



WorldCover

Product User Manual

Document Ref:	WorldCover_PUM_v1.0
Version:	1.0
Contract No:	4000128231/19/I-LG
Creation Date:	2020-09-01
Last Modified:	2020-10-15



WorldCover



DOCUMENT RELEASE SHEET

	Name	Signature	Date
Book Captain	Ruben Van De Kerchove (VITO)		15-10-2021
Contributing Authors	Daniele Zanaga (VITO) Wanda De Keersmaecker (VITO) Linlin Li (WU) Nandika Tsendbazar (WU) Myroslava Lesiv (IIASA)		
Approval (internal)	Ruben Van De Kerchove (VITO)		15-10-2021
Approval (ESA)	Olivier Arino		19-10-2021
Distribution	Public		

DOCUMENT CHANGE LOG

Issue	Issue date	Pages affected	Relevant information
0.9	2020-09-01	All	First version, 10% prototype
1.0	2021-10-15	All	Updated for the global product

Table of Content

TABLE OF CONTENT.....	3
LIST OF FIGURES.....	5
LIST OF TABLES	5
ACRONYMS AND ABBREVIATIONS.....	6
1 INTRODUCTION	7
1.1 Project objectives and approach	7
1.2 Purpose of the document	7
1.3 Overview	8
2 ALGORITHM	9
2.1 Summary	9
3 PRODUCT DESCRIPTION	11
3.1 File naming.....	11
3.2 File format.....	12
3.3 Product content	12
3.3.1 Map: Land cover classification.....	12
3.3.2 InputQuality: Quality layer.....	14
3.4 Product characteristics	14
3.4.1 Projection and grid Information	14
3.4.2 Spatial Information	14
3.4.3 Temporal information.....	14
4 KNOWN LIMITATIONS.....	15
5 DATA ACCESS AND CITATION	17
5.1 License.....	17
5.2 Citation.....	17
5.3 Delivery Mechanisms	17
5.3.1 Explore	17
5.3.2 Data Download	18
5.3.3 Google Earth Engine.....	20
5.4 Supporting files	20
5.4.1 Symbology.....	20
5.5 Contact.....	20

6 VALIDATION..... 21

7 REFERENCES..... 23

List of Figures

Figure 1: The WorldCover 2020 Map V100.....	8
Figure 2: Combined workflow to generate the WorldCover land cover product.....	9
Figure 3: Scheme of the 3 x 3 degree tiles.....	11
Figure 4: Screenshot of the WorldCover viewer.....	18
Figure 5: Spatial distribution of all validation sample sites for different (sub) continents	21

List of Tables

Table 1: Land Cover layer names in the filename.....	12
Table 2: Coding of the Map and definition of the classes.....	13
Table 3: Overall accuracy for the WorldCover product at the continental level and sample sites used (SSU- secondary sample units, PSU – primary sample units)	21

Acronyms and Abbreviations

CCI	Climate Change Initiative
CEOS	Committee on Earth Observation Satellites
COG	Cloud Optimized GeoTIFF
CGLS	Copernicus Global Land Service
CRS	Coordinate Reference System
DEM	Digital Elevation Model
EO	Earth Observation
ESA	European Space Agency
FAO	Food and Agriculture Organization of the United Nations
GIS	Geographic Information System
GLC	Global Land Cover
GRD	Ground Range Detected
LC	Land Cover
LCCS	Land Cover Classification System
UN	United Nations
PSU	Primary Sampling Unit
PUM	Product User Manual
SAR	Synthetic Aperture Radar
SCL	Scene Classification Layer
SSU	Secondary Sampling Unit
UTM	Universal Transverse Mercator
VI	Vegetation Indices

1 Introduction

1.1 Project objectives and approach

In an ever-changing environment, the need for accurate, timely and high-resolution information on land use/land cover and its changes has increased tremendously. However, until now, regional or continental land cover maps either used low-resolution images (>100 m) as input or were based exclusively on high-resolution optical Earth observation data, such as Sentinel-2 or Landsat. The use of SAR data such as Sentinel-1 to produce large area land cover maps is still in its infancy.

For this purpose and inspired by the 2017 WorldCover conference (<http://worldcover2017.esa.int>, attended by more than 400 participants) the European Space Agency (ESA) initiated the WorldCover project. The key outcome of this project was the release in October 2021 of a freely accessible global land cover product at 10 m resolution for 2020, based on both Sentinel-1 and Sentinel-2 data, containing 11 land cover classes and independently validated with a global overall accuracy of 74.4%.

A crucial aspect for WorldCover was the involvement of several end users active in different domains who provided primary input for all engineering aspects and followed the whole project workflow from design up to validation and uptake. Consequently, WorldCover intends to provide a substantial benefit to various user communities and expands the established global land cover base of users and the development of novel services.

1.2 Purpose of the document

This Product User Manual (PUM) describes the first version (V100) of the **WorldCover Land Cover products**, generated for **2020** and contains the required information to use these products. The PUM is the primary document that users have to read before handling the product. It gives an overview of the product characteristics, in terms of algorithm and technical characteristics and describes the known limitations and a summary of the independent validation results.

1.3 Overview



Figure 1: The WorldCover 2020 Map V100

In the following sections first the retrieval methodology of the WorldCover product is summarized (section 2) followed by a description of the technical properties of the product (section 3). Next, we list some identified known limitations of the product (section 4) as well as more information on how to access the product (section 5). Finally also a summary of the independent validation results are provided (section 6).

2 Algorithm

The ESA WorldCover global land cover product builds further on the GlobCover and CCI Land Cover experiences from the European Space Agency (Arino et al. 2008; ESA 2017). The algorithm used to generate the ESA WorldCover product is based on the algorithm to produce the dynamic yearly Copernicus Global Land Service Land Cover (CGLS-LC) map at 100 m resolution (Buchhorn et al., 2020a). The CGLS-LC workflow uses 100 m, 5-day, Proba-V data as an input which were re-processed to the Sentinel-2 UTM grid together with training data obtained at 10 m resolution. For the generation of the WorldCover product however both Sentinel-2 multi-spectral image data and Sentinel-1 C-band Synthetic Aperture Radar (SAR) data are used instead of Proba-V data.

