SOS-Water deliverable report

[D3.1, Data inventory and EO data needs for water resources monitoring]

Lead beneficiary	11 - EAWAG	Due Date	31 March 2023
WP no	3	New due date (if delay)	
Task no	1	Actual Delivery Date	
Dissemination level	PU – Public	Status	FINAL DRAFT

Authors

Authors	Partner no	Partner organisation	Name of author
Main author	11	EAWAG	Michael Brechbühler
Contributing author(s)	8	SIWRP	Nam Nguyen Trung
	7	GEOECOMAR	Sabin Rotaru
	6.1	FUTUREWATER-NL	Gijs Simons
	11	EAWAG	Daniel Odermatt

Review

Authors	Partner no	Partner organisation	Name of author
Technical review	2	UU	Jennie Steyaert Niko Wanders

Document history

Date	Version	Chapters affected	Description of change	Author	Document status
01/03/2023	0.1	All	Initial version outlining report	Michael Brechbühler	DRAFT
10/03/2023	0.2	All	Additions	Michael Brechbühler Daniel Odermatt Gijs Simons	FIRST DRAFT
03/10/202 06/03/2023	0.3	All	Technical review	Jennie Steyaert Niko Wanders	FIRST DRAFT
16/03/2023	0.4	In-situ data needs and mismatches	Additions	Nam Nguyen Trung	FIRST DRAFT
17/03/2023	0.5	In-situ data needs and mismatches	Technical review	Jennie Steyaert	FIRST DRAFT
23/03/2023	0.6	In-situ data needs and mismatches	Corrections and additions	Nam Nguyen Trung	FIRST DRAFT
28/03/2023	0.9	In-situ data needs and mismatches	Additions	Sabin Rotaro	FIRST DRAFT
28/03/2023	1	All	Final version	Michael Brechbühler	FINAL DRAFT

Publishable Executive Summary

The overall goal of the SOS-Water project is to develop the foundation for a holistic and participatory assessment framework of the Safe Operating Space (SOS) for the entire water resources system across five different case studies in Europe and abroad.

The "Data inventory and EO data needs for water resources monitoring" (D3.1) consists of a public data inventory and publicly accessible report. The purpose of the Data Inventory outlined in this report is to give an overview of all state-of-the-art Earth Observation (EO) datasets available that can be potentially used for water resource monitoring. It covers all relevant services available from Copernicus (CGLS, CLMS, C3S), the ESA CCI (lakes, land cover, soil moisture), EUMETSAT (H-SAF) and other national and international organizations (e.g., NASA, NOAA, USGS, FAO). Within the SOS-Water project, this deliverable is intended to be used by all the project partners, to ensure knowledge exchange regarding the availability and limitations of project-relevant state-of-the-art EO datasets.

The objective of the present document is to review the datasets collected in the data inventory and determine their potential usability for the SOS-Water project. All collected datasets have been reviewed and augmented with key attributes, including spatio-temporal scales, limitations, accessibility and data provenance. These attributes were further used to identify the gaps and mismatches between EO data availability and the data needs posed by the project relevant hydrological models: the Community Water Model (CWatM) and PCR-GLOBWB 2.0. The gaps identified and outlined in this report will be prioritized and addressed with dedicated development efforts as part of the upcoming task "Improved EO applications" (T3.2).

Table of Contents

Abbreviations

ABADL	Dobrogea – Litoral Water Bazinal Administration
AFDJ	Lower Danube River Administration
ANAR	National Administration of Romanian Waters
ANM	National Administration of Meteorology
ARBDD	Danube Delta Biosphere Reserve Administration
C3S	Copernicus Climate Change Service
CCI	Climate Change Initiative
CGLS	Copernicus Global Land Service
CLIMFILL	Climate Data Gap-fill
CLMS	Copernicus Land Monitoring Service
CNES	Centre national d'études spatiales
CSA	Canadian Space Agency
DINCAE	Data Interpolating Convolutional Auto-Encoder
DINEOF	Data Interpolating Empirical Orthogonal Functions
EO	Earth Observation
ESA	European Space Agency
EUMETSAT	European operational satellite agency for monitoring weather, climate and the environment from space
EWV	Essential Water Variable
FAO	Food and Agriculture Organization
GRACE	Gravity Recovery and Climate Experiment
GRACE-FO	Gravity Recovery and Climate Experiment Follow-On
H-SAF	Satellite Application Facility on Hydrology and Water Management
HYCOS	Hydrological Cycle Observing System
IM	Impact Model
INHGA	National Institute of Hydrology and Water Management
IWMS	Integrated Water Modelling System
KARIN	Ka-ban Radar Interferometer
LMCS	Landscape Change Monitoring System
LWLR	Lake Water-Leaving Reflectance
MRC	Mekong River Commission
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration.
RS	Remote Sensing

1 Description of deliverable

1.1 Objective and scope of deliverable D3.1

The overall goal of the SOS-Water project is to develop the foundation for a holistic and participatory assessment framework of the Safe Operating Space (SOS) for the entire water resources system across five different case studies in Europe and abroad. To achieve this goal an integrated water modelling system (IWMS) is designed that dynamically integrates water system models (WSMs) and impact models (IMs) with a focus on agricultural and energy production, drinking water supply and biodiversity. In Task 3.1 we collect and inventory available physical and socio-economic Earth observations (EO) data that will help advance the parameterization and benchmarking of the IWMS of water planning and management trends at local and regional scales. Furthermore, we will use this inventory to identify existing mismatches between both water resources challenges and EO data availability.

