GUIDELINES FOR VERIFICATION OF HIGH-RESOLUTION LAYERS PRODUCED UNDER GMES INITIAL OPERATIONS (GIO) LAND MONITORING 2011 – 2013

Supplement to the documents:
GMES Initial Operations (GIO) Land Monitoring 2011 – 2013 in the framework of regulation (EU) No 911/2010 Pan-EU Component Terms of Reference and template for national project plan

Framework service contract(s) for the provision of GMES Initial Operations 2011-2013 Land Monitoring Services: High Resolution land cover characteristics of 5 main land cover types (6 lots). Reference: Open call for tenders EEA/SES/11/004

Report information

Report title:	Guidelines for verification and enhancement of high-resolution layers produced under GMES initial operations (GIO) Land monitoring 2011 – 2013
Report number:	N/A
Report website:	N/A
Report version:	2
Date of preparation:	30.04.2012
Beneficiary organisation:	National Land Cover Teams, Service Providers
Prepared by:	György Büttner
Email address:	gyorgy.buttner@eea.europa.eu
Project coordinator:	Hans Dufourmont
Project coordinator organisation:	European Environment Agency
Project coordinator telephone number:	+45 2494 1115
Project coordinator email address:	hans.dufourmont@eea.europa.eu

DISCLAIMER

1 Contents

Revision history

Revision nr.	Date	Modified by	Comments
2	30.04.2012	G. Büttner	Comments by FI and NL are integrated. Definitions of imperviousness and forest are improved based on discussions with Service providers

2 Executive summary

The following five high-resolution layers will be produced as part of GIO Land (2011-2013) covering EEA39: imperviousness, forests (tree cover density and forest type), grasslands, wetlands and permanent water bodies. Service providers are in charge of production of the HRLs. The, so called, intermediate products will be verified by national teams (or if not interested, by Service Providers).

The task of verification is aiming at identifying systematic classification errors that are eligible for later improvement/enhancement. The 5 HRLs shall be checked for omission and commission errors. Verification shall be carried out on the intermediate products at full resolution (20m x 20m, in national projection), by visual comparison of the selected samples with existing reference data (e.g. topographic maps, aerial photography or others).

A verification procedure consisting of three parts is proposed: (1) General overview of data quality, (2) Checking “error prone” locations in each HRL by means of look-and-feel control (similarly to the approach used for the verification of the Soil Sealing product in the GMES FTS Land Monitoring Precursor[1]). (3) Applying an additional, statistically based quantitative[2] verification by using randomly selected samples to estimate commission and omission errors. The first two are mandatory activities. The third element is highly recommended.

The purpose of this document is to provide guidelines for the verification procedures to be applied for HRLs produced under GIO.

[1] Guidelines for verification of high resolution soil sealing layer - Qualitative assessment - Prepared by: C. Steenmans and A. Sousa, EEA, 2007

[2] Some member states have proposed quantitative verification methodology

3 Background and purpose of the document

The **Land monitoring service** of the Global Monitoring of Environment and Security (GMES) led by DG ENTR3 of the European Commission, has entered its Initial Operation (GIO) phase following the entry into application of Regulation (EU) n°911/2010 of 22 September 2010 of the European Parliament and the Council on the European Earth monitoring programme (GMES) and its initial operations (2011 to 2013).

Pursuant to the GMES Work Programme 2011 these activities will start with the production of the 5 High Resolution pan-European layers. To that aim, a call for tender has been organised, comprising 6 lots, following a distribution according to a combination of geographic and thematic criteria. On 5th October 2011 the EEA addressed the National Focal Points and National Reference Centres Land Cover in 39 European countries requesting for a letter of intent for participation in GIO Land. The request was accompanied with a Terms of Reference and template for national project plan.

Some of the task in the Technical Specification, namely "Verification of high-resolution layers (HRL)" and "Enhancement of HRLs" were not fully described in technical terms and "comprehensive guidelines" were promised to the countries as well as to Service Providers in a later stage. The main purpose of this document is to provide the guidelines for verification.

4 Products to be verified

The following definitions apply for the production of high resolution layers by Service Providers. Only the 20m x 20m products (the intermediate layers which are produced directly) are described, as these are the target of verification and enhancement.

4.1 HRL Imperviousness

Built-up areas are characterized by the substitution of the original (semi-) natural cover or water surface with an artificial, often impervious cover. These artificial surfaces are usually characterized by long cover duration[4]. A high resolution imperviousness dataset representing all artificially sealed areas will be produced using automatic derivation based on calibrated NDVI[5].

A per-pixel estimate of sealed soil is foreseen as the index for the degree of imperviousness (0-100%). Data for 2006 and 2009 were produced in the frame of GMES precursor activities and Geoland-2, respectively, and are provided by EEA. Imperviousness change layers shall be produced as a difference between the three dates and shall be presented as degree of imperviousness change (-100% -- +100%[6]).

The processing scheme is based on the experiences of the 2006 and 2009 exercises and the results of the product validation. Having those in mind, special attention shall be paid to:

(a) Objects corresponding to CLC classes[7] 1.3.x (mines, quarries, dump sites and construction sites), are to be classified as non-impervious surfaces.
(b) Not allowing false gaps in settlements (due to e.g. special roof types).
(c) Greenhouses should be classified as impervious surfaces.

[3] Directorate General Enterprise and Industry

[4] FAO Land Cover Classification System, 2005

[5] Normalised Difference Vegetation Index

[6] Negative change means a decrease of imperviousness, which sporadically might occur

[7] CLC classes listed below indicate only types of landscape objects / elements / structures and not necessarily suggest the use of CLC

(d) Beaches, dunes, sand (associated to CLC class 3.3.1), bare rocks (CLC 3.3.2) and sparsely vegetated areas (CLC 3.3.3), which should be classified as non-impervious surfaces and are common sources for misclassification.

(e) Airports and harbours are another common sources of misclassification and should be corrected regarding both commission and omission errors.

1. Definitions	
Soil Sealing **Imperviousness** (used as synonyms)	The covering of the soil surface with impervious materials as a result of urban development and infrastructure construction. http://eusoils.jrc.ec.europa.eu/library/themes/Sealing/ Destruction or covering of soils by buildings, constructions and layers of completely or partly impermeable artificial material (asphalt, concrete, etc.). It is the most intense form of land take. Sealed surfaces are part of built-up areas. An indicator of the intensity of land take is the proportion of the total built-up surface area which is sealed. http://ec.europa.eu/environment/soil/pdf/sealing/Soil%20sealing%20-%20Final%20Report.pdf. Soil sealing is also used to describe the covering or sealing of the soil surface by impervious materials by, for example, concrete, metal, glass, tarmac and plastic. http://glossary.eea.europa.eu
Built-up areas	Land consumed by settlements, infrastructure, and commercial and industrial areas.
Land Take	Land take is the increase of artificial surfaces (housing areas; green urban areas; industrial, commercial and transport units; road and rail networks; etc) over time. http://ec.europa.eu/environment/soil/pdf/sealing/Soil%20sealing%20-%20Final%20Report.pdf
Built-up area changes	Areas were the extent of the built-up areas has increased or decreased from one (observation) point/period to the other.
Sealing Changes	Change of the degree of soil sealing from one observation point/period to the other. Those changes can occur in both, a) previously built-up areas (e.g. urban densification), and in b) newly built-up areas

Elements included in the built-up area in **Imperviousness Mapping 2006, 2009 and 2012**	**Elements excluded from the built-up area in Imperviousness Mapping 2006, 2009 and 2012**
• Housing areas • Traffic areas (airports, harbours, railway yards, parking lots)	• Mines, quarries, peat production • Dump sites • Construction sites without discernible

• Industrial, commercial areas, factories • Amusement parks (excluding the pure green are-as associated with them) • Construction sites with discernible evolving built-up structures • Single (farm) houses (where possible to identify) • Other sealed surfaces, which are part of fuzzy categories, such as e.g. allotment gardens, ceme-teries, sport areas (visible infrastructure), camp sites (roads and infrastructure, possibly influenced by caravans), excluding green areas associated with them. • Roads and railways associated to other impervi-ous surfaces (no gaps manually filled, no roads manually digitized) • Water edges with paved borders	evolving built-up structures • Meadows used for sports of any kind • Bare soil, rock, sparsely vegetated areas • Sand, sand pits • Glaciers, snow, water • Railway lines

The built-up / non-built-up map derived from the 20m x 20m imperviousness layer will be used for verification. A density threshold of 30%[8] should be used to derive the built-up layer. This is not intended to be a separate product, but instead will be calculated for the verification only, because density products cannot be verified. Methodology and outcome of verification shall be clearly documented.

Where commission errors in Imperviousness2009 layer are detected during production, they will be labelled in the 2012 product and excluded from the changes in the change layer 2012-2009, but not used for reprocessing of the 2006 and 2009 layers.

Figure 1 shows the interrelation between elements of imperviousness time series products 2006, 2009, 2012, in order to achieve full consistency of the 100m time series products.

[8] http://www.eea.europa.eu/data-and-maps/data/eea-fast-track-service-precursor-on-land-monitoring-degree-of-soil-sealing-100m-1

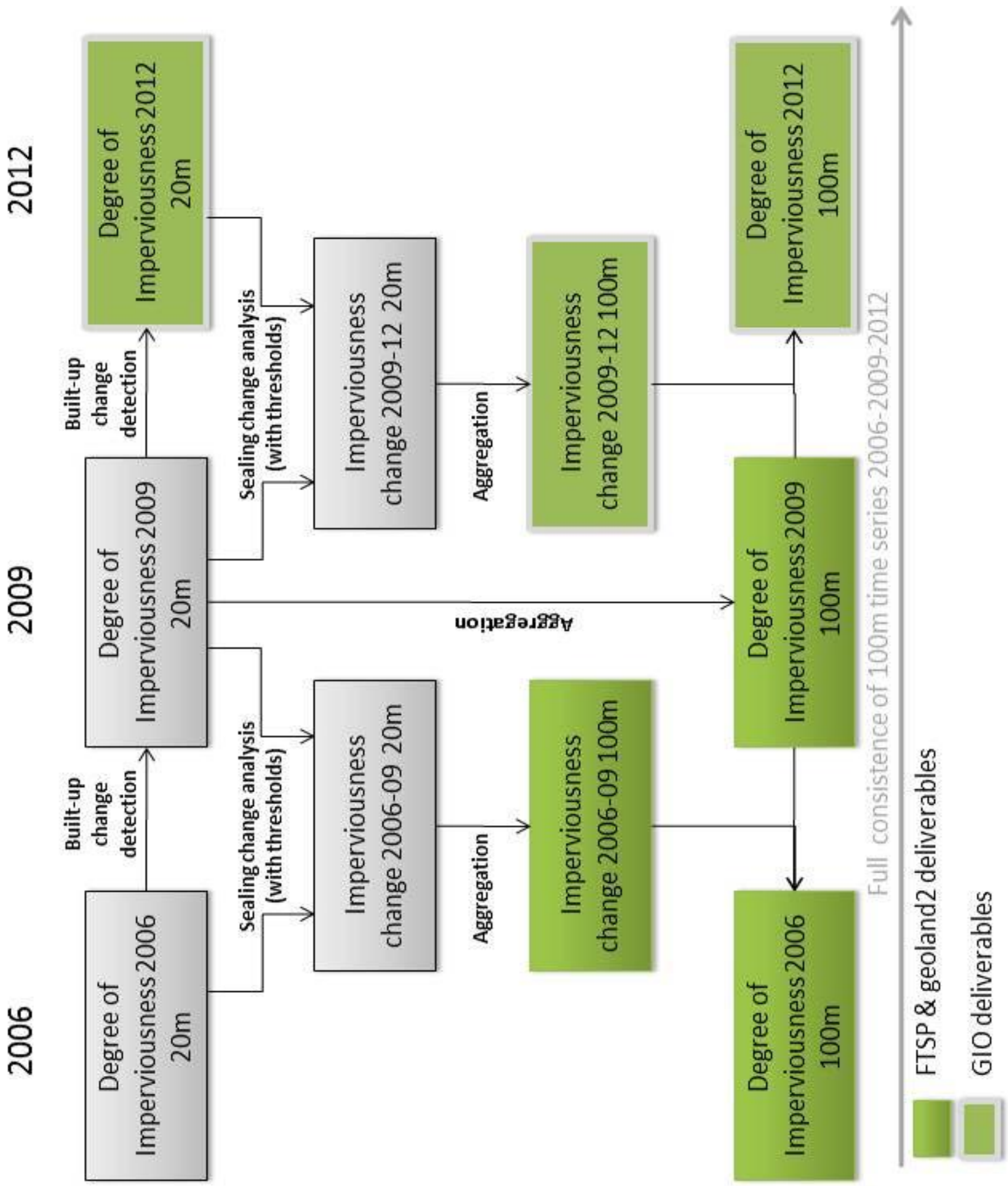


Figure 1: Interrelation between elements of Imperviousness time series products: 2006, 2009, 2012. (J. Weichselbaum, GeoVille)

