



D3.1 – Access to EO data

Project name

Asset Level Modelling of RISKS In the Face of Climate Induced Extreme Events and ADAPttation (RISKADAPT)

Horizon Innovation Actions | Project No. 101093939

HORIZON-MISS-2021-CLIMA-02-03



Funded by the
European Union

D3.1 – Access to EO data	
Dissemination level	Public
Type of deliverable	Report
Work package	WP3 – Climate Data, CC Forcing and Multi-Hazard Modelling
Status - version, date	V1.1, 06/04/2023
Deliverable leader	Marko Laine (FMI)
Contributing partners	
Contractual date of delivery	31/03/2023
Keywords	Earth observation, remote sensing, in-situ, data access, Copernicus

Quality Control

	Reviewer Name	Organisation	Date
Peer review 1	Mateja Škerjanec	ULFGG	05/04/2023
Peer review 2	Carlo Cintolesi	UNIBO	04/04/2023

Version History

Version	Date	Organisation	Summary of changes
0.1	10/02/2023		First draft of the document structure.
0.2	23/03/2023		First full version.
0.3	29/03/2023		Review comments.
1.0	29/03/2023		Text clean-up.
1.1	06/04/2023		Internal review.

Legal Disclaimer

Funded by the European Union's Horizon Europe Innovation Actions under grant agreement No. 101093939. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Commission. The information in this document is provided "as is", and no guarantee or warranty is given that it is fit for any specific purpose. The RISKADAPT project Consortium members shall have no liability for damages of any kind including without limitation direct, special, indirect, or consequential damages that may result from the use of these materials subject to any liability which is mandatory due to applicable law.

Table of Contents

Executive Summary	6
1. Introduction	7
1.1 Purpose of the deliverable	7
1.2 Structure of the deliverable	7
1.3 Intended audience	7
2. Copernicus Climate Data Store (CDS)	8
2.1 Introduction.....	8
2.2 ERA5	8
2.3 E-OBS	9
2.4 CMIP6	9
2.5 Accessing CDS data.....	9
3. Other data sources	9
4. Relation with other work packages/deliverables.....	10
5. Conclusions	12
References.....	13
Annexes.....	14

List of Figures

Figure 1: Plot produced by the example code. 15

List of Abbreviations and Acronyms

Abbreviation	Meaning
API	Application Programming Interface
CDS	Copernicus Climate Data Store
ECMWF	European Centre for Medium-Range Weather Forecasts
EO	Earth Observation
FMI	Finnish Meteorological Institute

Executive Summary

This deliverable report describes the Earth Observation (EO) data sources used in the RISKADAPT project. The most important data source is the Copernicus Climate Store. The data types include in-situ, remote sensed, and model-based data. Model-based data can be divided between historical reanalyses based on observations and climate model runs that simulate the climate system response to a scenario of future emissions. Local observational data from national data providers can be used to complement the Copernicus data.

1. Introduction

In this deliverable, the EO data refers to four different types of environmental data of the Earth: 1) remotely sensed data, which can be space-based satellite observations or from ground-based instruments; 2) in-situ data; 3) reanalysis data combining observations and models; and 4) climate model outputs. In reanalysis data, observational data are completed spatially and temporally using modelling. The advantage of reanalysis data is that it is complete and provided in a regular spatiotemporal grid. Thus, it makes it easier to study climatological changes in the phenomena of interest when compared with climate projections. The project will utilize open data sources, especially the Copernicus climate data.

1.1 Purpose of the deliverable

This deliverable will give an overview of the EO data directly available to the partners. As there has not yet been a comprehensive discussion within the consortium about the different data needs and data handling skills of the partners, the scope of this report will be very general. The deadline for this deliverable is month 3 of the project, so this report is a starting point for a discussion on data needs between project partners. In this respect, we can declare that the objectives related to this deliverable have been achieved in full and as scheduled.

1.2 Structure of the deliverable

The deliverable consists of this report which describes various ways to access Copernicus EO data. Section 2 describes Copernicus data sources that are relevant to the project, and Section 3 shortly comments on other sources of data. Section 4 explains the relation to other work in the project, and Section 5 has a conclusion. Annex 1 gives a short example of using Copernicus data, which is directly relevant to hydrological analyses in the work package WP3.

1.3 Intended audience

This report is meant primarily for other partners of the RISKADAPT project to understand the available data types that can be used in the analyses.

