



Ref. Ares(2023)697862 - 31/01/2023



# Data Annotation Report

Deliverable 3.5

Work package: WP3

Dissemination level: Public

Lead partner: KTH

Authors: Sebastian Hafner, Yifang Ban

Due date: 31.01.2023

Submission date: 31.01.2023

Deliverable	Data annotation
Deliverable No.	D3.5
Work Package	3
Dissemination Level	Public
Author(s)	Sebastian Hafner (KTH)
Co-Author(s)	Yifang Ban (KTH)
Date	31/01/2023
File Name	WP3_D3_5_Data_Anootation_Report_KTH_v03
Status	Revision
Revision	V03
Reviewed by (if applicable)	CTI
Information to be used for citations of this report	Hafner S., Yifang B. (2023): <i>Data annotation report</i> , D3.5, HARMONIA. Horizon 2020 Grant Agreement No 101003517, European Commission, 23 pp.

The sole responsibility for the content of this publication lies with the authors. It does not necessarily represent the opinion of the European Union. Neither the EASME nor the European Commission are responsible for any use that may be made of the information contained therein.



The HARMONIA project has received funding from the EU Horizon 2020 research and innovation programme under agreement No. 101003517.

## CONTACT

Email: [contact@harmonia-project.eu](mailto:contact@harmonia-project.eu)

Website: <http://harmonia-project.eu/>

# Table of content

Figures .....	4
Document revision history .....	5
Partners organisations.....	5
Abbreviations.....	5
Executive Summary .....	6
1. Objective of Task 3.5 Data Annotation.....	7
2. Data Description .....	7
2.1 Earth Observation Data .....	7
2.1.1 Sentinel-1 SAR .....	7
2.1.2 Sentinel-2 MSI.....	7
2.1.3 Landsat 8 OLI .....	7
2.2 Existing Datasets.....	8
2.2.1 Microsoft’s Building Footprints .....	8
2.2.2 SpaceNet7.....	8
2.2.3 Urban Atlas .....	9
2.2.4 CORINE Land Cover.....	9
2.2.5 Local Climate Zones .....	10
2.2.6 Cloud and Cloud Shadows Datasets .....	10
3. Earth Observation Data Processing .....	11
3.1 Sentinel-1 SAR .....	11
3.2 Sentinel-2 MSI .....	12
3.3 Landsat 8 OLI .....	12
4. Data Annotation Methods.....	12
4.1 Built-Up Area Labels .....	12
4.2 Built-Up Area Change Labels .....	14
4.3 Urban Land Cover Labels.....	15
4.3.1 Urban Atlas Labels .....	15
4.3.2 Local Climate Zone Labels.....	16
4.4 Urban Land Cover Change Labels.....	19
4.4.1 Urban Atlas Change Labels .....	19
4.4.2 CORINE Land Cover Change Labels.....	19
4.4.3 Cloud and Cloud Shadow Labels.....	19
4.5 Wildfire Labels.....	22

References ..... 23

## Figures

Figure 1: Train and test sites of the SpaceNet7 dataset (Van Etten, et al., 2021). ..... 8

Figure 2: Classification scheme used for the Urban Atlas. .... 9

Figure 3: Classification scheme used for the CORINE Land Cover product. .... 10

Figure 4: The local climate zone classifications scheme adopted from (Stewart, Oke, & Krayenhoff, 2014). 10

Figure 5: Overview of the data preparation workflow to retrieve satellite data from Sentinel-1 SAR, Sentinel-2 MSI, and Landsat 8 OLI. .... 11

Figure 6: Training and validation sites. Labelled data is only used in the United States, Canada and Australia. .... 13

Figure 7: Number of pixels in the training and validation sets..... 13

Figure 8: Sixty test sites grouped into the source domain and target domain. The target domain sites are further grouped into 5 regions, namely Europe (EU), Latin America (LA), Sub-Saharan Africa (SSA), Islamic World (IS), and Asia. Numbers in brackets denote the number of sites comprising a regional group..... 14

Figure 9: Number of pixels for the test set by region. .... 14

Figure 10: Built-up area change sample. NC (black) denotes not change..... 15

Figure 11: Overview of the pre-processing workflow used to rasterize and remap vector data using QGIS. 15

Figure 12: Local climate zone labels for Athens, Brussels and Milan (from left to right) for 2017. The legend was adopted from (Zhu, et al., 2022). .... 16

Figure 13: Sentinel-1 (S1) SAR data, Sentinel-2 (S2) MSI data and Urban Atlas (UA) data for 2018 for the four pilot cities. See legend in Figure 2. .... 17

Figure 14: Sentinel-1 (S1) SAR data, Sentinel-2 (S2) MSI data and Urban Atlas (UA) data for 2018 for the four pilot cities. See legend in Figure 2. .... 18

Figure 15: Landsat 8 (L8) OLI data, Sentinel-2 (S2) MSI data and Urban Atlas change labels between 2012 and 2018 for the four pilot cities. .... 20

Figure 16: Landsat 8 (L8) OLI data, Sentinel-2 (S2) MSI data and CORINE land cover change labels between 2012 and 2018 for the four pilot cities. .... 21

Figure 17: Overview of the pre-processing workflow for wildfire labels. .... 22

Figure 18: Example of the prepared EO data and the corresponding labels for a wildfire close to Athens. ... 22

## Document revision history

Version	Date	Modification reason	Modified by
1	12.01.2023	Initial draft	KTH
2	30.01.2023	Revision based on CTI's review	KTH
3	31.01.2023	KTH internal revision	KTH

## Partners organisations

No.	Name	Short name	Country
1	KTH Royal Institute of Technology	KTH	Sweden
2			
3			
4			

## Abbreviations

CH	Chapter
D	Deliverable
EU	European Union
GI	Green Infrastructure
NBS	Nature-Based Solutions
SDG	Sustainable Development Goal
WP	Work Package

## Executive Summary

Task 3.5 (M8-M20) aims to investigate various benchmark datasets that are freely available to obtain sufficient labelled data to train Deep Learning (DL) models. The trained DL models can be deployed to the four European pilot cities, namely Sofia, Milan, Athens (Piraeus), and Brussels (Ixelles), in order to derive urban land cover information and their changes from Earth observation (EO) data. In turn, the urban land cover and change maps can support subsequent tasks that address the user requirements for the pilot cities specified in D2.2.

To that end, KTH assessed the appropriateness of several different datasets for DL training according to the user requirements in D2.2. Specifically, built-up areas and their change, urban land cover and their changes, the lack of green, as well as extreme weather events, were identified as relevant for this task. It should also be noted that not all of the investigated datasets cover the pilot cities. However, once trained, models can be transferred to other geographical regions — including the pilot cities — for deployment. Moreover, the pilot cities may not provide enough data for the training of DL models which notoriously require large amounts of data.

Once identified as relevant, the datasets were pre-processed to establish a correspondence with EO data from the Sentinel-1 Synthetic Aperture Radar (SAR) mission and the Sentinel-2 MultiSpectral Instrument (MSI) mission. Specifically, KTH developed a pre-processing workflow to obtain satellite images for both sensor modalities. The labels leveraged from existing datasets were then pre-processed to establish correspondence with the satellite imagery

The final deliverables are a report which describes the applied methods to produce the annotated data, and the annotated data themselves. The report is structured as follows. First, the objective of the task is introduced. Then, various EO sensors (Section 2.1 Earth Observation Data) and existing datasets (2.2 Existing Datasets) are presented. Section 3 describes the EO data pre-processing workflow; and, finally, all the annotation activities KTH conducted as part of Task 3.5 are summarized in Section 4.

# 1. Objective of Task 3.5 Data Annotation

Annotating satellite imagery using standard practices is a time-consuming manual process. It requires an annotator with expertise in understanding the image content. Moreover, it's often necessary to have knowledge in satellite imaging systems and acquisition methods. The objective of this task is to investigate various benchmark datasets that are freely available to obtain sufficient labelled data for training the AI models developed and adapted during the project. In addition, we will harmonize both optical and Synthetic Aperture Radar (SAR) satellite data to produce consistent and precise training sets.

## 2. Data Description

### 2.1 Earth Observation Data

EO data were collected from three different satellites, namely Sentinel-1 SAR, Sentinel-2 MultiSpectral Instrument (MSI), and Landsat 8. Google Earth Engine (GEE), a popular cloud-based platform for geospatial big data analysis (Gorelick, et al., 2017), was used to access and, later on, pre-process the data. The following three sections (Section 2.1.1 Sentinel-1 SAR – Section 2.1.3 Landsat 8 OLI) describe the EO data in more detail.