In addition, the Data Inventory is intended as an information source for all project partners to keep them updated about the availability and limitations of project-relevant state-of-the-art EO datasets. For this purpose, the inventory will be continually updated throughout the project. For completeness, available in-situ datasets are also catalogued as part of the data inventory.

In this way, the specific objectives of this document (D3.1) and the provided data inventory are:

- Collect all relevant EO services and datasets from Copernicus (CGLS, CLMS, C3S), the ESA CCI (lakes, land cover, snow, soil moisture), EUMETSAT (H-SAF) as well as other national and international organizations (e.g., NASA, NOAA, USGS, FAO).

- Review all relevant in-situ datasets from the case study leaders and relevant organizations both nationally and internationally.

- Assess and compile all important dataset attributes, including spatio-temporal scales, limitations, accessibility and data provenance in a tabular data inventory.

- Determine mismatches between water resources challenges and EO data availability.

The mismatches identified in this report will be prioritized and addressed with dedicated development efforts as part of the upcoming Task 3.2 “Improved EO applications” This will include the up and down sampling of spatio-temporal resolution from standard geoinformation services and the gathering of additional parameters, such as evapotranspiration or cyanobacteria abundance retrieval, or value-added derivatives such as phenology, sedimentation, stratification or irrigation products. The development of selected applications may include algorithm development, but will more importantly, focus on the customization and up-scaling of information products to be used as reference data for the benchmarking of IWMS within the projects spatiotemporal scope.

1.2 Task management

Table 1 Outlines the distribution of work across the project partners involved in Task 3.1.

Table 1 – Task 3.1 Task management overview.

Partner organisation	Partner no	Description of work
EAWAG	11	Cataloging of all relevant (global and local) EO and in-situ datasets and assimilating them into the data inventory table Main author of this report
GEOECOMAR	7	Cataloging of relevant in-situ datasets within the Danube basin Co-author of report (Section about in-situ mismatches specific to Danube Delta)
SIWRP	8	Cataloging of relevant in-situ datasets within the Mekong basin Co-author of report (Section about in-situ mismatches specific to Mekong Delta)
FUTUREWATER-NL	6.1	Evaluation and cataloging of satellite-based altimetry (SWOT, Sentinel-3/6) and evapotranspiration (actual, potential) Co-author of this report (Section about potential of altimetry data)

2 Deliverable

2.1 Data inventory

2.1.1 Inventory methods

Prior to data collection, we first defined a list of potential data sources according to the objectives stated above. To structure and order the collection and sources, we defined the two main categories **Earth Observation** (Tab. 2) and **In-situ** (Tab. 3). For the purposes of this classification, EO are defined as remotely sensed products and consist of mainly datasets from satellites, while in-situ consist of datasets from ground measurements.

Tab. 2 – Earth Observation data sources.

Data Source	Description	URL
CEOS Database	Large database of EO data. Includes platform and timelines of historical, current and future missions.	http://database.eohandbook.com/data/dataactivity.aspx
ESA CCI Climate Dashboard	Overview and timelines of available datasets grouped by indicators	https://climate.esa.int/en/odp/#/dashboard
CEDA Archive	Data archive of various climatic datasets as well as ESA CCI datasets	https://catalogue.ceda.ac.uk/
EUMETSAT Data Store	Data storage of mostly meteorological datasets, including Copernicus products	https://www.eumetsat.int/eumetsat-data-store
Google Earth Engine	Cloud data provider for satellite products as well as other datasets (both local and global)	https://developers.google.com/earth-engine/datasets/
SPOT World Heritage Data Archive	SPOT 1-5 Data Archive (L1 data), cloud-processing to L2 at https://swh-2a-carto.fr/processing	https://regards.cnes.fr/user/swh/modules/60
Copernicus Open Access Hub	Data archive of Copernicus mission data (S1, S2, S3, S5)	https://scihub.copernicus.eu//
Copernicus Land Monitoring Service	Overview of Copernicus Land products (Global, Pan-European and Local services)	https://land.copernicus.eu/
Copernicus Local Land Service	Data library and access point for local products of the Copernicus Land Service	https://land.copernicus.eu/local
Copernicus Pan-European Land Service	Data library and access point for Pan-European products of the Copernicus Land Service	https://land.copernicus.eu/pan-european
Copernicus Global Land Service	Data library and access point for Global products of the Copernicus Land Service	https://land.copernicus.eu/global/
Copernicus Climate Change Service	Data library and access point for Copernicus Climate Change Service products	https://cds.climate.copernicus.eu/#!/home