Table 1: *Detailed specification of the HR Imperviousness to be verified*

Product
Degree of Imperviousness 2012, 20m x 20m, national projection
Methodology
Input data (in national projection): - *Degree of Imperviousness 2009 20m x 20m data* - *two coverages of IMAGE2012* *Production steps:* 1. *All IMAGE2012 NDVI images are calibrated with the* ***Degrees of Imperviousness 2009 20m x 20m*** 2. *NDVI images are mosaicked to Minimum and Maximum mosaics* 3. *The NDVI mosaics are compared with the Degrees of Imperviousness 2009 20m x 20m and change candidates are automatically derived with a rule-based approach* 4. *The derived built-up change candidates are visually corrected and supplemented, resulting in the final built-up changes 2009 – 2012 20 x 20m (no delivery)* 5. *For the total built-up area 2012, which is comprised of the 2009 built-up area plus the built-up changes 2009 – 2012, the final* ***Degrees of Imperviousness 2012 20m x 20m*** *are then modelled using the calibrated Minimum and Maximum NDVI mosaics and as reference the Degrees of Imperviousness 2009 20m x 20m.* *Imperviousness change 2009-2012 20m x 20m (no delivery; input to calculate the imperviousness change layer 2009-2012 100m x100m in a later step) is derived by subtracting Degree of Imperviousness 2009 from Degree of Imperviousness 2012, and applying a sealing difference threshold of 30 degrees and a MMU of 5 pixels to sealing changes within built-up areas 2009 and 2012 (built-up at both dates). No MMU or threshold is applied to sealing differences due to built-up increase or decrease, as these are visually edited (step 4) and thus represent interactively approved changes*
Geometric resolution
Pixel resolution 20m x 20m
Coordinate Reference System
National projection systems for country data sets
Geometric accuracy (positioning scale)
According to orthorectified satellite imagery delivered by ESA (target accuracy: 20m RMSE)
Verification
Verification to be performed by interested countries (or Service Provider). The built-up / non-built-up map will be verified. A threshold of 30% should be applied in transforming imperviousness to built-up. The built-up map will be produced as part of the verification process, i.e. by the country or the Service Provider
Data type
Raster

Raster coding
Thematic pixel values *0: all non-impervious areas* *1-100: imperviousness values* *254: unclassifiable (no satellite image available, or clouds, shadows, or snow)* *255: outside area*
Metadata
According to INSPIRE metadata standards

4.2 HRL Forest

Service element 1:

A high resolution forest dataset will be produced as a baseline for further monitoring. A per-pixel classification of tree cover[9] density and of dominant leaf type (broadleaf, coniferous) will be provided. A tree cover map (derived from tree cover density map by thresholding for the purposes of verification (not a separate product to be delivered by the SP)) and forest type map will be used for the verification task. Verification shall be carried out on the 20m x 20m maps (tree cover and forest type). Methodology and outcome of quality control has to be clearly documented. The following definitions shall be applied:

1. Tree cover density product:

The Tree Cover Density product shall be mapped with the following main specifications:

- No MMU (pixel resolution)
- Minimum Mapping Width: 20m
- Tree Cover Density range 0-100% (in the range 1-10%, the consortia will apply a regionally individual threshold as low as possibly can be detected from the available 20m resolution imagery).

The following table provides details on which features shall be included or excluded in the "tree cover" mapping.

Included Features	
Evergreen / non-evergreen broadleaved, sclerophyllous and coniferous trees	
Orchards, olive groves, fruit and other tree plantations, agro-forestry areas, forest nurseries, regeneration, transitional woodlands	
Alleys, wooded parks and gardens	
Groups of trees within urban areas	
Forest management/use features inside forests (forest roads, firebreaks, thinning, etc.)	Included if tree cover can be detected from the 20m resolution imagery per

[9] "tree cover" is used instead of "crown cover" in view of compliance with the global land cover initiative (GEO)

Forest damage features inside forests (partially burnt areas, storm damages, insect-infested damages, etc.)	pixel
Excluded Features	
Open areas within forests (roads, permanently open vegetated areas, clear cuts, fully burnt areas, other severe forest damage areas, etc.)	Excluded if no tree cover can be detected from the 20m resolution imagery per pixel
Shrubland	
Mediterranean bush lands (macchia, guarrigue etc.)	
Dwarf pine / green alder in high-mountainous areas	

2. Forest (following FAO definition[10]):

Contrary to the tree cover density product non-forest trees are excluded following the FAO forest definition. This is specified e.g. in the technical annex to the specific contracts (Phase I document), page 10: "*Forest (following FAO)"* with a link to: www.fao.org/docrep/006/ad665e/ad665e06.htm

Includes (FAO): forest nurseries and seed orchards that constitute an integral part of the forest; forest roads, cleared tracts, firebreaks and other small open areas < 0.5 ha. Forest in national parks, nature reserves and other protected areas such as those of specific scientific, historical, cultural or spiritual interest; windbreaks and shelterbelts of trees with an area of more than 0.5 ha and width of more or equal than 20 m; plantations primarily used for forestry purposes, including cork oak stands. For the EEA purpose forest cover in traditional agroforestry system such as Dehesa / Montado are included.

Excludes (FAO): land predominantly used for agricultural practices. In this sense fruit trees and olive groves are also excluded. Gardens and urban parks are also not considered as forest.

For this product, a minimum „Forest“ definition is applied, following the FAO definition:

- MMU = 0.5 ha
- Minimum Mapping Width (MMW) = 20m
- Tree Cover Density ≥10-100%

Considering different options for separating real forest areas as per the FAO definition from non-forest tree-covered areas (i.e. trees predominantly used for agricultural practices, trees in an urban context) for the 20m Forest Type product, the following strategy was agreed:

In order to avoid misinterpretation of the 20m x 20 m Forest Type product by potential users, specifically with respect to the contents / accuracies / information origins of the agreed additional information on "trees predominantly used for agricultural practices - broadleaved" (as derived from CLC classes 2.2.2 and 2.2.3) and "trees in urban context – broadleaved and

[10] http://www.fao.org/docrep/006/ad665e/ad665e06.htm

coniferous" (as derived from a filtering applied to HR Imperviousness), this additional information is made available through a labeling approach. In technical terms, this requires that in addition to the 20m x 20m Forest Type product's basic raster data set with the nomenclature of [0: all non-tree areas, 1: broadleaved trees, 2: coniferous trees – which would actually be more a dominant leaf type layer] an additional support raster data set is provided with information on the occurrence of "trees predominantly used for agricultural practices - broadleaved" and "trees in urban context – broadleaved and coniferous" – both only within the confines of broadleaved and coniferous tree areas of the basic product (not beyond!) – in the sense of additional, and not contradicting/conflicting information (as would be the case e.g. if CLC orchards would be identified in non-tree areas).

Such approach will allow providing to users the required real "Forest" information, which will then actually be contained in the combination of the 2 raster data sets described above. This will allow:

(i) to have a clear, traceable and explainable separation of the different sources of information, i.e. the 20m x 20m highly-accurate GIO classification vs. the 25 / 5 ha CLC-based overlay information, without already giving to users a "hard" intersection (with all associated problems of misinterpretation);

(ii) users nevertheless to simply derive the real "Forest" areas (close to the FAO definition) by simple GIS operations between these 2 raster data sets, i.e. by doing an intersection;

(iii) to preserve the mapped dominant leaf type information on pixel-level for various later applications, also e.g. in the climate change domain– without irreversibly casting away all non-Forest tree-based information.

***Table 2a**: Detailed specification of HR Forest products to be verified: tree cover density*

Product
Tree cover density 2012, 20m x 20m, national projection
Methodology
Automated analysis of biophysical parameters derived from multispectral and multitemporal Earth Observation data. Interactive editing at the end of the production chain.
Geometric resolution
Pixel resolution 20m x 20 m
Coordinate Reference System
National projection systems for country data sets
Geometric accuracy (positioning scale)
According to orthorectified satellite imagery delivered by ESA (target accuracy: 20m RMSE)
Verification
Verification to be performed by interested countries (or Service Provider). The tree cover map will be verified. In the lack of proper definition of threshold, for practical purposes a threshold of 30% shall be applied for transforming the tree cover density into the tree cover map. The tree cover map will be produced as part of the verification process, i.e. by the country or the Service Provider.
Data type
Raster
Raster coding
Thematic pixel values *0: all non-tree areas* *1-100: tree cover density* *254: unclassifiable (no satellite image available, or clouds, shadows, or snow)* *255: outside area)*
Metadata
According to INSPIRE metadata standards

Table 2b: *Detailed specification of HR Forest products to be verified: forest type*

Product
Forest type 2012, 20m x 20m, national projection, including two grids
Methodology
Classification of satellite imagery by using the tree-cover density layer to provide the "Dominant leaf type" layer. In an additional step, trees not used for forestry are identified, using, CLC and HRL Imperviousness. In order to avoid misinterpretation of the 20m Forest Type product by potential users, specifically with respect to the contents / accuracies / information origins of the agreed additional information on "trees predominantly used for agricultural practices - broadleaved" (as derived from CLC classes 2.2.2 and 2.2.3) and "trees in urban context – broadleaved and coniferous" (as derived from filtering applied to HR Imperviousness), this additional information is made available through a labeling approach in a separate grid. This requires that in addition to the 20m "Dominant leaf type" product an additional support raster data set is provided with information on the occurrence of "trees predominantly used for agricultural practices - broadleaved" and "trees in urban context – broadleaved and coniferous" – both only within the confines of broadleaved and coniferous tree areas of the basic product.
Geometric resolution
Pixel resolution 20m x 20m, MMU = 0.5 ha
Coordinate Reference System
National projection systems for country data sets
Geometric accuracy (positioning scale)
According to orthorectified satellite imagery delivered by ESA (target accuracy: 20m RMSE)
Verification
Verification to be performed by interested countries (or Service Provider).
Data type
Raster
Raster coding
Thematic pixel values for dominant leaf type layer 0: all non-forest areas *1: broadleaf forest* *2: coniferous forest* *254: unclassifiable (no satellite image available, or clouds, shadows, or snow)* *255: outside area*
Thematic pixel values for additional support raster dataset on non-forest trees *0: all non-tree areas* *3: trees predominantly used for agricultural practices – broadleaved* *4: trees in urban context – broadleaved and coniferous* *254: unclassifiable (no satellite image available, or clouds, shadows, or snow)* *255: outside area*
Metadata
According to INSPIRE metadata standards

4.3 HRL Grassland

A high resolution data set of permanent grassland shall be produced. The analysis will use the three reference years (2006, 2009, 2012) to detect the permanent presence of grassland. The discrimination of permanent grassland from other agricultural land areas such as arable land, bare soil, needs to take into account seasonal variations. A minimum of 3 (worst case), an average of 4-5 and up to maximum 8 seasonal images of AWiFS data (reference year 2012) will be used as additional information for the classification process. Verification shall be carried out on the 20m x 20m map (map of permanent grassland). Methodology and outcome of verification has to be clearly documented.

