



PROGRAMME OF
THE EUROPEAN UNION



Copernicus Land Monitoring Service – High Resolution Layer – Imperviousness 2021

ALGORITHM THEORETICAL BASIS DOCUMENT (ATBD)



Date: 30.07.2025

Doc. Version: 1.4



Document history

Revision	Date	Short description of changes
1.4	30.07.2025	Initial published version



Contents

1	Executive Summary.....	7
2	Background of the document	8
2.1	Scope.....	8
2.2	Content and structure	8
3	Project Overview.....	8
4	Product Methodology and Workflow	10
4.1	Input data collection and pre-processing	13
4.1.1	Satellite input data	13
4.1.2	Sentinel-2-based input features.....	14
4.1.3	VHR data	14
4.1.4	Reference databases	15
4.1.5	Other ancillary data.....	15
4.2	General methods used in production.....	15
4.2.1	Pre-processing	16
4.2.2	Root mean squared deviation	17
4.2.3	Deep learning	18
4.2.4	Aggregation methods	19
4.2.5	Intermediate layer: Change mask (CHM)	20
4.3	Production of the Imperviousness Density (IMD)	22
4.3.1	Method selection, training, and prediction.....	23
4.3.2	Mapping of Imperviousness Density	23
4.3.3	Mapping of United Kingdom in HRL Imperviousness 2021	26
4.4	Production of the Impervious Built-Up (IBU).....	27
4.4.1	Model training and prediction	27
4.4.2	Post-processing	28
4.5	Change layer generation.....	29
4.5.1	Generation of continuous change layers	29
4.5.2	Generation of the classified change layers	30
4.5.3	Imperviousness Change (IMDC) and Imperviousness Density Change Classified (IMCC)	30
4.5.4	Impervious Built-Up Change (IBUC)	31
4.6	Derived layers based on aggregation	31
4.6.1	Generation of the 100m status layers (IMD 2021, SBU 2021)	32
4.6.2	Generation of 100m change layers (IMDC 1821, SBUC 1821)	32
4.7	Confidence layers.....	32
5	Recognized technical issues	33



6 Terms of use and product technical support.....	35
6.1 Terms of use	35
6.2 Citation.....	35
6.3 Product technical support	35
7 References	36
List of abbreviations	37
Annex 1 - Cloud mask weighting	40

List of Figures

Figure 1: High-level overview of the input-data and final products and layers of HRL Imperviousness 2021	9
Figure 2: The HRL Imperviousness 2021 processing graph including all interim and final layers created during the processing	11
Figure 3: Simplified workflow for the calculation of quantiles	16
Figure 4: Simplified workflow for the calculation of harmonic coefficients	17
Figure 5: Subgraph showing the dependencies for the Root Mean Squared Deviation (RMSD) calculation from the pre-processed data.....	17
Figure 6: Calculation of the Root Mean Squared Deviation (RMSD) from the harmonic coefficients	18
Figure 7: Overview of the inputs used for the U-Net prediction	18
Figure 8: General approach for the mean status aggregation to lower resolutions (R)	19
Figure 9: General approach for the aggregation of changes to lower resolutions (R)	20
Figure 10: General approach for the maximum aggregation to a lower resolution (R)	20
Figure 11: Creation of a unified change mask (CHM). This mask is derived from the RMSD calculation by applying a thresholding approach followed by a sieving	21
Figure 12: Creation of the change mask using a thresholding approach and sieving.....	21
Figure 13: Processing graph for the IMD 2021.....	22
Figure 14: Derivation of a sealing mask based on a combination of IMD and CLCplus Backbone	24
Figure 15: Retrieval of regression coefficients to calibrate the imperviousness outline on the mean NDVI values (NDVI p_0).....	25
Figure 16: The Impervious Built-Up (IBU) 2018 layer is used as a mask to find the mean predicted sealing density of these high-density sealing areas.....	25
Figure 17: An optimization algorithm used to find a new minimum value for rescaling such that the mean predicted density of 2021 remains consistent with the values from 2018	26
Figure 18: Algorithm used for producing the IMD 2021 layer	26
Figure 19: The processing graph to produce the IBU layer.....	27
Figure 20: Examples of built-up prediction in a test area around Poiniers, France. From left to right: (1) The previous IBU product (2018), (2) a prediction carried out with a model trained on GHSL, (3) a prediction carried out with a model trained on OSM, (4) the ground truth, based on a rasterized version of OSM building footprints	28
Figure 21: IBU 2021 created using a threshold on the Raw Imperviousness Built-up (RIBUD) layer	28
Figure 22: General workflow to compute change and change support layers for adjustable resolutions.....	29
Figure 23: Full processing graph to obtain IMDC and IMCC products along with their support layers (IMDCS and IMCCS). The resampling resolution is set to R = 20m	31
Figure 24: Full processing graph to obtain IBUC along with the support layer (IBUCS). The resampling resolution is set to R = 20m.....	31
Figure 25: Processing graph of the 100m aggregated IMD 21 layer. The target resolution is set to R = 100m	32
Figure 26: Processing graph for the creation of the SBU layer as 100m mean aggregation from the 10m IBU layer.....	32



List of Tables

Table 1: Abbreviations and short description of all interim and final layers	12
Table 2: List of ancillary data used for the production of HRL Imperviousness 2021.....	15
Table 3: Overview of parameterizations used to create a specific reference year	16
Table 4: Overview of the model predictions.....	19
Table 5: Rule base to create the Imperviousness Change Classified (IMCC), Impervious Built-Up Change (SBUC) and Share of Built-Up Change Classified (SBCC) layers	30
Table 6: Summary of weights applied during the pre-processing. Class code and class name refer to the Sen2Cor Scene-classification Layer (SCL)	40



1 Executive Summary

Copernicus is the Earth Observation component of the European Union's Space Program. It offers space-based Earth observation and in situ (non-space) data as well as thematic data products. Copernicus data and products are freely and openly accessible to its users and six thematic Copernicus services (atmosphere monitoring, marine environment monitoring, land monitoring, climate change, emergency management and security) offer various operational thematic data products and services.

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

The High-Resolution Layer (HRL) Imperviousness is part of the pan-European CLMS portfolio that currently covers the EEA38+UK countries. It consists of two products: Imperviousness and Impervious Built-Up, along with their derived and supporting layers.

All layers contained here are derived from high-resolution optical satellite image time series (Sentinel-2) via automatic image processing methods and provide dedicated information on Impervious and Built-Up surfaces during the reference year 2021 (i.e. status) and detected dynamics between 2018 and 2021 (i.e. change). The aim of these layers is to provide reliable status and updates on artificial sealing characteristics across Europe to facilitate environmental monitoring applications, regional and transnational analyses and, generally, to support decision-making that is based on spatial evidence.

This Algorithm Theoretical Basis Document (ATBD) describes and justifies the algorithms used to generate the HRL Imperviousness 2021 by describing each element of the processing, allowing users to understand the details of the products.



2 Background of the document

2.1 Scope

The Algorithm Theoretical Basis Document (ATBD) provides in-depth technical insights and explanations of the underlying methodologies of HRL Imperviousness 2021. It is intended for users such as data scientists, developers, and researchers who need a detailed technical understanding. It does not serve general informational purposes like describing product characteristics, usage, or support details, these are covered in the Product User Manual (PUM) instead.

2.2 Content and structure

The document is structured as follows:

- Chapter 3 gives an overview of the project
- Chapter 4 specifies production methodology and its workflows
- Chapter 5 summarizes the recognized technical changes
- Chapter 6 lists the relevant references

3 Project Overview

The aim of this HRL Imperviousness 2021 production is the update and continuity of HRL Imperviousness, the contained Impervious Built-Up and their respective derived layers for the reference year 2021. A generalized overview of the products and layers from this production is given in Figure 1 below. The workflow encompasses data pre-processing, classification, and postprocessing to ensure the information remains accurate and relevant.

The technical aspects of the data processing and algorithmic methodologies used in this project are discussed in detail in the following chapters. These sections provide comprehensive insights into the scientific foundations, the computational steps involved, and the measures taken to maintain robustness and improve the accuracy of the production of the 2021 reference year.



PROGRAMME OF
THE EUROPEAN UNION

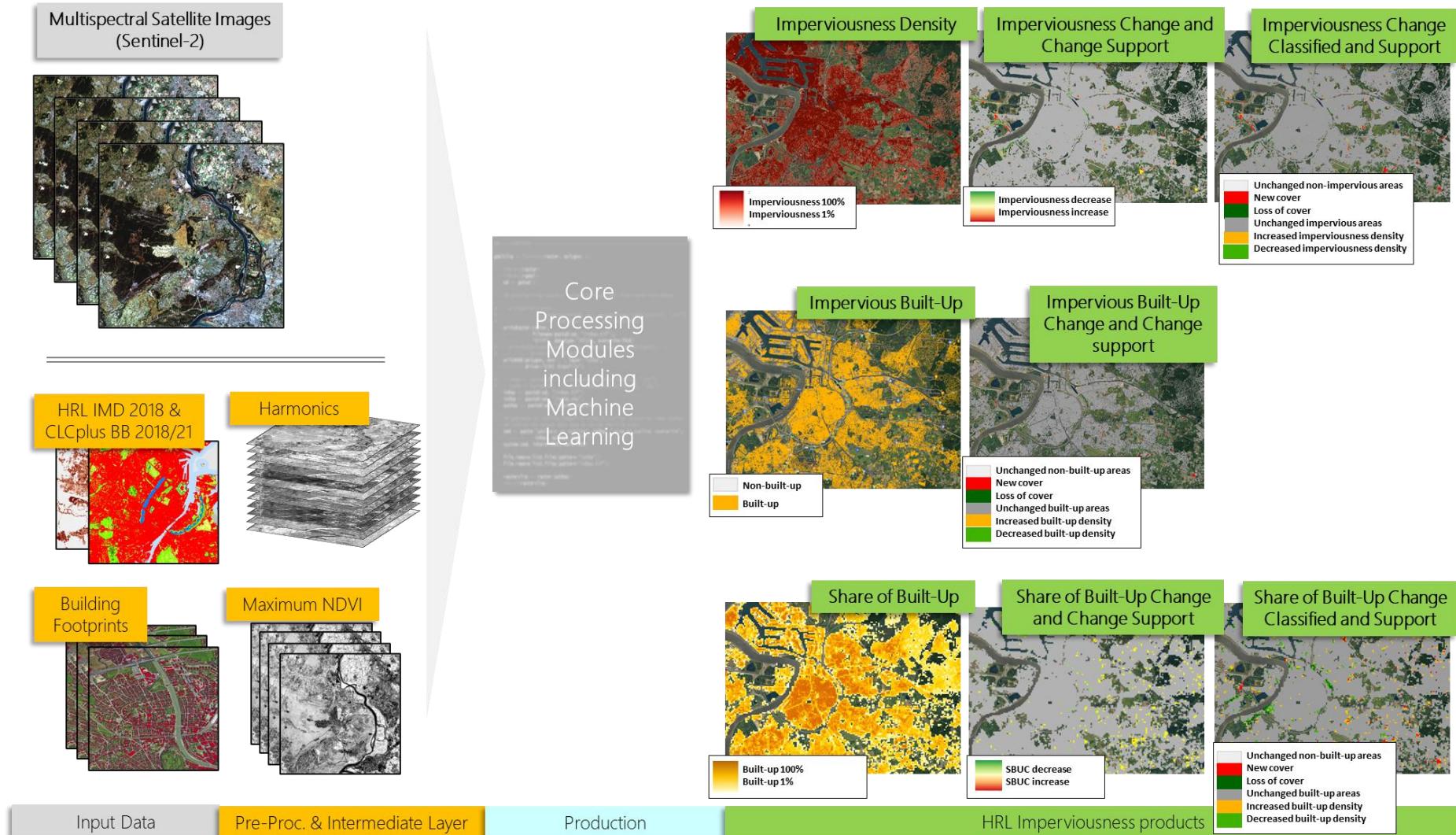


Figure 1: High-level overview of the input-data and final products and layers of HRL Imperviousness 2021



4 Product Methodology and Workflow

This chapter describes the applied approaches to derive the HRL Imperviousness 2021 products in detail. The production of HRL Imperviousness 2021 combines state-of-the-art image and time-series analysis techniques to leverage the full potential of the dense Sentinel-2 time series. While the overall generic production workflow is based on the process applied in 2018, notable methodological improvements have been introduced. These enhancements include a refined approach to classify sealing outlines and built-up areas, distinguishing this production from the previous HRL Imperviousness 2018 workflow. Figure 2 below provides a schematic representation of the workflow, which is further detailed in the subsequent chapters. It effectively illustrates the full dependency tree of the project. Each component of the workflow has been carefully designed to maximize the accuracy and reliability of the derived layers while maintaining a clear topological order of all processing steps. Subgraphs for target products are provided and explained in the respective chapters of this document. All abbreviations used in Figure 2 are described in Table 1 below.

