Time Features), which are a basic element of the analysis.

The Jaccard Index for the CLC+ Backbone Vector Product was measured with an average of 0.80 for continental Europe (0.78 incl. the DOMs), with 1.0 theoretically representing a perfect match with the reference polygons and 0.85 considered as a very good geometric delineation. The obtained figure may be interpreted such that the reached overall geometric delineation accuracy is appropriate.

Plausibility analysis of the context

Table 20 provides the result of the additional context analysis over entire Europe summarised from the reports prepared on the single biogeographical regions (Annex 6), whereas Table 21 provides the respective details per biogeographical region.

Table 20: Results of plausibility analysis of the 100m buffer region for EEA-38+UK

Sampling design	
No. of samples (equally distributed)	2,910
No. of samples within size classes	970 small/ 946 medium/ 994 large
Delineation plausibility within direct context (100m buffer)	
Sample characteristics (summary)	
Total area checked [ha] (buffer region without sample extent)	33,796 ha
... containing Hardbone [km (%)]	1,267 km (29 %)
... containing Softbone [km (%)]	2,775 km (69 %)
Average line density [km/ha]	0.12 km/ha
Plausibility check	
Plausible Softbone lines within buffer region [%]	2,263 km (80.5 %)

It can be stated that plausibility of the validation interpreters' Softbone delineation as compared to the automatically generated Softbone lines within the reviewed buffer regions ranges on average at 80%. The respective plausibility percentages for the individual biogeographical regions are generally high, with a maximum in the Macaronesian islands. Here, the Hardbone content (percentage) is relatively highest, and consequently, the Softbone content and plausibility is also comparably high. Detailed figures are given in Annex 6.

Table 21: Overview of buffer zone check of 100m buffer of approx. 20% of the samples

Code	No. of samples	Total area checked [ha]	Hardbone content		Softbone content		Average line density [km/ha]	Plausible Softbone lines within buffer	
			[km]	[%]	[km]	[%]		[km]	[%]
ALP	276	2,800	81.6	23%	266.7	77%	0.12	217.4	79.3%
ANA	259	2,770	64.1	1%	267.0	81%	0.12	234.4	86.4%
ARC	260	4,232	45.0	13%	297.2	87%	0.08	273.7	91.5%
ATL	283	3,150	155.6	37%	263.8	63%	0.13	191.8	72.2%
BLS	266	2,800	81.6	23%	266.7	77%	0.12	217.4	79.3%
BOR	260	2,995	71.3	21%	269.00	79%	0.11	231.2	85.9%
CON	276	2,910	145.00	35%	266.5	65%	0.14	203.00	74.5%
MAC	243	2,798	203.3	58%	146.5	42%	0.13	136.2	92.4%
MED	281	3,118	143	34%	281.5	66%	0.14	243.4	86.8%
PAN	254	2,923	125.1	34%	238.1	66%	0.12	176.8	76.0%
STE	252	3,300	151.3	42%	211.7	58%	0.11	137.5	60.7%
Total Europe	2,910	33,796	1,267	29%	2,775	69%	0.12	2,263	80.5%
DOM	79	781	48.1	19%	73.8	81%	0.16	62.4	79.9%
TRF	243	6,729	54.00	39%	230.9	61%	0.04	172.1	76.0%
Total	3,232	41,306	1,369	29%	3,079	69%	0.12	2,497	80.1%

5 Terms of Use and Product Technical Support

5.1 Terms of Use

The product(s) described in this document is/are created in the frame of the Copernicus programme of the European Union by the European Environment Agency (product custodian) and is/are owned by the European Union. The product(s) can be used following Copernicus full free and open data policy, which allows the use of the product(s) also for any commercial purpose. Derived products created by end users from the product(s) described in this document are owned by the end users, who have all intellectual rights to the derived products.

5.2 Citation

In cases of re-dissemination of the product(s) described in this document or when the product(s) is/are used to create a derived product it is required to provide a reference to the source. A template is provided below:

"© European Union, Copernicus Land Monitoring Service <year>, European Environment Agency (EEA)"

5.3 Product Technical Support

Product technical support is provided by the product custodian through Copernicus Land Monitoring Service helpdesk at copernicus@eea.europa.eu. Product technical support doesn't include software specific user support or general GIS or remote sensing support.

6 References

- Bosilj, P., Kijak, E., & Lefèvre, S. (2018). Partition and inclusion hierarchies of images: A comprehensive survey. *Journal of Imaging*, 4(2), 33.
- Cochran, W.G. (1977). Sampling Techniques. 3rd Edition, John Wiley & Sons, New York.
- Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1), 35-46.
- EUROSTAT (2018). LUCAS survey for 2018 <https://ec.europa.eu/eurostat/de/web/lucas/data/primary-data/2018> [2021-06-25]
- Havel, J., Merciol, F., & Lefèvre, S. (2019). Efficient tree construction for multiscale image representation and processing. *Journal of Real-Time Image Processing*, 16(4), 1129-1146.
- Jaccard, P. (1901). Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines. *Bull Soc Vaudoise Sci Nat*, 37, 241-272.
- Jenks, G. F. (1967). The data model concept in statistical mapping. *International yearbook of cartography*, 7, 186-190.
- Metzger, M. J., Bunce, R. G., Jongman, R. H., Sayre, R., Trabucco, A., & Zomer, R. (2013). A high-resolution bioclimate map of the world: a unifying framework for global biodiversity research and monitoring. *Global Ecology and Biogeography*, 22(5), 630-638.
- Olofsson, P., Foody, G. M., Stehman, S. V., & Woodcock, C. E. (2013). Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment*, 129, 122-131.
- Plourde, L., & Congalton, R. G. (2003). Sampling method and sample placement. *Photogrammetric Engineering & Remote Sensing*, 69(3), 289-297.
- Qiu, S., Zhu, Z., & He, B. (2019). Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4–8 and Sentinel-2 imagery. *Remote Sensing of Environment*, 231, 111205.
- Schindler, K. (2012). An overview and comparison of smooth labeling methods for land-cover classification. *IEEE Transactions on Geoscience and Remote Sensing*, 50(11), 4534-4545.

ANNEXES

ANNEX 1: NAMING CONVENTIONS

The following file naming convention was applied to both CLC+ Backbone Raster and Vector products. All letters except the THEME descriptor are in small (not capital) letters, and no points (".") and/or minus (" -") within file names. The file naming is based on the following descriptors:

CLMS_CLCplus_[PRODUCT ACRONYM]_[REFERENCE YEAR]_[RESOLUTION]_[EXTENT]_[EPSG]_[VERSION]

PRODUCT ACRONYM:

- Acronym of the respective product

CONF	Raster Confidence Layer
DSL	Data Score Layer
POST	Post-processing Layer
RASTER	Raster product 11 classes
VECTOR	Vector product

The Segmentation Confidence Layer is not listed here, as it was added as an attribute to the Vector product.

REFERENCE YEAR

- 2018 in four digits

RESOLUTION (in case of raster file)

- Four-digit (e.g. 010m)

EXTENT

- 5-digit unique id of the production / delivery unit
- "eu" for deliveries with full coverage of EEA-38 + UK

EPSG

- 5-digit EPSG code (geodetic parameter dataset code by the European Petroleum Survey Group), see <http://www.epsg-registry.org/>
- e.g. "03035" for the European LAEA projection

VERSION

- 4-digit qualifier of the version number, starting with "V1_0" for a first full final version, and allowing to capture re-processing/calculation of small changes as ("V1_1", "V1_2" etc.). In case of major changes, a second version should be used ("V2_0")

Example for first intermediate delivery:

CLMS_CLCplus_RASTER_2018_010m_PU117_03035_V1_0

Example for EEA-38 + UK delivery:

CLMS_CLCplus_VECTOR_2018_eu_03035_V1_0

ANNEX 2: COLOUR PALETTES

The CLC+ Backbone products are delivered with the following colour maps. For all raster the colour maps are embedded into the GeoTIFF and provided additional as *.clr for GIS colour palettes. For the Vector Product the colour maps are provided as ArcGIS-compatible *.lyr and *.lyrx files.

RASTER PRODUCT

Table 22: Colour palette for the Raster Product

Class name	Class code	Colour	
		R,G,B	Palette
Sealed	1	255, 0, 0	
Woody – needle leaved trees	2	34, 139, 34	
Woody – Broadleaved deciduous trees	3	128,255,0	
Woody – Broadleaved evergreen trees	4	0, 255, 8	
Low-growing woody plants (bushes, shrubs)	5	128,64,0	
Permanent herbaceous	6	204, 242, 77	
Periodically herbaceous	7	255, 255, 128	
Lichens and mosses	8	255, 128, 255	
Non- and sparsely-vegetated	9	191,191,191	
Water	10	0, 128, 255	
Snow and ice	11	0,255,255	
outside area	254	230, 230, 230	
NoData (NOT allowed in the final product)	255	0,0,0	

VECTOR PRODUCT
Table 23: Colour palette for the Vector Product

Class name	Class code	Colour	Palette
		R,G,B	
Very High Sealing Degree	11	230, 0, 77	
High Sealing Degree	12	255, 0, 0	
Pure needle leaved	21	0, 166, 80	
Dominantly needle leaved	22	0, 193, 80	
Pure broadleaved deciduous	31	109, 212, 0	
Pure broadleaved evergreen	32	79, 154, 0	
Dominantly broad leaved	33	128, 255, 0	
Shrubland	40	166, 242, 0	
Permanent herbaceous without trees	51	186, 187, 77	
Permanent herbaceous with few trees	52	211, 212, 77	
Permanent herbaceous with many trees	53	230, 230, 77	
Periodically herbaceous	60	255, 255, 168	
Lichens and Mosses	70	166, 166, 255	
Partly vegetated land - Low vegetation cover	81	204, 255, 204	
Partly vegetated land - Intermediate vegetation cover	82	147, 255, 147	
Non-vegetated land	90	204, 204, 204	
Water	100	128, 242, 230	
Snow and Ice	110	166, 230, 204	

DATA SCORE LAYER

Table 24: Colour palette for the Data Score Layer

Number of cloud free observations	Colour	
	R,G,B	Palette
> 200	0, 97, 0	
150	122, 171, 0	
100	255, 255, 0	
50	255, 153, 0	
0	255, 34, 0	

Compared to Europe, the input data availability for the DOMs is much lower in general. Therefore, a separate colour palette was created for the DOMs, which stretches the same colour palette over a smaller range of values (see Table 25).