Grassland includes the following landscape types:

- Pastures, grassland used for grazing or hay production (CLC classes 2.3.1[11], but also appears in classes 2.1.1 to 2.4.4).
- Cultivated or semi-natural grassland within forests, and grass covered surfaces within transitional woodland (appears in CLC classes 3.1.1-3.1.3, 3.2.4).
- Natural grassland in any surrounding (CLC class 3.2.1).
- Grassy areas with low (10%) fraction of scattered trees and shrubs.
- Alpine meadows with low fraction of bare rock or gravel.

Land covers not to be considered as grassland:

- Grassland in urban areas: parks, urban green in residential and industrial areas.
- Grass surfaces in sport and recreation areas, incl. golf courses.
- Clearcut areas, new forests.
- Areas of shrubs: areas dominated by moors and heathland (Atlantic, CLC class 3.2.2) or sclerophyllous vegetation (Mediterranean, CLC class 3.2.3).
- Surfaces covered exclusively by mosses and lichen (Subarctic).
- Peatland (either in natural condition or in excavation, CLC class 4.1.2).

Table 3: *Detailed specification for HRL grassland products to be verified*

Product
Map of permanent grassland 2006-2009-2012, 20m x 20m
Methodology
Use of seasonal time series of MR data to identify grasslands; combine HR data with MR data to achieve the required details; use of biophysical parameters in advanced classification.
Geometric resolution
Pixel resolution 20m x 20m
Coordinate Reference System
National projection systems for country data sets
Geometric accuracy (positioning scale)
According to orthorectified satellite imagery delivered by ESA (target accuracy: 20m

[11] CLC classes listed below indicate only types of landscape objects / elements / structures and not necessarily suggest the exclusive use of CLC

RMSE)
Verification
Verification to be performed by interested countries (or Service Provider).
Data type
Raster
Raster coding
Thematic pixel values *0: all non-grassland areas* *1: grassland areas* *254: unclassifiable (no satellite image available, or clouds, shadows, or snow)* *255: outside area*
Metadata
According to INSPIRE metadata standards

4.4 HRL Wetland

For wetland areas, the presence of surface water during the reference year (2012) will be mapped. Seasonal changes using AWiFS data will be used to map areas covered temporarily by water surfaces, as well as areas covered during the whole reference year.

Degree of wetness will be characterised by the Water Presence Index (WPI < threshold[12]). (WPI is calculated for each pixel as ratio of number of detected water occurrences and number of cloud free images; see definition of HRL Water under Ch. 4.5). A minimum of 3, and a maximum of 8 dates will be analysed depending on the availability of medium resolution data to provide a time series of Water Presence map. Mapping macrophytes (wetland vegetation) is considered a benefit.[13] Verification shall be carried out on the 20m x 20m wetland map. Methodology and outcome of verification has to be clearly documented.

Wetland should include:

- Wetlands associated to permanent water bodies
- Wetlands not associated to permanent water bodies
- Wetlands with vegetation (macrophyte) cover or without vegetation
- Peatlands (having presence of surface water)[14]
- Salt marshes, salines (coastal and inland), intertidal flats

Wetlands should not include areas of:

- temporary inundations
- temporary water-logging because of e.g. snow melt or heavy rains
- permanent water surfaces (rivers, lakes, lagoons, estuaries)
- fishponds
- rice fields

[12] Threshold is not known. It should be found during implementation.

[13] Mapping wetland vegetation was not requested in the Technical Specification, but offered by the SP.

[14] Peatland is defined in restricted way due to technological limitations. Some countries required broader definition.

Table 4: *Detailed specification for HRL Wetland product to be verified*

Product
Wetland map 2012, 20m x 20m
Methodology
The temporary water layer derived for HRL Water product is the main input. MR AWiFS data (2012) will be used to analyse seasonal patterns. HR RapidEye data can be used to adjust boundaries. Macrophytes associated to wet areas are mapped with a region-growing algorithm. The temporary water layer and the wetland vegetation layer will be combined.
Geometric resolution
Pixel resolution 20m x 20m
Coordinate Reference System
National projection systems for country data sets
Geometric accuracy (positioning scale)
According to orthorectified satellite imagery delivered by ESA (target accuracy: 20m RMSE)
Verification
Verification to be performed by interested countries (or Service Provider)
Data type
Raster
Raster coding
Thematic pixel values *0: all non-wetland areas* *1: wetland* *254: unclassifiable (no satellite image available, or clouds, shadows, or snow)* *255: outside area*
Metadata
According to INSPIRE metadata standards

4.5 HRL Water

Permanent water bodies shall be mapped at 20m x 20m spatial resolution. The analysis will use the three reference years (2006, 2009, 2012) to detect the permanent presence of surface water. Seasonal AWiFS data will be used to separate temporary water from permanent water bodies for the reference year 2012 and exclude effects of seasonal changes in water coverage. Verification shall be carried out on the 20m x 20m map of permanent water bodies. Methodology and outcome of verification has to be clearly documented.

The mapping is based on the Water Presence Index (WPI). WPI is calculated for each pixel as a ratio of detected water occurrences and the number of cloud free images. Permanent water layer determined as pixels having WPI>threshold. The proper value of threshold should be determined by the Service Provide).

Water surfaces which should be included in HRL water:

- Permanent lakes, ponds (artificial and man-made) including fish ponds
- Rivers, channels permanently with water
- Coastal water surfaces: lagoons, estuaries

Water surfaces which should be excluded from HRL water:

- Sea and ocean[15]
- Temporary inundation and water-logging
- Liquid dump sites

Table 5: *Detailed specification for HRL Permanent water bodies product to be verified*

Product
Permanent water bodies 2006-2009-2012, 20m x 20m
Methodology
Water Presence Index is calculated for each pixel as ratio of number of detected water occurrences and number of cloud free images by using multi-seasonal imagery. Permanent water layer is determined as pixels having WPI>threshold. This will be repeated for 2006, 2009 and 2012.
Geometric resolution
Pixel resolution 20m x 20m
Coordinate Reference System
National projection systems for country data sets
Geometric accuracy (positioning scale)
According to orthorectified satellite imagery delivered by ESA (target accuracy: 20m RMSE)
Verification
Verification to be performed by interested countries (or Service Provider)
Data type
Raster
Raster coding
Thematic pixel values *0: all non-permanent water areas* *1: all permanent water areas* *254: unclassifiable (no satellite image available, or clouds, shadows, or snow)* *255: outside area*)
Metadata
According to INSPIRE metadata standards

[15] Although it is water body, but taken out because "see and ocean" is not interested by GIO Land users.

Table 6: *Summary of HRL products to be verified*

Product (20m x 20m, national projection)	Year	Classification	Type of product
Degree of Impervious-ness	2012	Degree of Imperviousness: 0 – 100%	Density layer
Tree cover density	2012	Degree of tree cover densi-ty: 0 – 100%	Density layer
Forest type	2012	Broadleaf or coniferous (plus non-forest)	Map layer
Permanent grassland	2006 – 2009 – 2012	Permanent grassland or not permanent grassland (bina-ry)	Map layer
Wetland	2012	Wetland or non-wetland (bi-nary)	Map layer
Permanent water bodies	2006 – 2009 – 2012	Water or non-water (binary)	Map layer

5 Verification of HRL products: principles

According to the Technical Specification the task of verification is defined as:

- aiming at identifying systematic classification errors that are eligible for improvement (enhancement),
- done in randomly selected samples,
- by visual inspection of the selected samples with existing reference data (e.g. topographic maps, aerial photography or others),
- omission and commission errors are to be checked in all five HRLs,
- verification shall be carried out on the intermediate products at full resolution (20m x 20m), in national projection.

In-situ data shall be specified, made available and used to support the verification process. The best in-situ data would be the national, high-resolution data with proper thematic coverage (e.g. ground survey data or thematic maps). Simple statistical comparison of HRLs derived from satellite imagery and existing national data regarding HRLs is not sufficient. The reasons are manifold:

- not all HRLs have corresponding in-situ map (database) of sufficient quality,
- maps (databases) are always generalised to some extent,
- maps are frequently outdated, and do not represent the current situation.

The preferred strategy is the use of in-situ data to compare and analyse with the HRL. A set of samples should be selected for detailed evaluation of HRL data quality.

Verification is planned on three levels:

- General overview of data quality (obligatory): having the purpose to provide a general feeling about the data quality and orient the more detailed verification actions. The HRL should be compared to the best available in-situ data, and the major disagreements should be analysed.
- Look-and-feel verification in critical strata (obligatory): this will provide qualitative results in pre-determined locations, where classification problems are expected.
- Statistical verification[16] (highly recommended): by applying randomly selected samples the omission and commission errors are estimated.

5.1 Comparison of look-and-feel and quantitative verification

Both look-and-feel and statistical verification rely on sampling. Samples are evaluated based on relevant in-situ data by an expert (interpreter). The two approaches are compared in Table 7.

[16] Some member states proposed quantitative verification

Table 7: *Comparison of look-and-feel and statistical verification*

	Look-and-feel	Statistical
Sampling	Samples positioned on presumed critical areas (targeted sampling, non-random)	Random sampling of the entire study area (e.g. country)
Local expertise needed for sampling	Yes	No
Evaluating of samples	Visual comparison of the HRL map produced by the SP with relevant in-situ data in order to see if the HRL map is correct in the sampling location.	Estimation of the HRL map value at the sample point by using relevant in-situ data. HRL produced by the SP is not used in interpretation. Comparison of estimated HRL map value with the results of the SP is done only after all samples have been evaluated.
Are omission and commission errors checked?	Depending on sampling	Always intended to be checked. However, the uncertainty of the estimation of omission error may be very high if HRL population is small (e.g. imperviousness).
Characterisation of overall accuracy	Poor, because only critical areas are checked	Good, as samples cover all (or majority) of the area
Expression of results	Qualitative (subjective, limited comparability e.g. between countries)	Quantitative, comparable (between layers or countries)
Main use of results	Indicates performance in specific locations or in specific landscape types. Provides suggestions for enhancements.	Can predict classification accuracy and provide estimate for the volume of enhancement work. Below certain accuracy (e.g. 80% either in omission or commission) enhancement is not implemented.

6 Verification basics

6.1 Land Cover density layers

As seen in Table 6 land cover density layers (LC-DL) as well as Land Cover map layers (LC-Map) have to be verified. A LC-DL includes values in an interval (usually from 0% to 100%). The accuracy of a LC-DL can be estimated with reference values of the LC density, measured on high resolution reference imagery. The reference estimations, i.e. the verification are usually performed on random sample cells, or for practical reasons in a regular grid[17]. The results of the verification are visualised on a scatter-plot between the satellite based estimation and the in-situ data based estimation.

