# Geo-harmonizer PROJECT REPORTS

Last update: 30th of October, 2020

# Inventory and evaluation of existing map products

Prepared by: CTU in Prague, OpenGeoHub, Terrasigna

**www.opendatascience.eu/geoharmonizer-project**

**WORKING DOCUMENT PREPARED JOINTLY BY:**

CTU in Prague | OpenGeoHub | Terrasigna.com

**CONTACT:**


Martin Landa
martin.landa@fsv.cvut.cz

Tomislav Hengl
tom.hengl@opengeohub.org

# Document Information

| Document Reference | |
|---|---|
| Project: | Geoharmonizer:<br>EU-wide automated mapping system for harmonization of Open Data based on FOSS4G and Machine Learning |
| Title: | Inventory and evaluation of existing map products |
| Reference: | 2.b |
| Contract n°: | 2018-EU-IA-0095 |
| Version: | 1.0 |
| Distribution: | INEA, https://opendatascience.eu |

| Document Authors | | |
|---|---|---|
| Chapter 1 | Introduction | Lukas Brodsky, Lena Halounova, (CTU) |
| Chapter 2 | Land Cover Map Policies of Creation and Products | Codrina Maria Ilie (Terrasigna), Martijn Witjes (OGH) |
| Chapter 3 | Land Cover Products | Lukas Brodsky, Lena Halounova, Paulina Raeva (CTU) |
| Chapter 4 | Environmental Quality Maps | Saleem Ibrahim and Lena Halounova (CTU) - Air Quality Products<br>Carmelo Bonannella and Tomislav Hengl (OGH) - Water Quality Products and Soil Quality Products |
| Chapter 5 | Climate Change Indicators | Codrina Maria Ilie (Terrasigna) |
| Chapter 6 | Natural vegetation products | Carmelo Bonannella and Tomislav Hengl (OGH) |

# Project summary

Geo-harmonizer: EU-wide automated mapping system for harmonization of Open Data based on FOSS4G and Machine Learning

**Programme**: CEF Telecom (https://ec.europa.eu/inea/en/connecting-europe-facility/cef-telecom/2018-eu-ia-0095)

**Call year**: 2018

**Location of the Action**: Croatia, Czech Republic, Germany, Netherlands, Romania

**Implementation schedule**: September 2019 to June 2022

**Maximum EU contribution**: €1,423,864

**Total eligible costs**: €1,898,485

**Percentage of EU support**: 75%

**Coordinator**: Czech Technical University in Prague (Czech Republic) https://www.cvut.cz/en

**Summary**: The overall objective of the Action is to develop an original, web-based, scalable and modular system ("Geo-harmonizer") for hosting and accessing various thematic geospatial data layers (vector and raster GIS layers) to support cross-border services over the entire continental Europe. The beneficiaries will create a data portal and a software suite extending a wide variety of free and open-source software solutions for geospatial data (FOSS4G) in combination with state-of-the-art Machine Learning algorithms, and will be made available within EU-supported High Performance Computing (HPC)/Cloud computing infrastructures. The functionality of the system will be demonstrated vis-à-vis a list of new, added-value, pan-EU data sets including seamless continental Europe cover time-series (2000-2020), environmental quality indicators, potential natural vegetation maps and OpenStreetMap+ (improved continental Europe version of the OpenStreetMap). The data generated by the Action will be integrated into the European Data Portal. The output of the Action is to use the Geo-harmonizer to connect state-of-the-art remote sensing data sources (Sentinel-2, Landsat, and similar), machine learning framework, cloud, and High Performance Computing, and which is expected to result in a significant reduction of processing and delivery time.

**Project homepage**: https://opendatascience.eu/geoharmonizer-project

# Table of content

## List of Tables

# 1 Introduction

## 1.1 Purpose of the Report

This document is a report of the Task 2 of the Activity 2 called *Inventory and evaluation of data sources due to their usability within the projec*t in the Geo-harmonizer project implementation plan 2020–2022. The purpose of this document is to review the existing land cover and environmental map products.

The overall objective of this task is to search, review and assess existing land cover products, environmental quality maps, climate change indicators, and potential natural vegetation maps as related to the Geo-harmonizer project.

The above mentioned products have a wide range of differences among individual topics. Therefore the inventory was focused on analysis of the products and their attributes, and grouping of them according to their similarities.

## 1.2 Scope of the Report

This document covers 4 topics mentioned in 1.1 Chapter: land cover products, environmental quality maps, climate change indicators, and potential natural vegetation maps. The products and/or databases of individual thematic topics are summarized from various views to show the present situation of both the open, and not open data. The inventory was processed according to the following steps:

1. Initially, a search of all accessible data sets was performed. These products are thematically listed in tabulated form with identification of the product name, the link to the product, its spatial coverage and finally the maintainer is identified.
2. Next, a short introduction of the product is presented by the product abstract.
3. Finally, the main technical characteristics (for instance the spatial coverage, thematic resolution and its legend, spatial resolution or MMU, temporal coverage / updates, overall accuracy, product license, others) are tabulated and shortly assessed.

## 1.3 Legal Disclaimer

Information provided in this document should not be relied upon as legal advice or a substitute for legal advice. This information is being provided for general reference purposes only. The inventory information is compiled from the open internet sources. The authors have no means of the independent cross validation but solely rely upon the information from data providers and the associated publications.

## 1.4 Structure of the Document

Chapter 2 provides inventory of the Open Vector Map Coverage
Chapter 3 provides inventory and evaluation of the Land Cover products
Chapter 4 provides inventory and evaluation of the Environmental maps
Chapter 5 provides inventory and evaluation of the Climate Change indicators
Chapter 6 provides inventory and evaluation of the Natural vegetation products

## 1.5 Acronyms

AATSR Advanced Along-Track Scanning Radiometer
AOD Aerosol Optical Depth
AQI Air Quality Index
ASCAT Advanced SCATterometer
ASCII American Standard Code for Information Interchange
ATSR Along Track Scanning Radiometer
AVHRR Advanced Very High Resolution Radiometer
BA Burned Area
BC Black Carbon
CAMS Copernicus Atmosphere Monitoring System
CCI Climate Change Initiative
CDR Climate Data Records
CGLOPS-1 LST
Climate Change Indicators
CNRS Centre national de la recherche scientifique
CO  Carbon Monoxide
Copernicus EMS Copernicus Emergency Management Service
CTM  Chemical Transport Model
DEHM  Danish Eulerian Hemispheric Model
DFO Dartmouth Flood Observatory
DMSP Defense Meteorological Satellite Program
DUE Data User Element
EC European Commission
ECMWF European Centre for Medium-Range Weather Forecasts
ECV Essential Climate Variables
ECV Essential Climate Variables
EEA European Environment Agency
ENVISAT Environmental Satellite
ERC European Research Council
ERS European Remote Sensing
ESA European Space Agency
ESDC  European Soil Data Centre
ESDC European Soil Data Centre
ESDIS Earth Science Data and Information System
ETC/ICM European Topic Centre on Inland, Coastal and Marine waters
EU European Union
EUMETSAT European Organisation for the Exploitation of Meteorological Satellites
EURAD-IM  EURopean Air Pollution Dispersion Inverse Model
EUROS  EURopean Operational Smog
FD Flood Directive
FRP Fire Radiative Power
GCOS Global Climate Observing System
GEIA Global Emission InitiAtive
GEO Group on Earth Observations
GeoTIFF Geo Tag Image File Format

GH Geo-harmonizer
GIS Geographic Information System
GMS Japanese Geostationary Meteorological Satellite
GOES Geostationary Operational Environmental Satellite
GPCC Global Precipitation Climatology Centre
GridSat Gridded Satellite
GTS Global Telecommunication System
HDF4 Hierarchical Data Format 4
IR Infrared
IRWIN visible data, infrared window data
IRWVP infrared water vapor data
ISCCP International Satellite Cloud Climatology Project
ISD Integrated Surface Database
JAMI Japanese Advanced Meteorological Imager
JMA Japanese Meteorological Agency
JPEG Joint Photographic Experts Group
JRC Joint Research Centre
JSON JavaScript Object Notation
K Kelvin
KML Keyhole Markup Language
LC Land Cover
LCCS Land Cover Classification System
LOTOS Long Term Ozone Simulation
LSA-SAF EUMETSAT Satellite Application Facility on Land Surface Analysis
LST Land Surface temperature
LTDR Land Long Term Data Record
LUCAS Land Use and Coverage Area frame Survey
MAIAC Multi-angle Implementation of Atmospheric Correction
MATCH Multi-Scale Atmospheric Transport and Chemistry model
MetOp Meteorological Operational Satellite Program of Europe
MGP Experimental Merged Products
MMU Measurement and Monitoring Unit
MMU Minimal Mapping Unit
MODIS Moderate Resolution Imaging Spectroradiometer
MS Member States
MSG Meteosat Second Generation
MTSAT Multifunctional Transport Satellites
MW Microwave
NASA National Aeronautics and Space Administration
NCEI National Centers for Environmental Information
NCEO National Centre for Earth Observation
NDVI Normalised Difference Vegetation Index
NESDIS National Environmental Satellite, Data, and Information Service
NetCDF Network Common Data Form
NH3 amoniak
NO2 Nitrogen dioxide
NOAA National Oceanic and Atmospheric Administration
NPP National Polar-orbiting Operational Environmental Satellite System Preparatory Project

NRT Near Real Time
O3 Ozone
OSM OpenStreetMap
PDF Portable Document Format
PDU Power Distribution Unit
PFRA Preliminary Flood Risk Assessment
PM Particulate matter
POES Polar Operational Environmental Satellite
SAR Synthetic Aperture Radar
SDS Science Dataset
SEVIRI Spinning Enhanced Visible and Infrared Imager
SIA Secondary Inorganic Aerosol
SILAM System for Integrated modeLling of Atmospheric coMposition
SLSTR Sea and Land Surface Temperature Radiometer
SMHI Swedish Meteorological and Hydrological Institute
SO2 Sulfur dioxide
SSMIS Special Sensor Microwave Imager / Sounder
SST Sea Surface Temperature
SUOMI National Polar-orbiting Partnership
TCI Thermal Condition Index
TOA Top-of-Atmosphere
UA Urban Atlas
VIIRS Visible Infrared Imaging Radiometer Suite
WFD Water Framework Directive
WMS Web Map Service

# 2 Example of the Open Vector Map Coverage

## 2.1 OpenStreetMap as an Example of the Global Vector Map Coverage

### 2.1.1 OpenStreetMap

OpenStreetMap is an editable vector database of physical features on the ground, built and maintained by volunteers and distributed under the Open Data Commons Open Database License. The OpenStreetMap Foundation supports the hosting needs of the project through fund-raising activities. Regarding administrative borders, the contributors follow the “on the ground rule”, which states that in case of disputes, the information provided by the local editors has priority. Administrative regions in OSM represent “subdivisions of areas/territories/jurisdictions recognised by governments for administrative purposes.” and they can range from groups of nations to districts and suburbs.

The OSM database describes physical features with elements, which come in three types:

1. Nodes, points in space (e.g. traffic signs, phone booths, trees)
2. Ways, linear features and area boundaries (e.g. roads, buildings, and administrative boundaries);
3. Relations, characterizations of connections between elements (e.g. between the units between affected by an administrative administrative boundary).

The identity, functions, and other details of these features is described with tags. OSM's free tagging system allows the map to include an unlimited number of attributes describing each feature, however, in some cases in a layman's way.

A new version of the complete dataset is released every week on Planet OSM. Several mirrors are offered by universities and other institutions to share the networking and hosting load. Planet OSM and its mirrors provide users the ability to specify and download Extracts, which contain OSM data for individual continents, countries, and metropolitan areas. The most notable providers are:

1. www.overpass-api.de
2. www.download.openstreetmap.fr/
3. www.download.geofabrik.de
4. www.osm-internal.download.geofabrik.de
5. www.download.bbbike.org/osm/
6. www.protomaps.com/extracts

OSM data is used by several major companies, governments, universities, and NGO’s, such as United Nations, Cambridge University, Amazon, Apple, Air France, and the Bill and Melinda Gates Foundation. Furthermore, several companies and non-profit organisations use it to create derivative products, such as the following examples:

1. Mapillary is a service for sharing crowdsourced geotagged photos. It was developed by a Swedish startup, and sold to Facebook in June 2020. The service closely collaborates with OSM. For instance, user-provided images are analyzed with computer vision to extract

geotagged features such as traffic signs, which can easily be uploaded into the OSM database.

2. OSM Landuse Landcover is a derivative of OSM data specifically used for landuse and land cover analysis. Its nomenclature is similar to level 2 of the Corine land cover classification (29 classes). Land use was extracted from OSM tags, or predicted if unavailable (Schultz et al. 2017).

3. OpenTopoMap is a topographic map, automatically generated from OSM and SRTM data. It is maintained by a team of volunteers, and supported by the University of Erlangen-Nürnberg. It is freely available under the CC-BY-SA license.

Historical land cover can be generated by applying the method proposed by Schultz et al. (2017) on historical OSM data. Although this data is available yearly starting from 2013, its nomenclature is inconsistent. Creating harmonized land cover classifications from this historical data through the method of Schultz et al. (2017) will likely require specific approaches per year. The end result would, however, have a higher spatial (especially in rural areas) and temporal resolution than the Urban Atlas.

## 2.2 Examples of Methodologies of Land Use/Land Cover European Coverage data

### 2.2.1 INSPIRE Methodology

Infrastructure for Spatial Information in the European Community - INSPIRE - represents the EU initiative to allow a seamless access to environmental geospatial harmonized and interoperable data across all Member States. The initiative is supported by Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) and it stands on the following principles:

- Data should be collected only once and kept where it can be maintained most effectively.
- It should be possible to combine seamless spatial information from different sources across Europe and share it with many users and applications.
- It should be possible for information collected at one level/scale to be shared with all levels/scales; detailed for thorough investigations, general for strategic purposes.
- Geographic information needed for good governance at all levels should be readily and transparently available.
- Easy to find what geographic information is available, how it can be used to meet a particular need, and under which conditions it can be acquired and used.

To meet these principles, INSPIRE defines legislative norms, as well as technical implementation rules that address matters of data and metadata interoperability, network services as well as data and service sharing.
INSPIRE is based on the infrastructures for geospatial information established and operated by the Member States, according to the above mentioned implementation rules.

With regard to data specifications, the Directive addresses 34 spatial data themes needed for environmental applications, including administrative boundaries and transport. The INSPIRE datasets are in GML format, following the INSPIRE xml schema.

### 2.2.2 EuroGeographics Methodology

EuroGeographics is an international not-for-profit organization that aims to coagulate under its umbrella Europe's National Mapping, Cadastral and Land Registration Authorities. They claim to represent more than 90% of all official relevant institutions, enabling them to create and recommend guidelines and best practices, especially for the technical aspects regarding cross-national harmonization of boundaries. In that regard, EuroGeographics established a dedicated forum - the State Boundaries Forum.

Relevant for the harmonization of open data service is the work carried out through the European Location Services project and continued with OpenELS resulting in technical documents, guidelines and recommendations and a toolbox to assist national relevant authorities to participate in the pan-European infrastructures.

Through OpenELS, EuroGeographics has also published a pan-European small scale (1:1 000 000) topographic data covering **45 countries and territories**, containing five themes: Administrative Boundaries down to NUTS3 level and their administrative hierarchy; Water Network; Transport Networks; Settlements; and Named Locations - the **EuroGlobalMap**.

EuroGlobalMap is produced with the following vector data structure and content:

- The **administrative boundaries** theme refers to a boundary of an entity controlled by an administrative authority, this entity can be composed of several areas: international boundary and national administrative levels up to the NUTS3 level if it exists.**The water network** holds information on the watercourses, open water areas, seaside, and the associated hydrologic equipment (dams).
- The **transport network** holds information on roads and railways, ferry lines, and airports. Connections between ferry stations and other transport modes have been added.
- The **settlement** theme holds information on the built-up areas and populated places, mainly their name and number of inhabitants.
- The **vegetation and soils** theme holds features like wood, orchard or sandy area.
- The **named locations** are text objects which do not relate to vector data but are needed for cartographic purposes.
- The **miscellaneous** theme (MISC) holds various topographic elements like power lines, landmarks, and buildings

The primary data sources used for EuroGlobalMap used to be the national data collections of the mapping agencies, possibly at similar spatial resolution. Secondary data sources, either internal or external to the mapping agencies, could also be used to fill the required information. Since EGM v5.1 (2012), a new production process based on generalisation from EuroBoundaryMap and EuroRegionalMap has been progressively put into place. In the latest version, the Boundary, Hydrography and Transportation themes have been produced with this method. The full technical specifications of the dataset are available here.

With respect to the scope of Geo-harmonizer as detailed in the Implementation Plan, there are 2 main thematic topics of interest regarding the OpenStreetMap and the national thematic layers available: administrative boundaries and transport infrastructure. Both themes propose significant difficulties, be it legal as in the case of administrative boundaries or topological, as in the case of transport infrastructure. Datasets available at European level are scarce and highly variate, especially with respect to transport infrastructure.

Administrative boundaries represent imaginary lines outlined by national relevant authorities, such as national survey agencies, following historic land ownership heritage and that are compliant with relevant national laws and regulations. Each European country has its own administrative system, with different hierarchical lineage, different definitions and functions for the delineation of the administrative units.

*Table 2.1 Presents the main European aggregators of boundaries*

| ID | Type | Name | Countries not covered | Maintainer/Provider |
|---|---|---|---|---|
| 1 | Administrative boundary | EuroGlobalMap | Bosnia and Herzegovina, Montenegro | EuroGeographics |
| 2 | Administrative boundaries | INSPIRE | Albania, Belgium, Bosnia and Herzegovina Bulgaria, Croatia, Cyprus, Denmark, France Greece, Iceland, Italy, Kosovo, Latvia, Liechtenstein, Malta, Montenegro, North Macedonia, Slovakia, Sweden, Switzerland, United Kingdom | Each EU country is responsible for its INSPIRE datasets |
| 3 | Statistical boundaries | Nomenclature of territorial units for statistics - NUTS | Bosnia and Herzegovina, Kosovo | Geographic Information System of the Commission, The European Statistical Office |

*Table 2.2 General characteristics for the products of the main European aggregators of boundaries*

| Product / Characteristics | EuroGlobalMap | INSPIRE | Nomenclature of territorial units for statistics - NUTS |
|---|---|---|---|
| Complete over Geo-harmonizer region | No | No | No |
| Scale | 1:1 million | | 1:1 million |
| Geometric generalisations | The minimum accepted area size is 0.06 km2.<br>The matching tolerance of the geometry is 30 m.<br>The minimum length of an edge between two connected points should be 200 m. If connected points distances are less than 200 m, they have to be combined into one. | N/A | N/A |
| Current version | 2020 | Depending on the country | 2021 |
| Update frequency | annually | Depending on the country | every 3 years |
| Data Format | Personal Geodatabase and Shapefiles | GML format, following the INSPIRE xml schema | SHP, TopoJSON, geoJSON, Personal Geodatabase ,SVG |
| Product license | Open | Depending on the country | Open |

| | | | |
|---|---|---|---|
| Product access | download after registration | Inspire portal | direct download |
| Citation | N/A | N/A | N/A |

Transport infrastructure is the general term that encloses all components that are relevant for any kind of transport, irrespective of transport medium: land, water, air, such as roads, railroads, airports, inland waterways, marine shipping routes, bus stations, subway lines etc. The thematic is extensively complex and an exhaustive analysis of availability at European level is beyond the scope of this project. For analysis harmonization, we will consider the three following transport networks: road, rail and water.

*Table 2.3 List of existing open datasets with respect to transport infrastructure at the European level*

| ID | Type | Name | Countries not covered | Maintainer/Provider |
|---|---|---|---|---|
| 1 | roads, railways, ferry lines, airports | EuroGlobalMap | Bosnia and Herzegovina, Montenegro | EuroGeographics |
| 2 | airports, ports | Transport Networks | Europe | Geographic Information System of the Commission, The European Statistical Office |
| 3 | roads | Global Roads Open Access Data Set (gROADS) | Global | Center for International Earth Science Information Network (CIESIN) |
| 4 | road, rail, water, air, cableways | INSPIRE | EU | each EU country is responsible for its INSPIRE datasets |

*Table 2.4 General characteristics of existing open datasets with respect to transport infrastructure at the European level*

| Product / Characteristics | EuroGlobalMap | Transport Networks | Global Roads Open Access Data Set (gROADS) | INSPIRE |
|---|---|---|---|---|
| Complete over Geo-harmonizer region | No | Yes | yes | No |
| Scale | 1:1 million | 1:1 million | 1:250,000 | N/A |
| Geometric generalisations | The minimum accepted area size 0.06 sq.km. The matching tolerance of the geometry 30 m.<br>The minimum length of an edge between two connected points should be 200 m. If connected points distances are less than 200 m, they have to be combined into one. | | ~50m positional accuracy | N/A |
| Geometric | point/polygon | point/polygon | polyline | area objects, centreline objects, point objects |
| Current version | 2020 | 2013 | 1980s to 2010 | Depending on the country |

| | | | | |
|---|---|---|---|---|
| Update frequency | annually | N/A | N/A | Depending on the country |
| Data Format | ArcGIS Personal Geodatabase and Shapefiles | Personal Geodatabase | ArcGIS Personal Geodatabase and Shapefiles | GML format, following the INSPIRE xml schema |
| Product license | Open | Open | Open | Depending on the country |
| Product access | download after registration | direct download | download after registration | Inspire portal |
| Citation | N/A | N/A | Center for International Earth Science | N/A |

### References on Land map products

Center for International Earth Science Information Network-CIESIN - Columbia University, and Information Technology Outreach Services-ITOS - University of Georgia. 2013. "Global Roads Open Access Data Set, Version 1 (gROADSv1)." Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4VD6WCT.