NEO (NASA Earth Observations)	NASA EO dataset overview for all spheres and DAACs	https://neo.gsfc.nasa.gov/
NASA Earth-Data	Landing page for all available NASA associated EO DAACs.	https://www.earthdata.nasa.gov/
NASA/USGS Land Processes DAAC	Data library of NASA/USGS products	https://lpdaac.usgs.gov/product_search/
Alaska Satellite Facility	Data provider for various NASA missions including soil moisture (SMAP) and derived products (e.g., InSAR cloud processing)	https://asf.alaska.edu
Hydroweb Theia	Database of water level height (multi-sensor altimetry) from CLS under the coordination of LEGOS, IRD and CNES	https://hydroweb.theia-land.fr/
DAHITI (Database for Hydrological Time Series of Inland Waters)	Database of water level height (multi-sensor altimetry) from Universität München	https://dahiti.dgfi.tum.de/en/
G-REALM (Global Reservoirs and Lakes Monitor)	Database of water level height (multi-sensor altimetry) of global lakes > 100 km2	https://ipad.fas.usda.gov/cropexplorer/global_reservoir/
HydroSHEDS database	The HydroSHEDS database offers a suite of global digital data layers in support of hydro-ecological research and applications worldwide. Its various hydrographic data products include catchment boundaries, river networks, and lakes at multiple resolutions and scales.	https://www.hydrosheds.org/
ARC-Lake products	The ARCLake project have generated accurate and consistent LSWT time series from ATRS sensors for more than 250 lakes globally from 1995 to 2009.	http://www.laketemp.net/home_ARCLake/dataF/
GloboLakes products	The University of Reading contribution to the GloboLakes project is to develop robust algorithms for retrieving Lake Surface Water Temperature (LSWT) at global scale quantifying uncertainty estimates.	http://www.laketemp.net/home_GL/
Theia products	Provider of RS products and services for land surfaces.	https://www.theia-land.fr/en/products/#
HYDROLARE	Global in-situ database of hydrology of world lakes and reservoirs (including 1069 water bodies)	http://hydrolare.net/index.php
Global Terrestrial Network – Hydrology	GTN-H links existing networks and systems for integrated observations of the global water cycle.	https://www.gtn-h.info/
List of NOAA Open Data Dissemination	Overview of NOAA cloud datasets	https://www.noaa.gov/nodd/datasets

Program Datasets		
NOAA NCEI Products	Overview of NOAA products	https://www.ncei.noaa.gov/products
EUMETSAT HSAF	Overview of EUMETSAT Support to operational Hydrology and Water Management Products	https://hsaf.meteoam.it/

Tab. 3 – In-situ data sources.

Data Source	Description	URL
Open Development Mekong	Large open data hub with datasets from Mekong	https://data.opendevelopmentmekong.net/
Open Development Cambodia	Large open data hub with datasets from Cambodia	https://opendevelopmentcambodia.net/
Open Development Laos	Large open data hub with datasets from Laos	https://laos.opendevelopmentmekong.net/
Open Development Myanmar	Large open data hub with datasets from Myanmar	https://opendevelopmentmyanmar.net/
Open Development Thailand	Large open data hub with datasets from Thailand	https://thailand.opendevelopmentmekong.net/
Open Development Vietnam	Large open data hub with datasets from Vietnam	https://vietnam.opendevelopmentmekong.net/
DanubeGIS	Collection of Danube basin maps curated by the ICPDR	https://www.danubegis.org/
Danube River Basin Water Quality Database	Access to data from the TNMN (TransNational Monitoring Network)	http://www.icpdr.org/wq-db/
Danubis	ICPDR Information System with access to Datasets	https://danubis.icpdr.org/
BAFG - Hydrologische Daten	German federal data portal for water measurements of rivers/lakes	https://www.bafg.de/EN/06_Info_Service/01_WaterLevels/waterlevels_node.html
Pegelonline	German federal database of hydrological measurements of rivers/lakes and groundwater	https://www.pegelonline.wsv.de/gast/start
eHYD	Austrian federal database of hydrological measurements of rivers/lakes and groundwater	https://www.ehyd.gv.at/
Opendata.swiss	Swiss open data platform	https://opendata.swiss/de
GovData	German open data platform	https://www.govdata.de/
Open platform for French public data	French open data platform	https://www.data.gouv.fr/en/
BAFU - Hydrologische Daten und Vorhersagen	Swiss federal data portal for water measurements (e.g., flow, water level, temp., water quality) of Swiss rivers/lakes (~260 stations) and groundwater (~550 stations)	https://www.hydrodaten.admin.ch/de/