Table 25: Adapted colour palette for the Data Score Layer of the DOMs

Number of cloud free observations	Colour	
	R,G,B	Palette
85	0, 97, 0	
65	122, 171, 0	
42	255, 255, 0	
21	255, 153, 0	
0	255, 34, 0	

RASTER CONFIDENCE LAYER

Table 26: Colour palette for the Raster Confidence Layer

Confidence value [%]	Colour	
	R,G,B	Palette
100	12, 47, 122	
80	32, 153, 143	
60	0, 219, 0	
40	255, 255, 0	
20	237, 161, 19	
0	194, 82, 60	

POST-PROCESSING LAYER

Table 27: Colour palette for the Raster Post-processing layer

Class name	Class code	Colour	
		R,G,B	Palette
No change during post-processing	0	240, 240, 240	
Recoded during post-processing	1	112, 168, 0	
Outside area	254	0, 0, 0	
NoData	255	0, 0, 0	

SEGMENTATION CONFIDENCE LAYERS

Table 28: Colour palettes for the Segmentation Confidence Layer 1 (above) and 2 (below)

Confidence layer 1 (CONF_TEX)	Colour	
	R,G,B	Palette
0 - 20	11, 45, 122	
20 - 30	28, 137, 146	
30 - 40	17, 193, 76	
40 - 50	108, 235, 0	
50 - 60	249, 223, 7	
60 - 70	232, 146, 25	
70 - Max	196, 86, 57	
NoData	0, 0, 0	

Confidence layer 2 (CONF_SPEC)	Colour	
	R,G,B	Palette
0 - 10	11, 45, 122	
10 - 20	28, 137, 146	
20 - 30	17, 193, 76	
30 - 40	108, 235, 0	
40 - 50	249, 223, 7	
50 - 60	232, 146, 25	
60 - max	196, 86, 57	
NoData	0, 0, 0	

ANNEX 3: PROJECTION PARAMETERS

Except for French DOMs and the national products, the primary products are produced and delivered in the European LAEA projection (EPSG: 3035), which is defined according to the following WKT:

```
PROJCS["ETRS89-extended / LAEA Europe",
    GEOGCS["ETRS89",
        DATUM["European_Terrestrial_Reference_System_1989",
            SPHEROID["GRS 1980",6378137,298.257222101,
                AUTHORITY["EPSG","7019"]],
            AUTHORITY["EPSG","6258"]],
        PRIMEM["Greenwich",0,
            AUTHORITY["EPSG","8901"]],
        UNIT["degree",0.0174532925199433,
            AUTHORITY["EPSG","9122"]],
        AUTHORITY["EPSG","4258"]],
    PROJECTION["Lambert_Azimuthal_Equal_Area"],
    PARAMETER["latitude_of_center",52],
    PARAMETER["longitude_of_center",10],
    PARAMETER["false_easting",4321000],
    PARAMETER["false_northing",3210000],
    UNIT["metre",1,
        AUTHORITY["EPSG","9001"]],
    AXIS["Northing",NORTH],
    AXIS["Easting",EAST],
    AUTHORITY["EPSG","3035"]]
```

ANNEX 4: THEMATIC ACCURACIES FOR THE RASTER UNITS OF EEA-38 + UK

ZONE 1 – TURKEY

Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone1

Overall Accuracy (%)	ΔOA (%)
92.20	1.50

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone1

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer	84,48	92,03	90,32	73,39	88,87	96,99	87,69	-	88,14	100,00	100,00
Error on	9,78	3,96	5,12	17,87	6,77	1,09	3,98	-	6,58	-	-
Omission	15,52	7,97	9,68	26,61	11,13	3,01	12,31	-	11,86	-	-
User	92,86	97,56	95,94	86,06	86,67	89,05	96,65	-	87,88	97,14	58,57
Error on	4,33	2,03	2,68	5,58	5,32	3,10	2,07	-	4,29	2,80	10,66
Commission	7,14	2,44	4,06	13,94	13,33	10,95	3,35	-	12,12	2,86	41,43

ZONE 2 – FRANCE

Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 2

Overall Accuracy (%)	ΔOA (%)
94.90	1.10

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 2

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy	94,14	95,99	95,46	89,43	66,23	97,69	98,97	-	67,13	99,98	100,00
Error on Producer	5,27	2,99	2,02	13,49	9,06	1,43	0,96	-	13,83	0,04	-
Omission error (%)	5,86	4,01	4,54	10,57	33,77	2,31	1,03	-	32,87	0,02	-
User accuracy (%)	92,14	94,21	95,12	84,85	86,50	93,69	98,40	-	93,94	97,14	96,91
Error on User	4,51	3,32	2,43	5,90	5,01	2,70	1,38	-	3,29	2,79	3,50
Commission_error	7,86	5,79	4,88	15,15	13,50	6,31	1,60	-	6,06	2,86	3,09

ZONE 3 – SPAIN AND ANDORRA

Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 3

Overall Accuracy (%)	ΔOA (%)
93.20	1.20

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 3

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer acc	96,68	100,00	97,41	93,59	93,87	88,28	94,58	-	93,98	89,65	100,00
Error on Pro	4,54	-	2,67	4,17	2,66	2,57	2,07	-	6,05	12,94	-
Omission er	3,32	-	2,59	6,41	6,13	11,72	5,42	-	6,02	10,35	-
User accura	96,43	96,57	97,86	94,55	96,27	97,32	89,22	-	65,85	97,14	71,43
Error on Us	3,11	2,54	2,39	3,03	2,23	1,67	3,73	-	7,54	2,78	8,85
Commission	3,57	3,43	2,14	5,45	3,73	2,68	10,78	-	34,15	2,86	28,57

ZONE 4 – SWEDEN
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 4

Overall Accuracy (%)	ΔOA (%)
91.90	1.50

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 4

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	74,44	96,93	77,78		38,53	90,49	98,01	67,94	65,64	99,82	100,00
Error on Producer accuracy (%)	19,70	1,10	7,55		14,37	3,25	3,71	7,40	12,35	0,30	-
Omission error (%)	25,56	3,07	22,22		61,47	9,51	1,99	32,06	34,36	0,18	-
User accuracy (%)	89,39	94,14	83,43		63,79	88,96	96,72	88,66	75,32	99,51	82,01
Error on User accuracy (%)	5,44	2,17	5,67		13,06	3,41	3,20	5,80	6,70	0,94	7,05
Commission_error (%)	10,61	5,86	16,57		36,21	11,04	3,28	11,34	24,68	0,49	17,99

ZONE 5 – GERMANY
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 5

Overall Accuracy (%)	ΔOA (%)
95.50	1.00

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 5

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy	90,32	99,61	95,05	-	54,37	93,61	98,01	-	71,03	99,53	100,00
Error on Producer	5,31	0,76	2,49	-	16,43	2,34	1,22	-	27,21	0,56	-
Omission error (%)	9,68	0,39	4,95	-	45,63	6,39	1,99	-	28,97	0,47	-
User accuracy (%)	92,11	94,72	95,52	-	82,50	94,48	97,75	-	86,71	99,29	100,00
Error on User	3,78	2,76	2,45	-	6,23	2,45	1,52	-	5,60	1,39	-
Commission_error	7,89	5,28	4,48	-	17,50	5,52	2,25	-	13,29	0,71	-

ZONE 6 – FINLAND
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 6

Overall Accuracy (%)	ΔOA (%)
94.10	1.20

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 6

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accu	82,91	96,13	87,46		26,85	94,07	98,58	83,72	66,64	98,92	100,00
Error on Produ	19,18	1,17	6,28		10,53	2,80	2,69	22,44	17,98	1,87	-
Omission error	17,09	3,87	12,54		73,15	5,93	1,42	16,28	33,36	1,08	-
User accuracy	89,21	97,53	83,42		73,56	86,44	94,29	92,17	86,67	100,00	64,29
Error on User	5,33	1,39	5,30		9,49	4,30	3,94	4,20	4,60	-	9,90
Commission_error	10,79	2,47	16,58		26,44	13,56	5,71	7,83	13,33	-	35,71

ZONE 7 – NORWAY
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 7

Overall Accuracy (%)	ΔOA (%)
91.60	1.40

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 7

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	100,00	97,84	86,86	-	51,89	92,59	88,27	60,21	96,67	99,37	100,00
Error on Producer accuracy (%)	-	1,57	4,67	-	12,51	1,80	13,46	8,57	2,00	1,17	-
Omission error (%)	-	2,16	13,14	-	48,11	7,41	11,73	39,79	3,33	0,63	-
User accuracy (%)	87,14	98,47	90,63	-	88,04	91,20	97,86	83,03	81,68	99,50	88,57
Error on User accuracy (%)	5,94	1,47	3,73	-	3,96	2,70	2,41	5,69	4,81	0,97	5,60
Commission_error (%)	12,86	1,53	9,38	-	11,96	8,80	2,14	16,97	18,32	0,50	11,43

ZONE 8 – POLAND
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 8

Overall Accuracy (%)	ΔOA (%)
96.40	0.90

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 8

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	93,94	98,75	95,31	-	77,57	93,37	98,59	-	64,90	96,46	-
Error on Producer accuracy (%)	6,00	1,02	2,36	-	29,71	2,61	0,96	-	23,72	5,89	-
Omission error (%)	6,06	1,25	4,69	-	22,43	6,63	1,41	-	35,10	3,54	-
User accuracy (%)	82,14	99,18	97,63	-	51,68	95,45	96,82	-	82,42	100,00	-
Error on User accuracy (%)	6,64	1,10	1,75	-	10,95	2,18	1,92	-	6,26	-	-
Commission_error (%)	17,86	0,82	2,37	-	48,32	4,55	3,18	-	17,58	-	-

ZONE 9 – ITALY – MALTA
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 9

Overall Accuracy (%)	ΔOA (%)
89.80	1.50

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 9

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	87,88	87,78	96,56	81,49	68,26	89,57	93,01	-	94,47	89,59	100,00
Error on Producer accuracy (%)	6,92	6,95	1,82	6,25	8,03	2,54	2,68	-	5,87	9,40	-
Omission error (%)	12,12	12,22	3,44	18,51	31,74	10,43	6,99	-	5,53	10,41	-
User accuracy (%)	89,29	95,71	90,06	84,50	81,21	85,11	97,46	-	91,79	98,57	100,00
Error on User accuracy (%)	5,22	3,31	3,41	5,09	5,92	3,88	1,80	-	3,98	1,94	-
Commission_error (%)	10,71	4,29	9,94	15,50	18,79	14,89	2,54	-	8,21	1,43	-

ZONE 10 – UNITED KINGDOM – IRELAND
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 10

Overall Accuracy (%)	ΔOA (%)
91.30	1.50

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 10

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	88,03	95,11	77,52	-	58,70	95,04	96,51	1,85	55,35	98,99	-
Error on Producer	7,63	3,67	6,43	-	11,81	1,16	2,85	0,97	13,12	2,01	-
Omission error (%)	11,97	4,89	22,48	-	41,30	4,96	3,49	98,15	44,65	1,01	-
User accuracy (%)	78,99	89,63	89,30	-	86,18	94,23	88,69	36,36	66,24	93,57	-
Error on User accuracy	7,26	5,28	4,32	-	5,36	1,83	4,39	10,68	8,39	4,17	-
Commission_error (%)	21,01	10,37	10,70	-	13,82	5,77	11,31	63,64	33,76	6,43	-