Schultz, M., Voss, J., Auer, M., Carter, S., & Zipf, A. 2017. Open land cover from OpenStreetMap and remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 63, 206–213. https://doi.org/10.1016/j.jag.2017.07.014.

# 3 Land Cover Products

Land Cover (LC) is seen as the physical and biological cover of the Earth surface including artificial surfaces, agricultural areas, forests, (semi-)natural areas, wetlands and water bodies (INSPIRE directive). Di Gregorio and Jansen (2000) defined land cover as 'the observed (bio)physical cover on the Earth's surface', which is widely perceived as an important component of environmental and ecological systems, central to understanding global environmental change. Land cover (LC) is essential for a variety of social, scientific, etc., studies and environmental planning according to Jansen *et* al. (2008). The LC products are based on satellite imagery and reference field data. In order to better understand anthropogenic and environmental consequences in our land, a precise characterization of the land cover and land use is essential.

The problem of land cover harmonisation relates to spatial data integration and therefore needs to consider the data concepts adopted and the spatial, temporal, semantic and quality aspects of the data (Jansen *et* al. 2008). Differences in semantic concepts are often considered as the key obstacle for data integration and interoperability.

The variety of land cover classification systems limits the compatibility and comparability of land cover data, which makes analysis of multi-source land cover data a very difficult task. Harmonization of existing land cover classification systems is essential to improve their cross-comparison and validation. Harmonization of classifications enables the combination of existing heterogeneous land cover datasets to support global land cover data analysis (Hui Yang *et* al. 2017).

One important aspect of achieving harmonization in land cover classification systems, based on Hui Yang *et* al. (2017), is the notion of semantic interpretation. 'Semantic interpretation is defined as the process of matching the meaning of land cover classes between different classification systems. It helps to achieve harmonization by analyzing similarities and differences between definitions of land cover classes. The semantic interpretation between any two classification systems has been proved challenging, largely because of differences in definitions and related difficulties in creating one-to-one mapping relationships. One step towards achieving semantic interoperability of existing land cover databases and categorization is to develop methods for measuring the degree of semantic similarity between categories in different land use/land cover classification systems. Semantic interpretation aims to derive a solution for making different existing land cover classification systems interconnected and interoperable. Semantic interpretation is the process of determining the meaning of a land cover class, which provides the mechanism for combining, matching and translating the meaning of the class definitions from various land cover classification systems (Hui Yang *et* al. 2017).

Another important aspect of harmonization is the legend (nomenclature) translation. Nomenclature is essential for meaningful comparisons between different land cover classification systems. The translation of nomenclatures in a general schema is an essential step towards harmonization, which discovers the equivalent relations between two or more land cover classification systems (Hui Yang *et* al. 2017).

Hand in hand with harmonization comes the standardization. 'Despite the urgent need to establish and adopt a standard classification system, none of the current classifications has been

internationally accepted. Ideally, harmonization should be guided by existing or evolving standards and thus has to use a common structure for reference (Hui Yang *et* al. 2017).

The ISO defines two standards related to the land cover, a generic standard to define a set of rules creating a classification system (ISO 19144-1:2009) and a UML metamodel that allows any land cover classes to be expressed based on a rigorous logical framework (ISO 19144-2:2012). The aim of ISO 19144-2:2012 is to enable users the ability to compare and integrate information from existing different classification systems in a common reference language without replacing them.

This chapter provides inventory and evaluation of the existing land cover products (20) at global, continental (European) and national level (European countries). Table 3.1 provides the first-level reference to the evaluated land cover products with indication of the spatial coverage, link to the product web page and information on the product maintainer.

The variety of the land cover products are further assessed based on a set of parameters to be compared. These are namely: the spatial coverage, thematic resolution (number of classes), nomenclature model, spatial resolution or minimum mapping unit, temporal coverage (the updates), overall thematic accuracy, harmonization problem if identified, land cover product license, product access, product URL and citation.

To compare the global, continental and national land cover products, we assessed the main characteristics together with accuracy metrics as listed in the respective tables below (Chapters 3.1, 3.2., 3.3.) . The comparison focuses on these topics:

- overall accuracy of the products;
- comparison of the thematic resolution (number of classes) with overall accuracy of the products;
- comparison of the spatial resolution with the overall accuracy metric;
- and assessment of the temporal coverage (updates) of the products.

These four main assessment directions shall bring insights into the state-of-the-art of the global, European continental and national land cover mapping, while revealing the potential gaps in the land cover productions.

*Table 3.1 List of existing global, continental, national and regional Land Cover products*

| ID | LC name | Coverage | Maintainer |
|---|---|---|---|
| **GLOBAL LAND COVER PRODUCTS** | | | |
| 1. | CCI Land Cover (2017) | global | ESA |
| 2. | GlobeLand30 (2000) | global | National Geomatics Center of China (NGCC) |
| 3. | GLC - SHARE | global | Land and Water Division of FAO |
| 4. | GLC2000 (2000) | global | Joint Research Center (JRC) |
| 5. | NASA Land Cover (2011) | global | NASA |
| 6. | Land Cover (GLCNMO) | global | ISCGM, GSI of Japan |

| | | | |
|---|---|---|---|
| 7. | Copernicus Global Land Service | global | VITO |
| **EUROPEAN LAND COVER PRODUCTS** | | | |
| 8. | CORINE Land Cover (1990 -2018) | continental | EEA (EU Copernicus) |
| 9. | European Settlement Map (2016.17) | continental | EEA (EU Copernicus) |
| 10. | Urban Atlas | continental Large Urban Zone (LUZ) | EEA (EU Copernicus) |
| 11. | Land Cover Map of Europe 2017 (S2GLC, ESA) | continental | S2GLC ESA project |
| 12. | Pan-European land cover (30m) map of 2015 (based on Landsat and LUCAS data) | continental | Pflugmacher, Rabe, Peters, Hostert |
| 13. | High Resolution Layers | continental | EEA (EU Copernicus) |
| **NATIONAL LAND COVER PRODUCTS** | | | |
| 14. | LISA | national (AT) | AT consortium |
| 15. | LandCover DE | national (DE) | DLR |
| 16. | SIOSE | national (ES) | National Reference Center on Land Cover and on Land Use and Spatial Planning |
| 17. | LGN (Dutch land use datasets from 1995 to 2018) | national (NL) | Wageningen Environmental Research |
| 18. | OSO Land Cover | national (FR) | Theia Data and Services centre |
| 19. | LC Greece | national (GR) | Dimitrios Gounaridis, Anastasios Apostolou & Sotirios Koukoulas<br>SAGISRS, University of the Aegean |
| 20. | CLC50 | national (HU) | Institute of Geodesy, Cartography and Remote Sensing (FÖMI) |

## 3.1 Global Land Products

The following  seven global land cover products are presented in this chapter: GLC2000, GlobeLand30, NASA Land Cover, Land Cover (GLCNMO), ESA CCI Land Cover, GLC - SHARE, and CGLS. The inventory of the products, and their characteristics, are provided here in this chapter as available from the publicly available pieces of information.

### GLC2000

The general idea behind **GLC2000** is to provide a harmonized land cover database for the year of 2000 for the whole globe. It is produced by the Joint Research Center (JRC) of the EP. The year of 2000 is considered as a reference year for environmental assessment in relation with activities such

as the UN's Ecosystem-related International Conventions. The GLC2000 is based on the VEGA 2000 dataset, which consists of 14 months of pre-processed daily global data acquired by the VÉGÉTATION instrument on board of SPOT 4 satellite. The GLC2000 uses the LCCS system developed by FAO (JRC 2020).

### GlobeLand30

**GlobeLand30** was developed by the National Geomatics Center of China (NGCC) and its 2000 and 2010 versions became open-source in 2014. Later the Ministry of Natural Resources launched an updated GlobeLand30 in 2017. The latest version is GlobeLand30 2020. The product has 30 m resolution and consists of 10 classes. For the **GlobeLand30** creation multispectral imagery was used from TM5, ETM+ and Landsat. Pixel and object-based classification was applied and knowledge-based interactive verification of classification results was used to reach an 83.50% overall accuracy. The validation includes checking 150,000 points in 80 tiles of 853 tiles (Chen, J. *et* al. 2015). The product can be either downloaded from here or browsed here. This product is the first global product with the 30 m resolution. It was officially donated to the United Nations (GLOBELAND30 2015).

### Land Cover Classifications of NASA

NASA Earth Observations (NEO) together with the USGS Landsat programme, provide the longest continuous global archive of satellite Earth Observation (EO) data. Given that fact, Landsat imagery is a preferable image background for land cover and land changer assessment. The processing algorithm has 17 classes as per the International Geosphere-Biosphere Programme's global vegetation database. Input data is used from AVHRR as most of the LC data was based on NDVI. The programme ended in 2011. The temporal resolution of the LC products is annual from 2001 till 2011. The products can be downloaded as raster images directly from the NASA website here. The data is available in four spatial resolutions: 0.1º, 0.25º, 0.5º, 1.0º (NASA 2020).

### Global Land Cover by National Mapping Organizations

The **Global Land Cover by National Mapping Organizations** (GLCNMO) was created by the National Mapping Organization in collaboration with the Geospatial Information Authority (GIA) of Japan, Chiba University and national GIA's. The input data is based on MODIS Terra & Aqua. There are three LC products released in 2003, 2008 and 2013. The 2003 product has 1 km resolution as the last two - 500m. The products can be downloaded from github links located on the official webpage of GLCNMO (GLCNMO 2020).

### Global Land Cover - SHARE

FAO **Global Land Cover - SHARE** (GLC-SHARE) is a combination of the 'best available' (up to 2014) high resolution national and regional land cover databases produced by the Land and Water Division of FAO. MERIS and MODIS products were used with 250-300m resolution. Harmonization was based on LCCS. According to its creators, the approach adopted by FAO makes it possible to at least partially translate and harmonize various datasets and implement data fusion. The data fusion preserves the spatial consistency and semantic information.The harmonization criteria is to minimize the differences between various datasets. The Product has 11 thematic maps and a resolution approximately 1 km per tile. The goal behind this project is to produce a harmonized global land cover product which reduces the risk of information and provides quantifiable information on the land cover. A beta version of the LC product for 2012 was published in 2014 by FAO (FAO 2020).

### ESA Land Cover - CCI

ESA CCI Land Cover scope is to realize the full potential of EO data archives of ESA, together with the EU member states. The main objective is to thoroughly review all required algorithms for global land product generation. It is a part of the ESA Climate Change Initiative (CCI). The first global product with a 300m resolution was launched in 2017, with a data time-range from 1992-2015. Imagery was used from MERIS FR & RR, SPOT VÉGÉTATION. Methods based on FAO/UNEP LCCS were applied. The project is capitalized on GlobCover, GLC2000, GlobAlbedo and GloCorine. Since the launch of the first product, this project has been a part of the Copernicus Climate Change Service. A Land Cover Classification System (LCCS) was used to comply with the GLC2000, GlobCover 2005 and 2009 products. After the successful CCI LC project, its developers are currently creating LC maps from 2016 to 2020 for the Copernicus Climate Change Service (C3S). The latter were produced using Proba-V imagery with an overall accuracy of 71.1%. The creators state that accuracy depends on the thematic class and it may vary. This is due to spatial coverage of the MERIS FR in some areas, for example, south Argentina, Chile, etc. The LC products consist of annual land cover maps from 1992-2015, NDVI seasonality products, MERIS surface reflectance and open water bodies product. The consortium, which created the CCI LC products, consists of 11 members from different European universities and research institutions divided into three working groups: EO sciences team, Climate Research Group, and System Engineering. The final LC products could be viewed here (ESA 2017).

### Copernicus Global Land Operations

The **Copernicus Global Land Operations** (CGLOPS-1) was recently developed by the **VITO** team. In 2015, a global product was launched and an updated product is being developed annually for 2015-2019 time range. The product has a 100m resolution and uses the FAO LCCS nomenclature. All LC products use PROBA-V imagery and Random Forest classification algorithms. The products can be accessed through here. It is also possible to download the data in 20x20 degree tiles. The authors claim that the overall accuracy of the products is 80.2% as it might vary in different parts of the world (VITO 2019).

## 3.1.1 Technical Characteristics of the Global Land Cover Products

We have selected eleven of the main characteristics of the land cover products to compare the variety of the existing global databases (Table 3.2). The most important characteristics are the resolutions (spatial, thematic and temporal) in comparison to the product accuracy, while the others are descriptive pieces of information related to the potential of the use of the product in a process of land cover map harmonizations.

*Table 3.2 Global Land Cover characteristics*

| Product / Characteristics | ESA CCI LC | GlobeLand30 | GLC - SHARE | GLC2000 | NASA LC | Land Cover (GLCNMO) | CGLOPS-1 (VITO) |
|---|---|---|---|---|---|---|---|
| Spatial coverage | Global | Global | Global | Global | Global | Global | Global |
| Thematic resolution (No. of classes) | 22 | 10 | 11 | 23 | 17 | 20 | 23 |

| Nomenclature model | LCCS | GlobeLand30 | LCCS, FAO, SEEA | LCCS | IGBP | LCCS | LCCS |
|---|---|---|---|---|---|---|---|
| Spatial resolution / MMU | 300m | 30 m | 1 km (source) | 1 km (source) (source) | 1 deg 0.5 deg 0.25 deg 0.1 deg (source) | 500 m | 100 m |
| Temporal coverage / updates | 1992÷2015 | 2000, 2010, 2017, 2020 | 1998÷ 2012 | 2000 | 2001 ÷ 2011 | 2003, 2008,2013 | 2015 ÷ 2019 |
| Product release | 2017 | 2014, 2017, 2020 | 2014 | 2003 | - | 2003, 2008, 2013 | 2016-2020 |
| Overall accuracy | 71,7% (source) | 84% (2010) 86% (2020) (source) | 80% (source) | 70% (source) | 77% (source) | 76% (source) | 80% (source) |
| Harmonization problem | Not reported | Not reported | Not reported | Not reported | Not reported | Not reported | Not reported |
| Product license | © ESA Climate Change Initiative - Land Cover project 2017 | National Geomatics Center of China (NGCC) | Land and Water Division of FAO | JRC © European Communities | ©NASA | ISCGM, GSO of Japan, Chiba University, NGIAs | ©VITO Remote Sensing; Copernicus Service Information |
| Product access | Registration | Registration | Registration | JRC: open for scientific and non-commercial distribution | Free download | Free download | Open, web-viewer |
| Product URL | donwload | download | download | download | download | download | download |
| Citation | (ESA 2017) | (NGCC 2020) | (FAO 2020) | (JRC 2000) | (NASA 2011) | (GLCNMO 2020) | (VITO 2019) |

## 3.2 European Continental Land Products

In the following subchapter, six European land cover products are evaluated. These are obviously mainly the Copernicus program land products. These are namely CORINE Land Cover (Copernicus), Urban Atlas (Copernicus), European Settlement Map (Copernicus), High Resolution Layers (Copernicus), and than Pan-European Land Cover map (Humboldt University of Berlin) and S2GLC Map of Europe (ESA project). The Land Cover products are first introduced by a short abstract and next the inventory of the all products, their characteristics, are again tabulated together for their comparison.

### CORINE Land Cover - Copernicus

The most detailed complete land cover products are offered by **CORINE**. The CORINE programme was established by the European Commission (EC) to provide geographically harmonized information concerning the environment on the continent. CORINE Land Cover was launched in 1985 and uses the year of 1990 as reference one. Updates were released in 2000, 2012 and 2018. It consists of 44 classes of land cover with a minimal mapping unit of 25 ha. **CCL1990** used input data from Landsat-5 MSS/TM with a time range between 1986-1998 and spatial resolution of 50m.. The geometric accuracy of the LC product is claimed to be 100 m and thematic accuracy 85% which might not be achieved. The production time of **CCL1990** was 10 years and 27 countries participated in the production.

The **CCL2000** product used input data from Landsat-7 ETM with a spatial resolution of 25 m. The thematic accuracy of 85% is claimed to be achieved. The production time of **CCL2000** is 4 year and more than 30 countries participated.

The **CCL2006** product used input satellite data from SPOT - 4/5 and ORS P6 LISS with a spatial resolution of 25 m. The geometric and thematic accuracy of this product is stated to be the same as the previous two. This is the first CORINE product to be with free access for all users. Thirty-eight countries participated in the production.

The **CCL2012** product is based on IRS P6 LISS II and RapidEye satellite data with a spatial resolution of 25 m. The time range of the input data is within 2011-2012. Thematic and geometric accuracy are stated the same as in the previous products, namely, 85% and 100 m.

The **CLC2018** product is the first CORINE LC product to be based on Sentinel-2 imagery data; Landsat-8 was used as an additional satellite data provider for gap filling. The time consistency of the products is between 2017-2018 with a spatial resolution of the input data - 10 m. The product has the same geometric and thematic accuracy as the previous ones - 85%, 100 m. Thirty-nine countries participated in creating CLC2012 and CLC2018. Land cover products and Land change products could be either accessed or downloaded (Copernicus - EEA 2020).

### European Settlement Map - Copernicus

The **Copernicus European Settlement Map** (ESM) is a two-layer spatial raster dataset of the human settlements in Europe. The ESM products have been financed by the Directorate-General for Regional and Urban Policy (DG REGIO). It is based on input satellite data from SPOT5 and SPOT with a spatial resolution of 2.5m. The ESM was produced by the Joint Research Center of the European Commission with GHSL technology. This technology consists of a classification method based on machine learning. The first ESM was released in 2014 which is the successor of the built-up layer with an unprecedented detailed resolution of a European wall-to-wall coverage.In 2016 a new update was released - ESM2016. This product uses population data from GEOSAT 2011. ESM2016 has 12 classes unlike the ESM 2014 which has only 1 - built-up. The latest ESM product has a higher resolution than the previous ones - 2.5 m. This product is upgraded to complement the above mentioned 12 layers with greenery, watercourse, and railway layers divided into two groups: built-up areas and out of the built-up area. These products could be viewed here. The European Settlement Map 2017 (ESM2p5m) is the most recent product that was produced as part of the URBA project (The European Settlement Map 2017 Release 2017), (Copernicus 2020).

### Urban Atlas - Copernicus

The Urban Atlas is a product of the Copernicus Land Monitoring Service. It has several updates: 2006, 2012, 2018. Two land cover change products are also available: Urban Atlas Change 2006-2012, Urban Atlas Change 2012-2018. The Urban Atlas 2012 and 2018 consist of 17 urban classes with 0.25 ha resolution and 10 rural classes with a 1ha resolution. All products can be viewed or downloaded from this link. The latest technical report on creating the Urban Atlas is available here (Copernicus 2020).

### High Resolution Layers (HRL) - Copernicus

The **Copernicus High Resolution Layers** (HRL) are pan-European LC products which consist of 5 different LC layers: imperviousness, forests, grassland, water & wetness and small woody bodies. They complement the LC products from CORINE and are produced based on satellite imagery with a combination of automatic processing and interactive rule based classification. The year of 2015 is stated to be a reference year for the HRS products and its production is increasingly based on time series of satellite imagery and image fusion between optical and radar data. The input data is mainly from Sentinel;-1 and Sentinel-2 and since 2015 some high resolution imagery have been used.The spatial resolution of the HRL is 10m as of 2018. There are three production updates for the HRL: 2012, 2015 and 2018.

The 2012 production consists of the layers Imperviousness and Forests. The Imperviousness layer is based on 2006, 2009 and 2012 data with product spatial resolution of 20 m. The data for the forest layers is based on satellite data from 2012 with a 20m resolution.

In 2015, the Imperviousness layer overgone a whole preprocessing for the reference years 2006, 2009 and 2012 and a new 2015 product was added with a 20m resolution. The Forest layer was also updated with a 2015 data with 20m resolution as well. The Grassland land is newly added in 2015 with reference year 2015 and data with a 20m resolution. The Wetness & Water layer for 2015 production is a combination of a 7-year time series (2009-2015) but 2015 is considered as a reference year. The data used has 20m resolution. Small woody features were also added in the 2015 production which includes baseline data from VHR. The product is a vector map available in 5m and 100m resolution.

In the 2018 production, the HRL improved their product resolution to 10m. The Imperviousness layer added a 2018 update in addition to Impervious Built-up (IBU) and its corresponding 100m aggregate Share of Built-Up (SBU). The Forest layer also added a new 2018 product. The Grassland layer not only added a 2018 update but also an additional one for 2015-2018 grassland change product with a 20m resolution. The Wetness & Water layer added again a 7-year timeline based product (2012-2018). (Copernicus 2020).