Open data initiative of the Government of Spain	Open data collection from the Spanish government including some datasets of Jucar River	https://datos.gob.es/en
data.europe.eu	EU portal for European open data datasets	https://data.europa.eu/data/datasets?locale=en
Limnology models and datasets	Exhaustive list of hydrological datasets and models	https://github.com/jsta/limnology_models_data
GEMStat water quality data	Water quality data portal coordinated by the United Nations	https://gemstat.org/data/data-portal/

Once the data sources had been identified, we collected and acquired key metadata attributes of all the relevant datasets provided in each one of them. Most of the collected datasets within the data inventory are multivariate spatio-temporal EO datasets, resulting in a heterogenous collection. To ensure that the main contents (e.g., variables and units), dataset scales (e.g., spatial and temporal extent/resolution) and data access restrictions (e.g., licensing and availability) are captured, we developed a list of key attributes (Table 4) that we extracted for each dataset. These attributes ensure that the contents, limitations and the resulting potential use of a dataset can be quickly assessed by all project partners.

Table 4 – List of collected dataset attributes.

Attribute	Description	Example entry
Dataset name	Unique dataset name	ESA Lakes Climate Change Initiative (Lakes_cci): Lake products, Version 2.0.1
Description	Description of dataset/dataset-product	This dataset contains [..] based on [..].
Data provider	Name of data provider/portal	Copernicus Global Land Service
Variable	Main variable of dataset	Lake Surface Water Temperature
Unit	Unit of main variable	degrees Kelvin
Spatial Resolution	Spatial sampling used	20 meters
Spatial extent	Spatial coverage of dataset	Germany, Austria
Temporal resolution	Temporal sampling used	Weekly
Temporal extent	Temporal coverage of dataset	1992-09-26 to 2020-12-31
URL	URL to access dataset	https://climate.esa.int/en/projects/lakes/data/
DOI	Digital Object Identifier for dataset (if available)	10.5285/a07deacaffb8453e93d57ee214676304
References	Associated references as DOI or text (if available)	Messager, M., Lehner, B., Grill, G. et al. Estimating the volume and age of water stored in global lakes

		using a geo-statistical approach. Nat Commun 7, 13603 (2016).
Data type	Description of data type (e.g., tabular, vector, raster)	tabular
Data format	Description of data format (e.g., shapefile, csv, netCDF)	netCDF
Catchment	Associated SOS-Water case study	Mekong Basin
Type	Type of dataset (e.g., Earth Observation, In-Situ)	Earth Observation
License	Dataset terms of use	CC BY 4.0
Comments	Additional comments (e.g., accessibility restrictions, limitations of dataset, available ancillary variables)	Dataset is only accessible with registered DANUBIS ICPDR/DanubeGIS account.

To ensure that the datasets collected are appropriately labelled, documented, reusable, and easily accessible, we verified that the added datasets followed the usage of FAIR (Findable, Accessible, Interoperable, and Reusable) standards (see below for detailed information on this). This is crucial for the creation of a comprehensive and effective data inventory that can be readily searched, analysed, and shared across the project partners. To achieve this, all added datasets should be described in a detailed manner with rich metadata to fulfil the goals depicted in Figure 1.

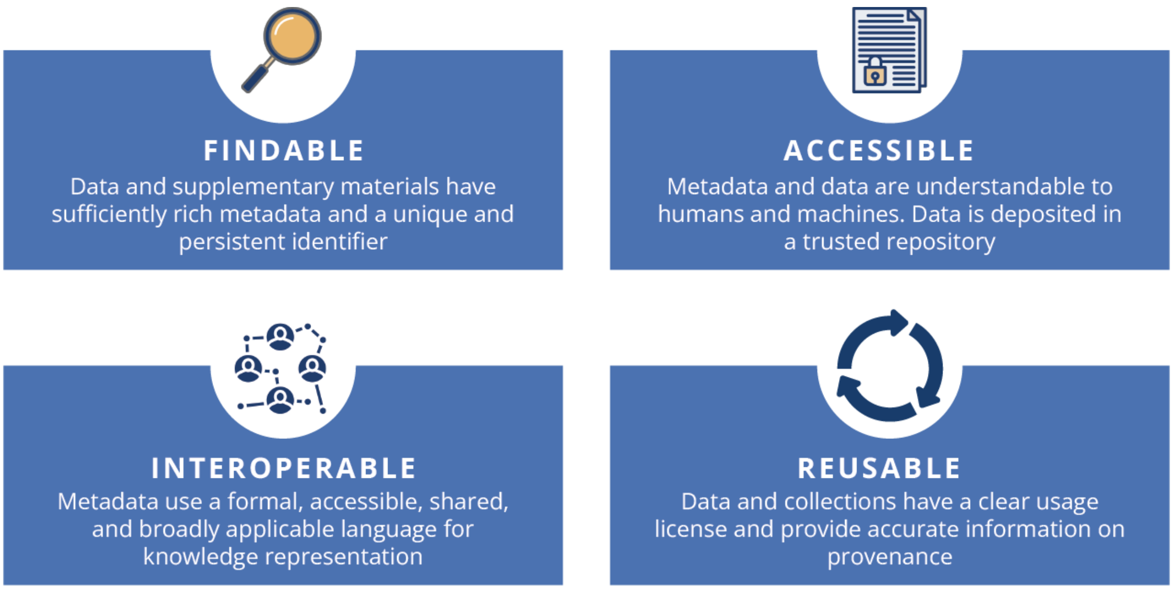

Figure 1 – Schematic representation and description of the FAIR data principles.