ZONE 11 – ROMANIA
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 11

Overall Accuracy (%)	ΔOA (%)
94.70	1.10

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 11

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy	81,39	97,60	96,03	-	80,04	93,01	97,66	-	49,45	99,79	-
Error on Producer	11,71	2,56	1,44	-	10,27	2,33	1,33	-	24,77	0,46	-
Omission error (%)	18,61	2,40	3,97	-	19,96	6,99	2,34	-	50,55	0,21	-
User accuracy (%)	85,93	97,73	98,20	-	73,29	93,62	95,32	-	76,07	89,05	-
Error on User	5,78	2,21	1,19	-	7,66	2,33	2,40	-	7,31	5,52	-
Commission_error	14,07	2,27	1,80	-	26,71	6,38	4,68	-	23,93	10,95	-

ZONE 12 – CYPRUS – GREECE
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 12

Overall Accuracy (%)	ΔOA (%)
87.90	1.80

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 12

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy	78,39	98,82	98,19	53,51	79,16	91,18	92,36	-	88,77	99,23	-
Error on Producer	12,91	1,81	1,66	8,40	5,18	2,96	4,18	-	7,33	1,47	-
Omission error (%)	21,61	1,18	1,81	46,49	20,84	8,82	7,64	-	11,23	0,77	-
User accuracy (%)	84,13	87,21	76,85	83,95	94,25	88,03	97,77	-	94,89	99,55	-
Error on User	6,74	5,31	6,30	6,89	3,15	4,13	2,12	-	3,65	0,89	-
Commission_error	15,87	12,79	23,15	16,05	5,75	11,97	2,23	-	5,11	0,45	-

ZONE 13 – BULGARIA
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 13

Overall Accuracy (%)	ΔOA (%)
94.10	1.30

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 13

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy	78,44	94,78	97,69	-	60,70	94,67	96,61	-	66,40	100,00	-
Error on Producer	13,58	4,76	1,30	-	10,38	2,71	1,89	-	24,20	-	-
Omission error (%)	21,56	5,22	2,31	-	39,30	5,33	3,39	-	33,60	-	-
User accuracy (%)	89,86	97,74	96,75	-	91,52	85,60	96,75	-	90,67	100,00	-
Error on User	5,07	2,19	1,96	-	4,14	4,41	2,08	-	4,82	-	-
Commission_error	10,14	2,26	3,25	-	8,48	14,40	3,25	-	9,33	-	-

ZONE 14 – ICELAND
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 14

Overall Accuracy (%)	ΔOA (%)
91.70	1.60

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 14

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	97,15	100,00	98,44	-	58,69	91,83	100,00	64,42	97,20	95,54	98,88
Error on Producer accuracy (%)	5,67	-	1,35	-	18,63	2,55	-	7,62	1,14	5,52	2,21
Omission error (%)	2,85	-	1,56	-	41,31	8,17	-	35,58	2,80	4,46	1,12
User accuracy (%)	86,23	63,70	69,57	-	85,89	96,34	93,57	86,13	87,95	96,43	98,01
Error on User accuracy (%)	6,17	10,16	8,89	-	5,61	1,72	4,20	4,98	3,29	3,09	1,95
Commission_error (%)	13,77	36,30	30,43	-	14,11	3,66	6,43	13,87	12,05	3,57	1,99

ZONE 15 – HUNGARY
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 15

Overall Accuracy (%)	ΔOA (%)
94.20	1.20

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 15

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	98,94	98,72	87,36	-	83,00	93,73	99,34	-	44,11	99,78	-
Error on Producer accuracy (%)	1,85	2,92	3,21	-	12,03	2,40	0,61	-	22,19	0,33	-
Omission error (%)	1,06	1,28	12,64	-	17,00	6,27	0,66	-	55,89	0,22	-
User accuracy (%)	77,87	90,07	98,20	-	65,47	91,98	96,79	-	82,78	91,43	-
Error on User accuracy (%)	8,14	5,18	1,38	-	9,37	3,31	1,86	-	6,44	4,81	-
Commission_error (%)	22,13	9,93	1,80	-	34,53	8,02	3,21	-	17,22	8,57	-

ZONE 16 – PORTUGAL
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 16

Overall Accuracy (%)	Δ OA (%)
89.90	1.50

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 16

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	91,50	94,26	84,08	95,96	88,05	89,46	82,50		82,83	94,65	
Error on Producer accuracy (%)	6,40	4,82	9,13	2,36	3,25	2,22	7,51		11,67	6,42	
Omission error (%)	8,50	5,74	15,92	4,04	11,95	10,54	17,50		17,17	5,35	
User accuracy (%)	97,14	82,86	90,71	83,86	95,16	93,16	80,00		76,97	97,08	
Error on User accuracy (%)	2,71	6,69	4,88	4,56	2,34	2,43	7,01		7,13	2,83	
Commission_error (%)	2,86	17,14	9,29	16,14	4,84	6,84	20,00		23,03	2,92	

ZONE 17 – AUSTRIA – SWITZERLAND – LIECHTENSTEIN
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 17

Overall Accuracy (%)	Δ OA (%)
95.90	1.10

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 17

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	92,01	98,27	94,52	-	70,85	97,53	94,57	-	97,21	100,00	100,00
Error on Producer accuracy (%)	6,78	1,36	3,57	-	12,61	1,23	3,69	-	3,15	-	-
Omission error (%)	7,99	1,73	5,48	-	29,15	2,47	5,43	100,00	2,79	-	-
User accuracy (%)	95,56	99,36	97,10	-	96,23	90,99	98,58	-	95,29	100,00	100,00
Error on User accuracy (%)	3,49	0,87	2,26	-	2,92	3,14	1,57	-	3,05	-	-
Commission_error (%)	4,44	0,64	2,90	-	3,77	9,01	1,42	100,00	4,71	-	-

ZONE 18 – BELGIUM – DENMARK – LUXEMBOURG – NETHERLANDS
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 18

Overall Accuracy (%)	Δ OA (%)
91.90	1.50

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 18

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	93,85	95,05	88,84		32,97	94,98	92,28		63,75	96,12	
Error on Producer accuracy	3,77	3,64	3,81		10,90	2,07	2,25		18,75	3,69	
Omission error (%)	6,15	4,95	11,16		67,03	5,02	7,72		36,25	3,88	
User accuracy (%)	91,92	89,86	94,51		85,71	84,37	96,85		85,63	99,55	
Error on User accuracy (%)	3,78	5,21	2,58		4,92	3,54	1,78		5,69	0,87	
Commission_error (%)	8,08	10,14	5,49		14,29	15,63	3,15		14,38	0,45	

ZONE 19 – ALBANIA – KOSOVO – NORTH MACEDONIA – MONTENEGRO – SERBIA
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 19

Overall Accuracy (%)	ΔOA (%)
94.70	1.20

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 19

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	83,67	99,09	98,46	100,00	60,91	93,70	97,38	-	93,18	100,00	-
Error on Producer accuracy (%)	12,01	1,14	0,94	-	8,98	2,36	1,87	-	5,18	-	-
Omission error (%)	16,33	0,91	1,54	-	39,09	6,30	2,62	-	6,82	-	-
User accuracy (%)	87,50	96,80	97,26	90,00	85,82	91,67	96,17	-	86,47	98,35	-
Error on User accuracy (%)	6,51	3,10	1,59	5,44	5,58	3,06	2,33	-	6,10	2,29	-
Commission_error (%)	12,50	3,20	2,74	10,00	14,18	8,33	3,83	-	13,53	1,65	-

ZONE 20 – SLOVENIA – CROATIA – BOSNIA & HERZEGOVINA
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 20

Overall Accuracy (%)	ΔOA (%)
90.40	1.60

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 20

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	86,41	87,27	96,41	99,42	62,87	94,44	83,02	-	73,27	99,43	100,00
Error on Producer accuracy (%)	11,78	6,04	1,40	1,16	8,02	2,24	5,72	-	17,48	0,81	-
Omission error (%)	13,59	12,73	3,59	0,58	37,13	5,56	16,98	-	26,73	0,57	-
User accuracy (%)	91,06	95,91	93,94	71,20	86,27	81,21	95,34	-	90,13	100,00	100,00
Error on User accuracy (%)	5,06	2,87	2,33	9,36	4,96	4,74	2,85	-	4,05	-	-
Commission_error (%)	8,94	4,09	6,06	28,80	13,73	18,79	4,66	-	9,87	-	-

ZONE 21 – CZECH REPUBLIC – SLOVAKIA
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 21

Overall Accuracy (%)	ΔOA (%)
97.20	0.80

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 21

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	95,88	99,57	95,93	-	79,21	96,29	98,85	-	70,00	99,72	-
Error on Producer accuracy (%)	4,67	0,70	1,87	-	15,88	1,97	0,91	-	22,89	0,56	-
Omission error (%)	4,12	0,43	4,07	-	20,79	3,71	1,15	-	30,00	0,28	-
User accuracy (%)	93,53	96,38	99,00	-	78,46	96,23	98,57	-	88,46	100,00	-
Error on User accuracy (%)	4,08	2,23	1,08	-	7,79	2,06	1,24	-	5,25	-	-
Commission_error (%)	6,47	3,62	1,00	-	21,54	3,77	1,43	-	11,54	-	-

ZONE 22 – LATVIA – LITHUANIA – ESTONIA
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 22

Overall Accuracy (%)	ΔOA (%)
91.60	1.40

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 22

	CODE_CLASS_BLIND										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	75,57	97,29	85,66	-	43,89	89,91	67,23	-	42,71	93,32	-
Error on Producer	19,07	1,67	3,69	-	14,57	3,29	3,67	-	9,29	7,39	-
Omission error (%)	24,43	2,71	14,34	-	56,11	10,09	32,77	100,00	57,29	6,68	-
User accuracy (%)	58,16	85,55	93,23	-	48,45	60,00	99,16	-	75,31	100,00	-
Error on User accuracy	9,92	4,41	2,81	-	10,21	5,70	0,93	-	6,28	-	-
Commission_error (%)	41,84	14,45	6,77	-	51,55	40,00	0,84	100,00	24,69	-	-

ZONE 23 – FRENCH OVERSEAS DEPARTMENTS
Overall Accuracy for CLC+ Backbone Raster Product 2018 in percent for zone 23

Overall Accuracy (%)	ΔOA (%)
97.60	0.60

CLC+ Backbone Raster Product - Producer's and User's accuracy in percent for zone 23