2.1 Summary

The methodology for producing the WorldCover land cover product can be split in 3 separate workflows which together generate the WorldCover product (Figure 2): **Data pre-processing**, ii) **Classification** and iii) **Map generation**. The classification workflow is further subdivided in a training & prediction phase:

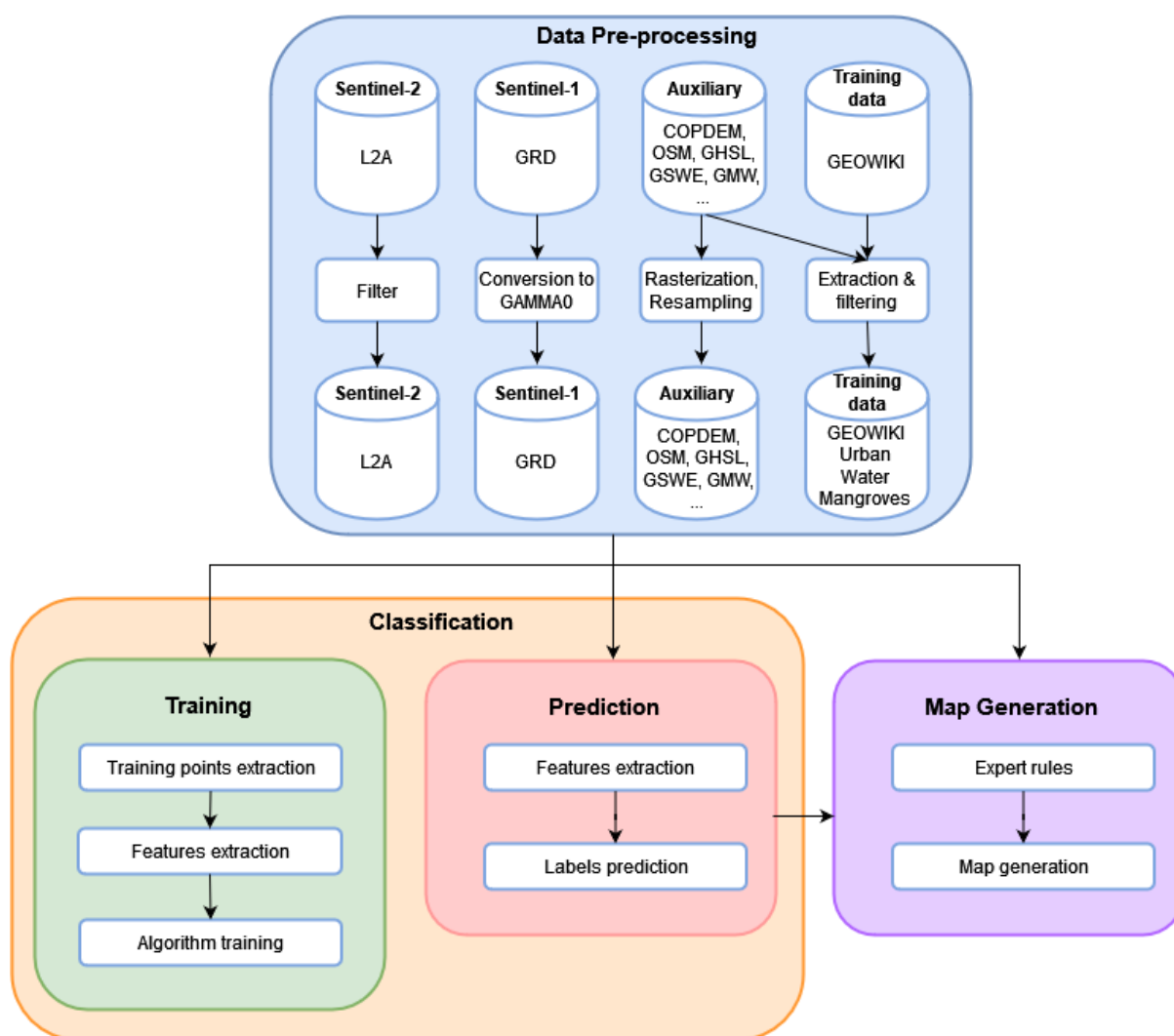


Figure 2: Combined workflow to generate the WorldCover land cover product

As a first step in the processing, Level 2A (L2A) and Ground Range Detected (GRD) products for Sentinel-2 (S2) and Sentinel-1 (S1) respectively, are selected and either filtered for cloud cover (Sentinel-2) or terrain corrected to Gamma0 backscatter time series (for Sentinel-1) which are geocoded to the Sentinel-2 grid. To reduce data redundancy, a maximum of one observation per day per tile is processed. In case of multiple observations per day, the first observation with the highest number of valid data pixels is kept.

Next, clouds, cloud shadows, cirrus and saturated pixels are removed in the Sentinel-2 reflectance bands using the scene classification layer (SCL) of the L2A product. Ten days median composites (using a 20 days moving window) are computed from the band time series, eliminating further noise, or pixels not correctly masked by the SCL layer. The remaining missing values are then filled by estimating their value from a linear interpolation between the closest available values in the time series. After the time series are pre-processed, additional vegetation indices (VI) for each time series step are calculated.

Starting from those cleaned Sentinel-2 surface reflectance and VI time series, descriptive statistics are computed for each band and VI. These includes the 10th, 50th, 90th percentiles and the interquartile range (75th percentile – 25th percentile). In addition to these statistics, long range averaged timestamps are used as features together with the quantiles to introduce time dynamics information in the classification. For this purpose, the NDVI timeseries is downsampled to 6 timestamps using Fourier methods. The same features are calculated on the S1 VV and VH backscatter time series as well. For S1 bands, an additional multitemporal speckle filter is used before compositing the timeseries. Altitude and slope are extracted from the Copernicus Global 30 m Digital Elevation Model (DEM) (Fahrland et al., 2019) and used as features as well. A series of spatial localizing features are also included: the position of each pixel (latitude & longitude) and the biome and realm features derived from the Ecoregions2017 layer (Dinerstein et al., 2017). Overall, 90 features are extracted from Sentinel-2, 16 features from Sentinel-1, 2 features from the DEMs and 23 positional features for a total of 131 features.

In the training phase, these features are extracted for each training location, from which training points are randomly sampled, with a maximum of 20 points per class per location. The training points selected are randomly splitted by location in training and test sets and used for training different models (scenarios) with a gradient boosting decision tree algorithm (CatBoost) (Prokhorenkova et al., 2018).

In the classification phase, exactly the same steps are carried out until the features are extracted, at that point the models/scenarios trained during the training phase are loaded and used to predict the labels of the input data together with the class probabilities.