[17] http://www.eea.europa.eu/data-and-maps/data/eea-fast-track-service-precursor-on-land-monitoring-degree-of-soil-sealing-100m-1

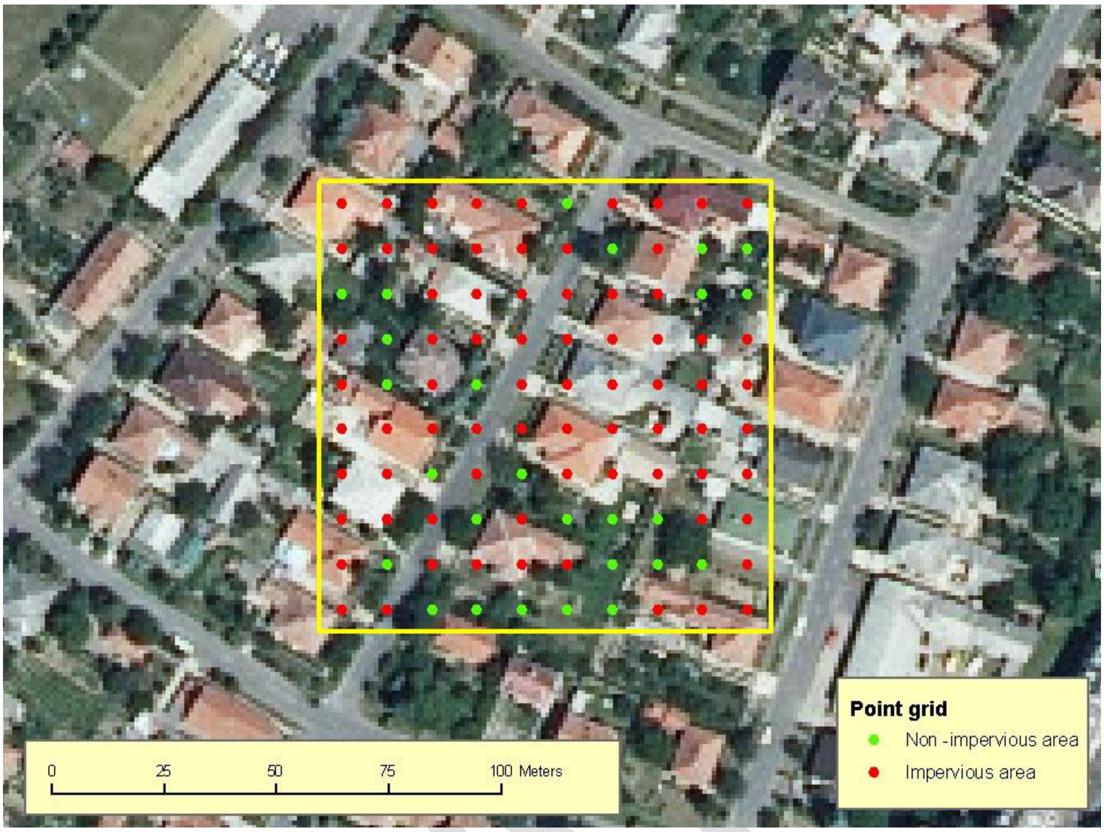


Figure 2 *Validation methodology of European SSE data. By counting the number of impervious points inside the 100x100 m grid cell (total of 100 points), the interpreter estimates sealing degree of the sample cell*[18].

To perform verification of LC-DLs at 20m x 20m resolution would need very-high resolution reference imagery (e.g. orthophotos), which will probably not be available in all EEA39 countries. Moreover, this kind of accuracy assessment is time-consuming, because several points have to be checked for each sample (Figure 2). Therefore verification of density layers is not proposed in these Guidelines. Instead, the corresponding LC-Map layers should be verified.

On the other hand, quality control of LC-DLs in 100m x 100m resolution is highly important and will be the target of the European validation.

6.2 Land Cover map layers

As seen in Table 6 we have three different kinds of HR Land Cover Maps for verification:

- forest type map
- map of permanent grassland
- wetland map
- map of permanent water bodies

Given that in the two HRLs (imperviousness, tree cover) we cannot verify densities (LC-DLs), the corresponding LC-Maps are to be verified. These binary LC-Maps are derived from

[18] http://www.eea.europa.eu/data-and-maps/data/eea-fast-track-service-precursor-on-land-monitoring-degree-of-soil-sealing-100m-1

the corresponding density layer through properly defined decision rules, which include a density threshold and occasionally an area threshold (MMU):

- Built-up map (derived from HRL imperviousness) decision rule: a 20m x 20m area is considered built-up, if imperviousness ≥ 30%[19]. MMU = 20m x 20m.
- Tree cover map (derived from HRL tree cover density) decision rule: a 20m x 20m area is considered tree-covered, if tree-density ≥ 30% (see Table 2a). MMU = 20m x 20m.

In practical terms the quantitative verification the reference interpretation should go on independently, without knowing the HRL. The 30% threshold (imperviousness) and the 30% threshold (tree density) will be applied by the photointerpreters in the best possible way, because accurate density estimation will not be possible (see Ch. 5.1)[20].

6.3 Omission and commission errors

The reference interpretation that supports the quantitative verification should be performed in randomly selected samples. The objective of the reference interpretation is to decide whether the cell belongs to the specific LC-Map class or not (e.g. is it grassland or not grassland). The result of the reference interpretation of the sample is compared to the HRL LC-Map. Basic accuracy parameters[21] are:

- Number of incorrectly classified samples
- Misclassification rate

By examining a single LC class, commission and omission error rates are the most informative quality parameters to be used to characterize classification accuracy. In this special case commission / omission error rates are defined as:

- Commission error rate = area classified erroneously as belonging to the HRL class (e.g. water, forest, etc.) divided by the real HRL class area (e.g. water, forest, etc.).
- Omission error rate = area classified erroneously as not belonging to the HRL class (e.g. non-water, non-forest, etc.) divided by the real HRL class area (e.g. water, forest, etc.).

Because the real area of a HRL class is not known (it is only estimated by the HRL), commission error should be calculated in practice as:

- Commission error rate = number of cases the samples cover an area classified erroneously as belonging to the HRL class, divided by the total number of samples distributed randomly within the mapped HRL class.

The estimation of omission error is more complex. First we estimate the commission error of the non-HRL class:

- Commission error rate for non-HRL class = number of cases the samples cover an area classified erroneously as not belonging to the HRL class (e.g. non-water, non-

[19]Validation has proven that 30% threshold has provided the optimal built-up classification accuracy in Soil Sealing 2006. Actually, the previously suggested 80% threshold captured only the densely built-up areas. http://www.eea.europa.eu/data-and-maps/data/eea-fast-track-service-precursor-on-land-monitoring-degree-of-soil-sealing-100m-1

[20]In the verification process the photointerpreter should check "presence of trees" or "absence of trees" in the selected sample-points, keeping in minds the 30% threshold as much as possible.

[21] ISO19157 Geographical Information – Data Quality

forest, etc.), divided by the total number of samples distributed randomly outside the mapped HRL class.

The area of commissions of a non-HRL (e.g. non-forest) class is equal to the area of omissions of a HRL (e.g. forest) class, but the calculated error rates would only be equal if the size of HRL and non-HRL class would be the same (i.e. both 50% of the total area). As omission error is defined relative to the HRL class area, the second step is to calculate the omission error for the HRL class from the commission error of the non-HRL class relational to the class areas:

$$E_{omission(HRLclass)} = E_{commission(non-HRLclass)} \cdot \frac{Area_{Total} - Area_{HRLclass}}{Area_{HRLclass}}$$ [22]

Estimating the uncertainty of accuracy values is presented in Annex 2.

6.4 Minimum sampled patch size

The geometric accuracy of remotely sensed imagery delimits the size of the patches sampled for verification. If the sampled patch size is too small, the registration error between IMAGE2012 (used to derive the HRLs) and the reference imagery (e.g. orthophoto[23]) will introduce significant uncertainty. According to ESA specification the expected registration accuracy of IMAGE2012 is 20 m[24]. In the optimal case the geometric shift between the two products is ½ satellite image pixel (i.e. 10 m), consequently a 20m x 20m pixel will cover significantly different area on IMAGE2012 and on the in-situ data (Figure 3).

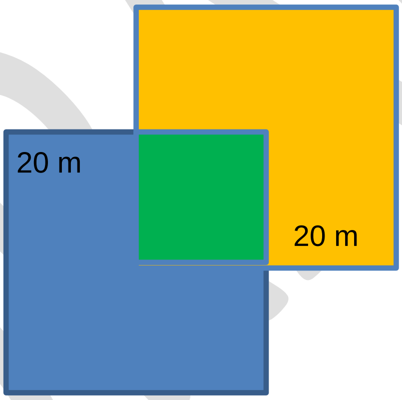


Figure 3 *Modelling thematic mistakes coming from locational inaccuracy of IMAGE2012*

In sampling for statistical verification it is proposed to select samples from patches having size of minimum 3x3 pixels around the sample. I.e. the valid sample should have a minimum of 3x3 neighbourhood belonging to the layer to be sampled (Figure 4). This way disagreement between the HRL and in-situ because of locational uncertainty will be reduced.

[22] Explanation:

$$E_{commission(non-HRL)} = \frac{Area_{commission(non-HRLclass)}}{Area_{Total} - Area_{HRLclass}}; E_{omission(HRLclass)} = \frac{Area_{omission(HRLclass)}}{Area_{HRLclass}}; Area_{omission(HRLclass)} = Area_{commission(non-HRLclass)}$$

[23] Orthophotos with regional coverage usually has positional accuracy in the range of 1 meter.

[24] http://gsc-prod.netcetera.ch/c/document_library/get_file?uuid=9f57e0f4-af57-43ca-aa26-b9418fbf40ea&groupId=10725

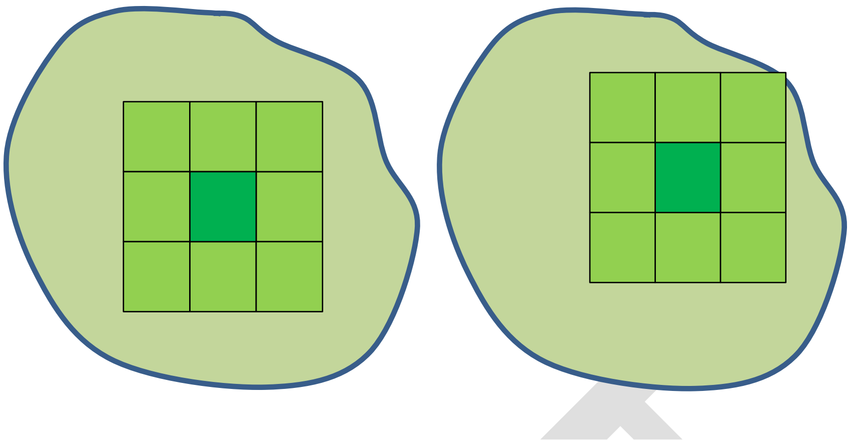


Figure 4a: *Example of a correct sample point selection. The randomly selected pixel (dark green) is surrounded by the same HRL value (e.g. forest) in the 3x3 window around the selected pixel. If we consider the half pixel locational instability, this point will be still fall on the same HRL value, so no error will be introduced into the verification statistics due to locational instability between the HRL and the in-situ data.*

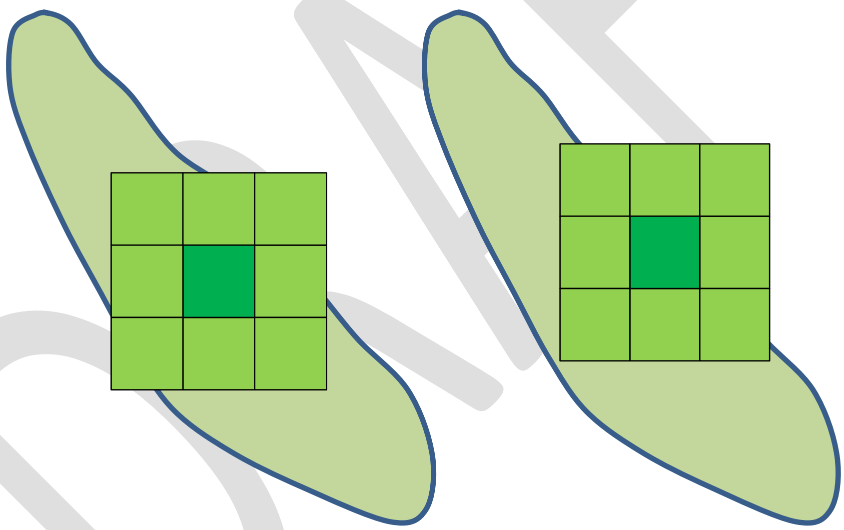


Figure 4b*: Example of an incorrect sample point. The randomly selected pixel (dark green) is not surrounded by the same HRL value (e.g. forest) around the 3x3 window around the selected pixel. If we consider a half pixel locational instability this point will fall not on the same HRL value, so an error will be introduced into the verification statistics, just because of the locational instabilitv between the HRL and the in-situ data*.

In case of HRL grassland, wetland and water, where MR satellite imagery is also applied to derive these layers, selecting samples from larger patches is even more justified.

7 Verification of HRL products: practice

The three levels of verification (Ch. 5) are discussed here more in detail.

7.1 General overview of data quality

It has the purpose to provide a general feeling about the data quality and orient the more detailed verification action(s). The HRL should be intersected with the best available in-situ database, and major areas of disagreements should be analysed (Figure 5). If the disagreements are modest in area, we can conclude that the HRL approximates the reality (represented by the in-situ) properly. If there are serious disagreements, the HRL quality might be questionable. In-depth analysis might need to understand the differences, because in-situ data themselves are not necessarily 100% accurate. The areal distribution of disagreements could be used to locate candidate strata for look-and-feel.