#### 2.1.1 Sentinel-1 SAR

The Sentinel-1 mission collects C-band SAR images at 20 m spatial resolution with dual polarization capability (HH+HV and VV+VH). Sentinel-1 images are available in GEE pre-processed to Ground Range Detected (GRD) images using the Sentinel-1 Toolbox. Pre-processing includes thermal noise removal, radiometric calibration and terrain correction. In addition, backscatter coefficient ( $\sigma$ ) were converted to decibels via log scaling ( $10 \log_{10} x$ ).

#### 2.1.2 Sentinel-2 MSI

The Sentinel-2 mission collects optical images at 13 spectral bands with various spatial resolutions. Band 2 (blue), Band 3 (green), Band 4 (red) and Band 8 (near-infrared) are provided at 10 m resolution, while Band 5 (red edge 1), Band 6 (red edge 2), Band 7 (red edge 3), Band 8a (red edge 4), Band 11 (short wave infrared 1), and Band 12 (short wave infrared 2) are provided at 20 m resolution. The remaining 3 bands provided at 60 m spatial resolution (Band 1, Band 9 and Band 10) were not considered for this task because they contain atmospheric information which is not considered relevant for land cover mapping and change detection. Sentinel-2 images are available in GEE as ortho-corrected images scaled by a factor of 10,000 (UTM projection) at two processing levels, Level-1A and Level-2A. Level-1A data representing Top-Of-Atmosphere (TOA) reflectance were chosen over Level-2C data representing surface reflectance due to the fact that not all early acquisitions of the Sentinel-2 mission are available at Level-2A in GEE.

#### 2.1.3 Landsat 8 OLI

The Landsat 8 satellite payload consists of two science instruments—the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). These two sensors provide seasonal coverage of the global landmass at a spatial resolution of 30 meters (visible, NIR, SWIR); 100 meters (thermal); and 15 meters (panchromatic). A detailed description of the data can be found at <https://landsat.gsfc.nasa.gov/satellites/landsat-8/>. Landsat 8 imagery are available in GEE as a Collection 2 Tier 1 calibrated TOA reflectance. Calibration coefficients are extracted from the image metadata. See (Chander, Markham, & Helder, 2009) for details on the TOA computation. Landsat scenes with the highest available data quality are placed into Tier 1 and are considered suitable for time-series processing analysis.

## 2.2 Existing Datasets

### 2.2.1 Microsoft's Building Footprints

Microsoft's computer generated building footprints are accessible on GitHub in separate repositories for the United States<sup>1</sup>, Canada<sup>2</sup> and Australia<sup>3</sup>. The approximately 145 million building footprints were automatically generated using the Microsoft Cognitive Toolkit (CNTK). Specifically, Microsoft's Bing team first applied Deep Neural Networks and the ResNet34 with RefinedNet up-sampling to segment buildings in VHR Bing satellite imagery. Satellite imagery for the segmentation in the United States was acquired between 2012 and 2015. Unfortunately, this information is not available in Canada and Australia. Segmentation outputs were then curated and a polygonization algorithm was applied to detect building edges and angles for the generation of building footprints. The quality of the produced building footprints was assessed using an evaluation set of approximately 15,000, 45,000 and 7,000 buildings for the United States, Canada and Australia, respectively. For all three datasets precision is above 0.980 and recall is above 0.650.

### 2.2.2 SpaceNet7

The SpaceNet7 Multi-Temporal Urban Development Challenge (Van Etten, et al., 2021) introduced a dataset with the track building footprints from satellite imagery. Specifically, the SpaceNet7 dataset contains temporal stacks (approximately 24 images) of very-high-resolution (approx. 4 m) monthly Planet composites, including corresponding manually annotated building footprints, ranging from 2018 to the beginning of 2020. The SpaceNet7 dataset covers 101 unique geographic sites over the span of two years, ranging from the beginning of 2018 to 2020. The sites are split into 60 training sites and 41 test sites, whereas building footprint labels are only available for the training areas. Figure 1 visualizes the locations of the SpaceNet7 sites.

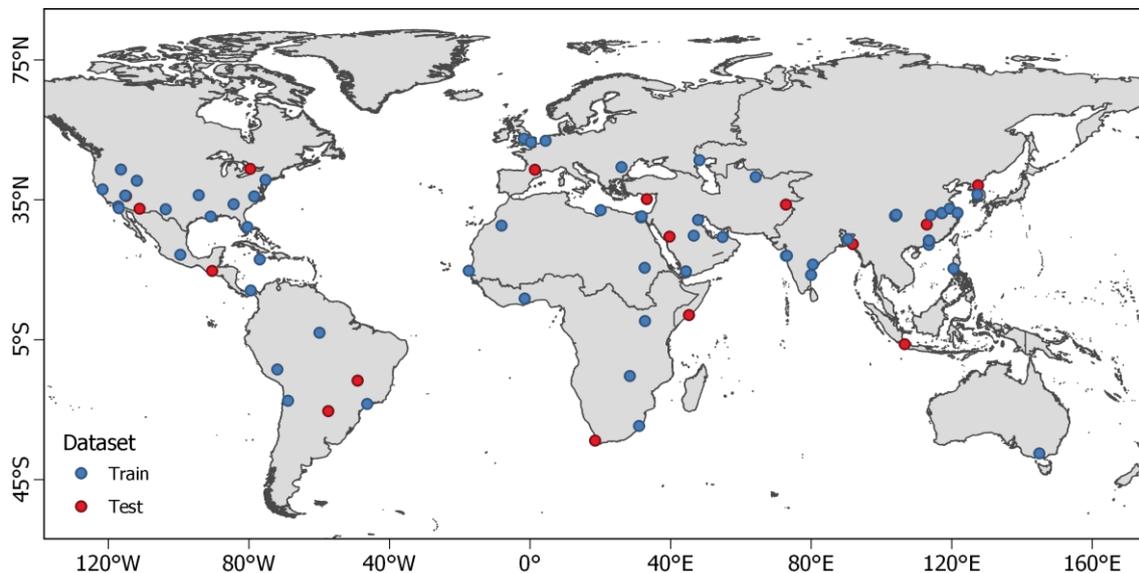


Figure 1: Train and test sites of the SpaceNet7 dataset (Van Etten, et al., 2021).

<sup>1</sup> <https://github.com/microsoft/USBuildingFootprints>

<sup>2</sup> <https://github.com/Microsoft/CanadianBuildingFootprints>

<sup>3</sup> <https://github.com/microsoft/AustraliaBuildingFootprints>

### 2.2.3 Urban Atlas

The European Urban Atlas<sup>4</sup> provides reliable, inter-comparable, high-resolution land use maps for over 300 Large Urban Zones and their surroundings (more than 100.000 inhabitants as defined by the Urban Audit) for the 2006 reference year in EU member states and for about 800 Functional Urban Area (FUA) and their surroundings (more than 50.000 inhabitants) for the 2012 and 2018 reference year in EEA39. Consequently, all four pilot cities are included in the Urban Atlas dataset. Two additional Urban Atlas layers were produced starting from the 2012 reference year: 1) Street Tree Layer within selected FUAs (depending on availability and suitability of satellite imagery) and 2) Building Heights for core urban areas of selected cities in EEA39. The first change layers were produced in 2012. The classification scheme used for the Urban Atlas is shown in Figure 2.

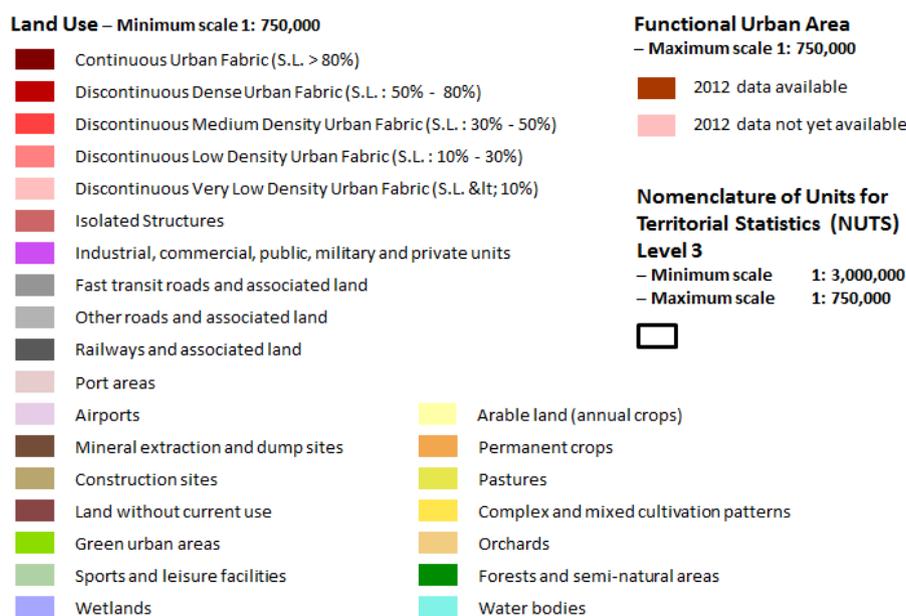


Figure 2: Classification scheme used for the Urban Atlas.