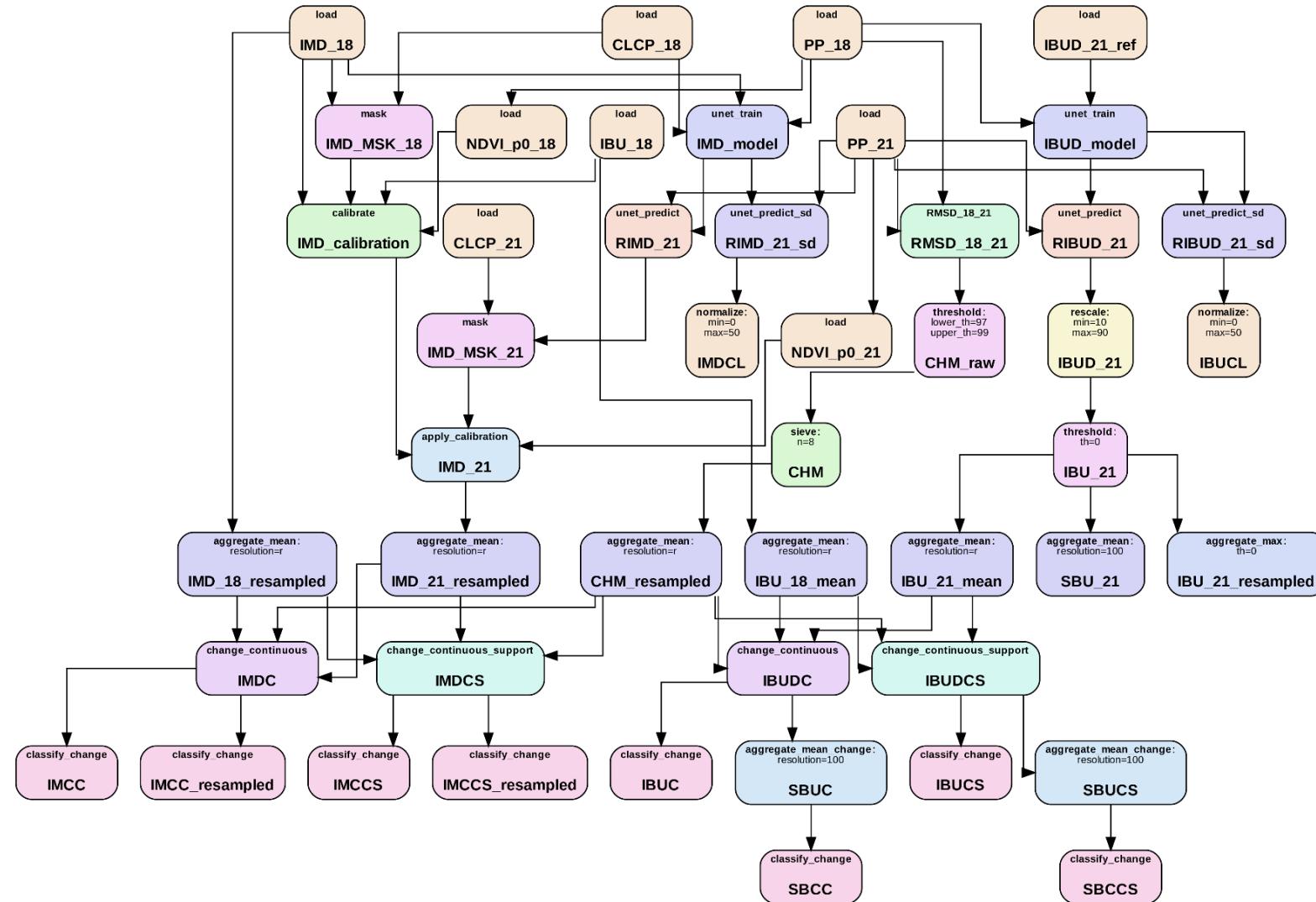


Figure 2: The HRL Imperviousness 2021 processing graph including all interim and final layers created during the processing

**Table 1: Abbreviations and short description of all interim and final layers**

Abbreviation	Description	Type
CHM	Refined change mask with sieving applied	Change detection input
CHM_raw	Threshold RMSD_18_21 to create a change mask	Change detection (interim)
CHM_resampled	Aggregated version of the change mask (mean aggregation)	Change detection input
CLCP_18	CLCplus Backbone 2018	Reference
CLCP_21	CLCplus Backbone 2021	Reference
IBU_18	2018 built-up (binary)	Reference
IBU_18_mean	Aggregated version of the binary built-up 2018 (mean)	Reference
IBU_21	Impervious Built-Up	Status, final output (10m)
IBU_21_mean	Mean aggregation of IBU_21 to lower resolution	Interim layer
IBUC	Impervious Built-Up Change	Classified change, final output (20m)
IBUCL	Confidence layer for Impervious Built-Up	Confidence, final output (10m)
IBUCS	Impervious Built-Up Change Support layer	Classified change, final output (20m)
IBUD_21	Calibrated built-up-densities derived from raw predictions	Interim layer
IBUD_21_ref	Reference built-up density	Reference
IBUD_21_resampled	Aggregate mean of IBUD to lower resolution	Interim layer
IBUD_model	Model coefficients for the U-NET model used to predict built-up degree	Model coefficients
IBUDC	Continuous change of built-up degrees	Continuous change
IBUDC_resampled	Aggregate mean of IBUDC to lower resolution	Continuous change
IBUDCS	Continuous change support of built-up degrees	Continuous change
IBUDCS_resampled	Aggregate mean of IBUDCS to lower resolution	Continuous change
IMCC	Imperviousness Change Classified	Final output (20m)
IMCC_resampled	Classify change using IMDC_resampled	Change using
IMCCS	Imperviousness Change Classified Support layer	Final output (20m)
IMCCS_resampled	Classify change support using IMDCS_resampled	Change support
IMD_18	2018 Imperviousness Density	Reference
IMD_21	Imperviousness Density	Status, final output (10m)
IMD_21_resampled	Imperviousness Density, 100m	Status, final output (100m)
IMD_MSK_18	Mask IMD_18 and CLCP_18 for 2021 imperviousness density calibration	Binary mask
IMD_MSK_21	Mask RIMD_21 and CLCP_21 for 2021 imperviousness density calibration	Binary mask
IMD_calibration	Calibration coefficients to map NDVI to IMD	Calibration Coefficients (NDVI vs IMD)
IMD_model	Model coefficients for the U-NET model used to predict imperviousness density	Model coefficients
IMDC	Calculate continuous change of imperviousness density	Continuous change
IMDC_resampled	Imperviousness Change	Continuous change, final output (20m, 100m)
IMDCL	Confidence layer for Imperviousness Density	Confidence, final output (10m)
IMDGS	Calculate continuous change support of imperviousness density	Continuous change, final output (20m, 100m)



IMDCS_resampled	Imperviousness Change Support layer	Mean of IMDCS
NDVI_p0_18	NDVI reference for 2018 (extracted from PP_18)	Timeseries parametrization
NDVI_p0_21	NDVI reference for 2021 (extracted from PP_21)	Timeseries parametrization
PP_18	Sentinel-2 pre-processing: Harmonic coefficients and quantiles for 2018	Timeseries parametrization
PP_21	Sentinel-2 pre-processing: Harmonic coefficients and quantiles for 2021	Timeseries parametrization
RIBUD_21	Direct predictions from the U-NET model for 2021 (built-up)	Raw prediction
RIBUD_21_sd	Standard deviations of RIBUD	Standard deviation
RIMD_21	Direct predictions from the U-NET model for 2021 (imperviousness)	Raw prediction
RIMD_21_sd	Standard deviations of RIMD	Standard deviation
RMSD_18_21	Root mean squared deviation between 2018 and 2021 (Harmonics)	Change detection (interim)
SBCC	Share of Built-Up Change Classified	Classified change, final output (100m)
SBCCS	Share of Built-Up Change Classified Support layer	Classified change, final output (100m)
SBU_21	Share of Built-Up	Status, final output (100m)
SBUC	Share of Built-Up Change	Continuous change, final output (100m)
SBUCS	Share of Built-Up Change Support layer	Continuous change, final output (100m)

4.1 Input data collection and pre-processing

The mapping and processing of HRL Imperviousness 2021 primary products heavily rely on high-resolution time series Earth Observation (EO) data from the Copernicus Sentinel missions. This section outlines the role of Sentinel-2 imagery, the challenges associated with optical data, and the rationale behind the chosen methodologies for product generation.

4.1.1 Satellite input data

High-Resolution time series EO data from the Copernicus Sentinel missions form the core for the mapping and processing of the HRL Imperviousness 2021 primary products. Sentinel-2 is nowadays a well-established source of optical time-series imagery. The combined dataset of the twin-satellite constellations provides a 5-day revisit time at the equator with even shorter revisit times towards the Poles. Clouds (and cloud shadows) remain a significant drawback for optical sensors, reducing the number of clear sky observations. Although a combination with high-resolution Synthetic Aperture Radar (SAR) data from the Sentinel-1 satellite constellation – specifically the Interferometric Wide (IW) swath Ground Range Detected (GRD) products – is possible when applying the processing methodology, only optical data was considered for the production. The reason for this is that the optical data already leads to highly accurate output products, reducing dependencies on the overall production workflow.

The HRL Imperviousness 2021 products were generated using Sentinel-2 Level-2A (L2A) data, which provides Bottom-Of-Atmosphere (BOA) reflectance images. The classification process utilized data from the reference period of January 1, 2021, to December 31, 2021, while training was conducted using data from the same period in 2018. The full annual time series for 2018/2021 was used, with all scenes downloaded from the Amazon Web Services (AWS) cloud platform. The L2A data was used directly as provided on AWS.



The L2A product is derived from the associated Level-1C (L1C) products using the Sen2Cor processor [1]. Sen2Cor has undergone several improvements over time, which are reflected in different processor versions and documented through corresponding Processing Baselines (PBs). As a result, the L2A dataset used in this production spans multiple Sen2Cor processing baselines. The reflectance value shift introduced with Processing Baseline 04.00 [2] was consistently accounted for by adjusting the affected data to align with the value range of pre-shift versions, thereby ensuring continuity across the time series.

Recently, ESA started a reprocessing campaign to create a harmonized L2A time series (Collection-1 Processing Baseline); however, this harmonized input product was not available at the time of production. Succeeding data reduction is done using statistical and time-series parameterization-based methodologies and described below in Section 4.1.2.