Finally, the different scenarios are combined into a final land cover map through the application of different expert rules and subsequently tiled into 3 x 3 degree tiles in geographic projection (EPSG:4326). Some of these expert rules use the following auxiliary datasets (OpenStreetMap (OpenStreetMap Contributors, 2020), Global Surface Water Explorer (Pekel et al., 2016), Global Mangrove Watch (Bunting et al., 2018) , Global Human Settlement Layer (Corbane et al., 2020a, 2020b), AgERA5 historic and near real time forcing data, CGLOPS permanent Ice mask (Buchhorn et al., 2020b) in order to modulate the probabilities provided by the classifier in order to produce a better final prediction. It is important to note that these auxiliary datasets are not imprinted in the final land cover map but only support the decision of the final class attribution.

3 Product description

The ESA WorldCover products are provided per 3 x 3 degree tile (Figure 3), 2651 in total. Each tile contains a set of 2 Cloud Optimized GeoTIFF (COG) files corresponding to the following data layers:

- **Map:** Land cover map with 11 classes
- **InputQuality:** Three band GeoTIFF providing three per pixel quality indicators of the Sentinel-1 and Sentinel-2 input data

Tiles are provided in EPSG:4326, geographic latitude/longitude CRS.

3.1 File naming

The GeoTIFF files follow this naming convention:

ESA_WorldCover_10m_<YEAR>_<VERSION>_<TILE>_<LAYER>.tif e.g.
ESA_WorldCover_10m_2020_v100_S48E036_Map.tif

where

- “ESA_WorldCover_10m” indicates this is the 10 m resolution ESA WorldCover product
- <YEAR> indicates the reference year (observation period) in four digits (e.g. 2020)
- <VERSION> shows the product version. The version denoted as vMmr (e.g. v100), with ‘M’ representing the major version (e.g. v1), ‘m’ the minor version (starting from 0) and ‘r’ the production run number (starting from 1)
- <TILE> the designation of the 3 x 3 degree tile, composed of the 2-digit latitude and 3-digit longitude of the lower-left corner (Figure 3)
Example: S48E036 for the tile covering the area from 36E to 39E and 48S to 45S.
- <LAYER> gives the name of the data layer (see Table 1; e.g. Map)

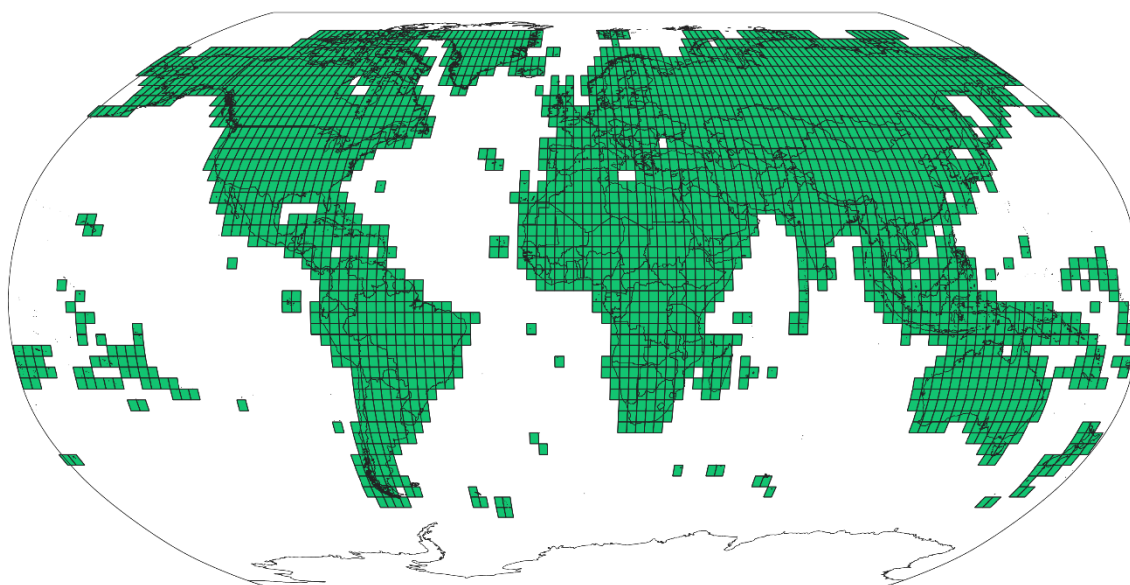


Figure 3: Scheme of the 3 x 3 degree tiles

Table 1: Land Cover layer names in the filename

Layer in filename	Description
Map	Main discrete classification according to FAO LCCS scheme, see Table 2
InputQuality	Per pixel quality indicator showing the quality of the EO input data 3 band GeoTIFF, see 3.3.2

3.2 File format

The data layers of the WorldCover product are provided as **Cloud Optimized GeoTIFF (COG)** files that are internally compressed (DEFLATE), have an internal tiling of size 1024, include overviews on levels 2, 4, 8, 16, 32 and 64 for faster loading in Geographic Information System (GIS) and include standard metadata attributes.

All land cover layers are stored as **single bytes (Map)** or **UINT16 (InputQuality)** per pixel, without scaling or offset.

3.3 Product content

3.3.1 Map: Land cover classification

The discrete classification **map** provides **11 classes** and is defined using the Land Cover Classification System (LCCS) developed by the United Nations (UN) Food and Agriculture Organization (FAO). The UN-LCCS system was designed as a hierarchical classification, which allows adjusting the thematic detail of the legend to the amount of information available