### 2.2.4 CORINE Land Cover

The CORINE Land Cover<sup>5</sup> (CLC) inventory was initiated in 1985 (reference year 1990). Updates have been produced in 2000, 2006, 2012, and 2018. It consists of an inventory of land cover in 44 classes. CLC uses a Minimum Mapping Unit (MMU) of 25 hectares (ha) for areal phenomena and a minimum width of 100 m for linear phenomena. The time series are complemented by change layers, which highlight changes in land cover with an MMU of 5 ha. Different MMUs mean that the change layer has higher resolution than the status layer. Due to differences in MMUs the difference between two status layers will not equal to the corresponding CLC-Changes layer.

<sup>4</sup> <https://land.copernicus.eu/local/urban-atlas>

<sup>5</sup> <https://land.copernicus.eu/pan-european/corine-land-cover>

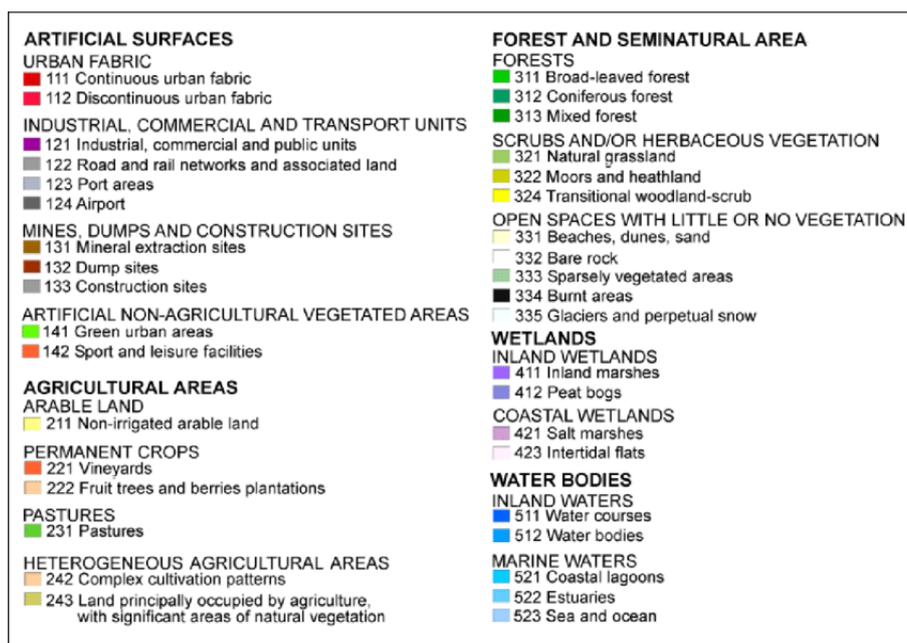


Figure 3: Classification scheme used for the CORINE Land Cover product.

### 2.2.5 Local Climate Zones

Local climate zones (LCZs) were originally developed for metadata communication of observational urban heat island studies (Stewart, Oke, & Krayenhoff, 2014). But soon they also showed potential in urban morphology mapping.

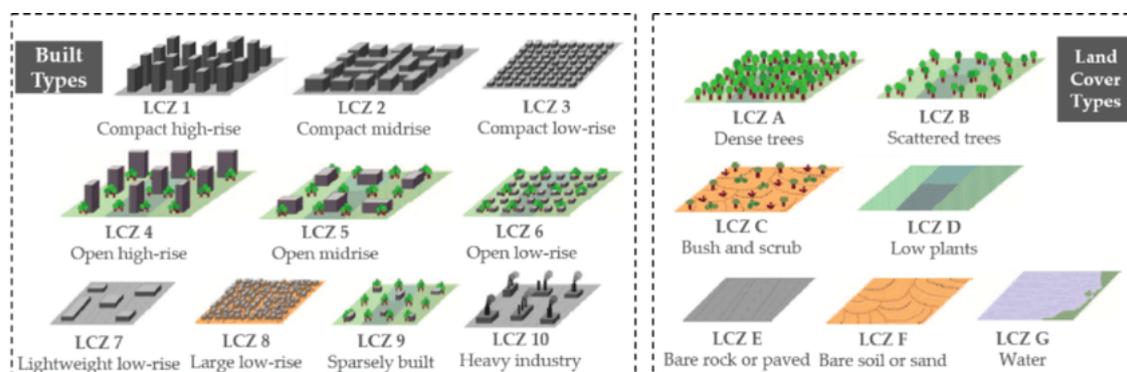


Figure 4: The local climate zone classifications scheme adopted from (Stewart, Oke, & Krayenhoff, 2014).

### 2.2.6 Cloud and Cloud Shadows Datasets

The SEN12MS-CR-TS dataset is a multi-modal and multi-temporal dataset for training and evaluating global and all-season cloud removal methods (Ebel, Xu, Schmitt, & Zhu, 2022). The dataset consists of 53 globally distributed Regions of Interest (ROIs) that are over 4000 × 4000 pixels each, covering about 40×40 km<sup>2</sup> of land. The total surface area covered by the data set is over 80000 km<sup>2</sup>. Of all collected ROI, 40 are defined as a training split and 13 as a hold-out split to evaluate cloud removal approaches on. For every ROI, 30 co-registered and paired Sentinel-1 SAR and Sentinel-2 MSI full-scene images evenly spaced in time throughout the year of 2018. Each acquired image was inspected and quality-controlled manually by the authors. The data is made available online<sup>6</sup>. It is about 2 TB in size.

<sup>6</sup> [https://patrickTUM.github.io/cloud\\_removal](https://patrickTUM.github.io/cloud_removal)

Another cloud and cloud shadow dataset considered in this task is the Sentinel-2 Cloud Cover Segmentation Dataset. It was generated as part of a crowdsourcing competition<sup>7</sup>, and later on, was validated using a team of expert annotators. The dataset consists of Sentinel-2 satellite imagery and corresponding cloudy labels stored as GeoTiffs. There are 22,728 chips in the training data, collected between 2018 and 2020. The dataset is available online<sup>8</sup>.

### 3. Earth Observation Data Processing

KTH started Task 3.5 by developing a methodology to pre-process satellite data in order to account for the fact that annotated data consists of two components — the labels and the corresponding data. To that end, a pre-processing workflow for Sentinel-1 SAR data, Sentinel-2 MSI data, and Landsat 8 data was developed in GEE (Gorelick, et al., 2017). Figure 5 illustrates an overview of the pre-processing workflow for Sentinel-1 SAR data, Sentinel-2 MSI data, and Landsat 8 data. The pre-processing steps for each sensor are described in the following paragraphs.

