

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

**Third Specific Contract - N°3436/R0-
COPERNICUS/EEA.57056**

HRL Water and Wetness 2015 FINAL VALIDATION REPORT



Document: CLMS_HRL_WaW_Validation_Report_SC03

Issue: 1.2

Date: 20/02/2019

Consortium Partners

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

Document Preparation and Release

	Affiliation	Name(s)	Date
Author	e-GEOS	Quirico D'amico Nicola Corsini	18/12/2018
Contributions	SIRS	Renaud Vandeputte	19/12/2018
Review	SIRS	Alexandre Pennec	20/12/2018
Approval	EEA		

Document Issue Record

Issue	Date	Author(s)	Description of Change
1.0	20/12/2018	Quirico D'Amico Nicola Corsini	First version
1.1	05/02/2018	Alexandre Pennec	Review
1.2	20/02/2019	Quirico D'Amico Nicola Corsini	Final version

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	Copernicus Land Monitoring Service – High Resolution Layer Water and Wetness Product Specifications
AD04	Third Specific Contract N°3436/R0-COPERNICUS/EEA.57056

Executive Summary

This report covers the validation of the Water and Wetness (WaW) product, within the Copernicus High Resolution Layer (HRL).

WaW products are extracted from High Resolution (HR) satellite data and other available data sources for the area of the EEA39 (39 member states and affiliated countries to the European Environment Agency). The original 20m products maps the classes of (0) No water / no wet area, (1) permanent water, (2) temporary water, (3) permanent wetness and (4) temporary wetness, with a pixel based MMU (Minimum Mapping Unit). To map water and wetness, multi-temporal imagery from 2009, 2012 and 2015, as well as a seasonal time series of Medium-Resolution (MR) and SAR images, are used. The production is based on unsupervised classification followed by visual improvement of the classification results and derivation of water frequencies based on seasonal spectral composites and different biophysical indices such as e.g. NDVI, NDWI and NMDI.

The final result is a raster dataset of permanent and temporary water as well as permanent and temporary wet surfaces with a spatial resolution of 20 x 20m. The 20m classified Water and Wetness layer is aggregated to 100m for the complete European LAEA layer according to the procedure described in the "Wetness_Water_Technical_Specs_Public_v1_GV_v3.pdf" (section 6. Aggregation concept).

This report provides the analysis of the fully aggregated 100m spatial resolution WaW product.

A comprehensive assessment of the results from the blind interpretation and the plausibility analysis was made, to identify input data limitations as well as other causes beyond the classification errors. This assessment includes a quantitative analysis of the mapped water and wetness (Permanent/Temporary) density against reference data in a stratified systematic sampling scheme.

The results from the blind analysis show that the (0) No water / no wet area, (1) permanent water classes meets the technical specifications, whereas the (2) temporary water, (3) permanent wetness and (4) temporary wetness fails to meet the target requirements. The (0) No water / no wet area, (1) permanent water classes require a minimum thematic accuracy of 85 %, whereas the (2) temporary water, (3) permanent wetness and (4) temporary wetness require a minimum thematic accuracy of 80 %.

Significantly, improved results were obtained using a plausibility approach, which allowed for some variation in all classes. There is some variability detectable between biogeographical regions and on country-based analysis.

In general, sometimes the blind interpretation results mismatch with the WAW product due to the approach used. In particular the methodology classifies the occurrence of water or moisture in a physical sense, independently from other relevant parameters such as the vegetation cover, the ecological concept and the biogeographic environment, as it is the case for the blind analysis. The difference in the results from production and validation are then reduced in the plausibility analysis, where the production methodologies are considered.

Compared to the blind interpretation the product shows some difference regarding the persistence of water in reservoirs under construction. Here the automatic classification recorded a presence of water only in a few images classifying them as intermittent, whilst the blind analysis resulted in the presence of water as permanent since the construction of the dam was completed. This discrepancy between automatic classification and blind analysis was considered and rectified in the plausibility analysis where the product specification and related extraction methodology were taken into account and the validation parameters aligned accordingly.

Regarding the temporary wet classification, a low commission accuracy was registered in the blind analysis. The plausibility analysis improved values of commission error into a very good levels of accuracy this confirm the limits of photo-interpretation limits to detect the presence of moisture during the blind interpretation.

The thematic accuracy has a worst trend going towards regions with colder climatic characteristics, that may be explained with the difficulty of compiling image time series free from snow, ice and clouds.

Table of Contents

1. Validation Framework.....	8
1.1. Products to be validated.....	8
1.2. Validation Criteria.....	10
2. Validation approach	11
2.1. Completeness	11
2.2. Logical consistency	11
2.2.1. Conceptual consistency.....	11
2.2.2. Domain consistency	11
2.2.3. Format consistency	11
2.2.4. Topological consistency.....	12
2.3. Positional Accuracy.....	12
2.4. Thematic Accuracy.....	12
2.4.1. Level of reporting	12
2.4.2. Stratification and sample design.....	12
2.4.3. Response Design.....	20
2.4.4. Estimation and analyses procedures.....	23
2.5. Temporal Quality.....	23
2.6. Usability	23
2.7. INSPIRE compliant metadata.....	23
3. Validation check list.....	24
4. Thematic accuracy.....	29
5. Conclusions and recommendations	41

List of Figures

Figure 1. Aggregation by countries or group of countries with an area greater than 90,000km ² level	12
Figure 2. Simple random (left) and random systematic (right) sampling designs	13
Figure 3. LUCAS points located on a regular grid.....	14
Figure 4. Number of sample points as a function of the expected error rate for two accepted standard error values (after Wack et al. 2012)	15
Figure 5. Replicates and sub-replicates used on LUCAS grid	16
Figure 6. Example of SSUs organised in a 5x5 20m grid.....	17
Figure 7. Interpretation tool for reference data labelling for the blind interpretation (implemented by e-GEOS).....	22
Figure 8. Interpretation tool for reference data labelling for the plausibility analysis (implemented by e-GEOS).....	23
Figure 9. Overall blind accuracy per country and group of countries	30
Figure 10. Overall plausibility accuracy per country and group of countries	30
Figure 11. Overall blind accuracy per biogeographic region	31
Figure 12. Overall plausibility accuracy per biogeographic region	31

List of Tables

Table 1: Detailed specification of the 100m Water and Wetness product	10
Table 2: WaW Distribution of sample units per strata and level of stratification	18
Table 3: Waw Distribution of sample units per countries or group of them.....	19
Table 4: WaW Weight factors to be applied to each stratum and level of stratification for constructing confusion matrices	20
Table 5: WaW product blind thematic producer's accuracy results per total European area covered, per bio-geographical regions and Countries and group of countries greater than 90,000km ² , for the WaW classes: No water / no wet area (NWaW), Permanent water (PWa), Temporary water (TWa), Permanent wet area (PWe), Temporary wet area (TWe). As for the Permanent water and Dry areas, the values in green indicate accuracies greater than the 85% threshold, yellow greater than 75% and red less than 75% considering the 95% CI. As for the Temporary water, Permanent wet areas and Temporary wet areas, the values in green indicate accuracies greater than the 80% threshold, yellow greater than 70% and red less than 70% considering the 95% CI.....	33
Table 6: WaW product blind thematic user's accuracy results per total European area covered, per bio-geographical regions and Countries and group of countries greater than 90,000km ² , for the WaW classes: No water / no wet area (NWaW), Permanent water (PWa), Temporary water (TWa), Permanent wet area (PWe), Temporary wet area (TWe). As for the Permanent water and Dry areas, the values in green indicate accuracies greater than the 85% threshold, yellow greater than 75% and red less than 75% considering the 95% CI. As for the Temporary water, Permanent wet areas and Temporary wet areas, the values in green indicate accuracies greater than the 80% threshold, yellow greater than 70% and red less than 70% considering the 95% CI.....	35
Table 7: WaW product plausibility thematic producer's accuracy results per total European area covered, per bio-geographical regions and Countries and group of countries greater than 90,000km ² , for the WaW classes: No water / no wet area (NWaW), Permanent water (PWa), Temporary water (TWa), Permanent wet area (PWe), Temporary wet area (TWe). As for the Permanent water and Dry areas, the values in green indicate accuracies greater than the 85% threshold, yellow greater than 75% and red less than 75% considering the 95% CI. As for the Temporary water, Permanent wet areas and Temporary wet areas, the values in green indicate accuracies greater than the 80% threshold, yellow greater than 70% and red less than 70% considering the 95% CI.....	37
Table 8: WaW product plausibility thematic user's accuracy results per total European area covered, per bio-geographical regions and Countries and group of countries greater than 90,000km ² , for the WaW classes: No water / no wet area (NWaW), Permanent water (PWa), Temporary water (TWa), Permanent wet area (PWe), Temporary wet area (TWe). As for the Permanent water and Dry areas, the values in green indicate accuracies greater than the 85% threshold, yellow greater than 75% and red less than 75% considering the 95% CI. As for the Temporary water, Permanent wet areas and Temporary wet areas, the values in green indicate accuracies greater than the 80% threshold, yellow greater than 70% and red less than 70% considering the 95% CI.....	39

List of Abbreviations

CI	Confidence Interval
CLC	CORINE Land Cover
EEA	European Environment Agency
ETC ULS	European Topic Centre on Urban Land and Soil Systems
EU-DEM	Digital Elevation Model over Europe
ESA	European Spatial Agency
GIO	GMES Initial Operations
GMES	Global Monitoring for Environment and Security
HR	High Resolution
HRL	High Resolution Layer
LAEA	Lambert Azimuthal Equal-Area
LUCAS	Land Use/Cover Area frame Survey
MMU	Minimum Mapping Unit
MR	Medium Resolution
PSU	Primary Sample Unit
PW	Permanent water
PWET	Permanent wet
SP	Service Provider
SSU	Secondary Sample Unit
TCD	Tree Cover Density
TW	Temporary water
TWET	Temporary wet area
UA	Urban Atlas
VHR	Very High Resolution
WaW	Water and Wetness

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

Pan-European High Resolution Layers (HRL) provide information on specific land cover characteristics and are complementary to land cover / land use mapping such as in the CORINE land cover (CLC) datasets. The HRLs are produced from 20 m resolution satellite imagery through a combination of automatic processing and interactive rule-based classification.

5 themes have been identified so far, corresponding mainly with the main themes from CLC, i.e. imperviousness degree (IMD), tree cover density (TCD) and forest type (FTY), grasslands (GRA), water and wetness (WaW) and small woody features (SWF). Except the SWF product, which is a vector layer, the pixels of 20 by 20 m are aggregated into 100 by 100 m grid cells for final products.

Pan-European wall to wall products cover all EEA39 countries. They were produced in a combined centralized and decentralized approach, involving service industry through market mechanisms and participating countries through grant agreements.