Table 2: Coding of the Map and definition of the classes

Map code	Land Cover Class	LCCS code	Definition	Color code (RGB)
10	Tree cover	A12A3 // A11A1 A24A3C1(C2)- R1(R2)	This class includes any geographic area dominated by trees with a cover of 10% or more. Other land cover classes (shrubs and/or herbs in the understorey, built-up, permanent water bodies, ...) can be present below the canopy, even with a density higher than trees. Areas planted with trees for afforestation purposes and plantations (e.g. oil palm, olive trees) are included in this class. This class also includes tree covered areas seasonally or permanently flooded with fresh water except for mangroves.	0,100,0
20	Shrubland	A12A4 // A11A2	This class includes any geographic area dominated by natural shrubs having a cover of 10% or more. Shrubs are defined as woody perennial plants with persistent and woody stems and without any defined main stem being less than 5 m tall. Trees can be present in scattered form if their cover is less than 10%. Herbaceous plants can also be present at any density. The shrub foliage can be either evergreen or deciduous.	255, 187, 34
30	Grassland	A12A2	This class includes any geographic area dominated by natural herbaceous plants (Plants without persistent stem or shoots above ground and lacking definite firm structure): (grasslands, prairies, steppes, savannahs, pastures) with a cover of 10% or more, irrespective of different human and/or animal activities, such as: grazing, selective fire management etc. Woody plants (trees and/or shrubs) can be present assuming their cover is less than 10%. It may also contain uncultivated cropland areas (without harvest/ bare soil period) in the reference year	255, 255, 76
40	Cropland	A11A3(A4)(A5) // A23	Land covered with annual cropland that is sowed/planted and harvestable at least once within the 12 months after the sowing/planting date. The annual cropland produces an herbaceous cover and is sometimes combined with some tree or woody vegetation. Note that perennial woody crops will be classified as the appropriate tree cover or shrub land cover type. Greenhouses are considered as built-up.	240, 150, 255
50	Built-up	B15A1	Land covered by buildings, roads and other man-made structures such as railroads. Buildings include both residential and industrial building. Urban green (parks, sport facilities) is not included in this class. Waste dump deposits and extraction sites are considered as bare.	250, 0, 0
60	Bare / sparse vegetation	B16A1(A2) // B15A2	Lands with exposed soil, sand, or rocks and never has more than 10 % vegetated cover during any time of the year	180, 180, 180
70	Snow and Ice	B28A2(A3)	This class includes any geographic area covered by snow or glaciers persistently	240, 240, 240
80	Permanent water bodies	B28A1(B1) // B27A1(B1)	This class includes any geographic area covered for most of the year (more than 9 months) by water bodies: lakes, reservoirs, and rivers. Can be either fresh or salt-water bodies. In some cases the water can be frozen for part of the year (less than 9 months).	0, 100, 200
90	Herbaceous wetland	A24A2	Land dominated by natural herbaceous vegetation (cover of 10% or more) that is permanently or regularly flooded by fresh, brackish or salt water. It excludes unvegetated sediment (see 60), swamp forests (classified as tree cover) and mangroves see 95)	0, 150, 160
95	Mangroves	A24A3C5-R3	Taxonomically diverse, salt-tolerant tree and other plant species which thrive in intertidal zones of sheltered tropical shores, "overwash" islands, and estuaries.	0, 207, 117
100	Moss and lichen	A12A7	Land covered with lichens and/or mosses. Lichens are composite organisms formed from the symbiotic association of fungi and algae. Mosses contain photo-autotrophic land plants without true leaves, stems, roots but with leaf-and stemlike organs.	250, 230, 160

3.3.2 InputQuality: Quality layer

Per pixel quality indicator showing the quality of the Earth Observation (EO) input data. The layer is a 3 band GeoTIFF with

- Band 1: Number of Sentinel-1 GAMMA0 observations used in the classification workflow
- Band 2: Number of Sentinel-2 L2A observations used in the classification workflow
- Band 3 Percentage (0-100) of invalid S2 observations discarded in the classification workflow (after cloud and cloud shadow filtering).

Combining Band 2 and Band 3 yields the total absolute number of valid Sentinel-2 L2A observations used in the classification workflow.

3.4 Product characteristics

3.4.1 Projection and grid Information

The WorldCover prototype products are delivered in a regular latitude/longitude grid (EPSG:4326) (Figure 3) with the ellipsoid WGS 1984 (Terrestrial radius=6378 km). The resolution of the grid is 1°/12000 or approximately 10 m at equator for the Map product. The InputQuality layer is provided at 1/2000 or approximately 60 m at equator

3.4.2 Spatial Information

The WorldCover product is constrained to all land areas observed by Sentinel-2 in 2020 except for areas in Antarctica (see Figure 3). In the north, the WorldCover product is limited to 82.75°N.

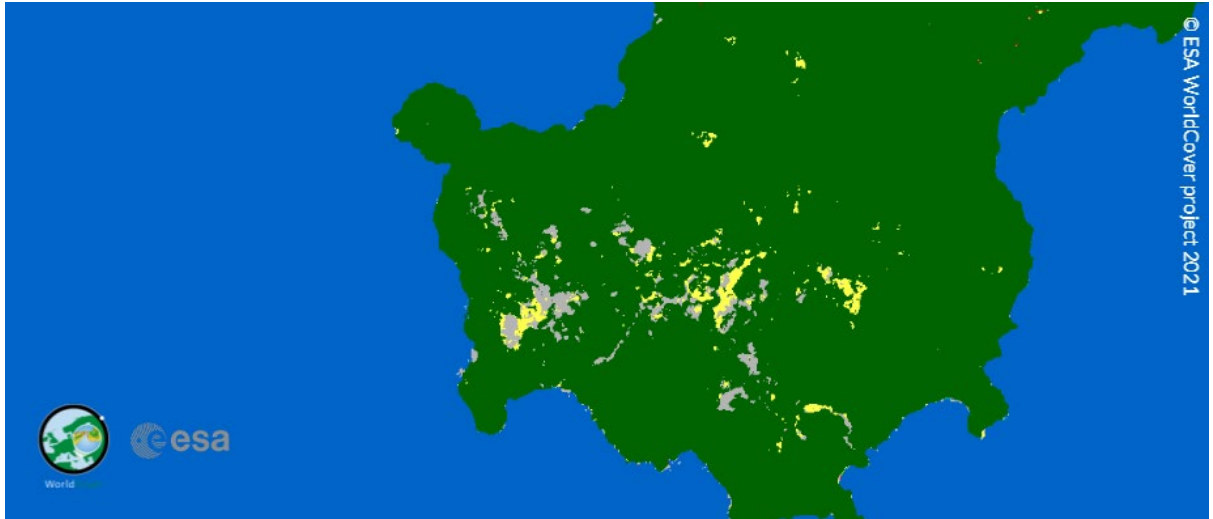
3.4.3 Temporal information

The products represent the **land cover for 2020** (from 01 January to 31 December). Sentinel-1 and Sentinel-2 data covering the reference year is used in its processing.