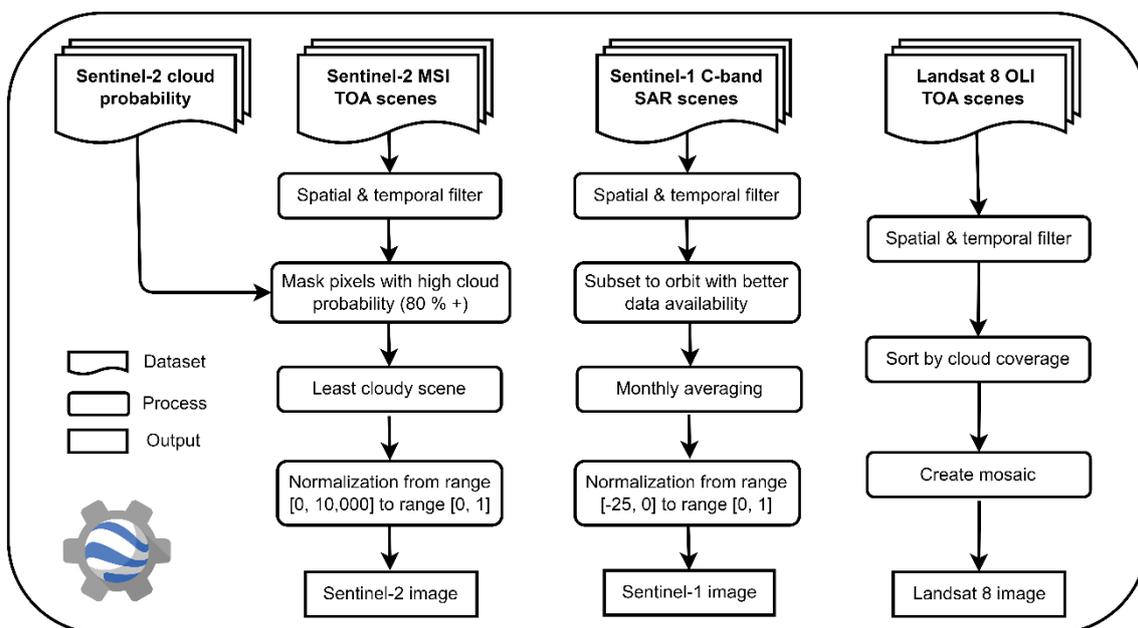


Figure 5: Overview of the data preparation workflow to retrieve satellite data from Sentinel-1 SAR, Sentinel-2 MSI, and Landsat 8 OLI.

#### 3.1 Sentinel-1 SAR

The proposed pre-processing chain for Sentinel-1 SAR data receives as a starting input Ground Range Detected (GRD) scenes from GEE. The collection of all available dual-band VV and VH GRD scenes acquired in Interferometric Wide Swath (IW) mode is then subset to a given time period and geographical region. Ascending and descending pass scenes were separated due to the strong influence of the incidence angle in the backscatter coefficient. Only scenes from the pass with better data availability in terms of absolute image counts were selected. Backscatter coefficients lower than -25 dB were then masked in each scene to remove noisy data. The remaining observations were used to compute the per-pixel temporal mean for both polarizations separately. Temporal mean is an effective method to remove speckle noise from SAR data

<sup>7</sup> <https://www.drivendata.org/competitions/83/cloud-cover/>{crowdsourcing competition

<sup>8</sup> <https://doi.org/10.34911/rdnt.hfq6m7>

without reducing the resolution (Chini, Pelich, Hostache, Matgen, & Lopez-Martinez, 2018). Finally, pixel values were normalized from the range of input values [-25, 0] to [0, 1].

### 3.2 Sentinel-2 MSI

Sentinel-2 MSI images are available in GEE as ortho-corrected images scaled by a factor of 10,000 (UTM projection) at two processing levels, Level-1A and Level-2A. Level-1A data represent TOA reflectance while Level-2C data represent surface reflectance. Due to the fact that not all early acquisitions of the Sentinel-2 mission are available at Level-2A in GEE, the Level-1A products were selected for the analysis. Deriving cloud-free imagery from a time series of S2 scenes is often based on temporal statistics. More specifically, the median of the pixel time series is computed after masking values that correspond with a high probability (80 % +) to clouds. Cloud probability is retrieved via the Sentinel Hub's cloud detector for S2 imagery<sup>9</sup>, which was recently added to GEE as a precomputed dataset. Finally, pixel values were normalized from the range [0, 10,000] to the range [0, 1]. In terms of spectral bands, we only processed the bands available at a spatial resolution of 10 m, i.e., blue (Band 2), green (Band 3), red (Band 4), and near-infrared (B8).

### 3.3 Landsat 8 OLI

The pre-processing chain for Landsat 8 OLI imagery starts from the GEE dataset "USGS Landsat 8 Collection 2 Tier 1 TOA Reflectance". In order to obtain a cloud-free Landsat 8 image from the collection, all scenes acquired over a specific geographic area within a predefined period were sorted by their cloud coverage. Thereafter, a mosaic was generated by using the image with the lowest cloud coverage first, and then filling possible gaps with subsequent images. The resulting image represents TOA reflectance scaled to the range [0, 1]. Like for Sentinel-2, we processed the blue (Band 2), green (Band 3), red (Band 4), and near-infrared (Band 5) bands.

## 4. Data Annotation Methods

### 4.1 Built-Up Area Labels

Built-up area labels were generated for 96 training and validation sites (Figure 6) and an additional 60 test sites covering unique geographies across the globe (Figure 8). The labels for the training and validation sites were obtained from Microsoft's open-access building footprints<sup>10</sup> by rasterizing the polygons and resampling them. On the other hand, the SpaceNet7 dataset (Van Etten, et al., 2021), featuring temporal stacks of monthly Planet composites including corresponding manually annotated building footprints, was exploited for building footprints for the test sites. In the next paragraph, the annotation method is described. However, for an in-depth description of the prepared built-up area labels, we refer readers to (Hafner, Ban, & Nascetti, Unsupervised domain adaptation for global urban extraction using sentinel-1 SAR and sentinel-2 MSI data, 2022).

First, Sentinel data were acquired for all 96 training and validation sites for the year 2016 using our pre-processing workflow described in Section 3. Earth Observation Data Processing. On the other hand, Sentinel data for the 60 test sites were acquired for a specific month in 2019.

---

<sup>9</sup> <https://github.com/sentinel-hub/sentinel2-cloud-detector>

<sup>10</sup> <https://blogs.bing.com/maps/2018-06/microsoft-releases-125-millionbuilding-footprints-in-the-us-as-open-data>

Thereafter, all EO data of the 30 sites located in the United States, Canada and Australia were annotated using Microsoft's building footprints. The label preparation started with transferring building footprints of the three datasets from the respective GitHub repository to GEE. We then converted all building footprint polygons laying within labelled training and validation sites to raster with a 10 m resolution. The resulting raster layer stores the percentage of overlap with buildings per-pixel. Finally, the layer was re-projected to UTM projection and resampled in order to establish a correspondence with the Sentinel data. Sentinel images for all sites and built-up area labels for the labelled training and validation sites were downloaded from GEE. Finally, the images were split into tiles of size 256 x 256 pixels. On the other hand, the EO data for the 66 sites located outside of the source domain, as defined in Figure 6, were not annotated, since in many other parts of the world, particularly in the Global South, this information is missing. Nonetheless, the unlabelled EO data can be leveraged for semi-supervised learning.

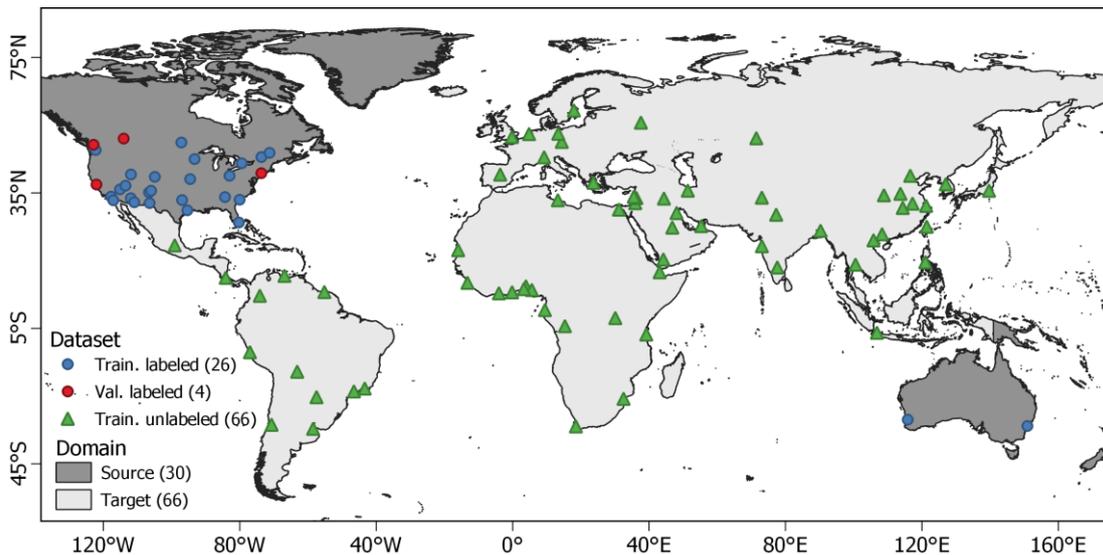


Figure 6: Training and validation sites. Labelled data is only used in the United States, Canada and Australia.