The current validation exercise addresses the combined Water and Wetness product, which shows the occurrence of water and wet surfaces over the period from 2009 to 2015. This layer is based on multi-temporal and multi-seasonal optical high-resolution Landsat data as well as image data from Copernicus space and contributing missions made available through the ESA DWH mechanism. Also, this layer is based on radar information from 10 m resolution Sentinel-1 data on a pan-European basis. In detail the following datasets were used:

- Landsat seasonal composites LS-5, -6, -7 and -8
- Main indices (mNDWI, NDWI, NDVI)
- Sentinel-1A (10m spatial resolution)
- ENVISAT-ASAR Wide Swath (75m spatial resolution)

In addition, the following ancillary data were used:

- Corine Land Cover 2012
- Riparian Zones (LCLU)
- EU-DEM (30m spatial resolution)
- Topographic Wetness Index (TWI)

7 years optical and SAR imagery time series were used to generate seasonal image composites aiming at capturing the intra-annual dynamics. In particular one image composite per season (each season covered by 3 months) was used to generate the following products:

- The main Water and Wetness (WAW) product with defined classes of (1) permanent water, (2) temporary water, (3) permanent wetness and (4) temporary wetness;
- The additional expert product: Water & Wetness Probability Index (WWPI).

Specifically, the object of the present validation exercise is the main Water and Wetness product. Here below some key information about the WaW product are given, however full specifications are available in AD03. The Water and Wetness product fully replaces the previous 2012 “water” and “wetland” products and should be used as baseline for future assessments and trends. It is available in full spatial resolution of 20m x 20m as well as aggregated to 100m x 100m spatial resolution in European projection ETRS LAEA, and in national projections. The WaW provides combined thematic information on water and wetness for the 2009-2015 period on the following classes:

- **permanent water** (e.g. rivers, lakes): always water (water in at least 80% of all observations);
- **temporary water** (e.g. temporarily inundated areas): alternation of dry and water or alternation of wet and water (water in >25% to 80% of all observations, with varying degrees of wet and dry; water dominates over wet);

- **permanent wet areas** (e.g. areas of permanently high soil moisture): always wet (wet in at least ~60% of all observations, region dependent);
- **temporary wet areas** (e.g. areas of changing soil moisture): alternation of dry and wet (wet in >25% to 60% of all observations, with varying degrees of wet and dry; wet dominates over dry);
- **dry areas**: always dry (dry in at least 75% of all observations): always dry (dry in at least 75% of all observations)

The product shows the occurrence of water and indicates the degree of wetness in a physical sense, assessed independently of the actual vegetation cover and is thus not limited to a specific land cover class and their relative frequencies.

The product is available for Albania, Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Kosovo, Latvia, Liechtenstein, Lithuania, Luxembourg, Macedonia the former Yugoslavian Republic of, Malta, Montenegro, Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom.

1.2. Validation Criteria

The detailed specifications of the Water and Wetness 100m product is shown below:

Table 1: Detailed specification of the 100m Water and Wetness product

Water and Wetness 100m	Acronym	Product category
	WaW	Aggregated status layer
Reference year		
<i>Performed for a time period of seven years from 2009 to 2015 (+/- one year)</i>		
Methodology		
<i>Aggregation of the 20m x 20m classified Water and Wetness layer</i>		
Geometric resolution		
<i>Pixel resolution 100m x 100m, fully conform with the EEA reference grid</i>		
Coordinate Reference System		
<i>European ETRS89 LAEA projection</i>		
Geometric accuracy (positioning accuracy)		
<i>Less than half a pixel. According to ortho-rectified satellite image base delivered by ESA (IMAGE2015).</i>		
Thematic accuracy		
<i>The thematic accuracy expected depends on the class:</i>		
<ul style="list-style-type: none"> • Permanent water: target accuracy 85% • Temporary water: target accuracy 80% • Permanent wet areas: target accuracy 80% • Temporary wet areas: target accuracy 80% • Dry areas: target accuracy 85% 		
Data type		
<i>8bit unsigned raster with LZW compression</i>		
Minimum Mapping Unit (MMU)		
<i>Pixel-based (no MMU)</i>		
Raster coding (thematic pixel values)		
<i>0: No water / no wet area 1: Permanent water 2: Temporary water 3: Permanently wet areas (wetness) 4: Temporary wet areas (wetness) 254: Unclassifiable 255: Outside production unit</i>		
Metadata		
<i>XML metadata files according to INSPIRE metadata standards</i>		
Delivery format		
<i>Geo-TIFF accompanied by INSPIRE Mapping tables in .xlsx format</i>		

The 85% and 80% accuracy values must be understood as follows:

- 15% for commission errors and 15% for omission errors for the permanent water and dry area classes.
- 20% for commission errors and 20% for omission errors for the temporary water, permanent/temporary wet areas classes.

2. Validation approach

The validation approach will provide guidance on how the products will be validated by defining suitable indicators or metrics. Detailed completeness and logical consistency checks are performed as part of the semantic checks undertaken by ETC ULS for most products. Therefore, the aim of this validation exercise is not to repeat these, but to review the existing documentation and perform additional checks if deemed necessary.

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

Thematic accuracy will represent the bulk of the work undertaken as part of this validation exercise.

2.1. Completeness

Description: Completeness provides an indication for missing data or omission within the intended area.

Indicators: the rate of excess items is used for areas mapped beyond the intended area and the rate of missing items is used to verify gaps in the intended area to be mapped.

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:

- Coordinate Reference System
- Pixel size and origin
- Compliance between 20 m and 100 m Water and Wetness and 20 m Additional Support layer (see Verification Guidelines)
- Additional attributes, symbology

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: Not applicable to raster data

2.3. Positional Accuracy

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

Visual checks were undertaken in relation to imagery used for validation and during the collection of sample units.

2.4. Thematic Accuracy

2.4.1. Level of reporting

The level of reporting for the validation results is pan-European. However, results are also provided at different levels of aggregation (as indicated in the Request for Services for the first specific contracts):

- Biogeographical regions;
- Countries or group of countries with an area greater than 90,000km² level. Countries < 90,000 km² shall be grouped into contiguous groups of countries > 90,000km² as much as possible.

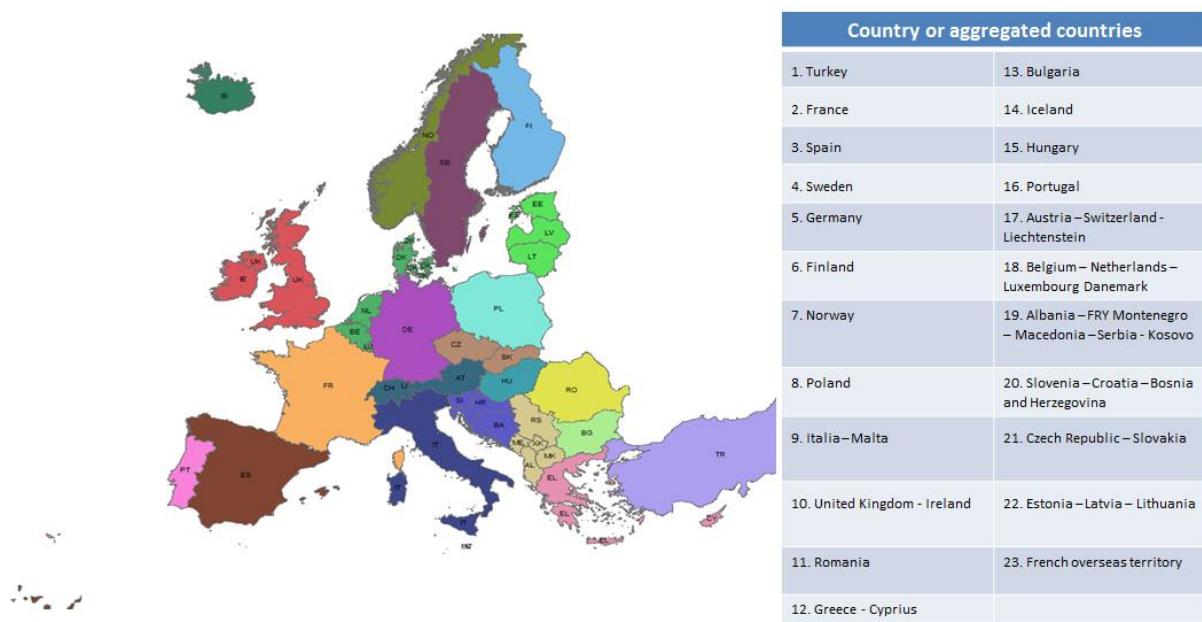


Figure 1. Aggregation by countries or group of countries with an area greater than 90,000km² level

Analysis of the results at a more disaggregated scale may contribute to assess regional differences, if any, and the causes of these differences.

2.4.2. Stratification and sample design

The following sub-sections provide a description of the procedure of a scientifically and statistically sound sampling scheme for assessing the thematic quality of the HRL Water and Wetness products. This comprises descriptions of the stratification approach, the sample size calculation procedure, and the strategy for ensuring representative sample distribution and sufficient regional spread.

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, 2004), in which case it is an area frame, or on a list of the features. According to (Gallego, 2004), area frames give a better representation of the population as the spatial dimension is kept.

In an area frame, sample units can be points, lines (often referred to as transects) or areas (often referred to as segments, described by Gallego, 1995). 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.*, 1998). 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.

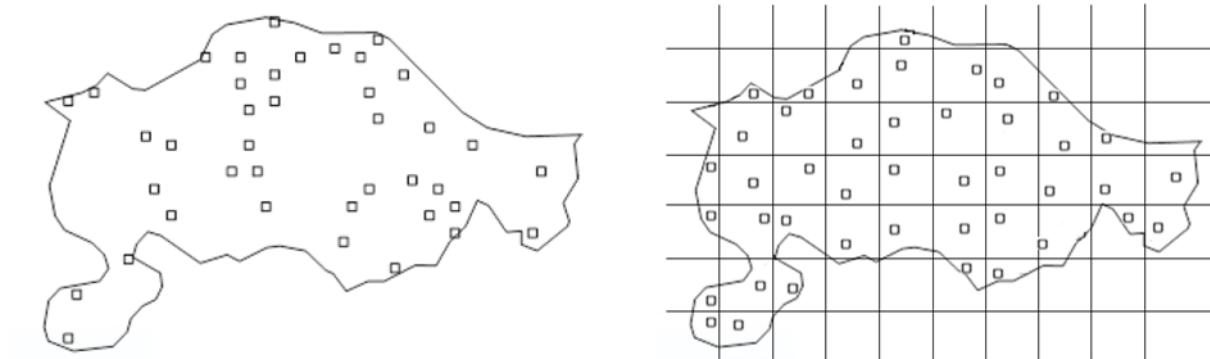


Figure 2. Simple random (left) and random systematic (right) sampling designs

The sampling and stratification design presented below is applicable to the HRL Water and Wetness product.

The validation approach of the Copernicus land Lot 1 external validation contract were adopted for the internal validation. It is based on a selected sample design for thematic accuracy assessment and combines systematic and stratified approaches and benefits from the advantages of both of them. It is based on the LUCAS (Land Use/Cover Area frame statistical Survey) sampling approach. LUCAS corresponds to a grid of approximatively 1,100,000 points throughout the European Union where land cover or land use type is observed. Using LUCAS points ensures traceability and coherence between the different layers.

