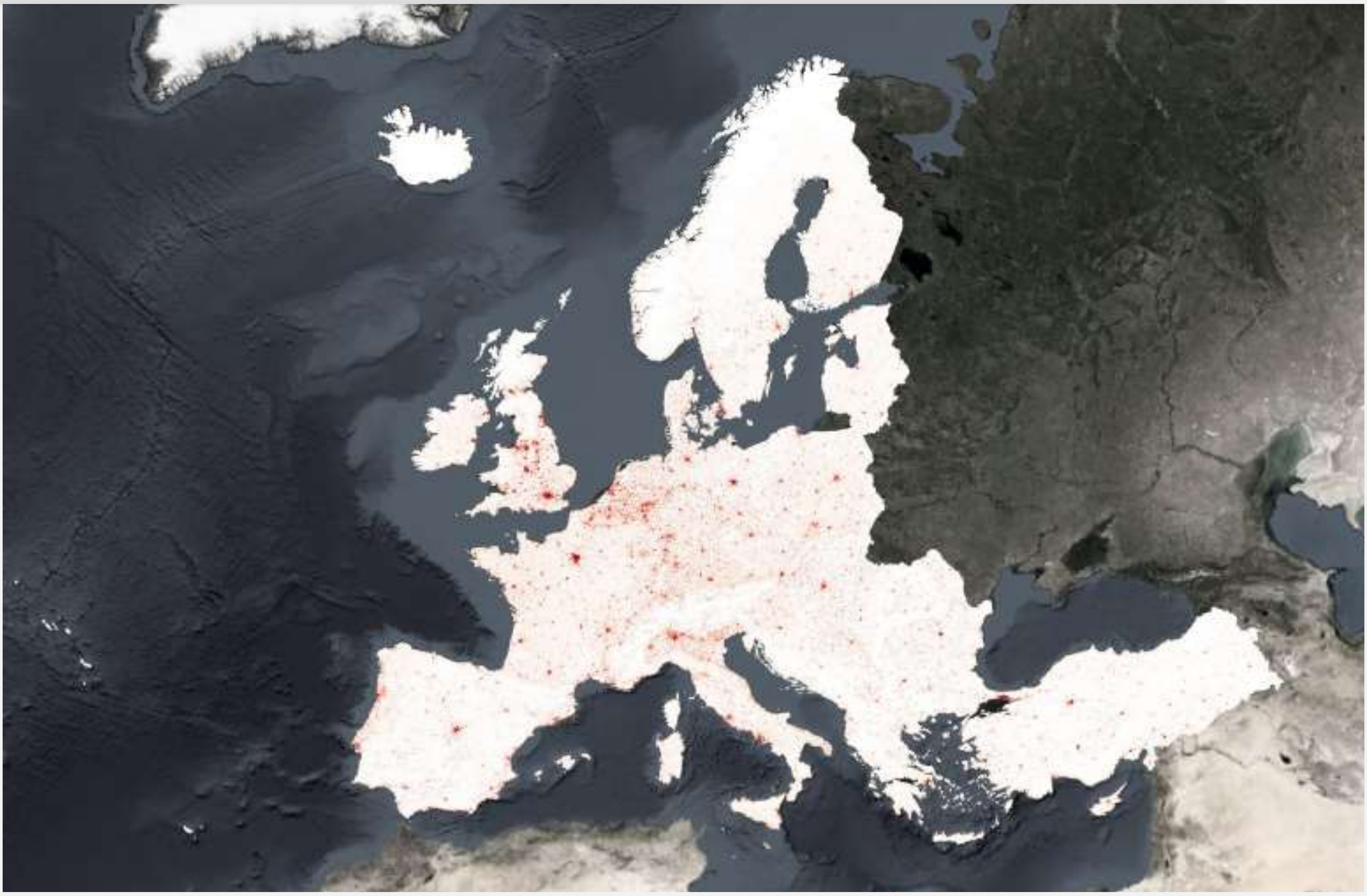
ETC-DI Report 2024/3

Analysis of usability of Imperviousness and CLC+ Backbone data for mapping sealed areas



Authors:
Gergely Maucha, Éva Kerékgyártó, Viktória Turos (Lechner Ltd.) Christophe Sannier (GAF), Jaroslav Dufek, Tomas Soukup (GISAT), Eva Ivits (EEA)

European Environment Agency
European Topic Centre
Data integration and digitalisation

Cover design: EEA
Cover image © ETC DI, Copernicus, ESRI, Maxar, Earthstar Geographics, and the GIS User Community
Layout: ETC DI / EEA

Version: 1

Publication Date 12/06/2024

ETC-DI coordinator: Umweltbundesamt GmbH (UBA) / Environmental Agency Austria (EAA)

ETC-DI partners: University of Malaga, Space4Environment Sarl, Universitat Autonoma de Barcelona, GISAT S.R.O., Epsilon-I, Sinergie, Lechner, Stiftelsen Norsk Institutt fof Luftforskning (NILU), Stichting Wageningen Research – Wageningen Environmental Research (WRWENR).

ISBN xxx-xx-xxxx-xxx-x
ISSN xxxx-xxxx
doi: xx.xxxx/xxxxx

European Topic Centre on
Data integration and digitalisation (ETC-DI)
https://www.eionet.europa.eu/etcs/etc-di

Contents

Acknowledgements

This study was supported by the Directorate-General for Regional and Urban Policy of the European Commission under Service Level Agreement 2021CE160AT130.

Summary

For EU policies to be efficiently planned, there is a need for a continental, harmonized, multitemporal and highly detailed indicator on soil sealing that allows the monitoring of the location and the degree of impacts.

The European Copernicus Land Monitoring Service has been producing datasets on imperviousness every 3 years since 2006, which are the only high-resolution datasets that enable European wide monitoring. However, after the 2015 reporting year, the input for the production of the imperviousness dataset was switched from mixed inputs to the European Sentinel satellites. While this led to an improvement in the spatial detail from 20 m to 10 m, the change in the input dataset also resulted in a break in the time series as the 2018 update was not comparable to the previous reference years.

Above issues have raised a discussion on statistical relevancy and appropriate use of areal statistics generated from CLMS products. To address these issues, we present a detailed analysis of similarities and differences in the specifications of sealing related CLMS datasets, as well as providing practical comparisons of sealed area estimations.

Based on the results of the analysis, a practical solution is presented in the form of a harmonized and bias-corrected continental soil sealing dataset for Europe for the entire observation period. This new dataset has been validated to be the best current dataset for monitoring imperviousness and soil sealing impacts as a direct input for European policies.

Background, scope and objectives

Artificially sealed areas are characterized by the substitution of the original (semi-)natural land cover or water surface with an artificial, often impervious cover. These artificial surfaces are usually maintained over long periods of time and such substitution is mostly irreversible. Due to the limited and largely unrenewable nature of land resources, amount, distribution, and dynamics of these substitution processes are crucial information for their sustainability assessment.

The number of spatial datasets documenting the extent and structure of artificially sealed areas is growing both within and beyond the Copernicus Land Monitoring Service (CLMS). New advances in quality, resolution, and availability of Earth observation data together with new processing approaches and (cloud-based) scalability of production workflows allow to create more quality and more detailed spatial data documenting the status and evolution of sealed surfaces more efficiently from local to global level.
On the other hand, based on our experience the actual amount and structure of detected sealed cover identified by different datasets varies due to several factors:

- Semantic content (class definitions are usually a mix of LC/LU parameters);
- Scale (raster resolution, Minimum Mapping Unit (MMU), Minimum Mapping Width (MMW));
- Surveying method (point sampling, visual mapping by photointerpretation, (semi-) automated image classification or feature extraction);
- Accuracy (random and possible systematic errors).

Sealing related datasets in CLMS portfolio

The Pan-European CLMS portfolio includes currently three main datasets serving information about artificial sealing:

CORINE Land Cover

CORINE Land Cover (CLC) is the oldest and most used dataset in CLMS portfolio, provides the longest time-series (2000-2006-2012-2018 and likely 2024) about the land cover of Europe. CLC is produced by national teams in the participating countries, mostly with traditional visual photo-interpretation methods.

Neither the 25ha MMU defined for status product, nor the 5ha MMU defined for change product do not allow the recognition of small sealed structures. In addition, class definitions of this categorical dataset mixes land cover and land use and other characteristics.

High Resolution Layer Imperviousness

The HRL Imperviousness (IMD) continuous datasets capture the spatial distribution of artificially sealed areas, including the level of sealing of the soil per area unit. The level of sealed soil (imperviousness degree 1-100%) is produced using an automatic algorithm based on calibrated Normalised Difference Vegetation Index (NDVI).

The latest upgrade to a 10m spatial resolution from 20m has unfortunately resulted in a break in the IMD based sealing statistics. While the former 20m resolution IMD time-series (2006-2009-2012-2015) have been successfully harmonized and have shown credible evolution in sealed cover, the upgraded IMD2018 data detect significantly more sealed structures than before, thus the amount of sealed cover is showing an unrealistic growth in the 2015-2018 period.

The update of HRL IMD is on-going, the IMD2021 status product is aimed to be comparable and consistent with the previous IMD2018 product. Imperviousness change data are produced in 20m spatial resolution.

CLC+ Backbone

The CLC+ Backbone Raster product provides a detailed pan-European wall to wall pure land cover reference layer in form of a 11-class 10 m spatial resolution raster product. The production was started with the reference year 2018, the update for year 2021 is on-going, further updates are planned in every 2 years. Currently no change product is specified in the CLC+ Backbone portfolio.

Scope and objectives

Various assessments and indicators require as critical input information about amount and distribution of sealed cover and the evolution of the sealed patterns. Thus, it is a burning question which dataset is to be used for a particular purpose in future analyses. Beside statistical comparison also structure and geometrical qualities and consistency need to be explored as reference sealing data are planned to be used not only for various land accounting activities, but also as essential input to several existing or planned spatial indicators (e.g. fragmentation, urban sprawl) where geometrical properties matter.

The scope of this task was to clarify on datasets fit-for-purpose for accounting and indicators activities in the future sealed areas mapping and analysis i.e.:

I. Assessment of similarities and differences of sealed features mapped by IMD vs CLC+ Backbone – beside statistical comparison also structural and geometrical qualities and time consistency;
II. Relation of both IMD/CLC+ Backbone to real sealing – statistical comparability with the results of point sampling-based surveys; and
III. Time-series consistency/validity of changes.

The task was divided into 10 activities:

1. Comparison of similarities and differences in dataset definitions;
2. Analysis and summary of available QC / QA assessments of both datasets;
3. Direct comparison of sealed content of Imperviousness degree (IMD) vs CLC+ Backbone data, scatter plot and difference maps;
4. Analysis of large deviations and comparison of geometrical properties by visual interpretation.
5. Creating an EU reference sealing level database by visual interpretation and statistical modelling. This kind of database will allow the comparison of EO based sealing level estimations with the point sampling (like LUCAS) based estimations in a direct and detailed way;
6. Comparison of sealed content of both datasets to ground truth represented by reference samples used for validation;
7. Harmonization of the HRL IMD time series for 2006, 2009, 2012, 2015 and 2018 on the highest possible spatial resolution but, as a minimum, on a 1km grid level;
8. Assessment of imperviousness vs CLC+ Backbone raster as input to the land take indicator, supporting the 8th EAP;
9. Prepare an outline of a scientific paper addressing the bias correction method for submission in 2024;
10. Detailed documentation of the work in a report describing different approaches, methodological elements and data sources used to produce harmonized sealing level (imperviousness) time-series data.

Link to previous work

Some of activities is focusing on the summary and extension of previous work performed under Copernicus service contract SC59032 Task 14:

- Subtask 1: Analysis of commonalities and synergies between CLC+ Backbone vs HRL2018 as well as RZ, CZ, N2K & UA performed[1];
- Subtask 4: Analysis of the threshold values that would allow comparability of the HRL 2015 and 2018 data were performed at three different aggregation level: 100m, 1km and 10km resolutions has been used for comparison[2];
- Subtask 9: “Pixel counting” vs. statistics discussion support provided by performing reference sealing level interpretation for Hungary as test area[3].

---

[1] Task 14 Subtask 1 Synergies between CLC+ Backbone and other CLMS products Final report v2

[2] Task 14 Subtask 4 Analysing the threshold values that would allow comparability of the HRL 2015 and 2018 data

[3] Task 14 Subtask 9 “Pixel counting” vs. statistics discussion support Final report

1. Comparison of similarities and differences in sealing related dataset definitions

Specification of Imperviousness data as well as soil sealing related part of specifications of CLC+ Backbone data is described in detail by highlighting commonalities and differences.

1.1 High Resolution Layer Imperviousness

Imperviousness presents the longest time-series in the portfolio of High Resolution land cover Layers (HRL). The production was started under the name of Soil Sealing 2006 and was updated in 2009, 2012 and 2015 in 20m resolution based on available (mainly IRS, Spot 5, Landsat) satellite imagery. The primary product was a 20m resolution Soil Sealing / Imperviousness Density (IMD) layer storing density values between 0-100%, representing the estimated percentage of the single raster cell covered by sealed surface. The 100m aggregated version was provided as an additional product.

Imperviousness change data were produced / published first only in 100m resolution. The imperviousness density change (IMC) layer was created as the difference of neighbouring status layers and presented the change of soil sealing density values in the range of 0-200, representing the difference of density values between -100% to +100%.

Introduction of classified changes

As the uncertainty of the sealing density values estimated for a 20m pixel is rather high, a derived Imperviousness Classified Change (IMCC) product was introduced by applying thresholds on IMC and additional rule-based GIS operations on previous IMD and IMC layers resulting following classes:

- 0: Unchanged areas (IMD = 0%)
- 1: New cover
- 2: Loss of cover
- 10: Unchanged areas (IMD>0 at both reference dates)
- 11: Increased imperviousness
- 12: Decreased imperviousness
- 254: Unclassifiable in any of parent status layers

Understanding main imperviousness products:

- **IMD:** Degree of Imperviousness status products estimate the ratio of the rectangular cell area covered by artificially sealed (impervious) surface between 0-100%. These products are created initially in 10m (previously 20m) resolution, aggregated in a second step to 100m;
- **IMC:** Degree of Imperviousness Change products estimate the change of the imperviousness degree between two statuses at corresponding reference years. These products are created initially still in 20m (10m foreseen for next update) resolution, aggregated in a second step to 100m;
- **IMCC:** Imperviousness Change Classified products are classified IMC products, presenting change features in 20m resolution only;
- **IBU:** Imperviousness Built Up product was introduced in 2018 production as additional derived product fully harmonized with the imperviousness component. No change product was provided during 2018 production;
- **SBU:** Share of Built up is a 100m resolution density layer aggregated from IBU. Values are showing the ratio of cell area (0-100%) covered by built up features.

Re-analysis of historical time-series

In frames of the Imperviousness 2015 update, the re-processing and the re-analysis of the 2006-2009-2012 historical layers was performed together with the production of 2015 status. The methodologies applied for both were the re-processing, and re-analysis of all existing density products, to assure properly calibrated HRL Imperviousness time-series, but did not constitute a complete re-production from the scratch.

The resulting 20m and 100m resolution IMD time-series for 2006-2009-2012-2015 years as well as the 20m and 100m resolution imperviousness density change (IMC) layers fulfil the criteria of entirely harmonized (accounting) time-series, namely that the difference of neighbouring pair of status layers corresponds to a change layer. Note, that although the accuracy of imperviousness data was significantly improved by the re-analysis, the IMC layers still contain technical changes (noise) beyond real changes. Additionally, the historical layers and consequently the change layers contain unclassifiable areas (coded with 254 value), because of cloud cover in satellite imagery.

Resolution upgrade introduced from year 2018

Like in case of all other HRLs, with the appearance of Sentinel-2 data in 2016, the spatial resolution was upgraded to 10m for 2018 production and beyond by keeping the 20m resolution for change data. Aggregated products are still derived from primary resolutions to 100m for IMD and IMC data.

Table 1.1.a: Overview of main HRL Imperviousness status and change products

Ref. year	**Status**		**Change**			
	Degree of Imperviousness	**Built up**	**Imperviousness change**		**Imperviousness change classified**	
2006	IMD2006 20m IMD2006 100m	-	IMC0609 20m IMC0609 100m	IMC0612 20m IMC0612 100m	IMCC0609 20m	IMCC0612 20m
2009	IMD2009 20m IMD2009 100m	-	IMC0912 20m IMC0912 100m		IMCC0912 20m	
2012	IMD2012 20m IMD2012 100m	-	IMC1215 20m IMC1215 100m		IMCC1215 20m	
2015	IMD2015 20m IMD2015 100m	-	IMC1518 20m IMC1518 100m		**IMCC1518 20m**	
2018	**IMD2018 10m** IMD2018 100m	IBU2018 10m SBU2018 100m				

Production basics

The production of IMD status layers may be simplified to two major steps:

I. Pre-classification: Separation of impervious / non-impervious surfaces based on the list of elements to be included[4] / excluded;
II. Calculation of degree of imperviousness values-based vegetation indices.

---

[4] Description of included land cover types is harmonized with the EAGLE data model’s terminology, which is formulated to be in line with INSPIRE data specification terminologies wherever possible. Description of included land cover types is harmonized with the EAGLE data model’s terminology, which is formulated to be in line with INSPIRE data specification terminologies wherever possible.

Separation of impervious / non-impervious surfaces

The list of elements to be included / excluded was slightly modified with the HRL2018 update.

Table 1.1.b: Elements included or excluded in the HRL2018 imperviousness products

Elements to be included in the HRL Imperviousness 2018	**Elements to be excluded in the HRL Imperviousness 2018**
• Housing areas (even with scattered houses) • Built-up traffic areas (airports, harbours, railway yards) • Non built-up traffic areas (airport runways, non built-up harbour areas, railway yards, parking lots) • Roads[5] • Railway tracks associated to other impervious surfaces (i.e. inside built-up area) • Industrial, commercial areas, factories, energy production and distribution facilities • Non built-up sealed surfaces, which are part of categories, such as e.g. allotment gardens, cemeteries, sport and recreation areas, camp sites, excluding green areas associated with them • Artificial grass-covered sport pitches • Green roofs • Construction sites with discernible evolving built-up structures. • Single (farm) houses (where possible to identify from satellite imagery) • Paved borders of water edges • Permanent greenhouses (covered through the year) • Solar panel parks	• Railway tracks not associated to other impervious surfaces (i.e. outside built-up area) • Construction sites without discernible evolving built-up structures • Non-permanent greenhouses (temporal plastic coverage) • Mines, quarries, peat extraction areas • Sand, sand pits • Dump sites • Natural, artificial and cultivated vegetated areas • Un-vegetated or sparsely vegetated areas • Un-vegetated agricultural fields, arable land • Vineyards, fruit plantations • Grass surfaces used for sports of any kind • Glaciers, snow, water

Black text: Specification unchanged

Red text: Specification changed, element was excluded in the specification of previous (2006-2015) IMD products

Blue text: Content was clarified for IMD2018 product, but the element remained in the list

Calculation of degree of imperviousness (IMD) values

The calculation of degree of imperviousness (or imperviousness density values is based on calibrated normalised difference vegetation index (NDVI). Non-impervious surfaces are characterized by definition with IMD = 0% value, while impervious areas may be characterized by IMD values 1-100%.

---

[5] Roads are captured with no requirement of creating un-interrupted linear features, which would require the use of additional road network data in the classification. This means that gaps in imperviousness values will continue to occur, in particular for smaller roads.

1.2 CLC+ Backbone

CLC+ Backbone is a spatially detailed, large scale, EO-based land cover inventory. The CLC+ Backbone Raster product provides a detailed pan-European wall to wall pure land cover reference layer in form of a 10 m spatial resolution raster product. For each pixel it shows the dominant land cover among the basic land cover classes.

The production was started with the reference year of 2018, based on Sentinel time series from July 2017 to June 2019 and auxiliary features. The update for year 2021 is on-going, further updates are planned in every 2 years.

Table 1.2.a: Thematic classes of CLC+ Backbone Raster product

Legend	CLC+ Backbone class	Comment
🟥	1: Sealed	
🟩	2: Woody needle leaved trees	
🟩	3: Woody Broadleaved deciduous trees	
🟩	4: Woody Broadleaved evergreen trees	
🟫	5: Low-growing woody plants	
🟨	6: Permanent herbaceous	
🟨	7: Periodically herbaceous	
🟪	8: Lichens and mosses	
⬜	9: Non and sparsely vegetated	
🟦	10: Water	Includes sea water buffer in 2018 product
🟦	11: Snow and ice	
🟦	253: Coastal seawater buffer	Appears only in 2021 raster product, formerly part of class 10.

Most of the class definitions for CLC+ Raster Product comprise a 50% area threshold to express that the dominant land cover should be assigned. While this is a plausible approach there are many situations where the land cover within a single 10m pixel comprises a spatial mix of different classes and no single class reaches an absolute majority of 50%. To address this issue a relative majority is used in most cases for the class assignment.

Sealed areas as specified by CLC+ Backbone

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

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

1.3 Comparison of dataset definitions concerning sealed surfaces

The definition of sealed surfaces is differing in few details when comparing the specifications of HRL Imperviousness and CLC+ Backbone.

Elements to be included to impervious / sealed areas

Table 1.3.a: Elements to be included in the HRL IMD 2018 and CLC+ Backbone 2018 'class 1: Sealed'

Elements to be included in the HRL Imperviousness 2018	**Elements to be included in the CLC+ Backbone 'class 1. Sealed' associated with HRL IMD2018 included features**
Housing areas (even with scattered houses)	Buildings
Built-up traffic areas (airports, harbours, railway yards)	Buildings, Specific structures and facilities, and open sealed surfaces
Non built-up traffic areas (airport runways, non built-up harbour areas, railway yards, parking lots)	Specific structures and facilities, and open sealed surfaces
Roads	Flat surfaces covered by any type of impervious material that is used for artificial surface pavements (e.g. asphalt, concrete, tarmacadam).
Railway tracks associated to other impervious surfaces (i.e. inside built-up area)	Railway tracks are also considered as part of this class since they typically comprise impervious structural elements and a highly compacted subsoil. – **Difference:** ALL railway tracks are to be included.
Industrial, commercial areas, factories, energy production and distribution facilities	Buildings, Specific structures and facilities, and open sealed surfaces
Non built-up sealed surfaces, which are part of categories, such as e.g. allotment gardens, cemeteries, sport and recreation areas, camp sites, excluding green areas associated with them	Specific structures and facilities, and open sealed surfaces
Artificial grass-covered sport pitches	Specific structures and facilities, and open sealed surfaces
Green roofs	Also vegetated rooftops are to be mapped under this class.
Construction sites with discernible evolving built-up structures.	Specific structures and facilities, and open sealed surfaces
Single (farm) houses (where possible to identify from satellite imagery)	Buildings
Paved borders of water edges	Flat surfaces covered by any type of impervious material that is used for artificial surface pavements (e.g. asphalt, concrete, tarmacadam).
Permanent greenhouses (covered through the year)	NOT listed in the definition, but mentioned in the User Manual on page 24 at Figure 19: *"Greenhouses should be classified as Sealed"* **Suggested to be included in the definition.**
Solar panel parks	**Difference:** NOT listed in the definition! Solar panel parks are often mapped in the dataset as 'class 6: Permanent herbaceous', sometimes 'class 1. Sealed' or 'class 9. Non- and sparsely-vegetated'.

Key differences in INCLUDED elements:

- The main difference is the mapping of solar panel parks. It is not mentioned in the CLC+ Backbone specifications. Although in practice in some location is still mapped as sealed, it is mapped more often as class 6: Permanent herbaceous;
- Greenhouses are not explicitly mentioned in CLC+ Backbone specifications, are only appearing at one sample screenshot in the User Manual (on page 24 at Figure 19), that "Greenhouses should be classified as Sealed";
- While according to CLC+ Backbone specifications all railways are to be included to sealed class, IMD includes only railway tracks associated to other impervious surfaces (i.e. inside built-up area).

Elements to be EXCLUDED from impervious / sealed areas

Table 1.3.b: Elements to be excluded from HRL IMD 2018 and CLC+ Backbone 'class 1: Sealed'

Elements to be excluded from HRL Imperviousness 2018	**Elements to be excluded from CLC+ Backbone 'class 1. Sealed' associated with HRL IMD2018 excluded features**
Railway tracks not associated to other impervious surfaces (i.e. outside built-up area)	**Difference:** ALL railway tracks are to be included.
Construction sites without discernible evolving built-up structures	non-sealed and semi-sealed artificial surfaces (e.g. nat. mat. displaced from original place, artificially consolidated, e.g. logistic and storage areas, festive squares, non-vegetated sport fields, grass pavers, permeable paving)
Non-permanent greenhouses (temporal plastic coverage)	NOT listed in the definition.
Mines, quarries, peat extraction areas	non-sealed and semi-sealed artificial surfaces (e.g. nat. mat. displaced from original place, artificially consolidated)
Sand, sand pits	NOT listed in the definition. Excluded elements contain only **artificial surfaces**, which are not considered to be included as sealed.
Dump sites	Waste materials (e.g. communal / industrial waste)
Natural, artificial and cultivated vegetated areas	NOT listed in the definition. Excluded elements contain only **artificial surfaces**, which are not considered to be included as sealed.
Un-vegetated or sparsely vegetated areas	NOT listed in the definition. Excluded elements contain only **artificial surfaces**, which are not considered to be included as sealed.
Un-vegetated agricultural fields, arable land	NOT listed in the definition. Excluded elements contain only **artificial surfaces**, which are not considered to be included as sealed.
Vineyards, fruit plantations	NOT listed in the definition. Excluded elements contain only **artificial surfaces**, which are not considered to be included as sealed.
Grass surfaces used for sports of any kind	grass pavers
Glaciers, snow, water	NOT listed in the definition. Excluded elements contain only **artificial surfaces**, which are not considered to be included as sealed.