	CODE_CLASS_PLAU										
	1	2	3	4	5	6	7	8	9	10	11
Producer accuracy (%)	91,66	100,00	-	98,99	52,10	85,95	96,36	-	53,30	96,93	-
Error on Producer accuracy (%)	7,49	-	-	0,18	15,94	9,25	5,72	-	13,17	1,85	-
Omission error (%)	8,34	-	-	1,01	47,90	14,05	3,64	-	46,70	3,07	-
User accuracy (%)	88,57	87,32	-	99,41	54,34	77,17	62,14	-	67,86	99,09	-
Error on User accuracy (%)	5,39	8,28	-	0,59	7,77	5,47	10,08	-	8,50	1,20	-
Commission_error (%)	11,43	12,68	-	0,59	45,66	22,83	37,86	-	32,14	0,91	-

ANNEX 5: THEMATIC ACCURACIES FOR THE VECTOR UNITS OF EEA-38 + UK

ZONE 1 – TURKEY

Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone1

Overall Accuracy (%)	ΔOA (%)
86.60	1.70

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone1

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	86,21	58,22	92,52	41,36	92,40	42,81	44,75	64,96	97,25
Error on Producer accuracy (%)	13,59	19,78	3,36	13,94	4,41	10,26	24,72	8,61	1,32
Omission error (%)	13,79	41,78	7,48	58,64	7,60	57,19	55,25	35,04	2,75
User accuracy (%)	92,65	90,00	90,86	83,33	88,08	90,24	50,00	77,67	82,29
Error on User accuracy (%)	6,20	7,29	4,28	14,91	5,31	7,64	28,29	7,75	3,53
Commission_error (%)	7,35	10,00	9,14	16,67	11,92	9,76	50,00	22,33	17,71
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	47,65	34,30	96,77	-	89,87	81,62	90,65	98,98	-
Error on Producer accuracy (%)	9,21	10,86	2,02	-	11,12	12,31	10,59	2,10	-
Omission error (%)	52,35	65,70	3,23	-	10,13	18,38	9,35	1,02	100,00
User accuracy (%)	71,93	70,59	96,36	-	79,17	100,00	84,62	100,00	-
Error on User accuracy (%)	9,85	13,78	2,33	-	12,28	-	7,91	-	-
Commission_error (%)	28,07	29,41	3,64	-	20,83	-	15,38	-	100,00

ZONE 2 – FRANCE

Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 2

Overall Accuracy (%)	ΔOA (%)
91.30	1.50

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 2

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	87,04	70,64	95,70	80,46	97,33	92,79	58,39	61,16	92,35
Error on Producer accuracy	11,19	11,74	3,61	13,47	2,35	13,66	13,61	8,68	3,79
Omission error (%)	12,96	29,36	4,30	19,54	2,67	7,21	41,61	38,84	7,65
User accuracy (%)	92,31	92,68	94,02	80,00	91,54	78,57	92,59	90,91	87,41
Error on User accuracy (%)	5,54	5,32	3,48	14,31	3,48	23,22	8,22	5,10	4,02
Commission_error (%)	7,69	7,32	5,98	20,00	8,46	21,43	7,41	9,09	12,59
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	85,73	76,57	99,49	-	68,58	59,35	89,49	100,00	100,00
Error on Producer accuracy	7,52	13,26	0,59	-	28,06	29,47	8,88	-	-
Omission error (%)	14,27	23,43	0,51	-	31,42	40,65	10,51	-	-
User accuracy (%)	79,27	74,36	97,76	-	82,35	75,00	100,00	100,00	100,00
Error on User accuracy (%)	9,11	14,07	2,50	-	18,12	21,22	-	-	-
Commission_error (%)	20,73	25,64	2,24	-	17,65	25,00	-	-	-

ZONE 3 – SPAIN AND ANDORRA
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 3

Overall Accuracy (%)	ΔOA (%)
90.40	1.50

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 3

	CODE_CLASS_PLAU									
	11	12	21	22	31	32	33	40	51	
Producer accuracy (%)	100,00	77,38	99,25	72,21	93,47	84,84	82,45	93,71	92,68	
Error on Producer accuracy (%)	-	17,90	1,42	21,72	4,90	6,06	20,56	3,17	2,84	
Omission error (%)	-	22,62	0,75	27,79	6,53	15,16	17,55	6,29	7,32	
User accuracy (%)	88,51	76,32	92,94	76,47	93,44	94,31	88,89	94,33	93,04	
Error on User accuracy (%)	7,12	14,09	5,61	20,16	6,11	3,91	14,52	2,63	2,80	
Commission_error (%)	11,49	23,68	7,06	23,53	6,56	5,69	11,11	5,67	6,96	
	52	53	60	70	81	82	90	100	110	
Producer accuracy (%)	77,64	63,60	99,16	-	50,35	66,77	85,15	87,90	-	
Error on Producer accuracy (%)	9,88	16,54	0,99	-	9,95	17,35	9,05	15,75	-	
Omission error (%)	22,36	36,40	0,84	-	49,65	33,23	14,85	12,10	-	
User accuracy (%)	71,67	63,33	88,33	-	100,00	73,81	94,59	98,41	-	
Error on User accuracy (%)	11,80	17,85	4,29	-	-	13,80	6,68	3,06	-	
Commission_error (%)	28,33	36,67	11,67	-	-	26,19	5,41	1,59	-	

ZONE 4 – SWEDEN
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 4

Overall Accuracy (%)	ΔOA (%)
94,2	1.30

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 4

	CODE_CLASS_PLAU									
	11	12	21	22	31	32	33	40	51	
Producer accuracy (%)	72,07	66,22	99,06	73,43	98,60	-	87,04	100,00	90,67	
Error on Producer accuracy (%)	25,69	32,37	0,72	13,36	2,55	-	11,75	-	3,88	
Omission error (%)	27,93	33,78	0,94	26,57	1,40	-	12,96	-	9,33	
User accuracy (%)	97,14	83,33	95,78	88,37	95,56	-	93,94	75,00	95,10	
Error on User accuracy (%)	5,37	15,23	1,99	9,16	4,33	-	7,91	49,00	2,90	
Commission_error (%)	2,86	16,67	4,22	11,63	4,44	-	6,06	25,00	4,90	
	52	53	60	70	81	82	90	100	110	
Producer accuracy (%)	87,42	74,16	98,33	100,00	64,74	71,98	91,72	98,54	100,00	
Error on Producer accuracy (%)	6,23	11,69	3,15	-	23,67	26,05	6,62	2,44	-	
Omission error (%)	12,58	25,84	1,67	-	35,26	28,02	8,28	1,46	-	
User accuracy (%)	87,61	86,27	86,54	85,71	70,00	81,82	97,37	100,00	90,91	
Error on User accuracy (%)	6,13	9,01	9,86	19,80	21,17	21,82	4,84	-	17,82	
Commission_error (%)	12,39	13,73	13,46	14,29	30,00	18,18	2,63	-	9,09	

ZONE 5 – GERMANY
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 5

Overall Accuracy (%)	ΔOA (%)
93.60	1.40

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 5

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	86,41	92,03	99,50	87,53	97,33	-	72,62	63,79	91,10
Error on Producer accuracy (%)	9,50	7,71	0,95	15,24	2,10	-	15,27	27,74	4,52
Omission error (%)	13,59	7,97	0,50	12,47	2,67	-	27,38	36,21	8,90
User accuracy (%)	87,78	75,56	96,69	92,31	92,25	-	89,29	76,92	88,89
Error on User accuracy (%)	6,92	13,66	3,23	10,24	4,49	-	10,55	17,22	4,30
Commission_error (%)	12,22	24,44	3,31	7,69	7,75	-	10,71	23,08	11,11
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	74,11	77,36	99,67	-	51,79	15,97	75,45	95,45	-
Error on Producer accuracy (%)	8,16	15,99	0,48	-	43,18	14,70	22,28	8,68	-
Omission error (%)	25,89	22,64	0,33	-	48,21	84,03	24,55	4,55	-
User accuracy (%)	96,67	72,73	98,65	-	70,83	70,00	96,88	100,00	-
Error on User accuracy (%)	3,91	16,21	1,85	-	19,92	27,08	4,14	-	-
Commission_error (%)	3,33	27,27	1,35	-	29,17	30,00	3,13	-	-

ZONE 6 – FINLAND
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 6

Overall Accuracy (%)	ΔOA (%)
95.60	1.00

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 6

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	60,42	94,16	99,80	93,73	99,18	-	94,14	64,23	92,97
Error on Producer accuracy (%)	29,41	4,85	0,24	9,11	1,77	-	4,97	26,83	5,18
Omission error (%)	39,58	5,84	0,20	6,27	0,82	-	5,86	35,77	7,03
User accuracy (%)	86,21	94,34	98,45	94,23	96,38	-	96,36	100,00	77,45
Error on User accuracy (%)	12,14	6,16	1,35	6,40	3,16	-	4,82	-	5,91
Commission_error (%)	13,79	5,66	1,55	5,77	3,62	-	3,64	-	22,55
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	95,02	98,31	100,00	81,03	9,23	16,81	31,10	99,95	-
Error on Producer accuracy (%)	3,75	3,61	-	23,62	2,96	11,89	21,93	0,09	-
Omission error (%)	4,98	1,69	-	18,97	90,77	83,19	68,90	0,05	-
User accuracy (%)	96,64	86,54	92,52	85,71	80,65	90,91	94,12	100,00	-
Error on User accuracy (%)	3,20	9,86	5,18	15,34	9,32	10,65	6,06	-	-
Commission_error (%)	3,36	13,46	7,48	14,29	19,35	9,09	5,88	-	-

ZONE 7 – NORWAY
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 7

Overall Accuracy (%)	ΔOA (%)
94,00	1,70

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 7

	CODE_CLASS_PLAU									
	11	12	21	22	31	32	33	40	51	
Producer accuracy (%)	100,00	60,17	99,51	100,00	100,00	-	86,70	74,27	94,10	
Error on Producer accuracy (%)	-	31,53	0,90	-	-	-	16,77	22,46	3,62	
Omission error (%)	-	39,83	0,49	-	-	-	13,30	25,73	5,90	
User accuracy (%)	89,06	100,00	98,85	97,30	93,75	-	94,74	100,00	91,91	
Error on User accuracy (%)	8,10	-	1,59	5,30	4,33	-	9,79	-	4,06	
Commission_error (%)	10,94	-	1,15	2,70	6,25	-	5,26	-	8,09	
	52	53	60	70	81	82	90	100	110	
Producer accuracy (%)	76,43	85,32	89,74	100,00	91,81	93,51	98,66	99,48	70,96	
Error on Producer accuracy (%)	8,95	11,42	18,03	-	10,22	8,14	1,79	1,02	28,10	
Omission error (%)	23,57	14,68	10,26	-	8,19	6,49	1,34	0,52	29,04	
User accuracy (%)	85,39	89,80	91,30	80,00	92,00	100,00	95,24	100,00	85,71	
Error on User accuracy (%)	7,18	8,48	11,77	39,20	10,23	-	6,44	-	15,34	
Commission_error (%)	14,61	10,20	8,70	20,00	8,00	-	4,76	-	14,29	