2.1.2 Data overview

In total, 101 project-relevant datasets were reviewed and compiled in the data inventory. The largest portion (31%) of entries were collected from the Copernicus Global Land Service (CGLS). The CGLS is a component of the Land Monitoring Core Service (LMCS) of Copernicus, the European flagship program on Earth Observation. The second largest part (17%) was collected from the ESA Climate Change Initiative (CCI) Open Data Portal.

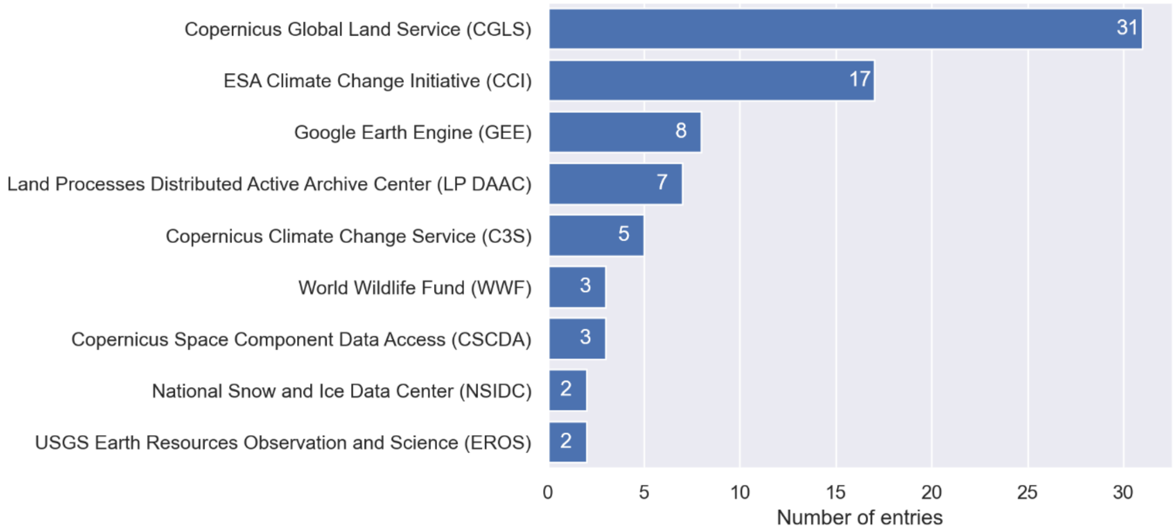

Figure 2 – Bar plot showing the number of dataset entries per data source. Sources with only one associated dataset entry (23 in total) are not shown.

When using EO data to measure bio-geophysical variables, the spatial and temporal resolution are two key factors to be aware of. Spatial resolution is typically measured in meters and represents the level of spatial detail. In most cases it can be equated to the pixel size and the smallest feature that can be detected. There are large variations in spatial resolution from several meters (e.g., 10m for Sentinel-1) to several kilometers (e.g., >300 km for GRACE). Temporal resolution refers to the time elapsed between two subsequent acquisitions of data from the same area on Earth (i.e., how long it takes a satellite to pass over the same point). Depending on the methods and platforms used, the temporal resolution can range from nearly continuous coverage to several days between revisits.

It is critical to match these two factors with the scales of the observed features or processes. For example, if the spatial resolution of EO imagery is too low, it may not be possible to distinguish between different types of land cover, such as wetlands and agricultural land. This could result in inaccurate estimates of the amount of water available in a particular area. SOS-Water intends to apply hydrological models up to daily temporal and one kilometer (5 arc-second) spatial resolution. Therefore, we assume these two thresholds as targeted resolutions.

Figure 3 gives an overview of grouped datasets according to the observed water variables and their spatial resolutions. Most collected EO datasets (68%) are provided at or below 1 km resolution (identified as on the graph as points on the left from the red line). There are 16% of EO datasets with point observations by station or aggregated by lake (depicted by datapoints on the black line).