The size of the training and validation set is shown in Figure 7. The labelled training set contains about  $1.32 \cdot 10^9$  pixels covering more than 132,000 km<sup>2</sup>. Approximately 11 % of that area corresponds to BUA. The unlabelled training is larger than the labelled training set with about  $1.79 \cdot 10^9$  covering more than 179,000 km<sup>2</sup>. The validation set is considerably smaller with about  $2.22 \cdot 10^8$  pixels covering more than 22,000 km<sup>2</sup> (approximately 10 % BUA).

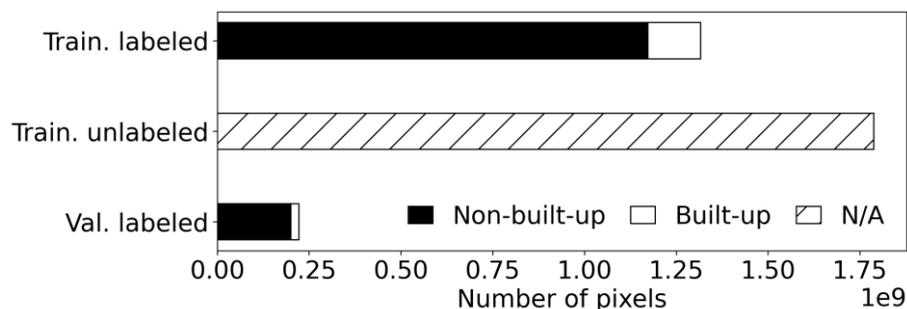


Figure 7: Number of pixels in the training and validation sets.

In contrast to the labelled training and validation sites, the 60 test sites were labelled using the building footprint labels from the SpaceNet7 Multi-Temporal Urban Development Challenge (Van Etten, et al., 2021). The label preparation workflow for the SpaceNet7 building footprints was identical to that for Microsoft's building footprints. Finally, the 60 sites were grouped into source and target domain according to their

geographical location. The target domain sites were further grouped into five cultural/geographical regions, namely Europe (EU), Latin America (LA), Sub-Saharan Africa (SSA), Islamic World (IW) and Asia.

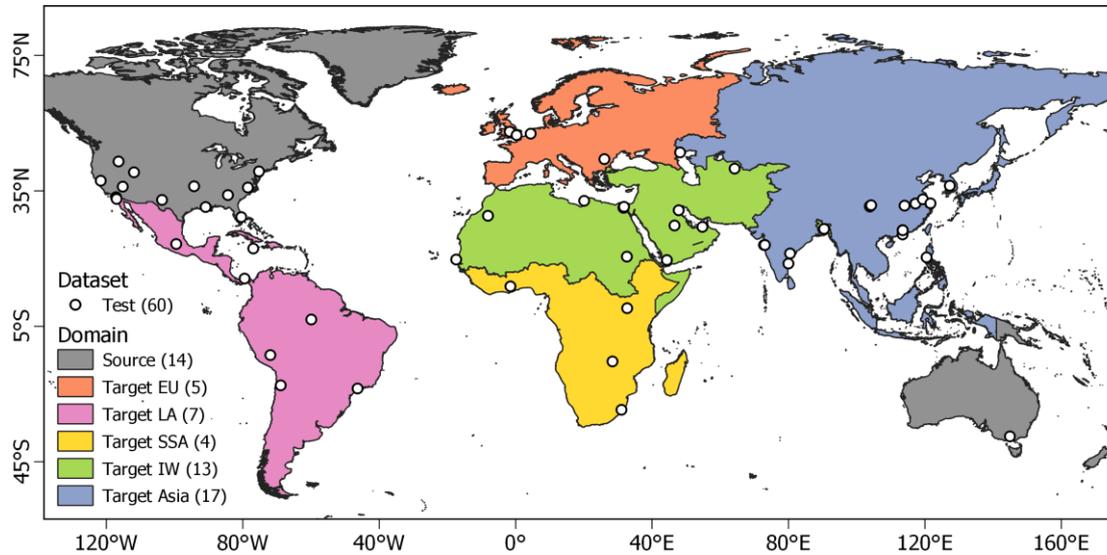


Figure 8: Sixty test sites grouped into the source domain and target domain. The target domain sites are further grouped into 5 regions, namely Europe (EU), Latin America (LA), Sub-Saharan Africa (SSA), Islamic World (IS), and Asia. Numbers in brackets denote the number of sites comprising a regional group.

The size of the test set is summarized in Figure 9. Across the six regions, it contains about  $1.11 \cdot 10^7$  pixels, where the fewest pixels are available for Europe ( $5.38 \cdot 10^5$ ) and the most pixels for Asia ( $3.34 \cdot 10^6$ ). BUA percentage ranges from 13 % (Source) to 21 % (LA). In total, the test set covers an area of  $1,112 \text{ km}^2$ .

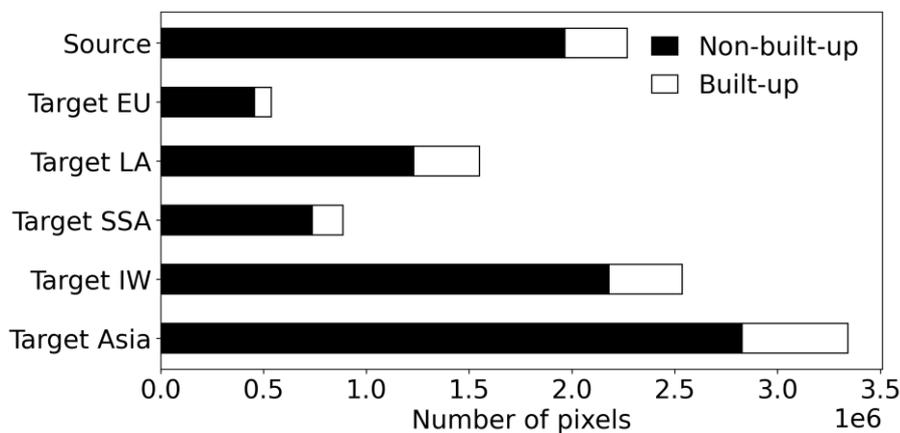


Figure 9: Number of pixels for the test set by region.

We named this dataset the SEN12 Global Urban Mapping Dataset and all its data is openly available on Zenodo<sup>11</sup>.

## 4.2 Built-Up Area Change Labels

Urban change labels were also generated for the 60 SpaceNet7 sites. Specifically, after pre-processing and downloading satellite data for the sites, building footprints for each timestamp at monthly time intervals were rasterized and resampled. However, due to the frequent occurrence of clouds in some regions, cloud-free data may not be available for each timestamp. Therefore, the Sentinel-2 MSI image time series for each

<sup>11</sup> <https://doi.org/10.5281/zenodo.6914898>

site were manually curated by removing images affected by clouds. Finally, a change date label was derived from the time series of building footprints. A sample of the generated annotated data is visualized in Figure 10. For an in-depth description of this annotated data, we refer readers to the publication (Hafner, Ban, & Nascetti, Multi-Modal Consistency Regularization Using Sentinel-1 SAR and Sentinel-2 MSI Data for Urban Change Detection, 2023).

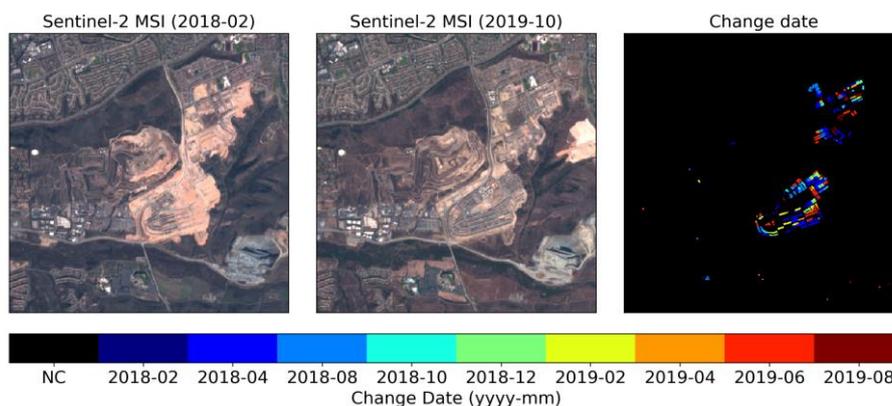


Figure 10: Built-up area change sample. NC (black) denotes not change.