4.1.2 Sentinel-2-based input features

The dense optical time series offered through the Sentinel-2 satellite constellation delivers unprecedented observation capacity. However, this also requires performant data reduction procedures to efficiently extract the relevant features associated with impervious areas. Two main approaches can be considered for data reduction: (1) extracting temporal statistics based on seasonal to annual data (this also refers to a cloud-free best-pixel image composite), and (2) applying a time-series modelling approach to reduce the information content into a pre-defined set of model coefficients.

For this HRL Imperviousness production, features derived from both approaches were used in combination. This dual strategy was adopted to take advantage of the complementary information provided by each method. Specifically, statistical features, such as band quantiles, were computed for the 10 m resolution input bands (i.e., B02, B03, B04, B08) and for B12 over the defined reference period, to capture spectral characteristics of the surface. Concurrently, the temporal behavior of the landscape was characterized by using a modelling approach based on a combination of polynomial and harmonic functions fitted to the time series of the Normalized Difference Vegetation Index (NDVI). This modeling captures the temporal dynamics of vegetation and phenology, which can be critical in distinguishing between impervious and non-impervious surfaces over time.

The features generated from both the statistical and time series modelling approaches were jointly fed into the classifier. This integrated feature set enhances the classification process by providing both robust spectral summaries and detailed temporal signatures, improving the overall accuracy and stability of the impervious surface classification.

4.1.3 VHR data

The ESA Copernicus Space Component Data Access (CSCDA) dataset VHR_IMAGE_2021 was mainly used for quality control and product verification. This dataset consists of one cloud-free Very-High Resolution (VHR) optical coverage of EEA38+UK territory (including all islands and French Overseas Departments) acquired within the vegetation seasons in 2020-2021-2022. This dataset is the result of VHR acquisition from several contributing missions:

- Primary satellite missions: Pleiades1A/1B, SuperView-1, WorldView-2, WorldView-3, Kompsat3/3/GeoEye-1
- Back-up satellite missions: SPOT-6/7 (TrueSharp), TripleSat, Geosat-2, Vision1

The dataset is available in two processing levels: Level 1 and Level 3 (ortho-rectified) provided in European projection (ETRS-89, EPSG:3035) except for French Overseas Department (Local



UTM projections) at 2-4m spatial resolution for four spectral bands (blue, green, red, near-infrared) [3]. For this activity, Level 3 data was used.

4.1.4 Reference databases

The **reference database**, (see Product User Manual 2015 [11] for more information) as used for the absolute calibration of the Imperviousness Densities 2015 and 2018 layers, was not updated and used to produce the 2021 layer.

However, the **product verification database**, which was set up independently for the 2015 and 2018 production, has been updated and used for the thematic assessment of HRL Imperviousness 2021 layers.

4.1.5 Other ancillary data

In the production of HRL Imperviousness 2021, most dependencies on various in-situ datasets were successfully reduced. For the Impervious Built-Up, only building footprints (such as OSM, Microsoft buildings, and Google building footprints) are required, streamlining the process while maintaining high quality.

Similarly, for the Imperviousness layers, reliance is solely on the HRL Imperviousness 2018 and CLCplus Backbone 2018 for classifier training as well as CLCplus Backbone 2021 for IMD calibration (see 4.3.2). In addition, the same building footprints as used for the Impervious Built-Up production serve as input for calibration to retrieve samples of 100% IMD.

This refinement has simplified the workflow and increased efficiency without compromising the quality of the output products. All datasets used in production are summarized in the Table 2.

Table 2: List of ancillary data used for the production of HRL Imperviousness 2021

Name	Type	Source/Reference
OSM buildings	Building footprints	OSM (from Geofabrik , accessed 12/2023)
MSB buildings	Building footprints	MSB (processing date 25.04.2023)
Google open buildings	Building footprints	Open-buildings v3
CLCplus Backbone 2018 (Class 1)	Imperviousness reference	CLCplus
Imperviousness Density (IMD) 2018	Imperviousness reference	HRL Imperviousness
CLCplus Backbone 2021 (Class 1)	Imperviousness calibration	CLCplus
Impervious Built-Up (IBU) 2018	Imperviousness calibration	HRL Imperviousness

4.2 General methods used in production

In the following chapters, detailed descriptions of the general processes required to derive all HRL Imperviousness 2021 products are provided. The overall production workflow, illustrated in Figure 2, presents an overview that employs generalized algorithms, which are described here in depth. The graph nodes contain the names of specific algorithms, which primarily include aggregation methods, U-Net training, and prediction. All methods are explained below and

further elaborated in the subsequent subchapters. Following this, derivations for individual products are presented.

4.2.1 Pre-processing

During satellite data pre-processing, quantiles and harmonic coefficients are calculated to support subsequent model training and prediction (as described below). The entire set of these features is utilized comprehensively in the modelling process. The pre-processed feature sets are referred to as **PP_18** and **PP_21**, depending on the respective year in the following description.

Pre-processing of a specific year includes the following parametrizations (Table 3).

Table 3: Overview of parameterizations used to create a specific reference year

Pre-processing Outputs	Details
Quantiles	q20* of bands B02, B03, B04, B05, B06, B07, B08, B8A, B11, B12; q90 of NDVI
Harmonics	NDVI, degree 3 (see Calculation of harmonic coefficients)

* Lower quantiles effectively reduce the remaining influence of clouds from imperfect cloud masks, while still preserving key spectral information.

Calculation of quantiles

As part of the pre-processing, quantiles are calculated in a weighted manner using a specific band from all available scenes, with weights based on the Sentinel-2 L2A Scene Classification Layer (SCL) mask (see Table 6 in Annex).

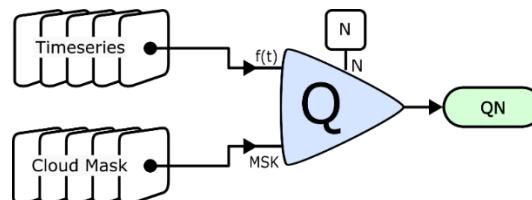


Figure 3: Simplified workflow for the calculation of quantiles

The process involves using a time series of a specific band or index, along with the corresponding cloud masks (SCL). This method ensures that only relevant data points, unaffected by clouds, contribute to the quantile calculation. The quantile number, N, indicates the specific quantile being calculated. For example, if N is set to 20%, the resulting output layer will be labeled as "Q20". This layer represents the 20th percentile of the data, providing a clear representation of the distribution of the values within the selected band (see Figure 3).

Calculation of harmonic coefficients

This pre-processing step transforms a variable number of input observations into a fixed set of model curve coefficients. This is done by fitting the following function to the input time series to retrieve the model curve coefficients p_0 and a_k (for a function of order N):

$$f(t) = p_0 + \sum_{k=0}^N a_k \sin(k\omega t + \phi_k)$$

This fitting process can be applied to any input band; in the context of HRL Imperviousness 2021, it is specifically applied to the NDVI time series. The coefficients are subsequently used for various operations in the process. The calculation requires input like those needed for quantile calculation. In this context, the number N represents the order of the harmonic series, with a default value of N=3. The model coefficients serve to approximate the full time series of a certain band or index. These coefficients are calculated using a weighted ordinary least squares approach. The output consists of a set of $3N+1$ coefficients, which includes coefficients for the amplitude and phase of each mode, plus a constant term. This set of coefficients is abbreviated as "CA ϕ " (Constant, Amplitude, Phase) (see Figure 4).

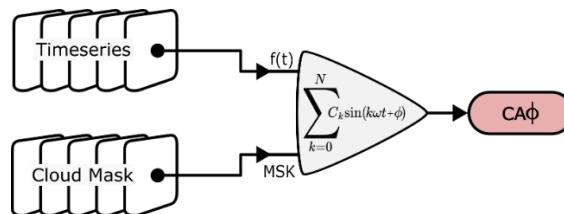


Figure 4: Simplified workflow for the calculation of harmonic coefficients

4.2.2 Root mean squared deviation

The first derived interim layer is obtained by calculating the Root Mean Squared Deviation (RMSD) using the harmonic coefficients of the NDVI from two distinct years. This RMSD serves as a proxy for detecting possible land-cover changes and is used to derive a change mask which is consistently used throughout the production of every change layer provided with HRL Imperviousness 2021.

Figure 5 below shows the dependencies for the RMSD calculation from the pre-processed data: The pre-processed data of 2018 and 2021 (PP18 and PP21) is used to derive the RMSD of the NDVI time series of two years.

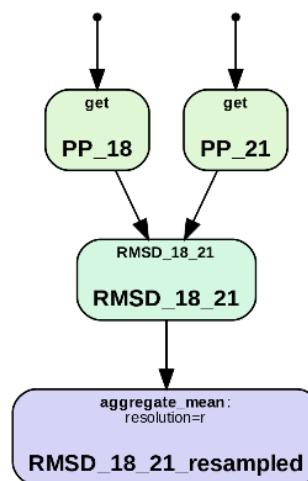


Figure 5: Subgraph showing the dependencies for the Root Mean Squared Deviation (RMSD) calculation from the pre-processed data

For identifying changes in built-up and impervious areas, the NDVI time series is particularly sensitive and effective.

The RMSD is mathematically represented as follows:

$$RMSD^2 = \frac{1}{T} \int_0^T (f_1(t) - f_2(t))^2 dt \propto \sum_{k=0}^N (C_{k,1} - C_{k,2})^2$$

By using the harmonic model coefficients (C_k) as input, this calculation simplifies determining the Euclidean distance between the coefficients of the two years. This simplification allows for a fast and efficient computation of an independent change indicator (see Figure 6).

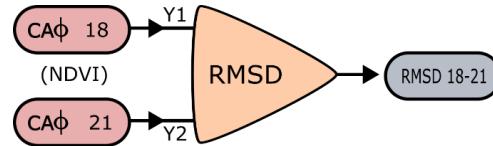


Figure 6: Calculation of the Root Mean Squared Deviation (RMSD) from the harmonic coefficients

4.2.3 Deep learning

Deep learning techniques are employed to develop models for classifying the fractional coverage of built-up and impervious pixels using a U-Net architecture.

Training Data

The models are trained using high-quality built-up vector footprints and imperviousness outlines (see Chapter 4.1.5):

- **Built-up vector footprints** are sourced from OpenStreetMap (OSM), Microsoft buildings, and Google building footprints.
- **Imperviousness outlines** come from the 2018 datasets of CLCplus Backbone and IMD.

Data and Model Transferability

All models are trained exclusively with information from Sentinel-2 satellite images. No additional data is needed for making raw predictions. It has been verified that these models are transferable between years. For instance, a model trained with 2018 image data can be applied to reliably and accurately predict IBU and IMD for the 2021 reference year.

Input Features

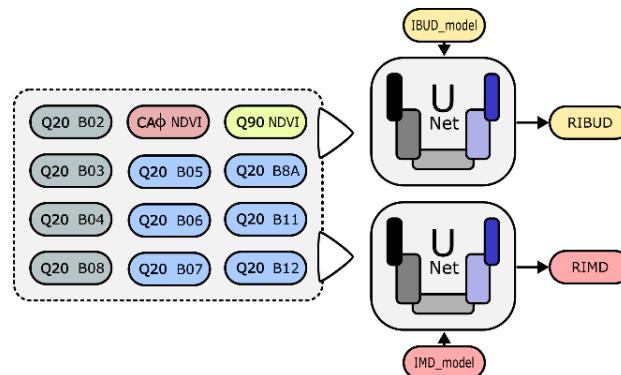


Figure 7: Overview of the inputs used for the U-Net prediction

All inputs are generated using data that has been pre-processed by the Quantiles or Harmonics approach. An overview of these inputs is shown in Figure 7 and the according methodology for these calculations is outlined in the previous section (Chapter 4.2.1).