Key differences in EXCLUDED elements:

- In CLC+ Backbone all railways are to be included with no exception, according to IMD specifications railway tracks not associated to other impervious surfaces (i.e. outside built-up area) are to be excluded from the dataset;
- While greenhouses are not explicitly mentioned in CLC+ Backbone specifications, according to IMD specifications non-permanent greenhouses (temporal plastic coverage) are to be excluded from impervious areas;
- Natural features are not explicitly listed in the excluded elements in CLC+ Backbone specifications only artificial features, which are not supposed to be mapped. On the other hand, inclusion rules defined for class 1: Sealed clarify the categories of features to be mapped, therefore natural elements are excluded by definition.

2. Analysis and summary of available QC / QA assessments

Copernicus LCLU products are the subject of a wide variety of QA/QC procedures. Terminology, methods and available QA/QC results in the context of HRL Imperviousness and CLC+ Backbone data were analysed, conclusions drawn for present work.

2.1 Terminology

The definition of many quality management terms used in the practice may be inconsistent. Often, the terms verification and validation are used even interchangeably.

Terminology of QA / QC processes used in this document

→ Verification is defined as QA/QC processes performed with corrective purpose during production phase, which includes:

1. Internal assessments of technical and thematic quality performed by a Service Provider after certain processing steps;
2. External assessment of technical quality and quick assessment of thematic accuracy by EEA or European Topic Centre (ETC) experts;
3. National "Post production verification" performed by national teams (NRC);

→ Validation of HRL products is defined as processes aiming to derive accuracy parameters characterizing final or semi-final products or after ending production or certain production steps:

4. Internal validation of HRL products of 50% or 100% readiness is performed by Service Providers;
5. External European scale validation of HRL products is performed by independent expert groups.

Terminology of QA / QC measures used in the context of land cover / land use datasets

- Technical conformity is defined as a conformance / non-conformance to pre-defined parameters or elements in the product specifications. Defined measures are grouped as completeness, logical consistency and positional accuracy. Many of technical quality checks may be automatized.
- Thematic quality quantifies the relationship of the product to" what was actually on the ground at that time", which is identified in the practice as a relationship of the product to a set of reference data. The characterization of the thematic quality is the main aim of HRL post production verification described in this document.
- Usability is referred often as "fitness-for-purpose" and analyses the applicability of a certain product considering a specific use case.

2.2 Summary of QA / QC measures

Basic principles of QA/QC methods and measures used in the context of HRL IMD and CLC+ Backbone are summarized below.

Technical conformity checks

Technical conformity checks are performed for all CLMS products centrally by OSS QC tool (introduced during the production HRL2018 data) provided by the EEA as part of the approval process. Most of basic technical conformity elements (e.g. correctness of CRS, code validity, file naming conventions, etc) are checked by the tool automatically.

Characterization of thematic quality

The main aim of thematic quality assessments (verification or validation) of HRL products is to identify and possibly quantify systematic errors in terms of misclassification or validity of other mapped parameters, like degree of imperviousness or tree cover density.

Four levels of processes and corresponding methodologies were identified and performed in previous exercises assessing thematic quality:

- General overview of data quality is aiming to provide a general feeling about the data quality and orient the more detailed verification actions. General overview includes a general visual impression and checking basic statistical parameters of HRL data as well as the physical comparison of HRL data to the best available in-situ data and analysis of major disagreements;
- Look and feel verification provides mostly qualitative results by checking pre-determined locations, where classification problems are expected;
- Statistical verification is aimed to provide quantitative accuracy parameters by applying randomly selected samples. Besides of a single overall accuracy value, statistical verification is able to provide estimations for user's and producer's accuracy (or for the corresponding commission and omission errors);
- Reference calibration is aimed to identify systematic calibration errors of land cover density values by gaining reference density values via visual interpretation of point samples.

Usability

Usability is referred often as "fitness-for-purpose" and analyses the applicability of a certain product considering a specific use case.

2.3 Summary of QA/QC results

Availability of QA/QC results in the context of HRL IMD and CLC+ Backbone is summarized in Table 2.3.a.

Table 2.3.a: Overview of external QA/QC checks in the context of HRL IMD and CLC+ Backbone data

QA/QC	ETC verifications	National verifications	European scale validations
Targets of QA/QC	Primary 20m/10m resolution data: IMD status and change 2015, 2018 CLC+ Backbone 2018	Primary 20m/10m resolution data: IMD status and change 2015, 2018 CLC+ Backbone 2018 V0.9 (not the final version)	Aggregated 100m resolution for IMD status and change 2006-2018 No results for CLC+ Backbone yet
Technical conformity check	Performed for HRL2015 only both for 20m and 100m data	Not performed	Yes
General overview	Yes	Yes	Not performed
Look & feel	Yes	Yes	Not performed
Statistical	Not performed	Was optional, results are available for some MS (Imperviousness status and change only)	Yes, performed on 100m resolution imperviousness status and change data
Reference calibration	Not performed	Was optional, early results are available for some MS (Imperviousness status only)	Yes, performed on 100m resolution imperviousness status data

2.3.1 European scale validations

External European validation results are available for HRL Imperviousness status and change data:

- HRL Imperviousness degree 2015 validation report provides validation results for products of re-analysed time-series of imperviousness status (IMD2006, IMD2009, IMD2012, IMD2015) and change (IMC0609, IMC0912 and IMC1215) data;
- HRL Imperviousness degree 2018 validation report provides validation results for products of re-analysed time-series of imperviousness status (IMD2018), share of built-up status (SBU2018) and imperviousness change IMC1518) data.

Methodological guideline for the European scale validation of CLC+ Backbone was prepared (not published yet), but **no European scale validation results are available for CLC+ Backbone**.

Validating imperviousness status data

In case of imperviousness status and change data the 100m resolution aggregated products were validated. The validation of the Imperviousness Degree Layers was done at two levels:

- *Reference calibration:* A scatterplot of the density values extracted from the sample units for both the reference and map data for each reference year was made with a view to assess the correlation between reference and map values and identify any systematic bias (slope and intercept of the regression line significantly different for 1 and 0 respectively);

- *Thematic accuracy:* A threshold was applied to the density values for reference and map data to produce binary attributes of built-up for both the reference and map data layers.

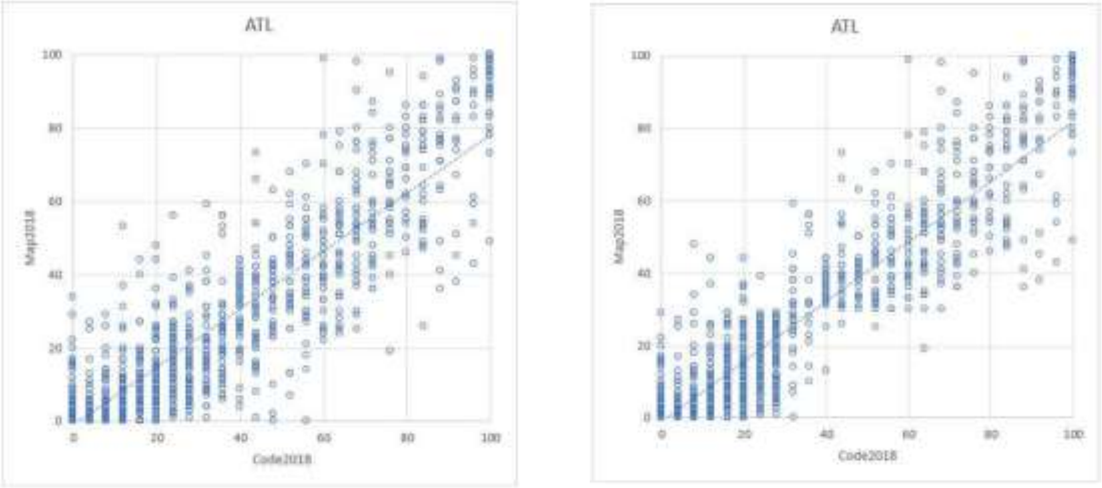
A double-blind approach was adopted as the initial process to guarantee the complete independence of the validation data from the map products. This may underestimate their accuracy where SSU points are uncertain. This was resolved by the plausibility approach for which the interpreter checks the map value to assess whether it can be considered correct or not, within the frame of accepted product specifications. Also, density values were adjusted based on experience of known uncertainties to allow a more realistic comparison.

Reference calibration

The validity of sealed area extent calculations by direct use of imperviousness density values was checked by a number of validation exercises estimating sealing level of aggregated 100x100m square sample units by visual interpretation of a point-grid.

Figure 2.3.1.a: Illustration of the estimation practice used during validation of imperviousness data by visual interpretation during HRL Imperviousness degree 2018 validation. An example of the PSU (Primary Sampling Unit) (yellow square) and the SSU (Secondary Sampling Unit) (green and red dots) for one sample location based on a LUCAS point.

The interpretation of each PSU at a particular point in time is based on the assessment (sealed or unsealed) of 5 x 5 grid of SSU points to derive a density value of soil sealing (imperviousness) at the PSU level. The values are thus represented by percentage values in 4% steps from 0 – 100%.

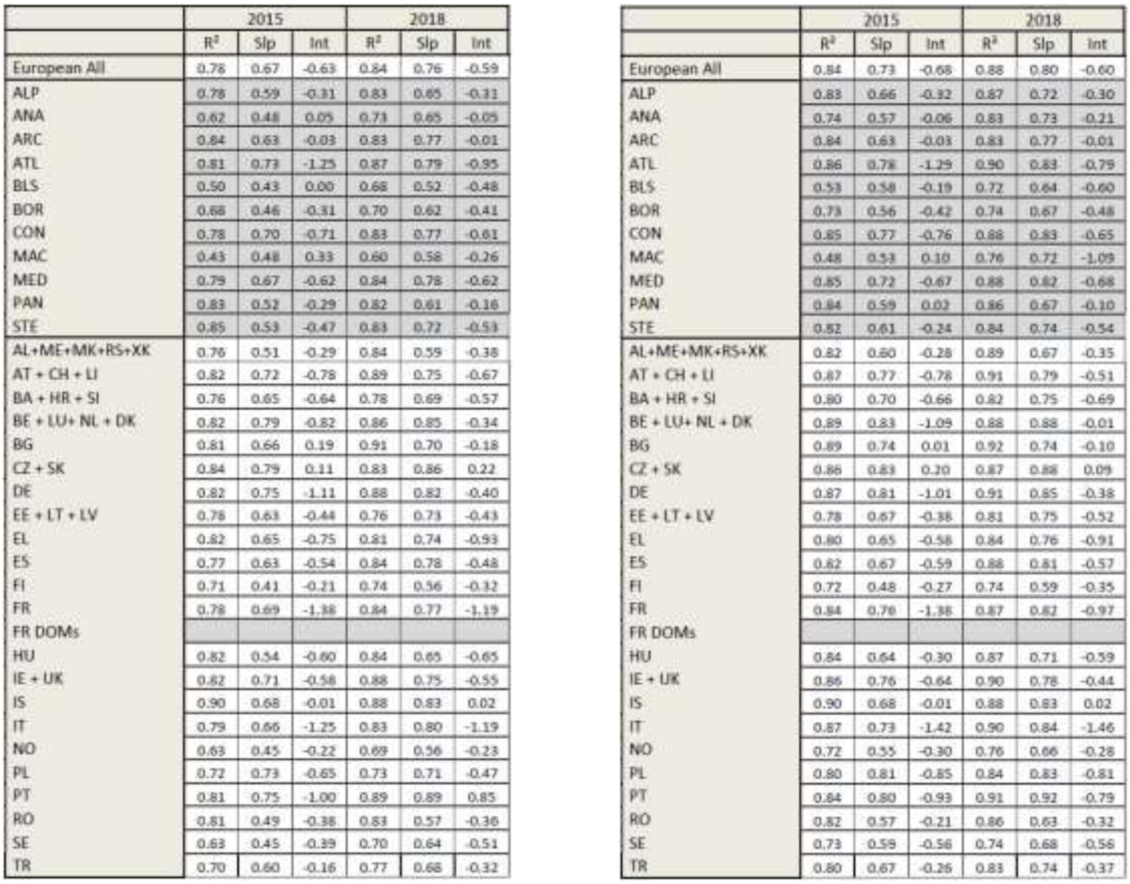


a) *blind interpretation*

b) *plausibility analysis*

Figure 2.3.1.b: Typical scatterplots of IMD2018 density validation (page 46 in IMD2018 validation report.)

Figure 2.3.1.b shows a typical result for imperviousness density value validation, significant systematic underestimation of sealing level values is indicated for both IMD2015 and IMD2018 data for all strata and countries (see Figure 2.3.1.c).



	2015			2018		
	R^2	Slp	Int	R^2	Slp	Int
European All	0.78	0.67	-0.63	0.84	0.76	-0.59
ALP	0.78	0.59	-0.31	0.83	0.65	-0.31
ANA	0.62	0.48	0.05	0.73	0.65	-0.05
ARC	0.84	0.63	-0.03	0.83	0.77	-0.01
ATL	0.81	0.73	-1.25	0.87	0.79	-0.95
BLS	0.50	0.43	0.00	0.68	0.52	-0.48
BOR	0.68	0.46	-0.31	0.70	0.62	-0.41
CON	0.78	0.70	-0.71	0.83	0.77	-0.61
MAC	0.43	0.48	0.33	0.60	0.58	-0.26
MED	0.79	0.67	-0.62	0.84	0.78	-0.62
PAN	0.83	0.52	-0.29	0.82	0.61	-0.16
STE	0.85	0.53	-0.47	0.83	0.72	-0.53
AL+ME+MK+RS+XK	0.76	0.51	-0.29	0.84	0.59	-0.38
AT + CH + LI	0.82	0.72	-0.78	0.89	0.75	-0.67
BA + HR + SI	0.76	0.65	-0.64	0.78	0.69	-0.57
BE + LU+ NL + DK	0.82	0.79	-0.82	0.86	0.85	-0.34
BG	0.81	0.66	0.19	0.91	0.70	-0.18
CZ + SK	0.84	0.79	0.11	0.83	0.86	0.22
DE	0.82	0.75	-1.11	0.88	0.82	-0.40
EE + LT + LV	0.78	0.63	-0.44	0.76	0.73	-0.43
EL	0.82	0.65	-0.75	0.81	0.74	-0.93
ES	0.77	0.63	-0.54	0.84	0.78	-0.48
FI	0.71	0.41	-0.21	0.74	0.56	-0.32
FR	0.78	0.69	-1.38	0.84	0.77	-1.19
FR DOMs						
HU	0.82	0.54	-0.60	0.84	0.65	-0.65
IE + UK	0.82	0.71	-0.58	0.88	0.75	-0.55
IS	0.90	0.68	-0.01	0.88	0.83	0.02
IT	0.79	0.66	-1.25	0.83	0.80	-1.19
NO	0.63	0.45	-0.22	0.69	0.56	-0.23
PL	0.72	0.73	-0.65	0.73	0.71	-0.47
PT	0.81	0.75	-1.00	0.89	0.89	0.85
RO	0.81	0.49	-0.38	0.83	0.57	-0.36
SE	0.63	0.45	-0.39	0.70	0.64	-0.51
TR	0.70	0.60	-0.16	0.77	0.68	-0.32

a) *blind interpretation*

	2015			2018		
	R^2	Slp	Int	R^2	Slp	Int
European All	0.84	0.73	-0.68	0.88	0.80	-0.60
ALP	0.83	0.66	-0.32	0.87	0.72	-0.30
ANA	0.74	0.57	-0.06	0.83	0.73	-0.21
ARC	0.84	0.63	-0.03	0.83	0.77	-0.01
ATL	0.86	0.78	-1.29	0.90	0.83	-0.79
BLS	0.53	0.58	-0.19	0.72	0.64	-0.60
BOR	0.73	0.56	-0.42	0.74	0.67	-0.48
CON	0.85	0.77	-0.76	0.88	0.83	-0.65
MAC	0.48	0.53	0.10	0.76	0.72	-1.09
MED	0.85	0.72	-0.67	0.88	0.82	-0.68
PAN	0.84	0.59	0.02	0.86	0.67	-0.10
STE	0.82	0.61	-0.24	0.84	0.74	-0.54
AL+ME+MK+RS+XK	0.82	0.60	-0.28	0.89	0.67	-0.35
AT + CH + LI	0.87	0.77	-0.78	0.91	0.79	-0.51
BA + HR + SI	0.80	0.70	-0.66	0.82	0.75	-0.69
BE + LU+ NL + DK	0.89	0.83	-1.09	0.88	0.88	-0.01
BG	0.89	0.74	0.01	0.92	0.74	-0.10
CZ + SK	0.86	0.83	0.20	0.87	0.88	0.09
DE	0.87	0.81	-1.01	0.91	0.85	-0.38
EE + LT + LV	0.78	0.67	-0.38	0.81	0.75	-0.52
EL	0.80	0.65	-0.58	0.84	0.76	-0.91
ES	0.82	0.67	-0.59	0.88	0.81	-0.57
FI	0.72	0.48	-0.27	0.74	0.59	-0.35
FR	0.84	0.70	-1.38	0.87	0.82	-0.97
FR DOMs						
HU	0.84	0.64	-0.30	0.87	0.71	-0.59
IE + UK	0.86	0.76	-0.64	0.90	0.78	-0.44
IS	0.90	0.68	-0.01	0.88	0.83	0.02
IT	0.87	0.73	-1.42	0.90	0.84	-1.46
NO	0.72	0.55	-0.30	0.76	0.66	-0.28
PL	0.80	0.81	-0.85	0.84	0.83	-0.81
PT	0.84	0.80	-0.93	0.91	0.92	-0.79
RO	0.82	0.57	-0.21	0.86	0.63	-0.32
SE	0.73	0.59	-0.56	0.74	0.68	-0.56
TR	0.80	0.67	-0.26	0.83	0.74	-0.37

b) *plausibility analysis*

Figure 2.3.1.c: Result of IMD2018 density validation (page 50-51 in IMD2018 validation report). The slope of the fitted linear (Slp) is smaller than 1 in all cases, indicating underestimation of sealed areas by imperviousness values.

The above results mean, that **imperviousness density values are in general underestimated,** and any calculations based on original definition of imperviousness density values results the **underestimation of sealed areas**.

Thematic accuracy

Thematic accuracy of imperviousness status data was performed on binary sealed / non-sealed maps created by applying a 30% threshold in IMD values. Density values lower than 30 % classified as 0 (non-sealed) and density values greater than or equal to 30 % classified as 1 sealed). The minimum acceptable thematic accuracy of 90 % should be reached for both omission and commission errors for class 1 (sealed).

IMD2018 data has shown the best quality in terms of user's and producer's accuracy compared to previous IMD data. Typically, **user's accuracy is higher** (corresponding to **lower commission**) in all cases (around 92% in case of blind interpretation, around 96% in case of plausibility analysis). **Producer's accuracy is lower** (corresponding to **higher omission**) in all cases (around 55% in case of blind interpretation for datasets IMD (2006-2015) 66.4% for IMD2018, while plausibility analysis shows around 85% producer's accuracy for datasets IMD (2006-2015), and 91.4% for IMD2018);

Table 2.3.1.d: Accuracy values for HRL Imperviousness status products by European scale validation

Product	Blind analysis		Plausibility analysis	
	Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy
IMD2006	54,25%	92,28%	84,57%	96,19%
IMD2009	54,66%	91,86%	84,41%	95,34%
IMD2012	55,38%	91,53%	84,75%	94,65%
IMD2015	55,10%	91,78%	84,50%	95,41%
IMD2018	66,40%	92,79%	91,37%	97,50%

Table 2.3.1.d provides summary results for pan-European area, detailed results for regional and country / country group strata are available in validation reports.

Validating imperviousness change data

The aggregated 100m resolution version of Degree of Imperviousness Change (IMC) products was selected for the European scale validation. For the purposes of European scale thematic accuracy assessment IMC products were re-classified to 3 thematic classes:

- 0: No change with zero imperviousness
- 1: Change in imperviousness
- 10: No change with imperviousness greater than zero

In case of first three periods (2006-2015) the reference data from the plausibility analysis were then processed to match the simplified change classes and a correspondence analysis performed. Both blind and plausibility results were published for the last period (201-2018). Table 2.3.1.e provides summary results for pan-European area, detailed results for regional and country / country group strata are available in validation reports.

Table 2.3.1.e: Accuracy values for HRL Imperviousness change products by European scale validation

Product	Blind analysis		Plausibility analysis	
	Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy
IMC 2006-2009			67,7%	56,9%
IMC 2009-2012			60,1%	52,7%
IMC 2012-2015			61,9%	61,9%
IMC 2015-2018	10,80%	37,48%	85,08%	51,01%

Available blind analysis results show extreme low producer's and user's accuracy for the 2015-2018 period, while plausibility analysis results show the highest producer's accuracy result (85,08%) for this period, while the user's accuracy value is similarly low for all periods. These results indicate that the **validity of identified changes is low**, but the lower user's accuracy indicate significant **omission of changes**.

2.3.2 National verifications

National verification of HRL2018 and CLC+ Backbone data was performed by many of EEA Member States under the Framework Service Contract EEA/IDM/R0/16/009/country. Specific contracts were established with countries for verifications.

Verification methodology

Primary, 10m (status) and 20m (changes) resolution raster data were verified during HR2018L verification process. Two steps of verification process were obligatory, one was optional:

1. General overview of data quality (obligatory);
2. Look & feel verification (obligatory);
3. Quantitative verification (optional, but supported).

National verification results were summarized by ETC/ULS in the frame of a Copernicus Service contract SC 58651 (Task 9).

General overview

The general overview included an overall visual impression and basic statistical parameters of HRL data as well as the physical comparison of HRL data to the best available in-situ data and analysis of major disagreements.

Look & feel verification

The look & feel verification has a corrective purpose, therefore the primary aim is to check potential error locations in an organized way. Identification of possible omission and commission sites can be helped by comparison with similar thematic datasets (e.g. national data).

- HRL look & feel assessment method is based on a controlled sampling by visiting locations of included/excluded elements listed in HRL specifications. If the country had available in-situ thematic data for the related stratum, the HRL dataset was compared to it. The arisen conflict areas were vectorized and sample locations were selected for verification, predominantly the largest polygons were examined. The evaluation was performed for the whole contradictory area represented by a polygon.
- If no appropriate in-situ data is available for direct comparison, sample locations are selected manually for the related stratum. The method of selection and the evaluation process are described in detail in the guidelines of verification.

Table 2.3.2.a: Accuracy values for HRL Imperviousness status products by European scale validation

NOTE	Meaning
excellent (5)	accuracy of the HRL is expected to reach almost 100%; practically no errors can be found in the verified area
good (4)	operator is confident that accuracy of the HRL is at least 85 %; only sporadic errors are encountered in the verified area
acceptable (3)	accuracy of the HRL is estimated to reach 85 % in most of the verified areas, minor errors can be detected in the verified area
insufficient (2)	accuracy of the HRL is not expected to reach the minimum 85%; several errors are encountered in the verified area
very poor (1)	operator is confident that accuracy of the HRL is bad and much below 85%; the verified area is wrongly mapped

On the location of the selected sample, HRL data was displayed on the top of VHR aerial or satellite imagery of appropriate acquisition date. Additional relevant in-situ data was suggested to be used as well for the visual evaluation of the sample. Each sample was evaluated by ordering one of the five-grade "notes" to the sample area (Table 2.3.2.a).

Verification results for IMD2018 status data

Imperviousness status data were verified on binary sealed / non-sealed maps created by applying a 30% threshold on density values of primary 10m resolution Degree of Imperviousness 2018 data. The results of Member States (MS) look & feel verifications included in the summary reports for HRL IMD2018 are presented in Table 2.3.2.b.

Table 2.3.2.b: Look and feel evaluation of omission and commission for HRL IMD2018 by countries

Country	**HRL IMD2018**	
	Look&feel	
	Omission	**Commission**
AT	4,4	4,6
BE	3	3
BG	2,8	1,3
CH	-	-
CZ	3,5	4,2
D	3	4
ES	3	3,5
FI	3,5	3
FR	3	3
GR	3,4	3,9
HR	-	-
HU	3,5	3,7
IE	2,9	4,1
IS	2,5	4
IT	2	2,5
LT	4	4
MT	3	4
NL	3	4
NO	3,7	4,4
PL	3	3
PT	3,7	2
RO	4	5
SE	2	2
SI	3,9	4,2
SK	3,3	3,6
UK	4	4,5
Average note:	**3,3**	**3,6**

Look&feel evaluation values	
Excellent	5
Good to Excellent	4,5
Good	4
Acceptable to Good	3,5
Acceptable	3
Insufficient to Acceptable	2,5
Insufficient	2
Very poor to Insufficient	1,5
Very poor	1

Key observations linked to general overview and look & feel analysis of IMD2018 data were:

- Countries reported slightly more omission errors than commission errors. In most countries, the IMD2018 performs slightly better in relation to the commission. However, the overall results are similar, being 3,3 for the omission and 3,6 for the commission.
- Omission errors are mainly present in a few strata for most countries. One of the main common issues is the stratum Single (farm) houses (where possible to identify from satellite imagery), which present many missing impervious areas. In some cases, this is explained by the high presence of trees/vegetation, but in others, the classification seems to be arbitrary.
- At Roads stratum, many missing parts and a lack of connectivity are present, even on highways. However, the stratum performed with acceptable results in most cases, but countries expressed many complaints in the evaluation reports. In most cases, the IMD2018 does not seem to be useful for highway and road network mapping, especially for statistics counting.
- Solar panel parks is a very problematic stratum and performed with insufficient or poor results in most countries. Some countries do not seem to agree with the technical guidelines when considering this stratum as sealed. The omission error is very high. Most solar panel parks were not included as an impervious cover or have many missing areas.
- Other omission issues are related to Artificial grass-covered sport pitches, which are often classified as non-impervious (the classification seems to be arbitrary); Paved borders of water edges, which present many missing parts and lack of continuity; and Greenhouses (covered through the year), which performed with acceptable results, but some countries reported many omission errors, being greenhouses totally omitted or with missing parts.
- Most common commission issues in the countries are related to with bare ground or lack of vegetation. The stratum Construction sites (without significant built-up structures) present high commission errors in most countries. This is also the case of Sand, sand pits and Dump sites but to a lesser extent. Mines, quarries and peat extraction areas performed better than these strata, almost reaching a good overall evaluation. Unvegetated or sparsely vegetated areas and Natural, artificial and cultivated vegetated areas present minor commission problems, especially the latter in areas where the vegetation density is low.