### 4.3 Urban Land Cover Labels

In addition to these large-scale datasets, KTH also generated annotated data covering the pilot cities. For example, the Urban Atlas, which provides pan-European comparable land use and land cover data for large urban zones with more than 100.000 inhabitants, was used to generate land cover labels, including built-up areas and urban green infrastructure, for the pilot cities. In addition, the recently published LCZ maps (Zhu, et al., 2022) were used to generate land cover labels. The following two subsections (Section 4.3.1 Urban Atlas Labels and Section 4.3.2 Local Climate Zone Labels) describe the data annotation.

#### 4.3.1 Urban Atlas Labels

In order to obtain urban land cover annotations, KTH first prepared EO data, namely Sentinel-1 SAR and Sentinel-2 MSI images, for the year 2018. Then, the Urban Atlas 2018 was leveraged to annotate the prepared EO data. The processing workflow is visualized in Figure 11. Specifically, KTH rasterized the Urban Atlas 2018 polygons for all four pilot cities using the Geographic Information System (GIS) QGIS<sup>12</sup>, and resampled them to a spatial resolution of 10 m in order to establish correspondence with the Sentinel data. The resulting data is visualized in Figure 13. Furthermore, the original Urban Atlas classes were remapped. The annotated data for the four pilot cities is shown in Figure 13 and Figure 14.

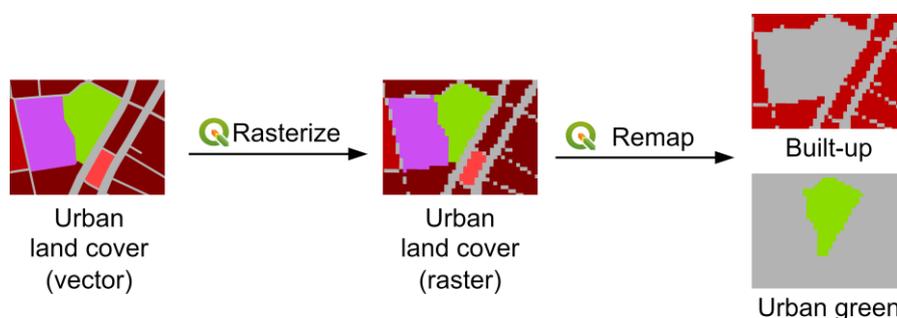


Figure 11: Overview of the pre-processing workflow used to rasterize and remap vector data using QGIS.

<sup>12</sup> <https://www.qgis.org>

### 4.3.2 Local Climate Zone Labels

The LCZ maps produced by (Zhu, et al., 2022) were downloaded<sup>13</sup> for the pilot cities Athens, Milan and Brussels. KTH resampled the LCZ maps to establish a correspondence with the satellite imagery. The resulting LCZ maps are visualized in Figure 12. Unfortunately, the LCZ data is not available for Sofia, since (Zhu, et al., 2022) only produced it for a selection of cities with more than 300,000 inhabitants.

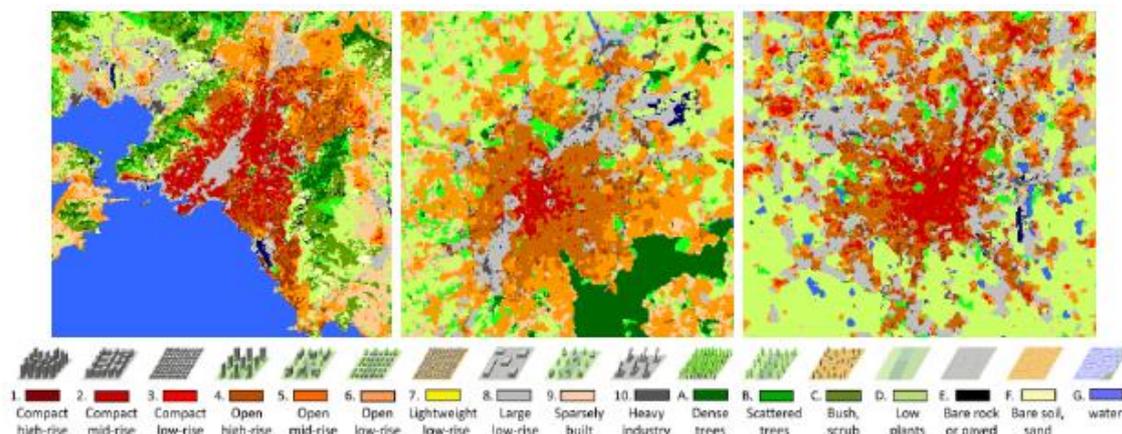


Figure 12: Local climate zone labels for Athens, Brussels and Milan (from left to right) for 2017. The legend was adopted from (Zhu, et al., 2022).

<sup>13</sup> <https://doi.org/10.14459/2021mp1633461>

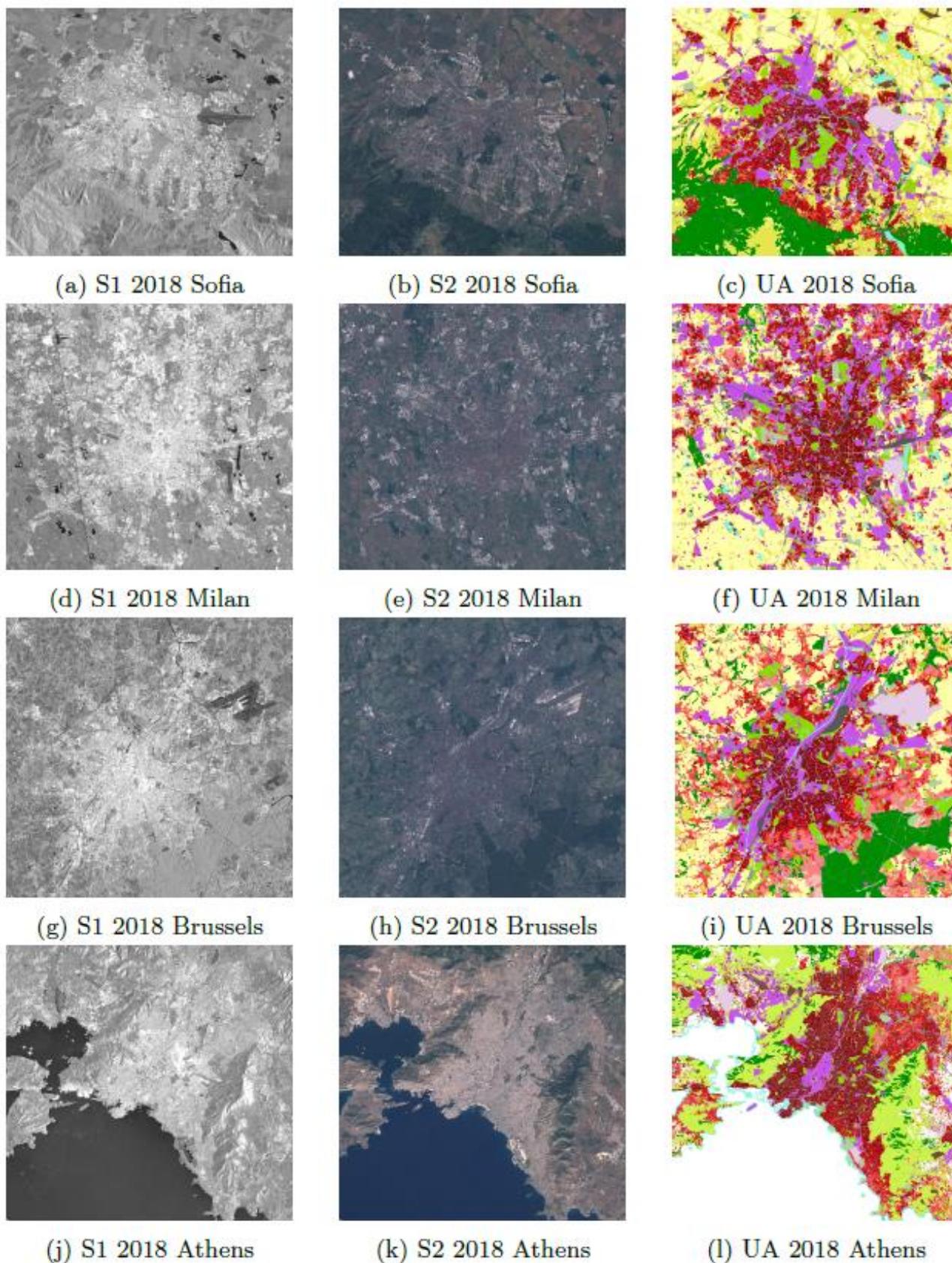


Figure 13: Sentinel-1 (S1) SAR data, Sentinel-2 (S2) MSI data and Urban Atlas (UA) data for 2018 for the four pilot cities. See legend in Figure 2.