### Pan-European Land Cover - HUB

The **pan-European land cover** map was developed by a team from the Humboldt University of Berlin (HUB).They claim that their project is the first attempt to use LUCAS and Landsat-8 for mapping pan-European land cover and land use. This product is a combination between LUCAS and Landsat-8 data. The Landsat data was acquired between 2014 and 2016. The best accuracy of 75% from the classification was achieved with data from three years 2014-2016. The LC product can be downloaded as a TIFF. There are 12 land cover and land use classes applied. The spatial resolution is 30 x 30 m. This product could be downloaded here. (Pflugmacher, D.,2019), (Pangea 2020).

### Land-cover Map of Europe of the Global Land Cover - (ESA project)

The **Land-cover Map of Europe** is a European LC product created as a part of the Global Land Cover - Sentinel 2 (**S2GLC)** project (S2GLC 2017). This product uses input multi-temporal satellite data from Sentinel-2 collected in 2017. The produced LC map has a 10m resolution and uses 13 thematic classes.

The project **S2GLC** (Sentinel-2 Global Land Cover) is a project funded by the European Space Agency (ESA) and produced by a consortium which consists of the Space Research Centre of Polish Academy of Sciences, IABG mbH, EOXPLORE, UG and University of Jena. The development of the LC product is divided into two parts: the first one deals with testing the best classification algorithm and the second with the classification itself applied on continental Europe.

The main goal of the project was to develop a classification methodology using only Sentinel-2 imagery for global mapping with a high-degree of automation and to extend the existing LC legend of Europe. Few nomenclature harmonization tests were carried out using the products CLC, HRL, Urban Atlas and GUF (Global Urban Footprint - *commercial product by DLR*). Other tests concerned data availability and number of cloudless images over certain European parts. Random Forest (RF) classification was used for the Sentinel-2 images along with training points from existing LC products (Gromny *et* al. 2019). The classification itself was computed on the cloud computation platform CREODIAS. The final product could be either viewed on the CREODIAS EO Browser (S2GLC must be checked) or download here. Full description of the project milestones can be found in the official technical report here.

## 3.2.1 Technical Characteristics of European Continental Land Cover Products

The same main set of the characteristics of the land cover products are compared in Table 3.3 to describe the variety of the existing continental European land cover databases. These are tabulated from the available public information.

*Table 3.3 European continental Land Cover characteristics*

| Product\<br><br>Characteristics | CORINE Land Cover (Copernicus) | Global Land Cover - Sentinel- 2 S2GLC (ESA) | Urban Atlas (Copernicus) | European Settlement Map (Copernicus) | Pan-European LC (Humboldt University of Berlin) | High Resolution Layers (HRL) (Copernicus) |
|---|---|---|---|---|---|---|
| **Spatial coverage** | Europe (EEA39) | Europe | Europe (EEA39) 319 FUAs | Europe | Europe | Pan-European |
| **Thematic resolution (No. of classes)** | 44 | 13 | 27 (17 urban classes 10 rural classes) | 1 | 12 | 5 |
| **Nomenclature model** | CORINE (LC/LU) | CORINE | CORINE compatible | - | CORINE compatible | CORINE compatible |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Spatial resolution (pixel size)** | | 10 m | | 2 m/10 m/100 m | 30 m | 10 m (20 m before 2018) |
| **MMU** | 25 ha | | 0.25 ha / 1 ha | | | |
| **Input data time coverage** | 1986÷1998; 2000±1year; 2006±1year 2011÷2012; 2017÷2018 | 2017 | 2006, 2012, 2018 | 2012, 2015, 2019 | 2015 ±1 yeap | 2006, 2009, 2012, 2015, 2018 |
| **Product release** | 1990, 2000, 2006, 2012, 2018 | 2019 | 2006, 2012, 2018 | 2014, 2016, 2017 | 2015 | 2012, 2015, 2018 |
| **Overall accuracy** | > 85% (source) | 86% (source) | 93% (source) | 96% (source) | 75% (source) | > 90 % (source) (> 85% before 2018) (source) |
| **Harmonization problem** | Not reported | Not reported | Not reported | Not reported | Not reported | Not reported |
| **Product license** | Copernicus: Free and open | Copernicus: Free and Panopen | Copernicus: Free and open | Copernicus: Free and open | Creative Commons Attribution-ShareAlike 4.0 International | Copernicus: Free and open |
| **Product access** | browse and download | browse download | browse and download | browse and download | download | browse and download |

## 3.3 European National Land Cover Products

In the following subchapter, there is a review of a subset of accessible European national land cover products based on publicly available internet information. Clearly, only several countries of the European Union have, or share, their information on national land cover activities. Fig. 4.9. illustrates the European countries with the national LC product. The most visible, and probably the most developed, are the Austrian LISA, German Land Cover DE, and Spanish SIOSE land cover products.

*Fig. 3.1 Map of Europe indicating the countries with national LC products*

### LISA - Austria

LISA (Land Information System Austria) provides Austria with land cover and land use information. The goal behind the project is to establish a public user participation of an interdisciplinary consensus, meeting the demand of the Land Information System and provide an operational system for land monitoring in Austria. LISA reduces the existing lack of information of different special fields such as spatial planning, forestry, agriculture, water and natural hazard management, as well as environmental protection and conservation. LISA is the key contributor of Austria to the LMCS (Land Monitoring Core Service) and therefore an important and essential element for this European Land Monitoring System. LISA was formed in four stages (LISA-1, LISA-2, CadasterENV Austria and CadasterENV). LISA-1 and LISA-2 have been funded by the Federal Ministry of Transport, Innovation and Technology through the Austrian Space Applications Programme of the Austrian Research Promotion Agency. These products are based on orthophotos and laser scanning data with limited use of satellite data and cover surface of 3 000 km². CadasterENV Austria and CadasterENV were founded by European Space Agency (ESA). CadasterENV Austria integrated data from Pléiades satellites and WorldView-2 satellites to provide VHR land cover of 10 000 km². Also a Land Cover Change (LCC) alert map for the whole Austria (84 000 km²) was created. In CadasterENV (fourth stage of LISA) Sentinel-2 imagery was used for enrichment of 10 000 km² of VHR Land Cover. Also preparations for the HR Land Cover for the entire Austria have been created (LISA 2020).

### Land Cover DE - Germany

The project Land Cover DE is a German contribution to the European programmes and systems for regular or on-demand surveys of the earth's surface. It was created using Sentinel-2 imagery from the years 2015 to 2017 and LUCAS 2015 in-situ reference data

(https://ec.europa.eu/eurostat/web/lucas). It contains seven land cover types: (1) artificial land, (2) open soil, (3) high seasonal vegetation, (4) high perennial vegetation, (5) low seasonal vegetation, (6) low perennial vegetation and (7) water. It provides Land Cover products for years 2015-2017 using the CORINE nomenclature. The spatial resolution of the product is 10 m and can be viewed on the DLR Earth Observation Center here. The overall accuracy is reported in (Weigand *et* al. 2020) to be 93.1%. More information could be found here.

### SIOSE - Spain

SIOSE is an information system on Land Cover of Spain, integrated within the National Plan of Observation of the Territory (PNOT) whose objective is to generate a database of Land Cover for the whole Spain at 1: 25 000 reference scale integrating information available from the Autonomous Communities and the General State Administration. It is produced in a decentralized and coordinated manner between different administrations following the INSPIRE principles; It is periodically updated. The organization of SIOSE has its antecedent in the development of the CORINE Land Cover project in Spain. The minimum mapping units for SIOSE are following: 0,5 ha for water, crops, wetlands, beaches, riparian vegetation and sea cliffs, 1 ha for urban areas and 2 ha for agriculture, forest and natural areas. SIOSE uses in its production besides orthophotos also satellite images from SPOT5, Landsat 5 TM and satellite images provided by the CCAA. SIOSE data model is a standardized, interoperable and harmonized model of land cover, according to ISO 19101 (Geographic Information - Reference Model) and ISO 19109 (Geographic Information - Rules for application schema). Land Cover is created by polygons. Each polygon can be considered as homogeneous, differentiating itself from the neighboring polygons by the coverage or combination of coverage that forms it. Currently SIOSE is working on the SIOSE High Resolution project with a reference date of 2014 and 2017 (SIOSE 2020). SIOSE covers the years of 2005, 2011 and 2017.

### OSO Land Cover - France

The OSO Land Cover product is an annual (since 2016) land cover map of continental France, created by the land cover expertise group of the Theia Data and Services Center. Annual versions have been produced since 2016, with an increasing number of classes (17 in 2016, 23 since 2018). For 2020, it is planned to extend nomenclature to a total of 30 classes including wetlands. It is freely available in raster format at 10 $m^2$ and in vector format at 20 $m^2$. Multi-temporal optical image series with high spatial resolution from Sentinel-2 satellites were used as input data. It is also planned to use data from SPOT-6/7 and Pléiades in the future. Older versions of OSO LC (2009 - 2011, 2014) are based on Landsat 5 and Landsat 8 (2014 map) satellite images (Theia, 2020).

### Land Cover of Greece

The Land Cover of **Greece** map for 2010 was created by the University of the Aegean due to lack of up to date LC maps at national scale. It uses Landsat 5TM and Landsat 7ETM+ images acquired between 2008 - 2011 and semi-automated classification method using random forest. Nomenclature is adopted from the CLC at level 2 and consists of 12 classes. Overall accuracy reaches a value of 83 % (Gounaridis *et* al. 2016).

### CLC50 project - Hungary

The **CLC50 project** is a Hungarian national LC product at scale 1:50 000 derived from CLC100 for local needs prior to entering the EU. The input data is imagery from SPOT-4 taken in **1998-1999**. The minimum mapping unit is 4 ha (1 ha for water). The product uses CLC level-3 nomenclature with

extensions to level 4 and 5 with a total of 79 classes. Thematic reliability is claimed to be 90 %. The Land Cover map was created by a 'computer assisted photo interpretation' method (Büttner *et* al. 2004).

### LGN - The Netherlands

**LGN** is a set of national land use raster land cover files of the Netherlands, created by the Wageningen Environmental Research, a research institute affiliated with Wageningen University. Nine versions have been produced since 1986, each representing land use in one particular year, with spatial resolution of 25 m (5 m since 2018). Since 2018 LGN Land Cover maps have been published annually. LGN1 and LGN2 are experimental products with limitations with 17 resp. 25 classes. LGN3 (25 classes) with extension of LGN3plus increases the number of classes to 39. Changes in land cover over time can be followed since LGN3. LGN7 is the first version which provides a nationwide map of land cover for one reference year. Previously, the Netherlands was divided in two due to limited availability of satellite images. LGN2018 came with spatial resolution of 5 m and with a total of 48 land cover classes.

### Land Cover Map 2015 - The United Kingdom

The **Land Cover Map 2015** (LCM2015) is the latest land cover product with a 25m spatial resolution produced at the Centre for Ecology & Hydrology and published in 2017. It distinguishes 21 thematic classes. A Joint Nature Conservation Committee nomenclature was used (JNCC Broad Habitats). Random Forest was the algorithm chosen for conducting the image classification. The input imagery is from 2014 and 2015. The project of creating land cover over Great Britain started in 1990 with LCM1990, LCM2000, LCM 2007 products (Land Cover Map 2015).

## 3.3.1 Technical Characteristics of European National Land Cover Products

The same main set of the characteristics of the land cover products are compared in Table 3.4 to describe the variety of the existing national European land cover products/ databases. These are tabulated from the available internet information as provided by the national teams.

*Table 3.4 European national Land Cover characteristics*

| Product / Characteristics | LISA | Land Cover DE | SIOSE | OSO LC | LC Greece | CLC50 HU | LGN | LCM2015 |
|---|---|---|---|---|---|---|---|---|
| Spatial coverage | Austria | Germany | Spain | France | Greece | Hungary | The Netherlands | The UK |
| Thematic resolution (No. of classes) | 13 and 12 attributes | 7 | 20 | 23 (17 before 2018) | 12 | 79 | 48 (39 before 2018) | 21 |
| Nomen | CORINE | CORINE | SIOSE | OSO LC | CORIN | CORINE | LGN | JNCC |

| clature model | compatible | | | | E | level-3 | | Broad Habitats |
|---|---|---|---|---|---|---|---|---|
| Spatial resolution | | 10 m | min 15 m | 10 m (raster) 20 m (vector) | | | 5m (25 m before 2018) | 25 m/ 1 km |
| MMU | 25 $m^2$ 50 $m^2$ | | 0.5 ha 1 ha 2 ha | | 0.1 ha | 4 ha (1 ha for water) | | 0.5 ha |
| Temporal coverage | 2009÷2010 2010÷2012 2013÷2015 2016÷2018 | 27/6/2015 ÷ 29/9/2017 | 2005, 2009, 2011, 2014, 2017 | 2009÷2011, 2014, 2016÷2020 | 2008 ÷ 2011 | 1998 ÷ 1999 | 1984÷1987, 1990÷1994, 1995÷1997, 1999÷2000, 2003÷2004, 2007÷2008, 2012, 2018, 2019 | 1988÷1989 1998÷1999 2014÷2015 |
| Product release | | | 2009, 2015, 2018 | 2016÷2020 | | | | 1990, 2000, 2007, 2015, 2017, 2018, 2019 |
| Overall accuracy | 95% (source) | 93.1% (source) | not available | 87.8% (2017) 89.0% (2018) (source) | 83% (source) | 90% (source) | 90,4 % (LGN4), 80,5 % (LGN5) (source) | 85% (LCM2000) 83% (LCM2007) |
| Harmonization problem | not reported | N/A | Fully harmonized | N/A | N/A | N/A | N/A | |
| Product license | | © DLR/EOC | © Instituto Geográfico Nacional | Land Cover Thea SEC | Gounaridis *et* al. (2016) | Büttner, *et* al. (2004) | Wageningen University & Research | |
| Product URL | download | browse download | browse | download | Not available | Not available | contact | browse |
| Product access | Registration | Free | Free | Registration | N/A | N/A | Paid Subscription | |

## References on Land Cover

Banko, Stemberger.2015. LISA 2.0: Enhancing the land information system of Austria (LISA) with temporal information form SENTINEL-2 data. INSPIRE & COPERNICUS Conference 24.11.2015, Bratislava.

Buchhorn, M., Smets, B., Bertels, L., Lesiv, M., Tsendbazar, N.-E., & Li, L. 2019. Copernicus Global Land Operations "Vegetation and Energy."

Buchhorn, M., Smets,B.,Bertels, L., Lesiv, M., Tsendbazar, N.-E., Masiliunas, D., Herold, M., Fritz, S. 2019. Copernicus Global Land Service: Land Cover 100m, Collection 2, epoch 2015 Global(Version V2.0.2) [Data set]. Zenodo. DOI: 10.5281/zenodo.3243509.

Büttner, G. & Maucha, G. & Bíró, M. & Kosztra, B. & Pataki, R. & Petrik, O. 2004. National land cover database at scale 1:50,000 in Hungary. EARSeL eProceedings. 3.

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., & Mills, J. 2015. Global land cover mapping at 30 m resolution: A POK-based operational approach. ISPRS Journal of Photogrammetry and Remote Sensing, 103, 7–27. https://doi.org/10.1016/j.isprsjprs.2014.09.002.

Copernicus Land Monitoring Service. 2020. *European Settlement Map*. Available from https://land.copernicus.eu/pan-european/GHSL/european-settlement-map.

Copernicus Land Monitoring Service. 2020. *Urban Atlas*. Available from https://land.copernicus.eu/local/urban-atlas.

Copernicus Land Monitoring Service. 2020. *High Resolution Layers*. Available from https://land.copernicus.eu/pan-european/high-resolution-layers.

Defourny, P., Lamarche, C., Bontemps, S., & De Maet, T. 2015. *Land Cover CCI Product User Guide*. https://climate.esa.int/media/documents/CCI_Land_Cover_PUG_v2.0.pdf.

Di Gregorio, A., and Jansen, L.J.M. 2000, Land Cover Classification System (LCCS): Classification Concepts and User Manual, Rome: FAO/UNEP, 177 pp.

European Environment Agency. 1995. CORINE Land cover - Part 2: Nomenclature, Published 01 Jan 1995.

European Space Agency. 2020. Land Cover. https://climate.esa.int/en/projects/land-cover/
Gromny E., Lewiński S., Rybicki M., Malinowski R., Krupiński M., Nowakowski A., Jenerowicz M., 2019. Creation of training dataset for Sentinel-2 land cover classification, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2019, doi:10.1117/12.2536773
"Global Land Cover - Sentinel 2." Home | Global Land Cover - Sentinel 2, s2glc.cbk.waw.pl/.

Ferri, S., Siragusa, A., Sabo, F., Pafi, M., & Halkia, S. 2017. The European Settlement Map 2017 Release; *Methodology and output of the European Settlement Map (ESM2p5m)*. https://doi.org/10.2760/780799.

Gromny, E., Lewinski, S., Rybicki, M., Malinowski, R., & Krupinski, M. 2019. Creation of training dataset for Sentinel-2 land cover classification. *Proceedings Volume 11176, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments*. https://doi.org/https://doi.org/10.1117/12.2536773.

Gounaridis, D., Anastasios A. & Sotirios K.. 2016. Land cover of Greece, 2010: a semi-automated classification using random forests, Journal of Maps, 12:5, 1055-1062, DOI: 10.1080/17445647.2015.1123656.

Jansen, L. J.M., Groom, G. and Carrai, G. 2008. Land-cover harmonisation and semantic similarity: some methodological issues, Journal of Land Use Science, 3:2-3, 131-160, DOI: 10.1080/17474230802332076.

JRC. (2000). Global Land Cover 2000. https://forobs.jrc.ec.europa.eu/products/glc2000/glc2000.php.

Land and Water Division of FAO. 2020. Global Land Cover - SHARE (GLC-SHARE) | Land & Water | Food and Agriculture Organization of the United Nations | Land & Water | Food and Agriculture Organization of the United Nations.

http://www.fao.org/land-water/land/land-governance/land-resources-planning-toolbox/category/details/en/c/1036355/.

Latham, J., Cumani, R., Rosati, I., & Bloise, M. 2014. Global Land Cover SHARE Global Land Cover SHARE (GLC-SHARE) database Beta-Release Version 1.0-2014. http://www.fao.org/uploads/media/glc-share-doc.pdf.

Louisa J.M. Jansen †, Groom, G. & Carrai, G. 2008. Land-cover harmonisation and semantic similarity: some methodological issues, Journal of Land Use Science, 3:2-3, 131-160, DOI: 10.1080/17474230802332076.

Malinowski, R., Lewinski, S., & Rybicki, M. 2019. *Final Report*. http://users.cbk.waw.pl/~mkrupinski/S2GLC_Phase2_FinalReport.pdf.

National Mapping Organization. 2013. *Land Cover (GLCNMO) - Global version*. Available from https://globalmaps.github.io/glcnmo.html.

NEO NASA Earth Observations. 2011. *Land Cover Classification (1 year).* Available fromhttps://neo.sci.gsfc.nasa.gov/view.php?datasetId=MCD12C1_T1.

National Geomatics Center of China. 2020. *GLOBELAND30*. Available from http://www.globallandcover.com/defaults_en.html?src=/Scripts/map/defaults/En/browse_en.html&head=browse&type=data.

Perera, K. and Moore, D. and Apan, A. and McDougall, K. 2010. Applying the global standard FAO LCCS to map rural Queensland land cover. In: Queensland Surveying and Spatial Conference 2010 (QSSC 2010), 1-3 Sept 2010, Brisbane, Queensland.

Pflugmacher, D., Rabe, A., Peters, M., & Hostert, P. 2019. Mapping pan-European land cover using Landsat spectral-temporal metrics and the European LUCAS survey. *Remote Sensing of Environment*, 221, 583-595. doi:10.1016/j.rse.2018.12.001.

S2GLC. 2020. *Global Land Cover - Sentinel 2*. Available from http://s2glc.cbk.waw.pl/ .

Theia. 2019. *Land cover SEC - THEIA*. Available from https://www.theia-land.fr/en/ceslist/land-cover-sec/

UK Centre for Ecology & Hydrology. 2015. *Land Cover Map 15*. Available from https://www.ceh.ac.uk/services/land-cover-map-2015

VITO. 2019. *GLOBAL LAND COVER*. Available from https://lcviewer.vito.be/2019

Weigand, M., Staab, J., Wurm, M., & Taubenbock, H. 2020. Spatial and semantic effects of LUCAS samples on fully automated land use/land cover classification in high-resolution Sentinel-2 data. *International Journal of Applied Earth Observation and Geoinformation*, *88*. https://doi.org/https://doi.org/10.1016/j.jag.2020.102065

Yang, Hui & Li, S. & Chen, J. & Zhang, X. & Xu, Sh. 2017. The Standardization and Harmonization of Land Cover Classification Systems towards Harmonized Datasets: A Review. *ISPRS International Journal of Geo-Information*. 6. 154. 10.3390/ijgi6050154.