Quantitative verification

The quantitative verification was optional, nineteen countries provided data. The overall accuracies were mostly high, between 80% and 95% in most cases. User's and producer's accuracies were very variable between classes and countries. In general, non-impervious and highest impervious classes had higher accuracies. In most cases, they were higher than 80%.

Verification results for Imperviousness change (2015-2018) data

The Imperviousness Change Classified (IMCC) layer was the target of national verifications. Only the two main change classes were checked:

- IMCC class 1: new impervious cover
- IMCC class 2: loss of impervious cover.

The corresponding raster-covered areas were vectorized and the 100 largest contiguous areas were selected for the look & feel assessment. Verification results by countries are summarized in Table 2.3.2.c.

Table 2.3.2.c: Look and feel evaluation results for HRL IMCC 2015-2018 by countries

Country	HRL IMCC_1518			HRL IMCC_1518	
	Look & feel (evaluation)			Look & feel (number of samples)	
	New impervious cover	Loss of impervious cover	Overall	New impervious cover	Loss of impervious cover
AT	4,4	4,5	4,5	100	100
BE	-	-	-	-	-
BG	4,7	1,4	3,1	27	22
CH	-	-	-	-	-
CZ	3,0	2,0	2,5	100	100
D	3,0	1,0	2,0	102	100
ES	5,0	5,0	5,0	100	100
FI	2,6	1,1	1,8	-	-
FR	4,0	2,5	3,3	100	100
GR	2,1	1,1	1,6	100	100
HR	-	-	-	-	-
HU	3,8	1,4	2,6	103	112
IE	2,3	1,2	1,8	100	100
IS	2,5	1,0	1,8	-	-
IT	4,0	3,0	3,5	100	100
LT	4,0	3,0	3,5	100	100
MT	3,0	4,5	3,8	-	-
NL	2,0	-	2,0	115	-
NO	2,0	1,0	1,5	15	16
PL	5,0	1,0	3,0	100	100
PT	3,2	1,5	2,4	100	100
RO	5,0	5,0	5,0	14	6
SE	4,0	1,0	2,5	100	100
SI	3,3	1,2	2,3	100	95
SK	-	-	-	-	-
UK	3,6	2,1	2,9	100	100
Overall	**3,5**	**2,2**	**2,8**	**1 676**	**1 551**

Key observations linked to general overview and look & feel analysis of IMCC1518 data were:

- The overall impression is that many of the reported changes in the impervious cover are errors produced by unreal changes but a better classification in the year 2018.
- The HRL IMCC detects changes in existing impervious surfaces that were correctly mapped in IMD 2018. This is the main source of error in the case of roads and railways, which in most cases do not show changes at all. In addition, there are specific errors associated with a better mapping of other classes such as single (farms) houses, industrial and commercial areas, and other buildings in areas with a high presence of vegetation.
- These false changes (technical changes) are extremely problematic regarding the loss of impervious cover class, and the main reason of the poor accuracy of the HRL IMCC1518 in most countries. Most changes detected reflect the correction of areas that were commission errors in

2015 (e.g. quarries, dump sites, sandpits, bare rocks, water bodies, etc.). This is the main reason of the poor accuracy of the HRL IMCC1518 in most countries.
- Misclassification in 2018 also produced several errors in change detection, mainly related to changes in cropland (new crops, grown trees, land management, etc.), construction sites, temporary removal of vegetation cover, arid deposits, or clearing of forests or even phenological changes that produce commission error and are detected as new impervious surface. To a lesser extent, sporadic classification errors in natural vegetated areas also contribute to false changes related to both gain and loss of impervious cover.
- In addition, some errors were caused by an issue identified as intermittent shading, where shadows of tall trees or tall parts of buildings affect the IMD classification because of the shading at certain times of the day and/or certain seasons of the year, or at times when the features are shaded by clouds, can lead to errors of omission in 2018, which are detected as a loss of imperviousness.
- Although in some countries and strata the HRL IMCC 15-18 seems to work with acceptable results at a general level, these problems together with limitations in the geometry of the sampled change polygons (missing parts, the inclusion of areas without change, edges accuracy additional parts, etc.) make the HRL IMCC performance to be very inconsistent and very poor in several of the countries, especially in relation to the loss of impervious coverage.

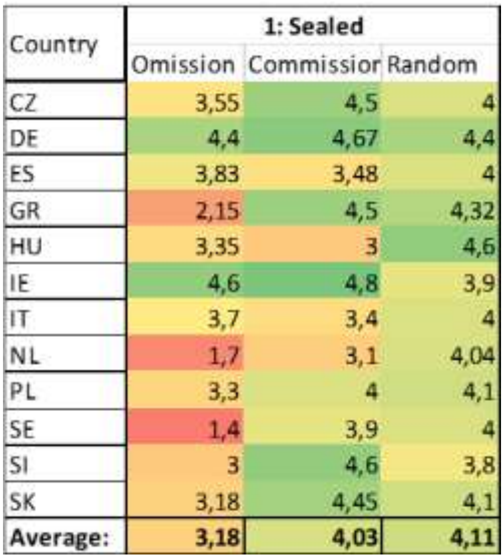
Verification results for CLC+ Backbone 2018 status data

National verification of CLC+ Backbone 2018 status data was performed in the frames of the Third specific contract under Framework Service Contract EEA/IDM/R0/16/009/xx: Copernicus Local Land monitoring services: NRCs LC (National Reference Centres for Land Cover) Copernicus supporting activities for the period 2017-2021.

Implementation of Task 1.2 was not compulsory, therefore only twelve countries delivered the CLC+ BB verification report: Czech Republic, Germany, Spain, Greece, Hungary, Ireland, Italy, The Netherlands, Poland, Slovenia, Sweden, and Slovakia.

It is important to mention, that the final (v1.1) version of CLC+ Backbone 2018 was not available in the time-frame of the verification, therefore the national verification was accomplished on CLC+ Backbone 2018 version 0.9.

Table 2.3.2.d: Look & feel evaluation results for CLC+ Backbone 2018 (v0.9) class 1 (sealed) by countries



Country	1: Sealed		
	Omission	Commissior	Random
CZ	3,55	4,5	4
DE	4,4	4,67	4,4
ES	3,83	3,48	4
GR	2,15	4,5	4,32
HU	3,35	3	4,6
IE	4,6	4,8	3,9
IT	3,7	3,4	4
NL	1,7	3,1	4,04
PL	3,3	4	4,1
SE	1,4	3,9	4
SI	3	4,6	3,8
SK	3,18	4,45	4,1
Average:	**3,18**	**4,03**	**4,11**

The verification task consisted of omission, commission, and random checks:

- Omission/commission checks: Specific CLC+ BB classes were compared to thematically comparable HRL2018 data and the largest vectorized conflict areas were subject to assessment.
- Random sampling: Random sampling was performed for all CLC+ Backbone classes. Extended number of random samples were checked for CLC+6 Backbone classes where comparable HRL data were not available for comparison.

National verification reports were collected by EEA and summarized in the ETC/DI CLC+ Backbone 2018 verification report. Results for CLC+ Backbone class 1 (Sealed) are presented in Tables d-e.

Table 2.3.2.e: Key observations linked to general overview and look & feel analysis CLC+ Backbone 2018 (v0.9) class 1 (sealed) by countries

CZ	Classification is overall **good**. Classification is worse or not complete in built up areas with higher percentage of sparse vegetation. Some samples are often classified as class 6. Industrial areas and halls (with white roofs) are sometimes not classified.
DE	Result is fairly **good**; the errors found belong in reality (according to imagery) to class 9. *Systematic errors:* Mining areas were often misclassified. According to the verification guidelines, mining areas should be classified in class 9: Non- and sparsely-vegetated. On the pictures you can see that only a small part is built on and the majority of the area is sand or just unsealed.
ES	**Good.** In general, its classification is good with minimal commission errors. In general, the sealing layer obtain a very good classification throughout the Spanish territory and it classify accurately.
GR	Overall classification was quite **good**. Serious commission errors except one weren't detected. The main errors detected deal with omission of sealed areas without any significant pattern. Major commission errors except one weren't detected either in any of the top 20 samples or in random samples.
HU	**Good to excellent**. Class is well captured. In discontinuous urban areas (villages and holiday settlements) some overestimation of sealing is still observed, but reduced compared to previous version. Source of this is mapping some gardens, unvegetated farm yards and dirt roads as sealed. Minor omissions (e.g. cemetery, small agro-industry) also found.
IE	Thematic accuracy is **overall good**. Few recurrent *omissions* occur in airports (e.g. Dublin and Cork) where some facilities such as runways and aircrafts stands are wrongly classified as 9 (non-and sparsely-vegetated). Likewise, parking plots and BMX/car trails, are partially omitted and often committed to class 9 as well as, industrial areas (e.g. Galway). *Commission* errors can be found in some areas of sand dunes (near Sligo) as well as in some storage areas within industrial, both should be classified in within class 9.
IT	**Good**. Sealed soil is well classified, main issues are the overestimation of green urban areas or low density urban areas, which are included in sealed class.
NL	**Good.** Often parts of greenhouses were not mapped as class 1 (irregular mapping). The ten selected samples for potential *omissions* were classified as class 3, 6, 7 or 9 in CLC+ BB. Only two of them were not considered as omissions. It was grasslands in urban areas that were mapped as sealed area in LGN. The other 8 samples were omissions. Those sealed areas were dam/construction to protect land from the sea, greenhouses or roads covered by trees. The ten selected samples for potential *commission* errors were all classified as permanent herbaceous in CLC+ SBB (class 6). Six of them were largely not commission errors. It is sealed areas in the harbour area of Rotterdam that in LGN were mapped as grass in built-up area (with some inclusions of permanent herbaceous). However, on the AP2018 the areas look like sealed areas (class 1). The other 4 examples are real commission errors as they are transhipment areas or areas not yet in use by any activities.
PL	Class 1 is generally correctly classified (**good**). Half of the analysed samples showed the missing parts and 6 unnecessary parts included. Average omission error is equal to 3.3 and commission to 4.0.
SE	**Good** result. Discrepancies are mainly due to difficulties to distinguish between sealed and non-sealed unvegetated surfaces.
SI	**Good.** Although smaller roads are better classified than in HRL_IMD, there are still some missing. Sometimes storage areas and railway tracks are wrongly included in class 1.
SK	**Good**. Generally, correct on random sample areas, including road surfaces, built-up areas and industrial sites according to reference images. Only small omissions were noticed and a few commissions on class 3, 6 or 9 areas within built-up or industrial areas.

3. Direct comparison of sealed content of Imperviousness versus CLC+ Backbone data

Artificially sealed areas are characterized by the substitution of the original (semi-)natural land cover or water surface with an artificial, often impervious cover. Two datasets of actual pan-European CLMS portfolio are mapping artificially sealed areas with very similar specifications: HRL Imperviousness and CLC+ Backbone by its first thematic class (sealed).

Similarities are:

- Both datasets are represented by a 10m resolution raster (from the reference year of 2018);
- Thematic specification of sealed areas is very similar (see details in chapter 1).

Main difference is:

- Primary product for HRL Imperviousness is density product, estimating the share of sealed area within a 10m raster cell as percentage value in the range of 0..100%, while CLC+ Backbone provides a flat map of land cover classes, for each raster cell it shows the dominant land cover among the basic land cover classes.

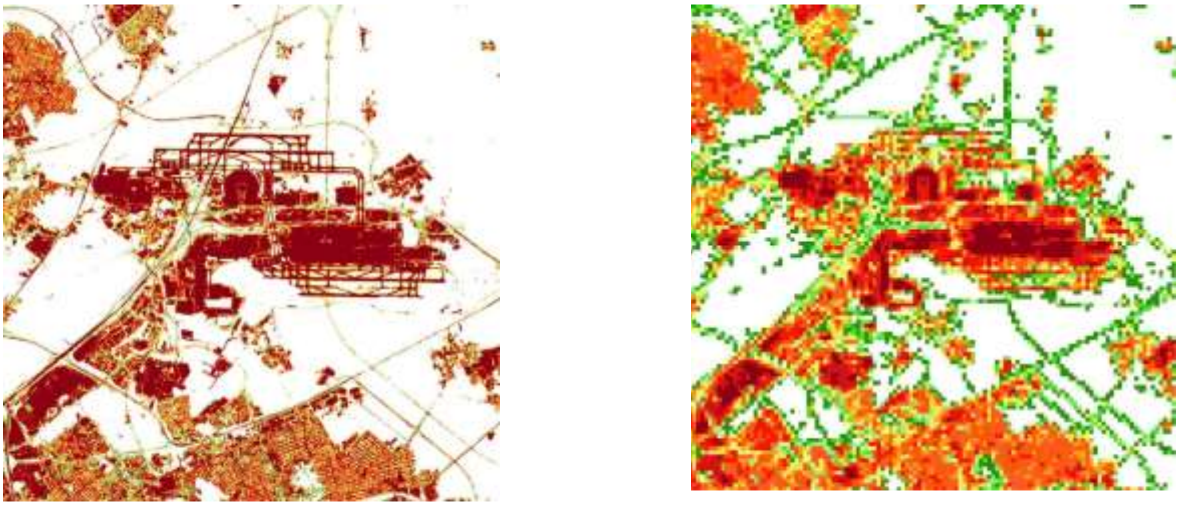
Slight differences in the specifications are realized in the practical information these datasets are providing in the context of artificial sealing.

3.1 Extracting sealing related information from raster datasets

Spatial indicators using sealing data as input are based on two main characteristics:

- Geometric structure of sealed features (fragmentation, connectivity);
- Extent of sealed areas within a certain spatial unit.

The geometric structure of sealed features delivers valuable input for certain environmental models and indicators. While the geometric information is significantly reduced by aggregation, sealed area statistics remain intact if creating density layers in higher aggregation levels.



a) Degree of imperviousness 2018 (10m resolution)

b) Degree of imperviousness 2018 (100m resolution)

Figure 3.1.a: Geometric structure captured by 10m primary and 100m aggregated IMD2018 products. Charles de Gaulle airport, Paris, France (x, y: 3777836,2899371 meters)

Sealed area calculation based on a binary map

CLC+ Backbone 2018 provides a “flat map” of land cover in 10m resolution. The CLC+ Backbone based calculation of aggregated sealed areas may be performed by the simple counting of all raster cells belonging to Class 1 (Sealed) within a statistical unit. The number of raster cells multiplied by the 100m^2 area of a single cell results the aggregated sealed area in square meters.

Sealed area calculation based on imperviousness degree values

Degree of Imperviousness (IMD) raster layers provide by definition “degree of imperviousness” values, meaning the estimated percentage of the share of sealed areas within a single raster cell. Imperviousness data-based calculation of the extent of sealed areas may be performed by two different ways:

a) Direct use of imperviousness density values;
b) Creating binary sealed maps by applying a threshold on initial imperviousness density values.

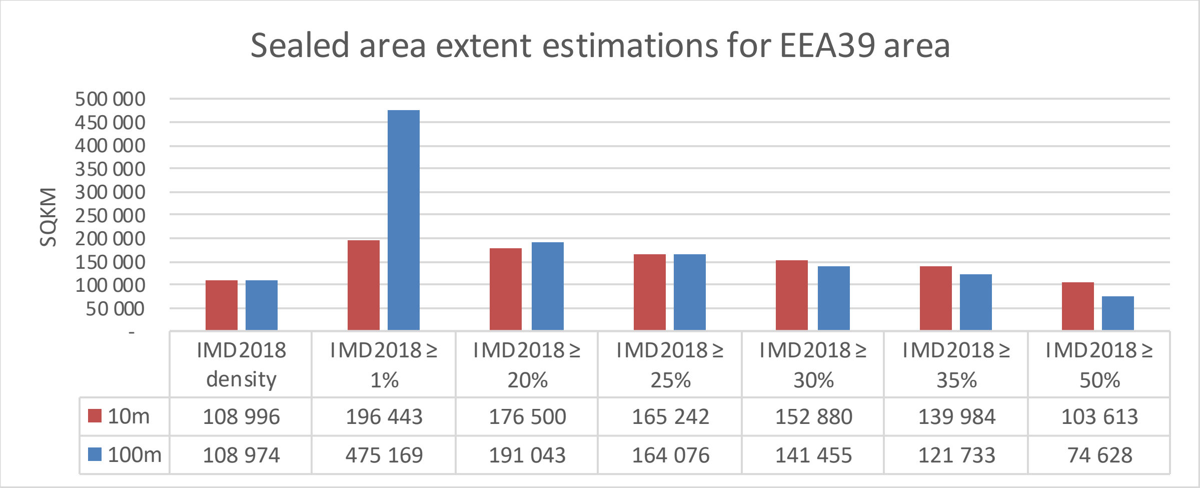
Direct use of imperviousness density values

Following the definition of imperviousness density values, the impervious (sealed) surface for the area represented by a single raster cell is calculated by the arithmetical product of the imperviousness value and the area of the cell (e.g. a single 10x10m cell with 1% imperviousness value includes 1m^2 = 0.0001 ha impervious surface). Sealed area extent for larger units may be calculated either by calculating the sealed area of each cell as described above and summarizing the sealed area values for the larger unit, or the mean sealing level calculated for the certain statistical unit is to be multiplied by the area of the unit.

Creating binary sealed maps by applying a threshold on initial imperviousness density values

By applying a certain threshold on imperviousness values, we can produce a binary map of sealed / non-sealed areas. It is important to mention, that the estimated extent of sealed areas is depending on at least two factors:

- The applied threshold values
- The spatial resolution of the density product



	IMD2018 density	IMD2018 ≥ 1%	IMD2018 ≥ 20%	IMD2018 ≥ 25%	IMD2018 ≥ 30%	IMD2018 ≥ 35%	IMD2018 ≥ 50%
🟥 10m	108 996	196 443	176 500	165 242	152 880	139 984	103 613
🟦 100m	108 974	475 169	191 043	164 076	141 455	121 733	74 628

Figure 3.1.b: Sealed area extent calculated for EEA39 by using IMD values directly applying various thresholds. Red columns show calculations based on primary 10m imperviousness density (IMD) product, while blue columns show calculations based on aggregated 100m IMD product.

While the aggregation to lower spatial resolution does not cause the loss of statistical information (only the loss of information concerning geometric structure), the application of a threshold on aggregated density layer means a spatial generalization and causes significant differences in sealed area statistics.

3.2 Methodology for the comparison of sealed content

In the frames of the Copernicus Service Contract SC59032 (Task14) various calculation methods were tested on HRL Imperviousness data, which may ensure the comparability between:

- Imperviousness data of different primary resolutions (i.e. 20m for IMD2015, 10m for IMD2018);
- HRL Imperviousness 2018 and CLC+ Backbone 2018 (both 10m resolution);
- HRL Imperviousness and point sampling based sealed area estimations.

Methodology for creating aggregated sealing data

The geometry of sealed features represented by IMD2018 and CLC+ Backbone 2018 data shows obvious similarities, especially when compared to lower resolution IMD2015 data. On the other hand the direct comparisons of primary 10m resolution data would cause extreme scattering, due to uncertainties caused by possible geometric shifts and uncertainties of spectral indices and classification.

To gain a better overview of the information content, thematic classes were extracted and aggregated into larger spatial units. Comparisons were performed at three different aggregation levels, namely 100m, 1km and 10km resolution LAEA statistical grids.

Main steps of the data preparation were:

1. Primary 10m resolution data were aggregated first to 100m resolution density layers, as 100m resolution allowed to perform pan-European calculations in a single process,
    a) The aggregated 100m resolution version of imperviousness degree data is available as part of standard CLMS product portfolio. These products were created by aggregating 10m / 20m resolution imperviousness density data to 100m imperviousness density data by calculation mean density value for aggregated 100m cells.
    b) Primary 10m / 20m resolution imperviousness data were converted to binary sealed / non-sealed maps by applying various thresholds. These layers were aggregated to 100m resolution density layers in the next step and used for further calculations.
    c) Class 1 (sealed) of the 10m resolution CLC+ Backbone data were re-classified to a 10m resolution binary (e.g. “sealed / non-sealed”) layer and aggregated to 100m resolution density layer;
2. Further aggregation of 100m resolution sealing density layers to lower resolutions (i.e. 1km or 10km statistical grid) by calculating mean density values expressed in percentages in the range of 0-100% partial sealed cover.

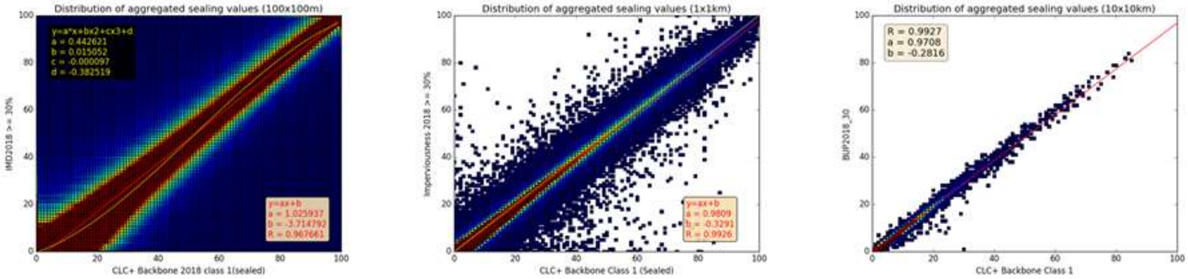
Determining optimal aggregation level for comparisons

The comparability between IMD and CLC+ Backbone was tested as cell-by cell comparison at various aggregation levels:

- The comparison of 100m x 100m resolution aggregated data has shown relatively high correlation, but large scattering;
- The comparison of 10km x 10km resolution aggregated data has shown very high correlation and low scattering. Additionally, 10km x 10km rectangle units are too large for proper examination of sealing data in the entire range of sealing levels, no aggregated density values

over 85% was found as no large contiguous sealed areas exist in Europe, which would cover a total area of a 10x10km rectangle.

- The comparison of 1km x 1km resolution aggregated data have shown high correlation and moderate scattering, this resolution was chosen as a good compromise.



a) *100m x 100m* b) *1km x 1km* c) *10km x 10km*

Figure 3.2.a: Aggregated sealing density values based on CLC+ Backbone 2018 and IMD 2018 (IMD ≥ 30%) compared at various resolutions

Determining optimal calculation methods for comparisons

The definition of the Degree of Imperviousness values would provide the ultimate solution for sealed area calculations, but many validation results indicate significant underestimation of sealed content, when applying the direct calculations based on IMD values. Still, this calculation was tested among others.

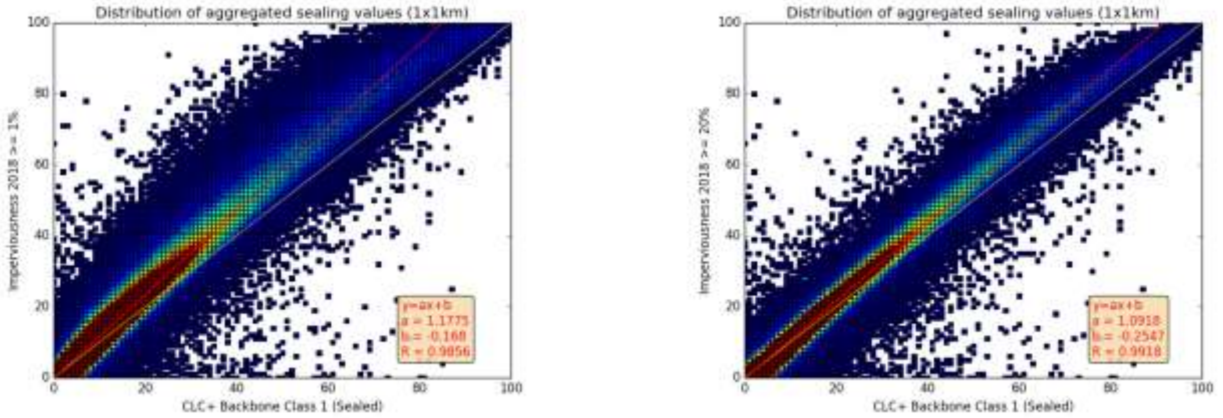
The threshold at 30% imperviousness is a commonly applied value for many practical applications when creation binary sealing maps from imperviousness degrees, therefore this value and values close to 30% were selected for first tests, other consideration were added later.

Finally, following calculation methods were tested:

- IMD: Direct calculation based on degree of imperviousness value;
- IMD ≥ 1%: All elements included to impervious areas by definition;
- IMD ≥ 20%: Commonly applied value - 10%;
- IMD ≥ 25%: Commonly applied value - 5%;
- IMD ≥ 30%: Commonly applied value for practical applications;
- IMD ≥ 35%: Commonly applied value + 5%;
- IMD ≥ 50%: Corresponding to theoretical majority rule (initial assumption for the comparability with CLC+ Backbone).