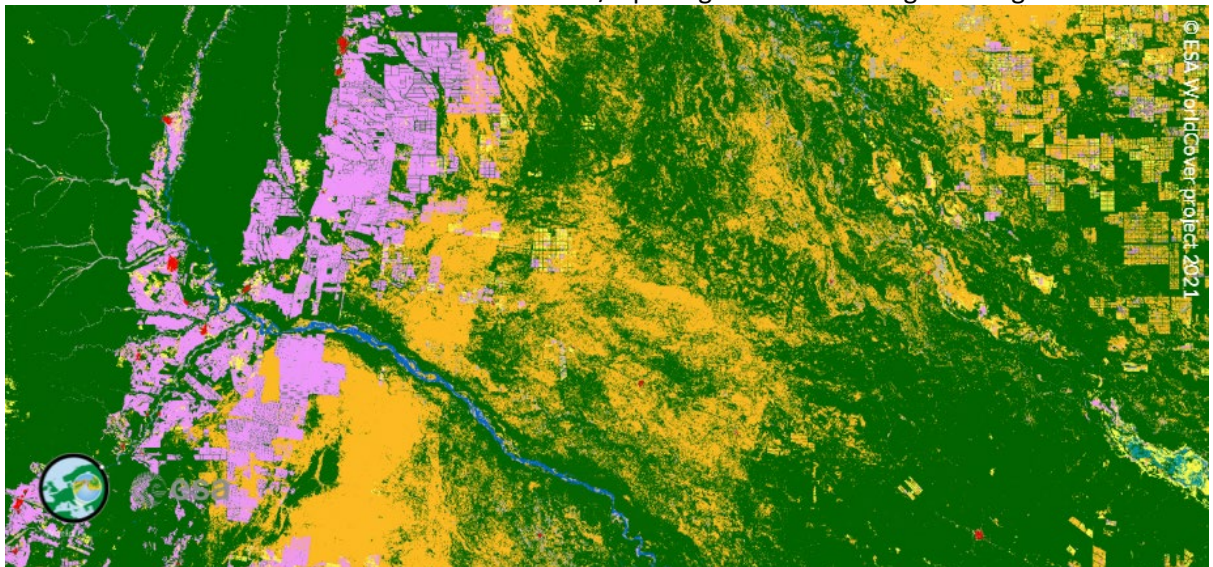
4 Known limitations

We identified a non-exhaustive list of limitations in the WorldCover map layer:

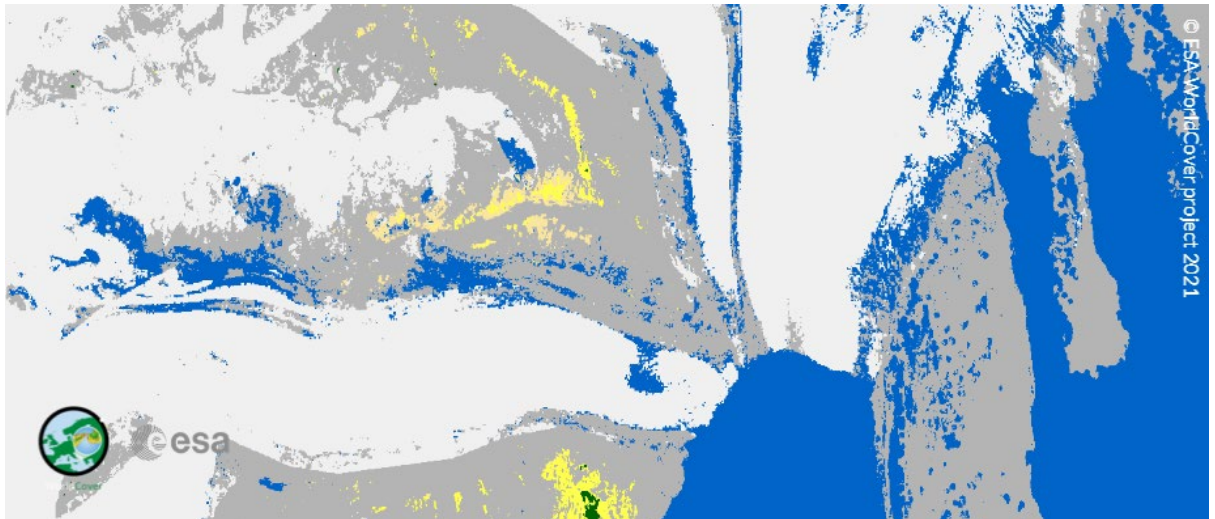
- Cloud artefacts in areas with a high cloud cover % (see e.g. 7.38°E; 1.58°N): Despite the use of Sentinel-1, cloud artefacts might still appear in areas with persistent cloud cover.



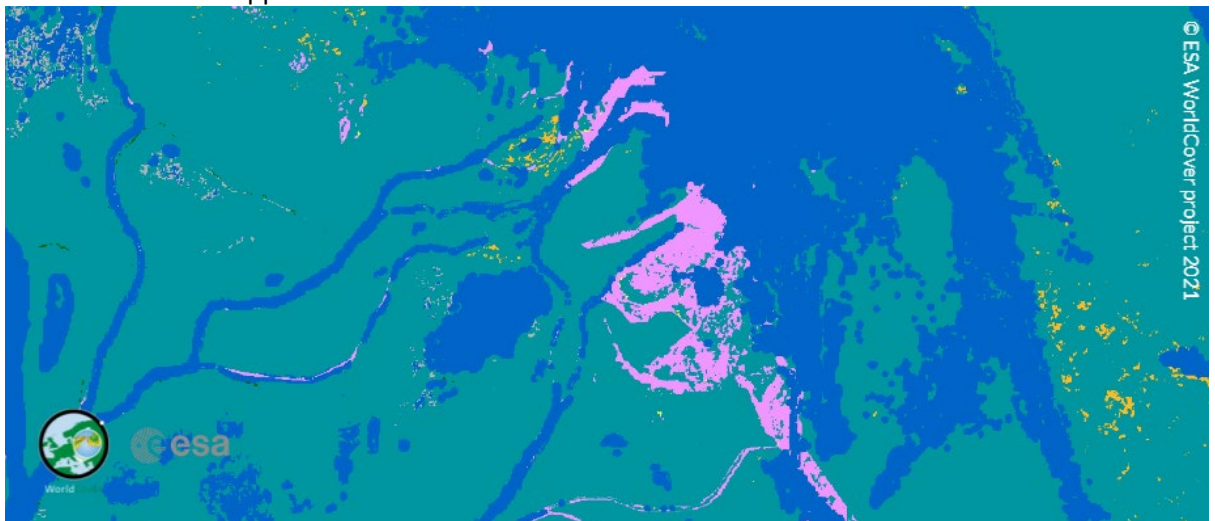
- Hard borders (see e.g. 63°W, 23°S): In areas with a high confusion between land cover classes, Sentinel-2 orbit borders or block borders might be visible due to sensitivity of the model inference to the different amount of observations/input signal between neighbouring areas.