2. Copernicus Climate Data Store (CDS)

2.1 Introduction

Copernicus Climate Data Store (CDS) is part of Copernicus Climate Change Service (C3S, 2023). The Copernicus program provides a wide range of data sets. The C3S provides access to climate data sets, including global climate projections, climate indices, and climate reanalysis data, such as ERA5. These data sets can be used for a variety of applications, including climate impact assessments, climate adaptation planning, and risk management. The C3S also provides tools and services to help users process and analyse the data, making it easier for them to extract the information they need. The Copernicus Data and Information Access Services (DIAS) platform provides access to these data and offers various tools and applications for data processing and analysis. One of these tools is the Copernicus CDS Toolbox, which is a web-based software package for processing and analysing Copernicus data.

Copernicus CDS Toolbox supports a wide range of output data formats, including netCDF, HDF5, and GeoTIFF. It has a user-friendly interface that allows easy access to Copernicus data. A wide range of processing and analysis functions, such as data visualization, data filtering, and data transformation is available in the Toolbox.

Important CDS data is provided by the Sentinel satellite missions. They currently include:

- **Sentinel-1 SAR Data:** Sentinel-1 is a radar imaging mission that provides all-weather, day-and-night imaging capabilities. The Sentinel-1 SAR data set provides high-resolution radar images that can be used for applications such as oil spill monitoring, ship detection, and land deformation monitoring.
- **Sentinel-2 Optical Data:** Sentinel-2 is a multispectral imaging mission that provides high-resolution optical imagery. The Sentinel-2 optical data set includes imagery in 13 spectral bands, ranging from the visible to the near infrared. This data set can be used for different applications such as land cover classification, vegetation monitoring, and urban planning.
- **Sentinel-3 Ocean and Land Colour Instrument (OLCI) Data:** Sentinel-3 is a mission that provides a range of ocean and land observations, including sea surface temperature, ocean colour, and vegetation index. The OLCI data set provides information on the colour of the ocean and land surface, which can be used to monitor water quality and detect algal blooms.
- **Sentinel-5P TROPOMI Data:** Sentinel-5P is a mission that provides data on atmospheric composition, including measurements of pollutants such as nitrogen dioxide, carbon monoxide, and sulfur dioxide. The TROPOMI data set provides high-resolution measurements of these pollutants, which can be used to monitor air quality and support air pollution control policies.

2.2 ERA5

The main source of historical meteorological data that is spatially and temporally complete will be ERA5 reanalysis available from Copernicus Climate Store (CDS)¹. There are several ways to access the data. The RISKADAPT project partners can access the data directly from the Copernicus service and possibly other providers, like Google Cloud. FMI as the leader of the WP3 will assist other partners with data access.

The European Centre for Medium-Range Weather Forecasts (ECMWF) produces the ERA5 data, a global climate reanalysis data set, that provides information on a wide range of atmospheric

¹ <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels>

parameters, such as temperature, wind, humidity, and precipitation, at a relatively high spatial and temporal resolution. The ERA5 data set is widely used by climate scientists, researchers, and policymakers around the world for climate change research, climate modelling, and weather forecasting. The data set covers the period from 1979 to the present day and is continually updated with new data as it becomes available.

2.3 E-OBS

E-OBS is a land-only gridded daily observational dataset for precipitation, temperature, sea level pressure, global radiation, wind speed, and relative humidity in Europe². This dataset is based on observations from meteorological stations across Europe provided by the National Meteorological and Hydrological Services (NMHSs) and other data-holding institutes. The station data network is accessible through the webpages of the European Climate Assessment & Dataset (ECA&D, www.ecad.eu, Klein Tank et al., 2002).

2.4 CMIP6

The important climate projection data set provided by CDS is from the sixth phase of the Coupled Model Intercomparison Project (CMIP6)³. It provides daily and monthly global climate projections data from different experiments, models, and time periods. The data underpins the Intergovernmental Panel on Climate Change 6th Assessment Report. The climate projection experiments follow the combined pathways of the Shared Socioeconomic Pathway (SSP) and Representative Concentration Pathway (RCP). The SSP scenarios provide different pathways for future climate forcing. The period covered is typically 2006-2100, while some extended RCP experimental data is available from 2100-2300.

2.5 Accessing CDS data

There are different ways to access data in CDS. The recommended way to use the CDS data is by Python language scripts, see e.g., information in <https://www.copernicus.eu/en/access-data>.