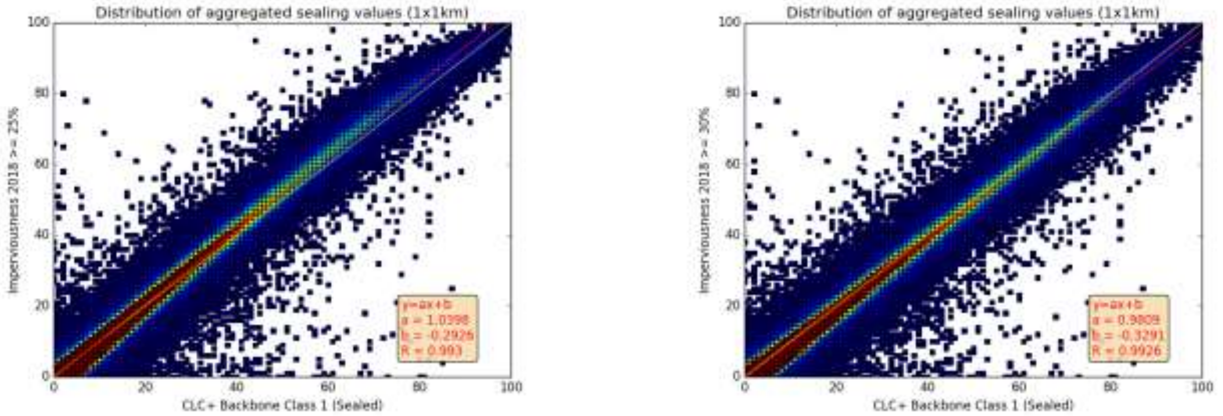
By comparing scatterplots of aggregated sealing densities between CLC+ Backbone and 10m resolution IMD2018 data following results were gained (Figure 3.2.b):

- The use of direct calculation based on degree of imperviousness 2018 values (f) has resulted significant underestimation of sealed areas compared to CLC+ Backbone sealed class;
- The use of all impervious cells (IMD2018≥1%; a) resulted significant overestimation of sealed areas compared to CLC+ Backbone sealed class;
- Best correlation values and close to linear relationships were found for IMD2018≥25% (c) and IMD2018 ≥ 30% (d).
- The use of thresholds over 30% including the majority rule (IMD2018≥50%) has resulted increasing underestimation of sealed areas compared to CLC+ Backbone sealed class.



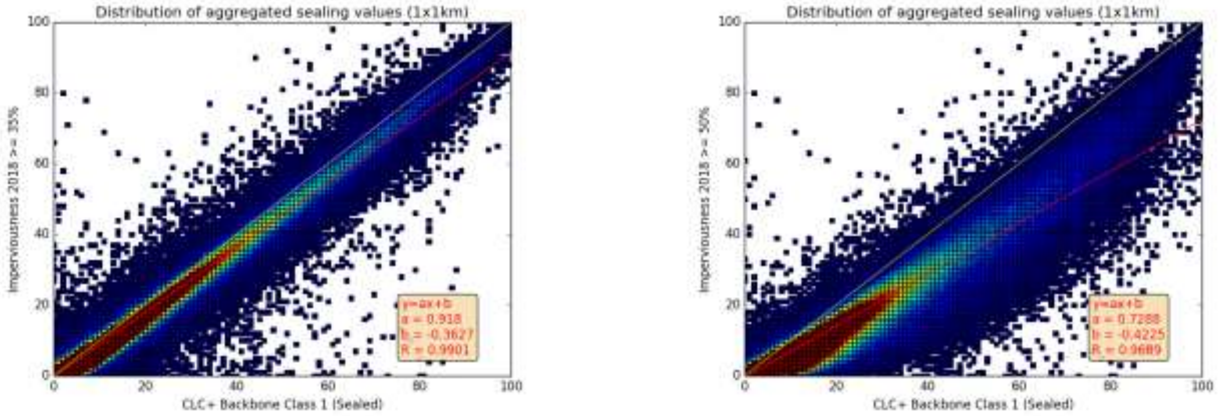
a) Backbone class 1 (Sealed) compared to HRL IMD2018 > 1%

b) Backbone class 1 (Sealed) compared to HRL IMD2018 > 20%



c) Backbone class 1 (Sealed) compared to HRL IMD2018 > 25%

d) Backbone class 1 (Sealed) compared to HRL IMD2018 > 30%



e) Backbone class 1 (Sealed) compared to HRL IMD2018 > 35%

f) Backbone class 1 (Sealed) compared to HRL IMD2018 (direct sealing level based calculation)

Figure 3.2.b: Comparison of aggregated sealing densities by 1km resolution sealing density layers for EEA39 area. Red line indicates fitted linear function, yellow line indicates main diagonal.

Based on the previous results, the threshold value of 30% - as the most commonly used value in the practice - was fixed for IMD2018 for further exploration. As a next step, various threshold values were explored for IMD2015 in order to find the best fit to a binary map created by IMD2018≥30% criteria. Based on comparisons of 1km x 1km as well as 10km x 10km aggregated data best fit was gained when the threshold of IMD2015 ≥ 25% was applied and compared to IMD2018 ≥ 30% data.

4. Analysis of large deviations and comparison of geometrical properties

Larger deviations of sealing level estimations were identified via difference maps and scatterplots were examined by visual interpretation.

4.1 Checking deviations between IMD2018 and CLC+ Backbone 2018

Difference maps were created based on differences of 1km x 1km aggregated sealing density values derived from various datasets to visualize regional and local effects.

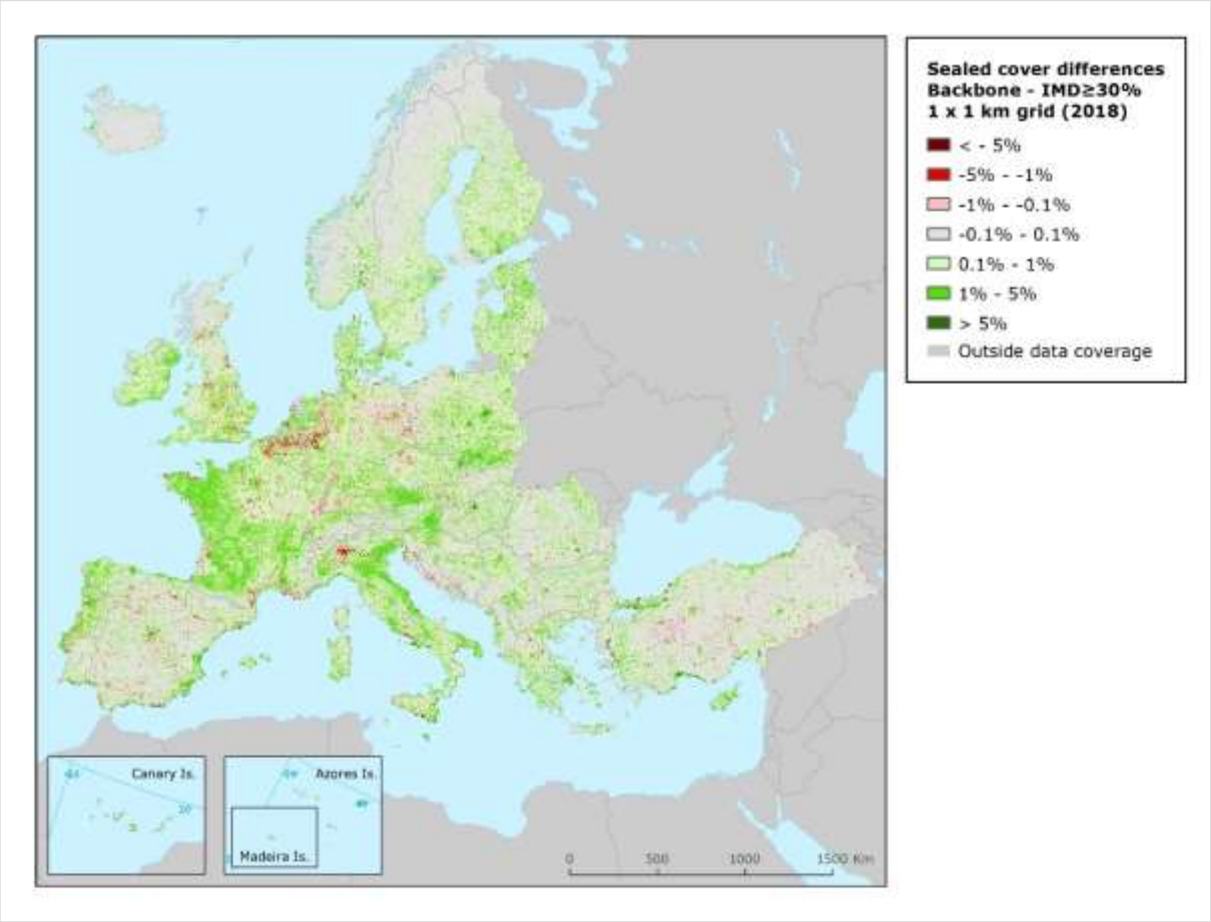


Figure 4.1.a: Local and regional differences of aggregated sealing estimated by CLC+ Backbone vs IMD2018 ≥ 30%. Green: More sealing is estimated by Backbone. Red: More sealing is estimated by IMD2018 ≥ 30%.

The overview of the difference map indicates slight surplus of sealed area estimated by CLC+ Backbone (green colours) for larger contiguous areas, while IMD2018 based sealing has shown surplus at specific locations. Locations of largest differences were checked with the help of VHR Bing or Google Earth historical imagery, typical examples are shown below.

4.1.1 Surplus of sealing by CLC+ Backbone 2018 (green cover on Figure 4.1.a overview map)

The surplus sealed areas estimated by CLC+ Backbone compared IMD2018 appears as green colour on the overview difference map. Typical examples are presented in following pages.

Figure 4.1.1.a: More detailed appearance of road network in CLC+ Backbone 2018.
Red: CLC+ Backbone 2018; Blue: IMD2018 ≥ 30%; Yellow: IMD2018: 1-30%
Purple: CLC+ Backbone 2018 & IMD2018 ≥ 30%; Orange: CLC+ Backbone 2018 & IMD2018: 1-30%
Türkiye, x,y= 6044976,2191979 (EPSG:3035)

Figure 4.1.1.b: More detailed appearance of secondary road network in CLC+ Backbone 2018.
Red: CLC+ Backbone 2018; Blue: IMD2018 ≥ 30%; Yellow: IMD2018: 1-30%
Purple: CLC+ Backbone 2018 & IMD2018 ≥ 30%; Orange: CLC+ Backbone 2018 & IMD2018: 1-30%
Italy, x,y= 4523451,2506710 (EPSG:3035)

Figure 4.1.1.c: More detailed appearance of secondary road network in CLC+ Backbone 2018.
Red: CLC+ Backbone 2018; Blue: IMD2018 ≥ 30%; Yellow: IMD2018: 1-30%
Purple: CLC+ Backbone 2018 & IMD2018 ≥ 30%; Orange: CLC+ Backbone 2018 & IMD2018: 1-30%
France, x,y= 3500754,2818803 (EPSG:3035)

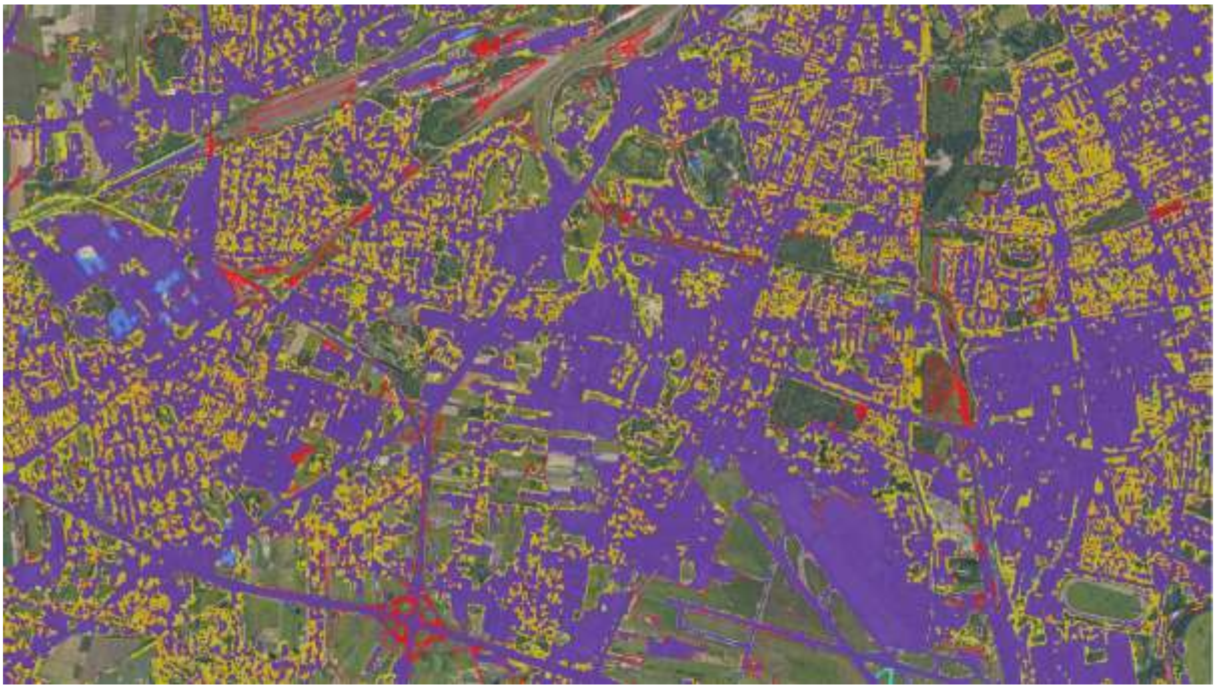


Figure 4.1.1.d: More detailed appearance of infrastructure in CLC+ Backbone 2018.
Red: CLC+ Backbone 2018; Blue: IMD2018 ≥ 30%; Yellow: IMD2018: 1-30%
Purple: CLC+ Backbone 2018 & IMD2018 ≥ 30%; Orange: CLC+ Backbone 2018 & IMD2018: 1-30%
Poland, x,y= 5068031,3284625 (EPSG:3035)

4.1.2 Surplus of sealing by IMD 2018 (red cover on Figure 4.1.a overview map)

The surplus sealed areas estimated by IMD2018 compared CLC+ Backbone appears as red colour on the overview difference map. Typical examples are presented in following pages.

Figure 4.1.2.a: Solar panel parks are not appearing in CLC+ Backbone 2018, only in IMD2018.
Red: CLC+ Backbone 2018; Blue: IMD2018 ≥ 30%; Yellow: IMD2018: 1-30%
Purple: CLC+ Backbone 2018 & IMD2018 ≥ 30%; Orange: CLC+ Backbone 2018 & IMD2018: 1-30%
Germany, x,y= 4594090,3164103 (EPSG:3035)

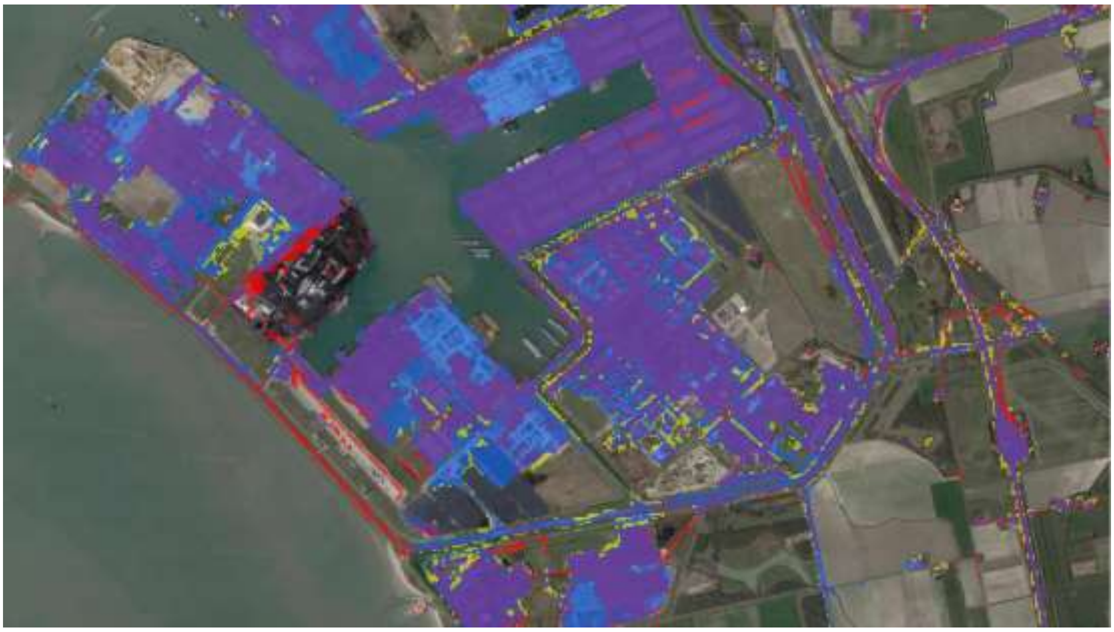


Figure 4.1.2.b: Commission error in IMD2018, overestimation of sealed areas.
Red: CLC+ Backbone 2018; Blue: IMD2018 ≥ 30%; Yellow: IMD2018: 1-30%
Purple: CLC+ Backbone 2018 & IMD2018 ≥ 30%; Orange: CLC+ Backbone 2018 & IMD2018: 1-30%
Netherland, x,y= 3884725,3165756 (EPSG:3035)

4.2 Checking deviations between IMD2018 and IMD2015

Difference maps were created based on differences of 1km x 1km aggregated sealing density values derived from various datasets to visualize regional and local effects.

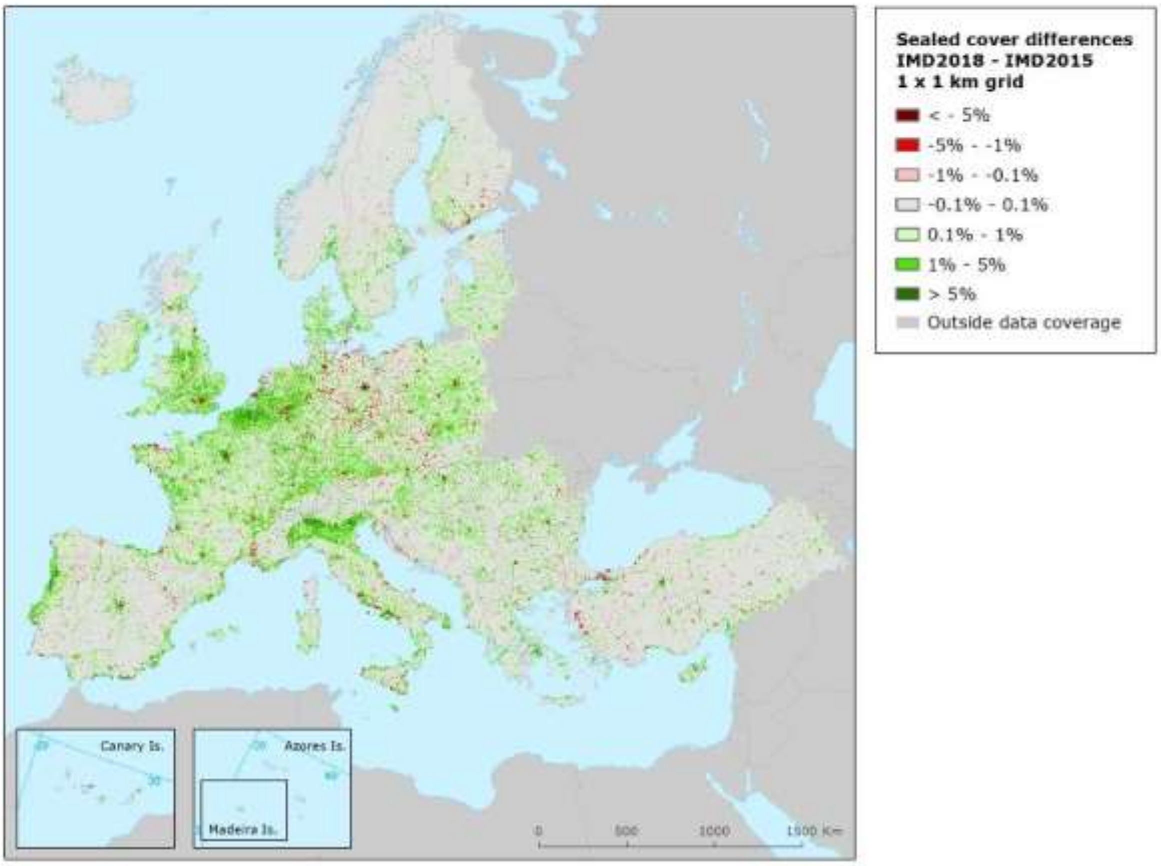


Figure 4.2.a: Local and regional differences of aggregated sealing estimated by IMD2018 ≥ 30% vs IMD2015 ≥ 30%. Green: More sealing is estimated by IMD2018. Red: More sealing is estimated by IMD2015

The overview of the difference map indicates surplus of sealed area estimated by IMD2018 (green colours) for larger contiguous areas, while IMD2015 based sealing has shown surplus at specific locations.

Locations of largest differences were checked with the help of VHR Bing or Google Earth historical imagery, typical examples are shown below.

Typical examples included missing details of road network, several details of city areas, greenhouses, farm houses and agroindustry, as well as commission in IMD2015 data at specific locations.

4.2.1 Surplus of sealing by IMD 2018 (green cover on Figure 4.2.a overview map)

Road network, several details of city areas, greenhouses, farm houses and agroindustry facilities are appearing in the 10m resolution IMD2018 with more detail.

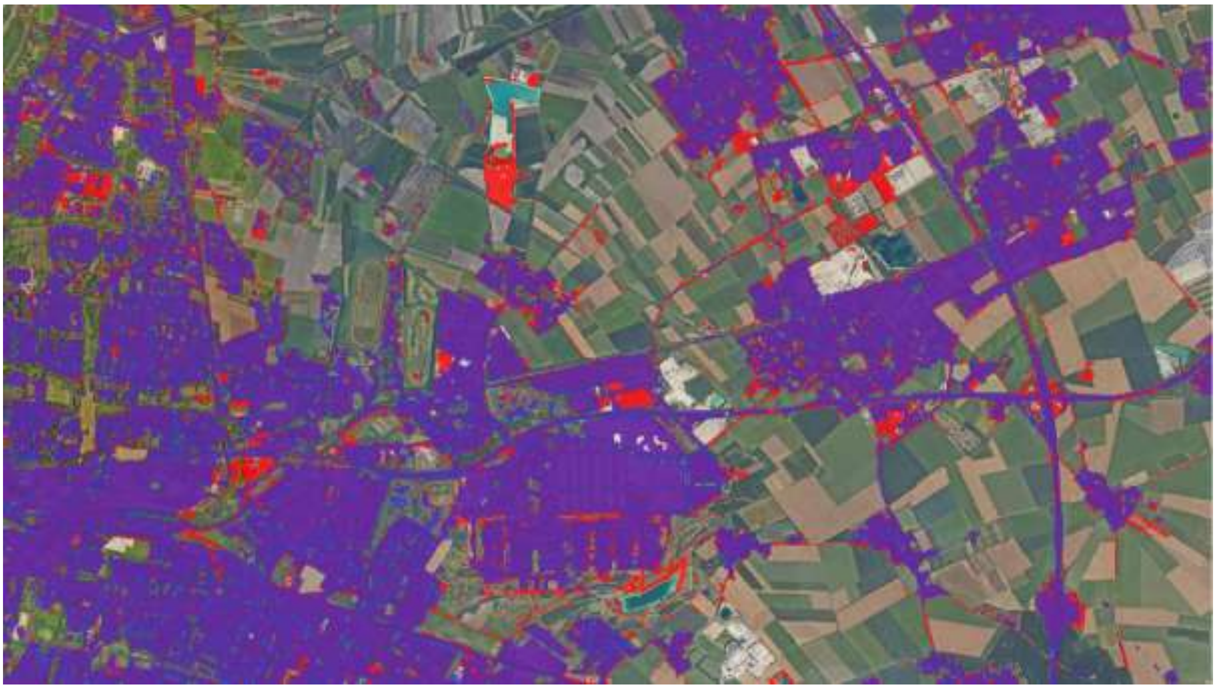


Figure 4.2.1.a: More details of road network and infrastructure in IMD2015 (red), but some commission on low density urban areas by IMD2015 (blue)
Red: IMD2018 ≥ 30%; Blue: IMD2015 ≥ 30%; Purple: IMD2018 ≥ 30% & IMD2018 ≥ 30%;
Germany (surrounding of Munich), x,y= 4446988,2782339 (EPSG:3035)

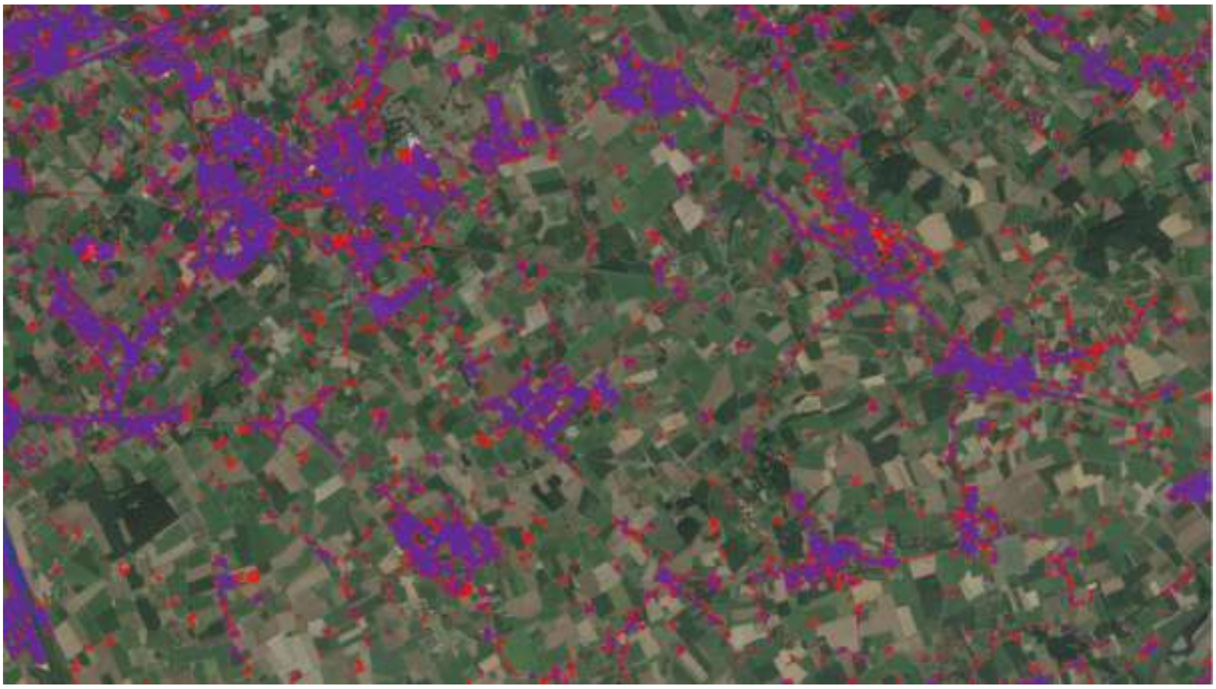


Figure 4.2.1.b: More details of road network, greenhouses, agroindustry (but some commission as well) in IMD2018 (red).
Red: IMD2018 ≥ 30%; Blue: IMD2015 ≥ 30%; Purple: IMD2018 ≥ 30% & IMD2018 ≥ 30%;
Belgium, x,y= 3863238,3096931 (EPSG:3035)

Surplus of sealing by IMD 2015 (red cover on Figure 4.2.a overview map)

Road network, several details of city areas, greenhouses, farm houses and agroindustry facilities are appearing in the 10m resolution IMD2018 with more detail, low density areas are more precisely differentiated.