The harmonic coefficients and quantiles of specific bands and indices are used as input features:

- **10 m harmonic coefficients:** NDVI
- **10 m quantiles (20%):** B02, B03, B04, B08
- **10 m quantiles (90%):** NDVI
- **20 m quantiles (20%):** B05, B06, B07, B08, B8A, B11, B12

Model Outputs

Depending on the training data, the models produce the outputs summarized in Table 4.

Table 4: Overview of the model predictions

Name	Trained on	Symbol/Abbreviation
RAW Impervious Built-Up density	building footprints	RIBUD
RAW Imperviousness density	CLCplus 18, IMD 18	RIMD

4.2.4 Aggregation methods

The following methods are used for specific steps in the production process where a reduction of the resolution is required. This is, for example, necessary to produce the change layers in a resolution of 20m, or to generate 100m aggregated products.

Mean status aggregation

Denoted as “**aggregate_mean**” in the processing graph (Figure 2). This type of aggregation calculates the mean value of all valid pixels contributing to a lower resolution pixel and is implemented in the following way:

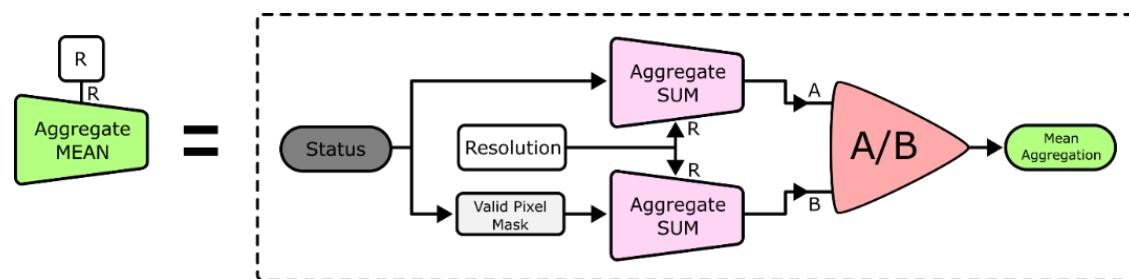


Figure 8: General approach for the mean status aggregation to lower resolutions (R)

Note that the mean value is only calculated using valid pixels. Valid pixels are all values other than the no data value (see Figure 8).

Mean change aggregation

Denoted as “**aggregate_mean_change**” in the processing graph (see Figure 2).

In principle, change layers can be aggregated using the mean aggregation on valid pixels as described above (even though the value range is shifted from [-100, 100] to [0,200]). However, there is a “special case” in the change layers, which makes the aggregation more complex. This case is denoted by the pixel value 201, marking pixels without change where IMD=0 in both contributing status layers (i.e., no imperviousness in both years) (see Figure 9):

- These pixels are first reset to the no change value (100)
- Then, the mean aggregation is applied
- Afterwards, aggregated pixels which had a 201 in all contributing 10m pixels are re-set to 201

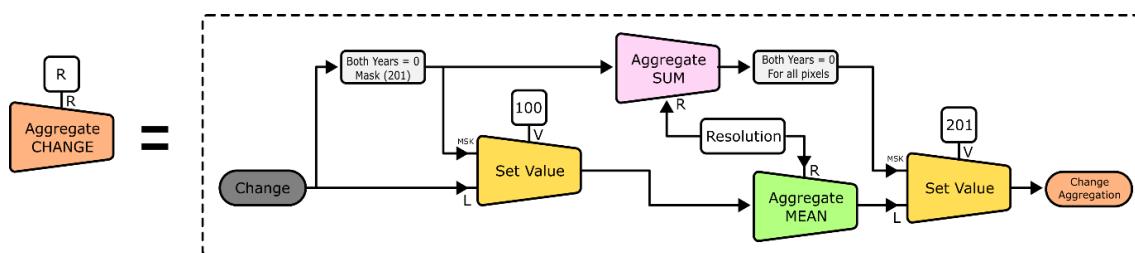


Figure 9: General approach for the aggregation of changes to lower resolutions (R)

Maximum status aggregation

Denoted as “**aggregate_max**” in the processing graph (see Figure 2). This type of aggregation is required to implement the “touching rule” for binary layers (e.g., IBU), i.e., a pixel that contains at least one true value results in a true output at lower resolution. Maximum aggregation is available as a resampling method in the Geospatial Data Abstraction Library (GDAL). A slight modification is needed to ensure that only valid pixels are considered (see Figure 10).

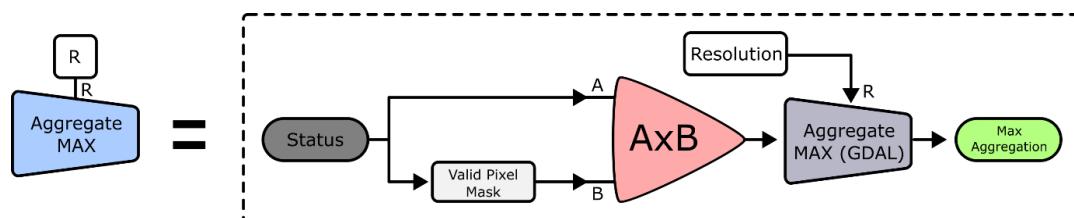


Figure 10: General approach for the maximum aggregation to a lower resolution (R)

4.2.5 Intermediate layer: Change mask (CHM)

A general indicator for potential changes between the two production years is needed to distinguish between real and technical changes. This indicator is derived from the RMSD by applying a thresholding approach. The subgraph of the processing workflow (see Figure 2), which leads to the final change mask, is shown in Figure 11.

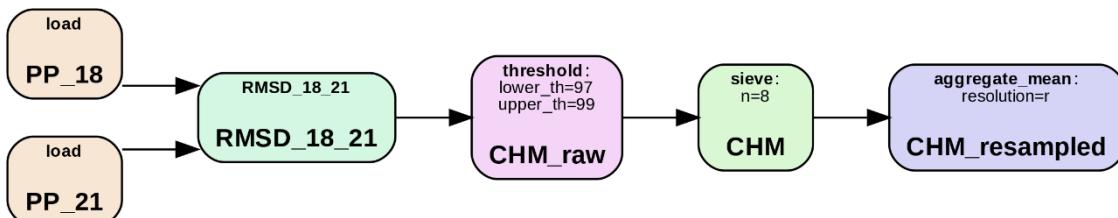


Figure 11: Creation of a unified change mask (CHM). This mask is derived from the RMSD calculation by applying a thresholding approach followed by a sieving

To derive this CHM, two types of thresholds are defined:

- **Lower Threshold:** This defines the extent of a change.
- **Upper Threshold:** Patches identified by the lower threshold are considered as a change only if at least one pixel within the patch exceeds this threshold.

The resulting binary mask is then filtered using a sieving approach to enhance change reliability for small patches, which have a higher probability to reflect noise inherent to EO time series. The sieving size S is set to 8 pixels in the full 10m resolution (i.e., equivalent to 2 pixels in 20m resolution).

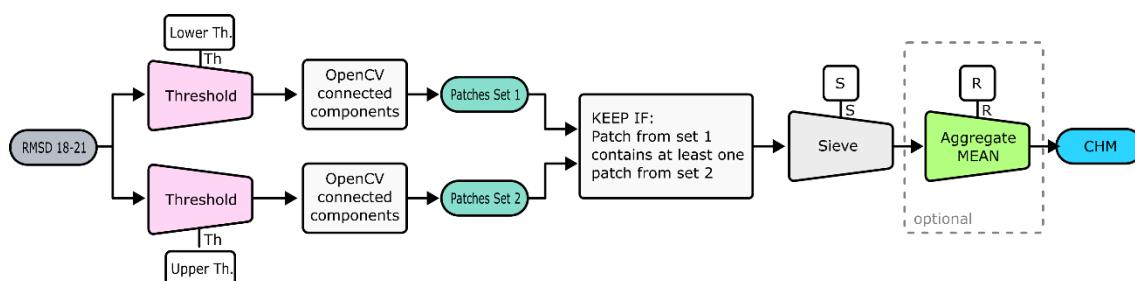


Figure 12: Creation of the change mask using a thresholding approach and sieving

The main algorithm used for this process is OpenCV’s “connected components” [4], which identifies all spatially connected pixel patches in a binary layer and assigns a unique label (or component ID) to each group of connected pixels. This allows each contiguous feature to be individually referenced and processed in subsequent steps (see Figure 12).

Implementation Details

1. Resolution and Aggregation:

- The CHM layer can be retrieved in any resolution down to 10 meters. This is achieved by incorporating the mean aggregation algorithm.
- The layer can have values between 0 and 1, being binary only at the native 10-meter resolution. At lower resolutions, it represents the percentage change per pixel.

2. Thresholds:

- Thresholds are determined individually for each tile (based on the EEA 100 km reference grid) by calculating the 97.5th and 99th percentiles (quantiles) of the RMSD values within that tile. This results in tile-specific upper and lower thresholds in pixel value space.
- The relative positions of the thresholds, set at the 97.5% (lower) and 99% (upper) levels, were determined empirically through iterative testing and visual

inspection. This semi-automatic approach ensures consistent and reliable separation of outliers across diverse tile conditions.

3. Mask and Change Type Identification:

- The resulting mask highlights pixels where a real change between reference years of any kind has been detected.

The change mask is not specific to changes in impervious surfaces and includes, e.g., agricultural fields or forests that are subject to changes. The specific type of change (e.g., IBU loss, IMD increase) is determined only after combining CHM layer with the corresponding status layers (see chapter 4.5).

4.3 Production of the Imperviousness Density (IMD)

The production steps to derive the IMD layer contain two main steps:

- (1) the retrieval of the non- / impervious surface mask through image classification and
- (2) the estimation and calibration of IMD 2021 values.

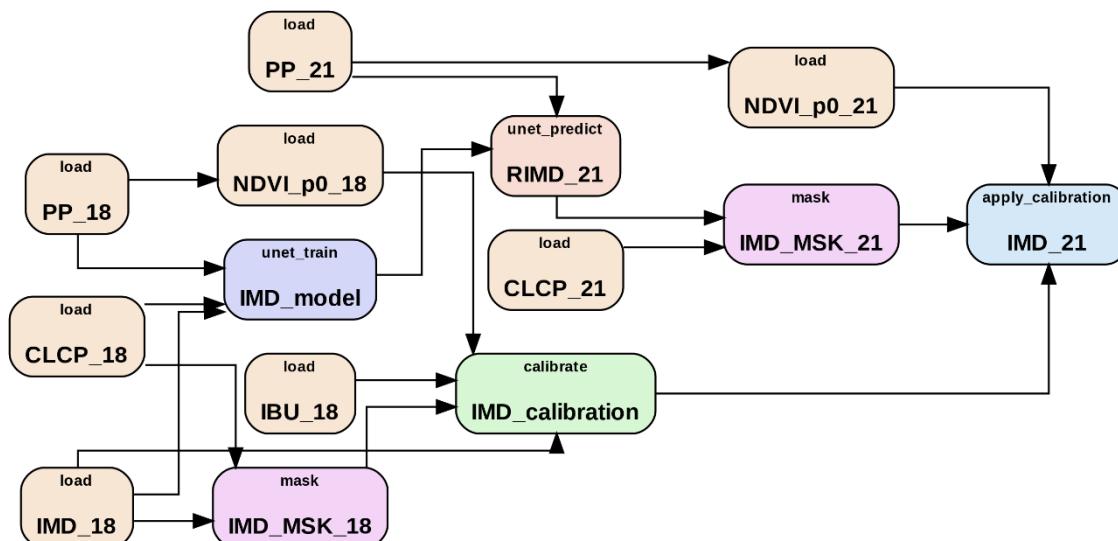


Figure 13: Processing graph for the IMD 2021

The processing graph to produce the final IMD layer is shown in Figure 13 above. The first essential step is the collection of relevant input data (i.e., EO data, auxiliary geospatial dataset) for training and verification input. These datasets are then prepared for the following data analytics and processing steps:

- Creation of a training input: Time-series statistics and parametrization (Sentinel-2) through percentiles and harmonic functions (PP_18, PP_21, see Chapter 4.2.1).
- Creation of a training mask: Using available, high-quality ground-truth data of a specific year for direct training and prediction of the imperviousness densities from the satellite input.
- Pre-processing: All Input features (see 4.2.3), as well as the training mask, are prepared for the U-Net classification.
- U-Net training: All input features and training masks are used to train a U-Net. To speed up the training effort and improve results, a transfer learning approach is implemented, using pre-trained models from neighboring tiles (if available) to provide a better initialization of the model weights and enhance generalization.



