

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

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

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

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

NATURA 2000 VALIDATION REPORT



Consortium Partners

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

Document Preparation and Release

	Affiliation	Name(s)	Date	Signature
Author	SIRS	Alexandre Pennec	24/07/2020	
Contributions	SIRS	Christophe Sannier	31/07/2020	
Review	SIRS	Alexandre Pennec	31/07/2020	
Approval	EEA	Hans Dufourmont		

Document Issue Record

Issue	Date	Author(s)	Description of Change
1.0	24/07/2020	Alexandre Pennec	First issue
1.1	31/07/2020	Christophe Sannier	Review and minor corrections
1.2	28/10/2020	Alexandre Pennec	Integration of feedbacks from EEA and SPs

Applicable Documents

ID	Title
AD01	Tender Specifications « Open Call for Tenders n° EEA/MDI/14/010 - Framework service contract for the GMES Initial Operations / Copernicus Land monitoring services – Validation Services for the geospatial products of the Copernicus land Continental and local components including in-situ data (lot 1)”
AD02	Framework service contract No EEA/MDI/14/010
AD03	Second Specific Contract N°3436/RO-COPERNICUS/EEA.56601
AD04	Third Specific Contract N°3436/RO-COPERNICUS/EEA.57056
AD05	Fourth Specific Contract - No 3436/RO-COPERNICUS/EEA.57889
AD06	Natura 2000 validation report - Second Specific Contract (https://land.copernicus.eu/user-corner/technical-library/n2k-validation-report)
AD07	Natura 2000 nomenclature and mapping guideline (https://land.copernicus.eu/user-corner/technical-library/N2K_Nomenclature_Guidelines.pdf)

Executive Summary

This report provides the evaluation results of the Natura 2000 data layers, specifically the Land Cover Land Use (LCLU) status and change products for 2006 and 2012. This analysis was performed over 100% of the N2K coverage for each of the Natura 2000 sites during the fourth Specific Contract of the validation project. The validation process consisted in the implementation thematic accuracy assessments. Detailed completeness and logical consistency checks are already performed as part of the semantic checks undertaken by the QC Tool during upload of the products.

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

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

Results for the thematic validation of the LCLU status products show that these datasets meet the validation requirement for both blind and plausibility analysis. During the second Specific Contract, the blind interpretation results showed substantially lower overall accuracies. Nevertheless, the revision of the nomenclature undertaken in 2017 which resulted in a reduction of classes and the experience of the validation team acquired over the years conduct to greater blind results during the Fourth Specific Contract. The analysis of the validation results at class and disaggregated level should provide insights on where the product could be improved.

In addition, the **MAES level 4 change product also exceeds the overall accuracy requirement of 80%**, but it should be noted that the amount of change only represents about 1% of the total area.

Based on the current exercise, several **recommendations** are given:

1. The **QC Tool reports** should be generated and made available openly as additional quality information to the users.
2. The N2K product was accompanied by layers which indicate the **scene name information** but the **EO data source could not be retrieved during the exercise and should be made available** to the validation team at the beginning of the bulk interpretation.
3. Some **auxiliary data sources** may have sometimes been used in the production, but this was not always entirely clear to the validation team and these could be detailed in an "Auxiliary Source Identification Layer".
4. The **analysis of the validation results at class and disaggregated geographical level** should provide insights on where the product could be improved thematically focusing on weaker classes and geographical regions.

Table of Contents

.....	0
1. Validation Framework	8
1.1. Products to be validated.....	8
1.2. Validation Criteria.....	11
2. Validation approach	12
2.1. Completeness	12
2.2. Logical consistency	12
2.2.1. Conceptual consistency.....	12
2.2.2. Domain consistency	12
2.2.3. Format consistency	12
2.2.4. Topological consistency.....	13
2.2.5. Additional logical consistency checks	13
2.3. Positional Accuracy.....	13
2.4. Thematic Accuracy.....	14
2.4.1. Level of reporting	14
2.4.2. Stratification and sample design.....	15
2.4.3. Response Design.....	19
2.4.4. Estimation and analyses procedures.....	21
2.5. Temporal Quality	21
2.6. Usability	21
2.7. INSPIRE compliant metadata	21
3. Validation check list.....	22
4. Additional Tables.....	30
4.1. Coordinate reference system description	30
4.2. Nomenclature	31
5. Thematic accuracy.....	33
5.1. LCLU 2006 and 2012 status layer	33
5.2. LCLU Change layer	37
6. Conclusions and recommendations	40
6.1. Main findings	40
6.2. Recommendations.....	40
Annex 1: N2K confusion matrix for the blind and plausibility interpretation for the 2006 and 2012 status layers.....	41
Annex 2: N2K confusion matrix for the blind and plausibility interpretation for the 2006-2012 change layer	46

List of Figures

Figure 1: Natura 2000 product over EEA39 for the reference year 2012	8
Figure 2: N2K pre-defined nomenclature on the basis of MAES typology of ecosystems (Level 1).....	9
Figure 3: Level of reporting at pan-European level.....	14
Figure 4: Level of reporting accordingly to the biogeographical regions 2016	14
Figure 5: Simple random (left) and random systematic (right) sampling designs	15
Figure 6: Replicates and sub-replicates used on LUCAS grid	16
Figure 7: N2K interpretation example: sample units (green dot) along with the polygons from the original dataset (yellow polygon) for the reference year 2006 (imagery dataset underlaid: N2000_Data_Procurement)	20

List of Tables

Table 1: Detailed specification of the Natura 2000 LCLU product.....	10
Table 2: N2K LC/LU distribution of sample units per strata.....	17
Table 3: N2K LC/LU weight factors (section 2.4.4) to be applied to each stratum for constructing confusion matrices	18
Table 4: Description of the N2K coordinate reference system.....	30
Table 5: Detailed Nomenclature for the LC/LU dataset.....	31
Table 6: N2K 2006 and 2012 status products thematic overall accuracy (OA) for Blind and Plausibility interpretation per bio-geographical region (n is the number of observations, green colour corresponds to OA greater than 85%, brown within 10% and red below 10% considering 95% CI).....	33
Table 7: N2K 2006 and 2012 status products thematic producer accuracy (PA) and user accuracy (UA) for Blind and Plausibility interpretations based on level 1 nomenclature groupings (green colour corresponds to OA greater than 85%, brown within 10% and red below 10% considering 95% CI)	34
Table 8: N2K 2006 and 2012 status products thematic producer accuracy (PA) and user accuracy (UA) for Blind interpretation based on level 4 nomenclature groupings (green colour corresponds to OA greater than 85%, brown within 10% and red below 10% considering 95% CI).....	34
Table 9: N2K 2006 and 2012 status products thematic producer accuracy (PA) and user accuracy (UA) for Plausibility interpretation based on level 4 nomenclature groupings (green colour corresponds to OA greater than 85%, brown within 10% and red below 10% considering 95% CI).....	35
Table 10: N2K 2006 - 2012 change product thematic overall accuracy (OA) for Blind and Plausibility interpretation per bio-geographical region at level 4 of the nomenclature (n is the number of observations, green colour corresponds to OA greater than 80%, brown within 10% and red below 10% considering 95% CI)	37
Table 11: N2K 2006 - 2012 change product thematic producer accuracy (PA) and user accuracy (UA)for Blind and Plausibility interpretation at level 1 of the nomenclature (green colour corresponds to OA greater than 80%, brown within 10% and red below 10% considering 95% CI).....	38

List of Abbreviations

ALP	Alpine region
ATL	Atlantic region
BLS	Black Sea region
BOR	Boreal region
CON	Continental region
EEA	European Environment Agency
EU-DEM	Digital Elevation Model over Europe
ESA	European Spatial Agency
ETC ULS	European Topic Center Urban land and soil system
EU	European Union
GIO	GMES Initial Operations
GMES	Global Monitoring for Environment and Security
HRL	High Resolution Layer
INSPIRE	Infrastructure for spatial information in Europe
LC/LU	Land Cover Land Use
LUCAS	Land Use/Cover Area frame Survey
MAES	Mapping of Assessment of Ecosystems and their Services
MED	Mediterranean region
MMU	Minimum Mapping Unit
MMW	Minimum Mapping Width
N2K	Natura 2000
NIR	Near Infra-Red
PAN	Pannonian region
PSU	Primary Sample Unit
STE	Steppic region
UA	Urban Atlas
VHR	Very High Resolution

1. Validation Framework

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

1.1. Products to be validated

Natura 2000 (an EU-wide network of nature protection areas established under the 1992 Habitats Directive) is the centrepiece of EU nature & biodiversity policy. The aim of the network is to assure the long-term survival of Europe's most valuable and threatened species and habitats that are listed under the Birds Directive and the Habitats Directive.

In 2015, a first iteration of the products was established. It was based on a selection of N2K grassland-rich sites (5 grassland habitats types 6210, 6240, 6250, 6510 and 6520), including a 2km buffer and covering approx. 160,000 km². The sites were mapped in order to assess their actual area, their condition and their development over time. The sites were analysed for the 2006 and 2012 reference years and a change analysis was performed. The mapping also included a 2km buffer zone where an analysis of pressures and threats was conducted. The analysis focused on a selection of grassland (semi-natural/species rich) habitat types.