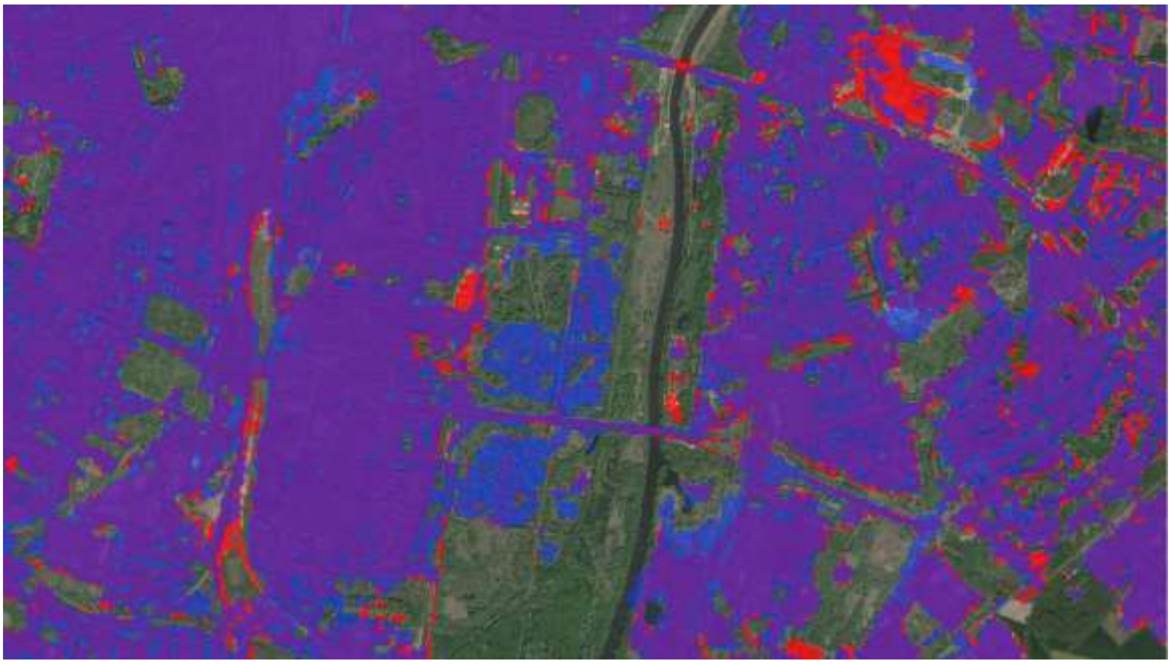


Figure 4.2.1.c: Commission errors in IMD2015 (blue spots) on low density urban and tree covered areas, as well as new sealed areas in IMD2018 (red spots)
Red: IMD2018 ≥ 30%; Blue: IMD2015 ≥ 30%; Purple: IMD2018 ≥ 30% & IMD2018 ≥ 30%;
Poland (City of Poznan), x,y= 4794141,3273218 (EPSG:3035)

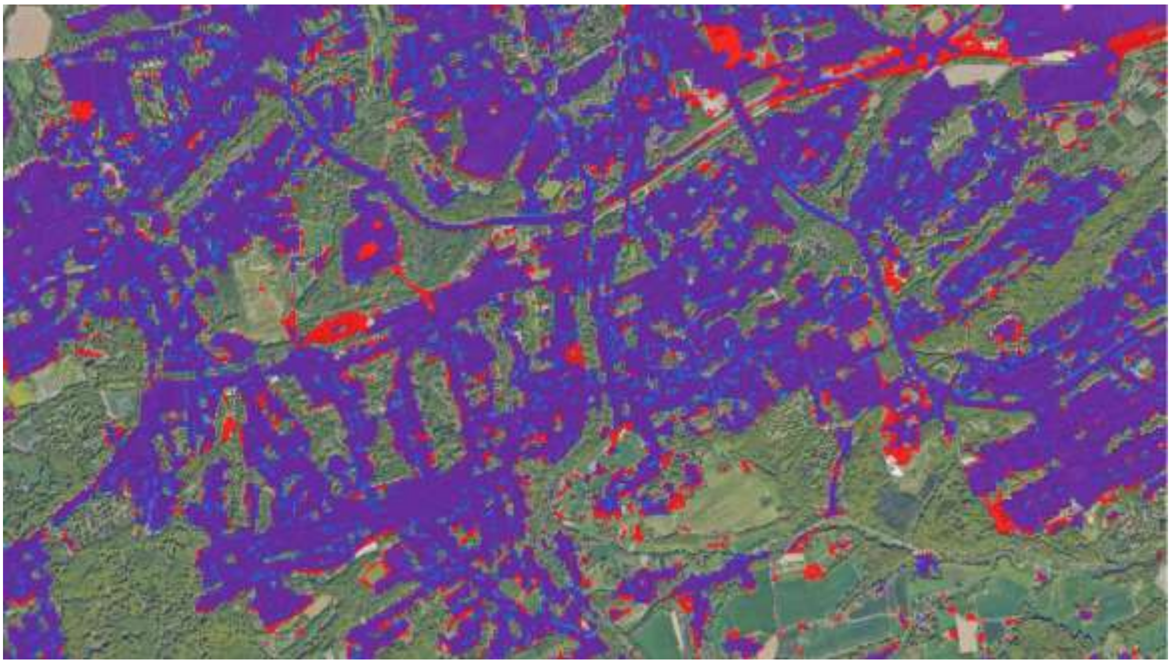


Figure 4.2.1.d: Commission errors in IMD2015 (blue spots) on low density urban areas, as well as new sealed areas in IMD2018 (red spots)
Red: IMD2018 ≥ 30%; Blue: IMD2015 ≥ 30%; Purple: IMD2018 ≥ 30% & IMD2018 ≥ 30%;
Germany (City of Bochum), x,y= 4128335,3149458 (EPSG:3035)

5. Creating an EU reference sealing level database

Besides of the appearance of potential misclassification errors, there are obvious principal differences between an Earth Observation (EO) based and an in-situ ground-truth based survey which arises doubts in statistical community whether area statistics could be directly extracted from a map. On the other hand, areal statistics derived from various EO based products show obvious differences as well. Thematic content, geometry of features as well as areal statistics derived on any spatial data are significantly influenced by several factors, like:

- Semantic content (class definitions);
- Scale (raster resolution, MMU, MMW);
- Surveying method and instructions;
- Accuracy (random and possible systematic errors).

The direct comparison of summary statistics derived from pan-European Earth Observation based CLMS products and the European LUCAS statistical survey is not an appropriate method to decide the discussion, as these are potentially influenced by all the above factors.

- The LUCAS survey is sample based, able to detect the appearance of very fine details at the exact sample location and delivers ground-truth, but able to provide statistically sound results for certain LCLU categories only at NUTS2 level and above;
- High resolution CLMS data like HRL Imperviousness or CLC+ Backbone are able to detect fine geometry down to a certain level, but are burdened by potential misclassification errors and the 10m resolution.

The comparison of areal statistics aggregated to a lower resolution grid helps to eliminate differences caused by local geometric uncertainties. Derived scatter-plots and difference maps provide easy overview of correlations as well as exceptions showing large differences in aggregated values. Estimated sealed areas were compared on aggregated levels in the frames of the Copernicus Service Contract SC59032 (Task14), see more details in task reports and the summary in Chapter 3.

The analysis of scatter-plots lead to following main conclusions:

- High correlation was found between estimated sealed content comparing CLC+ Backbone 2018 class1 (sealed) and HRL Imperviousness 2018 data, especially when a threshold of 30% was applied in IMD2018;
- The aggregation level of 1km x 1km was found as optimal for the comparison, as the 10km x 10km resolution was too coarse and the 100m x 100m comparisons have shown too large scattering.

The analysis led to the conclusion, that appropriate sample-based reference data are needed to be able to set-up a link between EO derived products and the “real-world” represented by sample based in-situ surveys. The idea was tested by visual interpretation for the test area of Hungary by interpreting 101 rectangular 1km x 1km sample units. Main conclusion of the study was, that obvious trend and correlation was experienced between EO based and “ground truth based” sealing level estimations. However, the few samples allowed by limited resources did not allow the proper estimation of the parameters of the trend.

5.1 Methodology for creating a reference sealing level database

European scale validation exercises of HRL Imperviousness data have set up a reference sealing level database on the aggregated 100m x 100m spatial resolution, by estimating the sealing level of a single Primary Sampling Unit (PSU), by visual interpretation of a 5 x 5 point-grid of altogether 25 Secondary Sampling Units (SSU). This validation exercise has resulted a large reference dataset, with thousands of

PSUs, and used for the validation of the degree of imperviousness values on 100m x 100m aggregated level. However, this reference dataset has its limitations, because:

- The uncertainties of EO based sealing level estimations cause a large scattering at 100m x 100m resolution due to possible geometric misallocations and other factors;
- The use of 25 SSUs per a rectangular unit does not allow high reliability of visual interpretation based sealing level estimations, the peak of confidence intervals at the confidence level of 95% around 50% sealing is CI=19,6%.

Sampling plan

Previous exercises of sealing level comparisons presented in Chapter 3 indicate the spatial resolution of 1km x 1km as optimal for comparing the sealed content of CLMS data like HRL Imperviousness and CLC+ Backbone. Therefore, sample grid cells of the European 1km x 1km ETRS89 LAEA statistical reference grid were selected as Primary Sampling Units for the new reference database.

Stratification

Multi-level stratified random sampling was applied for the selection of PSUs. Initial stratification of the 1km x 1km rectangular units covering EEA39 area was based on the appearance of sealing related changes within the potential PSU. Altogether 10 set PSUs were selected by this initial criterion.

Table 5.1.a: Overview of the number of PSUs resulted by the multi-level stratification

NAME of PSU set	Total number of potential PSUs in the set	Number of PSUs selected & interpreted
Inverse of CLC based land take greater than 1ha in period 2006-2012	4 980	86
Inverse of CLC based land take greater than 1ha in period 2012-2018	5 636	88
CLC based land take greater than 1ha in period 2006-2012	58 213	100
CLC based land take greater than 1ha in period 2012-2018	45 375	94
Loss of sealed cover (IMCC class 2) is greater than zero in period 2006-2015	185	59
New sealed cover (IMCC class 1) is greater than 1000 sqm both in periods 2015-2018 AND 2006-2009	158 383	100
New sealed cover (IMCC class 1) is greater than 1000 sqm both in periods 2015-2018 AND 2009-2012	130 325	101
New sealed cover (IMCC class 1) is greater than 1000 sqm both in periods 2015-2018 AND 2012-2015	79 578	101
Loss of sealed cover (IMCC class 2) is greater than 1000 sqm in period 2015-2018	17 650	99
No sealing change in any of periods between 2006-2018	4 584 092	101
TOTAL number of PSUs	**5 084 417**	**929**

The second level of the stratification was applied to ensure that similar number of PSUs are selected for each sealing levels between 0-100%. This second level of stratification was following criteria:

- EO based (CLC+ Backbone 2018 based in practice) estimation of aggregated sealing level was used for the random selection;
- One PSU was selected within each set based on each EO based sealing degree in the range of [0..100].

Each of the rectangular 1km x 1km sample grid cells (PSUs) was filled with a regular 10x10 point grid (SSUs). The status (sealed/non-sealed) of each of SSU was visually interpreted with the help of very-high resolution aerial ortho-imagery.

Figure 5.1.a: Reference interpretation of sealing within a 1x1km PSU for year 2018 (green: non-sealed, red: sealed).

Sealed status was interpreted for all five reference years corresponding to HRL IMD data in the time series (2006, 2009, 2012, 2015, 2018). Subjectivity factor of the interpretation was reduced by cross checking results by different interpreters.

5.2 Evaluation of interpretation results

Altogether 929 sample units (PSUs) were finally interpreted, distributed in 10 primary strata. The final number of interpreted SSUs and correspondingly the "valid" PSUs for a certain reference date was determined by the availability of appropriate VHR Google Earth historical imagery.

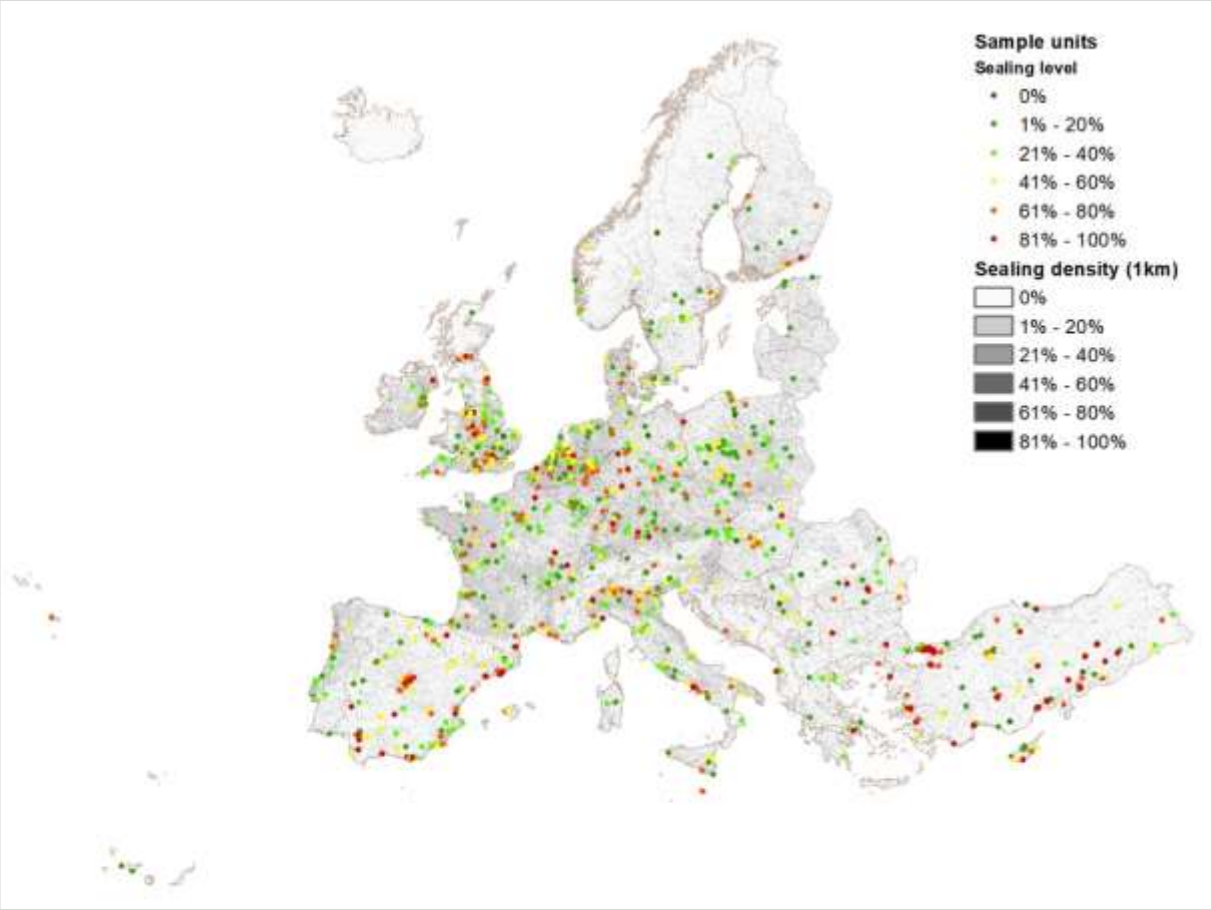


Figure 5.2.a: Location of interpreted PSUs in the reference database coloured by the CLC+ Backbone based sealing levels for the reference year of 2018

The number of valid SSUs (interpreted as "sealed" or "unsealed", but excluding locations where no appropriate reference imagery was found) was summarized for each PSU. PSUs with less than 100 valid SSUs were excluded from the further evaluation. Appropriate reference imagery was found for 927 samples for the reference year of 2018, while less samples could be interpreted for previous reference years – least samples could be interpreted for year 2006.

Table 5.2.a: Overview of the number of PSUs resulted by the multi-level stratification

Reference year	**2018**	**2015**	**2012**	**2009**	**2006**
Number of valid PSUs (100 valid SSUs)	927	825	900	619	512

The reference sealing level was calculated for each valid PSU by summarizing the number of SSUs of "sealed" status, this has resulted directly the estimated sealing level in the range of 0-100%.

The 100 SSUs within one PSU ensures higher reliability of estimations than the 25 SSUs used by European scale validation approaches for 100m x 100m PSUs. The peak of confidence intervals at the confidence level of 95% at 50% estimated sealing is CI=9,8%.

6. Comparison of sealed content of both datasets to ground truth

Reference sealing values interpreted during activity 5 were compared to Earth Observation based sealing level estimations, trends and deviations were analysed.

6.1 Methodology

Partial sealed cover expressed in percentages in the range of 0-100% for 1km x 1km PSUs were estimated both by visual interpretation (described in chapter 5) and by EO based land cover data as Degree of Imperviousness and CLC+ Backbone.

Sealing density values estimated by visual interpretation

The number of SSUs for "sealed" status provided a direct estimation of sealing density values for each PSU. PSUs with less than 100 valid (interpretable) SSUs were excluded from comparisons.

Sealing density values estimated by CLC+ Backbone

The CLC+ Backbone based estimation of 1km resolution sealing levels was calculated in following steps:

1. Class 1 (sealed) was extracted from 10m resolution CLC+ Backbone data resulting a 10m resolution binary sealed / non-sealed map;
2. Aggregation to 100m resolution by counting sealed 10x10m cells within each 100x100m cell, resulting sealing density values in a range 0-100%;
3. Aggregation to 1km resolution by calculation the mean sealing densities within each 1x1km cell, resulting sealing density values in a range 0.00-100.00%;

Sealing density values estimated by Degree of Imperviousness

Primary 10m resolution IMD data provide the estimation of sealing density values in a range of *0-100%.* Previous validation exercises have proven, that the direct use of these density values significantly underestimates real sealing densities. Best correlations between IMD based estimations and reference were shown when a threshold around 30% was applied on IMD data.

Therefore, the IMD based estimation of 1km resolution sealing levels was calculated in three steps:

1. Raster cells of IMD value ≥ 30% were extracted from 10m resolution IMD data resulting a 10m resolution binary sealed / non-sealed map;
2. Aggregation to 100m resolution by counting sealed 10x10m cells within each 100x100m cell, resulting sealing density values in a range 0-100%;
3. Aggregation to 1km resolution by calculation the mean sealing densities within each 1x1km cell, resulting sealing density values in a range 0.00-100.00%;

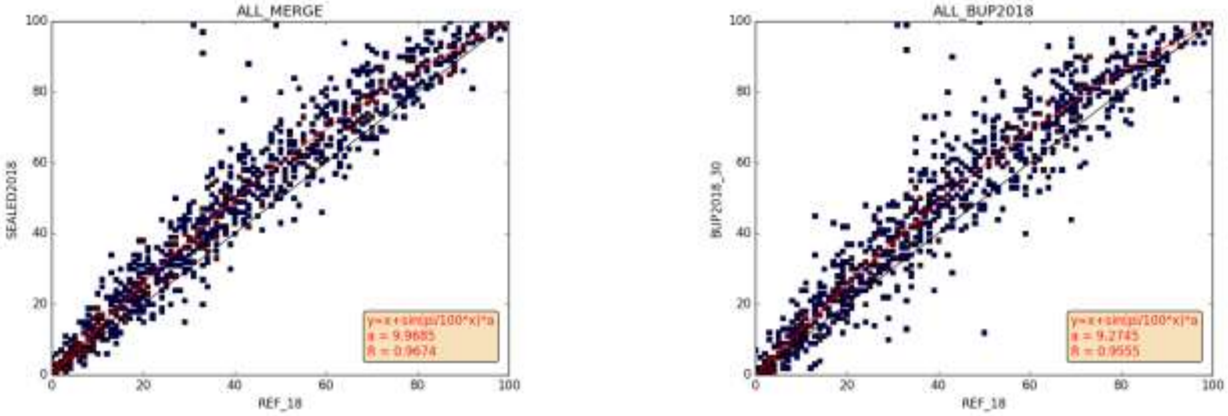
EO based sealing level estimations were rounded to integer values and compared to visual interpretation-based values by displaying these on scatterplots, fitting trend lines and calculating correlation coefficients.

6.2 Results

By examining scatterplots based on all PSUs collected (Figure 6.2.a) it became obvious, that a linear relationship is not really applicable to describe the shape of the trend when comparing CLC+ Backbone or IMD based estimations with reference:

- Typically, there is little or no difference between EO based and reference interpretation-based estimations close to sealing levels of 0% and 100%;
- Highest deviations are shown for medium sealing levels (around 50%);
- The correlation is high in both cases, the scattering is rather low;
- Some extreme deviations are shown by scatterplots, these were checked visually, see examples in chapter 6.3.

Finally, a non-linear trendline, described by the function **y=x+sin(pi/100*x)*a** was successfully applied to describe the relationship between EO based values and reference.



a) CLC+ Backbone 2018 vs reference

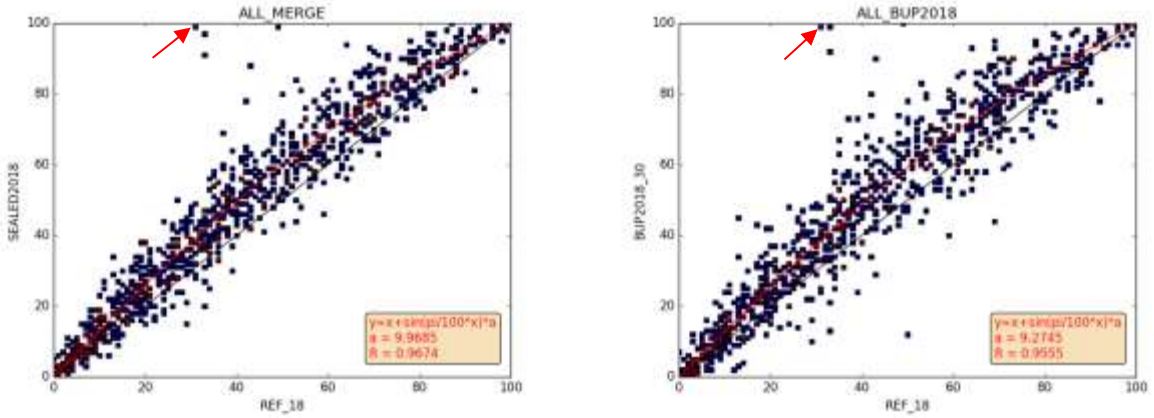
b) IMD2018 ≥ 30% vs reference

Figure 6.2.a: Comparison of 1x1 km aggregated sealing densities by a) Backbone 2018 class 1 (Sealed) and b) HRL IMD2018 ≥ 30% with sealing estimations gained by reference interpretation. Red line indicates fitted non-linear function, black line indicates main diagonal.

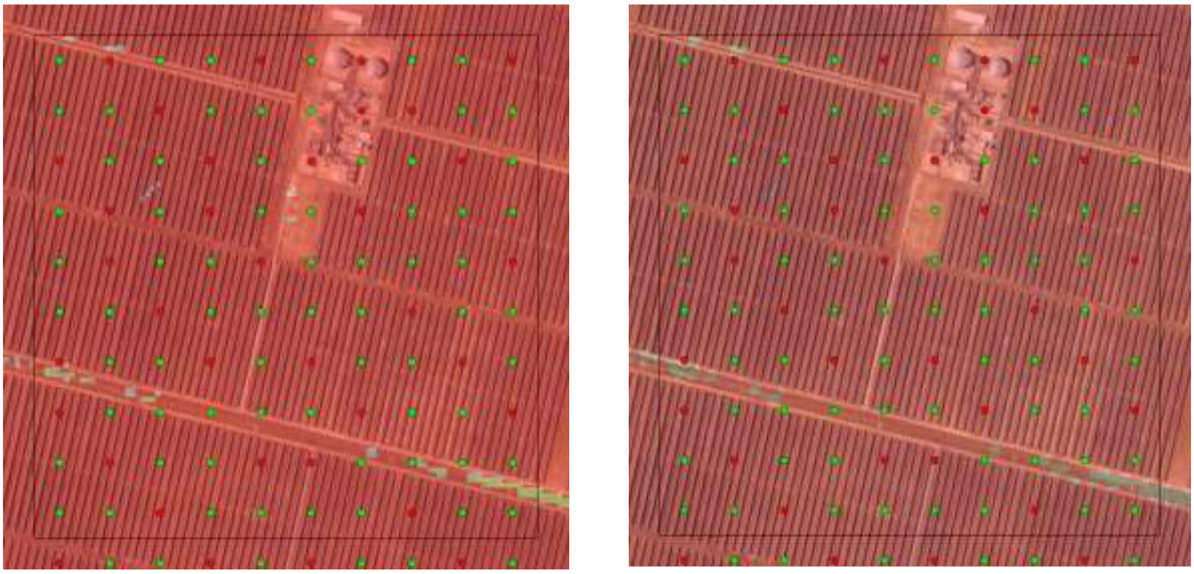
The fitted non-linear trend line is characterized by the one single coefficient "a". Both correlation values "R" and coefficients "a" have shown similar values, when comparing reference values to CLC+ Backbone and IMD based estimations.