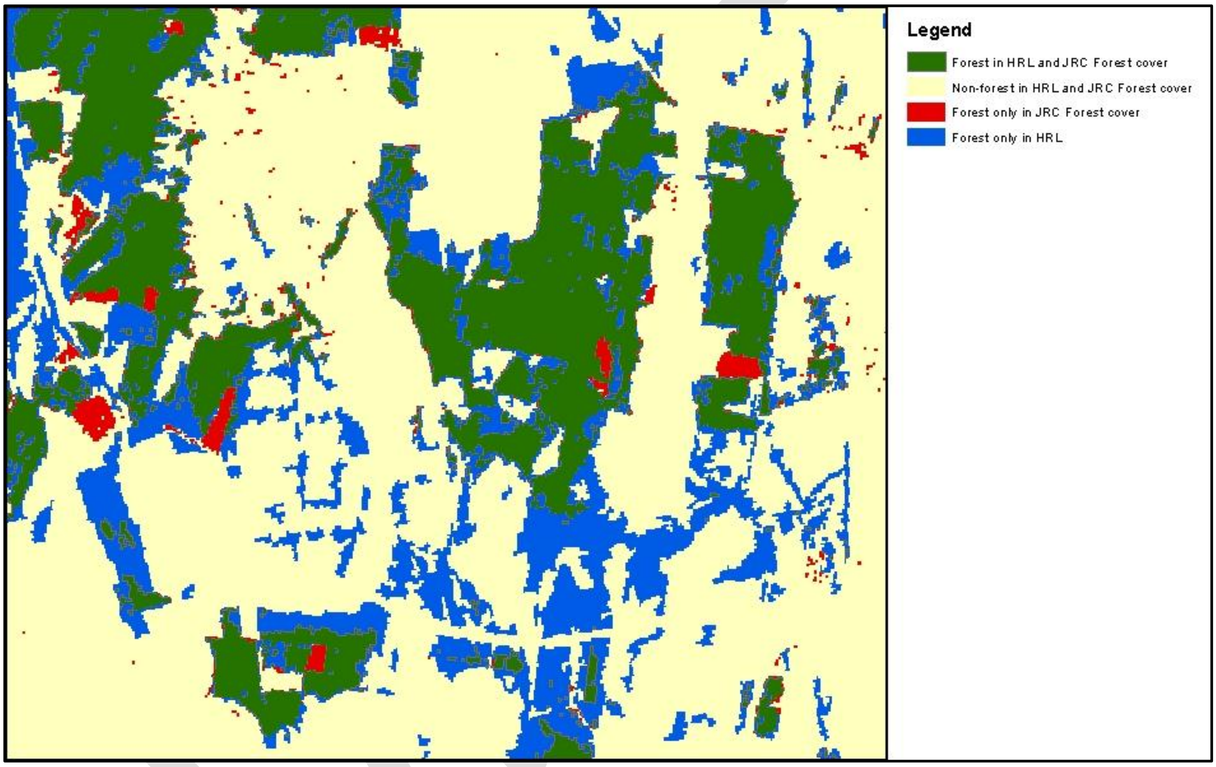


Figure 5 *Illustrating general overview of HRL tree cover data quality by means of comparing HRL Forest map (Geoland2) and JRC Forest cover map. Areas displayed in blue and red colours highlight the differences between the two datasets*

The following national databases can be used for providing the general overview:

- Imperviousness: database of buildings, road-database
- Forests: forest inventory or forestry field plots or forest layer of topographic map
- Grassland: grassland inventory, LPIS[25] data or grassland layer of topographic map
- Wetland: wetland inventory or wetland layer of topographic map
- Water: inventory of water bodies or water layer of topographic map

A national land cover database with resolution better than standard European-level CLC could help in case of all HRLs.

[25] LPIS = Land Parcel Identification System

7.2 Verification with look-end-feel method

Look-and-feel examples should be selected in a way, which ensures that areas of potential classification errors are checked. Selecting proper locations needs a-priory knowledge about the country or region and that of the expected performance of the classifier for each HRLs (e.g. we know that natural bright features, like quarries can be erroneously classified as impervious). Consequently, sample selection is not random in nature. A larger surrounding of the sample, not just a single pixel or a small group of pixels should be visually interpreted using all available in-situ data. The HRL should be displayed on the top of in-situ data layer to allow for easy overview of HRL quality, i.e. the evaluation is not blind. Sufficient number of samples should be checked for each sample type (strata) in order to be able to provide evaluation on five grades (Table 8):

Table 8: *Summarised qualitative evaluation of HRLs in five grades*[26]

excellent	meaning that you expect that the accuracy of the HRL is reaching almost 100%; no errors could be found in the areas that were verified
good	meaning that you are confident that the accuracy of the HRL is at least 85 %; only sporadic errors were encountered in the areas that were verified
acceptable	meaning that you estimate that the accuracy of the HRL in most of the verified areas will probably reach an accuracy of 85 %; some minor errors could be detected in the areas that were verified
insufficient	meaning that you do not expect that the accuracy of the HRL will reach the minimum 85 % accuracy; you encountered several errors in different regions
very poor	meaning that you are confident that the accuracy of the HRL is bad and much below the 85%; majority of the areas verified are wrongly mapped

Recommendations regarding the selection of “critical” strata for look-and-feel verification are shown in Tables 9 – 13. These lists are just indicative. National experts are encouraged to select any other strata, if relevant for the verification. The general overview of data quality (Ch. 7.1) can be used also to find strata for more detailed evaluation. Minimum 5 -10 samples are preferred to examine in each strata (lines of Tables 9-13), which can be a compromise between reliability and efforts of verification. In larger countries, or if the results are doubtful, more samples may be required. For larger countries, with a number of climatic zones (e.g. FR, NO, TR…) more than one look-and-feel exercise is proposed to be accomplished to characterise the different climatic zones separately.

Table 9: *Recommended selection of look-and-feel samples for HRL Imperviousness*

Strata to reveal commission errors	Explanation
Mineral extraction sites	Non-vegetated natural areas are often misclassified as impervious
Dump sites	
Construction sites	
Sand, gravel, dunes	
Bare rocks	
Sparsely vegetated areas	
Strata to reveal omission errors	
Major cities	Sometimes roofs with unusual colour are omitted

[26] Guidelines for verification of high resolution soil sealing layer - Qualitative assessment - Prepared by: C. Steenmans and A. Sousa, EEA, 2007

Continuity of highways	Sometimes highways are not continuous
Ports	Sometimes buildings and sealed surfaces in ports are not fully included
Major airports	Sometimes buildings and sealed surfaces in airports are not fully included

Table 10a: *Recommended selection of look-and-feel samples for HRL Forest (tree cover)*

Strata to reveal commission errors	Explanation
Moors and heathland and sclerophyllous vegetation	Moors and heathland and sclerophyllous vegetation (low shrubs) can be erroneously classified as forest
Strata to reveal omission errors	
Urban vegetation (trees in parks, cemeteries, etc)	Checking that all tree covered areas are classified properly
Trees in sport and recreation areas	
Orchards, fruit trees	
Olive plantations	
Lowland forests, broadleaved	
Lowland forests, coniferous	
Mountain forests (incl. forests on sunlit side and in shadow), broadleaved	
Mountain forests (incl. forests on sunlit side and in shadow), coniferous	
Forest along rivers & lakes	
Coastal forests	
Forest plantations (broadleaved and coniferous)	
Agricultural areas with scattered small forest patches	
Forest component of agroforestry areas (Dehesa/ Montado)	

Table 10b: *Recommended selection of look-and-feel samples for HRL Forest (forest type)*

Strata to reveal commission errors	Explanation
Major cities	Urban vegetation (line of trees, cemeteries, etc.) can be erroneously classified as forest
Sport and recreation areas	Group of trees in sport and recreation areas can be erroneously classified as forest
Orchards, fruit trees	Orchards, fruit trees can be erroneously classified as forest
Olives	Olives plantations can be erroneously classified as forest
Moors and heathland and sclerophyllous vegetation	Moors and heathland and sclerophyllous vegetation (low shrubs) can be erroneously classified as forest
Strata to reveal omission errors	

Lowland forests, broadleaved	Checking that all forests are classified properly
Lowland forests, coniferous	
Mountain forests (incl. forests on sun-lit side and in shadow), broadleaved	
Mountain forests (incl. forests on sun-lit side and in shadow), coniferous	
Forest along rivers & lakes	
Coastal forests	
Forest plantations (broadleaved and coniferous)	
Agricultural areas with scattered small forest patches	
Forest component of agroforestry areas (Dehesa/ Montado)	

Table 11: *Recommended selection of look-and-feel samples for HRL Grassland*

Strata to reveal commission errors	Explanation
Major cities	Urban vegetation (parks, cemeteries, etc.) can be erroneously classified as (agriculture) grassland
Major airports	Grassy vegetation in airports can be erroneously classified as (agriculture) grassland
Sport and recreation areas	Grass surface of sport and recreation areas (golf courses, grass-surfaced airports) can be erroneously classified as (agriculture) grassland
Clearcut areas	Clearcut areas can be erroneously classified as (agriculture) grassland
Strata to reveal omission errors	
Lowland grasslands	Checking that all grasslands are classified properly
Mountain grasslands	
Wet grasslands along rivers & lakes	
Coastal grasslands	
Pastures and hayfields	
Agroforestry areas including grass cover	

Table 12: *Recommended selection of look-and-feel samples for HRL Wetland*

Strata to reveal commission errors	Explanation
Temporary inundations by flood	Temporary inundations should not be classified as wetland
Temporary water-logging because of e.g. snow melt or heavy rains	Temporary water-logged areas should not be classified as wetland
Permanent water surfaces (rivers, lakes, lagoons, estuaries)	Permanent water surfaces (rivers, lakes, lagoons, estuaries) should not be classified as wetland
Fishponds	Fishponds should not be classified as wetland
Rice fields	Rice fields should not be classified as wetland
Liquid dump sites	Liquid dump sites should not be classified as wetland

Strata to reveal omission errors	
Wetlands associated to permanent water bodies	Checking that all wetlands are classified properly
Wetlands not associated to permanent water bodies	
Wetlands with vegetation (macrophyte) cover	
Wetlands without vegetation (macrophyte) cover	
Peatland (having presence of surface water[27])	
Coastal wetlands, salines	