ZONE 8 – POLAND
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 8

Overall Accuracy (%)	ΔOA (%)
96,10	2,10

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 8

	CODE_CLASS_PLAU									
	11	12	21	22	31	32	33	40	51	
Producer accuracy (%)	99,69	98,96	99,63	100,00	99,23	-	100,00	100,00	88,85	
Error on Producer accuracy (%)	0,58	2,02	0,73	-	1,47	-	-	-	8,68	
Omission error (%)	0,31	1,04	0,37	-	0,77	-	-	-	11,15	
User accuracy (%)	96,67	89,19	100,00	100,00	94,35	-	100,00	100,00	99,40	
Error on User accuracy (%)	4,58	10,44	-	-	3,49	-	-	-	1,14	
Commission_error (%)	3,33	10,81	-	-	5,65	-	-	-	0,60	
	52	53	60	70	81	82	90	100	110	
Producer accuracy (%)	89,31	82,48	100,00	-	15,42	50,44	55,47	97,82	-	
Error on Producer accuracy (%)	14,85	11,36	-	-	24,19	35,15	43,38	4,30	-	
Omission error (%)	10,69	17,52	-	-	84,58	49,56	44,53	2,18	-	
User accuracy (%)	94,55	100,00	92,68	-	88,89	100,00	97,50	100,00	-	
Error on User accuracy (%)	6,00	-	5,85	-	14,52	-	4,72	-	-	
Commission_error (%)	5,45	-	7,32	-	11,11	-	2,50	-	-	

ZONE 9 – ITALY – MALTA
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 9

Overall Accuracy (%)	ΔOA (%)
92.90	1.50

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 9

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	96,44	91,72	97,96	91,27	97,64	89,39	100,00	88,41	90,87
Error on Producer accuracy (%)	6,89	10,75	3,82	14,43	1,64	5,70	-	7,13	4,56
Omission error (%)	3,56	8,28	2,04	8,73	2,36	10,61	-	11,59	9,13
User accuracy (%)	100,00	98,51	94,83	92,31	91,79	92,86	81,25	94,19	94,09
Error on User accuracy (%)	-	2,88	5,80	14,49	3,37	5,00	21,22	4,86	3,01
Commission_error (%)	-	1,49	5,17	7,69	8,21	7,14	18,75	5,81	5,91
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	75,98	83,48	98,43	-	71,15	77,29	100,00	93,28	53,28
Error on Producer accuracy (%)	8,13	11,28	1,34	-	21,17	20,35	-	8,68	24,69
Omission error (%)	24,02	16,52	1,57	-	28,85	22,71	-	6,72	46,72
User accuracy (%)	89,74	85,11	93,26	-	92,86	94,74	89,13	100,00	100,00
Error on User accuracy (%)	6,23	10,29	3,76	-	9,22	9,55	9,53	-	-
Commission_error (%)	10,26	14,89	6,74	-	7,14	5,26	10,87	-	-

ZONE 10 – UNITED KINGDOM – IRELAND
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 10

Overall Accuracy (%)	ΔOA (%)
91.40	1.60

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 10

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	100,00	80,39	92,94	59,06	95,62	-	100,00	97,63	95,44
Error on Producer accuracy (%)	-	10,98	5,69	21,88	3,34	-	-	4,63	1,81
Omission error (%)	-	19,61	7,06	40,94	4,38	-	-	2,37	4,56
User accuracy (%)	89,22	83,05	92,59	66,67	93,94	-	91,67	71,43	92,86
Error on User accuracy (%)	6,37	9,42	4,48	24,69	4,10	-	16,33	22,14	2,24
Commission_error (%)	10,78	16,95	7,41	33,33	6,06	-	8,33	28,57	7,14
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	72,82	70,37	96,56	-	39,60	25,78	39,88	100,00	-
Error on Producer accuracy (%)	7,95	13,28	2,67	-	16,28	18,51	20,93	-	-
Omission error (%)	27,18	29,63	3,44	100,00	60,40	74,22	60,12	-	-
User accuracy (%)	84,54	72,09	94,57	-	87,18	52,94	84,00	100,00	-
Error on User accuracy (%)	6,82	13,57	4,61	-	9,66	24,46	12,32	-	-
Commission_error (%)	15,46	27,91	5,43	100,00	12,82	47,06	16,00	-	-

ZONE 11 – ROMANIA
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 11

Overall Accuracy (%)	ΔOA (%)
90.70	2.20

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 11

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	83,05	88,14	98,65	59,03	99,82	-	47,77	63,94	89,60
Error on Producer accuracy (%)	22,76	7,13	1,64	18,17	0,33	-	18,85	14,75	5,43
Omission error (%)	16,95	11,86	1,35	40,97	0,18	-	52,23	36,06	10,40
User accuracy (%)	80,56	89,74	91,30	82,35	90,38	-	85,19	86,79	94,14
Error on User accuracy (%)	13,50	9,07	5,96	15,93	4,20	-	12,51	8,79	2,87
Commission_error (%)	19,44	10,26	8,70	17,65	9,62	-	14,81	13,21	5,86
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	68,39	63,20	98,06	-	13,20	18,90	94,93	93,12	-
Error on Producer accuracy (%)	12,83	17,05	1,42	-	13,62	30,01	9,80	9,47	-
Omission error (%)	31,61	36,80	1,94	-	86,80	81,10	5,07	6,88	-
User accuracy (%)	86,27	83,78	90,00	-	86,36	77,78	80,00	100,00	-
Error on User accuracy (%)	8,64	11,42	5,48	-	13,73	20,37	21,74	-	-
Commission_error (%)	13,73	16,22	10,00	-	13,64	22,22	20,00	-	-

ZONE 12 – CYPRUS – GREECE
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 12

Overall Accuracy (%)	ΔOA (%)
81.60	2.40

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 12

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	83,52	71,62	98,26	68,09	100,00	70,21	64,08	68,15	81,81
Error on Producer accuracy (%)	15,68	18,63	4,05	17,37	-	12,48	18,02	6,27	5,76
Omission error (%)	16,48	28,38	1,74	31,91	-	29,79	35,92	31,85	18,19
User accuracy (%)	97,67	85,19	77,53	73,08	68,48	81,82	67,86	91,56	82,66
Error on User accuracy (%)	4,40	13,16	9,78	17,05	8,57	10,59	17,30	3,88	5,71
Commission_error (%)	2,33	14,81	22,47	26,92	31,52	18,18	32,14	8,44	17,34
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	71,51	47,04	93,33	-	75,66	72,24	67,58	100,00	-
Error on Producer accuracy (%)	11,02	16,04	3,80	-	22,85	16,98	21,06	-	-
Omission error (%)	28,49	52,96	6,67	-	24,34	27,76	32,42	-	-
User accuracy (%)	80,85	66,67	94,26	-	73,91	100,00	91,30	100,00	-
Error on User accuracy (%)	10,59	17,16	4,08	-	19,24	-	10,83	-	-
Commission_error (%)	19,15	33,33	5,74	-	26,09	-	8,70	-	-

ZONE 13 – BULGARIA
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 13

Overall Accuracy (%)	ΔOA (%)
93.10	1.60

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 13

	CODE_CLASS_PLAU									
	11	12	21	22	31	32	33	40	51	
Producer accuracy (%)	87,31	68,33	99,36	86,34	99,63	-	87,01	62,61	99,34	
Error on Producer accuracy (%)	23,65	17,41	1,15	17,21	0,43	-	16,62	13,39	1,30	
Omission error (%)	12,69	31,67	0,64	13,66	0,37	-	12,99	37,39	0,66	
User accuracy (%)	88,89	93,33	92,96	100,00	93,09	-	100,00	92,73	83,72	
Error on User accuracy (%)	14,94	10,93	6,13	-	3,27	-	-	6,36	6,01	
Commission_error (%)	11,11	6,67	7,04	-	6,91	-	-	7,27	16,28	
	52	53	60	70	81	82	90	100	110	
Producer accuracy (%)	78,66	43,40	97,51	-	100,00	48,37	92,59	100,00	-	
Error on Producer accuracy (%)	12,46	10,82	1,68	-	-	33,18	13,44	-	-	
Omission error (%)	21,34	56,60	2,49	-	-	51,63	7,41	-	-	
User accuracy (%)	81,36	86,67	99,09	-	90,91	100,00	100,00	100,00	-	
Error on User accuracy (%)	10,11	9,33	1,71	-	17,82	-	-	-	-	
Commission_error (%)	18,64	13,33	0,91	-	9,09	-	-	-	-	

ZONE 14 – ICELAND
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 14

Overall Accuracy (%)	ΔOA (%)
80.80	7.10

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 14

	CODE_CLASS_PLAU									
	11	12	21	22	31	32	33	40	51	
Producer accuracy (%)	100,00	87,98	100,00	100,00	100,00	-	-	86,76	99,94	
Error on Producer accuracy (%)	-	14,91	-	-	-	-	-	20,21	0,11	
Omission error (%)	-	12,02	-	-	-	-	-	100,00	13,24	0,06
User accuracy (%)	77,78	87,50	100,00	100,00	63,64	-	-	100,00	74,07	
Error on User accuracy (%)	30,80	21,61	-	-	35,64	-	-	-	18,74	
Commission_error (%)	22,22	12,50	-	-	36,36	-	100,00	-	25,93	
	52	53	60	70	81	82	90	100	110	
Producer accuracy (%)	30,66	13,23	-	59,09	55,90	53,32	96,55	61,38	92,20	
Error on Producer accuracy (%)	31,86	23,48	-	25,55	21,41	34,70	4,14	29,49	15,22	
Omission error (%)	69,34	86,77	-	40,91	44,10	46,68	3,45	38,62	7,80	
User accuracy (%)	87,50	66,67	-	100,00	76,47	62,50	77,78	100,00	100,00	
Error on User accuracy (%)	18,71	34,92	-	-	19,07	31,63	14,88	-	-	
Commission_error (%)	12,50	33,33	-	-	23,53	37,50	22,22	-	-	