These results mean, that sealing levels are still slightly overestimated by EO based datasets compared to reference, but the inverse of the fitted non-linear trend seems to be applicable to harmonize 1km resolution sealing level data (see more details in chapter 7.3).

6.3 Examples of significant deviations



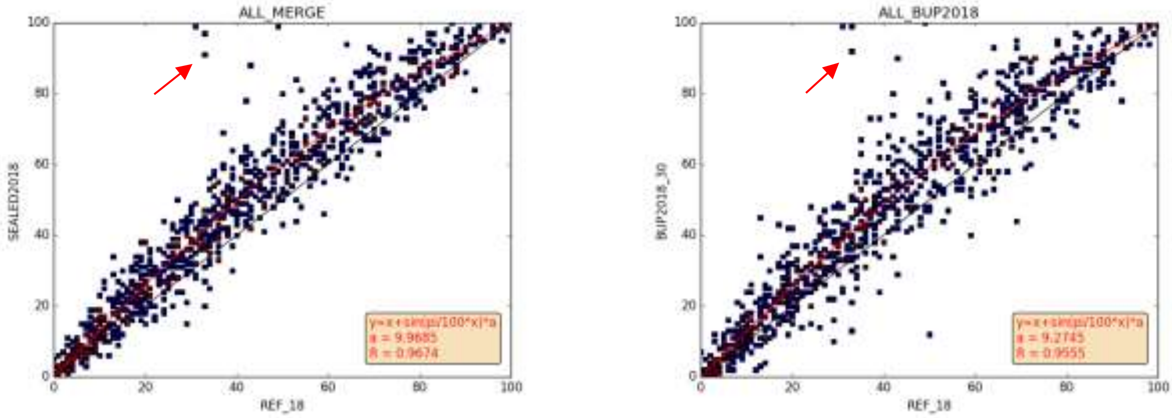
a) CLC+ Backbone 2018 vs reference *b) IMD2018 ≥ 30% vs reference*



c) CLC+ Backbone 2018 & GE imagery *d) IMD2018 ≥ 30% (red), IMD2018: 1-29% (yellow) & GE imagery*

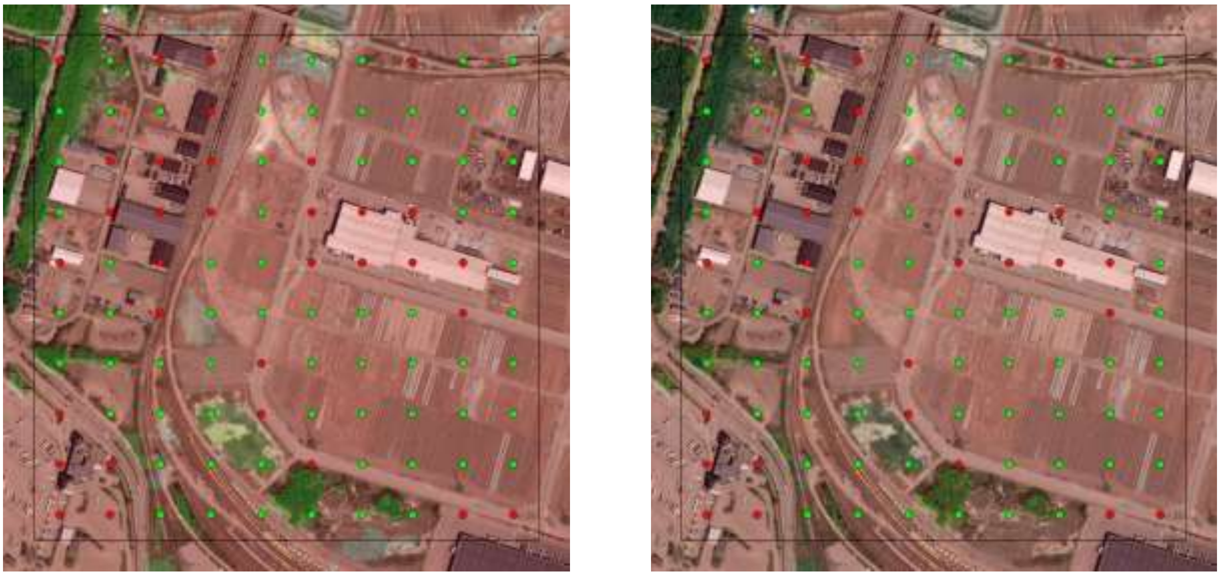
e) GE imagery & reference interpretation enlarged *f) GE Field photograph*

Figure 6.3.a: Example for extreme overestimation of sealing by both CLC+ Backbone and IMD2018. Solar panel park, Spain. x,y= 2866745,1888730 (EPSG:3035)



a) *CLC+ Backbone 2018 vs reference*

b) *IMD2018 ≥ 30% vs reference*



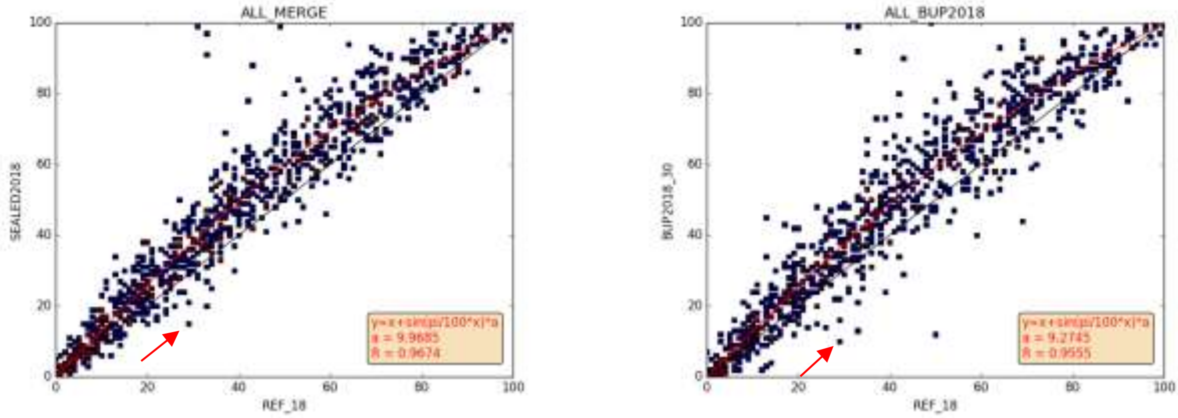
c) *CLC+ Backbone 2018 & GE imagery 2020*

d) *IMD2018 ≥ 30% (red), IMD2018: 1-29% (yellow)* & GE imagery 2020

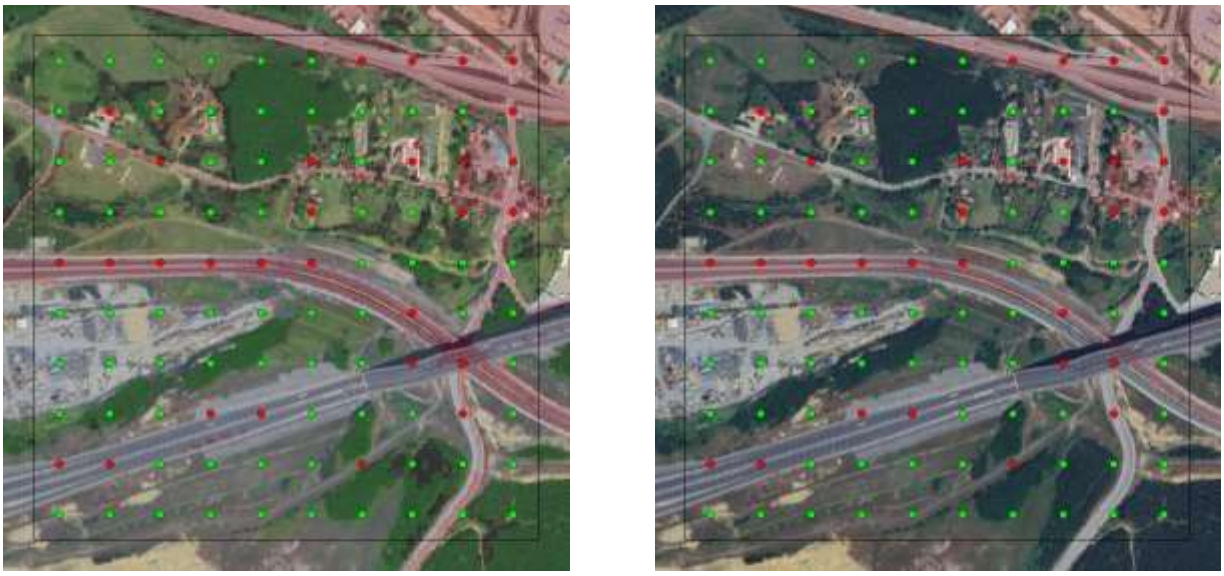
e) *GE imagery 11/10/2018*

f) *GE imagery 17/04/2019*

Figure 6.3.b: Example for extreme overestimation of sealing by both CLC+ Backbone and IMD2018. Container terminal, mostly unsealed but temporally loaded. Finland. x,y= 5243386, 4258980 (EPSG:3035)



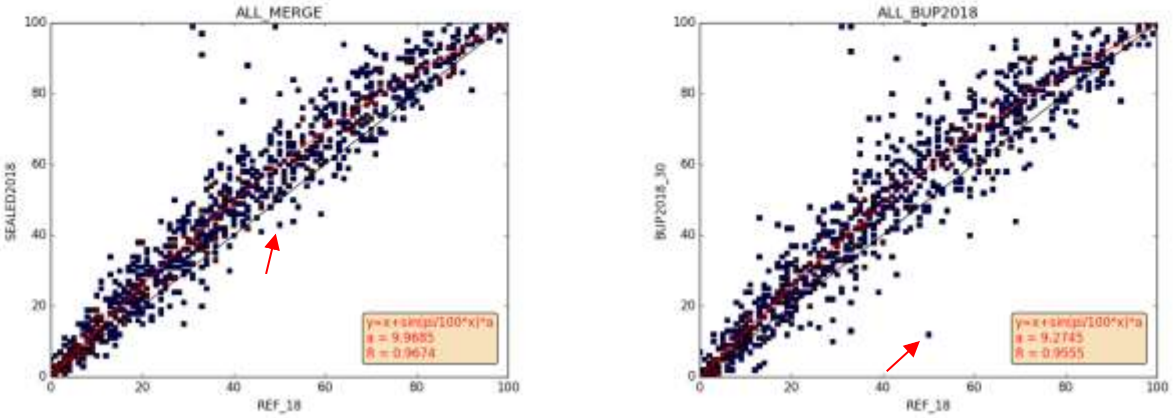
a) CLC+ Backbone 2018 vs reference *b) IMD2018 ≥ 30% vs reference*



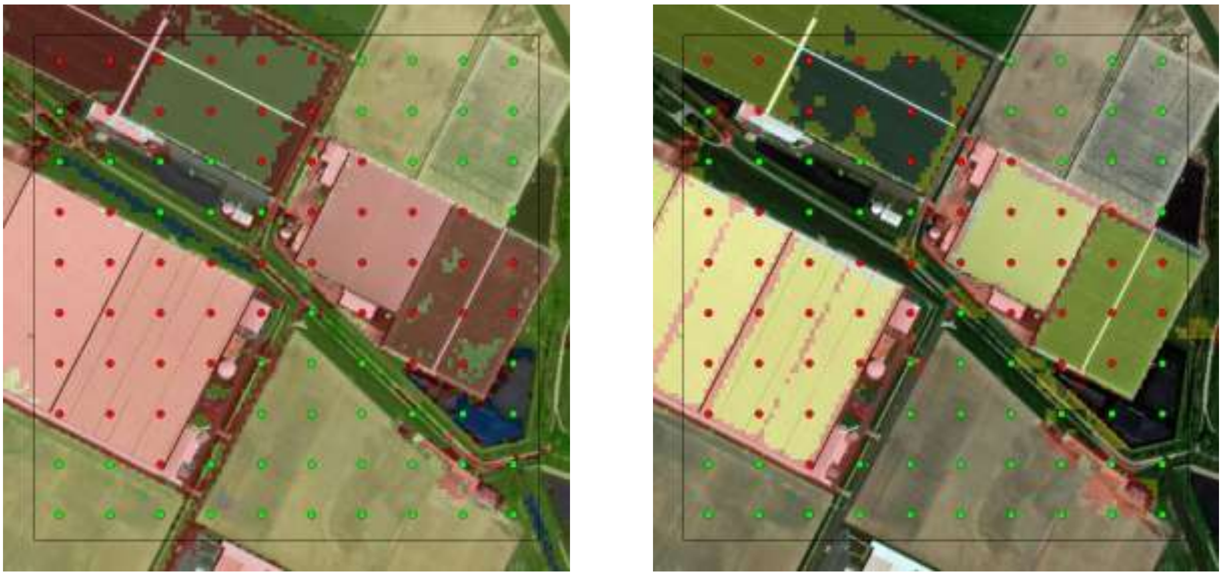
c) CLC+ Backbone 2018 & GE imagery 2021 *d) IMD2018 ≥ 30% (red), IMD2018: 1-29% (yellow)* & GE imagery

e) GE imagery 14/05/2018 *f) GE Street View 2019*

Figure 6.3.c: Example for significant underestimation of sealing by both CLC+ Backbone and IMD2018. New highway built in 2018 is omitted both by CLC+ Backbone and IMD2018. Türkiye. x,y= 5885244, 2210980 (EPSG:3035)



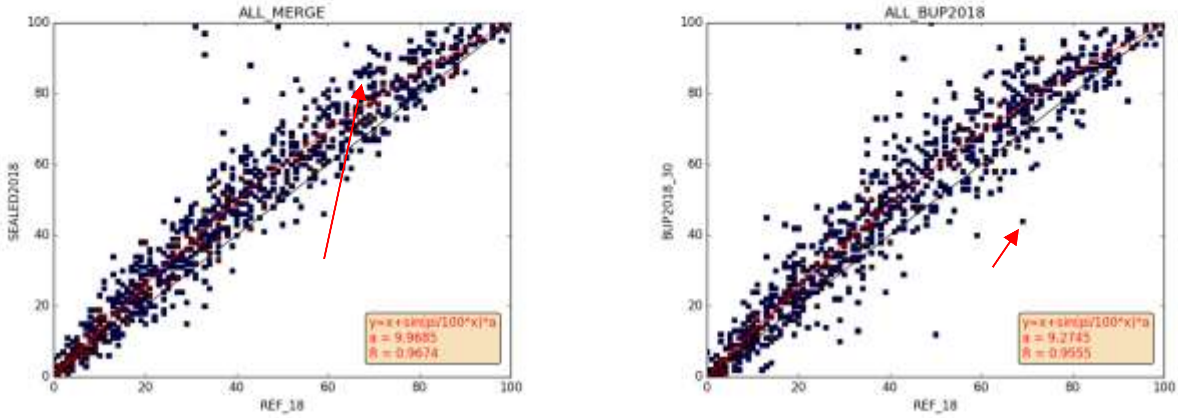
a) CLC+ Backbone 2018 vs reference *b) IMD2018 ≥ 30% vs reference*



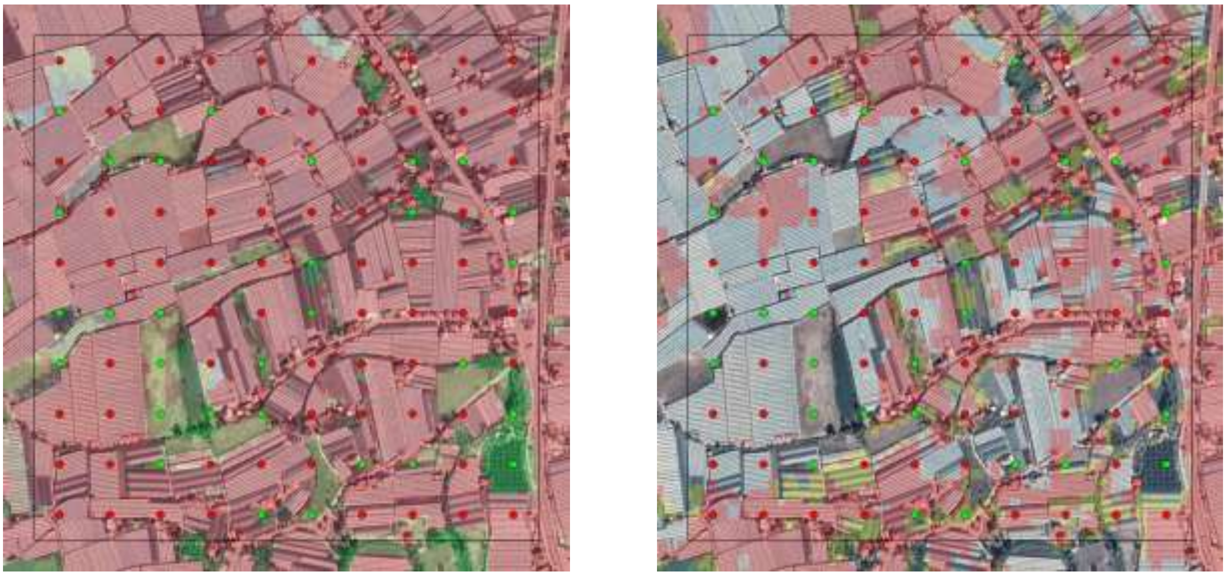
a) CLC+ Backbone 2018 & GE imagery *b) IMD2018 ≥ 30% (red), IMD2018: 1-29% (yellow)* & GE imagery

c) GE imagery 8/05/2018 *d) GE Street View 2012*

Figure 6.3.d: Example for significant (IMD2018 ≥ 30%) and slight (CLC+ Backbone) underestimation of sealing. Greenhouses are mapped mostly correctly by Backbone, and found by IMD2018, but IMD value is under 30%. The Netherlands. x,y= 3933666, 3182977 (EPSG:3035)



a) *CLC+ Backbone 2018 vs reference* b) *IMD2018 ≥ 30% vs reference*



e) *CLC+ Backbone 2018 & GE imagery* f) *IMD2018 ≥ 30% (red), IMD2018: 1-29% (yellow)* & GE imagery

g) *GE imagery 7/08/2018* h) *GE Street View 2020*

Figure 6.3.e: Example for significant underestimation by IMD2018, while slight overestimation by CLC+ Backbone. Greenhouses are mapped mostly correctly by Backbone, but many of these were not found by IMD2018. Türkiye. x,y= 6051478,1691983 (EPSG:3035)

7. Harmonization of the soil sealing time series

The continuing rapid expansion of sealed surfaces — often improperly planned and unjustified by the population's needs — leads to hydrological impacts, which influence run-off in urban catchments, and the loss of crop production, which affect the global carbon cycle. Sealing is also a key driver of the loss of biodiversity and ecosystem services. The European Commission's roadmap to a resource efficient Europe introduced a 'no net land take by 2050' initiative that aims to ensure that either all new urbanisation occurs on brownfields or any new land take is compensated for by reclaiming artificial land. Land take and sealing are also relevant to several targets of the EU biodiversity strategy for 2030, aimed at protecting and restoring nature[6].

The proper assessment of the spatial indicators and the quality of conclusions drawn requires high quality, harmonized time-series of input data, both in terms of geometry of sealed features and quantity of sealing.

7.1 History of soil sealing related information within the CLMS portfolio

High resolution soil sealing related information is provided within the CLMS portfolio by the time-series of the Degree of Imperviousness data from the reference year of 2006 - updated every 3 years - and by class 1 (sealed) of CLC+ Backbone data from the reference year of 2018.

Production of historical time-series

The production of pan-European high resolution soil sealing data was started under the name of Soil Sealing 2006 and was updated with similar specifications and parameters in 2009, 2012 and 2015 in 20m resolution based on available (mainly IRS, Spot 5, Landsat) satellite imagery. The primary product was a 20m resolution Soil Sealing / Imperviousness Density (IMD) layer storing density values between 0-100%, representing the estimated percentage of the single raster cell covered by sealed surface. The 100m aggregated version was provided as additional product.

Imperviousness change data were produced / published first only in 100m resolution. The imperviousness density change (IMC) layer was created as the difference of neighbouring status layers and presented the change of soil sealing density values in the range of 0-200, representing the difference of density values between -100% to +100%.

Re-analysis of historical time-series

Each of the soil sealing related status maps are created by individual complex classification procedures, providing detailed information about the structure and estimated quantity of artificially sealed surfaces. During the early calculations of imperviousness indicator, it turned out, that the derived pan-European indicators show unrealistic trends, as the consequence of missing harmonization between individual status maps.

Therefore, in frames of the Imperviousness 2015 update, the re-processing and the re-analysis of the 2006-2009-2012 historical layers was performed together with the production of 2015 status. The methodologies applied for both were the re-processing, and re-analysis of all existing density products, to assure properly calibrated HRL Imperviousness time-series, but did not constitute a complete re-production from the scratch.

---

[6] EEA indicators: Imperviousness and imperviousness change in Europe

The resulting 20m and 100m resolution IMD time-series for 2006-2009-2012-2015 years as well as the 20m and 100m resolution imperviousness density change (IMC) layers fulfil the criteria of entirely harmonized (accounting) time-series, namely that the difference of neighbouring pair of status layers corresponds to a change layer. Note, that although the accuracy of imperviousness data was significantly improved by the re-analysis, the IMC layers still contain technical changes (noise) beyond real changes. Additionally, the historical layers and consequently the change layers contain unclassifiable areas (coded with 254 value), because of cloud cover in satellite imagery.

Resolution upgrade in CLMS HRL portfolio

The appearance of the high quality 10m resolution Sentinel satellite imagery in 2016 has allowed the upgrade of upcoming CLMS high resolution land cover products from 20m to 10m spatial resolution. Unfortunately, the availability of more spatial details had confusing consequences to indicator calculations. While the former 20m resolution IMD time-series (2006-2009-2012-2015) has been successfully harmonized and have shown credible evolution in sealed cover, the upgraded IMD2018 data detect significantly more sealed structures than before, thus the amount of sealed cover is showing an unrealistic growth in the 2015-2018 period.

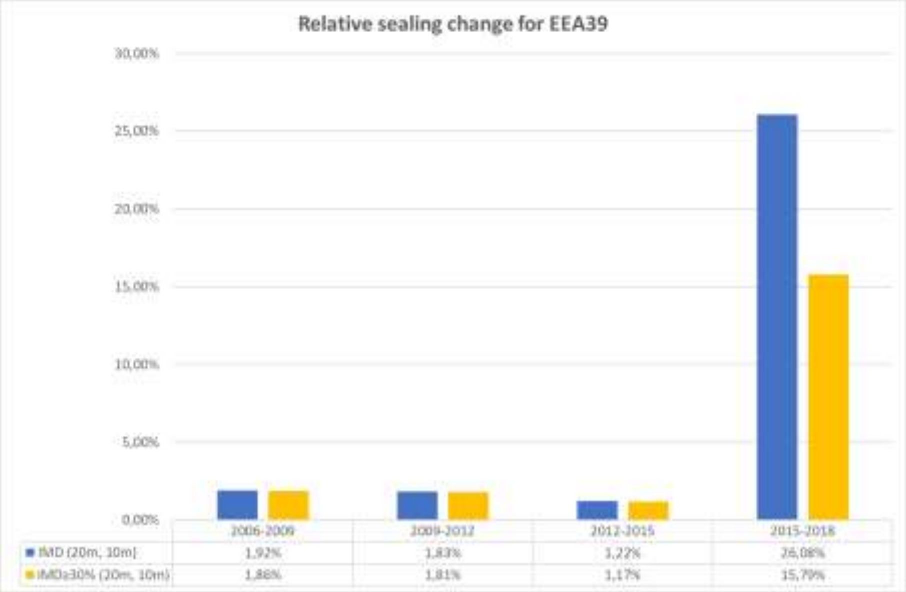
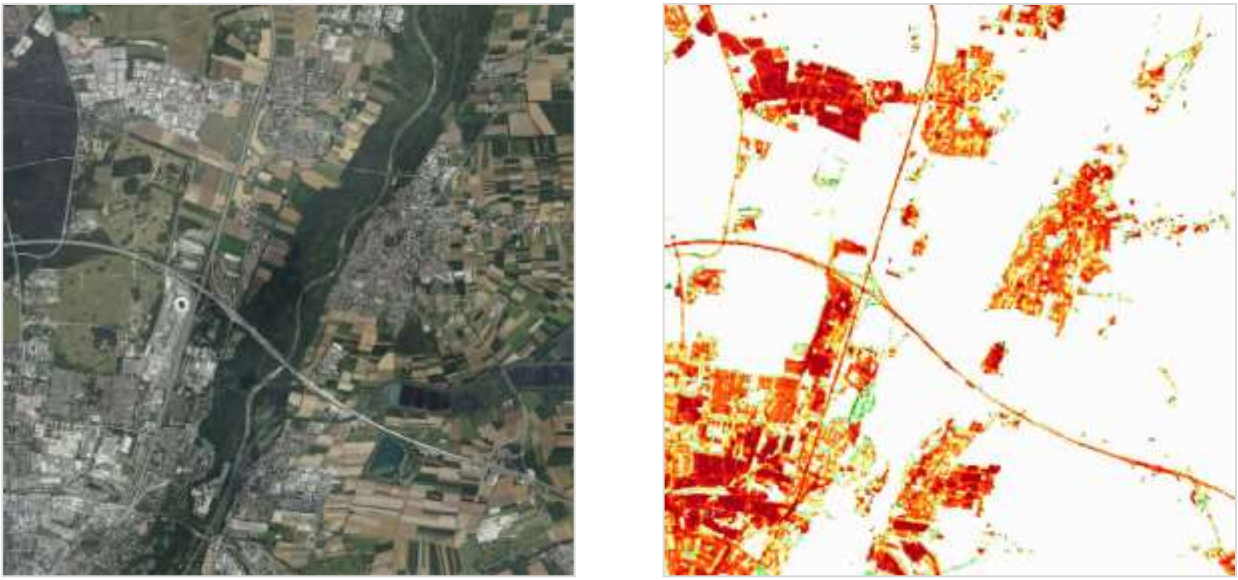


Figure 7.1.a: Changes of soil sealing relative to the previous reference year calculated for EEA39 area.