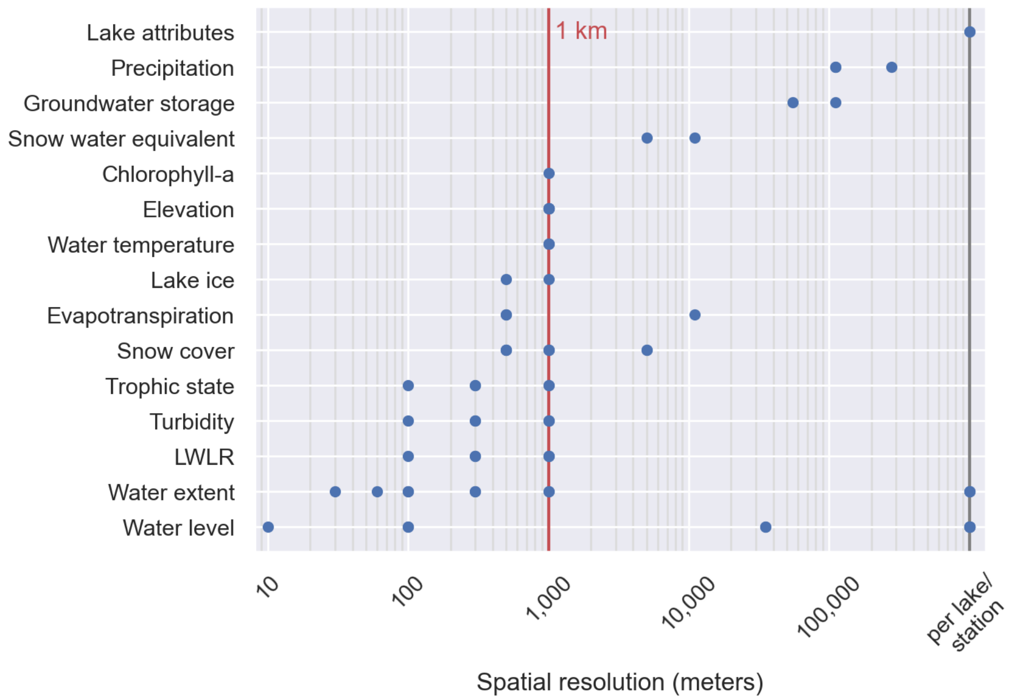

Figure 3 – Scatterplot of datasets and their reported spatial resolution (logarithmic axis) grouped into the associated water variables. The red vertical line marks the spatial resolution of one kilometer, which will be the targeted resolution for the hydrological models in the SOS-Water project. The black line marks datasets provided in point observations per station or aggregated per lake.

To complement the spatial resolution, Figure 4 gives an overview of datasets grouped according to the observed water variables and the corresponding temporal resolutions. Many EO products are provided as global multivariate datasets with daily temporal resolution. Thus, they fulfil the criteria for daily temporal resolution. However, this is often misleading and does not reflect the real data availability, since many of these datasets include data gaps that lower the real temporal resolution. These gaps are most often a result of either cloud cover or the limited spatial coverage due to orbital paths of the used satellite platforms.

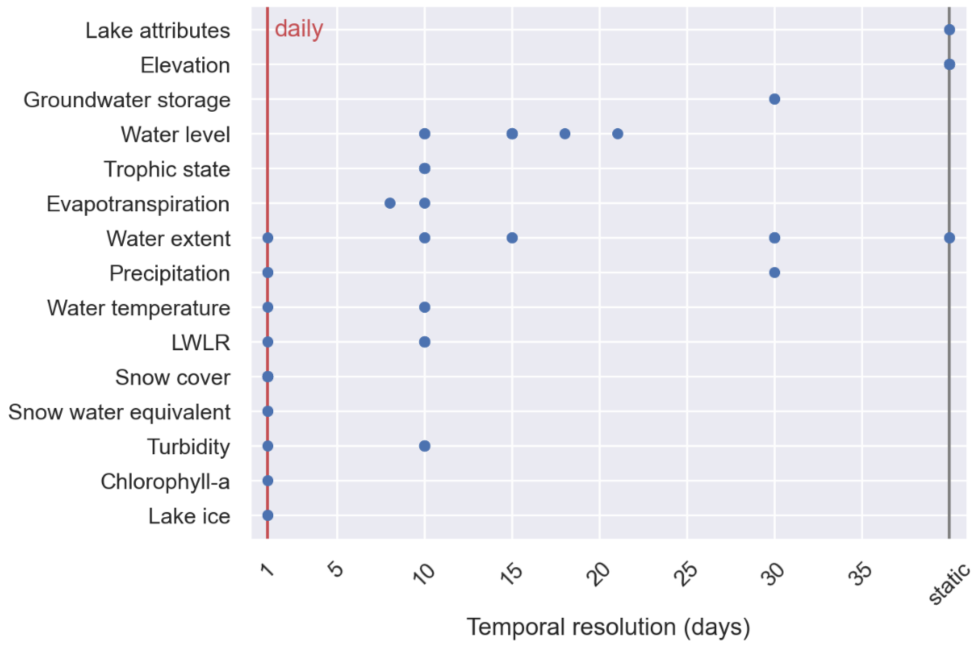

Figure 4 – Scatterplot of datasets and their reported temporal resolution grouped into the associated water variables. The red vertical line marks the daily temporal resolution, which will be the target resolution for the hydrological models in the SOS-Water project. The black line marks datasets provided as a static product without temporal information.