Table 13: *Recommended selection of look-and-feel samples for HRL Water bodies*

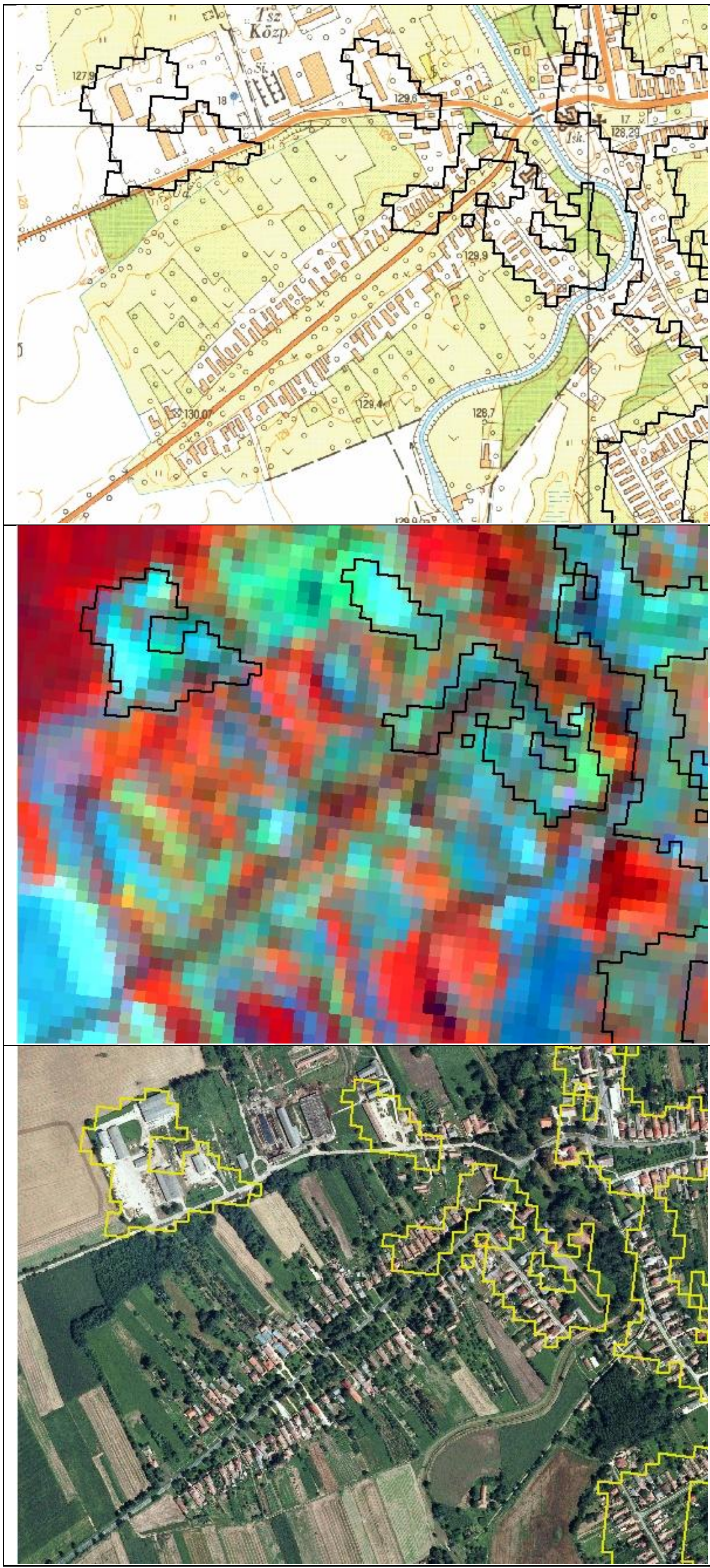
Reveal commission errors	Explanation
Liquid dump sites	Liquid dumpsites should not be classified as water body
Temporary water logged areas (e.g. next to main rivers)	Temporary water logged areas should not be classified as water bodies
Separation of lakes and sea & ocean	Sea & ocean (incl. fjords) should not be classified as water body
Strata to reveal omission errors	
Small lowland lakes (<25 ha area)[28]	Checking that all small water bodies are classified properly
Small mountain lakes (<25 ha area)	
Lakes in mining areas	
Lakes in recreation areas	
Fishponds	

In addition to the qualitative evaluation of each stratum some examples should be visualised as screen-shots to be placed into the HRL Verification Report (see Annex 1). In order to reduce the extent of the report the number of look-and-feel examples (see Examples 1 and 2) should be limited and only problematic cases should be shown. A short text attached to the example should describe the problem. Advices to select proper look-and-feel examples:

- Maximum 5-8 examples per each HRL, including both omission errors and commission errors. The examples should be representative. Avoid similar cases appearing more than once.
- The examples should clearly demonstrate the problem (i.e. attach multi-seasonal imagery in case of a wetland example).

[27] Peatland is defined in restricted way (Ch 4.4) due to technological limitations

[28] Assuming that >25 ha water bodies (identified by CLC already) is not worth to check



Example 1: *Example of omission error in HRL Imperviousness (2006)*

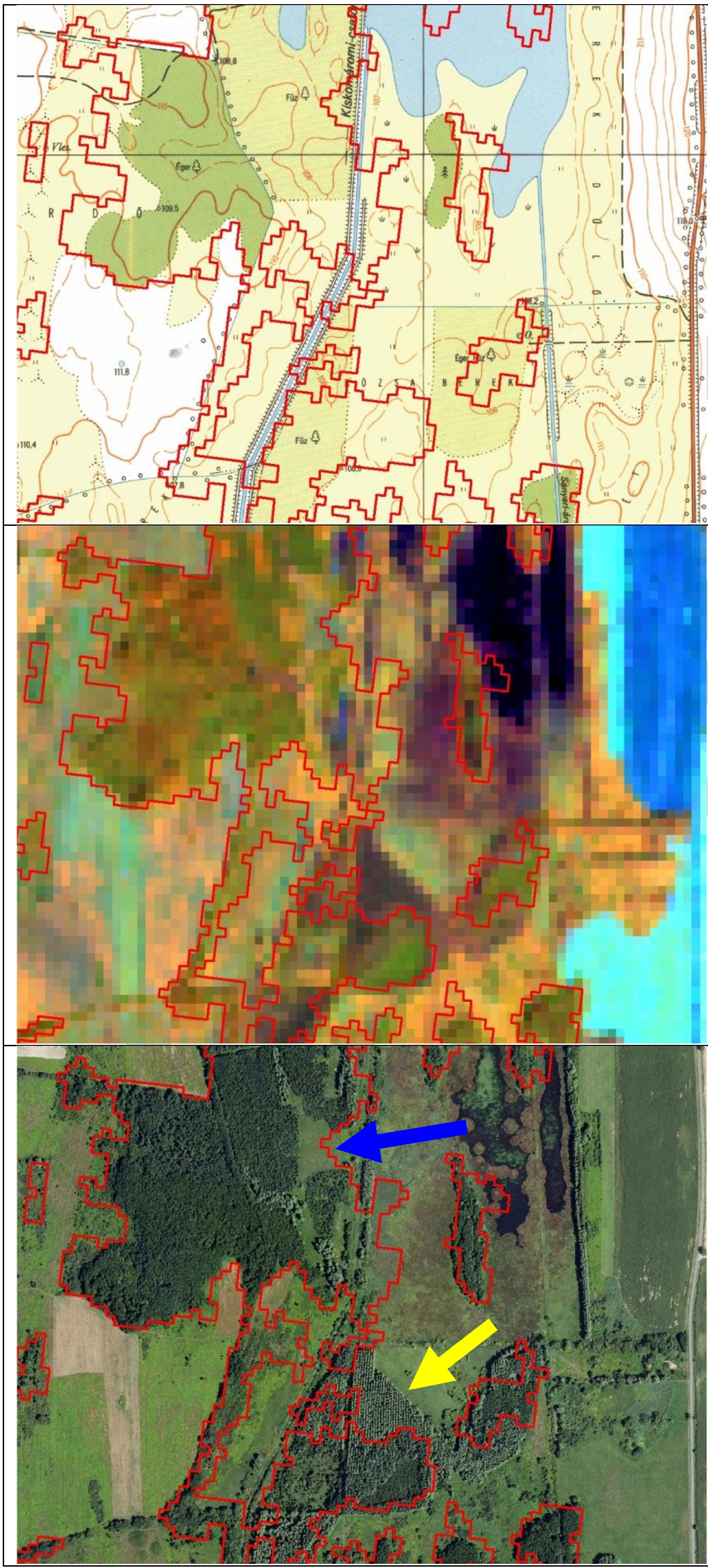
Top: Topographic map with HR built-up area map (produced with 30% threshold on Soil Sealing Enhancement 2006 data) overlaid (black/yellow outlines)

Middle: IMAGE2006

Bottom: VHR natural colour orthophoto

Location: Mihályi, Hungary

Interpretation: SW part of the built-up area is omitted from built-up map.



Example 2: *Example of commission and omission errors in HRL Forest (2008)*

Top: Topographic map with HR Forest area overlaid (red outlines)

Middle: IMAGE2006

Bottom: VHR natural colour orthophoto

Location: Near Zalakaros, Hungary

Interpretation: Commission (blue arrow) and omission (yellow arrow) errors.

7.3 Quantitative verification

To allow for comparison between verification results in different countries for the same HRL, or between different HRLs of the same country, the look-and-feel approach has only limited capabilities. Therefore it is highly recommended to derive quantitative measures to characterise accuracy. A minimum amount of statistically meaningful random samples should be selected to check commission as well as omission errors in each HRLs. Locations should be blindly interpreted (i.e. not knowing the value in the HRL) by using the best available in-situ data. This will be compared later with the HRL value produced by the Service Provider. Quantitative measures include the calculation of omission and commission error rates (see Ch. 6.3).

Commission error can be estimated with acceptable accuracy using relatively low amount of samples. The possibility (i.e. the efforts needed) of estimating omission error rates heavily depends on the size of the target class. As indication Table 14 includes the approximate size of HRLs in percentage of total area in five selected countries and in the total of EEA39. Imperviousness values were taken from European Soil Sealing (SSE2006) layer[29]. Forest, grassland and wetland statistics were derived from CLC2006[30], and should be considered as indicative, due to the low spatial resolution of CLC. The following CLC classes were used in computing these statistics:

- Forest: classes 3.1.1, 3.1.2, 3.1.3
- Grassland: classes 2.3.1, 3.2.1
- Wetland: classes 4.1.1, 4.1.2, 4.2.1, 4.2,2, 4.2.3
- Water: class 5.1.2

Table 14: *Estimation of per-cent coverage of high-resolution layer areas in selected countries and the EEA39*

Area	Impervious/ built-up	forests	grassland	wetland	water
Czech Republic	3.2	32.8	7.2	0.1	0.6
Cyprus	3.6	16.7	3.3	0.3	0.2
Finland	0.5	58.0	<0.5[31]	6.7	9.1
France	2.8	26.3	18.3	0.7	0.4
Romania	1.6	29.2	12.0	1.6	0.7
EEA39	1.8	28.5	10.6	2.4	2.2

Analysing Table 14 shows:

- Imperviousness / built-up is characterised by very low percentages, consequently verifying omissions will be difficult, because the non-impervious class is very large, and consequently very large number of samples would be needed to derive statistically meaningful verification results.
- Forests usually have relatively high percentage; therefore there is a possibility to verify omissions by using relatively small amounts of samples. However, there are exceptional countries, where forest cover is very low (e.g. in Iceland: < 1%).

[29] http://www.eea.europa.eu/data-and-maps/data/eea-fast-track-service-precursor-on-land-monitoring-degree-of-soil-sealing-100m-1

[30] Gy. Büttner, B. Kosztra, G Maucha and R. Pataki (ETC-LUSI): Implementation and Achievements of CLC2006, Final Draft, 2010

[31] Pekka Härmä: Value based on national land cover data (personal communication, 2012)

- Although grassland has usually smaller relative size than forest, there will be some countries where omission errors can be estimated with a not too large sample size.
- Wetland and water are both small sized layers, and will not provide an easy possibility to derive estimation for omission errors with acceptable number of samples. Regarding water bodies checking of large lakes (>25 ha, mapped by CLC) is not very meaningful. Efforts should concentrate on smaller water bodies (<25 ha, not mapped by CLC).

7.3.1 Number of samples

Commission error can be estimated with acceptable accuracy using relatively low amount of samples. Considering e.g. 100 samples drawn from any of the HR layers, the uncertainty of the estimation is ±3,57%[32] (see table in Annex 2). Uncertainties of the estimations based on random sampling is calculated as the standard deviation of a binomial distribution with (n, p) parameters, where n = number of all samples, p = proportion of erroneous samples)[33].

The case of omission error is complicated in case of small target population (e.g. imperviousness). Table 15 indicates the required number of samples calculated on the basis of mathematical statistics. If the target class covers < 10% of the total area, the number of samples required to estimate omission error is > 1000, i.e. needs lots of effort.

Table 15 *Number of samples needed to estimate omission error for target class with different sizes (expected omission error = 15%, uncertainty of estimation = ± 3.57%)*

P_{class} (%)	Number of samples
50	100
40	160
30	250
20	450
15	660
10	1050
5	2220
3	3800
1	11600
0.5	23500

In the case of commission error the uncertainty of the estimation (expressed as per-cent of the commission error) is independent from the area of the HRL class. The uncertainty of the estimation depends only on (n, p) parameters, where n = number of all samples, p = proportion of erroneous samples). In case of 100 samples the maximum uncertainty (maximum is always at p = 50%) is ± 5,0% (see Annex 2), in case of 250 samples the maximum uncertainty is 3,16%[34].