LUCAS points are located every 2 km on a regular grid, as illustrated below. A set of 81 points located on an 18x18 km square constitutes a group in which every point is associated with a number comprised between 1 and 81 (the numbers do not follow each other spatially). The same pattern with the same numbers allocation is repeated all over the grid. A replicate refers to the points with the same number selected on the whole LUCAS grid.

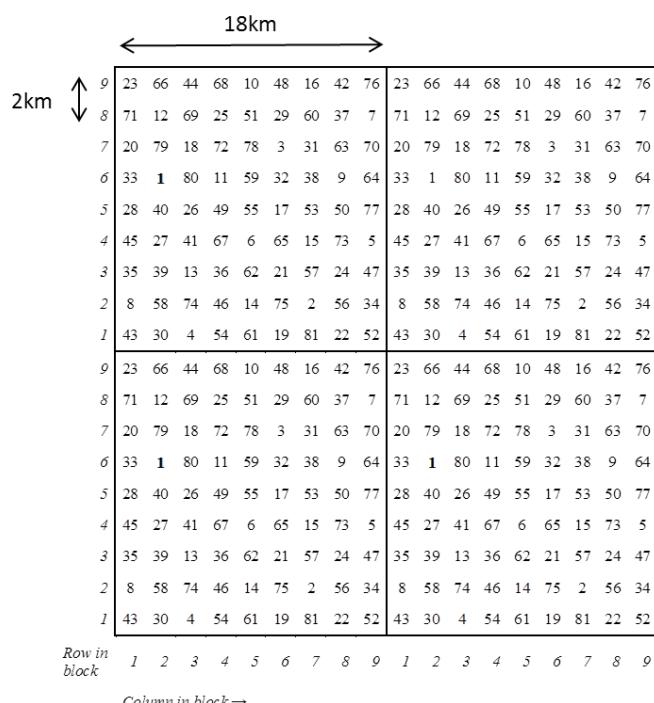


Figure 3. LUCAS points located on a regular grid

At first, the number of samples to allocate to each stratum (or thematic class) is calculated as a function of their area. In this manner the sampling design is not only systematic but also stratified. The number of sample units per stratum is to be defined to ensure sufficient level of precision at reporting level:

The determination of the number of sample units also considers the number of thematic classes.

It is possible to estimate a suitable sample size for each stratum based on the expected acceptable error rate.

The standard error of the error rate can be calculated as follows: $\sigma_h = \sqrt{\frac{p_h(1-p_h)}{n_h}}$ (1) where n_h is the sample size for stratum h and p_h is the expected error rate. This can be reworked to express the sample size n_h as a function of p_h and desired standard error σ_h : $n_h = \frac{p_h(1-p_h)}{\sigma_h^2}$. (2)

From Figure 4 it can be seen that for an expected 50% error rate, within a stratum, 100 sample units would be required to guarantee a standard error of 5%, whereas the number of samples would need to be increased by a factor of four if the accepted standard deviation is divided by a factor of 2. On the other hand if the expected error rate is 15%, only 51 samples would be necessary with a 5% standard error. A similar approach was adopted to determine the sample size for assessing the accuracy of CLC2006 and CLC2000-2006 changes (Büttner *et al.* 2012). This works well to assess commission errors, the definition of an appropriate number of sample units for omission errors is more difficult because it depends on the expected area of the theme to be mapped.

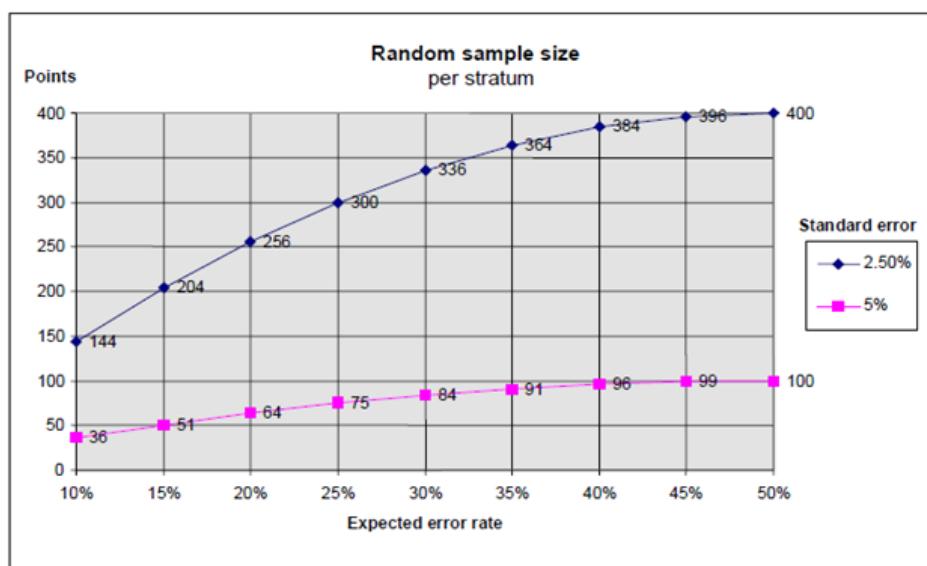


Figure 4. Number of sample points as a function of the expected error rate for two accepted standard error values (after Wack et al. 2012)

When using stratified sampling, the main issue to maximise the efficiency of the stratification (maximise the level of precision) is to optimize the sample allocation per stratum. A simple way is the use of equal allocation. Alternatively, the Neyman allocation algorithm is also often used for that purpose:

$$n_h = n * (N_h * \sigma_h) / [\Sigma (N_i * \sigma_i)], \quad (3)$$

where n_h is the sample size for stratum h , n is the total sample size, N_h is the population size for stratum h , and σ_h is the standard deviation of stratum h . According to Stehman (2012), Neyman optimal allocation should be preferred for estimating area of change as well as overall accuracy, whereas equal allocation is effective for estimating user's accuracy.

Moreover, in addition to the Neyman sample allocation, a minimum number of sample units per stratum is defined to ensure that even small strata are represented in the sample. As the expected error rate is expected to be less than 10%, a minimum of 50 sample units per reporting stratum should be sufficient.

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

For thematic 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 the Figure below.

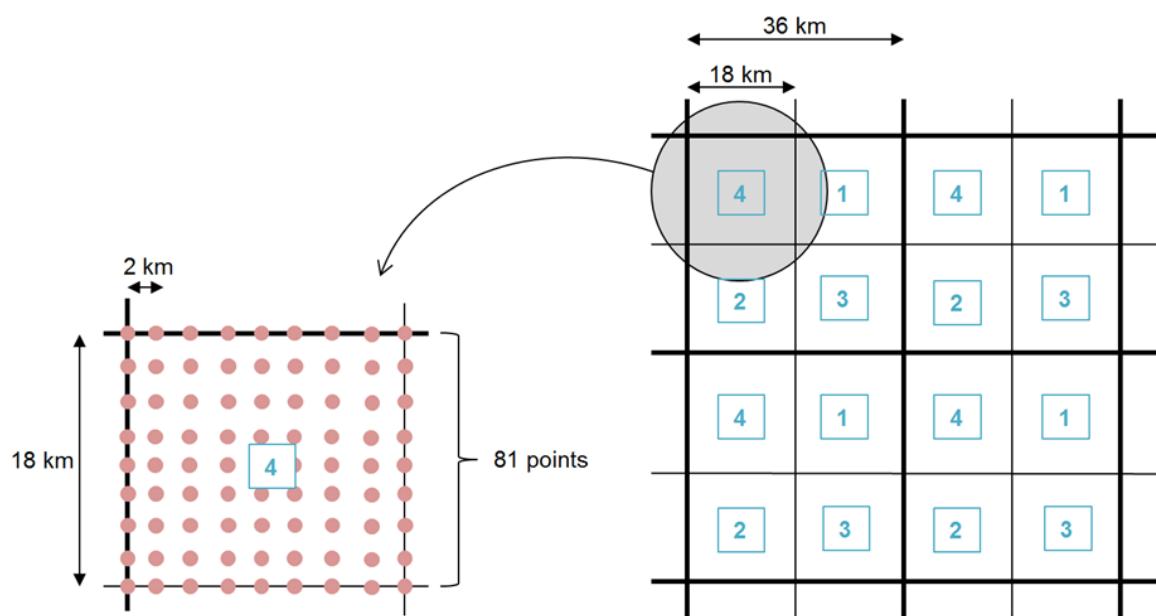


Figure 5. Replicates and sub-replicates used on LUCAS grid

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 could be densified by creating one point every 200 m.

2.4.2.2. Stratification approach

For the Water and Wetness product, a stratification based on a series of omission/commission strata is applied at two levels:

1. at pan-European level then
2. at countries or group of countries with an area greater than 90,000km² level if a minimum sample units of 20 is not reached for one omission/commission strata.

The number of primary sample units (PSUs) per stratum should be such to ensure a sufficient level of precision at reporting level. The minimum number of PSUs per stratum should be set at 20 if possible. Priority is given to strata which are known to be difficult to map.

The validation exercise covers the whole study area to be valid.

The first level of stratification is defined at pan-European level as follows:

- Commission “Permanent water”: Permanent water (minimum of 20 PSUs per country / group of countries)
- Commission “Temporary water”: Temporary water (minimum of 20 PSUs per country / group of countries)
- Commission “Permanent wet areas”: Permanent wet areas (minimum of 20 PSUs per country / group of countries)
- Commission “Temporary wet areas”: Temporary wet areas (minimum of 20 PSUs per country / group of countries)
- Omission High probability: No water / no wet area and CLC water and wet classes (minimum of 20 sample units per country / group of countries)

- Omission Low probability: Rest of the area (minimum of 20 sample units per country / group of countries)

The following CLC 2012 water and wet classes are defined as follows based on CLC2012:

- 4.1.1 = Inland marshes
- 4.1.2 = Peat bogs
- 4.2.1 = Salt marshes
- 4.2.2 = Salines
- 4.2.3 = Intertidal flats
- 5.1.1 = Water courses
- 5.1.2 = Water bodies
- 5.2.1 = Coastal lagoons
- 5.2.2 = Estuaries

If the minimum of 20 PSUs per country / group of countries is not reached at the first level of stratification, a second level is applied, per country to ensure the minimum number of PSUs.

Each PSU corresponds to one HRL pixel. Each PSU is then associated to secondary sampling units (SSUs) corresponding to a 5x5 grid with 20m between each SSU (Figure 6). The idea is that each SSU can then be associated with the corresponding HRL 20m layer pixel.



Figure 6. Example of SSUs organised in a 5x5 20m grid

Different sampling intensity are applied to focus on strata for which there is a higher probability that errors will be found. Weighting factors will be calculated based on the final sample selected to ensure that the different sampling intensities are accounted for when constructing confusion matrices to avoid the introduction of a bias toward these strata.