- Prediction: The trained model is used to predict imperviousness densities on newly processed data from the prediction year.
- Post-processing: Calibration coefficients are first derived by fitting IMD 2018 to NDVI p0 (2018) within the imperviousness 2018 outline (IMD_MSK_18). Samples for fully impervious areas (i.e. 100% IMD pixels) are retrieved from the IBU_18 dataset (green box “calibrate” in Figure 13). These coefficients are then applied to NDVI p0 (2021) to recalibrate raw predictions. CLCplus 2021 is used to refine and enhance the imperviousness outline.

Change and aggregated 100 m layers are described separately in sections 4.5 and 4.6.

4.3.1 Method selection, training, and prediction

To produce IMD, a highly efficient and noise-reducing time-series parameterization methodology (see Chapters 4.1.2 and 4.2.1) is combined with a modified deep convolutional neural network (CNN) called U-Net [5]. This has proven to work well for image segmentation tasks and has also successfully been demonstrated for previously released Global Human Settlement Layer (GHSL) products [6]. The U-Net model can assign classifications to each pixel while considering the surrounding shapes and edges. It has, therefore, proven superior to traditional pixel-based approaches like decision trees, Support Vector Machines, Random Forests, or other bootstrapping methodologies [7].

The U-Net, a convolutional auto-encoder network, is built in a classical way with equal resolution of input and output images. The training is carried out on an 80/20 train-validation-split of the data. Predictions are carried out on the reference year 2021 data, which are clipped into 256x256 pixel images. The U-Net architecture is optimized for an input set of 12 features, including spectral band statistics and derived features. To mitigate overfitting, a dropout rate of 10% is applied during training. The model employs binary cross-entropy as the loss function, enabling direct density prediction of imperviousness and built-up areas. Training and inference are implemented using TensorFlow, which provides an efficient and scalable framework for deep learning in high-dimensional remote sensing data. This setup ensures the network effectively captures both spatial patterns and sub-pixel structures.

4.3.2 Mapping of Imperviousness Density

The IMD layer is a high-resolution map of the percentage of surface area covered by artificial sealing for the reference year 2021, produced using automatic image processing methods. As for previous productions, calibration of the resulting IMD values and consistency with the historical layers (2006-2009-2012-2015-2018) are key requirements. The method presented is a further development of the algorithm used for the HRL Imperviousness 2018, with some major improvements regarding the classification approach (i.e., CNN-based image classification) of impervious surfaces.

The sealed area mask 2021 is extracted by combining the extracted satellite input features (Chapters 4.1.2 and 4.2.1) with the proposed classification approach (Chapter 4.2.3).

Since the training process is carried out with a combination of the CLCplus Backbone layer and the IMD status layer of the previous production (2018), an extended workflow to derive accurate and consistent imperviousness degrees is needed. The addition of the CLCplus Backbone layer is beneficial for improving the detection of low-density areas such as smaller roads. Since this

layer is binary only, a re-calibration of the density is needed. The necessary steps for re-calibration are described in the following sections.

Sealing mask creation

The first step in the calibration process is the creation of a binary sealing mask. This is needed for the production year 2018 to be able to calibrate only on values that are part of the sealed areas. Therefore, class 1 (Sealed areas) from CLCplus BB 2018 is combined with the sealing outline of the imperviousness 2018 layer to create the sealing mask for 2018 (IMD_MSK_18) (see Figure 14).

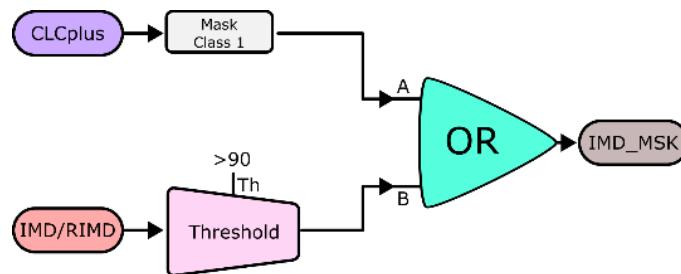


Figure 14: Derivation of a sealing mask based on a combination of IMD and CLCplus Backbone

The same procedure is used to obtain a sealing mask for 2021, using the Raw imperviousness prediction of 2021 (RIMD) and the CLCplus Backbone 2021 layer. This is used for the final step in producing the IMD 2021 layer.

The 2018 sealing mask is used specifically to calibrate the model on 2018 values, ensuring that the calibration is only applied to areas identified as sealed in that year. For the 2021 production, a similar procedure is followed, but the 2021 sealing mask, derived from the raw imperviousness prediction (RIMD) and the CLCplus Backbone 2021 layer, is used to apply the calibration that was based on the 2018 data.

Additionally, to prevent misclassification in areas with uncertain or borderline imperviousness, a threshold of 90% is applied to the inputs of the IMD layer (either the 2018 IMD layer or 2021 raw classification). This threshold is particularly important for the 2021 raw predictions, but to maintain consistency in the approach across both years, it is applied in the same manner to the 2018 data as well.

Absolute calibration

In previous productions, calibration was performed using a sample point reference database. However, this approach often faced challenges with an insufficient number of calibration points, necessitating frequent manual adjustments and the inclusion of points from nearby regions, which occasionally resulted in incomplete or unreliable calibrations. To address these issues, the method has transitioned to a more stable and robust process, directly calibrating the IMD 2018 layer.

This new method offers several key advantages:

- **Consistency:** By calibrating directly on previous IMD products, it ensures a higher level of consistency in the results. This method leverages the existing data infrastructure and aligns more closely with established production workflows.

- **Efficiency:** The new calibration process eliminates the need for time-consuming manual interactions associated with sample point reference databases. This streamlines the production process, reducing the likelihood of human error and increasing overall efficiency.
- **Data Utilization:** Direct calibration allows for a massive increase in the number of data points used. This leads to more robust and reliable models, as they are trained on a larger and more representative dataset.
- **Cost-Effectiveness:** By removing the dependency on the sample point reference database, it reduces costs associated with database maintenance and manual data handling. This makes the production process more economical and sustainable in the long term.

The calibration procedure is divided into multiple steps that are described in the following subsections.

Calibration on NDVI

The sealing mask (Figure 14) is used to extract density values from the IMD 2018 layer and calculate regression coefficients. For this, a linear regression model is used which maps NDVI values linearly to IMD. This mapping is performed on a densely sampled grid with a resolution of 60 meters. The output of this operation is the regression coefficient vector R (Figure 15).

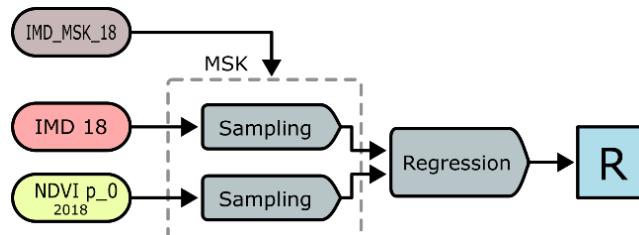


Figure 15: Retrieval of regression coefficients to calibrate the imperviousness outline on the mean NDVI values (NDVI p_0)

High-density sealing areas

In some regions, particularly in northern cities, large impervious areas such as facilities and parking lots were inaccurately assigned a value well below 100% in the IMD 2018 layer. This underrepresentation of high-density sealing areas, especially in places like Iceland, is addressed through the method described in this section.

To enhance high-density sealed surfaces, the IBU 2018 is used to sample from the IMD 2018 and estimate the maximum value of predicted densities based on known areas of 100% density (see Figure 16). This value is then used as a subsequent optimization step for automatic re-scaling of the prediction.

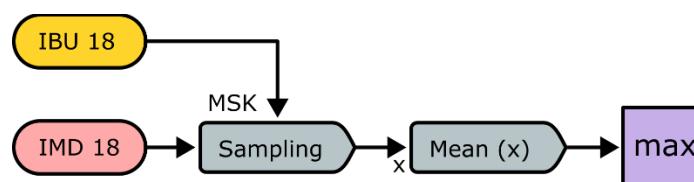


Figure 16: The Impervious Built-Up (IBU) 2018 layer is used as a mask to find the mean predicted sealing density of these high-density sealing areas

Automatic density re-scaling

This step uses an optimization algorithm to find a new minimum IMD value, ensuring that rescaling the predicted values does not change the average imperviousness density in that area. In simple terms, this step adjusts a threshold in a feedback loop to find a lower threshold on imperviousness densities for re-scaling, making sure the average density of all impervious pixels stays close to the values obtained in 2018 (see Figure 17).

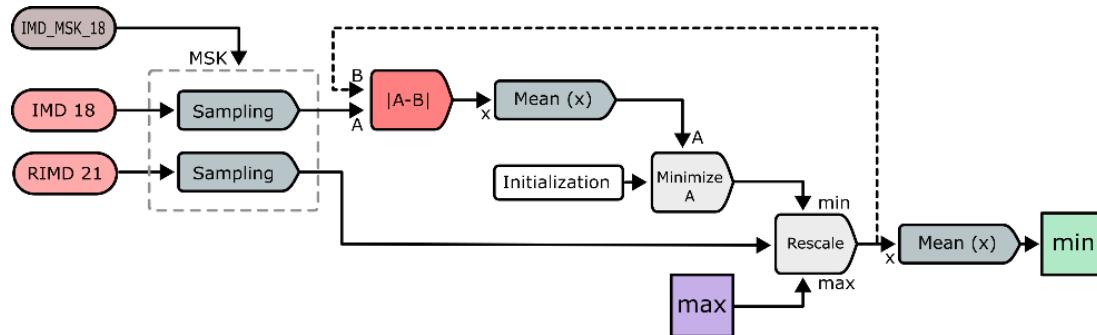


Figure 17: An optimization algorithm used to find a new minimum value for rescaling such that the mean predicted density of 2021 remains consistent with the values from 2018

Application to produce the IMD 2021 layer

The regression coefficients and re-scaling values obtained in the previous steps (R , \min , \max) are applied to data from 2021 to produce the IMD 2021 layer. For lower resolutions (e.g., IMD 100 m), mean aggregation is applied (see Figure 18). The parameters retrieved in the previous steps, i.e., the calibration coefficients (R) and the rescaling values (\min , \max), are used in the final steps to create IMD 2021.

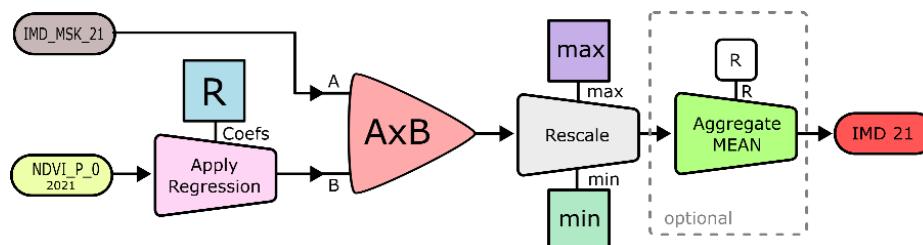


Figure 18: Algorithm used for producing the IMD 2021 layer

4.3.3 Mapping of United Kingdom in HRL Imperviousness 2021

Owing to the United Kingdom's return to the Copernicus programme at a progressed stage of this HRL Imperviousness 2021 production, the processing strategy followed for this territory was aligned with the main EEA38 workflow wherever possible. However, one key difference required an adaptation: the CLCplus Backbone 2021 dataset, which is used to derive the Imperviousness 2021 outline, does not include coverage for the UK. To address this, several additional steps were implemented:

- A distance map to potentially sealed surfaces was generated using OpenStreetMap (OSM) roads and buildings, based on a 2021 OSM snapshot.