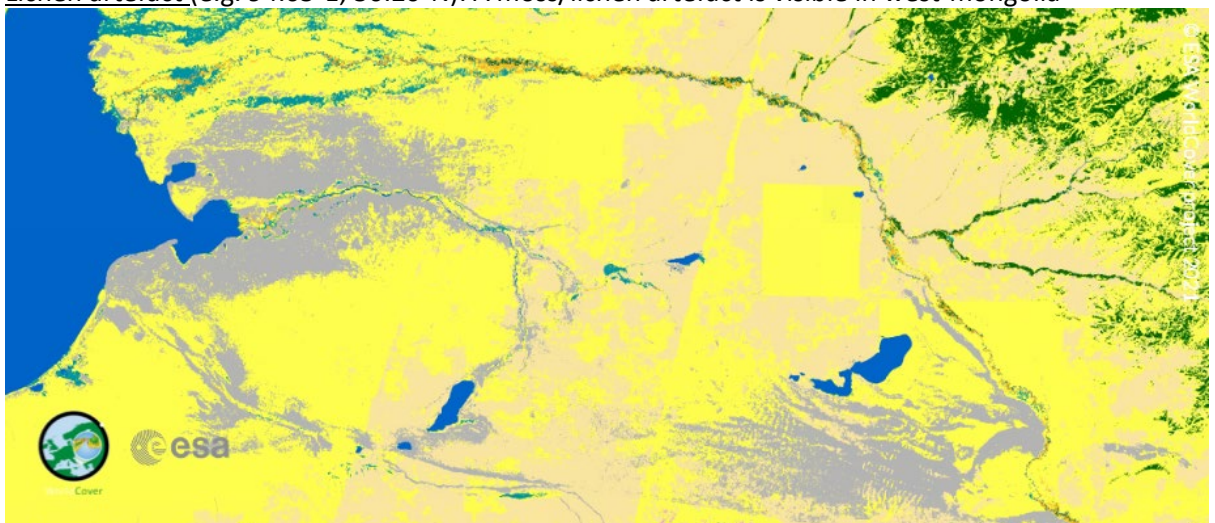
- Water on glaciers/ice sheets (e.g. 73,28°W , 49.89°S): Some areas in the ablation zone of glaciers are sometimes misclassified as water bodies.



- Irrigated agriculture vs wetland confusion (e.g. 116.30°E, 29.05°N): Due to the high spectral similarity between irrigated agriculture and herbaceous wetland, sometimes misclassifications happen between those classes



- Lichen artefact (e.g. 94.65°E; 50.20°N): A moss/lichen artefact is visible in west-mongolia



In order to improve the map, users are encouraged to give feedback on artefacts/misclassifications via a dedicated feedback tool embedded in WorldCover viewer (see 5.3.2)

5 Data Access and citation

5.1 License

The WorldCover product is provided free of charge, without restriction of use. For the full license information see the [Creative Commons Attribution 4.0 International License](#)

Publications, models and data products that make use of these datasets must include proper acknowledgement, including citing the datasets and the journal article as in the following citation.

5.2 Citation

A reference publication on the product will be available in the next month.

If you are using the data as a layer in a published map, please include the following attribution text: '© ESA WorldCover project 2020 / Contains modified Copernicus Sentinel data (2020) processed by ESA WorldCover consortium'

To cite these maps as data source in your publication, please add:

Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., Cartus, O., Santoro, M., Fritz, S., Georgieva, I., Lesiv, M., Carter, S., Herold, M., Li, Linlin, Tsendbazar, N.E., Ramoino, F., Arino, O., 2021. ESA WorldCover 10 m 2020 v100. <https://doi.org/10.5281/zenodo.5571936>

5.3 Delivery Mechanisms

The ESA WorldCover 10 m 2020 V100 product is delivered in **3x3 degree tiles** as **Cloud Optimized GeoTIFFs (COGs)** in **EPSG:4326 projection** (geographic latitude/longitude CRS). There are **2651 tiles**

Each tile contains 2 data layers:

- *Map*: Land cover map with 11 classes, a total of approximately **117 GB**.
- *InputQuality*: Three band GeoTIFF providing three per pixel quality indicators of the Sentinel-1 and Sentinel-2 input data

The dataset can be accessed and/or downloaded in different ways:

5.3.1 Explore

5.3.1.1 WorldCover viewer

The [WorldCover viewer](#) is a simple web-mapping tool that allows users to navigate the globe visualizing the ESA WorldCover product without installing any software. Furthermore, it also allows users to compare the map with Sentinel-1 and Sentinel-2 median composites for 2020, view land cover statistics per administrative area and provides the possibility to give feedback on the product. Finally, users can also use the viewer to download the WorldCover product either by administrative area, or customized region of interest.