Choosing a pragmatic approach, the sample sizes for commission and omission errors for each HRL is fixed in 250. To allow for 10% spare samples to compensate for missing samples (e.g. clouded satellite image), 280 samples is recommended to be selected for each HRL and for both error types. This can be considered as a trade-off between reduced uncertainty and acceptable interpretation workload.

[32]G. Maucha: Validation of GMES HR layers with respect to change detection (EAGLE meeting, Frankfurt): http://sia.eionet.europa.eu/EAGLE/EAGLE_5rdMeeting_g2_MONINA_FFM/index_html
[33] same source as above
[34]All this calculation refers to a confidence interval of ± 1σ (appr. 68,3%.)

In HRLs having smaller per-cent coverage (30% limit, see Table 15) sampling should be restricted to a subset of the area in order to receive a meaningful result.

7.3.2 Stratification to estimate omissions

Having selected 280 samples for estimating commission as well as omission errors (Ch. 6.3.1) will provide useful results for both error types if the HRL population is large enough. However, some land cover types (e.g. imperviousness, wetland, water in our case) cover just a small fraction of a country (Table 14). If we draw random samples from the entire area, most of the samples would fall on e.g. non-sealed areas. Consequently, the representativeness of the samples belonging to the small target class would be low. Increasing the total number of samples so much, as it would be required by the small target class would yield extremely large number of samples. Therefore stratification should be used to increase the efficiency of sampling. The aim of stratification is to reduce the area to be sampled, i.e. to reduce the size of the stratum, where samples to estimate omissions are distributed. Selecting the right way of stratification should be done especially carefully. If CORINE Land Cover is used for stratification, classes with large area (arable land and/or forests usually) should not be included in the selected stratum. In this Guideline some stratification strategies are proposed to estimate omission error for each HRL. In special cases, a strategy different from the proposed ones might be more useful.

When interpreting the omission error it should be kept in mind that it refers only to the sampled strata and it might be overestimated regarding the entire area.

7.3.2.1 *HRL Imperviousness*

Built-up map will be verified and not the imperviousness layer (see. Ch. 6.2). Checking omission errors for the small built-up class (e.g. 1.5 %) would need very large sample size, because errors (omitted built-up areas) have to be searched for inside the extreme large (e.g. 9.5% of total area) non-built-up class. To provide reliable results with limited efforts the sampling area has to be reduced. Two solutions are discussed:

1: Using the subset of CLC Artificial surfaces layer (selected from CLC2006 database), which should include the majority of impervious surfaces. Samples for interpretation should be placed into locations that fall into the following CLC classes and where no (= 0) built-up has been classified:

- 1.1.1 = continuous urban fabric
- 1.1.2 = discontinuous urban fabric
- 1.2.1 = industrial, commercial areas
- 1.2.2 = road and rail network
- 1.2.3 = ports
- 1.2.4 = airports

Omitted impervious areas inside the usually large agriculture and forest classes will not be checked this way.

2: Compare national built-up map (and/or road map) with HRL built-up map. Select samples in locations where HRL indicates non-built-up, while national data include built-up.

Soil Sealing 2006 and Imperviousness 2009 layers are not relevant to use in the verification, because these are expected to be used in the production of HRL Imperviousness 2012 (i.e. these data are not independent).

7.3.2.2 *HRL Forest*

There are two layers to be verified: (1) tree cover map derived from the tree cover density layer (see Ch. 6.2) and (2) forest type map. These are treated separately.

a) HRL Tree Cover map

Checking omission errors for Tree cover product will be feasible for several countries by using samples drawn from the entire non-tree class. If the HRL Tree cover is small (<30%) stratification should be applied. To provide reliable results with limited efforts the sampling area should be reduced. Two solutions are discussed:

1: Using the subset of CLC Forest and seminatural classes, plus four of the CLC agriculture classes, which include a significant trees and forest component. Samples should be placed into locations that fall into the following CLC classes and where (1) "non-tree" has been classified in HRL tree cover map.

- 2.2.2 = fruit trees, orchards
- 2.2.3 = olives
- 2.4.3 = agriculture with significant amount of natural vegetation
- 2.4.4 = agroforestry[35]
- 3.1.1 = broadleaf forest
- 3.1.2 = coniferous forest
- 3.1.3 = mixed forest
- 3.2.1 = natural grassland
- 3.2.4 = transitional woodland, shrub
- 3.3.3 = sparsely vegetated areas
- 3.3.4 = burnt areas

The classes above should include the majority of tree cover. Omitted tree cover inside the usually large arable land class (2.1.1) will not be checked this way.

2: Compare national forest map with HRL tree cover map. Select samples from areas where HRL includes no-trees, while national forest map shows the presence of trees.

3: Compare JRC Forest 2006 map[36] (an independent layer, not used in producing HRL Forest) with HRL tree cover map. Select samples in areas where HRL includes no-trees, while in-situ data indicate the presence of trees.

b) HRL Forest Type map

Checking omission errors for forest type product will be feasible for several countries by using samples drawn from the entire non-forest class. If the HRL Forest cover is small (<30%) stratification should be applied. To provide reliable results with limited efforts the sampling area should be reduced. Three solutions are discussed:

1: Using the subset of CLC Forest and seminatural classes, plus two of the CLC agriculture classes, which include a significant forest component. Samples should be placed into locations that fall into the following CLC classes, where "non-forest" has been classified in HRL forest type map:

- 2.4.3 = agriculture with significant amount of natural vegetation
- 2.4.4 = agroforestry[37]
- 3.1.1 = broadleaf forest
- 3.1.2 = coniferous forest
- 3.1.3 = mixed forest

[35] Dehesa (ES), Montado (PT): frequent land cover in the Iberian Peninsula: including agriculture (usually pasture) with significant percentage of forest trees

[36] http://efdac.jrc.ec.europa.eu/. The JRC Forest map 2006 does not include forest type. The JRC Forest type map 2006 has a lower resolution.

[37] Dehesa (ES), Montado (PT): frequent land cover in the Iberian Peninsula: including agriculture (usually pasture) with significant percentage of forest trees

- 3.2.1 = natural grassland
- 3.2.4 = transitional woodland, shrub
- 3.3.3 = sparsely vegetated areas
- 3.3.4 = burnt areas

The classes above should include the majority of forest cover. Omitted forests inside the usually large arable land class (2.1.1) will not be checked this way.

2: Compare national forest map with HRL Forest Type map. Select samples from areas where HRL includes no-forest, while national forest map shows the presence of forests.

3: Compare JRC Forest 2006 map[38] (an independent layer, not used in producing HRL Forest) with HRL forest type map. Select samples in areas where HRL includes no-forest, while in-situ data indicate the presence of trees.

7.3.2.3 ***HRL Grassland***

Checking omission errors for grassland product will be feasible for some countries by using samples drawn from the entire non-grassland class. If the HRL Grassland is small (<30%) stratification should be applied. To provide reliable results with limited efforts the sampling area should be reduced. Two solutions are discussed:

1: Using the subset of CLC classes that have significant grassland component: some of the agriculture classes and natural grassland. Samples should be placed into locations that fall into the following CLC classes and where “no-grassland” has been classified in HRL grassland:

- 2.3.1 = pastures
- 2.4.2 = complex cultivation (mixed agriculture)
- 2.4.3 = agriculture with significant amount of natural vegetation
- 2.4.4 = agroforestry
- 3.2.1 = natural grassland

The classes above should include the majority of grass cover. Omitted grasslands inside the usually large arable land (2.1.1) and forest (3.1.x) classes will not be checked this way.

2: Compare national grassland map with HRL grassland map. Select samples in areas where no grassland (= 0) has been classified, while in-situ data indicate the presence of grassland.

7.3.2.4 ***HRL Wetland***

Estimating omission error for the small wetland class would need very large sample size, because errors (omitted wetland areas) have to be searched for inside the large non-wetland class. To provide reliable results with limited efforts the sampling area should be reduced. Two solutions are discussed:

1: Using the subset of CLC Wetland classes, which should include the majority of wetland areas. Samples should be placed into locations that fall into the following CLC classes and where “no-wetland” is classified in HRL wetland:

- 4.1.1 = inland wetland,
- 4.2.1 = salt marshes
- 4.2.2 = intertidal flats
- 4.2.3 = salines

[38] http://efdac.jrc.ec.europa.eu/. The JRC Forest map 2006 does not include forest type. The JRC Forest type map 2006 has a lower resolution.

The classes above should include the majority of wetlands (without peatland[39]). Omitted wetlands inside the usually large arable land (2.1.1) and forest (3.1.x) classes will not be checked this way.

2: Compare national wetland map with HRL wetland map. Select samples in areas where no wetland (= 0) has been classified, while in-situ data indicate the presence of wetland.

7.3.2.5 *HRL Water*

Estimating omission error for the small water class would need very large sample size, because errors (omitted water areas) have to be searched for inside the large non-water class. Masking with CLC classes is difficult, because small lakes can appear in most of the classes. As omission of larger (>25 ha) water bodies is not probable, the CLC Water bodies class (5.1.2) is not worth to use in stratification. Two solutions are discussed:

1: Compare national map of water bodies with HRL water map. Select samples from areas where no water (= 0) is classified in HRL Water, while in-situ shows the presence of water.

2: Compare a European map of water bodies[40] with HRL water map. Select samples in areas where HRL has shown no water, while in-situ data indicate the presence of water.

7.3.3 Sampling in the practice

Considering the complexity of sample selection for verification (with minimum patch size (Ch. 6.4) and stratification (Ch. 7.3.2) if needed) samples for verification of HRLs is offered be provided by ETC-SIA, if the country (or Service Provider) requests it and if the required in-situ data are available for the ETC-SIA.

Samples will include set of coordinate pairs in national system as described in Chs. 7.3.1 and 7.3.2. Samples for verification of omission and commission errors will be provided separately.

Random sampling will include for each HRL:

- 280 samples to estimate commission error for each HRL
- 280 samples to estimate omission error for each HRL
- Description of the stratification methodology (if applied)

In the course of sample selection a minimum patch size of 3x3 pixels around the sampled pixel shall be considered to avoid errors coming from the expected locational instability of satellite imagery. For verification of forest type, samples will be placed inside larger (than 3x3=9 pixels) patches (0.5 ha = 12.5 pixel).

8 Implementing verification

Both types of verification (look-and-feel and statistics-based) rely on photointerpretation of selected samples. See Chs. 7.2 and 7.3 about details. Elements of the proposed verification are summarised for each HRL in Tables 16-21:

[39] Peatland is defined in restricted way due to technological limitations.

[40] ECRINS = European Catchments and Rivers Network System (EEA): http://eea.eionet.europa.eu/Public/irc/eionet-circle/ecrins/library?l=/hydrography/v1/ecrlakmdb/_EN_1.0

8.1 HRL Imperviousness

Table 16: *Verification of HRL Imperviousness*

	Explanation	Remark
In-situ data		
Optimal data for verification	VHR colour ortho-imagery (satellite or aerial). Target resolution 0.5m, minimum requirement 1m resolution	
	Map of built-up areas, database of buildings and roads	As recent as possible
Additional data for verification	IMAGE2012, used for production	
	Google Earth or equivalent	Be careful with geometry; year of acquisition should be known
	Eurostat/LUCAS2012	LU/LC codes, field photographs
Verification		
Map layer is to be verified, derived from HR density layer (Ch. 6.1)	Only the built-up map can be verified.	
Decision criteria to derive the binary map layer	The pixel (20m x 20m) is built-up if sealing percentage exceeds the threshold of 30%.	This threshold will be applied to compute the built-up map layer from imperviousness [41].
Photointerpretation	The 30% threshold will be approximated as much as possible.	Because sealing densities cannot be consistently estimated in 20m x 20m pixels
Minimum mapping unit (MMU)	20m x 20m (i.e. single pixels).	
Commission error	Recommended number of random samples: 250	
Omission error	Recommended number of random samples: 250. Stratification is needed (see Ch. 7.3.2.1).	Because of low percentage of imperviousness, very large number of samples would be needed if the whole area is sampled.
Look-and-feel	Max. 10 thematic strata, 5-10 samples in each. 50-100 examined locations (Table 9); The best 5-8 examples are presented in report to highlight typical mistakes.	Evaluation in five grades. Examples in report (as screen-shots) are clearly documented.