# 4 Environmental Quality Map Products

Environmental quality is the measure of a set of conditions of the environment relative to the requirements of one or more species (wildlife) and/or humans needs and preferences (Johnson *et* al. 1997). Specifically for the latter, those conditions are related to food production, prevention of land degradation, carbon sequestration, materials acquisition and production, wildlife and environment management, human health and well-being and ecosystem services in general.

In the context of climate change, the importance of these services has increased and so is the interest of governments in them. Therefore, there is a growing need for accurate, up-to-date, spatially referenced information on the quality of air, water and soil. For mapping purposes, in this section the evaluation of environmental quality products are then split in three different sections.

## 4.1 Air Quality Products

### 4.1.1 Introduction

It is not feasable to solely rely on ground observations to study air quality, since this process requires a great number of ground stations in order to cover all areas, which requires high costs and efforts (Ibrahim and Halounova, 2019). Many studies used satellite remote data to observe the suspended particles in the air. The most common one is the MODerate resolution Imaging Spectroradiometer (MODIS), which provides many products with different spatial and temporal resolutions to measure aerosol optical depth (AOD) on regional and global scales. MODIS products have many scientific data sets (SDS), and one of them is the quality assurance that defines conditions during observations. Generally, it is recommended to use high quality data when working with air quality.

Particulate matter (PM) with a diameter of less than or equal to 10 μm (PM10) is capable of infiltrating the tracheobronchial. Such particles become more dangerous as the diameter gets smaller, and thus PM2.5 particles are very harmful as they not only cause severe respiratory problems, but also reach systemic circulation. Many studies found that there was a relationship between AOD and PM2.5. This relationship can be explained by using some auxiliary data such as; wind speed, temperature, atmospheric pressure, in addition to some topographical and other data like elevation, Normalised Difference Vegetation Index (NDVI), etc.

Our goal is to produce open-source daily PM2.5 maps over Europe with high spatial resolution (~ 1 km) by using remote sensing data and machine learning algorithms.

### 4.1.2 Existing air quality models / products with open-source data

In order to communicate with the public of how clean or polluted the air is, agencies usually use the air quality index (AQI), which is computed using the air pollutant concentration over a specified averaging period. Some models visualize the values of one or more pollutants with other tools that help the user to understand the topic like charts and tables. In the following table, we will show some models which provide one or more open-source air quality data.

*Table 4.1 List of existing air quality products provided as open-source data*

| ID | AQ name | Coverage | Spatial resolution | Time resolution | Maintainer | Access |
|---|---|---|---|---|---|---|
| 1 | World air map | Global | - | 1 hour | Plume labs | Free / visualization with charts for some cities |
| 2 | ENSEMBLE MODEL | 25W/30N/45E/72N since 12th June 2019 - 25W/30N/45E/70N before | 0.1° x 0.1° | 1 hour | Copernicus | Free / forecast visualization |
| 3 | SILAM | Global | 0.1° x 0.1° | 1 hour | Finnish Meteorological Institute | Free/ Visualization and data download |
| 4 | Berkeley Earth | Global | 0.1° x 0.1° | 1 hour | Berkeley Earth Team | Free/ Visualization and data download |
| 5 | CHIMERE | Whole Europe from 10.4375°W to 29.9375°E | 0.1°x0.1° | 1 hour | INERIS (France) | Free register/ Visualization and data download |
| 6 | EMEP | Europe | 0.125° x 0.0625° | Daily | MET Norway (Norway) | Free/ Visualization and data download |
| 7 | EURAD-IM | 25°E to 45°W and 30°N to 70°N | 9 km on a Lambert conformal projection | Daily | FZJ-IEK8 (Germany) | Forecast visualization |
| 8 | LOTOS-EUROS | 15°W to 35°E and 35°N to 70°N | 0.1° x 0.1° | 1 hour | KNMI, TNO (The Netherlands) | Forecast visualization |
| 9 | ECCAD | Global | 0.1° x 0.1° | Daily/Monthly/Yearly | Laboratoire d'Aérologie/CNRS - France | Register / Visualization and download |
| 10 | CAMS | Global | 80 km | 1 hour | Copernicus | Register / download |
| 11 | European Air Quality Index | Europe | point representation | NRT | EEA | Free /data visualization with charts and tables |

### World Air Map (ID 1)

World Air Map uses machine learning algorithms and data fusion from a variety of sources including 12,000 government-operated monitoring stations, and a variety of satellite datasets, Copernicus Atmospheric Monitoring Service, atmospheric models and weather datasets to give information on current and future levels of air pollution. They provide a free mobile app called Plume Air Report to give information on current and future levels of air pollution. The information of the used algorithms is not available.

### ENSEMBLE Model (ID 2)

This model provides predictions of daily mean and maximum concentrations of aerosols in the air like (PM2.5, PM10, SO2, CO, NO2, Dust, SIA). Based on a sample of individual model members, the ENSEMBLE approach is suitable for air quality monitoring (Galmarini *et* al. 2004). The ENSEMBLE products have better performance than the individual model products. In addition, the spread between different members may be used to provide some insight on the uncertainty of the ENSEMBLE products. Consequently, the forecasts, analyses and re-analyses delivered as part of the CAMS Regional production are based on an ENSEMBLE approach. The ENSEMBLE model currently uses a median value approach (Marecal *et* al, 2015), of eight models: CHIMERE, EMEP, EURAD-IM, LOTOS-EUROS, MATCH, MOCAGE, SILAM, DEHM, GEM-AQ.

ENSEMBLE provides daily Air quality near real time (NRT) EPSgrams for 67 major European cities and urban areas and these graphics are displayed on the CAMS website for the Regional Air Quality. The chosen sites include 41 European capitals and 26 urban areas that are among the most populated and those areas highly polluted. The forecasts are based upon models that have resolutions of 10 km to 25 km, which is too coarse to take into consideration the quiet local and urban effects (high primary pollutants, titration of ozone, etc.) (METEO-FRANCE / G. Collin, 2020).

### SILAM Model (ID 3)

The System for Integrated modeLling of Atmospheric coMposition SILAM v.5.6 is a Eulerian chemical transport model (CTM), that contains a set of supplementary tools including a meteorological pre-processor, input-output adapters, reprojection and interpolation routines, etc. In the operational forecasts, these allowed direct forcing of the model by the European Centre for Medium-Range Weather Forecasts (ECMWF) IFS meteorological fields.

### CHIMERE Model (ID 4)

The multi-scale model is mainly intended to produce daily forecasts of ozone, aerosols and other pollutants, and to create long-term simulations for emission control scenarios. CHIMERE runs over a range of spatial scales from the regional to the urban scales, with resolutions from 1-2 Km to 100 Km. The chemical mechanism (MELCHIOR) is derived from the original EMEP mechanism. 6 aerosol sizes are represented as bins in the model. The numerical time solver is the TWOSTEP method.

### Berkeley Earth Data (ID 5)

Berkeley Earth is an independent U.S. noncommercial organization focused on environmental data science. Berkeley Earth supplies comprehensive open-source world air pollution data and highly user-accessible global temperature data that is timely, impartial, and verified. Berkeley applies Kriging interpolation method to derive pollution maps.

### EMEP MSC-W Model (ID 6)

This model is a chemical transport model designed by the Norwegian Meteorological Institute under the EMEP program (UN Convention on Long-range Transboundary Air Pollution). The EMEP MSC-W model system provides several possibilities concerning the chemical schemes used and the ability to include aerosol dynamics. Simpson et al. (2012) reviewed the EMEP MSC-W model in detail, as well as the main model updates since 2006. The forecast version of the EMEP MSC-W model (EMEP-CWF) has been functioning since June 2006. The scheduled model updates in CAMS_50 ensure that the model version stays as close as possible to the official EMEP Open Source version.

### EURAD-IM Model (ID 7)

EURopean Air Pollution Dispersion Inverse Model (EURAD-IM) is a three-dimensional, non-hydrostatic, multi-scale Chemistry and Transport Model. This system consists of 5 main components: the meteorological driver WRF, the pre-processors EEP and PREP for preparation of anthropogenic emission data, the EURAD-IM Emission Model EEM, and the chemistry transport model EURAD-IM (Hass et al. 1995, Memmesheimer et al., 2004). EURAD-IM involves advection, diffusion, chemical transformation, wet and dry deposition, sedimentation of tropospheric trace gases, and aerosols (METEO-FRANCE / G. Collin 2020). The system includes 3d-var and 4d-var chemical data assimilation (Elbern et al. 2007), and it serves both scientific system analyses and operational forecasts. In both modes the nesting technique is applicable from continental scale to local scale with 1 km resolution.

### LOTOS-EUROS Model (ID 8)

The LOTOS (Long Term Ozone Simulation) model that was originated from the US UAM (Urban Airshed Model)(Manders-Groot, *et* al. 2016), and the Eulerian air quality model EUROS (EURopean Operational Smog model) that was developed at RIVM for modeling winter smog periods over Europe (Egmond and Kesseboom, 1981) were unified in a single chemistry model known as LOTOS-EUROS version 1.0 (Schaap *et* al, 2005).

LOTOS-EUROS is a 3D chemistry transport model mainly designed to simulate air pollution in the lower troposphere. Many studies used this model to assess particulate air pollution and trace gases (METEO-FRANCE / G. Collin 2020). This model played a frequent role in international model comparisons addressing ozone and particulate matter (e.g. Solazzo *et* al. 2012b, Stern *et* al. 2008).

### ECCAD Data (ID 9)

ECCAD is the Global Emission InitiAtive (GEIA) data portal and is part of AERIS which is the French data service for atmosphere. ECCAD offers a vast number of datasets at global and regional scales, at different spatial and temporal resolutions and time periods, with tools to allow visualization, and on-line data analysis and comparison. Each dataset is provided with statistical information over geographical regions, in addition to the possibility of downloading data as Network Common Data Form (NetCDF) files with detailed metadata.

### CAMS Data (ID 10)

There are three main global production streams in the Copernicus Atmosphere Monitoring System (CAMS). Analyses and forecasts for aerosols and chemical species are produced twice a day; analyses and forecasts for greenhouse gases are produced once a day at increased spatial resolution; and global reanalyses are produced every few years.

### European Air Quality Index (ID 11)

The European Air Quality Index is one of the European Environment Agency products, which enables users to display up-to-the-minute data for each country, region and city in Europe.

## 4.1.3 Chemistry and Transport models / products (data upon request)

Table 4.2 shows four chemistry and transport models which provide air quality data upon request, all these models take part in the ENSEMBLE Model.

*Table 4.2 Chemistry and transport models which provide air quality data upon request*

| ID | AQ name | Coverage | Spatial resolution | Time resolution | Maintainer |
|---|---|---|---|---|---|
| 1 | MATCH | Sweden and the Baltic Sea region | 0.1° x 0.1° | - | SMHI (Sweden) |
| 2 | MOCAGE | Global | 0.1° regular lat-lon grid for forecast 0.2° regular lat-lon grid for assimilation | 1 hour | Météo-France (France) |
| 3 | DEHM | Northern Hemisphere | 18 km at 60°N | 1 hour | AARHUS UNIVERSITY (Denmark) |
| 4 | GEM-AQ | Global | 0.1° x 0.1° latitude/longitude spherical grid | Hourly, performed once a day for the previous day | IEP-NRI (Poland) |

### MATCH Model (ID 1)

The Multi-Scale Atmospheric Transport and Chemistry model (MATCH) is a three-dimensional, Eulerian model developed at the Swedish Meteorological and Hydrological Institute (SMHI) and used for Air Quality assessment in Sweden and the Baltic Sea region. MATCH includes different weather data fitted to different resolutions and projections and a range of alternative schemes for deposition and chemistry.

### MOCAGE Model (ID 2)

MOCAGE is a three-dimensional, multi-scale Chemistry and Transport Model that has been developed by Meteo-France for simulating atmospheric composition, includi ng both gases and aerosols. Since the year 2000, MOCAGE has been used to study the impact of pollutant anthropogenic emissions on climate change, and has been running daily since 2005. In 2004 Météo-France became part of consortium and operational platform 'PREV'AIR' (Rouil *et* al. 2009).

### DEHM Model (ID 3)

The Danish Eulerian Hemispheric Model (DEHM) is a three-dimensional, offline, large-scale, cutting-edge, atmospheric chemistry transport model developed to study long-range transport of air pollution in the Northern Hemisphere. DEHM was first developed in the early 1990's to monitor the atmospheric transport of sulphur-dioxide and sulphate into the Arctic (Christensen 1997; Heidam *et* al.,1999; Heidam *et* al. 2004). The model has been continuously modified, extended and updated since its launch. DEHM is a part of the integrated THOR system and the human exposure modelling system DEHM/UBM/AirGIS.

### GEM-AQ Model (ID 4)

GEM-AQ defined as an on-line chemical weather forecasting model (Kaminski et al. 2008). The model can be used to simulate atmospheric processes over a wide range of scales, from the global scale down to the meso-gamma (METEO-FRANCE / G. Collin, 2020). This model uses the meteorological fields of the operational IFS for initial and boundary. Data assimilation in the GEM-AQ modelling system is performed with Optimal Interpolation method (Robichaud and Menard 2014) and applied to the forecast. Currently, data assimilation is carried out at each forecast hour for O3, NO2, PM10 and PM2.5, using surface observations provided by Météo-France for the previous day.

## 4.1.4 Technical specification of products used in the air quality task in the project

Table 4.3 presents technical specifications of the products which will be used in the air quality task. All these products are publicly available.

*Table 4.3 List of products which will be used in the air quality task.*

| Product Characteristics | MCD19A2 | ERA5-Land hourly data | Particulate Matter (PM2.5) |
|---|---|---|---|
| Spatial coverage | Global | Global | Global |
| Spatial resolution | 1 km | 0.1 deg × 0.1 deg | Point measurements |
| Temporal resolution / updates | Daily | 6 Hours | Hourly |
| Data availability | 2010 - to present | 1981 - to present | Depends on location |
| Product license | NASA: open | Copernicus: open | The MIT License (MIT) Copyright © 2015 OpenAQ |
| Product URL | Download | Download | Download |
| Product access | Free access | Free access | Free for the last 90 days |
| Citation | Lyapustin, A., Wang, Y. (2018). *MCD19A2 MODIS/Terra+Aqua Land Aerosol Optical Depth Daily L2G Global 1km SIN Grid V006* [Data set]. NASA EOSDIS Land Processes DAAC. | Muñoz Sabater, J., (2019): ERA5-Land hourly data from 1981 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (<date of access>), 10.24381/cds.e2161bac | "OpenAQ, OpenAQ.org [online] Available from: https://openaq.org (Accessed DATE OF ACCESS." |

### MCD19A2

This product of MODIS collection 6 is generated using the Multi-angle Implementation of Atmospheric Correction (MAIAC) algorithm applied on both the Terra and the Aqua datasets. MCD19A2 is produced daily at 1 km pixel resolution.

The MCD19A2 product contains many Science Dataset (SDS) layers: blue band AOD at 0.47 µm, green band AOD at 0.55 µm, AOD uncertainty, fine mode fraction over water, column water vapor over land and clouds, smoke injection height, AOD QA, AOD model at 1 km, cosine of solar zenith angle, cosine of view zenith angle, relative azimuth angle, scattering angle, and glint angle at 5km. A low-resolution browse image is also included showing AOD of the blue band at 0.47 µm created using a composite of all available orbits. MCD19A2 is stored as Hierarchical Data Format 4 (HDF4), and each SDS layer contains a third dimension that represents the number of orbit overpasses.

### ERA5-Land hourly

ERA5-Land has been produced by replaying the land component of the ECMWF ERA5 climate reanalysis. Reanalysis combines model data with global observations into a globally complete and consistent dataset using the laws of physics. We will use this product to extract the meteorological variables which will be used as auxiliary data in our model.

### OpenAQ/PM2.5

OpenAQ is a not-for--profit organization that shares air monitoring station observations and research grade sources from 95 different countries around the world, users can download data according to the country or to specific locations. Data are aggregated for many pollutants like PM2.5, PM10, O3, SO2, NO2, CO, and BC.

### References on Air Quality


CAMS50_2018SC2_D2.0.2-U2_Models_documentation_202003_v2.

Christensen, J. H. 1997. The Danish Eulerian Hemispheric Model – a three-dimensional air pollution model used for the Arctic. *Atmospheric Environment*, 31, 4169–4191.

Egmond, N.D. and H. van Kesseboom. 1981. *Numerieke verspreidingsmodellen voor de interpretatie van de meetresultaten van het Nationaal Meetnet voor Luchtverontreiniging*. Tech. rep. Bilthoven: RIVM, RIVM report 227905048.

Elbern, H., A. Strunk, H. Schmidt, and O. Talagrand. 2007. Emission rate and chemical state estimation by 4-dimensional variational inversion. *Atmos. Chem. Phys.*, 7, 1-59.

Galmarini, S., R. Bianco, R. Addis, S. Andronopoulos, P. Astrup, J. C. Bartzis, R. Bellasio, R. Buckley, H. Champion, M. Chino, R. D'Amours, E. Davakis, H. Eleveld, H. Glaab, A. Manning, T. Mikkelsen, U. Pechinger, E. Polreich, M. Prodanova, H. Slaper, D. Syrakov, H. Terada and L. Van der Auwera, 2004. Ensemble dispersion forecasting - Part II: application and evaluation. *Atmospheric Environment*, 38, 28, 4619 – 4632.

Hass, H., H.J. Jakobs, and M. Memmesheimer. 1995. Analysis of a regional model (EURAD) near surface gas concentration predictions using observations from network. *Meteor. Meteor. Atmos. Phys.*, 57, 173-200.

Heidam, N. Z., Christensen, J., Wahlin, P., and Skov, H. 2004. Arctic atmospheric contaminants in NE Greenland: levels, variations, origins, transport, transformations and trends 1990-2001. *Science of the Total Environment*, 331, 5-28.

Heidam, N. Z., Wahlin, P., and Christensen, J. H. 1999. Tropospheric gases and aerosols in northeast Greenland. *Journal of the Atmospheric Sciences*, 56, 261-278.

Ibrahim, S. and Halounova, L. 2019. Statistical study of MODIS algorithms in estimating aerosol optical depth over the Czech Republic. *The civil engineering journal*, 4, article No:43.DOI 10.14311/CEJ.2019.04.0043.

Kaminski, J. W., Neary, L., Struzewska, J., McConnell, J. C., Lupu, A., Jarosz, J., Toyota, K., Gong, S. L., Cote, J., Liu, X., Chance, K., and Richter, A. 2008. GEM-AQ, an on-line global multiscale chemical weather modelling system: model description and evaluation of gas phase chemistry processes. *Atmos. Chem. Phys.*, 8, 3255-3281, doi:10.5194/acp-8-3255-2008.

Koble R. and G. Seufert. 2001. *Novel Maps for Forest Tree Species in Europe*.

Marecal *et* al. 2015. A regional air quality forecasting system over Europe: the MACC-II daily ensemble production. *Geosci. Model Dev.*, 8, 2777-2813.

Memmesheimer, M. *et* al. 2004. *Int. J. Environm. and Pollution*, 22, (1-2), 108-132.

Rouil, L., *et* al. 2009. PREV'AIR: an operational forecasting and mapping system for air quality in Europe, *Bull. Am. Meteor. Soc.*, 90(1), 73-83, doi:10.1175/2008BAMS2390.1.

Schaap, M., Roemer, M., Sauter, F., Boersen, G., Timmermans, R. M. A., and Builtjes, P. J. H. 2005. LOTOS-EUROS documentation, TNO report B&O 2005/297, TNO, Apeldoorn, the Netherlands.

Simpson, D. *et* al. 2012. The EMEP MSC-W chemical transport model – technical description, *Atmos. Chem. Phys.*, 12, 7825–7865, 2012.

Simpson, D., Fagerli, H., Jonson, J.E., Tsyro, S., Wind, P., Tuovinen J. P. 2003. Transboundary Acidification, Eutrophication and Ground Level Ozone in Europe, PART I. *Unified EMEP Model Description*, 104 pp.

Sofiev, M., Vira, J., Kouznetsov, R., Prank, M., Soares, J., Genikhovich, E. 2015. Construction of an Eulerian atmospheric dispersion model based on the advection algorithm of M.Galperin: dynamic cores v.4 and 5 of SILAM v.5.5. *Geosci. Model Developm.* Discuss., 8, 2905-2947, doi:10.5194/gmdd- 8-2905.

Solazzo, E., Bianconi, R., Pirovano, G., Matthias, V., Vautard, R., Moran, M.D., Wyat Appel, K., Bessagnet, B., Brandt, J., Christensen, J.H. 2012. Operational Model Evaluation for Particulate Matter in Europe and North America in the Context of AQMEII. *Atmos. Environ.,* 53, 75-92.

Stern, R., Builtjes, P., Schaap, M., R. Timmermans, R. Vautard, A. Hodzic, M. Memmesheimer, H. Feldmann, E. Renner R. Wolke and A. Kerschbaumer. 2008. A model inter-comparison study focusing on episodes with elevated PM10 concentrations. *Atmos. Environ.*, 42, 4567-4588.