ZONE 15 – HUNGARY
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 15

Overall Accuracy (%)	ΔOA (%)
90.80	2.00

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 15

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	97,61	78,08	96,69	58,72	98,55	-	48,60	67,10	87,70
Error on Producer accuracy	4,14	12,67	3,68	27,71	0,82	-	29,29	19,65	7,51
Omission error (%)	2,39	21,92	3,31	41,28	1,45	-	51,40	32,90	12,30
User accuracy (%)	79,76	82,29	92,96	88,14	90,38	-	89,26	61,32	82,11
Error on User accuracy (%)	9,35	7,56	4,24	5,71	4,63	-	5,56	11,17	6,75
Commission_error (%)	20,24	17,71	7,04	11,86	9,62	-	10,74	38,68	17,89
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	60,55	36,36	98,38	-	10,09	69,63	76,61	99,97	-
Error on Producer accuracy	10,46	11,10	1,04	-	7,92	32,50	7,83	0,06	-
Omission error (%)	39,45	63,64	1,62	-	89,91	30,37	23,39	0,03	-
User accuracy (%)	74,77	86,08	97,01	-	73,75	83,95	92,13	99,15	-
Error on User accuracy (%)	8,04	6,72	2,71	-	9,96	8,48	5,28	1,67	-
Commission_error (%)	25,23	13,92	2,99	-	26,25	16,05	7,87	0,85	-

ZONE 16 – PORTUGAL
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 16

Overall Accuracy (%)	ΔOA (%)
90.20	1.60

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 16

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	93,22	72,90	100,00	93,06	85,32	97,22	96,81	92,52	88,64
Error on Producer accuracy (%)	11,84	13,46	-	12,48	10,92	2,19	6,77	2,99	3,30
Omission error (%)	6,78	27,10	-	6,94	14,68	2,78	3,19	7,48	11,36
User accuracy (%)	88,89	80,95	93,33	73,08	88,00	87,89	87,50	94,22	95,44
Error on User accuracy (%)	8,16	11,47	6,53	19,44	12,26	3,97	12,04	2,99	2,41
Commission_error (%)	11,11	19,05	6,67	26,92	12,00	12,11	12,50	5,78	4,56
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	74,67	92,21	92,49	-	37,11	77,57	72,30	100,00	-
Error on Producer accuracy (%)	8,81	10,61	5,81	-	14,46	29,02	27,54	-	-
Omission error (%)	25,33	7,79	7,51	-	62,89	22,43	27,70	-	-
User accuracy (%)	80,00	62,50	94,32	-	81,25	61,54	85,71	100,00	-
Error on User accuracy (%)	9,11	23,01	4,84	-	14,72	30,15	14,97	-	-
Commission_error (%)	20,00	37,50	5,68	-	18,75	38,46	14,29	-	-

ZONE 17 – AUSTRIA – SWITZERLAND – LIECHTENSTEIN
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 17

Overall Accuracy (%)	ΔOA (%)
91.70	1.80

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 17

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	88,67	73,94	98,31	82,08	96,42	-	75,85	81,92	95,08
Error on Producer accuracy (%)	14,54	15,96	1,43	16,77	3,80	-	16,00	14,65	3,08
Omission error (%)	11,33	26,06	1,69	17,92	3,58	-	24,15	18,08	4,92
User accuracy (%)	100,00	82,14	95,18	86,67	93,04	-	85,00	86,67	92,11
Error on User accuracy (%)	-	13,94	3,29	12,37	4,75	-	14,59	12,16	4,32
Commission_error (%)	-	17,86	4,82	13,33	6,96	-	15,00	13,33	7,89
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	72,06	62,90	97,78	-	89,35	69,79	97,81	100,00	57,73
Error on Producer accuracy (%)	11,00	19,56	3,00	-	11,35	17,73	3,98	-	15,82
Omission error (%)	27,94	37,10	2,22	-	10,65	30,21	2,19	-	42,27
User accuracy (%)	75,41	61,90	98,15	-	94,44	90,00	88,16	100,00	100,00
Error on User accuracy (%)	10,99	21,28	3,56	-	10,30	15,18	7,68	-	-
Commission_error (%)	24,59	38,10	1,85	-	5,56	10,00	11,84	-	-

ZONE 18 – BELGIUM – DENMARK – LUXEMBOURG – NETHERLANDS
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 18

Overall Accuracy (%)	ΔOA (%)
91.20	1.70

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 18

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	89,69	65,14	87,78	67,97	99,13	-	75,46	22,24	94,41
Error on Producer accuracy (%)	6,81	12,09	7,30	22,91	1,58	-	22,94	11,55	3,11
Omission error (%)	10,31	34,86	12,22	32,03	0,87	-	24,54	77,76	5,59
User accuracy (%)	91,87	80,65	97,06	64,29	85,71	-	76,92	75,00	90,94
Error on User accuracy (%)	4,91	9,06	3,85	26,05	6,57	-	22,90	16,33	3,50
Commission_error (%)	8,13	19,35	2,94	35,71	14,29	-	23,08	25,00	9,06
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	85,29	70,83	98,96	-	36,39	27,62	44,84	98,68	-
Error on Producer accuracy (%)	8,96	14,88	1,05	-	26,29	20,63	22,09	1,79	-
Omission error (%)	14,71	29,17	1,04	-	63,61	72,38	55,16	1,32	-
User accuracy (%)	78,08	76,92	96,83	-	72,73	77,78	94,29	98,75	-
Error on User accuracy (%)	10,06	15,33	3,06	-	23,33	23,52	7,38	2,42	-
Commission_error (%)	21,92	23,08	3,17	-	27,27	22,22	5,71	1,25	-

ZONE 19 – ALBANIA – KOSOVO – NORTH MACEDONIA – MONTENEGRO – SERBIA
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 19

Overall Accuracy (%)	ΔOA (%)
90.60	2.00

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 19

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	78,10	78,30	92,24	82,17	98,87	66,30	43,88	51,29	91,47
Error on Producer accuracy (%)	28,62	18,36	9,06	20,64	1,36	27,98	20,37	12,37	4,94
Omission error (%)	21,90	21,70	7,76	17,83	1,13	33,70	56,12	48,71	8,53
User accuracy (%)	96,88	88,89	86,57	85,71	91,25	89,47	95,00	77,78	87,15
Error on User accuracy (%)	5,85	10,27	8,56	18,33	3,54	13,45	8,54	10,27	5,06
Commission_error (%)	3,13	11,11	13,43	14,29	8,75	10,53	5,00	22,22	12,85
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	79,12	76,68	97,87	-	72,73	69,72	76,14	93,89	-
Error on Producer accuracy (%)	8,45	13,06	1,76	-	30,39	25,67	20,69	7,48	-
Omission error (%)	20,88	23,32	2,13	-	27,27	30,28	23,86	6,11	-
User accuracy (%)	88,89	85,19	95,35	-	83,87	89,47	91,43	98,33	-
Error on User accuracy (%)	6,80	12,51	4,40	-	13,39	13,13	9,27	3,16	-
Commission_error (%)	11,11	14,81	4,65	-	16,13	10,53	8,57	1,67	-

ZONE 20 – SLOVENIA – CROATIA – BOSNIA & HERZEGOVINA
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 20

Overall Accuracy (%)	ΔOA (%)
94.50	1.60

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 20

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	100,00	98,29	96,12	94,08	99,09	62,90	82,17	86,44	92,72
Error on Producer accuracy (%)	-	3,29	4,49	11,03	0,81	23,10	16,81	10,34	4,39
Omission error (%)	-	1,71	3,88	5,92	0,91	37,10	17,83	13,56	7,28
User accuracy (%)	96,77	100,00	92,96	90,00	98,67	100,00	90,00	91,21	90,83
Error on User accuracy (%)	6,32	-	6,04	13,49	1,49	-	12,83	5,92	5,18
Commission_error (%)	3,23	-	7,04	10,00	1,33	-	10,00	8,79	9,17
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	84,71	91,43	93,13	-	95,39	95,49	95,31	100,00	-
Error on Producer accuracy (%)	9,76	10,66	5,18	-	8,44	8,62	8,97	-	-
Omission error (%)	15,29	8,57	6,87	-	4,61	4,51	4,69	-	-
User accuracy (%)	87,93	85,71	90,91	-	87,50	100,00	95,24	96,88	-
Error on User accuracy (%)	8,31	15,34	7,46	-	13,82	-	9,11	6,13	-
Commission_error (%)	12,07	14,29	9,09	-	12,50	-	4,76	3,13	-

ZONE 21 – CZECH REPUBLIC – SLOVAKIA
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 21

Overall Accuracy (%)	ΔOA (%)
92.20	1.90

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 21

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	100,00	69,31	99,26	81,12	96,65	-	78,96	52,63	87,13
Error on Producer accuracy (%)	-	15,75	0,96	13,90	2,13	-	13,51	18,07	7,43
Omission error (%)	-	30,69	0,74	18,88	3,35	-	21,04	47,37	12,87
User accuracy (%)	92,86	96,08	94,44	90,24	90,77	-	90,00	88,89	91,39
Error on User accuracy (%)	7,00	4,83	3,82	8,97	4,15	-	8,97	8,23	3,59
Commission_error (%)	7,14	3,92	5,56	9,76	9,23	-	10,00	11,11	8,61
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	86,48	55,49	98,97	-	55,46	32,84	53,20	93,31	-
Error on Producer accuracy (%)	7,38	16,70	0,97	-	50,01	30,48	28,68	9,79	-
Omission error (%)	13,52	44,51	1,03	-	44,54	67,16	46,80	6,69	-
User accuracy (%)	90,67	72,73	93,55	-	71,43	93,33	92,86	100,00	-
Error on User accuracy (%)	6,37	14,35	4,36	-	26,70	12,22	9,07	-	-
Commission_error (%)	9,33	27,27	6,45	-	28,57	6,67	7,14	-	-

ZONE 22 – LATVIA – LITHUANIA – ESTONIA
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 22

Overall Accuracy (%)	ΔOA (%)
91.10	1.90

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 22

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	68,60	57,10	95,77	79,63	95,69	-	87,37	91,27	87,27
Error on Producer accuracy (%)	19,44	21,81	3,20	16,47	3,21	-	12,99	16,25	5,78
Omission error (%)	31,40	42,90	4,23	20,37	4,31	-	12,63	8,73	12,73
User accuracy (%)	95,65	76,00	99,29	85,71	96,32	-	72,73	92,86	84,46
Error on User accuracy (%)	7,84	16,74	1,37	17,15	3,13	-	20,03	13,49	5,32
Commission_error (%)	4,35	24,00	0,71	14,29	3,68	-	27,27	7,14	15,54
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	78,84	69,41	96,28	-	50,89	51,51	95,19	100,00	-
Error on Producer accuracy (%)	7,94	14,26	2,54	-	18,86	27,90	7,92	-	-
Omission error (%)	21,16	30,59	3,72	-	49,11	48,49	4,81	-	-
User accuracy (%)	78,21	82,05	93,00	-	76,47	50,00	68,97	100,00	-
Error on User accuracy (%)	9,05	11,75	4,98	-	18,14	32,67	19,79	-	-
Commission_error (%)	21,79	17,95	7,00	-	23,53	50,00	31,03	-	-

ZONE 23 – FRENCH OVERSEAS DEPARTMENTS
Overall Accuracy for CLC+ Backbone Vector Product 2018 in percent for zone 23