The relative change of the quantity of sealed areas calculated both directly based on Degree of Imperviousness values (blue bars) and based on a binary sealing map derived by applying a threshold at IMD=30% (orange bars) shows extreme increase for the period 2015-2018 compared to all previous periods.

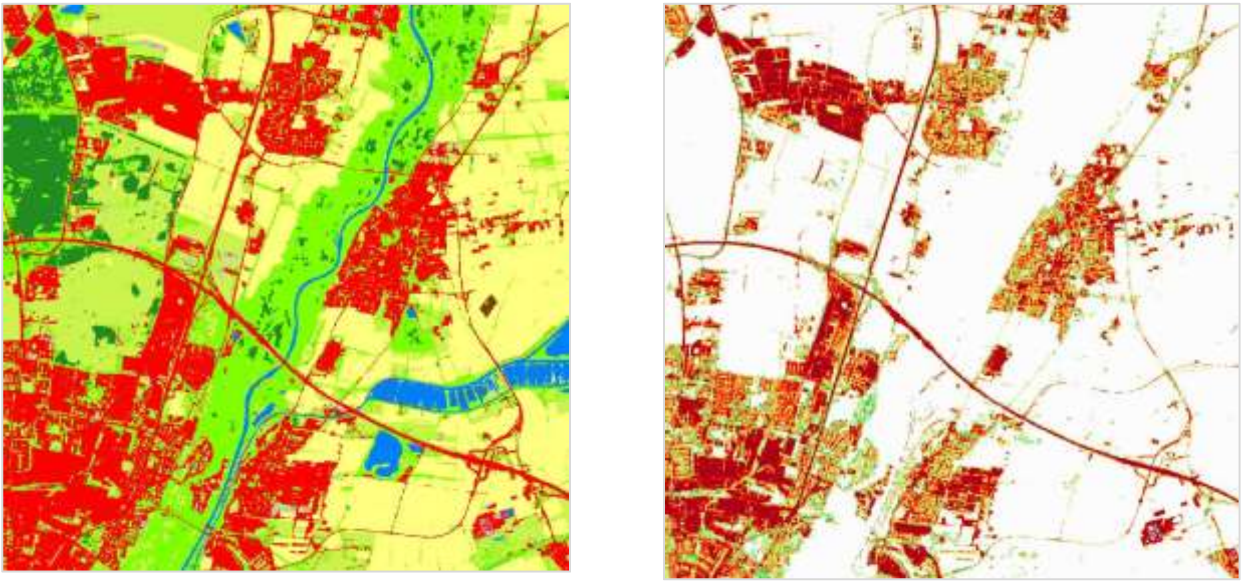
The extreme increase of relative differences mapped by IMD data is the consequence of two main reasons:

- Calibration differences between IMD2015 and IMD2018 data causing large difference between direct IMD (26,08%) and IMD≥30% (15,79%) based values;
- The effect of resolution upgrade, as more spatial structures were mapped by 10m resolution products, than previously in 20m resolution.



a) *VHR Google Earth imagery*

b) *Degree of imperviousness 2015 (20m resolution)*



c) *CLC+ Backbone 2018 (10m resolution)*

d) *Degree of imperviousness 2018 (10m resolution)*

Figure 7.1.b: Sealed areas captured by certain CLMS surveys. Hornbach, NW form Munich, Germany (x, y: 4443700,2786200 meters)

The geometric structure of sealed features is strongly determined by the spatial resolution, missing features and details are hardly to extract from low resolution data, as demonstrated on Figure 7.1.b.

Very similar details of sealed structures were mapped by the new pan-European land cover product of CLC+ Backbone appearing first in the reference year of 2018. While spatial structures are recognized with similar detail by 10m resolution CLC+ Backbone (c) and IMD2018 (d) maps, significantly less detail are shown by 20m resolution IMD2015 (b) data.

Additionally, there is a visible effect of calibration differences of imperviousness degrees between IMD2018 (d) and IMD2015 (b) maps, sealed surfaces are characterized with slightly different degrees of imperviousness values, represented by different shades of colours.

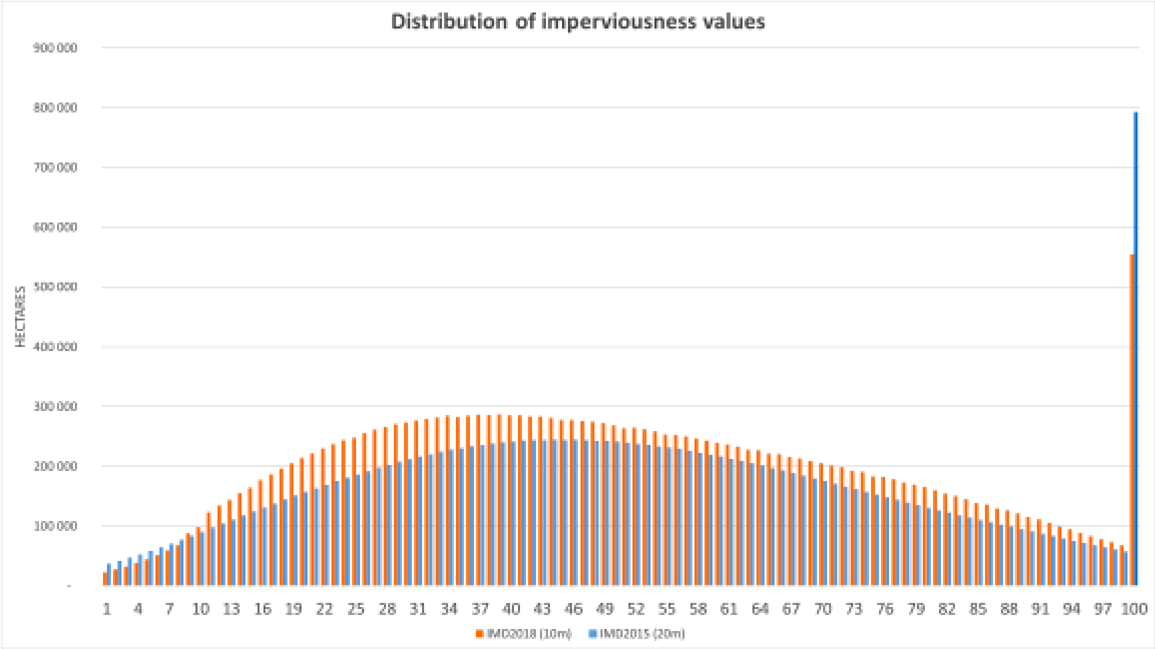


Figure 7.1.c: Histograms indicating calibration differences of imperviousness degrees between IMD2015 and IMD2018 data (IMD=0% values excluded). Calibration differences may have effect to estimated extent of sealed areas, depending on the calculation method applied.

Overall calibration differences between IMD2015 and IMD2018 data are illustrated by the histogram showing the distribution of degree of imperviousness data of primary IMD products for EEA39 area on Figure 7.1.c.

7.2 Harmonization of soil sealing information

Spatial indicators using sealing data as input are based on two main characteristics:

- Geometric structure of sealed features (fragmentation, connectivity);
- Quantity of sealed areas within a certain spatial unit.

The geometric structure of sealed features is strongly determined by the spatial resolution, missing features and details are hardly to extract from low resolution data.

On the other hand, the quantity of sealed areas may be estimated for larger spatial units by various methods including estimations based on spectral indices of lower resolution EO data or point sampling as described in chapter 5.

The accounting methodology applied previously for the harmonization of Corine Land Cover (CLC) time-series may provide a satisfactory solution for the harmonization of soil sealing time-series considering both aspects.

The accounting method

The accounting method is based on the idea of creating status maps for earlier reference years by combining the latest status map with change data. The method is based on following assumptions:

- The latest status map represents usually the highest quality, both in terms of spatial details and accuracy;
- The thematic content, as well as the spatial resolution (scale) of change maps is comparable with status maps. Furthermore, it is assumed, that change features included to the map represent real phenomena.

7.2.1 Initial harmonization of the soil sealing time-series

The accounting method was applied to create the initial version of a harmonised time-series of soil sealing data with following assumptions:

- The available highest quality soil sealing map is represented by the sealed class of CLC+ Backbone 2018 layer:
    - The thematic content shows very high correlation with IMD2018 data, especially with the binary sealing map derived from IMD2018 by applying the threshold at IMD=30%;
    - CLC+ Backbone 2018 and IMD2018 data may be characterized with the same (10m) spatial resolution and mapped area (EEA39);
    - Visual comparisons indicated, that although the sealed structure mapped by both layers is very close to each other, the thematic quality of CLC+ Backbone concerning sealed areas is slightly higher.
- Available highest quality sealing change information is provided by Imperviousness Classified Change (IMCC) layers:
    - Thematic classes “new sealed cover” (1) and “loss of sealed cover” (2) were validated as highest quality (although not perfect) available sealing change related data sources within CLMS portfolio;
    - IMCC layers are providing information for the entire EEA39 area, although some internal spots are empty (unclassifiable), because of missing data in one of source status layers;
    - The spatial resolution of IMCC layers is 20m, this is coarser than the resolution of 2018 status, but may be considered as an acceptable Minimum Mapping Unit applied on change features.

Main steps of the processing

1. Create **Sealed 2018 status** by extracting class 1 from CLC+ Backbone 2018 (10m) (set all other classes to 0).
2. Create the **Sealed 2015 accounting status** layer by combining **Sealed 2018 status** with **IMCC1518**
    a) Resample 20m resolution **IMCC1518** to 10m resolution
    b) Combine **IMCC1518** <u>new sealed cover</u> with **Sealed 2018 status**: IF **Sealed 2018 status** is sealed (value=1) and **IMCC1518** is class "new cover" THEN write new **Sealed 2015 accounting status** layer with value "unsealed" (unsealed=0) at corresponding grid cell.
    c) Combine **IMCC1518** <u>loss of sealed cover</u> with **Sealed 2018 status:** IF **Sealed 2018 status** is not sealed (value=0) and change is "loss of cover" THEN write new **Sealed 2015 accounting status** layer with value "sealed" (sealed=1) at corresponding grid cell.
    d) For the rest of the pixel copy the values (sealed/unsealed, i.e. 1/0) from the **Sealed 2018 status layer**
3. Create the **Sealed 2012 accounting status** layer by combining **Sealed 2015 status** with **IMCC1215**
    a) Resample 20m resolution **IMCC1215** to 10m resolution
    b) Combine **CLMS IMCC1215** <u>new sealed cover</u> with **Sealed 2015 accounting status**: IF **Sealed 2015 accounting status** is sealed (value=1) and **CLMS IMCC1215** is class "new cover" THEN write new **Sealed 2012 accounting status** layer with value "unsealed" (unsealed = 0) at corresponding grid cell.
    c) Combine **CLMS IMCC1215** <u>loss of sealed cover</u> with **Sealed 2015 status:** IF **Sealed 2015 status** is not sealed (value =0) and change is "loss of cover" THEN write new **Sealed 2015 status** layer with value "sealed" (sealed=1) at corresponding grid cell.
    d) For the rest of the pixel copy the values (sealed/unsealed, i.e. 1/0) from the **Sealed 2015 status layer**
4. Create the **Sealed 2009 accounting** status layer by combining **Sealed 2012 accounting status** with CLMS **IMCC0912** (steps are corresponding to previous a-d points);
5. Create the **Sealed 2006 accounting** status layer by combining **Sealed 2009 accounting** status with CLMS **IMCC0609** (steps are corresponding to previous a-d points);

Results of initial harmonization

The accounting process has resulted a harmonized 10m resolution binary "sealed / non sealed" maps (version "V1") for the entire EEA39 area, where the difference of any of these status layers is equal to the number of changes present in IMCC classes used.

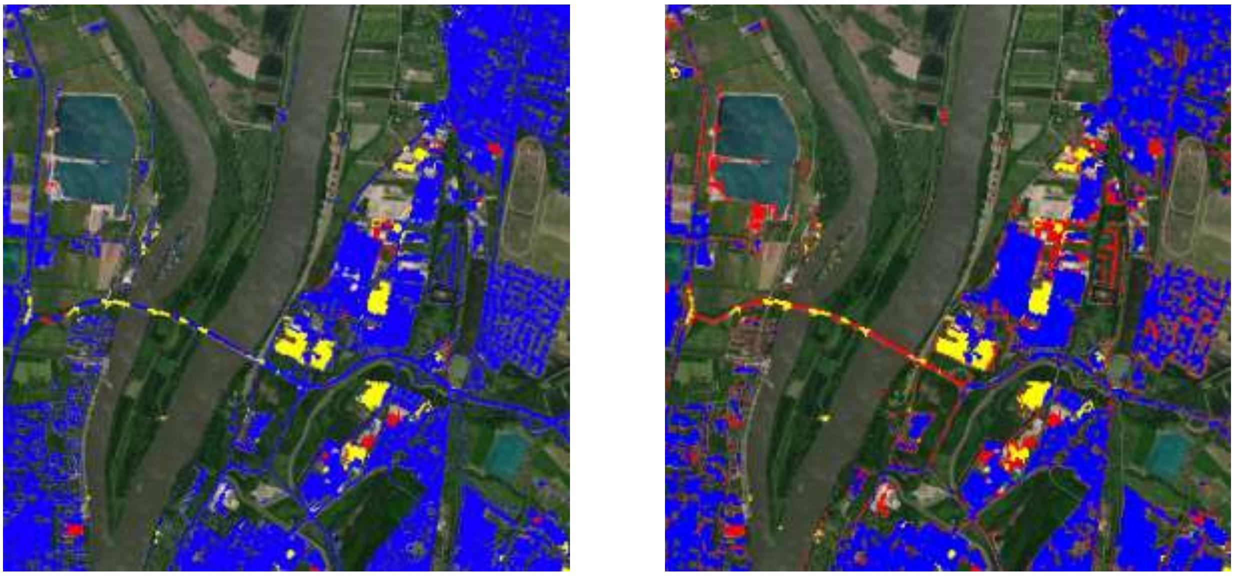
Visual inspection of the results has indicated significant improvements in the consistency of soil sealing time series (demonstrated on Figure 7.2.1.a).

- The harmonized time-series includes all the details, like smaller roads captured by CLC+ Backbone (a), but omitted by IMD2018 (b).
- The non-harmonized time-series of imperviousness (d) includes many differences not representing real changes (many red spots within settlement areas).

On the other hand, there are still some false changes are appearing in the harmonized time-series, like the lack of continuity of road network, and other small spots.

a) CLC+ Backbone 2018

b) IMD2018 (red: IMD≥30%, yellow: IMD 1-29%)



c) Harmonized time series of sealing (V1)

d) Non-harmonized time-series of imperviousness

CLC+ Backbone 2018 classes
🟥 1: Sealed
🟩 2: Woody needle leaved trees
🟩 3: Woody broadleaved deciduous trees
🟩 4: Woody broadleaved evergreen trees
🟫 5: Low-growing woody plants
🟩 6: Permanent herbaceous
🟨 7: Periodically herbaceous
🟪 8: Lichens & mosses
⬜ 9: Non and sparsely vegetated
🟦 10: Water
🟦 11: Snow & ice

IMD2018
☐ IMD value = 0%
🟨 IMD value: 1%-29%
🟥 IMD value: 30%-100%

Time-series of sealing
🟦 Sealed already in 2006
🟨 Built between 2006-2012
🟥 Built between 2012-2018

Figure 7.2.1.a: Comparison of harmonized (c) and non-harmonized (d) sealing time-series, as well as input status represented by CLC+ Backbone and Imperviousness 2018. New highway and bridge north part of Budapest. Hungary x,y= 5003535, 2762254

Typically, changes of linear elements at 20m resolution do not appear consistently, e.g. the new bridge (Megyeri bridge) was built in reality in 2008, should appear in yellow on (c) and (d) maps.

Besides visual checks, total sealed cover was calculated for the entire EEA39 area and compared to previous time-series statistics (Figure 7.2.1.b).

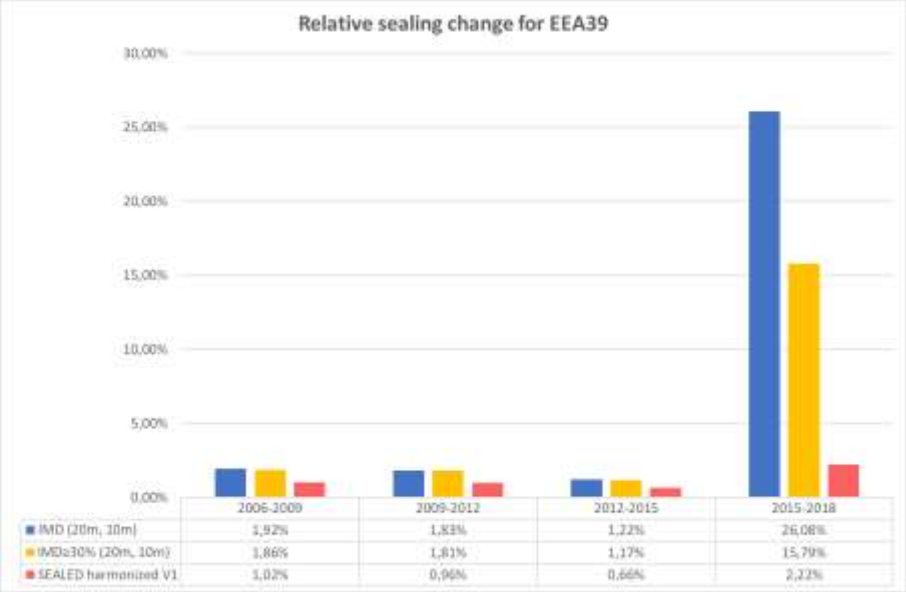


Figure 7.2.1.b: Changes of soil sealing relative to the previous reference year calculated for EEA39 area.

The comparison of pan-European statistics indicated promising improvements in the rate of changes:

- The relative change of soil sealing between 2015 and 2018 became significantly lower (2,22%), than by previous non-harmonized IMD based calculations;
- Additionally, the relative changes of sealing were showing slightly lower values the previously, still showing a decreasing rate over periods until 2015.

On the other hand, still a break is shown in decreasing rates between 2015 and 2018, even if significantly reduced.

7.2.2 Checking the credibility of soil sealing changes

All sealing related datasets considered (Imperviousness and CLC+ Backbone) are the result of individual complex classification procedure. While imperviousness changes are derived by a “post classification comparison” method including following main steps:

1. Individual classification of status considering certain reference years
2. Creating difference maps
3. Eliminating “noise” and keeping only meaningful changes

The last step may be supported by sophisticated GIS procedures, but by experience the production of reliable change information usually cannot exclude significant amount of manual (visual interpretation) work.

Previous validation and verification results (described with more detail in chapter 2) indicated, that the accuracy of Imperviousness change data is significantly lower, than the accuracy measured for imperviousness status.

Noise in land cover change data produced by “post classification comparison” method is typically appearing in small spots, as a consequence of small uncertainties in the classification of individual status. As no previous QA/QC results were found concerning the connection between the size of contiguous change patches and their credibility as a real change, additional visual quality checks were performed on Imperviousness Classified Change (IMCC) data in frames of present work.

All 20m resolution pan-European IMCC raster data were vectorized, and the distribution of the contiguous change areas were checked in the first step.

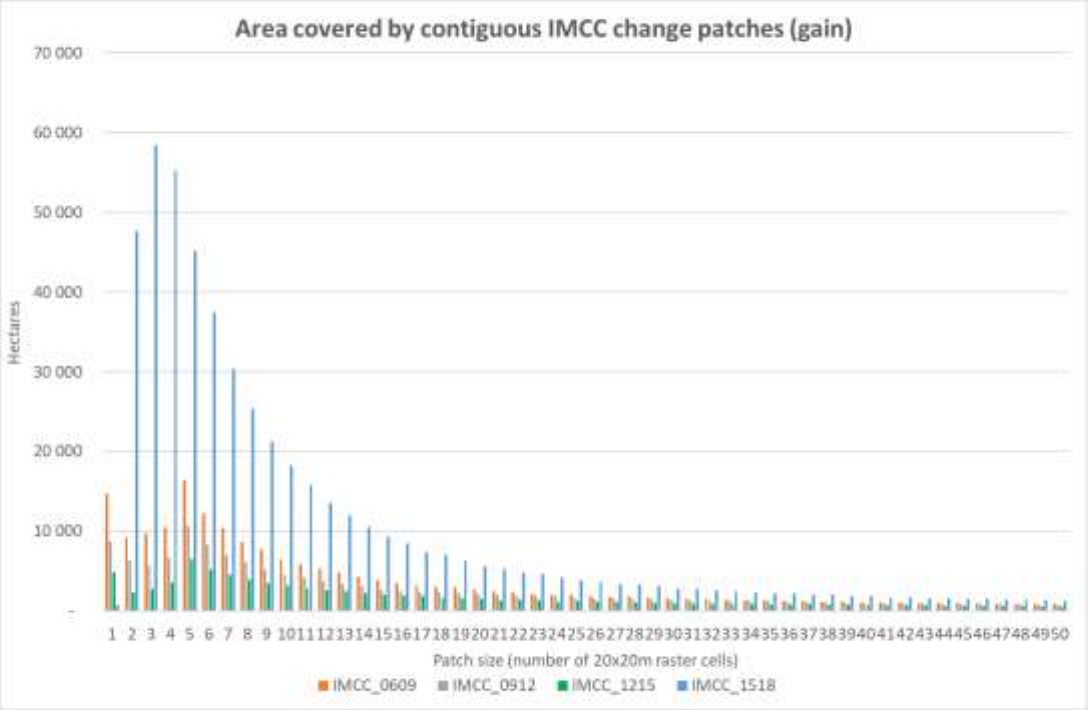


Figure 7.2.2.a: Distribution of contiguous IMCC change (new sealed cover) patches by patch size (enlarged, shows only patches smaller than or equal to 50 raster cells).

The examination of histograms leads to following main conclusions:

- IMCC changes in 2015-2018 period contain significantly more small change patches compared to all other periods. This may be explained by technical reasons because of resolution upgrade;
- In case of all periods, the largest total area is covered by small to very small contiguous change patches, while largest change patches cover only smaller part of the total changed area.

Limited remaining resources allowed to check only relatively few samples, therefore 100-100 random samples were checked in four ranges of patch size. Both IMCC class 1 (new sealed cover) and class 2 (loss of sealed cover) were checked, results are summarized in Table 7.2.a-b.

Table 7.2.2.a: Results of credibility checks for IMCC1518 new sealed cover class (4x100 samples)

New sealed cover 2015-2018	Number of contiguous patches	No real sealing change	Real sealing change	Only part of patch area is real change	No reference imagery found
Area < 0.5 ha	2 187 931	79,0%	6,0%	9,0%	6,0%
Area: 0.5ha - 1ha	127 703	55,0%	17,0%	25,0%	3,0%
Area: 1ha - 5ha	58 535	32,0%	33,0%	30,0%	5,0%
Area: > 5ha	4 831	13,0%	60,0%	25,0%	2,0%

Look & feel credibility check results indicate obvious link between IMCC patch size and quality, credibility of the small patches is significantly lower than the larger ones, while the largest number of contiguous new sealed cover patches is covered by small to very small contiguous change patches.

Only 6% of smallest new sealed cover spots (area < 0.5ha) proved to be real change, while largest patches (area > 5ha) proved to be more credible, 60% of these changes was indicated as realistic. Note, there is significant number of change patches, where part of the area is real change, another part is not a change.