[41] This threshold should be applied not only to compute the built-up map layer but also to determine the extent of the layer where samples will be selected.

8.2 HRL Forest / Tree-cover

Table 17: *Verification of HRL Tree density*

	Explanation	Remark
In-situ data		
Optimal data for verification	VHR colour / preferably colour infrared (CIR) ortho-imagery (satellite or aerial). Target resolution 0.5m, minimum requirement 1m resolution	Should be taken in summer (full canopy status)
	National forest map	Very useful, but not sufficient (see decision criteria)
	Ground survey (field plots)	
	Other maps on tree cover (urban, fruit tree plantations, olives)	
	JRC Forest 2006 map	Produced independently from GIO
	Eurostat/LUCAS2012	LU/LC codes, field photographs
Additional data for verification	IMAGE2012, used for production	
	Google Earth or equivalent	Be careful with geometry; acquisition date should be known
Verification		
Map layer is to be verified, derived from HR density layer (Ch. 6.1)	Only the tree cover map can be verified	
Decision criteria to derive the map layer	The pixel (20m x 20m) belongs to the tree cover map if tree cover density exceeds 30%. No forest definition (MMU, land use) is applied in deriving / verifying this product.	Due to the lack of any existing thresholds to compute tree cover map from HR tree density layer, 30% is recommended[42].
Photointerpretation	Due to practical reasons the majority rule will be applied in interpreting tree cover / non-tree-cover.	Because densities cannot be consistently estimated
Minimum mapping unit (MMU)	20m x 20m (i.e. single pixels).	
Commission error	Recommended number of random samples: 250	
Omission error	Recommended number of random samples: 250. If tree cover is larger than 30 %, samples are collected within HRL non-tree class. If tree cover is smaller than 30 %, stratification is need-	Because of low percentage of tree cover in some countries, very large number of samples would be needed if the whole area is sampled.

[42] CLC also applies 30% as the density limit for forest. The 10% (FAO) would be very difficult to apply consequently in photointerpretation. This threshold should be applied not only to compute the tree map layer but also to determine the extent of the layer where samples will be selected.

	ed (see Ch. 7.3.2.2).	
Look-and-feel	Max. 13 thematic strata, 5-10 samples in each. 65-130 examined locations (Table 10); The best 6-8 examples are presented in report to highlight typical mistakes	Evaluation in five grades. Examples in report (as screen-shots) are clearly documented.

8.3 HRL Forest / Forest type

Table 18: *Verification of HRL Forest type*

	Explanation	Remark
In-situ data		
Optimal data for verification	VHR colour ortho-imagery (satellite or aerial). Target resolution 0.5m, minimum requirement 1m resolution. CIR is obligatory.	Should be taken in summer (full canopy)
	Ground survey (field plots)	
	National forest map	Can be used if forest type is included. Problem if clear-cuts are not separated, but classified as forest
	Eurostat/LUCAS2012	LU/LC codes, field photographs
Additional data for verification	IMAGE2012, used for production	
	JRC Forest Type 2006 map	Produced independently from GIO
Verification		
The map layer produced by SP is directly applicable		
Minimum mapping unit (MMU)	0.5 ha (according to FAO forest definition)	
Photointerpretation	According to FAO forest definition (Ch. 4.2)	
Decision criteria	Due to practical reasons the majority rule will be applied in interpreting no-forest/broadleaf forest/coniferous forest.	
Commission error	Recommended number of random samples: 250	
Omission error	Recommended number of random samples: 250. If forest cover is larger than 30 %, samples are drawn from HRL non-forest layer. If forest cover is smaller than 30 %, stratification is needed (see Ch. 7.3.2.2).	Because of low percentage of forest cover in some countries, very large number of samples would be needed if the whole area is sampled.
Look-and-feel	Max. 13 thematic strata, 5-10 samples in each. 65-130 exam-	Evaluation in five grades. Examples in report (as

	ined locations (Table 10); The best 5-8 examples are presented in report to highlight typical mistakes	screen-shots) are clearly documented.

8.4 HRL Grassland

Table 19: *Verification of HRL Grassland*

	Explanation	Remark
In-situ data		
Optimal data for verification	National grassland inventory	Production date between 2006 and 2012
	IMAGE2006, 2009 and 2012 and AWiFS data (same as used for production)	Multi-year and multi-seasonal imagery is a must
Additional data for verification	VHR colour ortho-imagery (satellite or aerial), preferably CIR). Target resolution 0.5m, minimum requirement 1m resolution.	
	Eurostat/LUCAS2012	LU/LC codes, field photographs
Verification		
The map layer produced by SP is directly applicable		
Minimum mapping unit (MMU)	20m x 20m (i.e. single pixels).	Questionable, how the larger pixels (60m x 60m) of AWiFS will influence the quality.
Photointerpretation	According to grassland definition (see Ch. 4.3).	.
Decision criteria	Due to practical reasons the majority rule will be applied in interpreting grassland/non-grassland	
Commission error	Recommended number of random samples: 250	
Omission error	Recommended number of random samples: 250. If grassland cover is larger than 30% samples are drawn from HRL non-grassland. If grassland cover is smaller than 30% stratification is needed (see Ch. 7.3.2.3).	Because of low percentage of grassland cover in some countries, very large number of samples would be needed if the whole area is sampled.
Look-and-feel	Max. 9 thematic strata, 5-10 samples in each. 45-90 examined locations (Table 11); The best 5-8 examples are presented in report to highlight typical mistakes.	Evaluation in five grades. Examples in report (as screen-shots) are clearly documented.

8.5 HRL Wetland

Table 20: *Verification of HRL Wetland*

	Explanation	**Remark**
In-situ data		
Optimal data for verification	National wetland inventory	Production date between 2006 and 2012
	IMAGE2012 and AWiFS data (same used for production)	Multi-year and multi-season imagery is a must
Additional data for verification	VHR colour ortho-imagery (satellite or aerial), preferably CIR). Target resolution 0.5m, minimum requirement 1m resolution.	
	Eurostat/LUCAS2012	LU/LC codes, field photographs
Verification		
The map layer produced by SP is directly applicable		
Minimum mapping unit (MMU)	20m x 20m (i.e. single pixels).	It remains to be seen how the larger pixels (60m x 60m) of AWiFS will influence the quality.
Photointerpretation	According to wetland definition (see Ch. 4.4).	
Decision criteria	Due to practical reasons the majority rule will be applied in interpreting wetland/non-wetland.	
Commission error	Recommended number of random samples: 250	
Omission error	Recommended number of samples: 250. Stratification is needed (see Ch. 7.3.2.4).	Because of low percentage of wetlands, very large number of samples would be needed if the whole area is sampled.
Look-and-feel	Max. 11 thematic strata, 5-10 samples in each. 55-110 examined locations (Table 12); The best 5-8 examples are presented in report to highlight typical mistakes.	Evaluation in five grades. Examples in report (as screen-shots) are clearly documented.

8.6 HRL Water

Table 21: *Verification of HRL Permanent water bodies*

	Explanation	**Remark**
In-situ data		
Optimal data for verification	National inventory of water bodies / water layer of national topographic map	Production date preferably between 2006 and 2012
	European map of water bodies	ECRINS
	IMAGE2006, 2009 and 2012 and AWiFS data used for production	Multi-year and multi-season imagery is a must
Additional data for verification	VHR colour or b&w ortho-imagery (satellite or aerial). Target resolution 0.5m, minimum requirement 1m resolution.	
	Google Earth or equivalent	Be careful with geometry, year of acquisition should be known
	Eurostat/LUCAS2012	LU/LC codes, field photographs
Verification		
The map layer produced by SP is directly applicable		
Minimum mapping unit (MMU)	20m x 20m (i.e. single pixels).	It remains to be seen how the larger pixels (60m x 60m) of AWiFS will influence the quality.
Photointerpretation	According to definition of water bodies (see Ch. 4.5).	
Decision criteria (water body / non-water body)	Due to practical reasons the majority rule will be applied in interpreting water body/non-water body.	
Commission error	Recommended number of random samples: 250	Large water bodies (>25 ha) to be excluded
Omission error	Recommended number of samples: 250. Stratification is needed. (see Ch. 7.3.2.5).	Because of low percentage of water, very large number of samples would be needed if the whole area is sampled.
Look-and-feel	Max. 9 thematic strata, 5-10 samples in each. 45-90 examined locations (Table 12); The best 5-8 examples are presented in report to highlight typical mistakes	Evaluation in five grades. Examples in report (as screen-shots) are clearly documented.

9 Annex 1: HRL verification report template

The report shall be identical for all HRLs.

I. Administrative part

HRL	
Country (and region, if regions are analyzed separately)	
Institution carrying out the work	
General overview of data quality done by (name, position and e-mail)	
Look-and-feel analysis done by (name, position and e-mail)	
Statistical verification done by (name, position and e-mail)	
In situ data used. Replace *Data-1* with the proper type	*Data-1*
	Data-2
	….
	Data-n
Internal quality control by (name, position and e-mail)	
Date and place of writing the re-port	

II. General overview of data quality

Results of the general overview of data quality (descriptive text)	

III. Look-and-feel

Stratum	Name of the stratum (see Tables 9-13)	Number of locations analyzed	Results of the verification by strata (excellent, good, acceptable, insufficient, very poor)
1			
2			
3			
.			
.			
N			
Overall evaluation:			(excellent, good, acceptable, insufficient, very poor)
Any comment:			
If a stratum is not relevant in the country write “non-relevant” into column 3			
Feel free to introduce new strata if needed			

Provide examples with explanations

IV. Statistical verification

Stratification	*Describe the way of stratification or no stratification*
Stratification done by	*National institution or ETC-SIA*
Sampling done by	*National Institution or ETC-SIA*
Comment on sampling and stratification	
Number of random samples for commission error	*Total number of selected samples*
Number of valid samples for commission error	*Total number of valid (applicable) samples*
Accuracy, commissions (%)	
Comment on commissions	
Number of random samples for omission error	*Total number of selected samples*
Number of valid samples for commission error	*Total number of valid (applicable) samples*
Accuracy, omissions (%) with uncertainty	
Comment on omissions (%) with uncertainty	
Overall evaluation	*Remarks especially regarding the possibility / ways of en-hancing the product*

10 Annex 2: Estimating the uncertainty of accuracy values

By examining small databases we could check all the features (grid cells in case of raster HRL data) and calculate exact error values. Since it usually requires too much effort to check all the features, sampling is used to estimate the real errors in the database. A certain number of samples are needed to be able to provide representative estimation of the parameters of accuracy.

The binomial distribution as a statistical model is used to calculate the standard deviations for error values. Uncertainty of commission error calculations are calculated as:

$$UC_{commission} \approx \pm \frac{\sigma}{n} = \pm \frac{\sqrt{n \cdot E_{commission} \cdot (1 - E_{commission})}}{n}$$

Where n is the number of all samples distributed in the HRL class and $E_{commission}$ is the commission error value. The uncertainty has been expressed relative to the number of samples.

The uncertainty for omission error has to be calculated as:

$$UC_{omission(HRLclass)} = UC_{commission(non-HRLclass)} \cdot \frac{Area_{Total} - Area_{HRLclass}}{Area_{HRLclass}}$$

This calculation of uncertainty corresponds to a significance level of appr. 68,3%. Note, that the function of binomial distribution is not symmetrical (unlike normal distribution), the expression of the uncertainty with ± values is only an approximation.

Table 1. Illustration of the uncertainty calculated for various sample sizes and error probabilities

Number of samples (n)	Probability of error (p)	Calculated uncertainty
100	1%	± 0,99%
	15%	± 3,57%
	30%	± 4,58%
	50%	± 5,00%
500	1%	± 0,44%
	15%	± 1,60%
	30%	± 2,05%
	50%	± 2,24%
1 000	1%	± 0,31%
	15%	± 1,13%
	30%	± 1,45%
	50%	± 1,58%