## 4.2 Water Quality Products

Water quality refers to setting safe standard conditions of water bodies for surface water bodies (lakes, rivers, seas/oceans) and groundwater bodies, following the definitions given in the Water Framework Directive and the 1st and 2nd River Basin Management Plans (Directive 2000/60/EC of the European Parliament and of the Council, 2000). These standard conditions are mainly related to presence/absence and concentrations of chemical agents or microorganisms. Overall substances identified as pollutant agents or threats to aquatic environments can be categorized into four classes:

- Biological (i.e. algae, angiosperms),
- Microbiological (i.e., plankton, hormones),
- Physical (i.e., temperature, conductivity),
- Chemical (i.e., nitrogen, phosphorus, pH).

On the European level, the European Environment Agency (EEA) is the reference point for all relevant water data and information. The agency works closely with National Focal Points (NFPs) of the 32 member countries (EU-27 plus Iceland, Liechtenstein, Norway, Switzerland and Turkey) and cooperating countries (Albania, Bosnia and Herzegovina, North Macedonia, Montenegro, Serbia and Kosovo). Information from the NFPs is collected through the environment information and observation network (Eionet), which the agency is in charge of developing and coordinating.

The water quality products now available were produced as part of the implementation strategy of the Directive 2000/60/EC of the European Parliament, known also as EU Water Framework Directive (WFD).

*Table 4.4 List of products available for water quality*

| ID | Name | Coverage | Spatial resolution data type | Time resolution | Maintainer | Access |
|---|---|---|---|---|---|---|
| 1 | Nitrogen concentration in rivers | EU-28 | 1 km/raster | 2004 - 2009 | Corresponding author | Request to the authors |
| 2 | Phosphorus concentration in rivers | EU-28 | 1 km/raster | 2004 - 2009 | Corresponding author | Request to the authors |
| 3 | Pollution from urban runoff | EU-28 (minus Greece and Cyprus) | 1 km/raster | 2004 - 2009 | Corresponding author | Request to the authors |
| 4 | Phytoplankton (Water Framework Directive 2nd River Basin Management plans) | EU-28 | Vector data | 2016 - 2019 | European Environmental Agency | Open |
| 5 | Other aquatic flora (Water Framework Directive 2nd River Basin Management plans) | EU-28 | Vector data | 2016 - 2019 | European Environmental Agency | Open |
| 6 | Macroalgae (Water Framework Directive 2nd River Basin Management plans) | EU-28 | Vector data | 2016 - 2019 | European Environmental Agency | Open |
| 7 | Phytobenthos (Water Framework Directive 2nd River Basin Management plans) | EU-28 | Vector data | 2016 - 2019 | European Environmental Agency | Open |
| 8 | Hydrological or tidal regime (Water Framework Directive 2nd River Basin Management plans) | EU-28 | Vector data | 2016 - 2019 | European Environmental Agency | Open |

| | | | | | | |
|---|---|---|---|---|---|---|
| 9 | Oxygenation conditions (Water Framework Directive 2nd River Basin Management plans) | EU-28 | Vector data | 2016 - 2019 | European Environmental Agency | Open |
| 10 | Salinity conditions (Water Framework Directive 2nd River Basin Management plans) | EU-28 | Vector data | 2016 - 2019 | European Environmental Agency | Open |
| 11 | Acidification status (Water Framework Directive 2nd River Basin Management plans) | EU-28 | Vector data | 2016 - 2019 | European Environmental Agency | Open |

### 4.2.1 Product Descriptions

A brief description of each of the water quality products is provided down below. Most of the listed products available are provided by the EEA, which are aggregated and compiled the datasets based on the information gathered and provided by the relevant national agencies (NFPs). The spatial coverage of reference for each product is EU-28, but in some of them some countries are missing. Temporal coverage available is mostly recent (from 2016 to 2019), but there is no information for the whole Geo-harmonizer time span (2000-2020). The description is thematically-based, providing a detailed content of the dataset and the relevant references on the legal context that regulates the individual indicator. Full references are provided separately at the end of this chapter.

#### Nitrogen concentration in rivers

It is an estimated nitrogen concentration in rivers (mgN/l), based on the model GREEN (Grizzetti *et* al. 2011). Nitrogen concentration is one of 12 proposed indicators by Grizzetti *et* al. (2017) to assess and quantify the level of water pollution acting on European rivers. The paper tries to establish a set of objective indicators to quantify the major pressures on European rivers, pressures that were identified by a European Commission report (European Commission 2015).

#### Phosphorus concentration in rivers

It is estimated phosphorus concentration in rivers (mgN/l), based on the model GREEN (Grizzetti *et* al. 2011). Phosphorus concentration is one of 12 proposed indicators by Grizzetti *et* al. 2017 to assess and quantify the level of water pollution acting on European rivers. The paper tries to establish a set of objective indicators to quantify the major pressures on European rivers, pressures that were identified by a European Commission report (European Commission 2015).

#### Pollution from urban runoff

It is the relative intensity of the potential pollution load from urban runoff (dimensionless), which is estimated by the Heaney model (Heaney *et* al. 1976). The indicator is designed to reproduce potential pollution and not specific contaminants, based on urban land cover (CLC 2006), annual precipitation and population. It is one of 12 proposed indicators by Grizzetti *et* al. (2017) to quantify the major

pressures acting on European rivers according to a European Commission report (European Commission 2015).

### Phytoplankton

It is one of the **biological quality indicators** for **good ecological status** as defined in the Water Framework Directive and in the 2nd River Basin Management Plan. It shows the influence of pressures (e.g., pollution and habitat degradation) on the identified quality elements. The overall ecological status classification for a water body is determined, according to the 'one out, all out' principle, by the element with the worst status out of all the biological and supporting quality elements. Definitions for high, good, and moderate are found in Section 1.2, Annex V of directive 2000/60/EC.

### Other aquatic flora

It is one of the **biological quality indicators** for **good ecological status** as defined in the Water Framework Directive and in the 2nd River Basin Management Plan. See 'Phytoplankton' for more information.

### Macroalgae

It is one of the **biological quality indicators** for **good ecological status** as defined in the Water Framework Directive and in the 2nd River Basin Management Plan. See 'Phytoplankton' for more information.

### Phytobenthos

It is one of the **biological quality indicators** for **good ecological status** as defined in the Water Framework Directive and in the 2nd River Basin Management Plan. See 'Phytoplankton' for more information.

### Hydrological or tidal regime

It isone of the **hydromorphological quality indicators** for **good ecological status** as defined in the Water Framework Directive and in the 2nd River Basin Management Plan. It shows the influence of pressures (e.g., pollution and habitat degradation) on the identified quality elements. The overall ecological status classification for a water body is determined, according to the 'one out, all out' principle, by the element with the worst status out of all the biological and supporting quality elements. Definitions for high, good, and moderate are found in Section 1.2, Annex V of directive 2000/60/EC

### Oxygenation conditions

These are **chemical quality indicators** for **good chemical status** as defined in the Water Framework Directive and in the 2nd River Basin Management Plan. They show the influence of pressures (e.g., pollution and habitat degradation) on the identified quality elements. Good chemical status means that no concentrations of priority substances exceed the relevant EQS established in the Environmental Quality Standards (EQS) Directive 2008/105/EC (as amended by the Priority Substances Directive 2013/39/EU, Annex I and II). EQS aim to protect the most sensitive species from direct toxicity, including predators and humans via secondary poisoning.

### Salinity conditions

One of the **chemical quality indicators** for **good chemical status** as defined in the Water Framework Directive and in the 2nd River Basin Management Plan.

### Acidification status

One of the **chemical quality indicators** for **good chemical status** as defined in the Water Framework Directive (WFD) and in the 2nd River Basin Management Plan.

### References on Water Quality Products

European Commission. 2015. The water framework directive and the floods directive: actions towards the 'good status' of EU water and to reduce flood risks. COM (2015) 120.

Grizzetti, B., Pistocchi, A., Liquete, C., Udias, A., Bouraoui, F., Bund, W. V. D. 2017. Human pressures and ecological status of European rivers. Scientific Reports, 7(1). http://doi.org/10.1038/s41598-017-00324-3.

Grizzetti, B., Bouraoui, F., Aloe, A. 2011. Changes of nitrogen and phosphorus loads to European seas. *Global Change Biology*, 18(2), 769–782. http://doi.org/10.1111/j.1365-2486.2011.02576.x.

Heaney, J. P., Huber, W. C., Nix, S. J. 1976. Storm water management model: level I, preliminary screening procedures. U.S. Environmental Protection Agency, Office of Research and Development, Municipal Environmental Research Laboratory.

Panagos, P., Ballabio, C., Borrelli, P., Meusburger, K., Klik, A., Rousseva, S., Alewell, C. 2015. Rainfall erosivity in Europe. *Science of The Total Environment*, 511, 801–814. http://doi.org/10.1016/j.scitotenv.2015.01.008.

Tóth, G., Hermann, T., Szatmári, G., Pásztor, L. 2016. Maps of heavy metals in the soils of the European Union and proposed priority areas for detailed assessment. *Science of The Total Environment*, 565, 1054–1062. http://doi.org/10.1016/j.scitotenv.2016.05.115.

## 4.3 Soil Quality Products

Soil quality is best defined using four groups of variables:

1. Soil fertility and soil health indices (the Cornell University Framework for soil health),
2. Soil physical soil properties (e.g., water holding capacity, erosion risk, etc.),
3. Soil pollution measures (heavy metals and similar chemicals),
4. Soil biodiversity measures (concentrations of species: macro-fauna, fungi, nematodes, micro-organisms).

Soil fertility is typically quantified by availability or soil nutrients important for growth of specific crops, soil pH and Cation Exchange Capacity, i.e., capacity of soil to feed the plants with nutrients. Soil physical properties such as soil texture fractions, bulk density, water retention properties and , usually are important because they determine suitability of soil for different tillage

applications and the water holding capacity. As soil of high quality usually implies that it does not require intensive irrigation and that the tillage is inexpensive.

Most soil data (point or soil profile samples) are collected either via agricultural land soil surveys, forest land soil surveys and geochemical surveys (GEMAS project). Point data is then used to create predictions of soil variables over the whole space using some of the Predictive Soil Mapping techniques, most recently dominantly by using Machine Learning (e.g. Hengl and MacMillan 2018). In modern soil science, soil data is increasingly collected by Mid-Infrared soil spectroscopy methods (Sanderman *et* al, 2019, Orgiazzi *et* al. 2019). In Europe, most of the data products can be obtained via the European Soil Data Centre at JRC, or via the European Environmental Agency. There are also numerous country-based products which are reviewed in the Soil Geographical Databases compendium by David G. Rossiter.

*Table 4.5 List of products available for soil quality*

| ID | Name | Coverage | Spatial resolution | Reference year/period | Maintainer | Access |
|---|---|---|---|---|---|---|
| 1 | Topsoil organic carbon content | EU-25 | 1 km | 2009 | JRC | Open, by request |
| 2 | Topsoil physical properties | EU-25 | 500 m | 2015 | JRC | Open, by request |
| 3 | Soil erosion by water | EU-28 | 100 m | 2010 | JRC | Open, by request |
| 4 | Soil pH in Europe | EU-25 | 5 km | 2009 | JRC | Open, by request |
| 5 | Cover Management factor | EU-28 | 100 m | 2010 | JRC | Open, by request |
| 6 | Natural susceptibility to soil compaction | EU-27 | 1 km | 2008 | JRC | Open, by request |
| 7 | Saline and sodic soils | EU-27 | 1 km | 2008 | JRC | Open, by request |
| 8 | Rainfall erosivity | EU-28 + Switzerland | 500 m | 2000 - 2010 | JRC | Open, by request |
| 9 | Copper distribution in topsoil | EU-25 (minus Croatia, Cyprus, Malta) | 500 m | 2018 | JRC | Open, by request |
| 10 | Heavy metals in topsoil | EU-28 (minus Croatia) | 1 km | 2018 | JRC | Open, by request |

| | | | | | | |
|---|---|---|---|---|---|---|
| 11 | Soil Bulk Density | Global | 0.5 degree | 2005 | FAO | Registration |
| 12 | Heavy metals in topsoil (GEMAS) | EU-25 | 50 km | 2012 | Eurogeosurveys | By request |

### 4.3.1 Product Descriptions

A brief description of each of the soil quality products is provided down below. Most of the listed products available are provided by the European Soil Data Centre at JRC. Note that soil maps are constantly updated and are increasingly available at finer and finer spatial resolutions of better than 250 up to 100 and 30 m.

#### Topsoil organic carbon content

The layer  contains maps for Topsoil Soil Organic Carbon in EU-25 that are based on LUCAS 2009 soil point data through a generalized additive model. The map of predicted topsoil organic carbon content (g C kg-1) was produced by fitting a generalised additive model between organic carbon measurements from the LUCAS survey (dependent variable) and a set of selected environmental covariates; namely slope, land cover, annual accumulated temperature, net primary productivity, latitude and longitude. It also includes a Map of Standard Error of the OC model predictions (g C kg-1).

#### Topsoil physical properties

The layer contains 7 soil property maps that have been derived using soil point data from the LUCAS 2009 soil survey (around 20,000 points) for EU-25, using hybrid approaches like regression kriging. Properties included are (Ballabio *et* al. 2016; Ballabio *et* al. 2019):

- Clay content (%) in topsoil (0-20cm)
- Silt content (%) in topsoil
- Sand content (%) in topsoil
- Coarse fragments (%) content in topsoil
- Bulk density derived from soil texture datasets (obtained from the packing density and themapped clay content following the equation of Jones *et* al. 2003)
- USDA soil textural classes derived from clay, silt and sand maps
- Available Water Capacity (AWC) for the topsoil fine earth fraction

#### Soil erosion by water

The layer contains the Soil Loss by Water Erosion in Europe and is the result of applying a modified version of the Revised Universal Soil Loss Equation (RUSLE) model, RUSLE 2015, which delivers improved estimates based on higher resolution (100 m compared to 1 km) peer-reviewed inputs of rainfall, soil, topography, land use and management from 2010.

#### Soil pH in Europe

The layer contains a quantitative map of estimated soil pH values across Europe from a compilation of 12,333 soil pH measurements from 11 different sources, and using a geo-statistical framework based on regression kriging. Fifty-four (54) auxiliary variables in the form of raster maps at 5km

resolution were used to explain the differences in the distribution of soil pHCaCl2 and the kriged map of the residuals from the regression model was added. The lowest values correspond to the soils developed on acid rock (granites, quartzites, sandstones, etc.), while the higher values are related to the presence of calcareous sediments and basic rocks.

### Cover Management factor

The layer contains the Cover Management factor (C-factor), one of the input layers when calculating the Universal Soil Loss Equation (USLE) model, which is the most frequently used model for soil erosion risk estimation. The C-factor was estimated for a) arable lands based on crop composition and for b) all other land uses (non-arable) based on the vegetation density and land cover type. The management practices (reduced tillage/no till, plant residues and winter cover crops) were taken into account in estimating C-factor in arable lands.

### Natural susceptibility to soil compaction

The layer contains the natural susceptibility of agricultural soils to compaction if they were to be exposed to compaction. The evaluation of the soil's natural susceptibility is based on the creation of logical connections between relevant parameters (pedotransfer rules). The input parameters for these pedotransfer rules are taken from the attributes of the European soil database, e.g., soil properties: type, texture and water regime, depth to textural change and the limitation of the soil for agricultural use. Also auxiliary parameters have been used as impermeable layer, depth of an obstacle to roots, water management system, dominant and secondary land use. It was assumed that every soil, as a porous medium, could be compacted.

### Saline and sodic soils

The layer contains the area distribution of saline, sodic and potentially salt affected areas within the European Union. The accuracy of input input data only allows the designation of salt affected areas with a limited level of reliability (e.g. < 50 or > 50% of the area); therefore the results represented in the map should only be used for orientating purposes.

### Rainfall erosivity

The dataset contains rainfall erosivity for European Union (28 member States) and Switzerland based on REDES database with high temporal resolution rainfall measurements of 26,394 years. Gaussian Process Regression (GPR) model was used to interpolate the rainfall erosivity values of single stations and to generate the R-factor map. REDES is provided as a point database including R-factor (Rainfall erosivity factor) for each of the 1,675 stations. Monthly R-factor maps, R-factor detailed assessments for Greece and Switzerland and Future projections *(2050)* of R-factor are available as additional layers by request following the same procedure detailed for the whole Rainfall erosivity dataset.

### Copper distribution in topsoil

The layer contains copper distribution in the soils of 25 European Union (EU) Member States. Generalized Linear Models (GLM) were used to investigate the factors driving copper distribution in EU soils. Regression analysis shows the importance of topsoil properties, land cover and climate in estimating Cu concentration. Meanwhile, a copper regression model confirms our hypothesis that

different agricultural management practices have a relevant influence on Cu concentration. Besides the traditional use of copper as a fungicide for treatments in several permanent crops, the combined effect of soil properties such as high pH, soil organic carbon and clay, with humid and wet climatic conditions favours copper accumulation in soils of vineyards and tree crops.

### Heavy metals (HM) in topsoil

The dataset contains detailed maps of heavy metals based on topsoil HM data from LUCAS 2009. Maps are available for As, Cd, Cr, Cu, Hg, Pb, Mn, Sb, Co and Ni. Especially for copper, it is suggested to use this layer alongside the “Copper distribution in topsoil”, which provides a more complete and updated dataset.
The concentration maps were modeled by applying Regression Kriging and using the LUCAS dataset and several environmental (topography, geology, vegetation) and anthropic (human impact) covariates as input data, as described in Tóth *et* al. (2016). Risk areas were identified for each metal based on thresholds values proposed by the Finnish Ministry of Environment (MEF, 2007).

### Soil bulk density

The layer contains soil data for 3 depths (0-0.5 meters, 0.5-1.5 meters, and greater than 1.5 meters) for percent sand, clay, and rockiness for the entire globe at a half degree grid resolution. A single bulk density value and mineral depth are also provided for each gridcell. Bulk density values came from the International Geosphere Biosphere Programme . Mineral depth, clay, sand, and rockiness values came from the International Satellite Land Surface Climatology Project.

### References on Soil Products

Ballabio, C., Panagos, P., Lugato, E., Huang, J.-H., Orgiazzi, A., Jones, A., Montanarella, L. 2018. Copper distribution in European topsoils: An assessment based on LUCAS soil survey. *Science of The Total Environment,* 636, 282–298. http://doi.org/10.1016/j.scitotenv.2018.04.268.

Ballabio, C., Panagos, P., Monatanarella, L. 2016. Mapping topsoil physical properties at European scale using the LUCAS database. *Geoderma*, 261, 110–123. http://doi.org/10.1016/j.geoderma.2015.07.006.

Global Gridded Surfaces of Selected Soil Characteristics. 2005. International Geosphere Biosphere Programme. ftp://daac.ornl.gov/data/global_soil/IGBP-SurfaceProducts.

Johnson, D. L., Ambrose, S. H., Bassett, T. J., Bowen, M. L., Crummey, D. E., Isaacson, J. S., Winter-Nelson, A. E. 1997. Meanings of Environmental Terms. *Journal of Environmental Quality,* 26(3), 581–589. http://doi.org/10.2134/jeq1997.00472425002600030002x.

"Map for Europe of Natural Susceptibility of Soils to Compaction" 2008. European Commission - Joint Research Centre; available from ESDAC.jrc.ec.europa.eu.

Ministry of the Environment, Finland. 2007. Government Decree on the Assessment of Soil Contamination and Remediation Needs 214/2007.

Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., & Fernández-Ugalde, O. 2018. LUCAS Soil, the largest expandable soil dataset for Europe: a review. *European Journal of Soil Science*, 69(1), 140-153.

Panagos, P., Borrelli, P., Meusburger, K., Alewell, C., Lugato, E., Montanarella, L. 2015. Estimating the soil erosion cover-management factor at the European scale. Land Use Policy, 48, 38–50. http://doi.org/10.1016/j.landusepol.2015.05.021.

Panagos, P., Liedekerke, M. V., Jones, A., Montanarella, L. 2012. European Soil Data Centre: Response to European policy support and public data requirements. Land Use Policy, 29(2), 329–338. http://doi.org/10.1016/j.landusepol.2011.07.003.

Panagos, P., Borrelli, P., Poesen, J., Ballabio, C., Lugato, E., Meusburger, K., Alewell, C. 2015. The new assessment of soil loss by water erosion in Europe. *Environmental Science Policy*, 54, 438–447. http://doi.org/10.1016/j.envsci.2015.08.012.

# 5 Climate Change Indicators

## 5.1 Introduction

Climate change is an overarching topic with deep ramifications in all Geosciences and significant consequences for all aspects of life on Earth. Therefore, for the last decades there have been many initiatives attempting to understand the causes and, as much as possible, predict and mitigate the effects of climate change in different domains. In order for a consistent view on climate change, quantification is required and thus various indicators have been developed and improved over time, following technological, as well as scientific advancements. There are several organizations at international, as well as European level, that are doing relevant work in the climate change field, including activities related to climate change indicators, such as research, data gathering, computations etc. Such organizations are European Environmental Agency, the European Space Agency or the European Centre for Medium-Range Weather Forecasts.