- This distance map, along with the standard input features, was used as input to a lightweight XGBoost classifier.
- The classifier was trained using areas corresponding to CLCplus Backbone class 1 (sealed areas) from the 2018 dataset.
- The trained model was then applied to 2021 input features to derive an updated proxy dataset representing CLCplus Backbone class 1 for the UK.

This proxy data set served as the basis for generating the sealing outline, after which the standard workflow could be applied seamlessly. This approach enabled a harmonized integration of the UK data, ensuring compatibility with the existing HRL Imperviousness 2021 methodology while effectively compensating for the absence of CLCplus Backbone data.

4.4 Production of the Impervious Built-Up (IBU)

This section describes the approach to map Impervious Built-Up (IBU), which allows to differentiate above-ground built-up areas from all impervious and, specifically, the remaining flat impervious surfaces that are included in the IMD.

The processing graph of the production is shown in Figure 19, and the necessary steps for production are described in more detail in the following chapters.

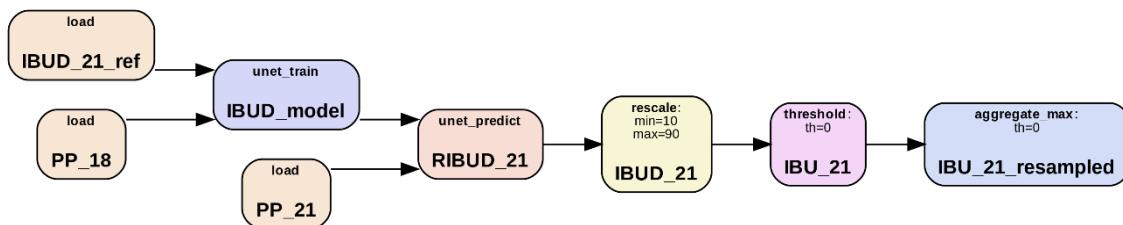


Figure 19: The processing graph to produce the IBU layer

The processing follows a methodology very similar to that used to produce IMD. For IBU, no re-calibration is necessary, as superior training data, derived from high-quality ground-truth sources (i.e. OSM, MSB, and Google Open Buildings), is available. Additionally, the final output of IBU is a binary mask in which pixels representing buildings and other above-ground constructions are assigned a value of 1. It is important to note that a continuous layer (IBUD 21) is generated as an intermediate step to ensure consistency in the processing scheme across both IMD and IBU prediction workflows.

4.4.1 Model training and prediction

Built-up areas for the IBU are derived with the same U-Net classification approach as impervious areas (Chapters 4.2.3 and 4.3.1). Tests have shown that the input training data quality strongly affects the resulting output product, as demonstrated in Figure 20. The figure focuses on a test area around Poiniers in France. The ground-truth image was generated by using OSM building footprints. On the left-hand side, the IBU 2018 layer is shown for reference.

From visual comparisons, training based on GHSL produces similar results to the IBU 2018 layer, while the model trained on OSM data captures reality more accurately.

Analysis of the test phase products, as well as the final production, confirmed that a consistent classification based on high-quality training data can be performed for the entire EEA38 region.

It has been observed through both visual inspection and technical analysis that the classifier generalizes well. The results consistently demonstrate that the prediction quality remains high even when training is performed using data from areas surrounding the target area, typically covering larger regions with similar characteristics. This ensures reliable performance across different spatial contexts. Furthermore, the available training datasets cover large portions of the European continent and span across a large variety of different biogeographical regions.

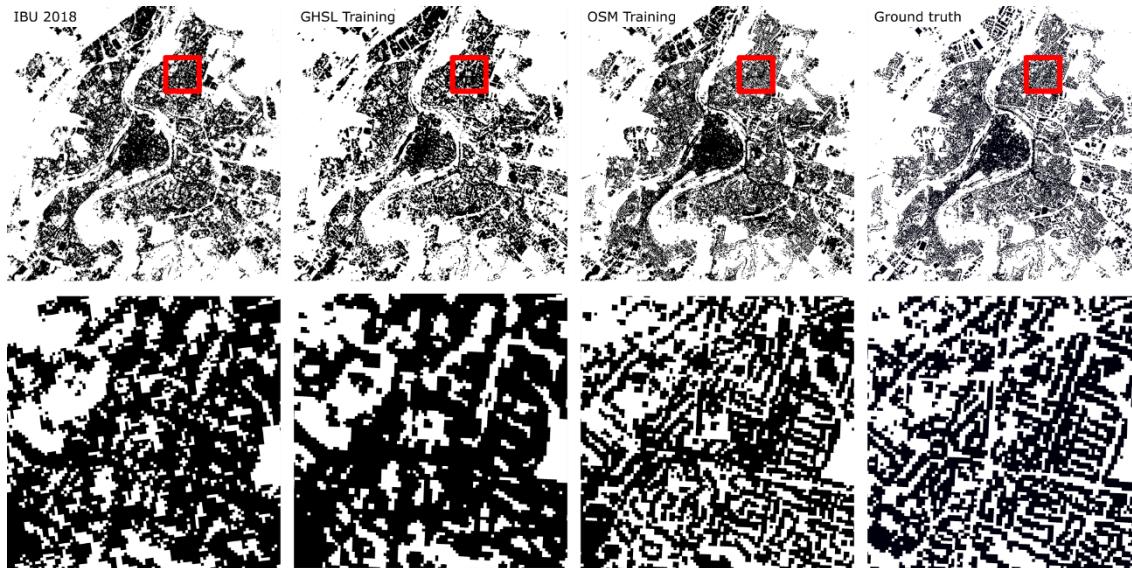


Figure 20: Examples of built-up prediction in a test area around Poiniers, France. From left to right: (1) The previous IBU product (2018), (2) a prediction carried out with a model trained on GHSL, (3) a prediction carried out with a model trained on OSM, (4) the ground truth, based on a rasterized version of OSM building footprints

4.4.2 Post-processing

The prediction output of the classification consists of built-up densities for each pixel (such as the IMD layer). Since the IBU product is binary, a simple threshold of 10% is applied to the predicted built-up densities. This is well seen in Figure 21 where the IBU 2021 is created using a threshold on the Raw Imperviousness Built-up (RIBUD) layer. Lower resolutions can be retrieved by applying maximum aggregation. This aligns with the definition that a pixel should be marked as built-up, if it contains a building.

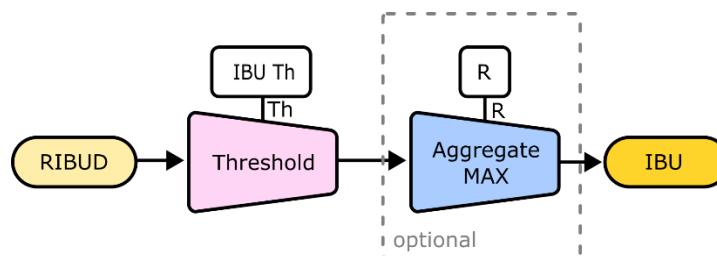


Figure 21: IBU 2021 created using a threshold on the Raw Imperviousness Built-up (RIBUD) layer

A threshold of 10% has been selected as it offers a balanced trade-off between maintaining classification accuracy and minimizing false positives. Lower thresholds, such as values just

above 0%, tend to overestimate built-up areas due to inherent numerical limitations, making 10% a practical and robust choice.

Note: The share of built-up (SBU) layer is produced by using the above approach with an output resolution of 100 m and an average resampling instead of the maximum resampling.

4.5 Change layer generation

This section describes the approach to generating the change layer between the 2018 and 2021 status products (both IMD and IBU), including the classified change layer. Both change layers (continuous and classified) are accompanied by a support layer which outlines the technical changes between both years. The generation of the change products follows a unified processing scheme which is described in the following sections.

4.5.1 Generation of continuous change layers

This generalized approach is denoted as “**change_continuous**” and “**change_continuous support**” in the processing graphs. The continuous change is derived from two status products (IMD, IBU) by applying the change mask (see Chapter 4.2.5), followed by an averaging operation (see Figure 22). The change support layer is then generated by subtraction, ensuring that the total change (e.g., IMD 2021 – IMD 2018) is given by the sum of real and technical changes.

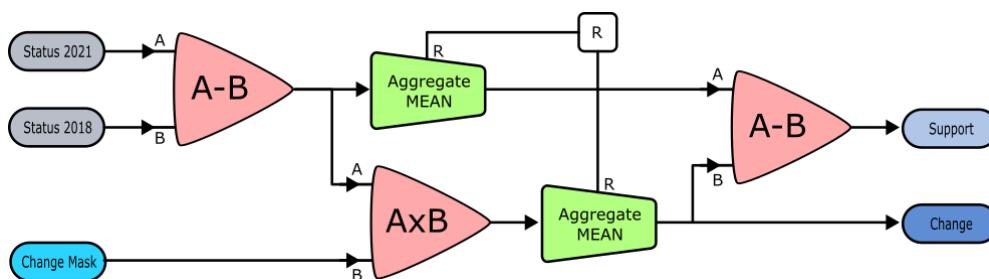


Figure 22: General workflow to compute change and change support layers for adjustable resolutions

The creation of the change and change support layers is done in two steps: Step 1 aggregates the differences of both status layers from 10 to 20 m spatial resolution by applying a mean aggregation method. Step 2 involves filtering the result for real and technical changes. This is achieved by intersection with the change mask.

The occurrence of imperviousness decrease is rare and, when it does happen, it is more likely due to re-greening (complete de-sealing) of an impervious surface rather than an actual decrease in imperviousness. Such changes typically produce a very high signal in the derived change mask. To address the issue, the sensitivity for detecting imperviousness losses was reduced by a factor of 10 by implementing an additional check based on the status layers. This adjustment lowers the model's sensitivity to smaller changes in the data, helping to minimize false positives, where non-impervious areas are incorrectly identified as having undergone imperviousness changes. Despite this reduction, the model still responds appropriately to clear instances of actual imperviousness losses, ensuring that significant changes are accurately detected while avoiding over-detection of minor or irrelevant changes.

Please note that instances of imperviousness losses were specifically searched for, but no significant examples were identified. This suggests that imperviousness losses were



overestimated in previous productions, despite being relatively rare in general. The reduced sensitivity applied in the current approach has helped to refine the change detection process, ensuring that only genuine imperviousness changes are captured. This adjustment minimizes the occurrence of false positives and leads to a more accurate representation of imperviousness dynamics, with a focus on real, substantial changes rather than isolated or marginal losses.

4.5.2 Generation of the classified change layers

In this step, the change is classified into the categories “new cover,” “loss of cover,” “unchanged,” “increased cover”, and “decreased cover.” This is done using a rule-based reclassification and applied to both the real- and technical change layers (Table 5). Classification of the change layers is denoted in processing graphs with the “**Classify_change**” operation.