There was a total of 17,991 sample units for the WaW products (Table 2 - 3):

Table 2: WaW Distribution of sample units per strata and level of stratification

LABEL	Com. Permanent Water	Com. Temporary Water	Com. Permanent Wet	Com. Temporary Wet	Omission High Probability	Omission Low Probability	Total
Pan-European level	1,500	1,100	1,100	2,000	1,300	10,500	17,500
Country level	67	57	148	92	127	0	75
AL+ME+MK+ RS+XK	0	0	16	11	16	0	43
AT + CH + LI	0	13	17	15	17	0	62
BA + HR + SI	8	2	19	11	15	0	55
BE + LU+ NL + DK	0	18	5	0	3	0	26
BG	13	0	19	16	14	0	62
CZ + SK	14	14	20	9	18	0	75
DE	0	0	7	0	0	0	7
EL	12	0	13	11	15	0	51
ES	0	0	2	0	2	0	4
HU	5	10	14	0	0	0	29
IT	0	0	16	9	11	0	36
PT	15	0	0	10	16	0	41
TOTAL	1,567	1,157	1,248	292	1,427	10,500	17,991

Table 3: WaW Distribution of sample units per countries or group of them

LABEL	Com. Permanent Water	Com. Temporary Water	Com. Permanent Wet	Com. Temporary Wet	Omission High Probability	Omission Low Probability	Total
AL+ME+MK+RS +XK	31	30	20	20	20	302	423
AT + CH + LI	23	20	20	20	20	227	330
BA + HR + SI	20	20	20	20	20	253	353
BE + LU+ NL + DK	35	20	20	34	20	205	334
BG	20	21	20	20	20	214	315
CZ + SK	20	20	20	20	20	250	350
DE	55	33	20	63	20	668	859
EE + LT + LV	53	23	34	53	42	314	519
EL + CY	20	23	20	20	20	251	354
ES	20	118	20	42	20	997	1,217
FI	345	42	122	132	185	540	1,366
FR	42	59	43	47	24	1,063	1,278
HU	20	20	20	29	21	173	283
IE + UK	35	33	62	211	345	523	1,209
IS	26	114	472	161	43	110	926
IT + MT	27	31	20	20	20	586	704
NO	150	135	32	682	165	440	1,604
PL	36	25	28	58	21	589	757
PT	20	24	20	20	20	173	277
RO	25	29	29	39	37	440	599
SE	421	113	83	344	279	689	1,929
TR	123	204	103	37	45	1,493	2,005
TOTAL	1,567	1,157	1,248	2,092	1,427	10,500	17,991

To ensure that unequal inclusion probabilities are accounted for in the construction of the error matrix, weights are applied to each stratum as shown in Table 4:

Table 4: WaW Weight factors to be applied to each stratum and level of stratification for constructing confusion matrices

LABEL	Com. Permanent Water	Com. Temporary Water	Com. Permanent Wet	Com. Temporary Wet	Omission High Probability	Omission Low Probability
Pan-European level	0,26304204	0,02710602	0,05058723	0,28842560	0,21769303	1,58042362
Country level						
AL+ME+MK+RS+X K			0,00798887	0,35646594	0,09171215	
AT + CH + LI		0,01146804	0,00709659	0,07996368	0,04701692	
BA + HR + SI	0,22529461	0,25904893	0,00398166	0,31723244	0,07218590	
BE + LU + NL + DK		0,01506869	0,14624948		1,09765172	
BG	0,16906466		0,00635444	0,04351122	0,05616531	
CZ + SK	0,18408188	0,01492342	0,00329416	0,33898556	0,04380910	
DE			0,09226985			
EL	0,22988224		0,03800002	0,18049930	0,09675803	
ES			0,49082631		1,39638051	
HU	0,79555894	0,04130793	0,02326776			
IT			0,03727305	0,38607136	0,13940287	
PT	0,11760767			0,32803825	0,04110763	

The weights are calculated based on the interpreted sample units, should the number of sample units be modified per stratum following the removal of interpreted sample units classified as uncertain, weights would need to be recalculated. However, if the number of samples removed is small, weight changes should be minimal considering the large sample overall.

The sample units were provided to the bulk interpretation team as two separate shapefiles (one for each delivery batch, according to AD04) in which all the information on strata was removed to ensure the independence of the interpretation.

2.4.3. Response Design

2.4.3.1. Overview

The sample design is based on the LUCAS sample frame and some sample units will coincide with some LUCAS points. However, the LUCAS points are re-interpreted based on available in situ data. LUCAS thematic information is not used directly.

A blind approach was applied to start with, i.e. the validation is undertaken without considering the map layer information, then a plausibility analysis was implemented for both layers considering the information of the map layer. This will contribute to provide an in-depth understanding of the causes of classification errors.

For identification of WaW classes, the specification from the European Environment Agency (EEA) in the document "Wetness_Water_Technical_Specs_Public_v1_GV_v3.pdf" is strictly observed.

The main Water and Wetness product provides a combined thematic information on water and wetness for the 2009-2015 period on the following classes:

0. Dry area (dry in at least 75% of all observations)
1. Permanent Water (water in at least 80% of all observations) includes the following landscape types
 - Permanent inland lakes (natural)
 - Artificial ponds (permanent fish ponds, reservoir)
 - Natural ponds (permanent open water surfaces of inland or coastal wetlands)
 - Rivers
 - Channels (permanently with water)
 - Coastal water surfaces: lagoons, estuaries
 - Liquid dump sites (permanent)
 - Water surfaces with floating vegetation
2. Temporary Water (water in >25% to 80% of all observations) includes the following landscape types
 - Temporary water surfaces associated to permanent water bodies (e.g. oscillating shoreline areas of reservoirs)
 - Temporary natural (e.g. steppe) lakes and temporary artificial lakes (e.g. cassettes of fishponds)
 - Intermittent rivers and temporarily flooded river banks
 - Flood areas
 - Water-logged areas
 - Temporary flooded agricultural fields e.g. rice fields
 - Intertidal areas
 - Temporarily inundated areas (due to snow melt, floods or rain)
3. Permanent Wet (wet in at least in 60% of all observations) includes the following landscape types
 - Reeds
 - Peat land
 - Inland and coastal wetlands (incl. salt marshes)
4. Temporary Wet (wet in >25% to 60% of all observations) includes the following landscape types
 - Inland saline marshes
 - Intermittent wetlands
 - Temporary wet agricultural fields
 - Temporary wet meadows

The following elements were excluded from WaW:

- Sea and ocean (border between sea water and fresh water in river estuaries and coastal lagoons is determined by "EEA Coastline for analysis V.2" dataset)
- Permanent snow and glaciers
- Small river channels and streams with widths less than approx. 30 to 40m (mixed pixel phenomenon)
- Elements below the 20x20m MMU

2.4.3.2. Interpretation procedure

1. Blind Interpretation

The blind interpretation was based not only on the presence of water or moisture but taking into account the vegetation cover, the ecological concept, the biogeographic environment.

The thematic interpretation was based on a comparison between different images with different dates and seasonality. This comparison allowed to identify the water and wet classes and in particular to detail the temporary and permanent attributes.

The selected primary source data set used as thematic and geometric reference:

- VHR-IMAGE 2015 in CIR/RGB (preferably summer).

The selected a source data set used as thematic **comparison** in order to check the persistence of water:

- VHR-IMAGE 2015 in CIR/RGB (spring or different season compared to the reference).

The selected a source data set used as **ancillary** data in order to solve the uncertain cases:

- VHR-Image 2012
- HR-Image 2009, 2012
- If required use supporting data in the interpretation process (Basemap ArcGIS-World Imagery / Hillshade / DEM / EU-Hydro Layer / GoogleEarth / Bing Maps)

These comprehensive reference datasets provided a sound basis for the interpretation. The interpretation took into consideration seasonal time series to improve the understanding of the permanence of water or wet in the different seasons. For labelling of the reference data, a tool was implemented by e-GEOS, which is shown in the figure below (Figure 7). During the WaW validation the following information were assessed for each of the 25 points in every sample:

- i. Thematic class: the interpreter assigns here the class value on which the point is located: 0 = dry areas; 1 = Permanent Water; 2 = Temporary Water; 3 = Permanent Wet; 4 = Temporary Wet areas
- ii. Accuracy: the interpreter indicates the confidence of blind interpretation: Accurate (certain interpretation) or Approximate (in case of uncertain interpretation)
- iii. Comments: free text information filled only if necessary

All sampling units flagged as “Approximate” were subject to a second interpretation by another interpreter, and the accuracy parameter was updated where necessary (both interpreters agreeing on it). Further, the interpretation cycle included internal quality checks to ensure the quality of validation process.

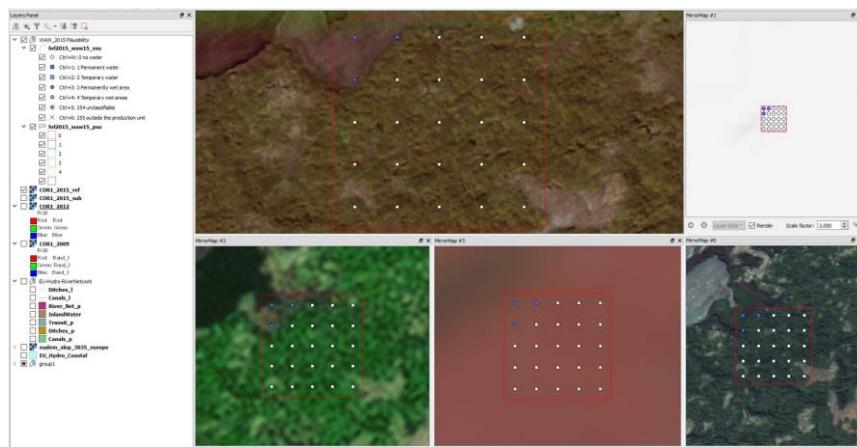


Figure 7. Interpretation tool for reference data labelling for the blind interpretation (implemented by e-GEOS)

2. Plausibility check

All plots showing a deviation between the WaW map product and the blind interpretation were re-interpreted. The same procedure as described above was applied, having the same image datasets as in the blind interpretation as reference plus the 20 m and 100m WaW map products (Figure 8. Interpretation tool for reference data labelling for the plausibility analysis (implemented by e-GEOS)). In particular a comparison with the 20 m and 100 m map products was made to evaluate the “plausible” errors or non-errors (not as strict as during the blind approach), eventually deciding which one between blind interpretation results and map product was plausible.

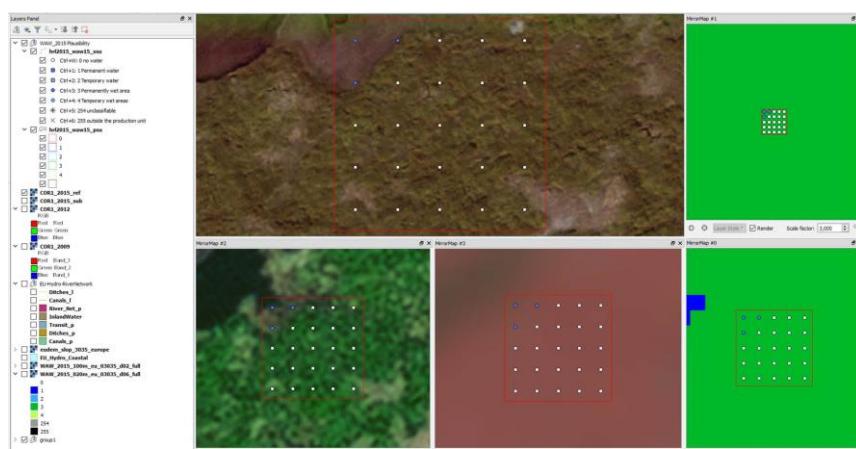


Figure 8. Interpretation tool for reference data labelling for the plausibility analysis (implemented by e-GEOS)

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

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. 2015 +/- 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 on the final semantic check report for the full pan-European 100m mosaic integrating some of the comments from the semantic check reports on lot deliveries.