With regard to the Geo-harmonizer scope, the Climate Change Initiative of ESA is of considerable interest. CCI was launched in 2010 with the main overarching scope of using the vast Earth Observation archives of ESA, built together with its Member States, to contribute to the Essential Climate Variables (Wmo Global Climate 2003) - ECV - framework developed by the Global Climate Observing System - GCOS (Spence and Townshend 1995) . An ECV is a physical, chemical or biological variable or a group of linked variables that critically contributes to the characterisation of Earth's climate. GCOS presently defines 54 ECVs, the CCI currently addresses 21 ECVs (Bojinski 2011). The projects run in parallel includes periodic processing of EO data to apply the newest algorithms as well as development activities of new algorithms that consider the emerging data sources for each of the ECVs. Additionally, the CCI also maintains a portal with all data, documentation up-to-date, a toolbox to facilitate the data analysis and a visualisation tool.

For the Geo-harmonizer project, 4 climate change indicators have been considered:

1. Land surface temperature
2. Precipitations
3. Flood events
4. Fire disturbance

## 5.2 Land surface temperature

According to the Global Climate Observing System (GCOS), **land surface temperature** (LST) is "Air temperature at a known height above surface, with the height specified in the metadata (K)". The requirements established by the Global Climate Observing System for this Essential Climate Variable (ECV) are: frequency: Hourly; Daily Tx/Tn; resolution: on site; required measurement uncertainty: 0.1K.

Table 5.1 lists the main land surface temperature products available.

*Table 5.1 Land surface temperature products*

| ID | Name | Coverage | Maintainer/Provider |
|---|---|---|---|
| 1 | Land Surface Temperature from MODIS LST | Global | Metz M., Andreo V., Neteler M. |
| 2 | LST: hourly LST from instantaneous observations - CGLOPS-1 LST | Global | Copernicus Global Land Services |
| 3 | LST10-DC: 10-day Land Surface Temperature with Daily Cycle - LST10-DC | Global | Copernicus Global Land Services |
| 5 | Sentinel 3 SLSTR - Level 2 LST | Global | ESA |

*Table 5.2 The main characteristics of the LST products.*

| Product / Characteristics | Land Surface Temperature from MODIS LST | CGLOPS-1 LST | LST10-DC | Sentinel 3 SLSTR - Level 2 LST |
|---|---|---|---|---|
| Spatial coverage | Global | Global | Global | Global |
| Spatial resolution / MMU | 3 arc-min | 5 km | 5 km | 1 km |
| Temporal coverage | 2003 - 2016 | Oct 2010 - present | 2017 - present | May 2016 - present |
| Temporal resolution | Monthly | Hourly | 10 days period | Daily |
| Data Format | GeoTiff | NetCDF 4 | NetCDF 4 | NetCDF 4 |
| Product license | Open Data Commons Open Database License (ODbL) v1.0 | N/A | N/A | Open data |
| Product access | Open | Open after registration | Open after registration | open |
| Citation | Metz *et* al. (2014) | Martins *et* al. (2019) | N/A | N/A |

## 5.2.1 Product description

### Copernicus Global Land Service LST products (MSG)

The global component of Land Service of Copernicus, the Earth Observation programme of the European Commission is providing the community with 2 LST Geo-harmonizer relevant products: Hourly Land Surface Temperature and LST10-DC 10-day Land Surface Temperature with Daily Cycle.

The research leading to the current version of the product has received funding from various European Commission Research and Technical Development programs. This product is based on MTSAT/HIMAWARI and GOES data, respectively owned by JMA and NOAA, and combined with the LST product from MSG under the EUMETSAT copyright, provided by LSA-SAF (Kogan 2001) .

The Copernicus Land Service LST is obtained from a constellation of geostationary (GEO) satellites, obtaining a final processed and fused global product (Freitas *et* al. 2013). For Europe, the (MSG) is used. LST is estimated from Top-of-Atmosphere (TOA) brightness temperatures of atmospheric window channels within the infrared range. More details regarding the algorithms used are available in the Product User Manual.

### A/ LST: hourly LST from instantaneous observations

CGLOPS-1 LST is produced every hour in Near Real Time after less than 3 hours of observation, consisting of instantaneous fields estimated every hour (00, 01, 02, …, 23 UTC). Its validation consists of a spatial consistency analysis at global scale, comparison with similar products from different sensors, and direct comparison with ground measurements. All results are available in the Validation Report. The accuracy assessments are done every year, using the most recent ground reference data available.

### B/ LST10-DC: 10-day Land Surface Temperature with Daily Cycle

The main input for this product is the hourly LST global product from the Copernicus Global Land Service - CGLOPS-1 LST. The 10-day LST Daily Cycle (LST10-DC) describes the LST daily cycle over each 10-day period, including minimum, maximum and median temperature values over the 10-days, for each hour. The product is generated at the end of each 10-day period - days 1, 11 and 21 of each month, with timeliness up to 3 days. Each of the main layers is a multidimensional dataset with 24 sub-layers, one for each hour of the day.

Additionally, the Thermal Condition Index (TCI) is also provided for each 10-day period. The TCI is obtained by comparing the median of the LST values observed during the compositing period around the local solar noon with the upper and lower values of maximum LST from a multi-year climatology. These extreme values are based on a long-term series 1985-2015, specifically built for this purpose. More details are provided in the Product User Manual.

### Land Surface Temperature from MODIS LST

The product represents a gap-free time series of gridded daily surface temperature from the MODIS collection 6 LST products (Metz, Andreo, and Neteler 2017). To eliminate the data gaps in the cloud covered areas, the authors developed a new method to reconstruct the daily global MODIS LST products MOD11C1 and MYD11C1 by combining temporal and spatial interpolation, using emissivity and elevation as covariates for the spatial interpolation. The data is monthly aggregated, with 3 main types of files for each month in each year:

- avg = average of daily averages
- min = minimum of daily minima
- max = maximum of daily maxima

### Sentinel 3 Sea and Land Surface Temperature Radiometer Level-2 Land Surface Temperature product

The Sea and Land Surface Temperature Radiometer (SLSTR) is a dual scan temperature radiometer, which has been selected for the low Earth orbit (800 - 830 km altitude) ESA Sentinel-3 operational mission as a part of the Copernicus. The principal objective of the SLSTR products is to provide global and regional Sea and Land Surface Temperature (SST, LST) to a very high level of accuracy (1K for

LST). The instrument was designed as a data continuation from previous instruments: ATSR on ERS-1, ATSR-2 on ERS-2 and AATSR on Envisat.

The SLSTR Level-2 LST product contains measurement file with Land Surface Temperature values (in Kelvin), as well as an Annotation Dataset that contains the associated parameters: Normalised Difference Vegetation Index (NDVI); GlobCover surface classification code (noted biome); fractional vegetation cover; total column water vapour. For each gridded pixel, the LST values as well as their estimated total uncertainties are given.

### Land surface Temperature Products under the umbrella of European Space Agency - Climate Change Initiative

The Climate Change Initiative (CCI) LST is building upon the experience of the ESA DUE GlobTemperature project, developed during 2013 - 2017, in which a series of LST products have been obtained and published. All datasets are available after registration.

*Table 5.3 GlobTemperature LST products*

| Available datasets | Period | Maintainer |
|---|---|---|
| LST from Advanced Along Track Scanning Radiometer | May 20, 2002 - Apr 8, 2012 | University of Leicester |
| Land Surface Temperature from Spinning Enhanced Visible and Infrared Imager | Jan 1, 2007 - Dec 31, 2016 | Portuguese Institute for Sea and Atmosphere |
| LST from Special Sensor Microwave/Imager Radiometer | Jan 1, 2003 - Dec 31, 2003 | Estellus |
| LST from Along Track Scanning Radiometer 2 | Jun 1, 1995 - Jun 22, 2003 | University of Leicester |
| LST from Multi-instrument | Jan 1, 2011 - Dec 31, 2013 | Portuguese Institute for Sea and Atmosphere and University of Leicester |

The ATSR GlobTemperature Level-2/Level-3 v1.0 LST were generated through the ESA DUE GlobTemperature Project with the support of National Centre for Earth Observation (NCEO) of the Natural Environment Research Council (Ghent, D. 2012).

The hourly LST data derived from SEVIRI/Meteosat is the LST product generated within the EUMETSAT Satellite Applications Facility on Land Surface Analysis (LSA SAF product LSA-001) reformatted and re-gridded within the GlobTemp project (Freitas *et* al. 2010), (Trigo *et* al. 2011).

The microwave LSTs have been produced with support from NASA, CNRS, and more recently by GlobTemperature (Aires *et* al. 2001),(Catherinot *et* al. 2011).

The ATSR GlobTemperature Level-2 v1.0 LST data were generated through the ESA DUE GlobTemperature Project with the support of National Centre for Earth Observation (NCEO) of the

Natural Environment Research Council (Ghent, D. 2012).

The Multi-instrument LST - Merged GEO+LEO - products are developed within the framework of the GlobTemperature Project with the support of the European Space Agency (ESA), LSA SAF, and National Centre for Earth Observation (NCEO) of the Natural Environment Research Council.

According to the Product Specification Document, published by the consortium in June 2020, the CCI LST initiative will produce a significant number of LST products using various sensors. The following LST products have been assumed by the CCI LST group.

*Table 5.4 LST products assumed by CCI LST group*

| Category of product and description | Satellite sensors and data to be used | Data Level produced (spatial resolution, temporal resolution) | Temporal range |
|---|---|---|---|
| LST ECV. Climate records formed from single satellite sensors. | ATSR series (ATSR-2, AATSR) (Envisat format) | L2P (1 km; orbit; One file per orbit per sensor (typically 14 orbits per day, one sensor operating 1995-2012 with overlap periods between sensors) | 01/08/1995 – 08/04/2012 |
| | | L3C (0.01°, 0.05°; daily, monthly, annual), data will be provided with day and night data (or descending and ascending orbits) in separate files. | |
| | MODIS series (Terra and Aqua) | L2P (1 km; orbit; one file per orbit granule per sensor (typically 288 granules per pay, two sensors operating) | 18/12/1999 - 31/12/2020 |
| | | L3C (0.01°, 0.05°; daily, monthly, annual), data will be provided with day and night data (or descending and ascending orbits) and separate files. | |
| | SEVIRI on MSG satellites 1-4 | L2P (5 km; disk; One file per scan one scan every 15 mins) | 01/01/2004 - 31/12/2020 |
| | | L3U (0.05°; disk; One file per scan one scan every 15 mins) | |
| | | L3C (0.05°; disk; monthly and annual for each sub-daily time stamp), data will be provided with day and night data (or descending and ascending orbits) and separate files.) | |
| | AVHRR on NOAA platforms 15—19 | L2P (1 km; orbit; One file per orbit per sensor (typically 14 orbits per day, one primary operational sensor operating at any one time) | 13/05/1998 - 31/12/2020 |
| | | L3C (0.01°, 0.05°; daily, monthly, annual), data will be provided with day and night data (or descending and ascending orbits) and separate files. | |
| | AVHRR on Metop platforms A to C | L2P (1 km; orbit; One file per orbit per sensor (typically 14 orbits per day, one primary operational sensor operating at any one time) | 01/01/2007 - 31/12/2020 |
| | | L3C (0.01°, 0.05°; daily, monthly, annual), data | |

| | | | |
|---|---|---|---|
| | | will be provided with day and night data (or descending and ascending orbits) and separate files. | |
| | SLSTR series (Sentinel 3A and Sentinel 3B) | L2P (1 km; orbit; 3 minute PDUs (granules)) | 16/02/2016 – 31/12/2020 |
| | | L3C (0.01°, 0.05°; daily, monthly, annual), data will be provided with day and night data (or descending and ascending orbits) and separate files. | |
| | IMAGER series on GOES satellites 12-16 | L2P (5 km; disk; One file per scan one scan every 30 mins.) | 01/01/2004 - 31/12/2020 |
| | | L3U (0.05°; disk; One file per scan one scan every 30 mins.) | |
| | | L3C (0.05°; monthly and annual for each sub-daily time stamp), data will be provided with day and night data (or descending and ascending orbits) and separate files. | |
| | JAMI series on the MTSAT-1 and 2 satellites | L2P (5 km; disk; One file per scan one scan every 30 mins) | 01/01/2009 – 31/12/2015 |
| | | L3U (0.05°; disk; One file per scan one scan every 30 mins) | |
| | | L3C (0.05°; monthly and annual for each sub-daily time stamp), data will be provided with day and night data (or descending and ascending orbits) and separate files. | |
| | SSM/I and SSMIS series (DMSP satellites F13, F17, F18) | L2P (25km; orbit; One file per orbit per sensor (typically 14 orbits per day, one sensor operating) | 01/01/1998 – 01/01/2008 |
| | | L3C (0.25°; daily, monthly, annual), data will be provided with day and night data (or descending and ascending orbits) and separate files. | |
| LST CDR. Climate records produced from combining data from different satellite sensors. | IR CDR. A merged IR CDR product produced from AATSR, SLSTR (Sentinel 3A and 3B), MODIS (Terra and Aqua), AVHRR (Metop A to C), SEVIRI (MSG satellites 1-4), IMAGER (GOES satellites 12-16) and JAMI (MTSAT1 and 2 satellites). | L3S (0.05°; daily, monthly, annual), data will be provided with day and night data (or descending and ascending orbits) and separate files. | 01/01/2009 - 31/12/2020 |
| | ATSR-SLSTR CDR. A merged IR CDR product produced from ATSR-2, | L3S (0.05°; daily, monthly, annual), data will be provided with day and night data (or descending and ascending orbits) and separate files. | 01/08/1995 - 31/12/2020 |

| | | | |
|---|---|---|---|
| | AATSR, SLSTR and TerraMODIS. | | |
| LST MGP. Experimental Merged products produced from combining data from different satellite sensors across the IR and MW. | An all-sky merged IR and MW product produced from AATSR, SLSTR (Sentinel 3A and 3B), MODIS (Terra and Aqua), AVHRR (Metop A to C), SEVIRI (MSG satellites 1- 4), IMAGER (GOES satellites 12-16) and JAMI (MTSAT-1 and 2 satellites), SSM/I and SSMIS (DMSP F11, F13 and F17). | L3S (0.25°; 3-hourly) | 01/01/2008 – 31/12/2008 |

More details regarding each foreseen product are given in the above mentioned document. However, there is no data of release indicated.

## 5.3 Precipitation

According to the Global Climate Observation System (GCOS), precipitation - liquid and solid - is one of the Essential Climate Variables (ECVs) with the most direct influence over mankind. Yet, it is still considered that even with all the national and international efforts, the observations are still not available at a sufficiently adequate density to define the distribution of precipitation in many parts of the globe, including the oceans and many land areas (Lahoz 2011)..

Precipitations are essential for several categories of domains, including agriculture, forestry and water resource management. Additionally, precipitation related extreme events - short-term such as flash floods or long-term such as tropical cyclones - have massive potential to cause life and economic loss.

Table 5.5 presents the minimum requirements as defined in the Global Climate Observation System (GCOS) Implementation Plan, published in 2016. The plan was developed to set the guidelines for implementation of a global observing system for climate, building on current actions and meeting the needs of the United Nations Framework Convention on Climate Change, adaptation planning and climate science.

*Table 5.5 Precipitation ECV requirements as set by the Global Climate Observation System*

| Product | Definition | Frequency | Spatial resolution | Requirement measurement uncertainty | Stability |
|---|---|---|---|---|---|
| Estimates of liquid and solid precipitation | Integration of solid and liquid precipitation rate reaching the ground over several time intervals. The | Monthly (resolving diurnal cycles and with | 25km | max(10% of daily totals; 0.1mm) | 5% of daily totals (regional |

| | | | | | |
|---|---|---|---|---|---|
| | reference requirement refers to integration over 24h. | statistics of 3 hr values) | | | scale) |

Table 5.6. presents a list of existing precipitations products.

*Table 5.6 List of existing precipitation products*

| ID | Name | Coverage | Maintainer/Provider |
|---|---|---|---|
| 1 | NOAA Climate Data Record (CDR) of Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN-CDR), Version 1 Revision 1 | 60 degrees S to 60 degrees N and 0 degrees to 360 degrees longitude | NOAA/NESDIS/NCEI |
| 2 | NOAA Climate Data Record (CDR) of Precipitation – GPCP Monthly, Version 2.3 | Global | NOAA/NESDIS/NCEI |
| 3 | Integrated Surface Database (ISD) of the National Centers for Environmental Information (NCEI) of the National Oceanic and Atmospheric Administration (NOAA) | Global | NOAA/NCEI |
| 4 | SM2RAIN-ASCAT | Global | National Research Council Italy - Research institute for geo-hydrological protection |
| 5 | SM2RAIN-CCI | Global | National Research Council Italy - Research institute for geo-hydrological protection |
| 6 | GPCC Full Data Daily Version.2018 at 1.0°: Daily Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data | Global | Global Precipitation Climatology Centre at Deutscher Wetterdienst |

Table 5.7 presents the main characteristics of the precipitations products listed in the table above.

*Table 5.7 Precipitation products' characteristics*

| Product / Characteristics | NOAA Climate Data Record (CDR) of Precipitation (PERSIANN-CDR), Ver 1 Rev 1 | NOAA Climate Data Record (CDR) of Precipitation – GPCP Monthly, Version 2.3 | Integrated Surface Database (ISD) of NCEI/NOAA | SM2RAIN-ASCAT | SM2RAIN-CCI | GPCC Full Data Daily Version.2018 at 1.0°: Daily Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data |
|---|---|---|---|---|---|---|
| Spatial coverage | 60 degrees S to 60 degrees N and 0 degrees to 360 degrees longitude | Global | Global | Global | Global | Global |
| Spatial resolution / MMU | 0.25 deg x 0.25 deg | 2.5 deg X 2.5 deg | in situ | 12.5 km | 0.25 deg x 0.25 deg | 1° × 1° |

| | | | | | | |
|---|---|---|---|---|---|---|
| Measurement units | mm | mm/day | mm | mm/day | mm/day | Daily |
| Temporal coverage | 1983 - present | 1979 - present | N/A | 2007-2020 | 1998 - 2015 | 1982 - 2016 |
| Temporal resolution | Daily | Monthly | Hourly, daily | | Daily | Daily |
| Data Format | NetCDF4 | NetCDF4 | ASCII | NetCDF, GeoTIFF | NetCDF | NetCDF |
| Product license | Open | Open | Open | Open | Open | Open |
| Product access | Free and unrestricted access. Registration is optional. | Free and unrestricted access. Registration is optional. | Free and unrestricted access. | Free and unrestricted access. | Free and unrestricted access. | Free and unrestricted access. |
| Citation | (Ashouri et al. 2015) | (Adler et al. 2016) | N/A | (Brocca et al. 2019) | (Ciabatta et al. 2018) | (Ziese et al. 2018) |

### 5.3.1 Product description

#### PERSIANN Precipitation Climate Data Record (PERSIANN-CDR)

The PERSIANN-CDR is a daily quasi-global precipitation product for the period of 1982 to the present, updated quarterly. The dataset is created using long wave infrared images from geosynchronous satellites, not the raw sensor data, but the infrared data gridded from GridSat-B1. The GridSat-B1 (Gridded Satellite) is the gridded derived product from ISCCP (International Satellite Cloud Climatology Project) B1 data. The ISCCP B1 data is comprised of all channel observations from a number of international GEO satellites including the Geostationary Operational Environmental Satellite (GOES) series, the European Meteorological satellite (Meteosat) series, the Japanese Geostationary Meteorological Satellite (GMS) series, and the Chinese Fen-yung 2C (FY2) series. The global ISCCP B1 IR brightness temperature from these GEO sources covers the time period from 1979 to present, at space and time resolutions of 10-km and 3-hour intervals. Better coverage began in 1983, albeit with a gap over the Indian Ocean due to the lack of GEO measurements. GridSat-B1 data provides data for three channels: visible data, infrared window (IRWIN) data, and infrared water vapor (IRWVP) data. The Infrared Window (IRWIN) data is the main input data for the PERSIANN model. The dataset doesn't cover the entire Geo-harmonizer region.

More information regarding the PERSIANN Precipitation CDR is available in the Climate Algorithm Theoretical Basis Document, published in 2014 (Ashouri et al. 2015).

The source code package for Precipitation - PERSIANN-CDR is also available for download.

#### NOAA Climate Data Record (CDR) of Precipitation

The Global Precipitation Climatology Project (GPCP) version 2.3 - GPCP V2.3 Monthly Analysis Product - is a global combined dataset for precipitations part of the Climate Data Record (CDR) Program of NOAA. The general approach of obtaining this dataset is to combine the precipitation information available from each of several satellite and in situ sources into a final merged product,

taking advantage of the strengths of each data type: passive microwave estimates are based on SSMI/SSMIS data; infrared precipitation estimates are included, using GOES (Geostationary Operational Environmental Satellites) data and POES (Polar Operational Environmental Satellite) data; as well as other low earth orbit data and in situ observations.