Table 5: Rule base to create the Imperviousness Change Classified (IMCC), Impervious Built-Up Change (SBUC) and Share of Built-Up Change Classified (SBCC) layers

Code	Class name	IMCC	IBUC	SBCC
0	Unchanged areas	Unchanged non-impervious areas	Unchanged non-built-up areas	Unchanged non-built-up areas
		IMD(n-1) = 0, IMD(n) = 0	IBU(n-1) = 0, IBU(n) = 0	SBU(n-1) = 0, SBU(n) = 0
1	New cover	Increased imperviousness density, no imperviousness at first reference date	Increased built-up density, no built-up at first reference date	Increased built-up density, no share of built-up at first reference date
		IMD(n-1) = 0, IMD(n) > 0	IBU(n-1) = 0, IBU(n) > 0	SBU(n-1) = 0, SBU(n) > 0
2	Loss of cover	Decreasing imperviousness density, no imperviousness at second reference date	Decreased built-up density, no built-up at second reference date	Decreased built-up density, no share of built-up at second reference date
		IMD(n-1) > 0, IMD(n) = 0	IBU(n-1) > 0, IBU(n) = 0	SBU(n-1) > 0, SBU(n) = 0
10	Unchanged cover	Unchanged impervious areas, imperviousness > 0 at both reference dates	Unchanged built-up area, built-up > 0 at both reference dates	Unchanged built-up areas, share of built-up > 0 at both reference dates
		IMD(n-1) > 0, IMD(n) > 0	IBU(n-1) > 0, IBU(n) > 0	SBU(n-1) > 0, SBU(n) > 0
11	Increased cover	Increased imperviousness density, imperviousness > 0 at both reference dates	Increased built-up density, built-up > 0 at both reference dates	Increased share of built-up share of built-up > 0 at both reference dates
		IMD(n-1) > 0, IMD(n) > 0 & IMD(n) > IMD(n-1)	IBU(n-1) > 0, IBU(n) > 0 & IBU(n) > IBU(n-1)	SBU(n-1) > 0, SBU(n) > 0 & SBU(n) > SBU(n-1)
12	Decreased cover	Decreased imperviousness density, imperviousness > 0 at both reference dates	Decreased built-up density, built-up > 0 at both reference dates	Decreased share of built-up, share of built-up > 0 at both reference dates
		IMD(n-1) > 0, IMD(n) > 0 & IMD(n) < IMD(n-1)	IBU(n-1) > 0, IBU(n) > 0 & IBU(n) < IBU(n-1)	SBU(n-1) > 0, SBU(n) > 0 & SBU(n) < SBU(n-1)
255	Outside area			

4.5.3 Imperviousness Change (IMDC) and Imperviousness Density Change Classified (IMCC)

The layers representing the change in imperviousness between 2018 and 2021 continue the legacy of the 20 m IMDC/IMCC products generated in the previous HRL implementations (2006, 2009, 2012, 2015, 2018). As such, the spatial resolution of 20 m for IMDC and IMCC is kept. Both the imperviousness change, and its classified version are derived as shown in Figure 23.

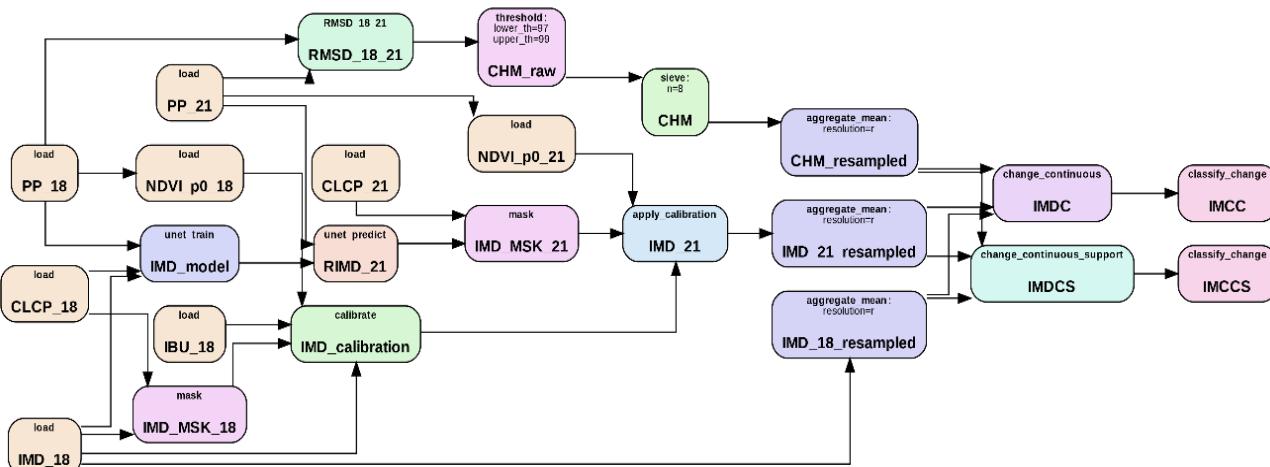


Figure 23: Full processing graph to obtain IMDC and IMCC products along with their support layers (IMDCS and IMCCS). The resampling resolution is set to R = 20m

The change and corresponding support layers are derived from the aggregated versions of the IMD 2018 and IMD 2021 with the help of the change mask (see 4.2.5). This is done in a consistent manner to ensure that the total change of each pixel value (i.e., IMD 2021 - IMD 2018) is equal to the sum of real changes (IMDC) and technical changes (IMDCS).

4.5.4 Impervious Built-Up Change (IBUC)

The IBUC is derived using a similar generalized workflow to that employed in the creation of the IMDC and IMCC layers (Figure 24).

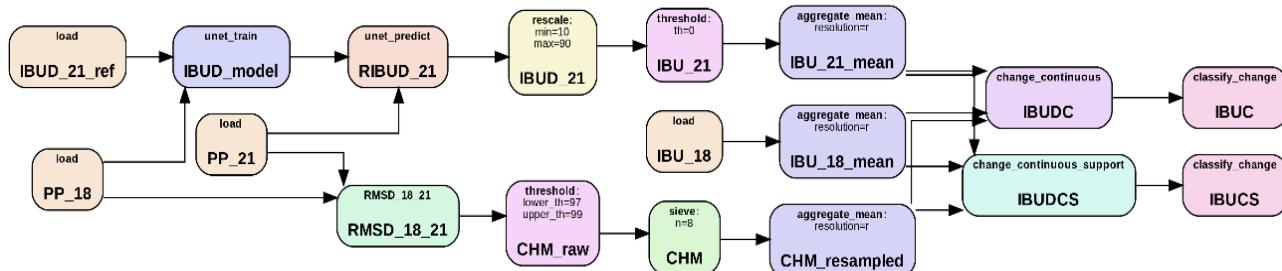


Figure 24: Full processing graph to obtain IBUC along with the support layer (IBUCS). The resampling resolution is set to R = 20m

This approach ensures consistency between the two products and allows for the unified implementation of the necessary methods. Additionally, the use of 20 m mean aggregation (see 4.2.4) supports a consistent definition of classified changes by considering the full 10 m resolution for detecting changes. This means that both classes, increase and decrease, are accessible and clearly defined.

4.6 Derived layers based on aggregation

This section describes the production of the aggregated layers IMD 2021, IMDC 1821, IBU 2021 with a 100 m spatial resolution. Geometric aggregation rules follow the aggregation approach applied to the HRL Imperviousness 2018 production. Please also refer to Chapter 4.2.4 for a general description of the aggregation methods.

4.6.1 Generation of the 100m status layers (IMD 2021, SBU 2021)

The 100m aggregated layers are generated using the mean aggregation method described in Chapter 4.2.4. The IMD 2021 100m dataset uses the mean aggregation method on the 10m IMD 2021 product (Figure 25).

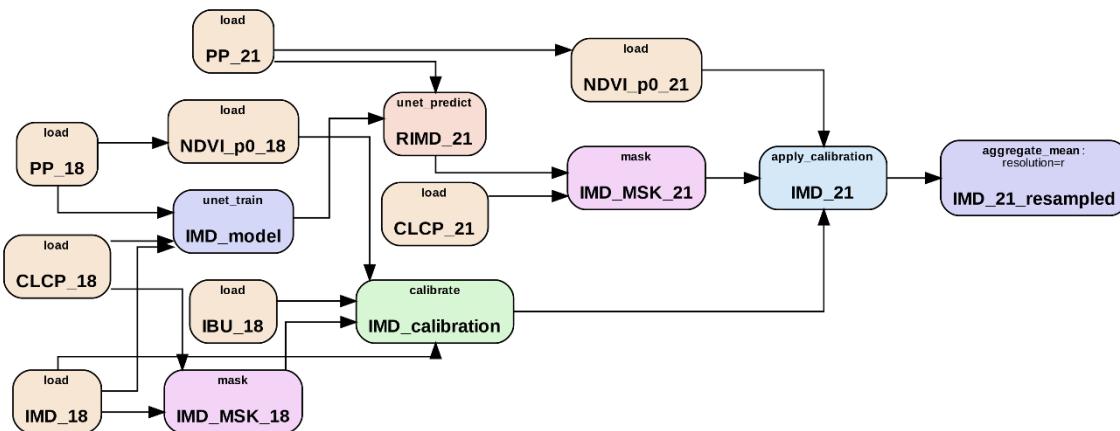


Figure 25: Processing graph of the 100m aggregated IMD 21 layer. The target resolution is set to $R = 100\text{m}$

Similarly, the SBU is derived from a mean aggregation to 100m from the binary IBU 2021 layer (see Figure 26).

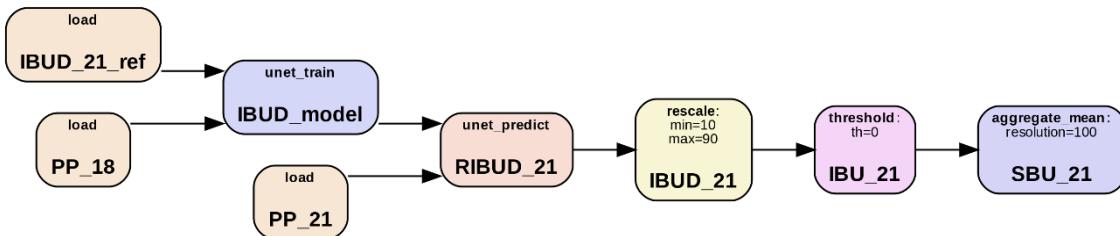


Figure 26: Processing graph for the creation of the SBU layer as 100m mean aggregation from the 10m IBU layer

4.6.2 Generation of 100m change layers (IMDC 1821, SBUC 1821)

A unified workflow for producing products at lower resolutions has been applied consistently across all products. Consequently, the aggregated 100m change products adhere to the same concept as detailed in sections 4.2.4 and 4.6.1.

4.7 Confidence layers

Confidence layers are derived from deviations in the 10m status layer predictions (IMD and IBU). For this, the U-Net classification is applied multiple times using different models from nearby tiles (up to 10 predictions, depending on the availability of trained models). The resulting standard deviation of the predictions is directly used as an indicator for the classification confidence per pixel.

To create an intuitive confidence scale ranging from 0 to 100, the standard deviation of each pixel's predictions is used as the basis. A standard deviation of 0 indicating full agreement across all model predictions is always mapped to the maximum confidence value of 100. Conversely,



the minimum confidence value of 50 is assigned when there is maximum disagreement, specifically when half of the binary classified predictions (based on a 50% density threshold) are 0 and the other half are 1. This ensures the confidence values reflect the level of consensus among models in a consistent and interpretable way.

5 Recognized technical issues

The production process of the HRL Imperviousness 2021 has encountered several technical challenges, which have been carefully analysed and addressed to improve the overall quality of the output. However, some limitations remain inherent to the data and tools available.

Firstly, the absence of harmonized high-quality ground-truth data for imperviousness (i.e. artificial imperviousness densities) presented a significant challenge. As a result, the focus was set on utilizing the previous IMD 2018 for calibration. While this approach provides a workable solution and has shown itself to be superior to using sparsely distributed sample points, it may also propagate existing limitations from earlier stages to an unknown extent.