PRODUCT:				HRL Water and Wetness (WaW)			
VALIDATION LEVEL:				Pan-European			
SERVICE PROVIDER: Multiple				SERVICE USER: EEA	ISSUE/REVISION: 1.0		
VALIDATION DATE: 21/12/2018				REVIEW DATE:			
CONDUCTED BY:				REVIEWED BY:	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	NOK	The following file are in excess compared to the delivery products described in WaW technical specification: WAW_2015_100m_eu_03035_d02_full.tif.vat.cpg			
1.2	Omission	Rate of missing items	OK				
2	LOGICAL CONSISTENCY						
2.1	Format consistency	File format/readable	OK	GeoTiff Readable in ArcGIS 10.x and QGIS.			
2.2		File naming conventions	NOK	Not fully compliant to the naming convention described in WaW technical specification. Version should be "4-digit qualifier of the version number, starting with "V1_1".			

PRODUCT:			HRL Water and Wetness (WaW)			
VALIDATION LEVEL:			Pan-European			
SERVICE PROVIDER: Multiple			SERVICE USER: EEA		ISSUE/REVISION:1.0	
VALIDATION DATE: 21/12/2018			REVIEW DATE:			
CONDUCTED BY:			REVIEWED BY:		APPROVED BY:	
No.	DATA QUALITY SUB-	DATA QUALITY MEASURE	DATA QUALITY RESULT	COMMENTS BY AUDIT TEAM	DRAFT AUDIT CONCLUSION	FINAL AUDIT CONCLUSION
2.3	Conceptual consistency	Attributes naming conventions	N/A			
2.4		Attributes types	OK			
2.5		Coordinate reference system	OK	European ETRS89 LAEA projection (EPSG 3035)		
2.6		Pixel size and origin	OK			
2.7		Compliance between 20 m and 100 m product	OK			
2.9		Additional attributes, symbology	OK			
2.10	Domaine consistency	Valid Codes	OK			
3	POSITIONAL ACCURACY					

PRODUCT:			HRL Water and Wetness (WaW)			
VALIDATION LEVEL:			Pan-European			
SERVICE PROVIDER: Multiple			SERVICE USER: EEA		ISSUE/REVISION:1.0	
VALIDATION DATE: 21/12/2018			REVIEW DATE:			
CONDUCTED BY:			REVIEWED BY:		APPROVED BY:	
No.	Data Quality Sub-	Data Quality Measure	Data Quality Result	Comments by Audit Team	Draft Audit Conclusion	Final Audit Conclusion
3.1	Absolute or external accuracy	RMSEP	n/a	This is dependent on the assessment of the 2015 CORE dataset		
3.2	Relative or internal accuracy	RMSEP	n/a	This is dependent on the assessment of the 2015 CORE dataset		
4	THEMATIC ACCURACY					
4.1	Classification correctness	Overall accuracy		See Section 4		
4.2		Min. producer's accuracy		See Section 4		
4.3		Min. user's accuracy		See Section 4		
4.4		Discrepancies along borderlines	OK			
5	TEMPORAL QUALITY					

PRODUCT:			HRL Water and Wetness (WaW)			
VALIDATION LEVEL:			Pan-European			
SERVICE PROVIDER: Multiple			SERVICE USER: EEA		ISSUE/REVISION: 1.0	
VALIDATION DATE: 21/12/2018			REVIEW DATE:			
CONDUCTED BY:			REVIEWED BY:		APPROVED BY:	
No.	Data Quality Sub-	Data Quality Measure	Data Quality Result	Comments by Audit Team		Draft Audit Conclusion
5.1	Temporal quality	Closeness of the acquired image data to the reference year	OK	The imagery used spans several years <i>from 2009 to 2015 (+/- one year)</i>		
6	USABILITY					
6.1	Usability	Usability description	OK			
7	METADATA					
7.1	INSPIRE compliant metadata	Presence	OK			
7.2		File format	OK			
7.3		File name	NOK	Not fully compliant to the naming convention described in WaW technical specification. Version should be "4-digit qualifier of the version number, starting with "V1_1".		
7.4		INSPIRE compliance	OK	Conformance class: XML encoding of ISO 19115/19119 metadata		

PRODUCT:			HRL Water and Wetness (WaW)			
VALIDATION LEVEL:			Pan-European			
SERVICE PROVIDER: Multiple			SERVICE USER: EEA	ISSUE/REVISION: 1.0		
VALIDATION DATE: 21/12/2018			REVIEW DATE:			
CONDUCTED BY:			REVIEWED BY:	APPROVED BY:		
No.	DATA QUALITY SUB-	DATA QUALITY MEASURE	DATA QUALITY RESULT	COMMENTS BY AUDIT TEAM	DRAFT AUDIT CONCLUSION	FINAL AUDIT CONCLUSION
			NOK	Conformance class: INSPIRE Profile based on EN ISO 19115 and EN ISO 19119 (Failed 2/8)		
			NOK	Conformance class: Metadata for interoperability (Failed 1/1)		

4. Thematic accuracy

For the thematic accuracy assessment, the WaW PSU value were derived from the SSU according to the aggregation rules product specifications.

The classes taken in account for the thematic accuracy index calculation are (section 1.2):

- 0: No water / no wet area
- 1: Permanent water
- 2: Temporary water
- 3: Permanently wet areas (wetness)
- 4: Temporary wet areas (wetness)

The overall, user's and producer's accuracies and the 95% confidence intervals, were calculated for each of the biogeographical regions and each countries or group of countries.

According to the product specifications, a thematic target accuracy of 85% is to be expected for the Permanent water and the Dry areas, whilst a thematic target accuracy of 80% is to be expected for the Temporary water, Permanent wet areas and Temporary wet areas.

The results in terms of blind and plausibility analysis at European level are shown in the Table 5 (biogeographical regions) and Table 6 (countries/group of countries). Accuracy figures were calculated following the procedure described in section 2.4.4, taking the weights of Table 4 into account.

The following considerations are made upon the results:

- Typical **omission errors within Permanent/Temporary water body** are likely related to:
 - Ponds and fragments of them partially covered by vegetation
 - Small lakes/reservoir or fragment of them partially covered by vegetation
- Typical **commission within Permanent/Temporary water body** are likely related to:
 - Liquid dump sites
 - Temporal water logged areas
 - Burnt areas
 - Coniferous forest stands in topographically influenced areas
- Typical **omission errors within the Permanent/Temporary wet** are likely related to:
 - Intertidal flats (e.g. Northern Germany)
 - Poljes in Balkan countries (temporary lakes in spring season)
 - River delta areas covered with forests (e.g. River Danube delta)
 - Peatlands in Scandinavia when temporary without surface water
 - Zones of fluctuation in lake and reservoirs
 - Zones of fluctuation in river beds (e.g. torrents in Southern Alps)
 - Lakes/Reservoirs with temporary vegetation cover
 - Coastal areas, salines and lagoons with non-permanent water cover
- Typical **commission errors within the Permanent/Temporary wet** are likely related to:
 - Steep shadowed slopes in mountain areas (e.g. Canary Islands, Turkey, Norway)
 - Dry rocky areas in Scandinavia
 - Coniferous forest stands in Scandinavia