In 2017, to allow for a harmonisation of the nomenclatures applied to the different local component products (Riparian Zones, N2K and the future Coastal Zone product), a revision of the nomenclature was undertaken which resulted in a reduction of classes. The N2K status layer now differentiates 55 thematic LC/LU classes (see section 4.2). Due to the revision of the nomenclature the original dataset had to be re-coded (the old dataset is discontinued) and the new dataset, with an increased area, offers currently a LC/LU product covering approximately 630,000 km².



Figure 1: Natura 2000 product over EEA39 for the reference year 2012

Land Cover/Land Use (LC/LU) classification is tailored to the needs of biodiversity monitoring. LC/LU is extracted from VHR satellite data and other available data in selected N2K sites (including a 2km buffer zone) for supporting biodiversity monitoring and mapping and assessment of ecosystems and their services. The classes follow the pre-defined nomenclature on the basis of MAES typology of ecosystems (Level 1 to Level 4) and Corine Land Cover. The Minimum Mapping Unit (MMU) is 0.5 ha; the Minimum Mapping Width (MMW) is 10m.

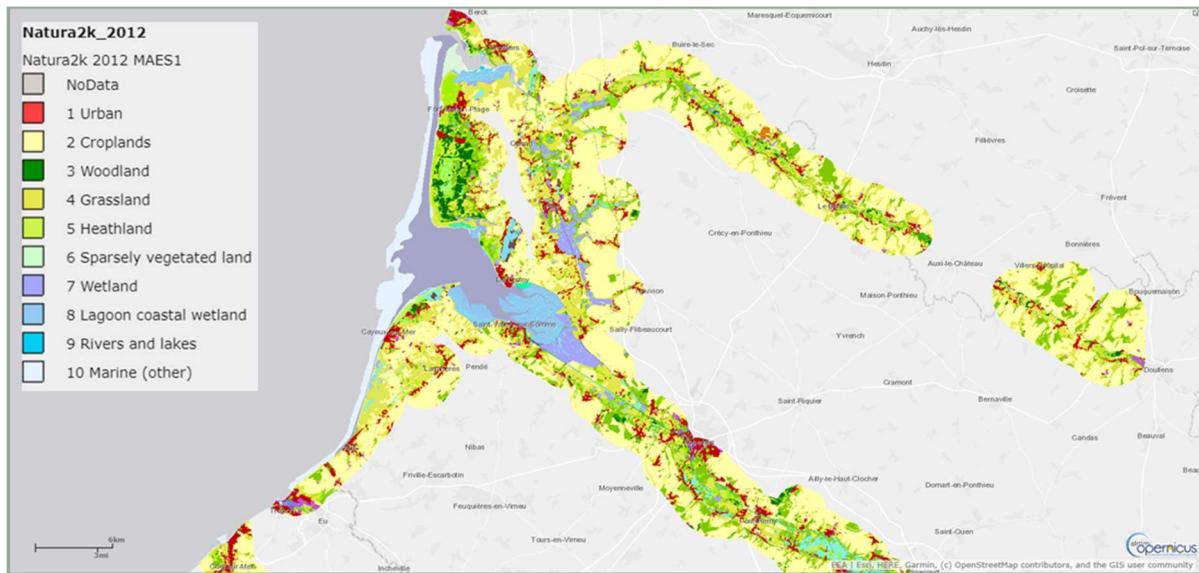


Figure 2: N2K pre-defined nomenclature on the basis of MAES typology of ecosystems (Level 1)

The dataset to be validated is composed of two products: Natura 2000 LC/LU mapping for 2006 and 2012 (status layers 2006 and 2012) and Natura 2000 LC/LU change mapping for 2006-2012 (change layer 2006-2012).

The detailed specifications of the Natura 2000 LC/LU products are shown below:

Table 1: Detailed specification of the Natura 2000 LCLU product

Product Title / Content	Product Short Name
Natura2000: LC/LU mapping of a selection of Natura2000 (N2K) sites.	LCLU
Product Definition	
<i>The Natura2000 product is providing a detailed LC/LU dataset for areas within buffer zone that comprises grassland habitats</i>	
Input Data Sources	
<p>1) Selected Natura2000 sites plus a 2km buffer zone</p> <p>2) Image data:</p> <p>Products:</p> <ul style="list-style-type: none"> • D2_MG2b_LOLA_011b • D2_MG2b_NARA_011b • DAP_MG2b_01 • DWH_MG2b_CORE_03 • DWH_MG2b_GEMS_ADD_003b • VHR_IMAGE_2015 • N2K_data_procurement <p>Missions:</p> <ul style="list-style-type: none"> • GeoEye1 (2m) • Pléiades (2.0m) • SPOT-5 HRG (2.5m) • SPOT-6 (1.5m) • WorldView-1 (1.6m) • WorldView-2 (2m) • QuickBird-2 (2.0m) 	
3) Additional data:	
<i>CLC 2006/2012/2018; Urban Atlas 2006/2012/2018; HR Layers; DWH_MG2_CORE_01 Coverage 1 (IRS 20m) & 2 (RapidEye, 5m); Landsat-8, Sentinel 2, Numerous reference and in-situ data sources.</i>	
Geographic Coverage)	
<i>EEA-39 (without Azores, Canarias and French DOMs):</i>	
Temporal Reference	
<i>Reference year 2006: 1.05.2004 – 30.09.2008</i>	
<i>Reference year 2012: 1.05.2010 – 29.09.2014</i>	
Projection	
<i>ETRS89 Lambert Azimuthal Equal Area (LAEA) (EPSG 3035)</i>	
Nomenclature	
<i>55 thematic classes</i>	
Minimum Mapping Unit	
<i>0.5 ha</i>	
Minimum Mapping Length	
<i>N/A</i>	
Minimum Mapping Width	
<i>10m</i>	

Product Title / Content	Product Short Name
Natura2000: LC/LU mapping of a selection of Natura2000 (N2K) sites.	LCLU
Thematic/Positional Product Accuracy	
<p>>85 %, taking into account the relative occurrence of the LC/LU classes</p>	
<p>>80 %, taking into account the relative occurrence of the LC/LU change classes.</p>	
<p><i>Positional accuracy is defined as < 5m.</i></p>	

1.2. Validation Criteria

The validation exercise focuses on thematic accuracy. The LC/LU expected overall accuracy is superior or equal to 85%, averaged at pan-European scale.

For the Change layer, changes expected overall accuracy is superior or equal to 80%, averaged at pan-European scale and for changes of Level 1. It should be noticed that the 80% requirement at Level 1 was not a product specification during the production phase but was applied to identify LC/LU change combinations, that might need further attention in future mappings

2. Validation approach

The validation approach will provide guidance on how the products will be validated by defining suitable indicators or metrics.

The elements of the quality assessment are presented below based INSPIRE Data Specifications for reference.

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

2.1. Completeness

Completeness should only be performed when the product is complete

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

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

2.2. Logical consistency

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

2.2.1. Conceptual consistency

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

Indicators:

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

2.2.2. Domain consistency

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

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

2.2.3. Format consistency

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

Indicators:

- File format conformance
- File name conformance

- Attribute names conformance
- Attribute types conformance

2.2.4. Topological consistency

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

Indicators:

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

2.2.5. Additional logical consistency checks

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

2.3. Positional Accuracy

Positional accuracy of Natura 2000 is directly related to the positional accuracy of the underlying VHR imagery.

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

2.4. Thematic Accuracy

2.4.1. Level of reporting

The level of reporting for the validation results is at pan-European level.

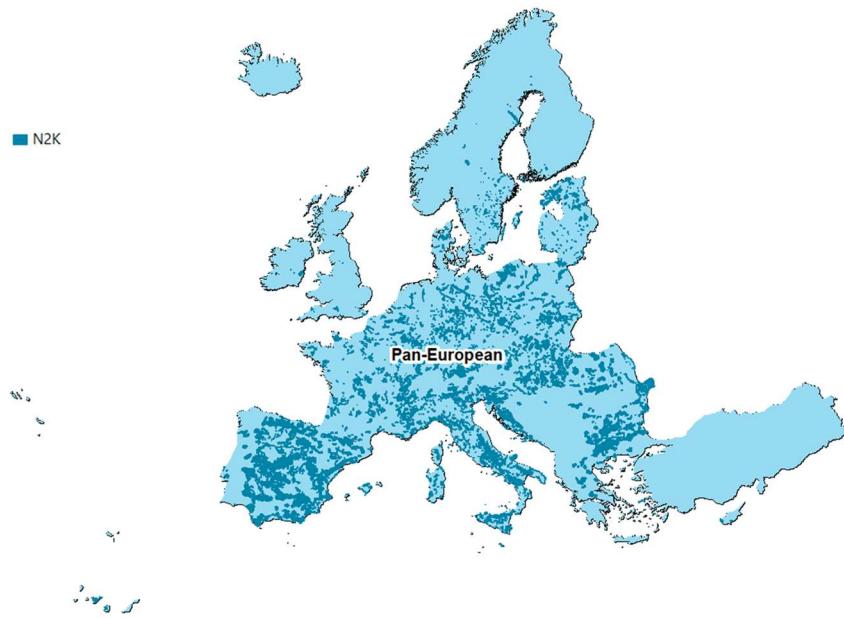


Figure 3: Level of reporting at pan-European level

However, results are also provided at different levels of aggregation:

- Results are provided by biogeographical regions 2016;

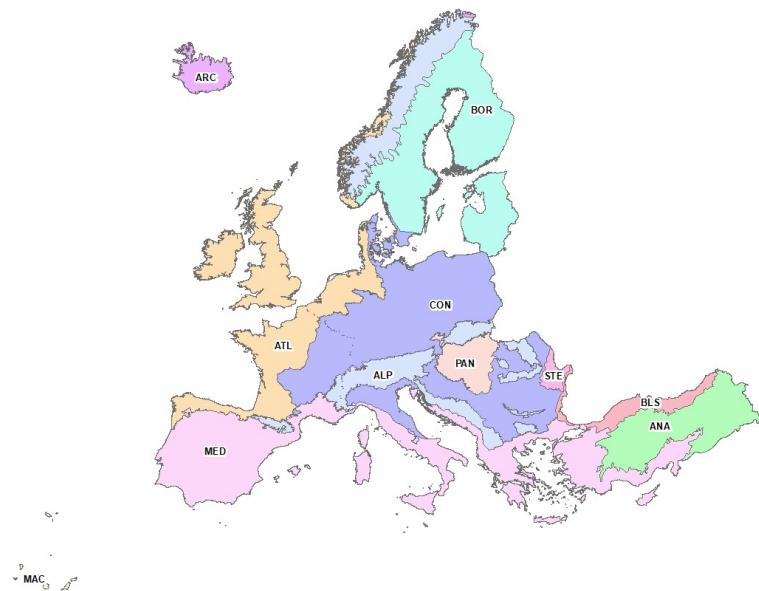


Figure 4: Level of reporting accordingly to the biogeographical regions 2016

Analysis of the results at a more disaggregated scale will contribute to assess regional differences if any and the causes of these differences. It should be noticed that no requirements are mandatory in the product specification for those results at disaggregated scale.

2.4.2. Stratification and sample design

2.4.2.1. Overview

The stratification and the sampling design primarily consists in selecting an appropriate sampling frame and sampling unit. The sampling units can either be “defined on a cartographic representation of the surveyed territory” (Gallego, 2004¹), in which case it is an area frame, or on a list of the features. According to (Gallego, 2004), area frames give a better representation of the population as the spatial dimension is kept.

In an area frame, sample units can be points, lines (often referred to as transects) or areas (often referred to as segments, described by Gallego, 1995²). The first step is to define the geographical area for which the accuracy assessment is to be reported and the type of sample units. For the majority of cases, point samples will be used, but areas or segments may be used in specific cases such as when not only thematic accuracy needs to be reported, but also the geometry of mapped objects. Points are considered as the most appropriate unit for our purpose. Polygons have also the drawback of being specific to a single map. In case of changes, the sample may not be adapted anymore.

Sampling design refers to the protocol whereby the samples are selected. A probability sampling design is preferred for its objectivity. “Simple random, stratified random, clustered random and systematic designs are all examples of probability sampling designs” (Stehman et al., 1998³). Even though a simple random design is easy to implement, its main drawback is that some portions of the population may not be adequately sampled. Cluster sampling is often used to reduce the costs of the collection of reference data but does not resolve geographic distribution problems. A systematic approach would solve this problem, yet it is not appropriate if the map contains cyclic patterns. A stratified approach consists in allocating a pre-defined number of samples per land-cover class. As explained in (Stehman et al., 1998), stratification ensures that each class is represented.

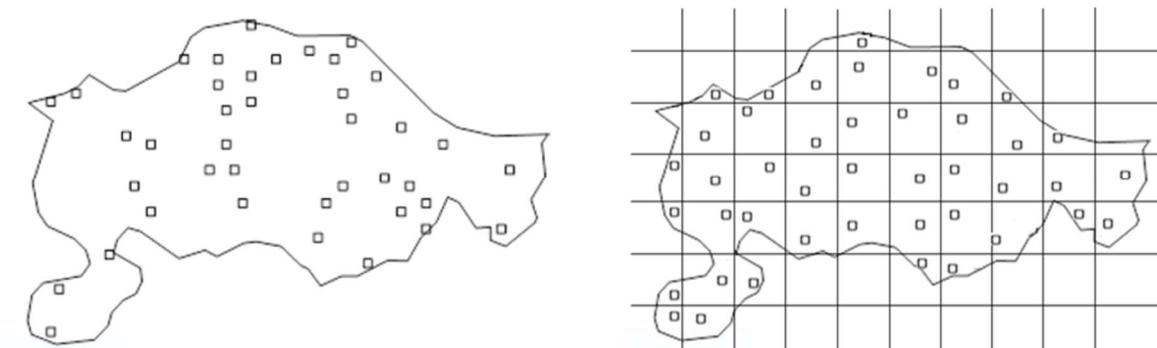


Figure 5: Simple random (left) and random systematic (right) sampling designs

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

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

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

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

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

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

A set of 81 points located on an 18x18 km square constitutes a group (red points shown in Figure 6) in which every point is associated with a number comprised between 0 and 81 (the numbers do not follow each other spatially). The same pattern with the same numbers allocation is repeated all over the grid. A replicate refers to the points with the same number selected on the whole LUCAS grid.

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

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

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

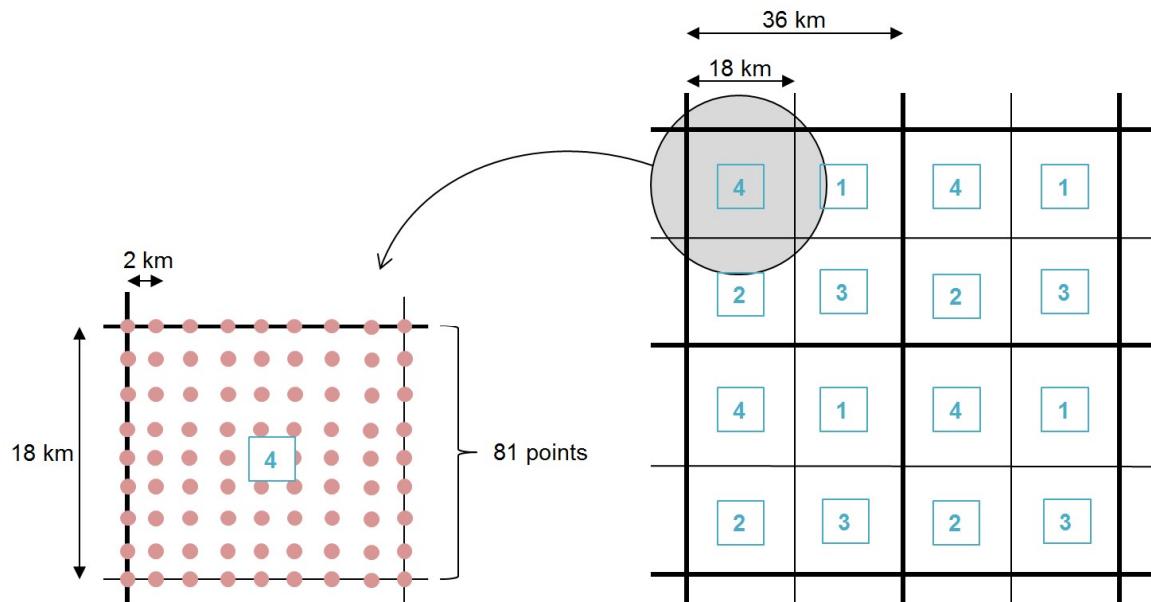


Figure 6: Replicates and sub-replicates used on LUCAS grid

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

2.4.2.2. Stratification

For Natura 2000, stratification is applied based on the LC/LU classes (Level 4) and LC/LU change classes based on the map products.

For the sampling design, LC/LU status and change classes mapped over small areas were aggregated respectively to one stratum called “Remain” and one stratum “Change” (resulting in one only stratum for all the changes). But this aggregation is only valid for the sampling design and each LC/LU classes are assessed individually (if sample units are drawn for the class considered) and thus included in the confusion matrix.

The number of sample units per stratum should be such to ensure a sufficient level of precision at reporting level. The minimum number of sample units per stratum should be set at 5 if possible.

The validation exercise covers the whole study area to be valid. There was a total of 34,932 sample units, covering about 625,000 km² distributed as shown in Table 2.

Table 2: N2K LC/LU distribution of sample units per strata

Strata	Class code	Number of sample units
1110	1110	938
1120	1120	280
1210	1210	189
1220	1220	72
1310	1310	99
1400	1400	135
2110	2110	5,872
2120	2120	59
2210	2210	577
2220	2220	478
2320	2320	155
2330	2330	164
2340	2340	935
3110	3110	6,261
3120	3120	66
3210	3210	4,446
3310	3310	1,134
3410	3410	1,051
3420	3420	63
4100	4100	2,629
4211	4211	446
4212	4212	2,412
4220	4220	500
5110	5110	654
5120	5120	156
5200	5200	1,700
6100	6100	624
6210	6210	62
6220	6220	74
6310	6310	339
6330	6330	68
7100	7100	304
7220	7220	111
8110	8110	77
8210	8210	88
9110	9110	209

Strata	Class code	Number of sample units
9210	9210	363
Remain	1230, 1240, 1320, 2310, 3220, 3320, 3500, 6320, 7210, 8120, 8130, 8220, 9120, 9130, 9220, 9230, 9240, 10000	413
Change	All change classes aggregated to one class	729
	Total	34,932

In order to ensure that unequal sampling intensities between strata are accounted for in the construction of the error matrix (section 2.4.4), weights are applied to each stratum Table 3).