The *in situ* data is represented by the Global Precipitation Climatology Centre (GPCC) gauge analysis. Precipitation gauge reports are archived from a time-varying collection of over 70,000 stations around the globe, both from Global Telecommunications System (GTS) reports and from other world-wide or national data collections.

The source code package for the GPCP CDR is available for download.

### Integrated Surface Database (ISD) of the National Centers for Environmental Information (NCEI) of the National Oceanic and Atmospheric Administration (NOAA)

The Integrated Surface Database (ISD) consists of global hourly and synoptic observations compiled from global sources into a single common ASCII format and common data model. The initiative started in 1998 integrating data from over 100 original data sources, including various data formats that were manually entered from paper forms during the 1950s–1970s time frame. ISD contains over 35k stations worldwide and even though some data goes as far back as 1901, the substantial increase in volume happens in the 1940s and again in the early 1970s. Currently, there are over 14,000 "active" stations updated daily in the database (Smith, Lott, and Vose 2011). Although ISD has geographic gaps in data, the Geo-harmonizer region is not affected, Europe, alongside North America, being one of the best represented regions.

ISD includes numerous parameters such as wind speed and direction, wind gust, temperature, dew point, cloud data, sea level pressure, altimeter setting, station pressure, present weather, visibility, precipitation amounts for various time periods, snow depth, and various other elements as observed by each station.

The ISD goes through a well documented quality check, that is fully described in The Quality Control of the Integrated Surface Hourly Database (Lott 2004).

The 4 derived ISD products are: (1) Global Hourly, (2) Local Climatological Data, (3) ISDLite (this is a sub-product of the ISD, prepared so it is easier to work with it; it contains eight common surface parameters in a fixed-width format free of duplicate values, sub-hourly data, and complicated flags), (4) Global Summary of the Day.

### SM2RAIN-ASCAT

The SM2RAIN-ASCAT is a global daily satellite rainfall from ASCAT soil moisture for precipitation product for rainfall data exclusively over land, obtained applying the SM2RAIN algorithm to ASCAT soil moisture observations (Luca Brocca *et* al. 2019). The known limitations identified by the authors consist of the underestimation of peak rainfall events and the presence of spurious rainfall events due to high-frequency soil moisture fluctuations that might be corrected in the future with more advanced bias correction techniques.

The algorithm, in Python, R and Matlab is freely available for download. The description of the method and of its performance can be found in (L. Brocca *et* al. 2013, Luca Brocca *et* al. 2014).

### SM2RAIN-CCI global daily rainfall dataset

The SM2RAIN-CCI (1 Jan 1998 – 31 December 2015) is a product obtained within the framework of ESA Climate Change Initiative and eartH2Observe projects and it is for rainfall data, specifically for land. It has been obtained by applying the SM2RAIN algorithm to the ESA CCI soil moisture Active and Passive products at version 03.1 separately (Luca Brocca *et* al. 2014). The algorithm has been calibrated during three different periods (1998-2001, 2002-2006 and 2007-2013) against the Global Precipitation Climatology Centre Full-Data daily dataset (Schamm et al. 2015). The quality flag provided within the raw soil moisture observations has been used to mask out low quality data, as well as the areas characterized by high topographic complexity, high frozen soil and snow probability and presence of tropical forests.

### GPCC Full Data Daily Version.2018 at 1.0°: Daily Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data

The Full Data Daily Product V.2018 (V.2) is based on near- as well as non-real time data specifically for land. It was extended backward in time to cover the period from 1982 to 2016 and is based on up to more than 35,000 stations per month. Anomalies of the daily precipitation relative to the monthly total at that station were interpolated. Thus, this analysis uses only the stations that also feature a monthly precipitation total.

More information regarding the GCPP products is available here.

## 5.4 Flood events

The flood is the most frequent natural disaster and it can be caused by heavy rainfall, rapid snowmelt or a storm surge from a tropical cyclone or tsunami in coastal areas and it can simplistically be defined as an overflow of water that is not usually covered by it, a dray area. Although it is not an ECV it is directly influenced by precipitation.

There are 3 common types of floods:

- Flash floods - caused by rapid and excessive rainfall that raises water heights quickly, and rivers, streams, channels or roads may be overtaken;
- River floods - caused when consistent rain or snow melt forces a river to exceed capacity;
- Coastal floods - caused by storm surges associated with tropical cyclones and tsunami.

Table 5.8 presents a list of relevant existing flood events products.

*Table 5.8 List of relevant flood events products*

| ID | Name | Coverage | Maintainer/Provider |
|---|---|---|---|
| 1 | A European Flood Database | Europe | Institute of Hydraulic Engineering, TU Wien |
| 2 | European Past Floods | Europe | European Environmental Agency |
| 3 | EEA potential flood-prone area extent | Europe | European Environmental Agency |
| 4 | Global Active Archive of Large Flood Events | Global | Dartmouth Flood Observatory |
| 5 | Copernicus Emergency Management Service - Mapping | Global | European Commission |

Table 5.9 presents the main characteristics of the flood events products listed in the table above.

*Table 5.9 List of the characteristics of the flood event products listed above*

| Product / Characteristics | A European Flood Database | European Past Floods | EEA potential flood-prone area extent | Global Active Archive of Large Flood Events | Copernicus EMS - Mapping |
|---|---|---|---|---|---|
| Spatial coverage | Europe | Europe | EEA39 | Global | Europe |
| Spatial resolution / MMU | N/A | N/A | N/A | N/A | <10m |
| Measurement units | N/A | N/A | | N/A | ha |
| Temporal coverage | N/A - 2015 | 1980 - 2015 | 2011 - 2016 | 1985 - present | 2012 - present |
| Temporal resolution | N/A - 2015 | N/A | N/A | N/A | N/A |
| Data Format | N/A | csv, Microsoft Access db zip | Personal Geodatabase | Personal Geodatabase, WMS | geospatial PDF, GeoJPEG, ESRI Shapefiles, Google Earth KML, GeoJSON |
| Product license | closed | open | open | open | open |
| Product access | N/A | Free and unrestricted access. | Free and unrestricted access. | Free and unrestricted access. | direct download from website |
| Citation | Hall et al. 2015 | N/A | N/A | (G.R.Brakenridge, "Global Active Archive of Large Flood Events", Dartmouth Flood Observatory) | "Copernicus Emergency Management Service." Directorate Space, Security and Migration, European Commission Joint Research Centre (EC JRC) |

## 5.4.1 Flood Products and Databases Description

### The European Flood Database

The European Flood Database represents an initiative under the project ERC Advanced Grant "FloodChange", Project No. 291152. The database contains hydrological data of annual maximum and daily mean discharge series, from over 7000 hydrometric stations of various data series lengths, from over 50 different data sources (Hall et al. 2015).

Due to different data exchange protocols, the database is available only to the partners in the consortium.

### EEA European Past Floods

The dataset has been assessed and reported by the ETC on Inland, Coastal and Marine Waters (ETC/ICM) and EEA. It contains information on past floods in Europe based on the reporting of EU Member States for the EU Floods Directive (2007/60/EC) and combined with information provided by relevant national authorities and global databases on natural hazards. The geospatial information is also stored in a csv table with specific codes, such as flood location code, which is a unique code for the flood location that can also be used as a identifier for multiple surface water bodies designated under the WFD which the flood location is represented by (as reported under FD). It can be linked with corresponding PFRA spatial data (points, polylines or polygons).

The datasets are structured with respect to data format, however, the quality of the data can vary.

### EEA potential flood-prone area extent

The EEA potential flood prone area extent delineates the area that is flooded once every 100 years, i.e., the probability of flooding is 1% assuming that flooding is unrestricted. The potential flood prone area encapsulates the river channel and floodplain. In reality, the floodplain is split into an active floodplain where flooding still occurs, and a former floodplain where flooding is restricted due to flood protection. The former floodplain could be flooded again either if a flood exceeds the capacity of flood protection, or if factors that control the presence of water were removed. These factors include channel and floodplain morphological alterations as well as structural flood protection measures. The dataset is vectorial.

### Global Active Archive of Large Flood Events

The dataset represents an archive derived from a wide variety of news and governmental sources. The quality and quantity of information available about a particular flood is not always in proportion to its actual magnitude, and the intensity of news coverage varies from nation to nation. In general, news from floods in low-tech countries tend to arrive later and be less detailed than information from 'first world' countries. Each flood event has an associated GIS polygon representing the area affected by the flooding in that event. The same news and governmental sources that are used to complete the entries in the tables are used to determine an approximate geographic area that is affected by the flood event; however this is not the actual areas of inundation. When DFO has obtained satellite data and produced a Rapid Response Inundation Map for an event a link to that map is provided in the Country column of the Register of Major Flood Events.

## 5.5 Fire Disturbance Products

With respect to fire, there are three types of products considered outlined by the Global Climate Observation System: Burned Area, Active Fire Map, and Fire Radiative Power (FRP). The following table presents the requirements for each, as given by the GCOS.

*Table 5.10 List of the products defined by the Global Climate Observation System in the Implementation Plan 2016*

| Product | Definition | Frequency | Spatial resolution | Requirement measurement uncertainty |
|---|---|---|---|---|
| Burned Area | Area affected by fire, irrespective of land cover, it can include natural vegetation or croplands. | 24 h | 30 m | 15% (error of omission and commission), compared to 30 m observations |

| | | | | |
|---|---|---|---|---|
| Active Fire Map | Presence of a temporal thermal anomaly within a grid cell. Generally, the active fire maps are defined by the date/hour when the thermal anomaly was detected. | 6 hours at all latitudes from Polar-Orbiting and 1 hour from Geostationary | 0.25-1 km (Polar); 1-3 km (Geo) | 5% error of commission; 10% error of omission; Based on per-fire comparisons for fires above target threshold of 5 MW/km² equivalent integrated FRP per pixel (i.e. for a 0.5 km² pixel the target threshold would be 2.5 MW, for a 9 km² pixel it would be 45 MW). |
| Fire Radiative Power | Amount of energy released by area unit. Commonly it is expressed in W/m2 . This variable is a function of actual temperature of the active fire at the satellite overpass and the proportion of the grid cell being burned. | 6 hours at all latitudes from Polar-Orbiting and 1 hour from Geostationary | 0.25-1 km (Polar); 1-3 km (Geo) | 10% integrated over pixel. Based on target detection threshold of 5 MW/km² equivalent integrated FRP per pixel (i.e. for a 0.5 km² pixel the target threshold would be 2.5 MW, for a 9 km² pixel it would be 45 MW).and with the same detection accuracy as the Active Fire Maps. |

Fire disturbance is highly heterogeneous, in space and time. Fires can be diurnally, seasonally and even inter-annually and also have a patchy structure. That is why for BA the preferred spatial requirements is 30 m, the minimum accepted is 500 m.

There are several products available, summarised in Table 5.11.

*Table 5.11 List of fire disturbance related products*

| ID | Name | Coverage | Maintainer/Provider |
|---|---|---|---|
| 1 | MODIS Fire_cci Burned Area Pixel (/GRID) product, version 5.1 | Global | European Space Agency - Climate Change Initiative |
| 2 | LTDR Fire_cci version 1.0 | Global | European Space Agency - Climate Change Initiative |
| 3 | Sentinel-1 Burned Area product - FireCCIS1SA10 | South America - Amazon forest | European Space Agency - Climate Change Initiative |
| 4 | MODIS Active fires | pan-European | NASA - Fire Information for Resource Management System |
| 5 | VIIRS Active fires | pan-European | NASA - Fire Information for Resource Management System |
| 7 | Country totals | pan-European without Belarus, Ukraine, Rep of Moldova | JRC - European Forest Fire Information System |
| 8 | Copernicus Emergency Management Service - Mapping | Europe | European Commission |

Table 5.12 presents the main characteristics of the Fire Disturbance products:

*Table 5.12 List of relevant product for the fire disturbance*

| Product / Characteristics | MODIS Fire_cci Burned Area Pixel (/GRID) product, version 5.1 | LTDR Fire_cci version 1.0 | Sentinel-1 Burned Area product - FireCCIS1SA10 | MODIS Active fires | VIIRS Active fires | Country totals | Copernicus Emergency Management Service - Mapping |
|---|---|---|---|---|---|---|---|
| Spatial coverage | Global | Global | South America - Amazon forest | Global | Global | pan-European without Belarus, Ukraine, Rep of Moldova | Europe |
| Spatial resolution / MMU | 250 m | 0.25 x 0.25 degrees | 40 m | 1 km | 375 m | per country | ha |
| Temporal coverage | 2001-2019 | 1982-2017 | 2017 | 2000 - ongoing | 2012-ongoing | varies depending on the country | 2012 - present |
| Temporal resolution | Monthly | monthly | monthly | 1, 7, 30 days, fire season aggregation | 1, 7, 30 days, fire season aggregation | yearly | na |
| Data Format | GeoTIFF | NetCDF | GeoTIFF | shapefile, KML, txt, WMS, WMS-T | shapefile, KML, txt, WMS, WMS-T | txt | geospatial PDF, GeoJPEG, ESRI Shapefiles, Google Earth KML, GeoJSON |
| Product license | Open data | Open Data | Open Data | Open Data | Open Data | open data | open |
| Product access | Open | open after registration | open after registration | through written data request | through written data request | direct download from website | direct download from website |
| Citation | Pettinari et al. 2020 | (Otón et al. 2020) | (Chuvieco et al. 2019) | This data set was provided by LANCE FIRMS operated by NASA ESDIS with funding provided by NASA | (Schroeder et al. 2014) | N/A | Copernicus Emergency Management Service |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | Headquarters. 10.5067/FIRMS/MODIS/MCD14DL.NRT.006 | | | |

## 5.5.1 Product description

### MODIS Fire_cci Burned Area Pixel (/GRID) product, version 5.1

The MODIS Fire_cci version 5.1 products (FireCCI51) comprise maps of global burned area developed based on the MOD09GQ Collection 6 images, acquired by the Terra satellite. The BA product files consist of 3 layers containing the date of detection, the confidence level and the land cover in the pixel detected as burned. Table 5.13 details each layer.

*Table 5.13 List of what each layer of the fire disturbance product describes*

| Layer name | Physical values | Unit | Potential values |
|---|---|---|---|
| Date of the first detection (JD) | [1, 365/366] | Day of the year | • 0 (zero): when the pixel is not burned.<br>1 to 366: day of the first detection when the pixel is burned.<br>-1: when the pixel is not observed in the month.<br>-2: used for pixels that are not burnable: water bodies, bare areas, urban areas, permanent snow and ice. |
| Confidence level (CL) | [0, 100] | % | - 0 (zero): when the pixel is not observed in the month, or it is not burnable (not vegetated).<br>- 1 to 100: Probability values. The closer to 100, the higher the confidence that the pixel is actually burned. This value expresses the uncertainty of the detection for all pixels, even if they are classified as unburned. |
| Land cover of burned pixels (LC) | [0/ 180] | N/A | 0 (zero): when the pixel is not burned in the month, either because it was observed and not classified as burned, or because it is non burnable or was not observed.<br>10 to 180: land cover code when the pixel is burned. Land cover of the pixel detected as burned, extracted from the Land Cover CCI maps. |

When the pixel is characterized as burned, it is assumed that the complete pixel was burned. The date of the burned pixel may not be coincident with the actual burning date, but most probably taken from one to several days afterwards, depending on image availability and cloud coverage. For areas with low cloud coverage, the detected date of burn should be very close to the actual date of burn, while for areas with high cloud coverage the date may be from several days or even weeks after the fire is over. For each pixel, it is assumed that there is only one land cover category, thus errors in the land cover product are reflected in the burnt area product as well.

More information regarding the FireCCI is available in the Product User Guide.

### LTDR Fire_cci v1.0 product - FireCCILT10

The FireCCILT10 is based on the Land Long-Term Data Record product based on images acquired by the Advanced Very High Resolution Radiometer (AVHRR 2-3) sensor on board the National Oceanic and Atmospheric Administration (NOAA 7-19) satellites.

Table 5.14 presents the layers contained by the grid BA product:

*Table 5.14 List of what each layer of the grid BA product describes*

| Layer name | Physical values | Unit | Potential values |
|---|---|---|---|
| Burned_area | [0,n] | Square meters | Sum of burned area of all pixels detected as burned within each grid cell. |
| Standard Error | [0,n] | Square meters | This value is the standard error of the estimation of BA in each grid cell. |
| Fraction of burnable area | [0,100] | | The fraction of area in the grid that corresponds to vegetated land covers that could be affected by fire. |
| Fraction of observed area | [0,100] | | The fraction of area in the grid that was observed for the whole monthly period. |
| burned_area_in_vegetation_class | [6, 23] | Square meters | Sum of all burned pixels of each land cover as defined by the LC_cci version 2.0.7.. |

Due to the large LTDR pixels, an additional step is made to calculate how much of each pixel was burned. As a result, the LTDR pixels can be partially or totally burned. The algorithm is described in the Algorithm Theoretical Basis Document.

For the value of burned area in vegetation class, it is assumed that each burned pixel that adds to the total burned area in a grid cell corresponds to only one land cover, as in most land cover maps. Therefore, the BA product depends heavily on the land cover product.

More information regarding the FireCCILT10 is available in the Product User Guide.

### Sentinel-1 Burned Area product - FireCCIS1SA10

The product is a demonstrator to test the suitability of radar data for identifying burned areas, located in tropical South America for one year, 2017. The product was obtained employing an algorithm that uses Sentinel-1 SAR information, with MODIS hotspots and land cover as auxiliary information, to detect areas that have burned due to wildfires.

### MODIS Active fires

The MODIS active fire product detects fires in 1km pixels that are burning at the time of overpass under relatively cloud-free conditions using a contextual algorithm (Giglio *et* al. 2016; Kaufman *et* al.

1998) . The MODIS Active Fire Product User's Guide provides more detailed information about the MODIS active fire product suite.

### VIIRS Active fires

The (VIIRS) 375 m thermal anomalies / active fire product provides data from the VIIRS sensor aboard the joint NASA/NOAA Suomi National Polar-orbiting Partnership (Suomi NPP) and NOAA-20 satellites. The 375 m data complements Moderate Resolution Imaging Spectroradiometer (MODIS) fire detection; they both show good agreement in hotspot detection but the improved spatial resolution of the 375 m data provides a greater response over fires of relatively small areas and provides improved mapping of large fire perimeters. The 375 m data also has improved nighttime performance. Consequently, these data are well suited for use in support of fire management (e.g., near real-time alert systems), as well as other science applications requiring improved fire mapping fidelity.

More details on the algorithm used are described in (Schroeder *et* al. 2014).

### Country totals

The country totals refers strictly to forest fires. The data collection was done according to Regulation EEC No 804/94 (now expired) that established a Community system of information on forest fires for which a systematic collection of a minimum set of data on each fire occurring, the “Common Core”.

Since 2004 the forest fire data provided each year by individual EU Member States and other European countries are checked, stored and managed by JRC within EFFIS. At present the database contains fire data from 22 countries: Bulgaria, Croatia, Cyprus, Czech Republic, Estonia, Finland, France, Germany, Greece, Hungary, Italy, Latvia, Lebanon, Lithuania, Poland, Portugal, Romania, Slovakia, Spain, Sweden, Switzerland, and Turkey.

### Copernicus Emergency Management Service - Mapping

The general description of the service has been written in Chapter 5.4.

### References on Fire Disturbance Products


Adler, R., Jian-Jian W., Sapiano, M., Huffman, G., Long Chiu, Ping Ping X., Ferraro, R. 2016. Global Precipitation Climatology Project (GPCP) Climate Data Record (CDR), Version 2.3 (Monthly). NOAA National Centers for Environmental Information. https://doi.org/10.7289/V56971M6.

Aires, F., C. Prigent, W., Rossow, B., and M. Rothstein. 2001. A new neural network approach including first guess for retrieval of atmospheric water vapor, cloud liquid water path, surface temperature, and emissivities over land from satellite microwave observations. *J. Geophys. Res.*, 106 (D14), 14887–14907, doi:10.1029/2001JD900085.

Ashouri, H., Kuo-Lin H., Soroosh S., Braithwaite, D.K., Knapp, K.R., Cecil, L.D., Nelson, B.R., and Prat, O.P. 2015. “PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies.” *Bulletin of the American Meteorological Society* 96 (1), 69–83.

Brakenridge, G.R., "Global Active Archive of Large Flood Events", Dartmouth Flood Observatory, University of Colorado, http://floodobservatory.colorado.edu/Archives/index.html.

Bojinski, S. 2011. “Response by ESA to GCOS—Results of the Climate Change Initiative Requirements Analysis. ”*Eur. Space Agency*.

Brocca, L., Moramarco, T., Melone, F., and Wagner, W. 2013. “A New Method for Rainfall Estimation through Soil Moisture Observations.” *Geophysical Research Letters* 40 (5), 853–58.

Brocca, L., Ciabatta, L., Massari, Ch., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W., and Levizzani, V.. 2014. “Soil as a Natural Rain Gauge: Estimating Global Rainfall from Satellite Soil Moisture Data.” *Journal of Geophysical Research: Atmospheres*. https://doi.org/10.1002/2014jd021489.