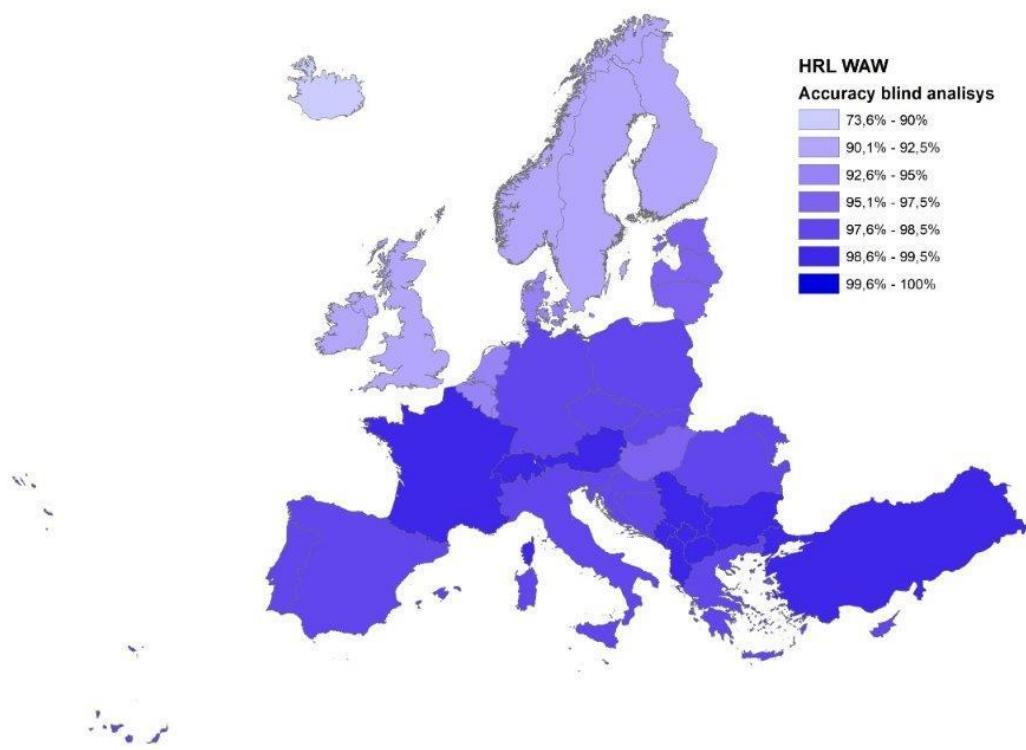


Figure 9. Overall blind accuracy per country and group of countries

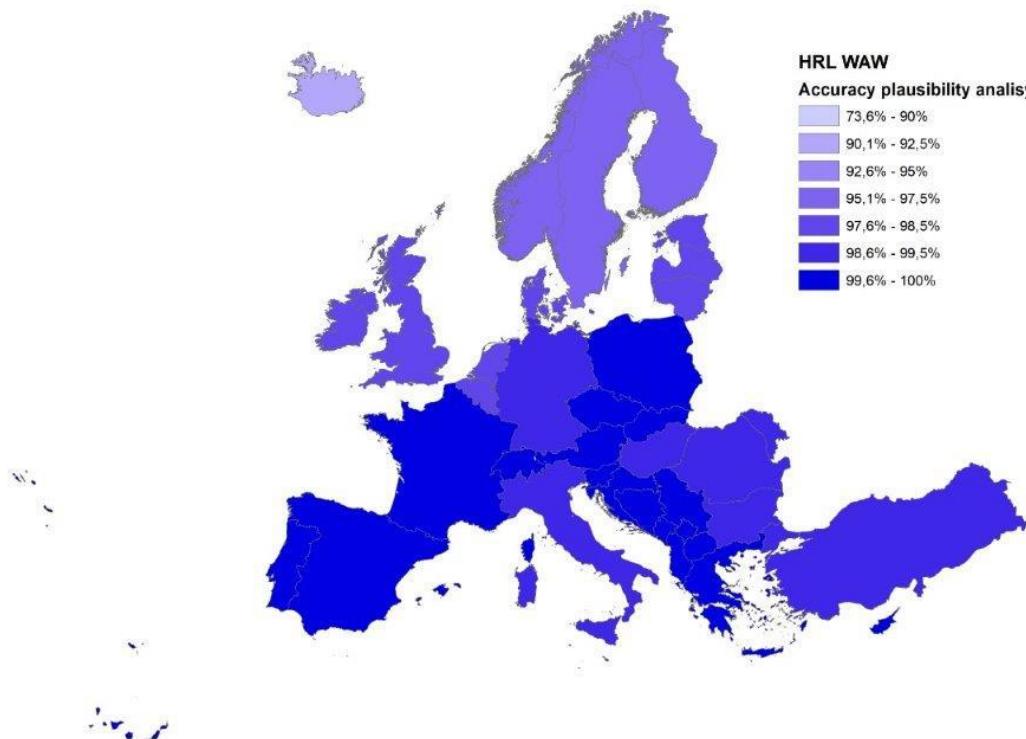


Figure 10. Overall plausibility accuracy per country and group of countries

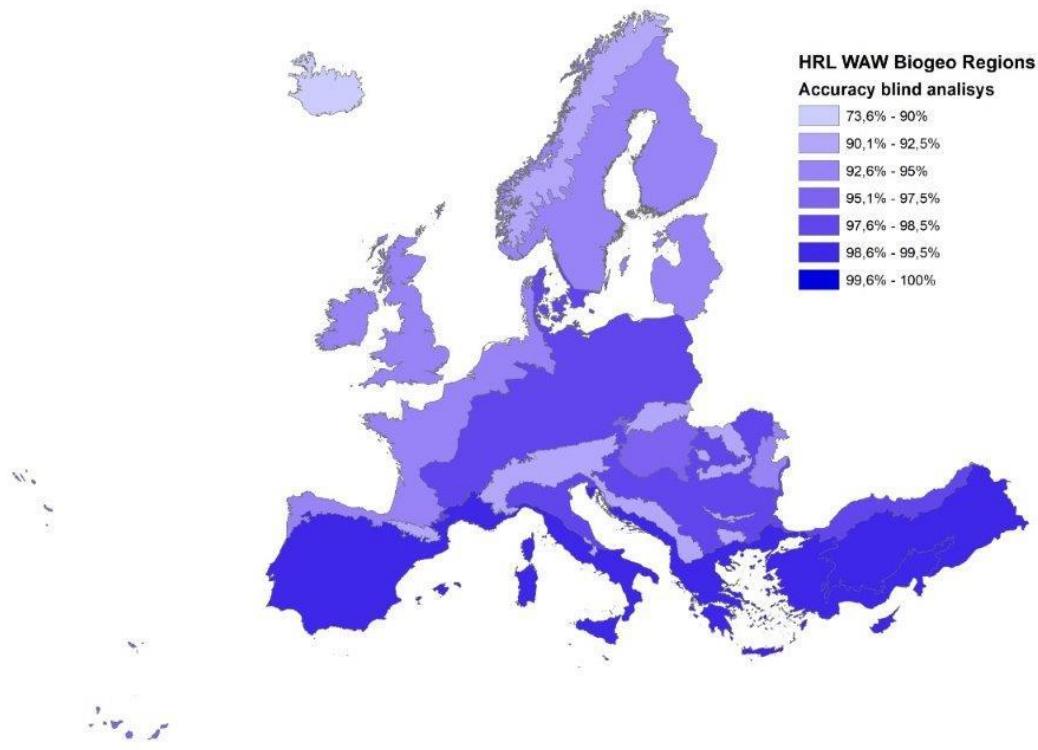


Figure 11. Overall blind accuracy per biogeographic region

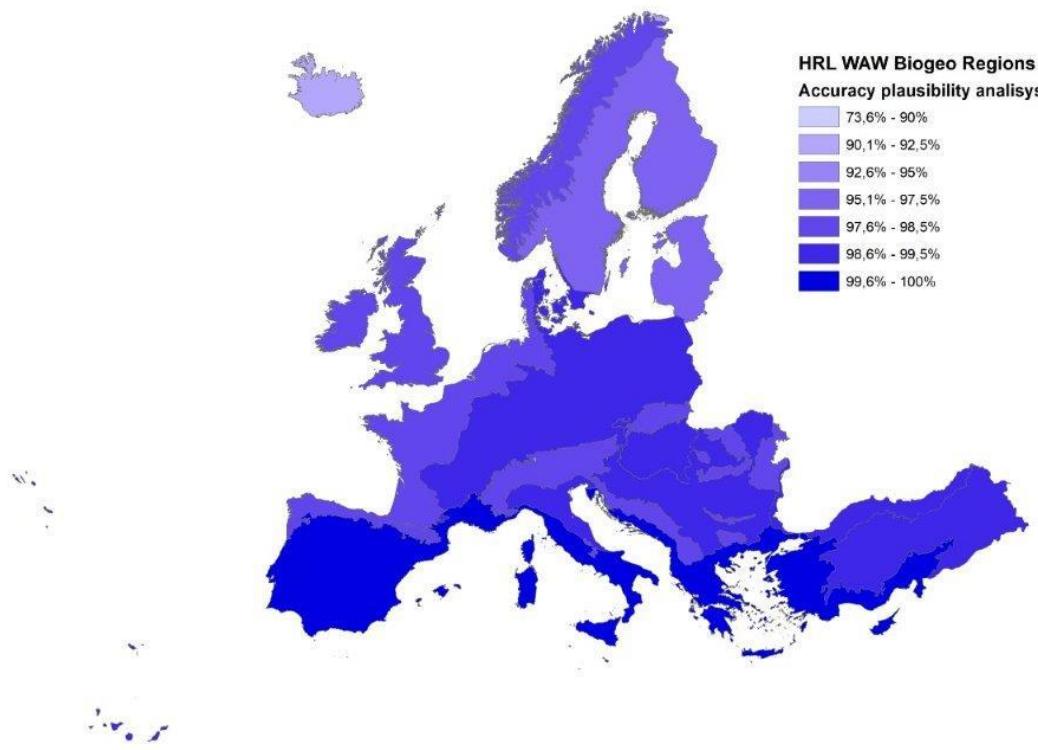


Figure 12. Overall plausibility accuracy per biogeographic region

Figure 11-12 above show Overall accuracy values. It is observable a gradient of accuracy values, with higher values in the temperate regions that decrease moving toward the colder areas.

The accuracy for the individual thematic classes has also been analysed. With the following results:

- **Dry Areas Accuracy:** the general accuracy is significantly above the target of the 85%. The critical values are correlated with the biogeographical gradient seen above. The producer accuracy is generally slightly below than the user accuracy. The accuracy is improved by the plausibility analysis.
- **Permanent water Accuracy:** the user general accuracy is significantly above the target of the 85%. The critical values are correlated with the biogeographical gradient seen above. The producer accuracy is significantly below than the user accuracy and below the target of 85%. The user's accuracy is improved by the plausibility analysis. In the same way the plausibility analysis raises the producer's accuracies into an acceptable range.
- **Temporary water and wet, Permanent Wet Accuracy:** the producer and user accuracies are significantly below the target. This does not change throughout the bio-geographical regions and the country analysis. In general, the temporary water thematic accuracies are higher in the producer rather than user and the permanent wetness accuracies are higher in the user rather than producer. The accuracy is improved by the plausibility analysis, raising the user accuracies into an acceptable range for the temporary water and permanent wet area, but remains far below the expected producer's accuracies. Regarding the temporary wet area, the plausibility analysis produced a good accuracies for both user and producer.

Table 5: WaW product blind thematic producer's accuracy results per total European area covered, per bio-geographical regions and Countries and group of countries greater than 90,000km², for the WaW classes: No water / no wet area (NWaW), Permanent water (PWa), Temporary water (TWa), Permanent wet area (PWe), Temporary wet area (TWe). As for the Permanent water and Dry areas, the values in green indicate accuracies greater than the 85% threshold, yellow greater than 75% and red less than 75% considering the 95% CI. As for the Temporary water, Permanent wet areas and Temporary wet areas, the values in green indicate accuracies greater than the 80% threshold, yellow greater than 70% and red less than 70% considering the 95% CI.

	Producer Accuracies									
	Prod NWaW	CI 95%	Prod PWa	CI 95%	Prod TWa	CI 95%	Prod PWe	CI 95%	Prod TWe	CI 95%
EEA39	97,64%	0,02%	81,82%	0,48%	25,42%	0,85%	21,71%	0,73%	56,94%	0,74%
Alpine	93,44%	0,03%	76,21%	0,53%	38,02%	0,04%	1,80%	0,88%	70,73%	0,72%
Anatolian	99,84%	0,06%	86,98%	0,17%	38,70%	1,10%	50,25%	0,59%	24,14%	0,52%
Arctic	83,22%	0,28%	45,50%	0,02%	51,92%	1,18%	52,35%	0,48%	50,86%	0,28%
Atlantic	96,49%	0,10%	72,06%	0,12%	3,90%	0,05%	9,44%	0,77%	47,74%	0,06%
Black Sea	99,68%	0,13%	82,04%	0,00%	36,06%	1,18%	12,45%	1,25%	39,85%	0,29%
Boreal	97,45%	0,34%	91,04%	0,39%	14,23%	1,27%	14,20%	1,38%	48,69%	0,83%
Continental	98,89%	0,02%	70,66%	0,93%	18,84%	0,39%	16,31%	1,39%	57,75%	1,48%
Macaronesia	96,07%	0,59%	N/A	N/A	N/A	N/A	0,00%	0,00%	100,00%	0,00%
Mediterranean	99,17%	0,19%	75,71%	0,47%	39,35%	1,11%	43,42%	0,95%	48,01%	0,35%
Pannonician	97,13%	0,32%	87,51%	0,17%	24,07%	1,25%	27,37%	0,95%	43,73%	0,25%
Steppic	96,73%	0,06%	90,03%	0,36%	100,00%	0,00%	23,02%	1,71%	62,62%	0,00%
AL+ME+MK+RS+XK	99,01%	0,00%	91,17%	0,00%	40,75%	0,03%	23,42%	0,02%	98,04%	0,00%
AT + CH + LI	99,30%	0,00%	89,30%	0,00%	32,79%	0,01%	63,53%	0,01%	98,11%	0,00%
BA + HR + SI	98,70%	0,00%	90,53%	0,00%	100,00%	0,00%	4,40%	0,01%	74,45%	0,01%
BE + LU+ NL + DK	98,36%	0,01%	83,19%	0,02%	1,13%	0,40%	5,36%	0,85%	39,67%	0,76%
BG	99,70%	0,00%	64,67%	0,82%	74,33%	0,01%	14,05%	0,00%	69,54%	0,00%
CZ + SK	98,51%	0,00%	64,29%	0,81%	89,16%	0,00%	5,51%	0,01%	0,00%	0,00%
DE	98,52%	0,01%	76,49%	0,52%	33,25%	0,01%	21,27%	0,04%	42,93%	0,93%
EE + LT + LV	97,86%	0,02%	83,62%	0,52%	100,00%	0,00%	15,60%	0,70%	42,03%	0,80%
EL	99,06%	0,00%	85,65%	0,01%	70,51%	0,01%	28,76%	0,02%	56,99%	0,01%