Table 3: N2K LC/LU weight factors (section 2.4.4) to be applied to each stratum for constructing confusion matrices

Strata	Sample weight
1110	1.02813217
1120	0.893653101
1210	0.802596346
1220	0.333354445
1310	0.54782235
1400	0.690559961
2110	1.075632901
2120	0.207907444
2210	0.991437423
2220	0.971368182
2320	0.737975399
2330	0.759127442
2340	1.027935974
3110	1.076222971
3120	0.271088619
3210	1.072772917
3310	1.037514919
3410	1.033615227
3420	0.231784089
4100	1.064280396
4211	0.963315077
4212	1.062505101
4220	0.977878832
5110	1.002232643
5120	0.737674808
5200	1.053096584
6100	0.998771732
6210	0.225930557
6220	0.356644353
6310	0.925908411
6330	0.297776745
7100	0.908394296
7220	0.596189158
8110	0.386551754
8210	0.470615131
9110	0.828197762
9210	0.936549754
Change	0.225770884
Remain	0.235535435

2.4.3. Response Design

2.4.3.1. Overview

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

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

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

The double-blind approach was first applied, where the product information was not available during interpretation.

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

- VHR Image 2006 NIR with spatial resolution equal or greater than 2.5m (imagery used for the production and provided by GAF):
 - N2000_Data_Procurement
 - D2_MG2b_NARA_011b
 - DAP_MG2b_01
 - DWH_MG2_CORE_02_Image2006
- VHR Image 2012 NIR with spatial resolution equal or greater than 5m (imagery used for the production and provided by GAF):
 - DWH_MG2b_CORE_03
 - D2_MG2b_NARA_011b
 - D2_MG2b_LOLA_011b
 - VHR_IMAGE_2015
 - DAP_MG2b_01
 - DWH_MG2_CORE_01
 - Landsat_8
 - Landsat_5

Significant delays in the EO data source provision by the service provider have to be highlighted. It causes major delays in the beginning of the bulk interpretation which has been done under time pressure. Moreover, all the scenes used during the production were not retrieved by the service provider and so not providing to the validation team. Finally, 281 sample units were not covered with the 2006 EO data (corresponding to 19 scenes) and 180 not covered for 2012 (corresponding to 25 scenes). Nevertheless, the use of the imagery used during the production is required due to the complexity of the nomenclature of this product. For future validation exercise, EO data source should be made available to the validation team at the beginning of the bulk interpretation along with a PSIL file which would help the validation exercise.

As complementary data:

- LUCAS field survey photos
- Bing maps (ArcGIS Basemap layer, RGB imagery with varying spatial resolution)
- Google Earth imagery (Google commercial ArcGIS plugin, RGB imagery with varying spatial resolution)
- Google StreetView
- National and regional web mapping services (RGB and/or CIR imagery with varying spatial resolution)

The sample units were provided to the bulk interpretation team as one shapefile along with the polygons from the original dataset in which all the information on strata or Land-Cover/Use was removed to ensure the independence of the interpretation (see next section).

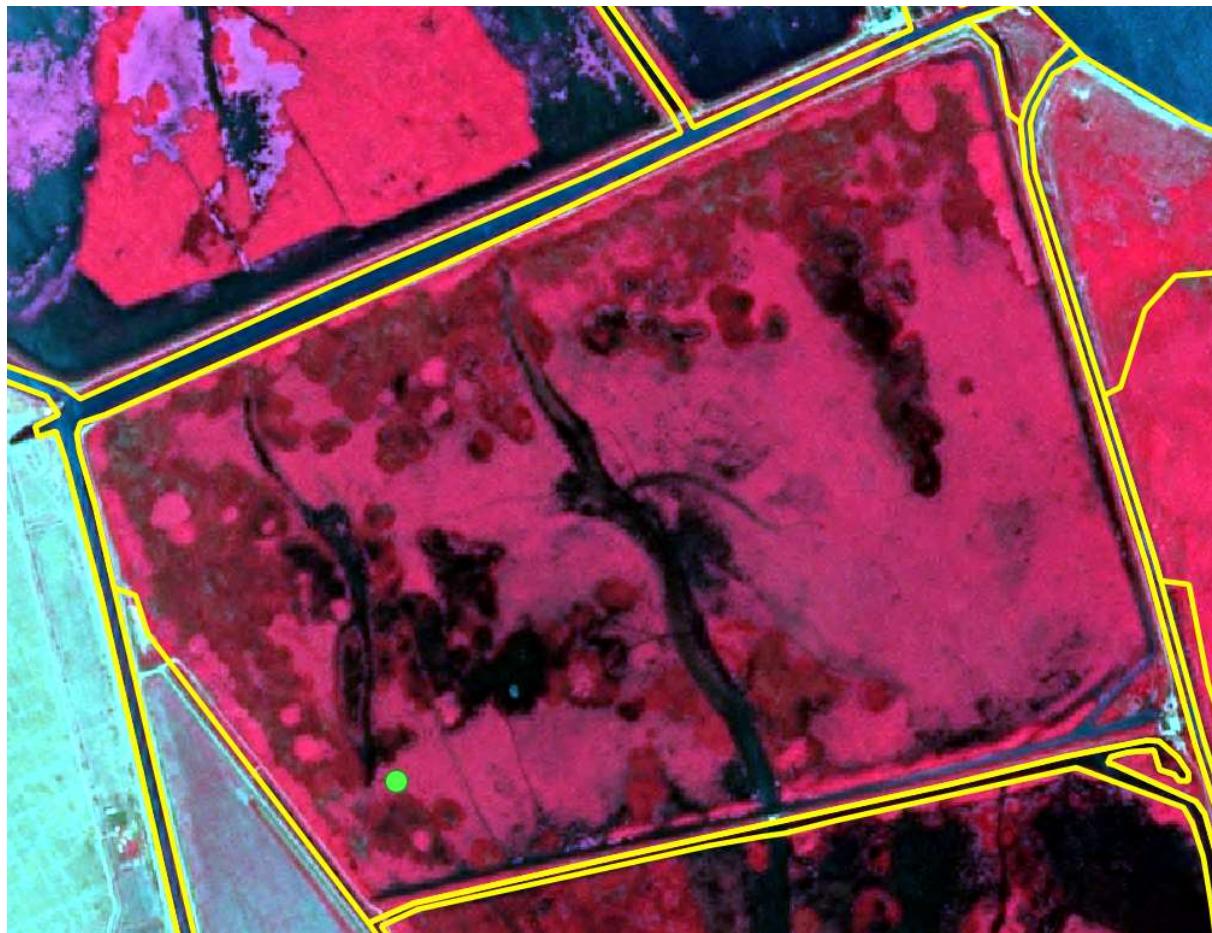


Figure 7: N2K interpretation example: sample units (green dot) along with the polygons from the original dataset (yellow polygon) for the reference year 2006 (imagery dataset underlaid: N2000_Data_Procurement)

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

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

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

The Plausibility Analysis is calculated with results from this second interpretation.

2.4.4. Estimation and analyses procedures

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

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

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

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

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

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

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

2.5. Temporal Quality

Temporal quality is evaluated by providing an indication of the closeness of the acquired image data to the reference year, e.g. the percentage area covered outside the accepted reference period as defined in the tender/product specification i.e. 2006/2012 +/- 1-2 year(s).

2.6. Usability

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

2.7. INSPIRE compliant metadata

Presence of INSPIRE compliant metadata should be verified.

3. Validation check list

This preliminary validation exercise was performed based on the 5 Natura 2000 sites randomly selected as part of the Look-and-Feel check reports. Detailed completeness and logical consistency checks were not performed as part of the semantic checks undertaken by the QC Tool (not available during the production phase and during the upload of the products). Therefore, the aim of this validation exercise is to perform some checks. We recommend however to produce the QC Tool reports and make them available openly as additional quality information to the users.