The inclusion of CLCplus Backbone as a reference for delineating Imperviousness 2021 outlines has significantly increased overall accuracy and the “look-and-feel” of the IMD 2021. However, it has also introduced some drawbacks. Specifically, any commission errors present in CLCplus Backbone are effectively transferred to IMD 2021. This issue is compounded by the fact that model training also partially relies on CLCplus Backbone, making it difficult to eliminate these errors. Addressing this problem would require a better, independent, and large-scale ground-truth dataset to ensure higher fidelity in both calibration and prediction processes.

Another technical issue encountered relates to the use of the GDAL library, which persists across all versions. Specifically, GDAL tends to use pre-generated pyramids, if available, rather than accessing the full-resolution data. This behaviour led to incorrect resampling of low-resolution layers and, consequently, to erroneous outputs. To rectify these errors, affected datasets required re-processing at full resolution (i.e., intermediate layers without pyramids had to be written) to ensure consistent processing. Users are therefore advised to exercise caution when resampling layers containing internal pyramids, such as those in HRL Imperviousness 2021 - to avoid introducing unintended errors into derived datasets or analyses.

While the classifier has significantly improved, certain areas remain challenging. Residual commission errors can still occur, particularly in regions with spectral characteristics like impervious surfaces, such as beaches, gravel fields, bare soil, or construction sites. Additionally, as previously mentioned, the training data introduces some uncertainty; unclear or imprecise boundaries in the reference data can lead to increased confusion in the classification results.

Change detection remains a complex task despite notable improvements in the methodology. The process relies on previous status layers that are not re-processed with updated methods. The cumulative nature of errors in different operations (change layer creation) poses additional challenges. This accumulation can affect the reliability of change maps and require continuous refinement and validation.

Detecting and calibrating imperviousness and built-up losses is particularly difficult. These types of changes are rare and hard to capture accurately, leading to challenges in ensuring precise calibration and modelling. An additional complication arises from the presence of semi-impervious surfaces, such as cobblestones, permeable pavements and other porous surfaces-which are increasingly used in urban design to reduce surface sealing. These materials introduce sub-pixel level variability and spectral ambiguity, making them difficult to detect and distinguish



using remote sensing. Moreover, there is often a lack of consistent reference data to guide the calibration of such surfaces. This poses a challenge when assessing imperviousness decreases, especially since urban redesign strategies tend to replace fully sealed surfaces with partially permeable alternatives rather than removing them entirely.

Despite these challenges, ongoing efforts continue to refine the methodology and seek innovative solutions to enhance the accuracy and reliability of the product.



6 Terms of use and product technical support

6.1 Terms of use

The Terms of Use for the product(s) described in this document, acknowledge that:

“Free, full and open access to the products and services of the Copernicus Land Monitoring Service is made on the conditions that:

1. When distributing or communicating Copernicus Land Monitoring Service products and services (data, software scripts, web services, user and methodological documentation and similar) to the public, users shall inform the public of the source of these products and services and shall acknowledge that the Copernicus Land Monitoring Service products and services were produced “with funding by the European Union”.
2. Where the Copernicus Land Monitoring Service products and services have been adapted or modified by the user, the user shall clearly state this.

Users shall make sure not to convey the impression to the public that the user's activities are officially endorsed by the European Union.

The user has all intellectual property rights to the products he/she has created based on the Copernicus Land Monitoring Service products and services.”

6.2 Citation

When **planning to publish a publication (scientific, commercial, etc.)**, it shall explicitly mention:

“This publication has been prepared using European Union's Copernicus Land Monitoring Service information; <insert all relevant DOI links here, if applicable>” [8]

When developing a **product or service using the products or services of the Copernicus Land Monitoring Service**, it shall explicitly mention:

“Generated using European Union's Copernicus Land Monitoring Service information; <insert all relevant DOI links here, if applicable>”

When **redistributing a part of the Copernicus Land Monitoring Service (product, dataset, documentation, picture, web service, etc.)**, it shall explicitly mention:

“European Union's Copernicus Land Monitoring Service information; <insert all relevant DOI links here, if applicable>” [8]

6.3 Product technical support

Product technical support is provided by the product custodian through Copernicus Land Monitoring Service desk [9]. Product technical support does not include software specific user support or general GIS or remote sensing support.

More information on the products can be found on the Copernicus Land Monitoring Service website [10].



7 References

- [1] European Space Agency. Sentinel-2 Processing Documentation (2023). <https://sentiwiki.copernicus.eu/web/s2-processing>
- [2] Copernicus Sentinel Hub. Sentinel-2 Processing Collection: Processing Baseline (2023). <https://sentiwiki.copernicus.eu/web/s2-processing#S2Processing-Collection-1ProcessingBaselineS2-Processing-Collection-Processing-Baseline>
- [3] Copernicus Space Component Data Access. Optical VHR Coverage Over Europe: VHR Image 2021 (2023). <https://dataspace.copernicus.eu/explore-data/data-collections/copernicus-contributing-missions/collections-description/VHR-IMAGE-2021>
- [4] OpenCV Documentation. Image Processing: Shape Transformation Functions (Version 3.4). OpenCV (2023). https://docs.opencv.org/3.4/d3/dc0/group__imgproc__shape.html
- [5] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In the International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.
- [6] Corbane, C., Syrris, V., Sabo, F., Politis, P., Melchiorri, M., Pesaresi, M., ... & Kemper, T. (2021). Convolutional neural networks for global human settlements mapping from Sentinel-2 satellite imagery. *Neural Computing and Applications*, 33(12), 6697-6720.
- [7] Tong, X.-Y., Xia, G.-S., Lu, Q., Shen, H., Li, S., You, S., & Zhang, L. (2020). Land-cover classification with high-resolution remote sensing images using transferable deep models. *Remote Sensing of Environment*, 237(111322), 111322. doi:10.1016/j.rse.2019.111322.
- [8] Copernicus Land Monitoring Service. Data Policy (2025). <https://land.copernicus.eu/en/data-policy>
- [9] Copernicus Land Monitoring Service (CLMS). Contact Service Helpdesk (2025) <https://land.copernicus.eu/en/contact-service-helpdesk>
- [10] Copernicus Land Monitoring Service. Homepage (2025). <https://land.copernicus.eu/>
- [11] Imperviousness density 2015 (Raster 20 m and 100 m), Europe, 3-yearly. Copernicus Land Monitoring Service. <https://land.copernicus.eu/en/products/high-resolution-layer-imperviousness/imperviousness-density-2015>



List of abbreviations

Abbreviation	Name
AI	Artificial Intelligence
ATBD	Algorithm Theoretical Basis Document
AWS	Amazon Web Services
BOA	Bottom-Of-Atmosphere
CAφ	Harmonic coefficients (Constant, Amplitude, phase. Pre-processing output)
CHM	Refined change mask with sieving applied
CHM_raw	Threshold RMSD_18_21 to create a change mask
CHM_resampled	Aggregated version of the change mask (mean aggregation)
CL	Confidence layer
CLC	CORINE Land Cover
CLCP_18	CLCplus Backbone 2018
CLCP_21	CLCplus Backbone 2021
CLMS	Copernicus Land Monitoring Services
CNN	Convolutional Neural Network
CSCDA	Copernicus Space Component Data Access
EEA	European Environment Agency
EEA38+UK	The 32 member and 6 cooperating countries of the EEA, and UK
EO	Earth Observation
ESA	European Space Agency
GDAL	Geospatial Data Abstraction Library
GHSL	Global Human Settlement Layer
GIS	Geographic information system
GRD	Ground Range Detected
HRL	High Resolution Layer
IBU	Impervious Built-Up
IBU_18	Impervious Built-Up 2018 update
IBU_18_mean	Aggregated version of the binary built-up 2018 (mean)
IBU_21	Impervious Built-Up 2021 update
IBU_21_mean	Aggregated version of the binary built-up 2021 (mean)
IBUC	Impervious Built-Up Change
IBUCL	Confidence layer for Impervious Built-Up (10m)
IBUCS	Impervious Built-Up Change Support layer
IBUD_21	Calibrated built-up-densities derived from raw predictions
IBUD_21_ref	Reference built-up density
IBUD_21_resampled	Aggregate mean of IBUD to lower resolution
IBUD_model	Model coefficients for the U-NET model used to predict built-up degree
IBUDC	Continuous change of built-up degrees



IBUDC_resampled	Aggregate mean of IBUDC to lower resolution
IBUDCS	Continuous change support of built-up degrees
IBUDCS_resampled	Aggregate mean of IBUDCS to lower resolution
IMCC	Imperviousness Change Classified
IMCC_resampled	Classify change using IMDC_resampled
IMCCS	Imperviousness Change Classified Support layer
IMCCS_resampled	Classify change support using IMDCS_resampled
IMD	Imperviousness Density
IMD_18	Imperviousness Density 2018 update
IMD_21	Imperviousness Density 2021 update
IMD_21_resampled	Imperviousness Density, 100m
IMD_MSK_18	Mask IMD_18 and CLCP_18 for 2021 imperviousness density calibration
IMD_MSK_21	Mask RIMD_21 and CLCP_21 for 2021 imperviousness density calibration
IMD_calibration	Calibration coefficients to map NDVI to IMD
IMD_model	Model coefficients for the U-NET model used to predict imperviousness density
IMDC	Imperviousness Change
IMDC_resampled	Imperviousness Change
IMDCL	Confidence layer for Imperviousness Density (10m)
IMDCS	Imperviousness Change Support layer
IMDCS_resampled	Imperviousness Change Support layer
JRC	Joint Research Center
L1C	Sentinel-2 Level-1C
L2A	Sentinel-2 Level-2A
MSB	Microsoft Building Footprints
NDVI	Normalized Difference Vegetation Index
NDVI_p0_18	NDVI reference for 2018
NDVI_p0_21	NDVI reference for 2021
OSM	Open Street Map
PP_18	Sentinel-2 pre-processing: Harmonic coefficients and quantiles for 2018
PP_21	Sentinel-2 pre-processing: Harmonic coefficients and quantiles for 2021
PUM	Product User Manual
Q[N]	N% Quantile (e.g. Q20) (Pre-processing output)
R	Resolution
RIBUD	RAW imperviousness built-up density (AI model output)
RIBUD_21	Direct predictions from the U-NET model for 2021 (built-up)
RIBUD_21_sd	Standard deviations of RIBUD
RIMD	RAW imperviousness density (AI model output)
RIMD_21	Direct predictions from the U-NET model for 2021 (imperviousness)
RIMD_21_sd	Standard deviations of RIMD
RMSD	Root mean squared deviation



RMSD_18_21	Root mean squared deviation between 2018 and 2021
SAR	Synthetic-Aperture Radar
SBCC	Share of Built-Up Change Classified
SBCCS	Share of Built-Up Change Classified Support layer
SBU	Share of Built-Up
SBU_21	Share of Built-Up 2021 update
SBUC	Share of Built-Up Change
SBUCS	Share of Built-Up Change Support layer
SCL	Scene Classification Layer
U-Net	Convolutional Neural Network
UK	United Kingdom
UTM	Universal Transverse Mercator
VHR	Very High Resolution



Annex 1 - Cloud mask weighting

All processes relying on weighted approaches (Ordinary least squares and quantiles calculations) use the following assigned weighting based on the Sen2Cor cloud masks that come along with Sentinel-2 L2A products:

Table 6: Summary of weights applied during the pre-processing. Class code and class name refer to the Sen2Cor Scene-classification Layer (SCL)

Class code	Class name	Weight value
0	No data	0
1	Saturated or defective	0
2	Dark area pixels	0
3	Cloud shadows	10
4	Vegetation	50
5	Bare soils	50
6	Water	50
7	Cloud low probability	10
8	Cloud medium probability	5
9	Cloud high probability	1
10	Thin cirrus	1
11	Snow	1