	Producer Accuracies									
	Prod NWaW	CI 95%	Prod PWa	CI 95%	Prod TWa	CI 95%	Prod PWe	CI 95%	Prod TWe	CI 95%
ES	99,38%	0,00%	60,43%	0,01%	42,89%	0,01%	27,41%	0,92%	34,08%	0,00%
FI	98,21%	0,03%	96,05%	0,03%	29,73%	0,21%	17,66%	0,59%	53,91%	0,60%
FR	99,31%	0,00%	75,57%	0,01%	9,69%	0,93%	22,19%	0,93%	41,43%	0,93%
HU	97,48%	0,02%	90,88%	0,01%	47,48%	0,05%	43,36%	0,05%	27,72%	0,91%
IE + UK	94,83%	0,06%	65,42%	0,70%	5,44%	0,79%	9,55%	0,63%	38,38%	0,69%
IS	82,87%	0,24%	45,70%	0,67%	52,57%	0,30%	52,35%	0,54%	49,62%	0,64%
IT	99,39%	0,00%	65,20%	0,66%	11,13%	0,85%	21,22%	0,01%	0,00%	0,51%
NO	83,82%	0,30%	73,00%	0,52%	43,85%	0,08%	1,64%	0,80%	82,31%	0,40%
PL	98,65%	0,01%	63,76%	0,76%	100,00%	0,00%	52,37%	0,03%	67,00%	0,81%
PT	98,30%	0,01%	73,11%	0,01%	45,75%	0,01%	70,11%	0,01%	100,00%	0,00%
RO	99,05%	0,01%	80,57%	0,02%	13,92%	0,81%	14,70%	0,04%	45,88%	0,03%
SE	94,85%	0,07%	86,44%	0,34%	6,57%	0,23%	5,74%	0,65%	48,91%	0,71%
TR	99,72%	0,00%	84,83%	0,39%	35,06%	0,03%	53,11%	0,73%	32,19%	0,89%

Table 6: WaW product blind thematic user's accuracy results per total European area covered, per bio-geographical regions and Countries and group of countries greater than 90,000km², for the WaW classes: No water / no wet area (NWaW), Permanent water (PWa), Temporary water (TWa), Permanent wet area (PWe), Temporary wet area (TWe). As for the Permanent water and Dry areas, the values in green indicate accuracies greater than the 85% threshold, yellow greater than 75% and red less than 75% considering the 95% CI. As for the Temporary water, Permanent wet areas and Temporary wet areas, the values in green indicate accuracies greater than the 80% threshold, yellow greater than 70% and red less than 70% considering the 95% CI.

	User Accuracies									
	User NWaW	CI 95%	User PWa	CI 95%	User TWa	CI 95%	User PWe	CI 95%	User TWe	CI 95%
EEA39	98,53%	0,16%	97,01%	0,02%	30,44%	0,01%	52,57%	0,01%	26,16%	0,06%
Alpine	96,99%	0,25%	98,95%	0,01%	9,62%	0,00%	26,83%	0,01%	34,75%	0,07%
Anatolian	99,42%	0,05%	91,84%	0,56%	51,72%	0,00%	93,75%	0,10%	21,43%	0,52%
Arctic	86,51%	0,18%	84,62%	1,03%	67,83%	0,00%	50,42%	0,09%	36,69%	0,38%
Atlantic	97,90%	0,00%	93,67%	1,02%	13,39%	0,13%	36,48%	0,18%	18,78%	0,00%
Black Sea	99,30%	0,05%	100,00%	0,00%	20,69%	0,00%	70,94%	0,00%	8,97%	0,70%
Boreal	97,22%	0,24%	98,51%	0,00%	10,07%	0,05%	54,93%	0,15%	28,99%	1,03%
Continental	99,37%	0,16%	97,70%	0,00%	25,89%	0,10%	32,61%	0,44%	14,49%	0,00%
Macaronesia	99,28%	0,12%	N/A	N/A	N/A	N/A	N/A	N/A	21,02%	1,24%
Mediterranean	99,76%	0,03%	95,97%	0,25%	39,29%	0,16%	56,02%	0,03%	8,38%	0,59%
Pannonian	99,16%	0,04%	91,38%	0,54%	35,53%	0,04%	62,56%	0,03%	13,99%	0,68%
Steppic	99,35%	0,31%	88,89%	0,00%	28,57%	0,00%	64,71%	0,69%	11,76%	0,00%
AL+ME+MK+RS+XK	99,88%	0,00%	93,55%	0,02%	56,67%	0,00%	50,00%	0,00%	20,84%	0,01%
AT + CH + LI	99,88%	0,00%	95,65%	0,02%	7,01%	0,00%	65,80%	0,00%	14,38%	0,00%
BA + HR + SI	99,87%	0,00%	100,00%	0,00%	44,61%	0,00%	22,08%	0,00%	10,43%	0,01%
BE + LU+ NL + DK	96,01%	0,30%	94,12%	0,04%	23,15%	0,00%	19,21%	0,00%	27,27%	0,05%
BG	99,52%	0,13%	100,00%	0,00%	28,57%	0,00%	59,20%	0,00%	2,35%	0,00%
CZ + SK	99,49%	0,12%	100,00%	0,00%	7,30%	0,00%	30,00%	0,00%	0,00%	0,00%
DE	99,58%	0,10%	96,36%	0,01%	12,12%	0,00%	28,32%	0,00%	7,94%	0,02%
EE + LT + LV	98,15%	0,17%	98,11%	0,02%	8,70%	0,00%	75,76%	0,01%	15,09%	0,04%
EL	99,72%	0,00%	89,86%	0,01%	56,52%	0,00%	73,14%	0,00%	6,30%	0,00%

	User Accuracies									
	User NWaW	CI 95%	User PWa	CI 95%	User TWa	CI 95%	User PWe	CI 95%	User TWe	CI 95%
ES	99,75%	0,06%	100,00%	0,00%	25,42%	0,01%	61,30%	0,00%	4,88%	0,01%
FI	97,33%	0,15%	97,81%	0,07%	23,08%	0,00%	61,74%	0,02%	41,60%	0,08%
FR	99,59%	0,10%	95,24%	0,01%	16,95%	0,00%	40,00%	0,00%	8,70%	0,01%
HU	99,04%	0,14%	96,68%	0,01%	38,11%	0,00%	73,47%	0,00%	10,34%	0,03%
IE + UK	96,58%	0,20%	100,00%	0,00%	16,67%	0,00%	36,54%	0,01%	19,31%	0,11%
IS	85,69%	0,42%	84,62%	0,04%	69,03%	0,03%	50,53%	0,23%	36,65%	0,33%
IT	99,28%	0,14%	100,00%	0,00%	29,03%	0,00%	43,01%	0,00%	0,00%	0,00%
NO	95,02%	0,25%	97,95%	0,03%	11,11%	0,01%	21,43%	0,00%	30,69%	0,37%
PL	99,35%	0,13%	97,22%	0,01%	16,00%	0,00%	39,29%	0,00%	24,14%	0,03%
PT	99,80%	0,00%	100,00%	0,00%	25,00%	0,01%	60,00%	0,01%	20,64%	0,02%
RO	99,15%	0,09%	92,00%	0,01%	37,93%	0,00%	65,52%	0,00%	12,82%	0,02%
SE	94,79%	0,25%	98,75%	0,04%	4,55%	0,00%	31,25%	0,01%	32,29%	0,16%
TR	99,65%	0,08%	91,80%	0,02%	45,54%	0,01%	86,41%	0,00%	13,89%	0,01%

Table 7: WaW product plausibility thematic producer's accuracy results per total European area covered, per bio-geographical regions and Countries and group of countries greater than 90,000km², for the WaW classes: No water / no wet area (NWaW), Permanent water (PWa), Temporary water (TWa), Permanent wet area (PWe), Temporary wet area (TWe). As for the Permanent water and Dry areas, the values in green indicate accuracies greater than the 85% threshold, yellow greater than 75% and red less than 75% considering the 95% CI. As for the Temporary water, Permanent wet areas and Temporary wet areas, the values in green indicate accuracies greater than the 80% threshold, yellow greater than 70% and red less than 70% considering the 95% CI.

	Producer Accuracies									
	Prod NWaW	CI 95%	Prod PWa	CI 95%	Prod TWa	CI 95%	Prod PWe	CI 95%	Prod TWe	CI 95%
EEA39	99,89%	0,00%	88,32%	0,40%	56,60%	0,64%	35,94%	0,67%	88,90%	0,35%
Alpine	99,95%	0,00%	82,22%	0,48%	66,46%	0,05%	8,94%	0,85%	94,66%	0,31%
Anatolian	100,00%	0,00%	98,75%	0,10%	69,60%	0,73%	44,62%	0,51%	49,31%	0,46%
Arctic	97,68%	0,12%	49,00%	0,02%	78,98%	0,30%	78,79%	0,35%	87,48%	0,22%
Atlantic	99,84%	0,04%	79,55%	0,08%	54,27%	0,08%	21,36%	0,02%	87,42%	0,04%
Black Sea	99,95%	0,00%	96,70%	0,00%	71,73%	1,11%	15,35%	1,25%	69,46%	0,28%
Boreal	99,77%	0,06%	94,06%	0,35%	54,69%	1,10%	18,99%	1,37%	79,80%	0,59%
Continental	99,94%	0,01%	81,35%	0,90%	37,72%	0,63%	19,37%	1,27%	91,70%	0,73%
Macaronesia	100,00%	0,00%	N/A	N/A	N/A	N/A	0,00%	0,00%	100,00%	0,00%
Mediterranean	99,96%	0,01%	92,90%	0,35%	57,98%	0,97%	52,31%	0,87%	94,40%	0,17%
Pannonian	99,93%	0,05%	93,47%	0,15%	61,11%	0,84%	59,88%	0,92%	84,68%	0,15%
Steppic	99,97%	0,00%	93,27%	0,36%	100,00%	0,00%	31,43%	1,74%	92,72%	0,00%
AL+ME+MK+RS+XK	99,93%	0,00%	97,44%	0,00%	73,99%	0,01%	50,76%	0,01%	99,59%	0,00%
AT + CH + LI	100,00%	0,00%	92,52%	0,00%	87,44%	0,00%	100,00%	0,00%	100,00%	0,00%
BA + HR + SI	99,98%	0,00%	94,41%	0,00%	100,00%	0,00%	7,31%	0,01%	96,04%	0,00%
BE + LU+ NL + DK	99,80%	0,00%	86,38%	0,02%	100,00%	0,00%	50,82%	0,06%	74,03%	0,02%
BG	99,99%	0,00%	68,29%	0,84%	89,66%	0,00%	30,53%	0,00%	100,00%	0,00%
CZ + SK	99,76%	0,00%	89,82%	0,00%	100,00%	0,00%	6,37%	0,01%	99,70%	0,00%
DE	99,84%	0,00%	79,98%	0,53%	91,84%	0,00%	20,96%	0,03%	86,68%	0,54%
EE + LT + LV	99,93%	0,00%	97,74%	0,01%	58,50%	0,06%	23,56%	0,07%	79,70%	0,53%