Figure 5 gives the number of grouped datasets according to the reported spatial extent. Most of the collected EO products (83%) are provided on a global extent, resulting in a coverage of all case studies. Due to limitations in spatial resolution, many of the global datasets are limited to large lakes and reservoirs. For example, the state-of-the-art multivariate ESA CCI Lakes v2.02 dataset covers key lake Essential Climate Variables (ECVs) on a global scale but is limited to 2024 lakes in total. Thus, the applicability and usefulness for the specific study area must be further evaluated.

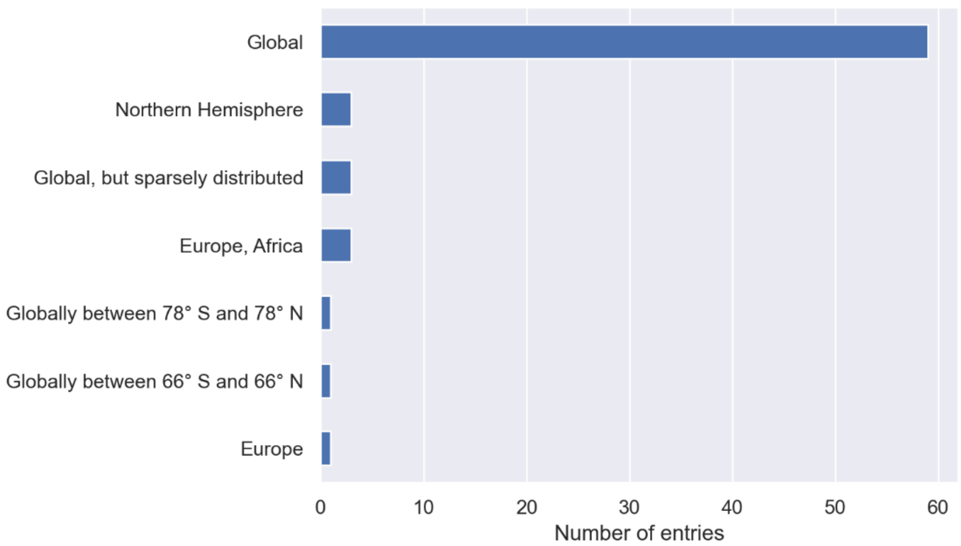

Figure 5 – Bar plot showing the number of dataset entries grouped by spatial extent.

2.2 Gap analysis

2.2.1 Water resources challenges

Water resource challenges are often difficult to monitor and manage due to a lack of reliable, timely and extensive data. By linking the challenges at hand with associated Essential Water Variables (EWVs), we can define the measurable variables necessary to better understand and manage our freshwater resources, which are essential for human well-being and ecosystem health. Many of these EWVs can be assessed using EO data. However, despite the potential benefits of EO data, there are remaining gaps in data availability, which limit our ability to fully understand and address these challenges. Some of these gaps include:

- **Spatial and temporal resolution**: EO data may not have the necessary spatial and temporal resolution to capture the complexity of freshwater resource challenges. For example, remote sensing data may not be able to detect small-scale changes in water quantity and quality or may not provide sufficient temporal resolution to capture changes in water use over time. Specifically, studies looking at water quality, availability and use may have to use proxy variables such as chlorophyll-a concentrations and lake water level.
- **Spectral Resolution**: Some EO datasets may not have enough spectral resolution (defined as the ability of a sensor to define fine wavelength intervals) to differentiate between different types of freshwater features or assess water quality variables accurately.
- **Data gaps**: There may be gaps in EO data coverage that limit our ability to monitor freshwater resources in certain areas, such as poorly mapped (e.g., high latitude) or cloudy regions (e.g., European Alps). In addition to regional data gaps, there may also be temporal gaps due to climatic patterns and seasons such as monsoons or continuously cloudy weather. This can make it difficult to accurately assess water availability, quality, and use in these areas.
- **Data accessibility**: While EO data is becoming more widely available, there are still barriers to accessing and using this data effectively. This includes issues such as data privacy, data sharing agreements, and the need for specialized technical expertise. EO data is often complex and requires advanced analytical techniques to extract meaningful information. Furthermore, published products are delivered in a wide array of formats and processing levels. As a result, elaborate pre-processing steps are often needed to integrate data from different sources.

In addition to the challenges associated with EO, there are also significant freshwater resource challenges that EO data alone cannot address. Addressing these challenges requires either new technologies or a comprehensive approach that incorporates EO data alongside other types of data and information, as well as effective policies and governance frameworks. The SOS-Water Framework is meant to use an interdisciplinary group to address these challenges.