PRODUCT: N2K						
VALIDATION LEVEL: EEA39						
SERVICE PROVIDER: EFTAS		SERVICE USER: EEA		ISSUE/REVISION: 1.0		
VALIDATION DATE: 17/07/2020		REVIEW DATE: 24/07/2020				
CONDUCTED BY: SIRS		REVIEWED BY: Alexandre PENNEC		APPROVED BY:		
No.	Data Quality Sub-	Data Quality Measure	Data Quality Result	Comments by Audit Team	Draft Audit Conclusion	Final Audit Conclusion
1	COMPLETENESS					
1.1	Commission	Rate of excess items		No information provided by Look-and-Feel check reports.	N/A	
1.2	Omission	Rate of missing items		No information provided by Look-and-Feel check reports.	N/A	
2	LOGICAL CONSISTENCY					
2.1	Format consistency	File format	100%	All files are conformed to ESRI geodatabase	Accepted	
2.2		File name	100%	All File Names conforming	Accepted	

PRODUCT: N2K						
VALIDATION LEVEL: EEA39						
SERVICE PROVIDER: EFTAS		SERVICE USER: EEA		ISSUE/REVISION: 1.0		
VALIDATION DATE: 17/07/2020		REVIEW DATE: 24/07/2020				
CONDUCTED BY: SIRS		REVIEWED BY: Alexandre PENNEC		APPROVED BY:		
No.	Data Quality Sub-	Data Quality Measure	Data Quality Result	Comments by Audit Team		Draft Audit Conclusion
2.3		Attribute names	100%	All attribute names in file conforming		Accepted
2.4		Attribute types	All attribute types conforming according	All attribute types conforming		Accepted
2.5	Conceptual consistency	Feature type	N/A	No information provided by Look-and-Feel check reports		-
2.6		MMU		Checked by EEA QC tool during product upload. QC reports are not open access.		-

PRODUCT: N2K						
VALIDATION LEVEL: EEA39						
SERVICE PROVIDER: EFTAS		SERVICE USER: EEA		ISSUE/REVISION: 1.0		
VALIDATION DATE: 17/07/2020		REVIEW DATE: 24/07/2020				
CONDUCTED BY: SIRS		REVIEWED BY: Alexandre PENNEC		APPROVED BY:		
No.	Data Quality Sub-	Data Quality Measure	Data Quality Result	Comments by Audit Team	Draft Audit Conclusion	Final Audit Conclusion
	2.7	MMW		Checked by EEA QC tool during product upload. QC reports are not open access.	-	
2.7		Coordinate reference system	100%	File format conforms to specifications. File format consistency is checked by EEA QC tool during product upload. QC reports are not open access.	Accepted	
2.8		Unique identifier	100%	OK	Accepted	
2.9		Nomenclature	100%	Nomenclature conforms to specifications. Nomenclature consistency is checked by EEA QC tool during product upload. QC reports are not open access.	Accepted	
2.10	Domaine consistency	Value domain non-conformance	100%	Domaine consistency conforms to specifications	Accepted	

PRODUCT: N2K						
VALIDATION LEVEL: EEA39						
SERVICE PROVIDER: EFTAS		SERVICE USER: EEA		ISSUE/REVISION: 1.0		
VALIDATION DATE: 17/07/2020		REVIEW DATE: 24/07/2020				
CONDUCTED BY: SIRS		REVIEWED BY: Alexandre PENNEC		APPROVED BY:		
No.	Data Quality Sub-	Data Quality Measure	Data Quality Result	Comments by Audit Team	Draft Audit Conclusion	Final Audit Conclusion
2.11	Topological consistency	Overlaps		Checked by EEA QC tool during product upload. QC reports are not open access.	-	
2.12		Gaps		Checked by EEA QC tool during product upload. QC reports are not open access.	-	
2.13		Multipart features		Checked by EEA QC tool during product upload. QC reports are not open access.	-	
2.14		Neighbouring features		Checked by EEA QC tool during product upload. QC reports are not open access.	-	

PRODUCT: N2K						
VALIDATION LEVEL: EEA39						
SERVICE PROVIDER: EFTAS		SERVICE USER: EEA		ISSUE/REVISION: 1.0		
VALIDATION DATE: 17/07/2020		REVIEW DATE: 24/07/2020				
CONDUCTED BY: SIRS		REVIEWED BY: Alexandre PENNEC		APPROVED BY:		
No.	Data Quality Sub-	Data Quality Measure	Data Quality Result	Comments by Audit Team	Draft Audit Conclusion	Final Audit Conclusion
2.15		Self-intersections		Checked by EEA QC tool during product upload. QC reports are not open access.	-	
2.16		Null geometry		Checked by EEA QC tool during product upload. QC reports are not open access.	-	
2.17		Unclosed rings		Checked by EEA QC tool during product upload. QC reports are not open access.	-	
2.18		Duplicate vertex		Checked by EEA QC tool during product upload. QC reports are not open access.	-	
2.19		Pseudo nodes		Checked by EEA QC tool during product upload. QC reports are not open access.	-	

PRODUCT: N2K						
VALIDATION LEVEL: EEA39						
SERVICE PROVIDER: EFTAS		SERVICE USER: EEA		ISSUE/REVISION: 1.0		
VALIDATION DATE: 17/07/2020		REVIEW DATE: 24/07/2020				
CONDUCTED BY: SIRS		REVIEWED BY: Alexandre PENNEC		APPROVED BY:		
No.	Data Quality Sub-	Data Quality Measure	Data Quality Result	Comments by Audit Team	Draft Audit Conclusion	Final Audit Conclusion
2.20		Non matching nodes		Checked by EEA QC tool during product upload. QC reports are not open access.	-	
3	POSITIONAL ACCURACY					
3.1	Absolute or external accuracy	Positional accuracy of polygon	From 99% to 100%	Correct or shifted positional accuracy of polygons	Accepted	
4	THEMATIC ACCURACY					
4.1	Classification correctness	Overall accuracy	From 76% (IT9220135.) to 92% (PTCON0014)		Conditionally accepted	
4.2		Correctness of object delineation (i.e. omissions / commissions)	From 51% (FR2500088) to 78% (LV0536600)	The percentage of correctness of sampling delineation is much below the acceptable threshold.	Not accepted	

PRODUCT: N2K						
VALIDATION LEVEL: EEA39						
SERVICE PROVIDER: EFTAS		SERVICE USER: EEA		ISSUE/REVISION: 1.0		
VALIDATION DATE: 17/07/2020		REVIEW DATE: 24/07/2020				
CONDUCTED BY: SIRS		REVIEWED BY: Alexandre PENNEC		APPROVED BY:		
No.	Data Quality Sub-	Data Quality Measure	Data Quality Result	Comments by Audit Team		Draft Audit Conclusion
4.3		Polygon delineation detail	From 67% (FR2500088) to 83% (LV0536600)			Conditionally accepted
5	TEMPORAL QUALITY					
5.1	Temporal quality	Closeness of the acquired image data to the	N/A	Dependant on the VHR 2018 data set accuracy. No accuracy report available for VHR 2018 mosaic.		
6	USABILITY					
6.1	Usability	Usability description	N/A	N/A		
7	METADATA					
7.1	INSPIRE compliant metadata	Presence	100%	Metadata files are present		Accepted
7.2		File format	100%	All files are in XML format		Accepted
7.3		File name	100%	All files are named correctly		Accepted

PRODUCT: N2K							
VALIDATION LEVEL: EEA39							
SERVICE PROVIDER: EFTAS			SERVICE USER: EEA	ISSUE/REVISION: 1.0			
VALIDATION DATE: 17/07/2020			REVIEW DATE: 24/07/2020				
CONDUCTED BY: SIRS			REVIEWED BY: Alexandre PENNEC	APPROVED BY:			
No.	DATA QUALITY SUB-	DATA QUALITY MEASURE	DATA QUALITY RESULT	COMMENTS BY AUDIT TEAM	DRAFT AUDIT CONCLUSION	FINAL AUDIT CONCLUSION	
7.4		INSPIRE compliance	33%	INSPIRE compliant	Accepted		

4. Additional Tables

4.1. Coordinate reference system description

Table 4 lists the projection parameters of the N2K product.

Table 4: Description of the N2K coordinate reference system

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

4.2. Nomenclature

For information purpose, **Erreur ! Source du renvoi introuvable.** Table 5 below describes the latest LCLU nomenclature used for the N2K product.

Table 5: Detailed Nomenclature for the LC/LU dataset

Code	Class name
1110	Urban fabric (predominantly public and private units)
1120	Industrial, commercial and military units
1210	Road networks and associated land
1220	Railways and associated land
1230	Port areas and associated land
1240	Airports and associated land
1310	Mineral extraction, dump and construction sites
1320	Land without current use
1400	Green urban, sports and leisure facilities
2110	Arable land
2120	Greenhouses
2210	Vineyards, fruit trees and berry plantations
2220	Olive groves
2310	Annual crops associated with permanent crops
2320	Complex cultivation patterns
2330	Land principally occupied by agriculture with significant areas of natural vegetation
2340	Agro-forestry
3110	Natural and semi-natural broadleaved forest
3120	Highly artificial broadleaved plantations
3210	Natural and semi-natural coniferous forest
3220	Highly artificial coniferous plantations
3310	Natural and semi-natural mixed forest
3320	Highly artificial mixed plantations
3410	Transitional woodland and scrub
3420	Lines of trees and scrub
3500	Damaged forest
4100	Managed grassland
4211	Semi-natural grassland with woody plants (C.C.D. $\geq 30\%$)
4212	Semi-natural grassland without woody plants (C.C.D. $\leq 30\%$)
4220	Alpine and sub-alpine natural grassland
5110	Heathland and Moorland
5120	Other scrub land
5200	Sclerophyllous vegetation
6100	Sparsely vegetated areas
6210	Beaches and dunes
6220	River banks
6310	Bare rocks and rock debris
6320	Burnt areas (except burnt forest)
6330	Glaciers and perpetual snow
7100	Inland marshes
7210	Exploited peat bog
7220	Unexploited peat bog
8110	Coastal salt marshes
8120	Salines
8130	Intertidal flats
8210	Coastal lagoons
8220	Estuaries

Code	Class name
9110	Interconnected water courses
9120	Highly modified water courses and canals
9130	Separated water bodies belonging to the river system
9210	Natural water bodies
9220	Artificial standing water bodies
9230	Intensively managed fish ponds
9240	Standing water bodies of extractive industrial sites
10000	Sea and ocean

5. Thematic accuracy

5.1. LCLU 2006 and 2012 status layer

Table 6 summarizes overall accuracies obtained at pan-European level and by biogeographical region, for the blind and plausibility interpretations.