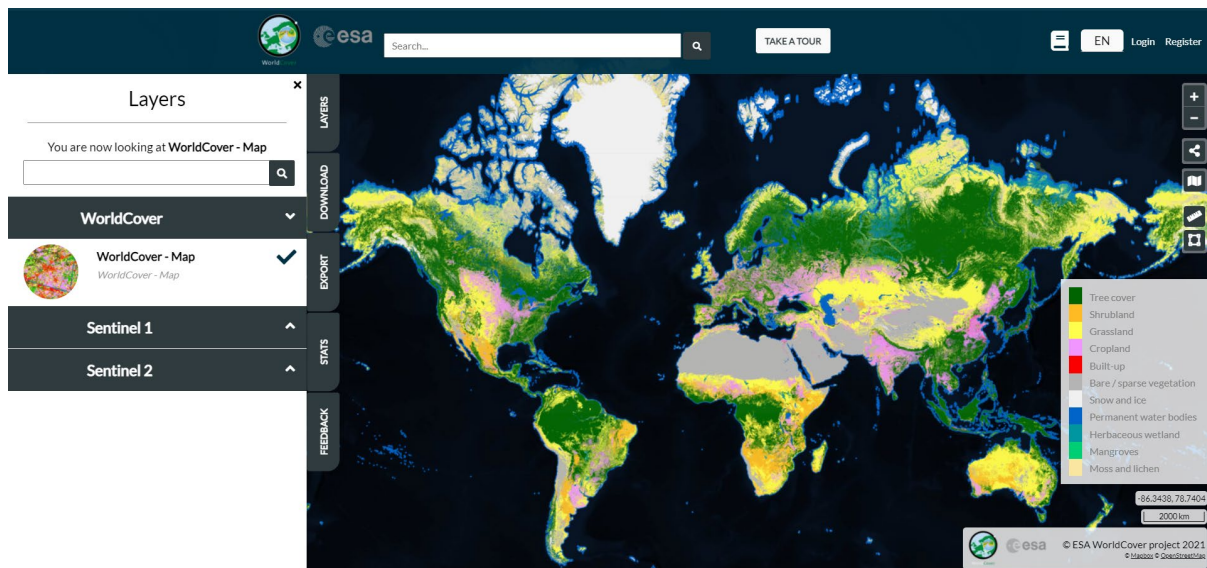


Figure 4: Screenshot of the WorldCover viewer

5.3.1.2 Google Earth Engine App

A dedicated [Earth Engine App](#) has been made which allows users to visualize the ESA WorldCover product and compare with Sentinel-1 and Sentinel-2 composites for 2020.

5.3.1.3 Web Map Services

The ESA WorldCover product can also be used within other websites or GIS clients by 'Web Map Services'. These services provide a direct link to the cached images and are the best option if you simply want to map the data and produce cartographic products. They are not suitable for analysis as the data are represented only as RGB images.

WMTS: <https://services.terrascope.be/wmts/v2>

WMS: <https://services.terrascope.be/wms/v2>

Layer: WORLDCOVER_2020_MAP

5.3.2 Data Download

The dataset can be downloaded from:

1. **WorldCover viewer** (Login required)
 - Navigate to the [WorldCover viewer](#)
 - Register for VITO's Terrascope platform (see the link in the top right)
 - Log in the [WorldCover viewer](#) with your Terrascope credentials
 - From the 'Download' tab in the WorldCover viewer, zoom to an area of interest or draw a polygon to retrieve the list of tiles covering that area. The platform will zip the tiles and initiate the download.
2. **Terrascope with Python** (Login required)

Tiles can be downloaded programmatically using Python and the `terrascatalogueclient` package. Please use your Terrascope credentials to download (see 1)

 - Install the `terrascatalogueclient`, see <https://vitobelgium.github.io/terrascatalogueclient/installation.html> or run:

```
pip install --extra-index-url https://artifactory.vgt.vito.be/api/pypi/python-packages/simple terrascatalogueclient
```

- Customize and run the following python script:

```
from terracatalogueclient import Catalogue

# create catalogue object and authenticate
catalogue = Catalogue().authenticate()

# search for products in the WorldCover collection
products =
catalogue.get_products("urn:eop:VITO:ESA_WorldCover_10m_2020_V1")

# download the products to the given directory
catalogue.download_products(products, "downloads")

from shapely.geometry import Polygon
from terracatalogueclient import Catalogue

### Authenticate to the Terrascope platform (registration required)
# create catalogue object and authenticate interactively with a browser
catalogue = Catalogue().authenticate()

# or authenticate with username and password
# catalogue = catalogue.authenticate_non_interactive(username, password)

### Filter catalogue
# search for all products in the WorldCover collection
# products =
catalogue.get_products("urn:eop:VITO:ESA_WorldCover_10m_2020_V1")

# or filter to a desired geometry, by providing it as an argument to
`get_products`

bounds = (3, 33, 15, 45)
geometry = Polygon.from_bounds(*bounds)
products =
catalogue.get_products("urn:eop:VITO:ESA_WorldCover_10m_2020_V1",
geometry=geometry)

### Download
# download the products to the given directory
catalogue.download_products(products, "downloads")
```

3. AWS S3 bucket

The products are hosted on the `esa-worldcover` S3 bucket in `eu-central-1` region. To download all the products:

- install the AWS command line interface, see <https://docs.aws.amazon.com/cli/latest/userguide/cli-chap-install.html> or run `pip install awscli` in a python environment.
- run `aws s3 sync --no-sign-request s3://esa-worldcover/v100/2020/map/ /local/path` (please modify `/local/path` to the desired download location)

4. Zenodo

The dataset can be downloaded as a single `tar.gz` file from [Zenodo](#).

Please note that the *Map* data layer can be downloaded from all 4 options. The *InputQuality* layer is so far only available from options 1 or 2.

5.3.3 Google Earth Engine

The WorldCover data can also be accessed and used in the Google Earth Engine platform - for more information see [here](#). The WorldCover data is available with Asset ID `ESA/WorldCover/v100`

5.4 Supporting files

5.4.1 Symbology

The WorldCover map layer contains a colormap which will display the files in desktop GIS tools (such as QGIS or ArcGIS) using the symbology that has been used in the WorldCover viewer. However, these colormaps do not contain the labels for the values. These can be added to the files by using the symbology files for either [QGIS](#) or [ArcGIS](#).

- To use the *.qml file in QGIS: Double click on the layer to open the Layer properties dialog. At the bottom of the dialog, click Style | Load Style... Navigate to the location of the *.qml file and click Open.
- To use the *.lyr file in ArcGIS: Double click on the layer to open the Layer properties dialog. Click on the Symbology tab and change the rendered to Unique Values. Click 'Yes' when prompted to build an attribute table. Click on the Import Symbology button in the top right. Navigate to the location of the *.lyr file and click Add

Besides the standard WorldCover colormap also colormap files with a symbology used in the CCI land cover maps is provided. [QGIS](#) or [ArcGis](#)

There are no symbology files for the InputQuality layer.

5.5 Contact

More information and documentation about the product is available from the WorldCover web site at <https://esa-worldcover.org>

If you have any feedback on the ESA WorldCover product please contact remotesensing@vito.be

6 Validation

The WorldCover prototype product was statistically validated using an independent validation dataset. The validation followed the CEOS (the Committee on Earth Observation Satellites) Validation guidelines (Stage 3).

The statistical validation used the multi-purpose Global Land Cover Validation dataset developed for Copernicus Global Land Service- Land Cover 100 product (the CGLS-LC validation dataset) which is suitable for validation of 10m resolution global land cover maps. This dataset was updated for the reference year 2020 by revisiting random and targeted (for land cover change) subsets.