Overall Accuracy (%)	ΔOA (%)
91.60	3.50

CLC+ Backbone Vector Product - Producer's and User's accuracy in percent for zone 23

	CODE_CLASS_PLAU								
	11	12	21	22	31	32	33	40	51
Producer accuracy (%)	87,08	27,82	100,00	-	-	99,79	100,00	13,71	54,54
Error on Producer accuracy (%)	19,10	37,92	-	-	-	0,18	-	7,95	21,26
Omission error (%)	12,92	72,18	-	-	-	0,21	-	86,29	45,46
User accuracy (%)	100,00	85,71	100,00	-	-	92,74	100,00	70,37	78,38
Error on User accuracy (%)	-	24,25	-	-	-	3,90	-	15,82	11,90
Commission_error (%)	-	14,29	-	-	-	7,26	-	29,63	21,62
	52	53	60	70	81	82	90	100	110
Producer accuracy (%)	31,41	100,00	100,00	-	-	27,89	100,00	100,00	-
Error on Producer accuracy (%)	28,30	-	-	-	-	25,08	-	-	-
Omission error (%)	68,59	-	-	-	100,00	72,11	-	-	-
User accuracy (%)	60,00	66,67	25,00	-	-	100,00	100,00	100,00	-
Error on User accuracy (%)	27,72	46,20	42,44	-	-	-	-	-	-
Commission_error (%)	40,00	33,33	75,00	-	100,00	-	-	-	-

ANNEX 6: GEOMETRIC VALIDATION RESULTS FOR THE VECTOR PRODUCT PER BIOGEOGRAPHIC REGION
ALP – Alpine Biogeographical Region

Results of geometric validation for Alpine biogeographical region

Sampling design		
No. of samples (equally distributed)	1438	
No. of samples within size classes	475 small/ 488 medium/ 475 large	
Delineation accuracy of geometric units:		
Appropriate size	Defined threshold of landscape elements that can be 15% larger or smaller	Validation result
too large: (+/- 5m tolerance)	≤10% (≤ 144 Softbones)	9.9 % (143) 1.7 % (24)
too small: (+/- 5m tolerance)	≤15% (≤ 216 Softbones)	20.9 % (301) 8.3 % (120)
Appropriate delineation		
Shift of border within 30m buffer(at least once beyond threshold)	41.1 % (591)	
Perimeter shift beyond 25m buffer (proportion located outside)	24.8 % (357)	
Jaccard Index		
Overall Index (range of single landscape elements)	0.79 % (0.22 – 1 %)	

Results of plausibility analysis of the 100m buffer region for Alpine biogeographical region

Sampling design				
No. of samples (equally distributed)	276			
No. of samples within size classes	90 small/ 89 medium/ 97 large			
Delineation plausibility within direct context (100m buffer)				
Sample characteristics (summary)				
Total area checked [ha] (buffer region without sample extent)	2800 ha			
... containing Hardbone [km (%)]	81.6 km (23 %)			
... containing Softbone [km (%)]	266.7 km (77 %)			
Average line density [km/ha]	0.12 km/ha			
Plausibility check				
Plausible Softbone lines within buffer region [%]	217.4 km (79.3 %)			

ANA – Anatolian Biogeographical Region

Results of geometric validation for Anatolian biogeographical region

Sampling design		
No. of samples (equally distributed)		1437
No. of samples within size classes		484 small/ 471 medium/ 482 large
Delineation accuracy of geometric units:		
Appropriate size	Defined threshold of landscape elements that can be 15% larger or smaller	Validation result
too large: <i>(+/- 5m tolerance)</i>	≤10% (≤ 144 Softbones)	5.6 % (81) 0.9 % (13)
too small: <i>(+/- 5m tolerance)</i>	≤15% (≤ 216 Softbones)	21.9 % (314) 5.1 % (73)
Appropriate delineation		
Shift of border within 30m buffer(at least once beyond threshold)		35.1 % (505)
Perimeter shift beyond 25m buffer (proportion located outside)		15.9 % (228)
Jaccard Index		
Overall Index (range of single landscape elements)		0.78 % (0.32 - 1)

Results of plausibility analysis of the 100m buffer region for Anatolian biogeographical region

Sampling design		
No. of samples (equally distributed)		259
No. of samples within size classes		88 small/ 84 medium/ 87 large
Delineation plausibility within direct context (100m buffer)		
Sample characteristics (summary)		
Total area checked [ha] (buffer region without sample extent)		2770 ha
... containing Hardbone [km (%)]		64.1 km (19 %)
... containing Softbone [km (%)]		267 km (81 %)
Average line density [km/ha]		0.12 km/ha
Plausibility check		
Plausible Softbone lines within buffer region [%]		234.4 km (86.4 %)

ARC – Arctic Biogeographical region

Results of geometric validation for Arctic biogeographical region

Sampling design		
No. of samples (equally distributed)		1370
No. of samples within size classes		459 small/ 454 medium/ 457 large
Delineation accuracy of geometric units:		
Appropriate size	Defined threshold of landscape elements that can be 15% larger or smaller	Validation result
too large: <i>(+/- 5m tolerance)</i>	≤10% (≤ 0 Softbones)	16.3 % (223) 3.9 % (54)
too small: <i>(+/- 5m tolerance)</i>	≤15% (≤ 0 Softbones)	18.4 % (252) 6 % (82)
Appropriate delineation		
Shift of border within 30m buffer(at least once beyond threshold)		55.7 % (763)
Perimeter shift beyond 25m buffer (proportion located outside)		35.8 % (491)
Jaccard Index		
Overall Index (range of single landscape elements)		0.82 % (0.33 - 1)

Results of plausibility analysis of the 100m buffer region for Arctic biogeographical region

Sampling design		
No. of samples (equally distributed)		260
No. of samples within size classes		87 small/ 87 medium/ 86 large
Delineation plausibility within direct context (100m buffer)		
Sample characteristics (summary)		
Total area checked [ha] (buffer region without sample extent)		4232 ha
... containing Hardbone [km (%)]		45 km (13 %)
... containing Softbone [km (%)]		297.2 km (87 %)
Average line density [km/ha]		0.08 km/ha
Plausibility check		
Plausible Softbone lines within buffer region [%]		273.7 km (91.5 %)

ATL – Atlantic Biogeographical region

Results of geometric validation for Atlantic biogeographical region

Sampling design		
No. of samples (equally distributed)		1432
No. of samples within size classes		479 small/ 470 medium/ 483 large
Delineation accuracy of geometric units:		
Appropriate size	Defined threshold of landscape elements that can be 15% larger or smaller	Validation result
too large: <i>(+/- 5m tolerance)</i>	≤10% (≤ 143 Softbones)	16 % (229) 2.8 % (40)
too small: <i>(+/- 5m tolerance)</i>	≤15% (≤ 215 Softbones)	19.5 % (279) 9.1 % (130)
Appropriate delineation		
Shift of border within 30m buffer(at least once beyond threshold)		44.1 % (631)
Perimeter shift beyond 25m buffer (proportion located outside)		29.3 % (420)
Jaccard Index		
Overall Index (range of single landscape elements)		0.81 % (0.17 - 1)

Results of plausibility analysis of the 100m buffer region for Atlantic biogeographical region

Sampling design		
No. of samples (equally distributed)		283
No. of samples within size classes		93 small/ 95 medium/ 95 large
Delineation plausibility within direct context (100m buffer)		
Sample characteristics (summary)		
Total area checked [ha] (buffer region without sample extent)		3150 ha
... containing Hardbone [km (%)]		155.6 km (37 %)
... containing Softbone [km (%)]		263.8 km (63 %)
Average line density [km/ha]		0.13 km/ha
Plausibility check		
Plausible Softbone lines within buffer region [%]		191.8 km (72.2 %)

BLS – Black Sea Biogeographical region

Results of geometric validation for biogeographical region Black Sea

Sampling design		
No. of samples (equally distributed)		1436
No. of samples within size classes		481 small/ 474 medium/ 481 large
Delineation accuracy of geometric units:		
Appropriate size	Defined threshold of landscape elements that can be 15% larger or smaller	Validation result
too large: <i>(+/- 5m tolerance)</i>	≤10% (≤ 144 Softbones)	8.4 % (120) 0.6 % (8)
too small: <i>(+/- 5m tolerance)</i>	≤15% (≤ 215 Softbones)	17.8 % (255) 3.2 % (46)
Appropriate delineation		
Shift of border within 30m buffer(at least once beyond threshold)		31.3 % (450)
Perimeter shift beyond 25m buffer (proportion located outside)		12.7 % (182)
Jaccard Index		
Overall Index (range of single landscape elements)		0.8 % (0.42 - 1)

Results of plausibility analysis of the 100m buffer region for biogeographical region Black Sea

Sampling design		
No. of samples (equally distributed)		266
No. of samples within size classes		84 small/ 91 medium/ 91 large
Delineation plausibility within direct context (100m buffer)		
Sample characteristics (summary)		
Total area checked [ha] (buffer region without sample extent)		2800 ha
... containing Hardbone [km (%)]		81.6 km (23 %)
... containing Softbone [km (%)]		266.7 km (77 %)
Average line density [km/ha]		0.12 km/ha
Plausibility check		
Plausible Softbone lines within buffer region [%]		217.4 km (79.3 %)

BOR – Boreal Biogeographic region

Results of geometric validation for Boreal biogeographical region

Sampling design		
No. of samples (equally distributed)		1375
No. of samples within size classes		459 small/ 448 medium/ 468 large
Delineation accuracy of geometric units:		
Appropriate size	Defined threshold of landscape elements that can be 15% larger or smaller	Validation result
too large: <i>(+/- 5m tolerance)</i>	≤10% (≤ 138 Softbones)	11.7 % (161) 2.1 % (29)
too small: <i>(+/- 5m tolerance)</i>	≤15% (≤ 206 Softbones)	36.3 % (499) 18.8 % (259)
Appropriate delineation		
Shift of border within 30m buffer(at least once beyond threshold)		56 % (770)
Perimeter shift beyond 25m buffer (proportion located outside)		38.8 % (533)
Jaccard Index		
Overall Index (range of single landscape elements)		0.74 % (0.28 - 1)

Results of plausibility analysis of the 100m buffer region for Boreal biogeographical region

Sampling design		
No. of samples (equally distributed)		260
No. of samples within size classes		89 small/ 84 medium/ 87 large
Delineation plausibility within direct context (100m buffer)		
Sample characteristics (summary)		
Total area checked [ha] (buffer region without sample extent)		2995 ha
... containing Hardbone [km (%)]		71.3 km (21 %)
... containing Softbone [km (%)]		269 km (79 %)
Average line density [km/ha]		0.11 km/ha
Plausibility check		
Plausible Softbone lines within buffer region [%]		231.2 km (85.9 %)