- Using the CDS Toolbox. The data is stored and accessed in the Copernicus platform, calculations are defined through a web interface, and the results of the calculations can be visualised inside the interface or downloaded for further processing locally. The Toolbox allows users to develop applications that use the CDS data content. The user of the Toolbox can write Python scripts to access, manipulate, and export climate data made available through the CDS.
- Using the CDS API call by local Python programs. This way, the data is downloaded using the CDS API interface, and computations are then done locally.
- Using a cloud service, such as Google Cloud, that stores parts of Copernicus data. Now the computations can be done in the cloud or using a mode where the data extraction is done in the cloud and computations are performed locally⁴.

3. Other data sources

There are many observational data sources provided by national research institutes, such as national meteorological offices. Nowadays, data providers typically set up web-based API that uses open

² <https://cds.climate.copernicus.eu/cdsapp#!/dataset/insitu-gridded-observations-europe>

³ <https://cds.climate.copernicus.eu/cdsapp#!/dataset/projections-cmip6>

⁴ <https://developers.google.com/earth-engine/datasets/catalog>

standards for data access. For example, Finnish Meteorological Institute (FMI) has made its data sets freely available for public use, available in machine-readable, digital format. Additionally, computer programs were developed at the Institute, released as open-source code, and made available in the GitHub code repository. The technical implementation of the online service complies with the requirements laid down in the INSPIRE Directive. The search service will share the open data according to the Open Geospatial Consortium standards⁵. During the project, the partner's data requirements will determine which data sources will be used. These will be described in detail in the project's data management plan.

3.1.1 GEOSS

The international Group on Earth Observations (GEO) community has created a Global Earth Observation System of Systems (GEOSS)⁶ to better integrate observing systems and share data by connecting existing infrastructures using common standards. There are more than 400 million open data resources in GEOSS from more than 150 national and regional providers such as NASA and ESA, international organizations such as WMO and the commercial sector such as Digital Globe.

The GEOSS Portal (<https://www.geoportal.org>) offers a single Internet access point for users seeking data, imagery, and analytical software packages relevant to all parts of the globe. It connects users to existing data bases and portals and provides reliable, up-to-date, and user-friendly information. In addition to web browser-based interface to the Portal and various data sets, GEOSS offers APIs, which expose the discovery and access functionalities of the GEOSS and as such can be exploited by client applications.

4. Relation with other work packages/deliverables

Many WPs utilize EO data. One important use will be the construction of extreme value climatology for various meteorological variables, such as temperature, precipitation, humidity, and wind speed affecting the structures studied in this project, such as bridges and tall buildings. To assess the climate change effect, historical climatology will be compared to the one obtained with regional climate model outputs. The results will be conditional on assumed emission pathways. For example, in WP4, "Multi-Hazard Vulnerability and Adaptation, Structural Resistance Integration in Lifecycle Analyses", task 4.4 assesses atmospheric icing conditions in relation to corona losses in energy transfer lines and the related climate change mitigation measures. The atmospheric icing model will be adjusted to calculate the icing from regional climate model outputs, which involves investigating the available European Coordinated Regional Climate Downscaling Experiments (CORDEX) outputs and developing ways to utilize those outputs for the icing calculation. The results will show how icing conditions evolve with changing climate.

The following WPs and deliverables are directly relevant:

- WP2 – User Requirements, Architecture
- WP3 – Climate Data, CC Forcing, Multi-Hazard Modelling
- WP4 – Multi-Hazard Vulnerability and Adaptation, Structural Resistance Integration in Lifecycle Analyses
- D1.2 Quality Management Plan
- D1.3 Data Management Plan v1
- D3.2 Extreme value distributions-present climate

⁵ <http://www.opengeospatial.org/standards>

⁶ <https://earthobservations.org/geoss.php>

- D3.3 Extreme value distributions-future climate
- D3.4 Hydrologic/hydraulic modelling regarding floods
- D3.5 Model of high wind loading on a high rise building
- D4.2 Struct. resistance integration in lifecycle analyses
- D5.1 Data Management System

5. Conclusions

This report describes the intended types of EO data for RISKADAPT, including in-situ, remote sensed, and model-based data. Model-based data can be divided between historical reanalyses based on observations and climate model runs that simulate the climate system response to a scenario of future emissions. The most important data source is the Copernicus Climate Store which provides all the above-mentioned data types. Depending on the intended application, local observational data from national data providers might be needed to complement the Copernicus data. Currently, the efficient, least error-prone, and most reproducible way to access data in various forms is to use application programming interfaces (API) based on open standards.