The validation dataset includes more than 21 000 primary sampling units (PSUs) spread over seven (sub)continents selected based on a stratified random sampling (Figure 5). Each PSU consists of 100 secondary sample units (SSU) which correspond to Sentinel-2 10m resolution pixels.

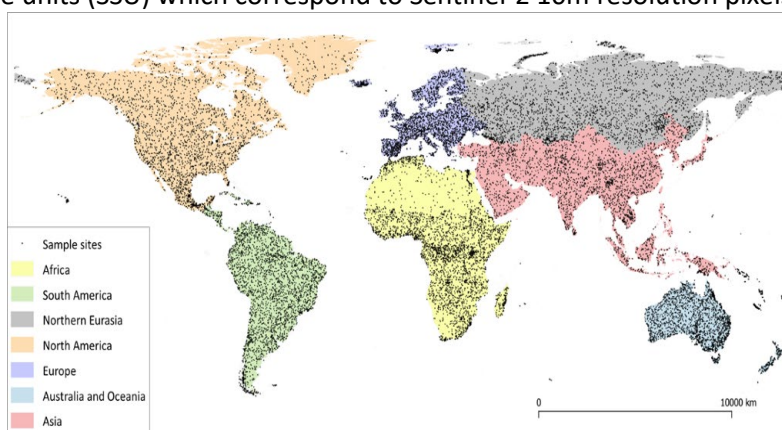


Figure 5: Spatial distribution of all validation sample sites for different (sub) continents

The validation results showed that the overall accuracy of the WorldCover product is $74.4 \pm 0.1\%$ for 2020. In terms of land cover types, tree cover and snow/ice, cropland, water body, and bare/sparse vegetation classes had high accuracies, while shrubs, herbaceous wetland, and moss/lichen classes were mapped with lower accuracies.

Overall accuracy at the continental level is mostly above 70%, with the highest accuracy of 80.7% for Asia and the lowest accuracy of 67.5% for Oceania (Table 3). The lower accuracy for Oceania could be due to high shrubland and grassland confusions in the open shrublands of Australia.

Table 3: Overall accuracy for the WorldCover product at the continental level and sample sites used (SSU- secondary sample units, PSU – primary sample units)

	Overall accuracy	Total SSU	Total PSU
Global	74.4 ± 0.1	1,935,650	21,624
Africa	73.6 ± 0.2	313,709	3,599
Europe	76.8 ± 0.2	282,703	3,118
Eurasia	70.2 ± 0.2	274,280	3,003
Asia	80.7 ± 0.1	284,481	3,064
Oceania	67.5 ± 0.2	248,205	2,961
North America	72.2 ± 0.2	258,486	2,880
South America	76.1 ± 0.2	273,786	2,999

The product was also assessed in terms of spatial accuracy to highlight areas with higher overall quality and areas that have higher uncertainty. Detailed information on the validation of this product and comparison results can be found in **the product validation report**. Overall, the WorldCover product shows promising results in characterizing the World's land cover at 10m resolution.

7 References

Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R., Weber, J.L. (2008). GlobCover the most detailed portrait of Earth. *ESA Bulletin*, 136, 25-31

Buchhorn, M., Lesiv, M., Tsendbazar, N.-E., Herold, M., Bertels, L., Smets, B. (2020). Copernicus Global Land Cover Layers—Collection 2. *Remote Sens.* 12(6), 1044.

Buchhorn, M., Bertels, L., Smets, B., De Roo, B., Lesiv, M., Tsendbazar, N. E., Masiliunas, D., & Li, L. (2020). Copernicus Global Land Service: Land Cover 100m: Version 3 Globe 2015-2019: algorithm theoretical basis document.

Bunting P., Rosenqvist A., Lucas R., Rebelo L-M., Hilarides L., Thomas N., Hardy A., Itoh T., Shimada M. and Finlayson C.M. (2018). The Global Mangrove Watch – a New 2010 Global Baseline of Mangrove Extent. *Remote Sensing* 10(10): 1669. doi: 10.3390/rs1010669.

Corbane, Christina; Sabo, Filip; Politis, Panagiotis; Syrris Vasileos (2020): GHS-BUILT-S2 R2020A - built-up grid derived from Sentinel-2 global image composite for reference year 2018 using Convolutional Neural Networks (GHS-S2Net). European Commission, Joint Research Centre (JRC) PID: <http://data.europa.eu/89h/016d1a34-b184-42dc-b586-e10b915dd863>, doi:10.2905/016D1A34-B184-42DC-B586-E10B915DD863

Corbane, C., Syrris, V., Sabo, F. et al. Convolutional neural networks for global human settlements mapping from Sentinel-2 satellite imagery. *Neural Comput & Applic* (2020). doi:10.1007/s00521-020-05449-7

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., Hahn, N., Palminteri, S., Hedao, P., Noss, R., Hansen, M., Locke, H., Ellis, E. C., Jones, B., Barber, C. V., Hayes, R., Kormos, C., Martin, V., Crist, E., Sechrest, W., Price, L., Baillie, J. E. M., Weeden, D., Suckling, K., Davis, C., Sizer, N., Moore, R., Thau, D., Birch, T., Potapov, P., Turubanova, S., Tyukavina, A., de Souza, N., Pintea, L., Brito, J. C., Llewellyn, O. A., Miller, A. G., Patzelt, A., Ghazanfar, S. A., Timberlake, J., Klöser, H., Shennan-Farpón, Y., Kindt, R., Lillesø, J. B., van Breugel, P., Graudal, L., Vogé, M., Al-Shammari, K. F., Saleem, M. (2017). An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm, *BioScience*, 67(6), 534–545.

ESA. Land Cover CCI Product User Guide Version 2. Tech. Rep. (2017). Available at: maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf

Fahrland, E., Jacob, P., Schrader, H., Kahabka, H. (2019). Copernicus DEM Product Handbook. AIRBUS Defence and Space, Campaign ID: GEO.2018-1988-2, Version 1.0, pp. 27.

Jean-Francois Pekel, Andrew Cottam, Noel Gorelick, Alan S. Belward, High-resolution mapping of global surface water and its long-term changes. *Nature* 540, 418-422 (2016). (doi:10.1038/nature20584)

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems*, 6638-6648.

OpenStreetMap contributors. (2020) Planet dump retrieved from <https://planet.openstreetmap.org>.