CON – Continental Biogeographic region

Results of geometric validation for Continental biogeographical region

Sampling design		
No. of samples (equally distributed)		1434
No. of samples within size classes		480 small/ 476 medium/ 478 large
Delineation accuracy of geometric units:		
Appropriate size	Defined threshold of landscape elements that can be 15% larger or smaller	Validation result
too large: <i>(+/- 5m tolerance)</i>	≤10% (≤ 149 Softbones)	10.6 % (152) 2.4 % (35)
too small: <i>(+/- 5m tolerance)</i>	≤15% (≤ 215 Softbones)	17.2 % (247) 5.6 % (80)
Appropriate delineation		
Shift of border within 30m buffer(at least once beyond threshold)		34.9 % (501)
Perimeter shift beyond 25m buffer (proportion located outside)		18.1 % (259)
Jaccard Index		
Overall Index (range of single landscape elements)		0.82 % (0.12 - 1)

Results of plausibility analysis of the 100m buffer region for Continental biogeographical region

Sampling design		
No. of samples (equally distributed)		276
No. of samples within size classes		96 small/ 88 medium/ 92 large
Delineation plausibility within direct context (100m buffer)		
Sample characteristics (summary)		
Total area checked [ha] (buffer region without sample extent)		2910 ha
... containing Hardbone [km (%)]		145 km (35 %)
... containing Softbone [km (%)]		266.5 km (65 %)
Average line density [km/ha]		0.14 km/ha
Plausibility check		
Plausible Softbone lines within buffer region [%]		203 km (74.5 %)

MAC – Macaronesian Biogeographical region

Results of geometric validation for Macaronesian biogeographical region

Sampling design		
No. of samples (equally distributed)	668.00	
No. of samples within size classes	224 small/ 222 medium/ 223 large	
Delineation accuracy of geometric units:		
Appropriate size	Defined threshold of landscape elements that can be 15% larger or smaller	Validation result
too large: <i>(+/- 5m tolerance)</i>	≤10% (≤ 67 Softbones)	8.8 % (59) 1.2 % (8)
too small: <i>(+/- 5m tolerance)</i>	≤15% (≤ 100 Softbones)	14.4 % (96) 3.4 % (23)
Appropriate delineation		
Shift of border within 30m buffer (at least once beyond threshold)	32.9 % (220)	
Perimeter shift beyond 25m buffer (proportion located outside)	16.9 % (113)	
Jaccard Index		
Overall Index (range of single landscape elements)	0.85 % (0.36 - 1)	

Results of plausibility analysis of the 100m buffer region for Macaronesian biogeographical region

Sampling design				
No. of samples (equally distributed)	243			
No. of samples within size classes	77 small/ 83 medium/ 83 large			
Delineation plausibility within direct context (100m buffer)				
Sample characteristics (summary)				
Total area checked [ha] (buffer region without sample extent)	2798 ha			
... containing Hardbone [km (%)]	203.3 km (58 %)			
... containing Softbone [km (%)]	146.5 km (42 %)			
Average line density [km/ha]	0.13 km/ha			
Plausibility check				
Plausible Softbone lines within buffer region [%]	136.2 km (92.4 %)			

MED – Mediterranean Biogeographical region

Results of geometric validation for Mediterranean biogeographical region

Sampling design		
No. of samples (equally distributed)		1468
No. of samples within size classes		500 small/ 474 medium/ 494 large
Delineation accuracy of geometric units:		
Appropriate size	Defined threshold of landscape elements that can be 15% larger or smaller	Validation result
too large: <i>(+/- 5m tolerance)</i>	≤10% (≤ 147 Softbones)	8.6 % (126) 0.9 % (13)
too small: <i>(+/- 5m tolerance)</i>	≤15% (≤ 220 Softbones)	22.7 % (333) 7.2 % (106)
Appropriate delineation		
Shift of border within 30m buffer(at least once beyond threshold)		38.8 % (569)
Perimeter shift beyond 25m buffer (proportion located outside)		19.1 % (281)
Jaccard Index		
Overall Index (range of single landscape elements)		0.81 % (0.11 - 1)

Results of plausibility analysis of the 100m buffer region for Mediterranean biogeographical region

Sampling design		
No. of samples (equally distributed)		281
No. of samples within size classes		93 small/ 94 medium/ 94 large
Delineation plausibility within direct context (100m buffer)		
Sample characteristics (summary)		
Total area checked [ha] (buffer region without sample extent)		3118 ha
... containing Hardbone [km (%)]		143 km (34 %)
... containing Softbone [km (%)]		281.5 km (66 %)
Average line density [km/ha]		0.14 km/ha
Plausibility check		
Plausible Softbone lines within buffer region [%]		243.4 km (86.8 %)

PAN – Pannonian Biogeographic region

Results of geometric validation for Pannonian biogeographical region

Sampling design		
No. of samples (equally distributed)		1437
No. of samples within size classes		480 small/ 479 medium/ 478 large
Delineation accuracy of geometric units:		
Appropriate size	Defined threshold of landscape elements that can be 15% larger or smaller	Validation result
too large: <i>(+/- 5m tolerance)</i>	≤10% (≤ 144 Softbones)	5.4 % (77) 0.3 % (4)
too small: <i>(+/- 5m tolerance)</i>	≤15% (≤ 216 Softbones)	18.6 % (268) 4.7 % (68)
Appropriate delineation		
Shift of border within 30m buffer(at least once beyond threshold)		35.8 % (515)
Perimeter shift beyond 25m buffer (proportion located outside)		16.6 % (238)
Jaccard Index		
Overall Index (range of single landscape elements)		0.81 % (0.37 - 1)

Results of plausibility analysis of the 100m buffer region for Pannonian biogeographical region

Sampling design		
No. of samples (equally distributed)		254
No. of samples within size classes		82 small/ 86 medium/ 86 large
Delineation plausibility within direct context (100m buffer)		
Sample characteristics (summary)		
Total area checked [ha] (buffer region without sample extent)		2923 ha
... containing Hardbone [km (%)]		125.1 km (34 %)
... containing Softbone [km (%)]		238.1 km (66 %)
Average line density [km/ha]		0.12 km/ha
Plausibility check		
Plausible Softbone lines within buffer region [%]		176.8 km (76 %)

STE – Steppic Biogeographic region

Results of geometric validation for Steppic biogeographical region

Sampling design		
No. of samples (equally distributed)		1439
No. of samples within size classes		479 small/ 478 medium/ 482 large
Delineation accuracy of geometric units:		
Appropriate size	Defined threshold of landscape elements that can be 15% larger or smaller	Validation result
too large: <i>(+/- 5m tolerance)</i>	≤10% (≤ 144 Softbones)	11.5 % (166) 0.6 % (8)
too small: <i>(+/- 5m tolerance)</i>	≤15% (≤ 216 Softbones)	16.2 % (233) 4.2 % (61)
Appropriate delineation		
Shift of border within 30m buffer (at least once beyond threshold)		38.1 % (548)
Perimeter shift beyond 25m buffer (proportion located outside)		19.5 % (281)
Jaccard Index		
Overall Index (range of single landscape elements)		0.81 % (0.06 - 1)

Results of plausibility analysis of the 100m buffer region samples for Steppic biogeographical region

Sampling design		
No. of samples (equally distributed)		252
No. of samples within size classes		85 small/ 82 medium/ 85 large
Delineation plausibility within direct context (100m buffer)		
Sample characteristics (summary)		
Total area checked [ha] (buffer region without sample extent)		3300 ha
... containing Hardbone [km (%)]		151.3 km (42 %)
... containing Softbone [km (%)]		211.7 km (58 %)
Average line density [km/ha]		0.11 km/ha
Plausibility check		
Plausible Softbone lines within buffer region [%]		137.5 km (60.7 %)

DOM – French Overseas Departments

Results of geometric validation for French Overseas Departments (islands)

Sampling design		
No. of samples (equally distributed)	552	
No. of samples within size classes	185 small/ 182 medium/ 185 large	
Delineation accuracy of geometric units:		
Appropriate size	Defined threshold of landscape elements that can be 15% larger or smaller	Validation result
too large: <i>(+/- 5m tolerance)</i>	≤10% (≤ 55 Softbones)	33.9 % (187) 17.2 % (95)
too small: <i>(+/- 5m tolerance)</i>	≤15% (≤ 83 Softbones)	15.6 % (86) 8.2 % (45)
Appropriate delineation		
Shift of border within 30m buffer (at least once beyond threshold)	53.3 % (294)	
Perimeter shift beyond 25m buffer (proportion located outside)	44 % (243)	
Jaccard Index		
Overall Index (range of single landscape elements)	0.76 % (0.01 - 1)	

Results of plausibility analysis of the 100m buffer region for French Overseas Departments (islands)

Sampling design				
No. of samples (equally distributed)	79			
No. of samples within size classes	27 small/ 26 medium/ 26 large			
Delineation plausibility within direct context (100m buffer)				
Sample characteristics (summary)				
Total area checked [ha] (buffer region without sample extent)	781 ha			
... containing Hardbone [km (%)]	48.1 km (39 %)			
... containing Softbone [km (%)]	73.8 km (61 %)			
Average line density [km/ha]	0.16 km/ha			
Plausibility check				
Plausible Softbone lines within buffer region [%]	62.4 km (76 %)			

TRF – Tropical Rainforest

Results of geometric validation for Tropical Rainforest biogeographical region

Sampling design		
No. of samples (equally distributed)		1265
No. of samples within size classes		411 small/ 429 medium/ 425 large
Delineation accuracy of geometric units:		
Appropriate size	Defined threshold of landscape elements that can be 15% larger or smaller	Validation result
too large: <i>(+/- 5m tolerance)</i>	≤10% (≤ 127 Softbones)	30 % (380) 19.5 % (247)
too small: <i>(+/- 5m tolerance)</i>	≤15% (≤ 190 Softbones)	28.4 % (359) 13 % (164)
Appropriate delineation		
Shift of border within 30m buffer(at least once beyond threshold)		77.2 % (977)
Perimeter shift beyond 25m buffer (proportion located outside)		68 % (860)
Jaccard Index		
Overall Index (range of single landscape elements)		0.68 % (0.01 - 1)

Results of plausibility analysis of the 100m buffer region for Tropical Rainforest biogeographical region

Sampling design		
No. of samples (equally distributed)		243
No. of samples within size classes		81 small/ 79 medium/ 83 large
Delineation plausibility within direct context (100m buffer)		
Sample characteristics (summary)		
Total area checked [ha] (buffer region without sample extent)		6729 ha
... containing Hardbone [km (%)]		54 km (19 %)
... containing Softbone [km (%)]		230.9 km (81 %)
Average line density [km/ha]		0.04 km/ha
Plausibility check		
Plausible Softbone lines within buffer region [%]		172.1 km (79.9 %)