References

Copernicus Climate Change Service (C3S) (2023): ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). 10.24381/cds.adbb2d47 (Accessed on 07-MAR-2023)

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2018): ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). 10.24381/cds.adbb2d47 , (Accessed on 07-MAR-2023)

Klein Tank, A.M.G., et al. (2002) Daily Dataset of 20th-Century Surface Air Temperature and Precipitation Series for the European Climate Assessment. *International Journal of Climatology*, 22, 1441-1453. 10.1002/joc.773

Annexes

Annex 1 – Example of the use of ERA5 reanalysis data.

The following Python code shows a typical example of accessing and visualising ERA5 reanalysis data using Copernicus API. It uses a predefined area definition in a shapefile to select the area of interest. After loading monthly average precipitation data for 2001-2021, it selects a grid point inside the nonrectangular area and calculates the average over that area. Lastly, the monthly time series is plotted. The produced time series plot is presented in Figure 1. More information on using Copernicus CDS data can be found on the CDS home page.

<https://cds.climate.copernicus.eu/cdsapp#!/home>

Code example

```
import xarray as xr
import matplotlib.pyplot as plt
import geopandas as gpd
import regionmask
import cdsapi

# load area definition shapefile of the area of interest
# extract lonlat bounding box
shape = gpd.read_file('Drainage_basin/basin.shp')
bb = shape.to_crs('EPSG:4326').bounds.values[0]

# retrieve ERA5 data of total precipitation over the area
cdsparams = {
    'format': 'grib',
    'variable': 'total_precipitation',
    'product_type': 'monthly_averaged_reanalysis',
    'year': [f'{n}' for n in range(2001, 2022)],
    'month': [f'{n:02d}' for n in range(1, 13)],
    'time': '00:00',
    'area': [bb[3]+0.1, bb[0]-0.1, bb[1]-0.1, bb[2]+0.1],
}
c = cdsapi.Client()
c.retrieve(
    'reanalysis-era5-single-levels-monthly-means',
    cdsparams,
    '/tmp/download.grib')
```

```
# open data set
ds = xr.open_dataset('/tmp/download.grib')

# select total precipitation
da = ds['tp'] * 1000
da.attrs['units'] = 'mm'

# mask for pixels inside given region
mask = regionmask.mask_geopandas(shape.to_crs("EPSG:4326"),
                                  ds.longitude.values, ds.latitude.values,
                                  lon_name='longitude', lat_name='latitude')

# plot monthly average precipitation over the region
fig, ax = plt.subplots(1, 1, figsize=(10, 5))
da.where(mask == 0).mean(dim=['latitude', 'longitude']).plot(marker='o')
plt.title("Monthly average total precipitation over the drainage area")
plt.ylabel("Total precipitation [mm]")
plt.xlabel("")
plt.grid()
plt.show()
```

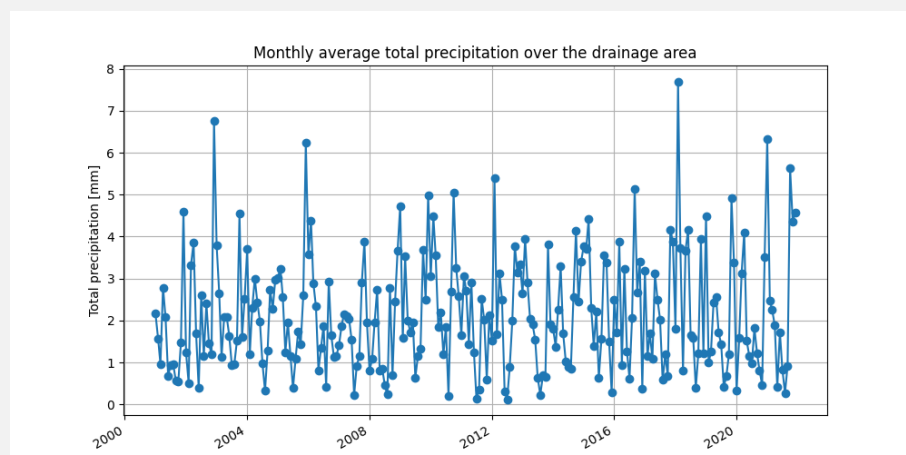


Figure 1: Plot produced by the example code.