Brocca, L., Filippucci, P., Hahn, S., Ciabatta, L., Massari, Ch., Camici, S., Schüller,L., Bojkov, B., and Wagner., W. 2019. “SM2RAIN–ASCAT (2007–2018): Global Daily Satellite Rainfall Data from ASCAT Soil Moisture Observations.” *Earth System Science Data*. https://doi.org/10.5194/essd-11-1583-2019.

Catherinot, J., C. Prigent, R. Maurer, F. Papa, C. Jiménez, F. Aires, and W. B. Rossow (2011), Evaluation of “all weather” microwave-derived land surface temperatures with in situ CEOP measurements. *J. Geophys. Res.*, 116 D23105, doi:10.1029/2011JD016439.

Chuvieco, E., Pettinari, M.L., Tanse, M., Belenguer Plomer, M. Á., and Lizundia-Loiola, J. 2019. “FireCCIS1SA10: ESA Fire Climate Change Initiative (Fire_cci) - Sentinel-1 Burned Area Product for a Large Demonstrator Area in South America.” e-cienciaDatos. https://doi.org/10.21950/VTDZ1L.

Ciabatta, L., Massari, Ch., Brocca, L., Gruber, A., Reimer, Ch., Hahn, S., Paulik, Ch., Dorigo, W. , Kidd, R., and Wagner, W.. 2018. *SM2RAIN-CCI (1 Jan 1998 – 31 December 2015) Global Daily Rainfall Dataset*. https://doi.org/10.5281/zenodo.1305021.

Csiszar, I. A., Morisette,J.T., and Giglio, L. 2006. “Validation of Active Fire Detection from Moderate-Resolution Satellite Sensors: The MODIS Example in Northern Eurasia.” *IEEE Transactions on Geoscience and Remote Sensing*. https://doi.org/10.1109/tgrs.2006.875941.

Freitas, S. C., Trigo, I.F., Macedo, J., Barroso, C., Silva, R., and Perdigão, R. 2013. Land Surface Temperature from Multiple Geostationary Satellites. *International Journal of Remote Sensing*. https://doi.org/10.1080/01431161.2012.716925.

Freitas, S. C., Trigo, I. F., Bioucas-Dias, J. M., Goettsche, F.-M., 2010: Quantifying the Uncertainty of Land Surface Temperature Retrievals From SEVIRI/Meteosat, *IEEE Trans. Geosci. Remote Sens.* DOI: 10.1109/TGRS.2009.2027697.

Ghent, D. 20212. Land Surface Temperature Validation and Algorithm Verification (Report to European Space Agency).

Giglio, L., Schroeder, W. and Justice, Ch.O. 2016. The Collection 6 MODIS Active Fire Detection Algorithm and Fire Products. *Remote Sensing of the Environment* 178 (June): 31–41.

Hall, J., Arheimer, B., Aronica, G.T., Bilibashi, A., Boháč, M., Bonacci, O., Borga, M. 2015. A European Flood Database: Facilitating Comprehensive Flood Research beyond Administrative Boundaries. *Proceedings of the International Association of Hydrological Sciences* 370: 89–95.

Kaufman, Y. J., Justice, Ch.O., Flynn, L.P., Kendall, J.D., Prins E.M., Giglio, L., Ward, D.E., Menzel, W.P., and Setzer, A.W. 1998. Potential Global Fire Monitoring from EOS-MODIS. *Journal of Geophysical Research: Atmospheres*. https://doi.org/10.1029/98jd01644.

Kogan, F. N. 2001. Operational Space Technology for Global Vegetation Assessment. *Bulletin of the American Meteorological Society*. https://doi.org/2.3.co;2">10.1175/1520-0477(2001)082<1949:ostfgv>2.3.co;2.

Lahoz, W. 2011. Systematic Observation Requirements for Satellite-Based Products for Climate, 2011 Update, Supplemental Details to the Satellite-Based Component of the Implementation Plan for the Global Observing

System for Climate in Support of the UNFCCC (2010 Update). *World Meteorological Organization (WMO): Geneva, Switzerland*.

Lott, J. N. 2004. The Quality Control of the Integrated Surface Hourly Database. 14th Conference on Applied Climatology. *Am. Meteorol. Soc. , Seattle, Wash*, 10–16.

Metz, M., Andreo, V.„ and Neteler, M. 2017. A New Fully Gap-Free Time Series of Land Surface Temperature from MODIS LST Data. *Remote Sensing* 9 (12): 1333.

Otón, G. Pettinari, M.L., Chuvieco, E. 2020. ESA CCI ECV Fire Disturbance: D3.3.4 Product User Guide - LTDR, version 1.1. Available at http://www.esa-firecci.org/documents.

Pettinari, M.L., Lizundia-Loiola, J., Chuvieco, E. 2020. ESA CCI ECV Fire Disturbance: D4.2 Product User Guide - MODIS, version 1.0. Available at: https://www.esa-fire-cci.org/documents

Rivero, A. W.. 2018. Product User Manual for Copernicus EMS Rapid Mapping. https://emergency.copernicus.eu/mapping/sites/default/files/files/CopernicusEMS_RapidMapping_ProductUserManual_v1.pdf.

Schamm, K., M. Ziese, K. Raykova, A. Becker, P. Finger, A. Meyer-Christoffer, and U. Schneider. 2015. GPCC Full Data Daily Version 1.0 at 1.0◦: Daily Land-Surface Precipitation from Rain-Gauges Built on GTS-Based and Historic Data. *Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory.* Available at https://doi. org/10. 5065/D6V69GRT (accessed on 1 March 2017).

Schroeder, W., Oliva, P., Giglio, L., and Csiszar, I.A. 2014. The New VIIRS 375m Active Fire Detection Data Product: Algorithm Description and Initial Assessment. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2013.12.008.

Smith, A., Lott, N., and Vose, R. 2011. The Integrated Surface Database: Recent Developments and Partnerships. *Bulletin of the American Meteorological Society* 92 (6): 704–8.

Spence, T., and Townshend, J. 1995. The Global Climate Observing System (GCOS). *Climatic Change*. https://doi.org/10.1007/bf01095141.

Trigo, I. F., DaCamara,C.C., Viterbo, P., Roujean, J.-L., Olesen, F., Barroso, C., Camacho-de Coca, F., Carrer, D., Freitas, S., García-Haro, C.J., Geiger, B., Gellens-Meulenberghs, F., Ghilain, N., Meliá, J., Pessanha, L., Siljamo, N., and Arboleda, A. 2011. The Satellite Application Facility on Land Surface Analysis. *Int. J. Remote Sens*., 32, 2725-2744, doi: 10.1080/01431161003743199.

WMO Global Climate. 2003. *The Second Report on the Adequacy of the Global Observing Systems for Climate in Support of the UNFCCC*.

# 6 Natural Vegetation Products

In the framework of current economic, social and environmental context, sustainable forest management (SFM) and forest policy need structured, precise and harmonized data. Comprehensive interpretations of humankind impact on forest land, how these alterations affected forest change trends and concerning issues such as acid deposition, biodiversity and forest sustainability have been recently investigated through reconstructions of historic vegetation maps and the study of Potential Natural Vegetation (PNV).

## 6.1 Historical vegetation products

Historical reconstructions of vegetation maps are available from a regional to a global scale: on a regional scale, they are derived from old topographic maps and land use data records; constraints such as the scarcity of historical data and the need for harmonization for the available sources, require instead the use of model-based approaches on a national, continental or global scale. As such, the products this way derived are subject to several degrees of uncertainty (Klein Goldevijk & Vergburg, 2013). Furthermore, existing products do not meet the requirements needed for greenhouse gas and climate assessment, due to insufficient spatial and thematic detail (Fuchs et al., 2013).

Harmonization and inclusion of historic statistics and topographic maps coming from different sources (national cartographic institutes, cadastres, independent communities) into model-based reconstruction have recently proved useful to improve the accuracy of the predictions.

*Table 6.1 List of publications about available for historical reconstruction of vegetation*

| ID | References | Dataset name | Spatial coverage | Spatial resolution | Temporal coverage | Thematic coverage |
|---|---|---|---|---|---|---|
| 1 | Hurtt et al., 2006 | Biomass density and wood harvest patterns | Global | 0.5 degree | 1700 - 2100 | Cropland<br>Primary land (forest)<br>Pasture<br>Secondary land<br>Urban land |
| 2 | Kaplan et al., 2009 | Forest cover estimates by region | Pan-European | 5 arc minutes | 1000 BC to 1850 AD | Forest |
| 3 | Fuchs et al., 2015 | Europe's land cover based on HILDA+ | European | 1 km | 1900 - 2010 | Forest<br>Settlement<br>Crops<br>Grassland<br>Other land<br>Water |

### 6.1.1 Product description

A brief description of each of the historical vegetation products is provided down below. The description is thematic-based, providing a detailed content of the dataset, technical specifications about the input data, the methodology used and application contexts. Full references are provided separately.

#### Biomass density and wood harvest pattern

Hurtt et al. focused on developing reconstructions of the spatial patterns of land conversions (land-use transitions), wood harvesting and resulting secondary lands annually, for the period 1700–2000, using data-based historical cases. They used land-use states (**HYDE** dataset), potential biomass density and recovery rates and national annual wood harvest statistics aggregated at a continental scale to model land-use transitions and secondary land area and age. Biomass density was derived from harvest statistics using the **Miami LU model** (Hurtt. et al., 2002). The final products show spatial patterns of forest cover change in time, both primary/undisturbed forests (Primary lands) and secondary forests (included into the 'Secondary lands' class, generated by abandonment of cropland or pasture).

#### Forest cover estimates by region

Kaplan et al. created annually resolved time series of anthropogenic deforestation in Europe over the past three millennia by digitizing and synthesizing a database of population history for Europe and surrounding areas, developing a model to simulate anthropogenic deforestation based on population density that handles technological progress, and applying the database and model to a gridded dataset of land suitability for agriculture and pasture to simulate spatial and temporal trends in anthropogenic deforestation.

#### Europe's land cover based on HILDA+

Fuchs et al. produced two land cover maps which show the difference of historic land cover in Europe between 1900 and 2010. Both maps are results of a model reconstruction performed by the HIstoric Land Dynamics Assessment model (HILDA, Version 2.0). **HILDA** combines multiple harmonized and consistent data streams of historic land cover information in one approach, like remote sensing products, national inventories, aerial photographs, land cover statistics, old encyclopaedias and historic land cover maps.The added value of using historical data was evaluated using historical maps by performing a reconstruction with and without the historic information.

## 6.2 Potential natural vegetation products

Potential natural vegetation refers instead to the hypothetical status of vegetation cover in equilibrium with climate, natural disturbances, lithology and hydrologic conditions at a particular location if not

impacted by human activities. PNV changes according to the long-term conditions above mentioned, so it is mandatory to reference the maps in a specific temporal context (GH context: 2000-2020).

Such as for historic vegetation maps, PNV maps find a limitation in their application in their coarse spatial resolution and the limited amount of data available for modeling: historical ground truth data on vegetation status, and specifically pristine vegetation, on a both small and large scale is lacking.

*Table 6.2 List of publications about potential natural vegetation maps*

| ID | References | Spatial coverage | Spatial resolution | Temporal coverage | Thematic coverage |
|---|---|---|---|---|---|
| 1 | Ramankutty and Foley, 1999 | Global | 5’ | 1700 - 1992 | Biomes, 15 classes |
| 2 | Pongratz *et* al., 2008 | Global | 5’ | 800 - 1992 | Biomes, 11 classes |
| 3 | Potapov *et* al., 2011 | Global | 1 km | 0 - 2000 | Forest, 3 classes |
| 4 | Levasseur *et* al. 2012 | Global | 10’ | 0 - 2000, mainly focused on 1960 - 2000 | Biomes, 10 classes |
| 5 | Lin *et* al., 2012 | Global | 1 km | 1950 - 2000 with projection until 2050 | Biomes, 8 classes |
| 6 | Hengl *et* al. 2018 | Global | 1 km | 2000 - 2015 | Forest, 67 species |

The European Commission too has recently shown interest in forests and their management, especially concern about the impact of climate change on them. That’s why in 2016, by combining data collected in the last century from regional or national forest atlases, geodatabases and scientific literature, the European Atlas of Forest Tree Species (San-Miguel-Ayanz *et* al. 2016) was published: an accurate assessment of forests, forests resources and their distribution is the first step to tackle the issue at hand. The atlas is the first available product for Europe that provides geo-referenced species distribution maps and probability distribution maps for 76 forest tree species.

### 6.2.1 Product description

A brief description of each of the PNV products is provided down below. The description is thematic-based, providing a detailed content of the dataset, technical specifications about the input data, the methodology used and application contexts.

On average, available products have very coarse resolution, generic (forest/non-forest, biome-level) thematic accuracy and global coverage: only the most recent one (Hengl *et* al., 2018)

provides a finer thematic level of detail (tree species-level). Full references are provided separately at the end of the Chapter 6.2.

Ramankutty and Foley prepared a global PNV map at 5 minutes spatial resolution. The main goal was to compare the PNV map with a croplands cover map to assess the consequences of cultivation. The product is derived from a reclassification and interpolation of the **DISCover** dataset (Loveland and Belward, 1997) with the **BIOME3** dataset (Haxeltine and Prentice, 1996). The dataset assigns as dominant vegetation type (DPV) the one observed from satellite where regions are not dominated by human land use, while in regions dominated by human land use vegetation types come from the BIOME3 dataset.

Limitations in the methodology and the product are the broad thematic accuracy (formation level) and the precision in distinguishing vegetation types from satellite (DISCover dataset in this case).

Pongratz *et* al. produced a new PNV map based on the previous work of Ramankutty and Foley and increased its temporal coverage by 900 years into the past. They included statistics on population in the process to propagate back in time both fractions of land cover usage, particularly croplands, and potential natural vegetation. It is the only global product with this level of thematic accuracy which covers a time span this long.

Potapov *et* al. produced a map of potential forests, representing an estimate of where forests would grow under current climate conditions and without human influence. The main source of data for defining potential forest extent is the terrestrial ecoregions of the world (Olson *et* al., 2001). Each ecoregion was classified as belonging to one of four categories: dense forests, open forests, woodlands, or non-forest, depending on its description (including current and potential vegetation) and its proportion of different forest types, with additional input from the following datasets: current forest extent; bioclimatic zoning and original forest cover extent (FAO, 1999; Bryant *et* al., 1997; Zomer *et* al., 2008); and a forest distribution map produced by modeling based on global climate variables and elevation.

Levasseur *et* al. modelled a high-resolution PNV distribution over Europe, entirely based on vegetation and climatological data. Their approach consists in using a multinomial logistic regression (MLR). MLR builds statistical relationships between vegetation data and climatological variables independently from models and without any subjective corrections. These relationships allow to model the occurrence probabilities of each PNV type. The 'vegetation–climate' relationship estimated by MLR from **BIOME 6000** modern data is implicitly constrained by an atmospheric $CO_2$ concentration of about 360 ppm. As a prospect, exporting this relationship in different climatic conditions leads to a distribution ignoring the crucial effect of $CO_2$ on vegetation. The PNV modelled by MLR cannot claim to be fully independent of human influences: MLR is mainly based on climatological data between 1961 and 1990, impacted by human activities through climate change. However, for the modern period, BIOME 6000 can be confidently considered as reference data collected in areas with less possible human activity, although ensuring data not impacted by humans is difficult.

Lin *et* al. evaluated three net primary productivity (NPP)–climate models, i.e. the Miami model, the Schuur model and the classification indices-based model, to understand the sensitivity of grassland to climate change and the effect of climate changes on the grassland ecosystems. To do su they used the integrated orderly classification system of grassland (**IOCSG**), which classifies grasslands according to climatic factors, and can therefore be used to predict climate-linked spatial

or temporal succession from an original class to a new class as these climatic factors change. The result comprises the spatial distribution of grassland biomes during the period of 1950–2000.

Hengl *et* al. provide the most up to date PNV maps with global coverage, with a total of 67 forest tree species maps derived by using a tree occurrence point dataset (**GBIF**, **EU Forest**) and 160 environmental covariates (DEMs, temperatures, precipitation, lithology) as input for different machine learning algorithms. Goal of the product was to improve the mapping accuracy of PNV products (1 km), compare the results with previous PNV products obtained with more traditional approaches and establish a fully reproducible protocol for PNV mapping based on open data. Spatial distributions of forest species were also compared with the ones produced for the European Atlas of Forest Tree Species.

### References on Natural Vegetation products


Bryant, D., Nielsen, D., Tangley, L. 1997. *T*he last frontier forests. *Issues in Science and Technology*, *14*(2), 85.

Fuchs, R., Herold, M., Verburg, P. H., Clevers, J. G. P. W. 2013. A high-resolution and harmonized model approach for reconstructing and analysing historic land changes in Europe. *Biogeosciences*, 10(3), 1543–1559. http://doi.org/10.5194/bg-10-1543-2013.

Fuchs, R., Verburg, P. H., Clevers, J. G., Herold, M. 2015. The potential of old maps and encyclopaedias for reconstructing historic European land cover/use change. *Applied Geography*, 59, 43–55. http://doi.org/10.1016/j.apgeog.2015.02.013.

Goldewijk, K. K., Verburg, P. H. 2013. Uncertainties in global-scale reconstructions of historical land use: an illustration using the HYDE data set. *Landscape Ecology*, 28(5), 861–877. http://doi.org/10.1007/s10980-013-9877-x.

Haxeltine, A., Prentice, I. C. 1996. BIOME3: An equilibrium terrestrial biosphere model based on ecophysiological constraints, resource availability, and competition among plant functional types. *Global Biogeochemical Cycles*, *10*(4), 693–709. http://doi.org/10.1029/96gb02344

Hengl, T., Walsh, M. G., Sanderman, J., Wheeler, I., Harrison, S. P., Prentice, I. C. 2018. Global mapping of potential natural vegetation: an assessment of machine learning algorithms for estimating land potential. *PeerJ*, 6. http://doi.org/10.7717/peerj.5457.

Hurtt, G. C., Pacala, S. W., Moorcroft, P. R. *et* al. 2002. Projecting the future of the US carbon sink. Proceedings of the National Academy of Sciences of the United States, 99, 1389–1394.

Hurtt, G. C., Frolking, S., Fearon, M. G., Moore, B., Shevliakova, E., Malyshev, S., Houghton, R. A. 2006. The underpinnings of land-use history: three centuries of global gridded land-use transitions, wood-harvest activity, and resulting secondary lands. *Global Change Biology*, *12*(7), 1208–1229. http://doi.org/10.1111/j.1365-2486.2006.01150.x.

Kaplan, J. O., Krumhardt, K. M., Zimmermann, N. 2009. The prehistoric and preindustrial deforestation of Europe. *Quaternary Science Reviews*, 28(27-28), 3016–3034. http://doi.org/10.1016/j.quascirev.2009.09.028.

Levavasseur, G., Vrac, M., Roche, D. M., Paillard, D. (2012). Statistical modelling of a new global potential vegetation distribution. *Environmental Research Letters*, 7(4), 044019. http://doi.org/10.1088/1748-9326/7/4/044019.

Lin, H., Feng, Q., Liang, T., Ren, J. 2012. Modelling global-scale potential grassland changes in spatio-temporal patterns to global climate change. International Journal of Sustainable Development; *World Ecology*, 20(1), 83–96. http://doi.org/10.1080/13504509.2012.749815.

Loveland, T., Belward, A. 1997. The International Geosphere Biosphere Programme Data and Information System global land cover data set (DISCover). *Acta Astronautica*, *41*(4-10), 681–689. http://doi.org/10.1016/s0094-5765(98)00050-2.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., Kassem, K. R. (2001). Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience*, *51*(11), 933. http://doi.org/10.1641/0006-3568(2001)051[0933:teotwa]2.0.co;2.

Pongratz, J., Reick, C., Raddatz, T., Claussen, M. 2008. A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles, 22*(3). http://doi.org/10.1029/2007gb003153.

Potapov, P., Laestadius, L., Minnemeyer, S. (2011). Global map of potential forest cover. World Resources Institute: Washington, DC. Online at www.wri.org/forest-restoration-atlas.

Ramankutty, N., Foley, J. A. (1999). Estimating historical changes in global land cover: Croplands from 1700 to 1992. Global *Biogeochemical Cycles, 13(*4), 997–1027. http://doi.org/10.1029/1999gb900046.

San-Miguel-Ayanz, J., De Rigo, D., Caudullo, G., Houston Durrant, T., Maur,i A. 2016. European atlas of forest tree species. Brussels: European Commission, Joint Research Centre.

Zomer, R. J., Trabucco, A., Verchot, L. V., Muys, B. 2007. Land Area Eligible for Afforestation and Reforestation within the Clean Development Mechanism: A Global Analysis of the Impact of Forest Definition. *Mitigation and Adaptation Strategies for Global Change*, *13*(3), 219–239. http://doi.org/10.1007/s11027-007-9087-4.