	Producer Accuracies									
	Prod NWaW	CI 95%	Prod PWa	CI 95%	Prod TWa	CI 95%	Prod PWe	CI 95%	Prod TWe	CI 95%
EL	99,99%	0,00%	89,70%	0,01%	80,19%	0,01%	47,06%	0,01%	94,38%	0,00%
ES	99,96%	0,00%	90,37%	0,01%	57,80%	0,01%	27,59%	0,92%	88,32%	0,00%
FI	99,87%	0,01%	96,33%	0,04%	67,57%	0,19%	22,23%	0,61%	76,84%	0,43%
FR	99,95%	0,00%	90,51%	0,00%	37,83%	0,92%	36,15%	0,93%	98,39%	0,00%
HU	100,00%	0,00%	94,27%	0,01%	70,34%	0,05%	66,89%	0,05%	77,71%	0,66%
IE + UK	99,83%	0,01%	77,77%	0,58%	29,27%	0,23%	19,74%	0,23%	83,39%	0,15%
IS	97,53%	0,07%	49,09%	0,67%	78,80%	0,27%	78,74%	0,38%	86,91%	0,33%
IT	99,99%	0,00%	87,09%	0,00%	17,50%	0,88%	28,38%	0,01%	75,49%	0,79%
NO	99,75%	0,02%	81,41%	0,43%	80,28%	0,26%	8,89%	0,71%	95,73%	0,21%
PL	99,99%	0,00%	76,55%	0,62%	100,00%	0,00%	25,34%	0,91%	94,59%	0,01%
PT	99,98%	0,00%	84,91%	0,01%	77,13%	0,01%	75,77%	0,01%	100,00%	0,00%
RO	100,00%	0,00%	83,86%	0,02%	29,16%	0,84%	21,79%	0,04%	81,96%	0,02%
SE	99,65%	0,01%	88,71%	0,35%	41,07%	0,16%	10,55%	0,69%	87,62%	0,33%
TR	99,98%	0,00%	98,83%	0,00%	61,03%	0,02%	48,44%	0,87%	71,41%	0,68%

Table 8: WaW product plausibility thematic user's accuracy results per total European area covered, per bio-geographical regions and Countries and group of countries greater than 90,000km², for the WaW classes: No water / no wet area (NWaW), Permanent water (PWa), Temporary water (TWa), Permanent wet area (PWe), Temporary wet area (TWe). As for the Permanent water and Dry areas, the values in green indicate accuracies greater than the 85% threshold, yellow greater than 75% and red less than 75% considering the 95% CI. As for the Temporary water, Permanent wet areas and Temporary wet areas, the values in green indicate accuracies greater than the 80% threshold, yellow greater than 70% and red less than 70% considering the 95% CI.

	User Accuracies									
	User NWaW	CI 95%	User PWa	CI 95%	User TWa	CI 95%	User PWe	CI 95%	User TWe	CI 95%
EEA39	99,03%	0,12%	99,27%	0,01%	80,53%	0,00%	71,47%	0,01%	92,49%	0,04%
Alpine	98,45%	0,16%	100,00%	0,00%	63,61%	0,00%	53,43%	0,01%	98,68%	0,02%
Anatolian	99,43%	0,05%	98,98%	0,21%	96,55%	0,00%	97,50%	0,06%	64,29%	0,61%
Arctic	92,23%	0,15%	96,15%	0,55%	81,74%	0,00%	79,03%	0,07%	94,67%	0,18%
Atlantic	98,87%	0,00%	99,21%	0,37%	68,83%	0,23%	49,56%	0,19%	92,89%	0,00%
Black Sea	99,37%	0,05%	100,00%	0,00%	93,10%	0,00%	72,66%	0,00%	46,19%	1,23%
Boreal	97,93%	0,21%	99,46%	0,00%	51,08%	0,19%	63,85%	0,14%	85,51%	0,80%
Continental	99,47%	0,15%	99,42%	0,00%	88,40%	0,12%	47,70%	0,47%	89,14%	0,00%
Macaronesia	99,28%	0,12%	N/A	N/A	N/A	N/A	N/A	N/A	100,00%	0,00%
Mediterranean	99,81%	0,03%	98,96%	0,13%	96,43%	0,08%	71,75%	0,03%	85,86%	0,75%
Pannonic	99,16%	0,04%	97,13%	0,32%	98,16%	0,02%	89,54%	0,02%	96,26%	0,37%
Steppic	99,35%	0,31%	88,89%	0,00%	71,43%	0,00%	82,35%	0,55%	82,35%	0,00%
AL+ME+MK+RS+XK	99,88%	0,00%	96,77%	0,01%	96,67%	0,00%	57,26%	0,00%	100,00%	0,00%
AT + CH + LI	99,88%	0,00%	100,00%	0,00%	100,00%	0,00%	76,22%	0,00%	96,88%	0,00%
BA + HR + SI	99,87%	0,00%	100,00%	0,00%	100,00%	0,00%	37,85%	0,00%	90,05%	0,01%
BE + LU+ NL + DK	99,01%	0,00%	100,00%	0,00%	86,11%	0,00%	46,29%	0,01%	84,85%	0,04%
BG	99,52%	0,13%	100,00%	0,00%	95,24%	0,00%	85,16%	0,00%	66,46%	0,00%
CZ + SK	99,90%	0,00%	100,00%	0,00%	75,90%	0,00%	35,00%	0,00%	79,84%	0,01%
DE	99,60%	0,10%	98,18%	0,01%	63,64%	0,00%	37,46%	0,00%	80,95%	0,03%
EE + LT + LV	98,95%	0,10%	100,00%	0,00%	65,22%	0,00%	78,79%	0,01%	83,02%	0,04%
EL	99,72%	0,00%	100,00%	0,00%	95,65%	0,00%	82,11%	0,00%	89,76%	0,01%

	User Accuracies										
	User NWaW	CI 95%	User PWa	CI 95%	User TWa	CI 95%	User PWe	CI 95%	User TWe	CI 95%	
ES	99,75%	0,06%	100,00%	0,00%	94,07%	0,00%	61,30%	0,00%	78,05%	0,01%	
FI	97,72%	0,14%	99,38%	0,04%	71,79%	0,00%	69,57%	0,02%	83,20%	0,06%	
FR	99,70%	0,08%	100,00%	0,00%	91,53%	0,00%	62,50%	0,00%	93,48%	0,01%	
HU	99,04%	0,14%	100,00%	0,00%	100,00%	0,00%	92,60%	0,00%	93,10%	0,03%	
IE + UK	98,13%	0,09%	100,00%	0,00%	56,67%	0,00%	48,08%	0,01%	94,06%	0,06%	
IS	91,75%	0,32%	96,15%	0,02%	82,30%	0,02%	78,98%	0,18%	94,41%	0,16%	
IT	99,28%	0,14%	100,00%	0,00%	96,77%	0,00%	77,57%	0,00%	79,71%	0,01%	
NO	97,17%	0,18%	100,00%	0,00%	60,32%	0,01%	35,71%	0,00%	98,65%	0,09%	
PL	99,55%	0,11%	97,22%	0,01%	76,00%	0,00%	50,00%	0,00%	96,55%	0,01%	
PT	99,80%	0,00%	100,00%	0,00%	100,00%	0,00%	80,00%	0,01%	95,32%	0,01%	
RO	99,19%	0,09%	96,00%	0,01%	82,76%	0,00%	86,21%	0,00%	74,36%	0,03%	
SE	96,68%	0,18%	99,50%	0,03%	37,61%	0,01%	48,75%	0,01%	93,42%	0,08%	
TR	99,66%	0,08%	98,36%	0,01%	96,53%	0,00%	92,23%	0,00%	69,44%	0,01%	

5. Conclusions and recommendations

This report represents a thorough examination of the WaW product layer of the HRL product line. It provides an independent validation assessment of the performance of the products against their design specifications with blind interpretations and plausibility analyses, clarifying the thematic quality of the products then made available to the wider community.

The Validation Check list (Section 3) highlight minor shortcomings with data packaging, naming convention, symbology and metadata contents, which should be considered for future updates:

- Introduction of a correct file naming convention which is also applicable for the attributed structure of the product
- Provision of correct assigned colormap (symbology)
- A parent scene identification layer (PSIL) which includes details on the image data used in the production (e.g. time stamps, sensor, ancillary information etc.) would be helpful in the validation and further investigations and should be considered to be part of the product
- INSPIRE compliance should be ensured in future: 100% INSPIRE conformity of metadata

Overall the plausibility analysis, raises the accuracy values in all classes since key input parameters and product specifications that cannot be ignored in such assessment are taken into account.

The validation exercise shows different results depending on the classes analysed. In particular for Dry and Permanent Water, the thematic accuracy is generally in line with the target for commission errors and slightly below for the omission errors. The thematic accuracy is influenced by the biogeographical regions, with higher values in the temperate regions that decrease moving toward the colder areas. This probably happens for the difficulty to get images clear from cloud and ice/snow. For future updates the suggestion could be to select a subset of images without the presence of snow or ice.

Even if the identification of Dry and Permanent Water classes is generally good, the discrimination between the permanent and temporary attributes results highly critical. This could be explained by the fact that only the water occurrence in the images is considered by the automatic procedure, without considering situations where the water is covered by the vegetation or to the fact that the image packages do not cover continuous time frame.

Mismatch in the persistence of water was found on the reservoirs under construction. In particular the automatic classification identified presence of water only in few images and classified them as intermittent whilst the blind interpretation classifies them as perennial. On the other hand the plausibility analysis agrees with the automatic classification as it takes into account the product specifications and the extraction methodology, increasing the accuracy accordingly, both user and producer.

Commission errors in the temporary wetness class are often related to the significant presence of humidity detected by the WAW classification in cultivated areas, where the blind interpretation tends to classifies them as dry areas. This is mostly due to the difficulty to identify soil moisture by blind visual interpretation of optical data. In this specific case, the accuracy is increased in the plausibility analysis, since other relevant parameters such as the product classification and the related extraction criteria and methodology are taken into account.

During the validation assessment some minor issues were detected on the water and wetness classes present in the high slope areas, which can be reduced by the use of a more accurate DTM.