At pan-European level, the Natura 2000 LCLU status products the accuracy meets the minimum requirement (85%) for the overall accuracy for both blind and the plausibility analysis. The width of the 95% confidence intervals depends on (i) the number of sample units selected, (ii) the accuracy reached and (iii) the efficiency of the stratification. The stratification appears to be very efficient with 95% CI less than 1% achieved at pan-European. Indeed, based on the experience of previous Specific Contracts (stratification formerly based on the product class and the biogeographical regions) the stratification was simplified to reach this efficiency.

In addition, these results are relatively homogenous across the European bio-geographical regions. For the blind analysis, the accuracy requirement is met with substantial differences between the blind and plausibility results for the Steppic biogeographical region. The variability at bio-geographical region and country/group of countries level is also very low with 95% CI less than 1%. This suggests that the stratification is more efficient at disaggregated levels contrary to previous Specific Contracts resulting from the adaptation of the stratification strategy.

Table 6: N2K 2006 and 2012 status products thematic overall accuracy (OA) for Blind and Plausibility interpretation per bio-geographical region (n is the number of observations, green colour corresponds to OA greater than 85%, brown within 10% and red below 10% considering 95% CI)

	n	Blind				Plausibility			
		2006		2012		2006		2012	
		OA	95% CI	OA	95% CI	OA	95% CI	OA	95% CI
EEA39	34,932	92,09%	0,09%	91,07%	0,09%	95,65%	0,06%	95,82%	0,06%
ALP	4,849	89,97%	0,02%	87,64%	0,02%	94,63%	0,01%	94,55%	0,01%
ATL	2,704	90,29%	0,01%	89,58%	0,01%	94,78%	0,01%	94,68%	0,01%
BLS	373	97,89%	0,00%	96,57%	0,00%	98,47%	0,00%	98,79%	0,00%
BOR	1,351	90,26%	0,01%	91,31%	0,01%	96,77%	0,00%	97,05%	0,00%
CON	10,378	94,06%	0,03%	93,15%	0,04%	97,38%	0,02%	97,42%	0,02%
MAC	69	93,13%	0,00%	93,13%	0,00%	94,42%	0,00%	94,42%	0,00%
MED	12,977	91,32%	0,04%	91,32%	0,04%	94,58%	0,03%	95,01%	0,03%
PAN	1,662	94,32%	0,00%	91,28%	0,01%	96,18%	0,00%	96,13%	0,00%
STE	569	94,35%	0,00%	78,21%	0,01%	95,32%	0,00%	95,74%	0,00%

At biogeographical regions, for the plausibility analysis, the weakest results are obtained for the Macaronesia, Alpine and Atlantic regions with overall accuracies for the reference years 2006 and 2012 under the 95%. The best results are obtained for the Black Sea and Continental regions for which the accuracies are greater than 97%. Nevertheless, the results are all very satisfying and very homogenous.

Detailed confusion matrices are available in the Annex 1 for the status layers.

The results at European level based on level 1 nomenclature groupings are shown in Table 7.

Table 7: N2K 2006 and 2012 status products thematic producer accuracy (PA) and user accuracy (UA) for Blind and Plausibility interpretations based on level 1 nomenclature groupings (green colour corresponds to OA greater than 85%, brown within 10% and red below 10% considering 95% CI)

	Blind				Plausibility			
	2006		2012		2006		2012	
	PA	UA	PA	UA	PA	UA	PA	UA
1 - Urban	97,87%	98,97%	98,52%	99,03%	98,63%	99,12%	98,86%	99,33%
2 - Cropland	97,51%	98,86%	97,73%	98,67%	98,34%	99,22%	98,36%	99,23%
3 - Woodland and forest	98,49%	99,09%	98,15%	98,98%	98,77%	99,40%	98,75%	99,43%
4 - Grassland	95,57%	93,49%	94,86%	93,46%	96,61%	95,92%	96,55%	95,98%
5 - Heatland and scrub	93,22%	91,76%	93,30%	91,40%	95,14%	93,56%	95,25%	93,60%
6 - Sparsely vegetated land	91,50%	85,45%	88,81%	82,71%	95,44%	85,44%	95,82%	85,63%
7 - Wetland	90,17%	95,98%	89,20%	93,75%	89,49%	96,34%	90,13%	95,24%
8 - Lagoons, coastal wetlands and estuaries	96,69%	97,53%	94,68%	97,22%	96,91%	99,53%	96,90%	99,52%
9 - Rivers and lakes	98,77%	98,77%	99,66%	97,33%	99,11%	97,54%	99,74%	97,75%

The lower results are obtained for the class “6 - Wetland” with producer and user accuracies around 90% (10% of commission errors). The “9 – Rivers and lakes” class shows the better results user and the producer accuracies that nearly reach 100% in a plausibility approach (very few omission or commissions errors). The classes “1 – Urban”, “2 – Cropland” and “3 - Woodland and forest” also show very high results with both user and producer accuracies greater than 98%. Regarding the “3 - Woodland and forest” classes, it should be noted that the former density classes were dropped out which lead to greater results and which explains the large difference between the different iteration of the product. The “4 – Grassland”, “5 - Heatland and scrub”, “6 - Sparsely vegetated land” and “8 - Lagoons, coastal wetlands and estuaries” classes shows intermediate results with producer and user accuracies mostly greater than 95%.

The results at European level based on level 1 nomenclature groupings are shown in Table 8 and Table 9.

Table 8: N2K 2006 and 2012 status products thematic producer accuracy (PA) and user accuracy (UA) for Blind interpretation based on level 4 nomenclature groupings (green colour corresponds to OA greater than 85%, brown within 10% and red below 10% considering 95% CI)

n		Blind			
		2006		2012	
		PA	UA	PA	UA
1110	974	97,88%	99,57%	97,72%	99,37%
1120	332	98,03%	94,91%	98,08%	92,58%
1210	208	90,72%	99,47%	92,44%	99,34%
1220	72	100,00%	88,89%	96,20%	84,72%
1230	13	100,00%	100,00%	81,60%	100,00%
1240	33	100,00%	96,88%	100,00%	100,00%
1310	197	96,95%	91,85%	94,94%	93,88%
1320	46	49,39%	65,78%	84,62%	89,05%
1400	144	90,99%	90,18%	90,23%	89,61%
2110	5,947	96,59%	97,82%	96,82%	96,18%
2120	61	90,27%	88,75%	93,12%	95,10%
2210	583	94,97%	87,77%	96,71%	85,13%
2220	480	98,01%	93,11%	96,66%	93,47%
2310	24	7,59%	87,50%	4,83%	79,17%
2320	155	77,56%	83,06%	41,81%	75,48%

n	Blind				
	2006		2012		
	PA	UA	PA	UA	
2330	165	55,76%	76,03%	50,71%	73,83%
2340	938	98,78%	97,95%	98,73%	95,94%
3110	6,268	98,40%	95,12%	98,26%	94,59%
3120	66	55,02%	82,81%	45,82%	83,33%
3210	4,449	97,81%	90,99%	97,76%	90,91%
3220	6	6,12%	83,33%	3,57%	83,33%
3310	1,134	66,87%	92,49%	64,88%	90,74%
3320	1	14,88%	100,00%	17,96%	100,00%
3410	1,151	76,60%	87,88%	76,87%	88,26%
3420	63	42,49%	85,93%	36,23%	90,48%
3500	17	25,96%	81,44%	100,00%	76,47%
4100	2,676	98,00%	80,79%	97,63%	78,97%
4211	451	68,64%	84,94%	64,60%	81,73%
4212	2,475	82,46%	87,91%	81,15%	88,06%
4220	501	83,31%	98,00%	79,40%	97,40%
5110	679	95,95%	84,07%	95,20%	77,16%
5120	157	62,66%	87,65%	52,94%	89,12%
5200	1,736	92,43%	92,22%	92,88%	92,56%
6100	640	83,47%	71,56%	78,50%	68,22%
6210	64	80,57%	98,39%	73,36%	96,87%
6220	78	92,56%	96,11%	95,48%	94,77%
6310	339	86,83%	92,60%	89,31%	91,15%
6320	53	83,37%	48,32%	100,00%	94,35%
6330	68	97,48%	100,00%	88,97%	98,53%
7100	312	89,01%	96,83%	87,53%	94,36%
7210	34	70,08%	100,00%	70,08%	100,00%
7220	111	98,73%	90,99%	100,00%	89,19%
8110	77	100,00%	94,84%	100,00%	94,81%
8120	30	100,00%	100,00%	100,00%	96,67%
8130	3	46,42%	100,00%	100,00%	100,00%
8210	88	96,13%	98,86%	90,24%	98,86%
8220	9	100,00%	100,00%	100,00%	100,00%
9110	212	99,39%	99,74%	99,12%	99,39%
9120	45	100,00%	100,00%	100,00%	97,78%
9130	30	100,00%	93,67%	87,94%	96,67%
9210	391	98,47%	97,18%	99,28%	93,38%
9220	39	59,85%	93,33%	55,99%	94,82%
9230	37	80,02%	91,89%	79,58%	83,78%
9240	25	100,00%	79,05%	100,00%	95,93%
10000	45	100,00%	82,59%	100,00%	82,22%