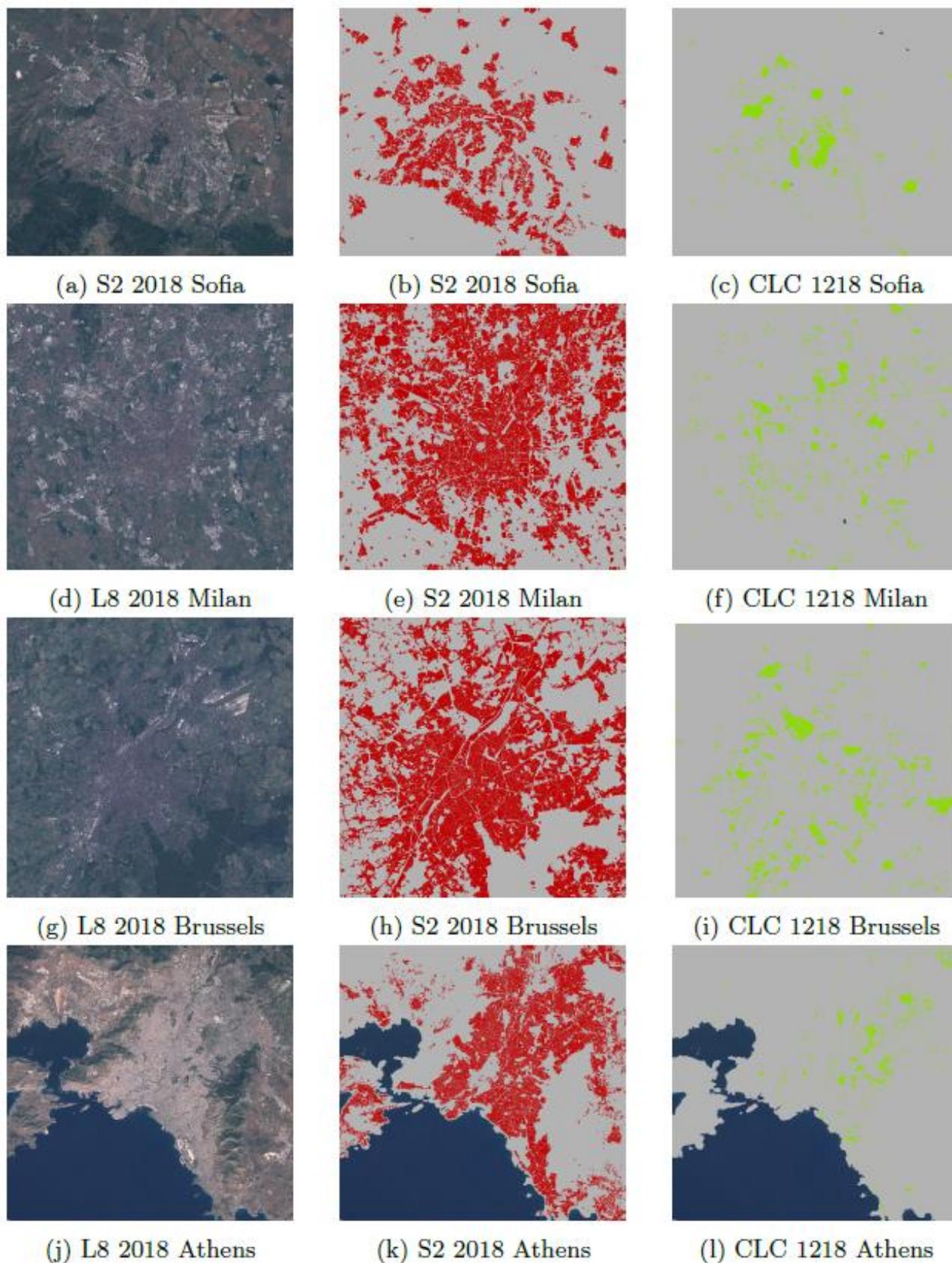


Figure 14: Sentinel-1 (S1) SAR data, Sentinel-2 (S2) MSI data and Urban Atlas (UA) data for 2018 for the four pilot cities. See legend in Figure 2.

## 4.4 Urban Land Cover Change Labels

### 4.4.1 Urban Atlas Change Labels

The Urban Atlas was also exploited for urban land cover change labels. Specifically, the Urban Atlas 2012 was pre-processed using the workflow outlined in Figure 11. Consequently, we obtained the same land cover labels for 2012 and 2018. By comparing the labels, we were able to obtain an urban land cover change product for the period 2012 -- 2018. Unfortunately, the Sentinel-1 and Sentinel-2 missions started collecting data several years after 2012. Therefore, we decided to use EO data collected by the Landsat mission for 2012. Specifically, the earliest acquisitions of the Landsat 8 mission, launched in 2013, were used. On the other hand, for 2018 the Sentinel-2 data was kept. To unify the spatial resolution of the Landsat 8 and Sentinel-2 imagery, we up-sampled the former imagery, Landsat 8, to a spatial resolution of 10 m. The resulting EO data and the corresponding urban land cover change labels are visualized in Figure 15. While Figure 15 only differentiates between change and no-change, it should be noted that the produced change labels also contain information specific land cover change types, including urban green change and built-up area change.

### 4.4.2 CORINE Land Cover Change Labels

Land cover change labels were also obtained from the CORINE land cover maps of 2012 and 2018. It should be mentioned that the Copernicus Land Service provides a dedicated change product for CORINE land cover which accounts for the fact that not all differences between the CORINE land cover maps of 2012 and 2018 correspond to real changes. On the other hand, the CORINE land cover change product<sup>14</sup> includes the real land cover changes mapped at the higher resolution of 5 ha minimal mapping unit, while the intersect includes the difference between two generalized, lower resulting datasets with 25 ha minimal mapping unit. The resampled land cover change labels and the corresponding EO data (see Section 4.4.1 Urban Atlas Change Labels) are visualized in Figure 16.

### 4.4.3 Cloud and Cloud Shadow Labels

Although the preparation of cloud and cloud shadow labels was initially not on the list of activities for task 3.5, upon a request from ICCS as part of task 3.6, we investigated several datasets containing relevant annotations, including a recently published multi-modal multi-temporal cloud removal dataset (SEN12MS-CR-TS) (Ebel, Xu, Schmitt, & Zhu, 2022) and the Sentinel-2 Cloud Cover Segmentation Dataset<sup>15</sup>. For a more detailed description of the datasets, we refer readers to Section 2.2.6 Cloud and Cloud Shadows Datasets.

---

<sup>14</sup> <https://land.copernicus.eu/pan-european/corine-land-cover/lcc-2012-2018>

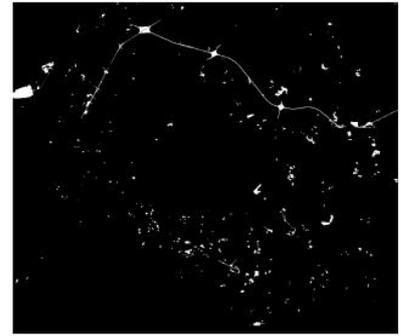
<sup>15</sup> <https://doi.org/10.34911/rdnt.hfq6m7>