Table 7.2.2.b: Results of credibility checks for IMCC1518 loss of sealed cover class (4x100 samples)

Loss of sealed cover 2015-2018	Number of contiguous patches	No real sealing change	Real sealing change	Only part of patch area is real change	No reference imagery found
Area < 0.5 ha	106 298	94,0%	3,0%	1,0%	2,0%
Area: 0.5ha - 1ha	715	73,0%	24,0%	2,0%	1,0%
Area: 1ha - 5ha	232	82,0%	11,0%	4,0%	3,0%
Area: > 5ha	6	66,7%	0,0%	0,0%	33,3%

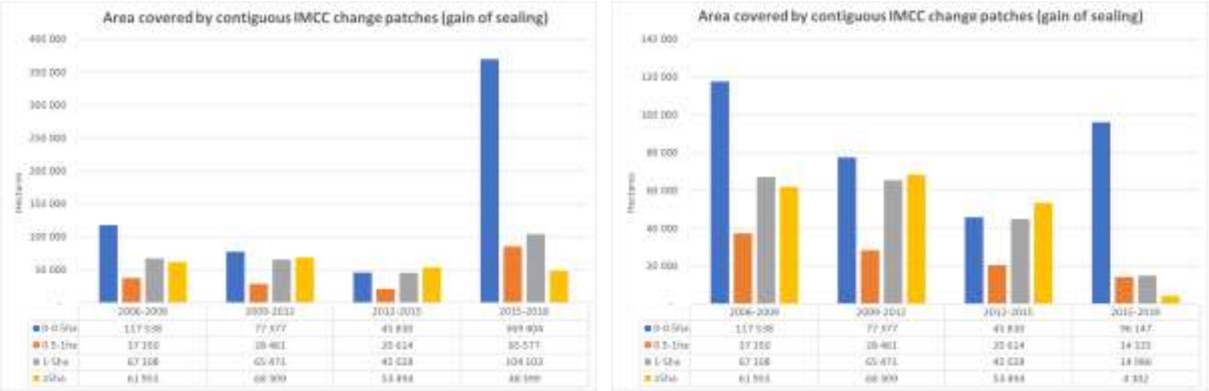
While the dominant process in EEA39 is the increase of sealed surfaces, the loss of sealed surfaces (like recultivation) is present, but rather rare. This is reflected in the small number of contiguous losses of sealed area patches as well, but especially notable the very low credibility of these changes shown by the look & feel check. Larges patches (area > 5ha) have appeared only at six locations, and none of these are real change. In case of IMCC1518 data there is relatively large number (106 298) of very small (area < 0.5ha) loss of sealed cover patches, but only 3% proved to be real change.

7.2.3 Improved harmonization of the soil sealing time-series

The significant statistical bias caused by the extreme amount of non-real changes (noise) in IMCC1518 data has made necessary the repetition of harmonization process with improved quality change data.

Improving the quality of IMCC1518 change data by filtering out non-real changes

Non-real changes appear similarly in IMCC data for all periods, even if the number of contiguous change patches is lower, than in 2015-2018 period. In order to filter out many non-realistic changes a complex filtering method was designed by Christophe Sannier. This methodology was applied only IMCC_1518 class "new sealed cover", key statistical results are demonstrated in Figure 7.2.3.a.



a) Original distribution of IMCC new cover

b) Complex filtering applied on IMCC 1518 new cover

Figure 7.2.3.a: Distribution of IMCC new cover (gain) areas by size ranges and periods (all EEA39 area).

The filtering algorithm has resulted the reduction of the sum of change area in all ranges. Total new sealed cover area was reduced by 78,6% from 607 483 ha to 129 771 ha.

Repeating the accounting process with filtered IMCC1518 data

The accounting process was applied again with the same conditions, except that IMCC_1518 new sealed cover was replaced by the filtered version. The process has resulted a new 10m resolution time-series of binary sealing maps signed as version 3 (V3)[7].

The complex filtering algorithm has resulted statistically more credible results than the previous V1 version, relative changes of soil sealing were further decreased from 2,2% to 0,7% for period 2015-2018.

Table 7.2.3.a: Total sealed areas expressed in SQKM calculated for EEA39 by comparing V1 and V3 results

YEAR	SEALED V1 status (sqkm)	SEALED V1 change (%)	SEALED V3 status (sqkm)	SEALED V3 change (%)
2018	175 664	2,22%	175 664	0,70%
2015	171 856	0,66%	174 437	0,65%
2012	170 730	0,96%	173 310	0,95%
2009	169 103	1,02%	171 684	1,01%
2006	167 388		169 968	

This time-series of 10m resolution binary soil sealing maps is one of final results of the present work, but not an ultimate solution, as:

- Only IMCC changes (new cover) in 2015-2018 period were filtered, while non-real changes appear in other periods and within IMCC1518 class 2 (loss of sealing) as well;
- Although the complex filtering algorithm applied on IMCC1518 class 1 data delivers statistically improved (more credible) results, still local issues were found by look & feel fast checks.

The reassuring solution for the future improvement of soil sealing time-series would be a general revision of all imperviousness change data, including possible visual control and manual corrections.

Aggregation of 10m binary maps to 100m resolution sealing density layers

The time-series of 100m resolution sealing density maps was created by aggregating the harmonized V3 10m binary soil sealing maps. The sealing density value corresponds to the sum of 10m resolution "sealed" cells covering the 100m resolution cell. Only fully covered 100m cells were considered, if NoData values exist for any of the cells the resulting 100m resolution cell was set to "No data".

This time-series (SEALING_DENSITY_100m.zip) is one of the deliveries aimed to be used as input by Task 2 (Landscape fragmentation), was uploaded to Teams folder.

Aggregation of 100m sealing density layers to 1km resolution sealing density layers

The time-series of 1km resolution sealing density layers was created by aggregating the harmonized V3 100m sealing density layers. The aggregation to 1km resolution was performed by calculation the mean sealing densities within each 1x1km cell, resulting sealing density values in a range 0.00-100.00%.

Note, while sealing density values of 100m resolution layers stored as integer values are enough to not loose statistical information compared to 10m resolution binary maps, 1km sealing density maps need to be stored as floating point numbers to keep the consistency of statistical information.

---

[7] Version 2 time-series was an experimental version, discontinued, not discussed in this report.

7.3 Performing statistical bias correction

The analysis of the reference sealing data collected on the basis of 1km x 1km Primary Sampling Units has concluded, that a non-linear trendline, described by the function **y=x+sin(pi/100*x)*a** is applicable to describe the relationship between aggregated EO based sealing values and reference estimates collected by visual interpretation.

The non-linear trend line is describing a relationship, where there is few or no difference between EO based and reference interpretation-based estimations close to sealing levels of 0% and 100%, while highest deviations from the main diagonal are shown for medium sealing levels (around 50%). The fitted non-linear trend line is characterized by the one single coefficient "a" showing the magnitude of the deviation – and consequently the estimated magnitude of the overestimation of the sealing over the entire area of interest (EEA39). This description of the relationship made possible to perform a statistical bias correction of EO based sealing data by applying the inverse of the fitted non-linear function.

7.3.1 Fitting non-linear trend line to all status layers in soil sealing time-series

Reference interpretation results interpreted for all five reference years were compared to corresponding aggregated 1km resolution status layers derived from the harmonized V3 version of soil sealing time-series, individual trend lines were fitted to describe the non-linear relationships. The value of coefficients "a" and the number of interpreted PSUs used is shown in Table 7.3.a.

Table 7.3.1.a: Coefficients "a" gained by fitting the non-linear function y=sin(pi/100*x) * a to model the relation between sealing level gained by V3 sealing time-series layers and sample based visual estimation of sealing levels

Reference year	2018	2015	2012	2009	2006
Coefficient "a"	9.9857	10.9639	11.6828	11.0839	11.2790
Number of samples used	927	825	900	619	512

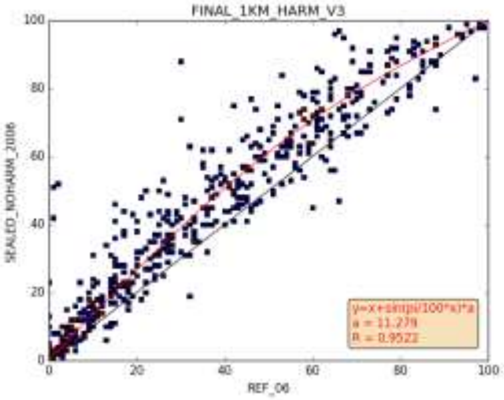
While the availability of appropriate reference imagery was satisfactory for year 2018, no reference imagery was available for a number of samples in case of earlier reference years, as only samples were used for comparison, where all the 100 points within the 1x1 km rectangle could be interpreted. Although the magnitude of individual coefficients "a" is similar, values show some variation partly caused by the number of interpretable PSUs.

7.3.2 Performing statistical bias correction by using individual correction parameters

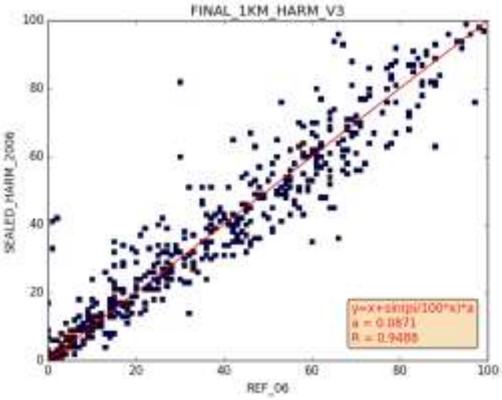
The inverse of the non-linear functions was used to correct sealing density values of 1km layers. This step was aimed to fit 1km sealing density values closer to reference.

As a non-linear function *y=x+sin(x)* has no direct analytical inverse, a numerical method was applied to calculate bias corrected sealing values based on individual corrections coefficients "a". Bias corrected raster layers were compared again to reference by fitting the same non-linear trendline.

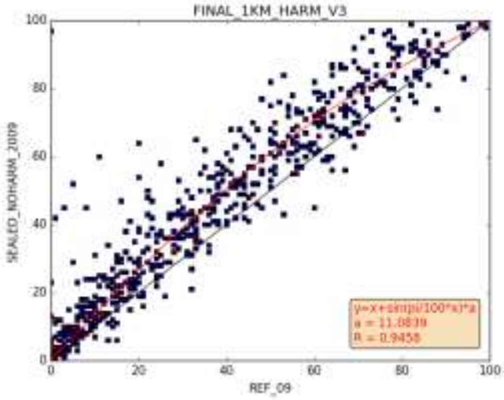
The applied bias correction has resulted close to linear relationships with low average deviation from the main diagonal (Figure 7.3.2.a).



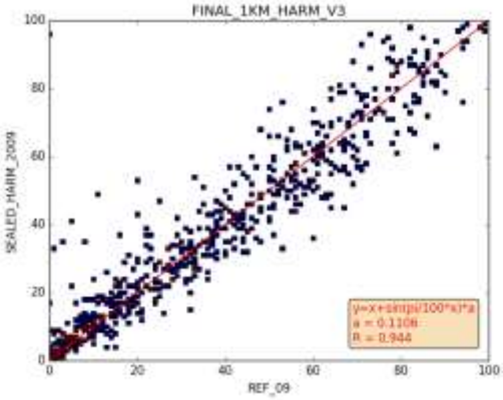
a) SEALED 2006 vs reference



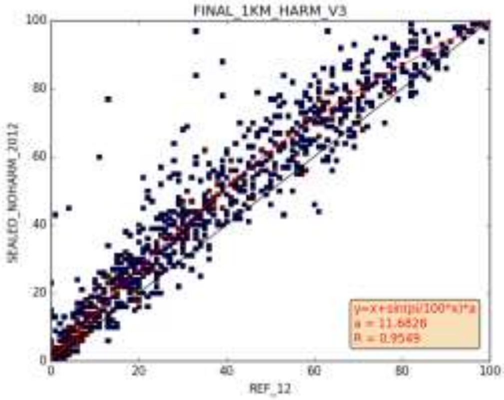
b) SEALED 2006 bias corrected



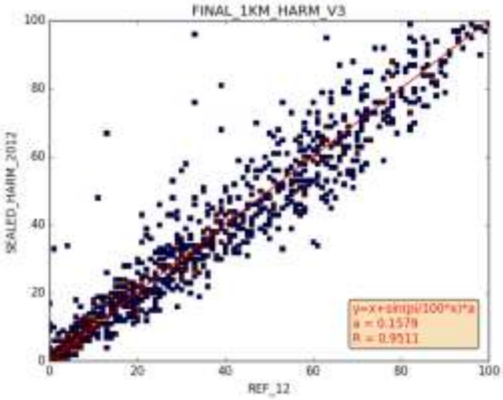
c) SEALED 2009 vs reference



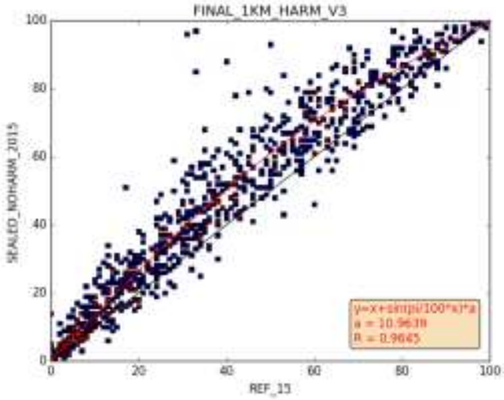
d) SEALED 2009 bias corrected



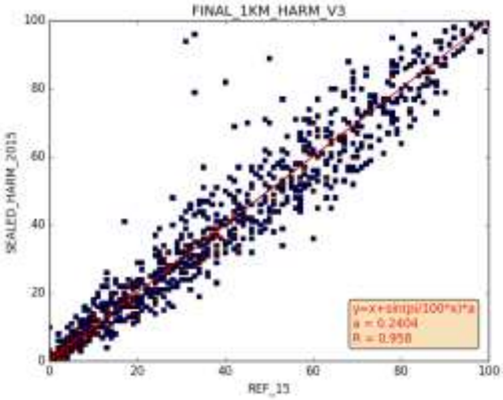
e) SEALED 2012 vs reference



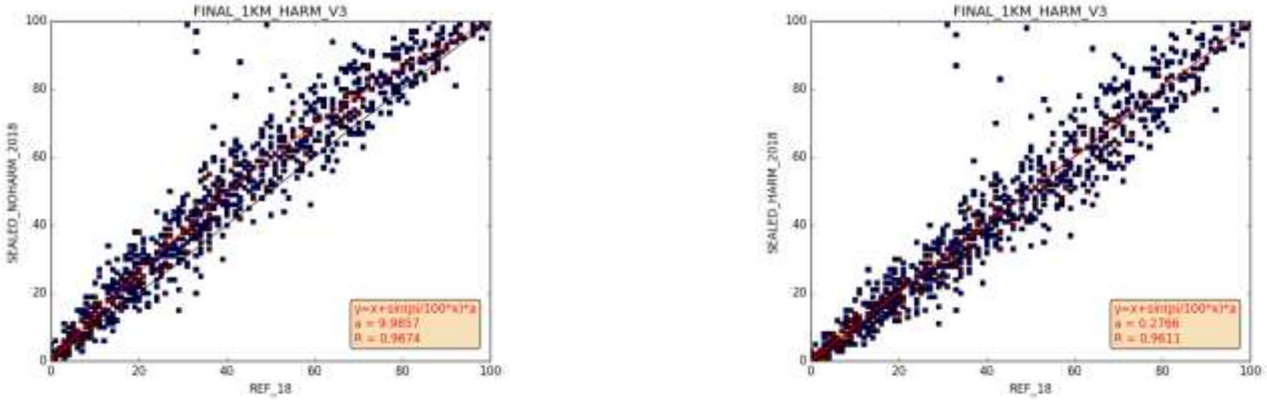
f) SEALED 2012 bias corrected



g) SEALED 2015 vs reference



h) SEALED 2015 bias corrected



i) SEALED 2018 V3 vs reference

j) SEALED 2018 V3 bias corrected

Figure 7.3.2.a: Comparison of 1x1 km aggregated sealing densities (left: without and right: with bias correction) to sealing estimations gained by reference interpretation. Red line indicates fitted non-linear function, black line indicates main diagonal.

Total sealed area and relative sealing changes were calculated for EEA39 area on the basis of bias corrected 1km resolution soil sealing maps and compared to previous non-bias corrected calculation results (Table 7.3.3.a).

As the consequence of the bias correction the total sealed area was reduced by 21-23% (Sealed V3 status compared to Sealed V3, bias corrected). On the other hand, the use of individual bias correction coefficients has resulted unlikely fluctuation of total sealed areas especially considering 2009 and 2012 reference years (sealed V3 change, bias corrected).

7.3.3 Performing statistical bias correction by using one single correction parameter

Finally, as a compromise the decision was to apply the same bias correction coefficient (a=9.9857, result gained for the reference year of 2018) for all reference years. This process has resulted the V3.1 harmonized, bias corrected soil sealing time-series in 1km resolution.

Table 7.3.3.a: Total sealed areas expressed in SQKM calculated for EEA39 from various versions of sealing related time-series.

YEAR	No bias correction applied				Bias correction with individual coefficients for every reference year		Bias correction, same coefficient was applied to all reference years	
	IMD≥30% status (sqkm)	IMD≥30% change (%)	SEALED V3 status (sqkm)	SEALED V3 change (%)	SEALED V3 status (sqkm)	SEALED V3 change (%)	SEALED V3.1 status (sqkm)	SEALED V3.1 change (%)
2018	152 880	15,79%	175 664	0,70%	139 306	2,90%	139 306	0,74%
2015	132 027	1,17%	174 437	0,65%	135 386	2,26%	138 279	0,68%
2012	130 504	1,81%	173 310	0,95%	132 396	-0,29%	137 348	1,00%
2009	128 183	1,86%	171 684	1,01%	132 786	1,48%	135 988	1,05%
2006	125 846		169 968		130 850		134 574	

IMD≥30%: Estimations based on binary sealing maps derived from imperviousness time-series by applying 30% threshold
SEALED V3: Similar to SEALED V1, but IMCC_1518 gain filtered by complex method (Christophe Sannier)
SEALED V3 corrected: Bias correction applied at 1x1km resolution (individual coefficients for each reference year)
SEALED V3.1 corrected: Bias correction applied at 1x1km resolution (same coefficients for all reference years)

Summary statistics based on the V3.1 version of harmonized soil sealing time-series shows a continuous increase of total sealed areas without a significant break. The rate of relative changes is decreasing between 2006-2015. There is still a slight increase in the rate between 2015-2018 (from 0,68% to 0,74%).

This 1km resolution V3.1 time-series was considered as a final delivery aimed to be used to derive aggregated sealing statistics. The time series was uploaded to Teams folder in two different, but statistically equivalent and consistent versions:

- SEALING_DENSITY_HARMONIZED_V3_1.zip: Sealing density layers in 1km resolution (floating point)
- SEALED_HARMONIZED_SQM_1km_V3_1.ZIP: Raster layers containing values of estimated sealed area in SQM for each 1x1km raster cell.

8. Assessment of Imperviousness vs CLC+ Backbone as input to land take indicator, supporting the 8th EAP

Within the European context, two high level policy targets directly deal with the issues of soil sealing and land take:

- the EU Roadmap to Resource Efficient Europe, which demands “no net land take until 2050”;
- the UN Sustainable Development Goal 15.3, which aims to “halt and reverse land degradation” until 2030 and which in 2017 introduced the concept of “Land Degradation Neutrality”, which the EU and its Member States have pledged to integrate.

While these targets highlight the issue of efficient land use, they do not involve specific implementation mechanisms.

Land take is also known as “urbanisation”, “increase of artificial surfaces” and represents an increase of settlement areas (or artificial surfaces) over time, usually at the expense of rural areas. This process can result in an increase of scattered settlements in rural regions or in an expansion of urban areas around an urban nucleus (urban sprawl). A clear distinction is usually difficult to make (Prokop et al., 2011).

Land take indicator for entire EEA39 area was calculated based on Corine Land Cover (CLC) data.

By the latest definition of EEA indicator list the land take indicator in cities and commuting zones is currently calculated from the Urban Atlas dataset of the Copernicus Land Monitoring Service for the years 2012 and 2018. Changes from agriculture, forest and semi-natural/natural land, wetlands or water to urban areas are grouped and expressed in km2 of converted area.

Soil sealing can be defined as the 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 and is essentially an irreversible process (Prokop et al., 2011).

The indicator of Imperviousness and imperviousness change was initially calculated from the time-series of 100m resolution HRL Imperviousness data.

8.1 Key differences between land take and imperviousness change indicators

Land take and imperviousness change indicators are aiming to measure similar processes and include overlaps in the definitions, but there are some key differences. While imperviousness change is aiming to map the changes in artificial sealed cover status only, land take includes additional cases where the reduction of natural soil cover is realized e.g. as new open mine areas, or even as new permanent water surfaces.

Moreover, land take is considering not only the change of the artificial status of the land, but provides more information about the processes e.g. by source LCLU status of consumption.

Additionally, any indicator values are influenced substantially by the characteristics of input data as thematic, spatial and temporal resolution, surveying method, accuracy and other parameters. For example, CLC based land take indicator maps the total area of a new airport or other industrial unit as land take considering the 5ha Minimum Mapping Unit defined for CLC changes, while imperviousness change indicator maps only the area of a new runway, buildings and other sealed surfaces, but does not consider remaining non-sealed parts.

8.2 Assessment of Imperviousness and CLC+ Backbone data as input

Strengths and weaknesses of available CLMS products as input to land take indicator are summarized in Table 2.8.a.

Table 8.2.a: Strengths and weaknesses of available CLMS products as input to land take indicator

Dataset	Strengths	Weaknesses
Traditional CLC	- Full EEA39 cover - Long time-series (back to 2000) - High quality, visually interpreted changes - High thematic resolution - Land use information included	- Low spatial resolution (5ha MMU for changes) - Only 6 years update-frequency
Urban Atlas	- Relatively high spatial resolution - Relatively long time-series (back to 2006) - 3 years update period form 2018 (planned) - High quality, visually interpreted changes - Relatively high thematic resolution - Land use information included	- Available only for priority areas - Only 6 years update period between 2006-2018 - Different spatial resolution used by change types: • Urban to urban: 0.1 ha; • Rural/natural to urban: = 0.1 ha; • Rural/natural to rural/natural: 0.25 ha; • Urban to rural/natural: 0.25 ha
CLC+ Backbone	- Full EEA39 / EEA38 cover - High spatial resolution (10m) - High update frequency (3→2 years) - Relatively high thematic resolution	- Short time-series (back to 2018 only) - Only land cover information (no land use) - No dedicated change layer specified
HRL Imperviousness	- Full EEA39 / EEA38 cover - High spatial resolution (20m / 10m) - Relatively long time-series (back to 2006) - High update frequency (3 years) - Dedicated change data available	- Maps only one single land cover theme (artificially sealed areas) - Time-series harmonization issues due to resolution upgrade by 2018 - Relatively low credibility of change information
New harmonized soil-sealing time-series	- Full EEA39 cover (up to 2018) - High spatial resolution (down to 10m) - Relatively long time-series (back to 2006) - High update frequency (3 years) - Harmonized time-series - Statistical bias-correction applied (only at 1km resolution)	- Maps only one single land cover theme (artificially sealed areas) - Feasibility of future (after 2018) harmonized and bias corrected update not clarified yet

Based on the strength and weaknesses of currently available CLMS products one can conclude, that none of the products is able to provide alone a pan-European high-resolution time-series of land take.

On the other hand, the wise combination of these datasets may provide a satisfactory short-term solution.

9. Preparation of an outline of a scientific paper addressing the bias correction method for submission in 2024

A first draft of the scientific paper was submitted to contributing authors on 5 January. There is a pending issue on the selection of an appropriate journal to which the paper should be submitted. Initially it had been decided to submit the manuscript to the "journal of Land Use Policy" https://www.sciencedirect.com/journal/land-use-policy, but looking at the scope and time from submission to acceptance (**350 days**, 7.1 Impact Factor), it was now decided to submit it to "Ecological Indicators" (**94 days** on average from submission to acceptance and 6.9 impact factor) https://www.sciencedirect.com/journal/ecological-indicators which seems to be a better choice also from the scope of the journal. In addition, "Ecological Indicators" also have a partner journal "Environmental and Sustainability Indicators" https://www.sciencedirect.com/journal/environmental-and-sustainability-indicators which could be an alternative should the editor find our article not suitable for "Ecological Indicators" with similar lead time to acceptance and still a reasonable impact factor (**105 days** and 4.3 impact factor).

Authorship was discussed and Joachim Maes, Lewis Dijkstran and Javier Gallego have agreed to be included as authors in addition to Christophe Sannier, Gergely Maucha and Eva Ivits.

References

CLMS data specifications

Product specification for re-analysed time series of Imperviousness 2006-2015: High Resolution Layer Imperviousness Date: 10/03/2016

Product user manual Imperviousness Lot1: Imperviousness 2018, Imperviousness Change 2015 – 2018 and Built-up 2018 Document version: 2.2

Product specification and user manual: CLC+ Backbone raster product Issue: 2.1a Date: 06/09/2022

CLMS validation reports

HRL Imperviousness degree 2015 validation report Issue: 1.3 Date: 28/02/2019

HRL Imperviousness degree 2018 validation report Issue: 1.3 Date: 24/11/2020

ETC reports

Guidelines for verification of High Resolution Layers produced by the CLMS (Copernicus Land Monitoring Service) as part of the 2018 reference year productionSC58651 (contract 9) Task 9 Assessment of MS verification results HRLs 2018) Version 1.3 Date: 10/02/2021 ETC/ULS report

SC58651 Task 9 Assessment of MS verification results (HRLs 2018) - IMD 2018 Date: 17.12.2021 ETC/ULS report

SC58651 Task 9 Assessment of MS verification results (HRLs 2018) - HRL IMCC 2015-2018 Date: 17.12.2021 ETC/ULS report

SC59032 TASK 14 - Subtask 1 Synergies between CLC+ Backbone and other CLMS products Final report v2 Date: 07/07/2023 ETC/DI report

SC59032 TASK 14 - Subtask 4 Analysing the threshold values that would allow such comparability of the HRL 2015 and 2018 data Final report Date: 13/06/2023 ETC/DI report

SC90032 TASK 14 - Subtask 9 Pixel counting" vs. statistics discussion support Final report Date: 13/06/2023 ETC/DI report

EEA indicators

Imperviousness and imperviousness change in Europe

Net land take in cities and commuting zones in Europe

Articles

Prokop G, Jobstmann H, Schonbauer A (2011) Report on best practices for limiting soil sealing and mitigating its effects. European Commission, Brussels. doi:10.2779/15146

European Topic Centre on
Data integration and digitalisation
https://www.eionet.europa.eu/etcs/etc-di

The European Topic Centre on Data integration and digitalisation (ETC-DI) is a consortium of European institutes under contract of the European Environment Agency.

European Environment Agency
European Topic Centre
Data integration and digitalisation