Table 9: N2K 2006 and 2012 status products thematic producer accuracy (PA) and user accuracy (UA) for Plausibility interpretation based on level 4 nomenclature groupings (green colour corresponds to OA greater than 85%, brown within 10% and red below 10% considering 95% CI)

n	Plausibility				
	2006		2012		
	PA	UA	PA	UA	
1110	974	98,37%	99,57%	98,79%	99,47%
1120	332	98,40%	95,63%	98,48%	95,99%
1210	208	96,36%	99,47%	96,54%	99,20%

n	Plausibility				
	2006		2012		
	PA	UA	PA	UA	
1220	72	100,00%	98,61%	100,00%	98,61%
1230	13	100,00%	100,00%	100,00%	100,00%
1240	33	100,00%	100,00%	100,00%	100,00%
1310	197	98,12%	94,02%	94,09%	96,20%
1320	46	69,49%	74,60%	84,85%	91,24%
1400	144	91,30%	93,45%	91,50%	94,45%
2110	5,947	97,85%	98,36%	98,05%	98,37%
2120	61	96,14%	96,14%	100,00%	96,73%
2210	583	96,42%	91,82%	97,80%	95,33%
2220	480	98,84%	93,94%	99,03%	96,40%
2310	24	9,04%	79,17%	11,22%	79,17%
2320	155	85,63%	87,57%	88,40%	86,45%
2330	165	61,00%	88,80%	60,43%	89,65%
2340	938	99,12%	98,18%	98,95%	97,97%
3110	6,268	98,75%	98,46%	98,83%	98,63%
3120	66	73,28%	84,38%	69,09%	83,33%
3210	4,449	98,90%	96,85%	99,11%	96,85%
3220	6	7,17%	83,33%	4,90%	83,33%
3310	1,134	89,14%	95,82%	88,72%	95,86%
3320	1	6,88%	100,00%	9,88%	100,00%
3410	1,151	84,45%	92,12%	86,69%	93,64%
3420	63	65,85%	96,87%	65,50%	96,83%
3500	17	42,24%	81,44%	100,00%	76,47%
4100	2,676	98,91%	90,97%	98,81%	90,99%
4211	451	83,05%	88,74%	82,47%	88,66%
4212	2,475	91,57%	93,80%	91,67%	93,90%
4220	501	84,21%	99,60%	84,21%	99,60%
5110	679	99,27%	87,54%	99,57%	87,62%
5120	157	73,37%	93,60%	72,55%	93,60%
5200	1,736	94,43%	94,62%	94,54%	94,61%
6100	640	90,31%	74,28%	90,64%	74,43%
6210	64	81,09%	96,77%	82,73%	96,87%
6220	78	92,74%	98,70%	95,65%	98,69%
6310	339	92,41%	93,20%	92,38%	92,63%
6320	53	95,66%	70,81%	100,00%	92,47%
6330	68	100,00%	100,00%	95,63%	100,00%
7100	312	87,28%	97,72%	87,93%	96,24%
7210	34	97,26%	100,00%	97,26%	100,00%
7220	111	100,00%	90,09%	100,00%	90,09%
8110	77	96,93%	98,71%	96,91%	98,70%
8120	30	100,00%	100,00%	100,00%	100,00%
8130	3	100,00%	100,00%	100,00%	100,00%
8210	88	96,17%	100,00%	96,17%	100,00%
8220	9	100,00%	100,00%	100,00%	100,00%
9110	212	99,19%	99,74%	99,79%	99,87%
9120	45	100,00%	100,00%	100,00%	100,00%
9130	30	100,00%	93,67%	100,00%	96,67%
9210	391	99,05%	94,70%	99,68%	94,73%
9220	39	56,23%	93,33%	59,61%	94,82%
9230	37	81,34%	100,00%	90,05%	97,30%
9240	25	100,00%	79,05%	100,00%	100,00%
10000	45	100,00%	82,59%	100,00%	82,22%

Regarding the producer accuracy, the lower results are obtained for the classes “2310 - Annual crops associated with permanent crops”, “3220 - Highly artificial coniferous plantations” and “Highly artificial mixed plantations” with accuracies lower than 10% which indicates very high amount of omission errors (90%). The classes “2330 - Land principally occupied by agriculture with significant areas of natural vegetation”, “9220 - Other scrub land”, “3120 - Highly artificial broadleaved plantations”, “3420 Lines of trees and scrub” and “5120 - Other scrub land” also show very weak results with producer accuracies from 30 to 40% (60 to 70% of omission errors). The lower user accuracies are reached for the classes “2310 - Other scrub land”, “3500 - Damaged forest”, “6100 - Other scrub land” with results around 75-80% which indicates a level of commission errors around 20-25%. Nevertheless, most of the individual classes that exhibit lower accuracies tend to be limited to classes covering very small areas (See Annex 1).

The classes “4211 - Semi-natural grassland with woody plants (C.C.D. $\geq 30\%$)”, “4220 - Alpine and sub-alpine natural grassland”, “6210 - Beaches and dunes” show intermediate results regarding the producer accuracies (70-80%; 20-30% of commission errors). Concerning the user accuracies, intermediate precisions are obtained for the classes “3120 - Highly artificial broadleaved plantations”, “3220 - Highly artificial coniferous plantations” and “10000 - Sea and ocean” with accuracies around 82-83% (less than 20% of omission errors).

The classes “1220 - Railways and associated land”, “1230 - Port areas and associated land”, “1240 - Airports and associated land”, “8120 – Salines”, “8130 - Intertidal flats”, “8210 - Coastal lagoons”, “8220 –“ Estuaries”, “9110 - Interconnected water courses” and “9120 - Highly modified water courses and canals” show the better results user and the producer accuracies that nearly reach 100% in a plausibility approach (very few omission or commissions errors).

5.2. LCLU Change layer

The assessment of the change layer at level 4 of the nomenclature exhibits even higher overall accuracy results and exceeds the accuracy requirement threshold of 80% for both the blind (with the exception of the Black Sea region where marshes were interpreted as salt marshes in the blind interpretation, see above) and plausibility analysis (see Table 6). However, it should be noted that the amount of change detected is very small:

- 0.5% of the total area based on the map data

This means that the overall accuracy of the change layer is dominated by the unchanged area.

Table 10 summarizes overall accuracies obtained at pan-European level and by biogeographical region, for the blind and plausibility interpretations.

Table 10: N2K 2006 - 2012 change product thematic overall accuracy (OA) for Blind and Plausibility interpretation per biogeographical region at level 4 of the nomenclature (n is the number of observations, green colour corresponds to OA greater than 80%, brown within 10% and red below 10% considering 95% CI)

	n	Blind		Plausibility	
		OA	95% CI	OA	95% CI
EEA39	34,932	89,52%	0,10%	95,30%	0,06%
ALP	4,849	86,27%	0,02%	94,31%	0,01%
ATL	2,704	87,55%	0,01%	94,26%	0,01%
BLS	373	95,10%	0,00%	97,82%	0,00%
BOR	1,351	89,10%	0,01%	96,28%	0,00%
CON	10,378	91,79%	0,04%	97,01%	0,02%
MAC	69	93,13%	0,00%	94,42%	0,00%
MED	12,977	89,66%	0,04%	94,34%	0,03%

PAN	1,662	90,27%	0,01%	95,62%	0,00%
STE	569	76,08%	0,01%	94,69%	0,00%

At pan-European level, the Natura 2000 LCLU change products the accuracy meets the minimum requirement (80%) for the overall accuracy for both blind and the plausibility analysis.

At biogeographical regions, the lower results are obtained for the Alpine and Atlantic regions for the reference years 2006 and 2012 and the best results are obtained for the Black Sea and Continental regions. These results tend to confirm the results obtained for the status layers.

Table 11 summarizes producer and user accuracies obtained at pan-European level for the blind and plausibility interpretations at Level 1 of the nomenclature. Results are presented at Level 1 because too many combinations are possible at Level 4 which map small surfaces.