(a) L8 2018 Sofia



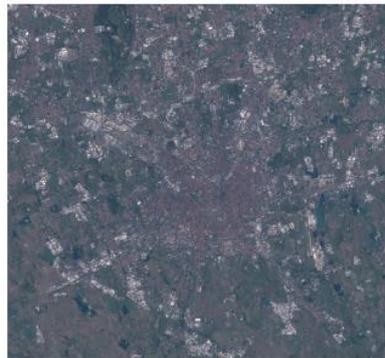
(b) S2 2018 Sofia



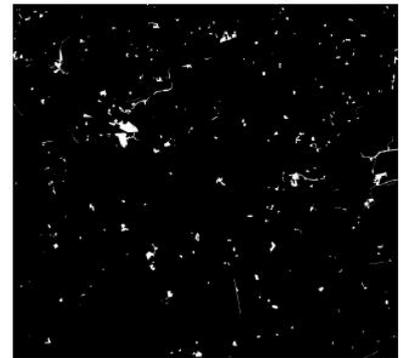
(c) CLC 1218 Sofia



(d) L8 2018 Milan



(e) S2 2018 Milan



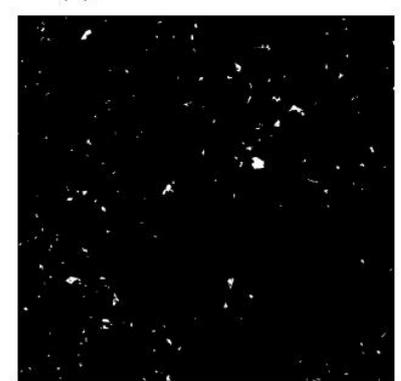
(f) CLC 1218 Milan



(g) L8 2018 Brussels



(h) S2 2018 Brussels



(i) CLC 1218 Brussels



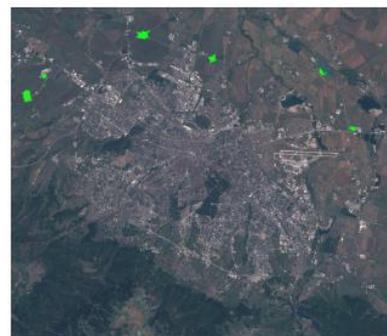
Figure 15: Landsat 8 (L8) OLI data, Sentinel-2 (S2) MSI data and Urban Atlas change labels between 2012 and 2018 for the four pilot cities.



(a) L8 2018 Sofia



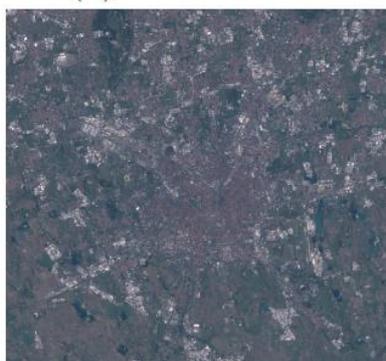
(b) S2 2018 Sofia



(c) CLC 1218 Sofia



(d) L8 2018 Milan



(e) S2 2018 Milan



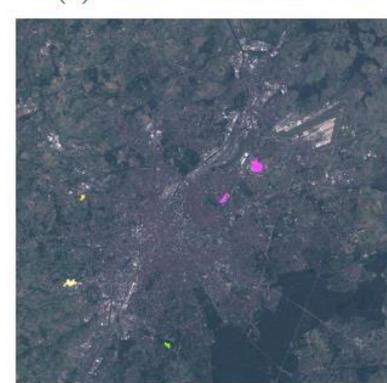
(f) CLC 1218 Milan



(g) L8 2018 Brussels



(h) S2 2018 Brussels



(i) CLC 1218 Brussels

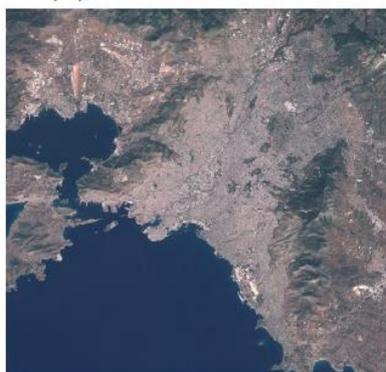


Figure 16: Landsat 8 (L8) OLI data, Sentinel-2 (S2) MSI data and CORINE land cover change labels between 2012 and 2018 for the four pilot cities.

## 4.5 Wildfire Labels

Wildfire labels were obtained from the Copernicus Emergency Management Service<sup>16</sup>. Figure 17 shows how the labels were queried for a wildfire event that occurred close to Athens in 2021.

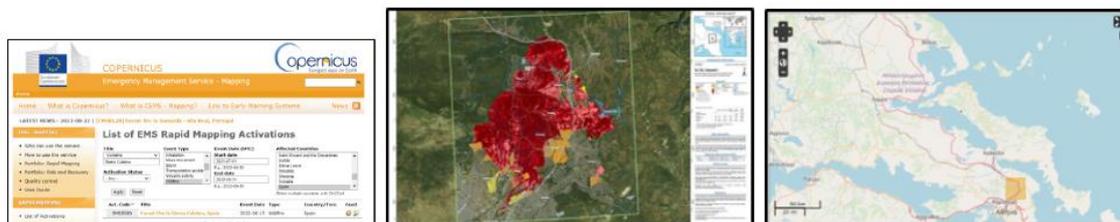


Figure 17: Overview of the pre-processing workflow for wildfire labels.

However, due to the unique characteristics of wildfires, the EO data pre-processing workflow deviated from the earlier introduced one (i.e., the one described in Figure 5). We refer readers to (Zhang, Hu, & Ban, 2022) for an in-depth description of the Sentinel-1 SAR and Sentinel-2 MSI pre-processing chains. Figure 18 exemplifies the obtained EO data and the corresponding wildfire labels for a fire close to the city of Athens in 2021.

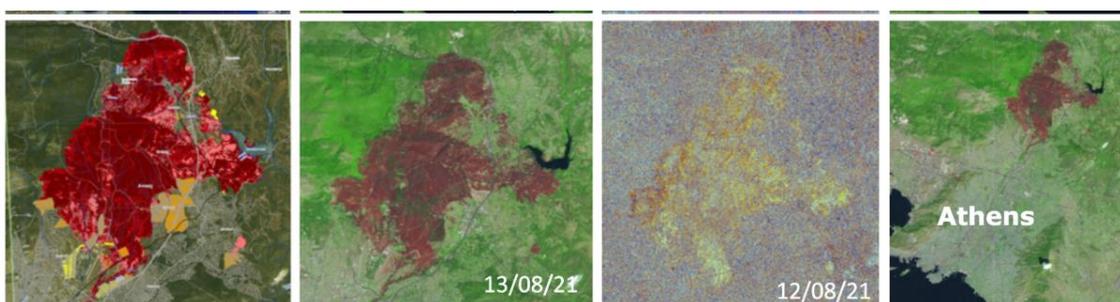


Figure 18: Example of the prepared EO data and the corresponding labels for a wildfire close to Athens.

<sup>16</sup> <https://emergency.copernicus.eu/>

## References

- Chander, G., Markham, B. L., & Helder, D. L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote sensing of environment*, 113(5), 893-903.
- Chini, M., Pelich, R., Hostache, R., Matgen, P., & Lopez-Martinez, C. (2018). Towards a 20 m global building map from Sentinel-1 SAR data. *Remote Sensing*, 10(11), 1833.
- Ebel, P., Xu, Y., Schmitt, M., & Zhu, X. X. (2022). SEN12MS-CR-TS: A Remote-Sensing Data Set for Multimodal Multitemporal Cloud Removal. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1-14.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, 202, 18-27.
- Hafner, S., Ban, Y., & Nascetti, A. (2022). Unsupervised domain adaptation for global urban extraction using sentinel-1 SAR and sentinel-2 MSI data. *Remote Sensing of Environment*, 280, 113192.
- Hafner, S., Ban, Y., & Nascetti, A. (2023). Multi-Modal Consistency Regularization Using Sentinel-1 SAR and Sentinel-2 MSI Data for Urban Change Detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, (under review).
- Stewart, I. D., Oke, T. R., & Krayenhoff, E. S. (2014). Evaluation of the 'local climate zone'scheme using temperature observations and model simulations. *International journal of climatology*, 34(4), 1062-1080.
- Van Etten, A., Hogan, D., Manso, J. M., Shermeyer, J., Weir, N., & Lewis, R. (2021). The multi-temporal urban development spacenet dataset. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (pp. 6398-6407).
- Zhang, P., Hu, X., & Ban, Y. (2022). Wildfire-S1S2-Canada: A Large-Scale Sentinel-1/2 Wildfire Burned Area Mapping Dataset Based on the 2017--2019 Wildfires in Canada. *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*, (pp. 7954-7957).
- Zhu, X. X., Qiu, C., Hu, J., Shi, Y., Wang, Y., Schmitt, M., & Taubenböck, H. (2022). The urban morphology on our planet--Global perspectives from space. *Remote Sensing of Environment*, 269, 112794.