Table 11: N2K 2006 - 2012 change product thematic producer accuracy (PA) and user accuracy (UA) for Blind and Plausibility interpretation at level 1 of the nomenclature (green colour corresponds to OA greater than 80%, brown within 10% and red below 10% considering 95% CI)

n	Blind		Plausibility	
	PA	UA	PA	UA
1-1	1,842	98,35%	98,76%	98,24%
1-2	2	12,16%	50,00%	100,00%
1-3	5	0,00%	0,00%	0,00%
1-4	15	54,02%	66,67%	100,00%
1-5	3	29,57%	66,67%	100,00%
1-6	1	5,35%	100,00%	100,00%
1-9	4	100,00%	100,00%	100,00%
2-1	89	89,62%	88,76%	95,13%
2-2	8,264	98,25%	98,31%	97,54%
2-3	22	13,72%	59,09%	51,37%
2-4	42	6,82%	76,19%	35,03%
2-5	7	17,83%	28,57%	100,00%
2-6	3	100,00%	33,33%	100,00%
2-9	8	100,00%	100,00%	100,00%
3-1	25	78,46%	84,00%	95,45%
3-2	15	23,01%	73,33%	53,16%
3-3	13,048	98,73%	98,70%	98,14%
3-4	37	6,97%	51,35%	23,39%
3-5	14	6,67%	57,14%	59,60%
3-6	3	9,52%	66,67%	100,00%
3-7	1	9,85%	100,00%	17,93%
3-9	10	100,00%	80,00%	100,00%
4-1	98	77,55%	83,67%	77,05%
4-2	59	11,84%	69,49%	45,17%
4-3	43	8,89%	67,44%	27,69%
4-4	5,988	97,03%	91,98%	93,02%
4-5	15	4,35%	46,67%	100,00%
4-6	11	4,65%	45,45%	33,03%
4-7	3	17,51%	66,67%	29,82%

n	Blind		Plausibility	
	PA	UA	PA	UA
4-9	17	59,98%	82,35%	62,79%
5-1	8	100,00%	62,50%	100,00%
5-2	9	6,25%	11,11%	33,33%
5-3	29	11,73%	68,97%	57,30%
5-4	13	4,46%	53,85%	77,78%
5-5	2,510	94,17%	89,55%	91,05%
5-6	51	31,17%	84,31%	79,50%
6-3	8	24,25%	87,50%	73,16%
6-4	7	1,29%	85,71%	16,10%
6-5	23	11,63%	56,52%	37,38%
6-6	1,170	91,81%	85,33%	90,70%
6-9	2	35,57%	100,00%	49,47%
7-2	3	31,04%	100,00%	100,00%
7-4	1	1,67%	100,00%	2,80%
7-7	449	92,16%	95,28%	89,35%
7-9	9	19,42%	100,00%	10,93%
8-1	1	100,00%	100,00%	100,00%
8-6	1	100,00%	100,00%	100,00%
8-8	207	97,93%	99,52%	93,73%
9-1	1	48,94%	100,00%	100,00%
9-2	1	17,35%	100,00%	17,35%
9-6	2	0,00%	0,00%	N/A
9-7	4	8,53%	75,00%	36,62%
9-9	729	99,33%	97,06%	99,21%
				97,06%

It should be emphasized that most of the change classes that exhibit lower accuracies tend to be limited to classes covering very small areas. So, the following analysis of the plausibility results stresses on change classes that map significant areas (excluding the unchanged Level 1 classes).

Regarding the producer accuracy, the lower results are obtained for the classes “2-4 – Cropland-Grassland”, “3-4 – Woodland-Grassland”, “4-2 – Grassland-Cropland”, “4-3 – Grassland-Woodland” and “5-3 – Heathland-Woodland” which indicates very high amount of omission errors. The lower user accuracies are reached for the almost the same classes “3-4 – Woodland-Grassland”, “4-2 – Grassland-Cropland” and “4-3 – Grassland-Woodland”.

The classes “4-1 – Grassland-Urban” and “5-6 – Heathland-Sparsely vegetated land” show intermediate results regarding the producer accuracies (70-80%; 20-30% of commission errors). Concerning the user accuracies, intermediate precisions are obtained for the class “2-4 – Cropland-Grassland” with accuracies around 75% (less than 25% of omission errors).

The better results user and the producer accuracies in a plausibility approach are obtained for the class “2-1 – Cropland-urban” with accuracies that reached the 80% threshold (very few omission or commissions errors). The classes “4-1 – Grassland-urban” and “5-6 – Heathland-Sparsely vegetated land” show good user accuracies.

6. Conclusions and recommendations

6.1. Main findings

The N2K dataset, **both status and change products**, show **high overall thematic accuracy** for all the analyzed bio-geographical regions as well as the two interpreted dates (2006 and 2012), namely higher than the 90% threshold for the plausibility validation accuracy. This dataset is also homogeneous across biogeographical regions, with overall accuracies greater than the requirements (both 2006 and 2012).

Blind accuracy assessment results for both status and change layers are much greater than accuracies obtained in previous Specific Contract, but it is important to note that the very detailed and complex nomenclature with classes, difficult to characterize based on visual interpretation firstly, is now assimilated and has been simplified during the last revision of the product resulting in a simplification of the nomenclature (and harmonization with LC/LU products such as Riparian Zones).

Even though the MAES level 1 change product exceeds the overall accuracy requirement of 80%, the amount of change only represents about 0.5% of the total area. Therefore, the overall accuracy is dominated by the unchanged areas.

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

6.2. Recommendations

Based on the current exercise, several **recommendations** are given:

1. The **QC Tool reports should be generated** and made available openly as additional quality information to the users. Indeed, the QC Tool was not available during the production phase and during the upload of the products. These reports contain valuable information for the users such as detailed completeness and logical consistency checks.
2. The N2K product should be accompanied by layers which indicate the **scene name information** per polygon in the attribute table of the vector layers. Nevertheless, the **EO data source could not be retrieved during the exercise and should be made available to the validation team** at the beginning of the bulk interpretation. Significant delays in the EO data source provision by the service provider have to be highlighted. It causes major delays in the beginning of the bulk interpretation which has been done under time pressure. Moreover, scenes could not be retrieved by the service provider and so providing to the validation team. Finally, 281 sample units were not covered with the 2006 EO data (corresponding to 19 scenes) and 180 not covered for 2012 (corresponding to 25 scenes). Nevertheless, the use of the imagery used during the production is required due to the complexity of the nomenclature of this product. For future validation exercise, EO data source should be made available to the validation team at the beginning of the bulk interpretation along with a PSIL for the identification of source imagery file which would help the validation exercise.
3. Some **auxiliary data sources** may have sometimes been used in the production, but this was not always entirely clear to the validation team and these could be detailed in an “Auxiliary Source Identification Layer” similar to a PSIL used for the identification of source imagery.
4. The analysis of the **validation results at class and disaggregated geographical level** should provide insights on where the product could be improved thematically focusing on weaker classes and geographical regions and cases for which some thematic classes could be regrouped to improve their identification if required.

Annex 1: N2K confusion matrix for the blind and plausibility interpretation for the 2006 and 2012 status layers

Plausibility 2012	Reference																																																								
	1110	1120	1210	1220	1230	1240	1310	1320	1400	2110	2120	2210	2220	2310	2320	2330	2340	3110	3120	3210	3220	3310	3320	3410	3420	3500	4100	4211	4212	4220	5110	5120	5200	6100	6210	6220	6310	6320	6330	7100	7210	7220	8110	8120	8130	8210	8220	9110	9120	9130	9210	9220	9230	9240	10000	Total	User
Map	1110	967,38	1,03			1,03		1,03																																		972,52	99%	1%	1110												
	1120	2,68	251,46	0,89		2,91		2,01	0,89																															261,96	96%	4%	1120														
	1210		154,73			0,23		0,80																																	155,98	99%	1%	1210													
	1220	0,33	23,67																0,23																						24,00	99%	1%	1220													
	1230			3,06																																							3,06	100%	0%	1230											
	1240			7,77																																							7,77	100%	0%	1240											
	1310	1,00	0,23			73,46	0,23	0,77	0,23																														76,36	96%	4%	1310															
	1320	0,24	0,24			9,81		0,24																																10,76	91%	9%	1320														
	1400	3,45	0,23			89,97																																			95,26	94%	6%	1400													
	2110	1,08			1,08	6230,08		2,15	1,08	13,98	10,76	45,18	1,08																														6333,05	98%	2%	2110											
	2120					0,42	12,30																																		12,72	97%	3%	2120													
	2210					7,93		546,65		11,90	1,98	1,98		0,99					0,99																							573,41	95%	5%	2210												
	2220					0,97		3,14	448,03	5,83		0,97	0,97						2,91		0,97	0,97																					464,77	96%	4%	2220											
	2310	0,24						0,24	4,48		0,24	0,24	0,24																									5,65	79%	21%	2310																
	2320	0,74				6,64			2,95	98,89	3,69								0,74		0,74																114,39	86%	14%	2320																	
	2330	1,52				2,28		0,76	0,76		111,82	0,76																									124,72	90%	10%	2330																	
	2340							1,03		942,27	3,08								10,28		1,03	4,11														961,80	98%	2%	2340																		
	3110	2,15						1,08		1,08									3,23	6647,26	5,38	3,23	1,08	45,20	1,08	13,99	5,38			1,08	3,23												6739,81	99%	1%	3110											
	3120							0,27											1,36	14,91	0,27	0,54			0,27												17,89	83%	17%	3120																	
	3210					2,15					1,07	16,09		4620,04	15,02	84,75	1,07	12,87																										4770,23	97%	3%	3210										
	3220							0,24											1,04		1,18																			1,41	83%	17%	3220														
	3310																		1,04	14,53	23,86	2,08	1127,78		4,15																						1176,54	96%	4%	3310							
	3320																			0,24																						0,24	100%	0%	3320												
	3410					0,45	0,23	1,49		2,52	0,23		1,06		17,03	1,06				3,42	2555,63	30,86	101,56																				1108,91	94%	6%	3410											
	3420																		0,96	5,04	3,85		27,20		0,96	381,92	4,82		0,23	4,82	0,96		</																								

Annex 2: N2K confusion matrix for the blind and plausibility interpretation for the 2006-2012 change layer