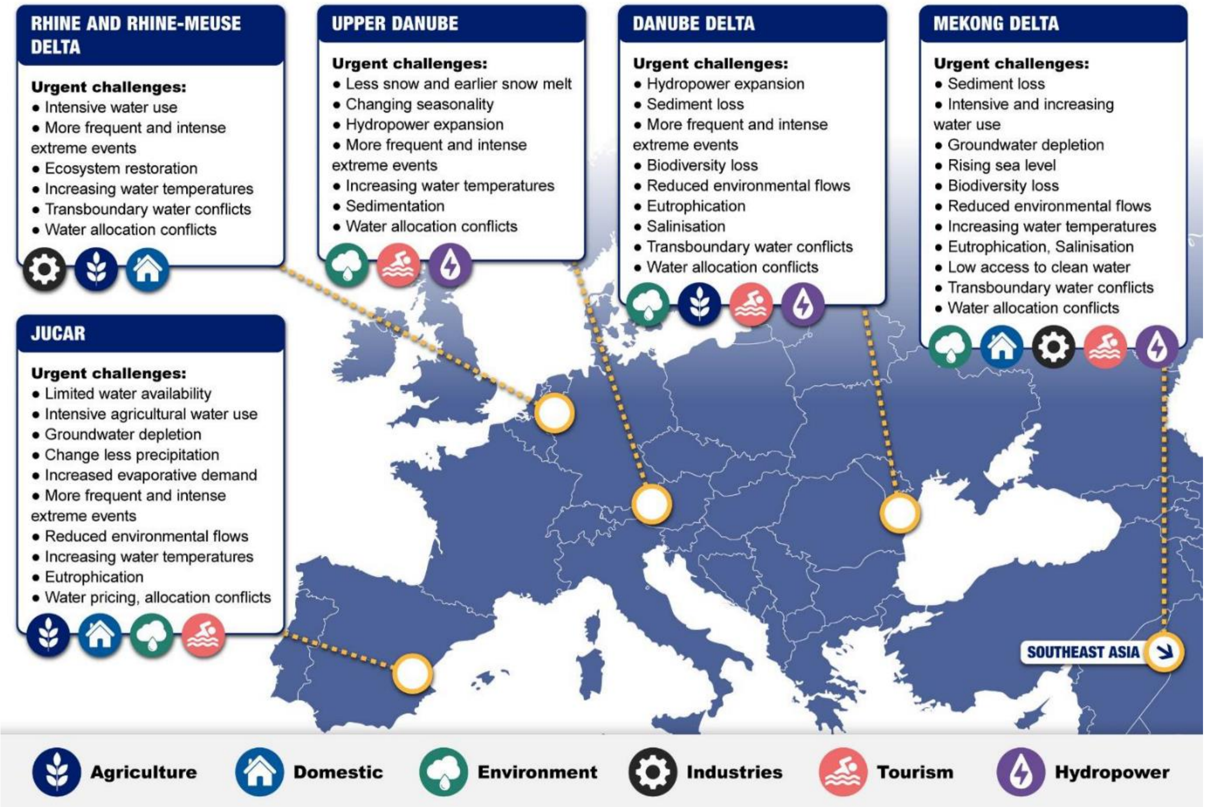

Figure 6 – Overview of SOS-Water case studies and their urgent water challenges and key sectors impacted.

Several urgent water challenges have been identified within the SOS-Water case studies as part of a preliminary examination as shown in Figure 6. The main challenges fall into six water use categories (agriculture, domestic, environment, industries, tourism and hydropower) and encompass key issues related to climate change such as decreased precipitation, decreased water availability and quality. In addition to these environmental issues, each basin has transboundary water allocation conflicts related to current and predicted changes in water availability and demand.

EO can be a key tool to monitor and promote sustainable solutions to these key challenges from Figure 5. Table 5 denotes the specific ways in which EWVs can be used to monitor urgent water challenges such as biodiversity, water availability, temperature, demand, dam building, climatic shifts and water allocations.

Table 5 – Urgent water challenges, associated EWVs and EO monitoring opportunities.

Urgent water challenge	Associated EWVs[1]	EO monitoring opportunities
Ecosystem restoration	Land cover, vegetation and land use	Surface water extent, vegetation productivity, land use/cover
Increasing water temperatures	Water Quality	Surface water temperature
Limited water availability	Lakes/reservoir levels and aquiver volumetric change, terrestrial water storage, Soil moisture/temperature, Groundwater	Satellite altimetry, groundwater storage estimation, soil moisture
Intensive agricultural water use	Water use/demand, Land cover, vegetation and land use, Soil moisture/temperature	Land use/cover, soil moisture
Groundwater depletion	Groundwater	Groundwater storage estimation
Change less precipitation	Precipitation, Snow cover (including snow water equivalent, depth, freeze thaw margins)	Precipitation monitoring, snow monitoring
Increased evaporative demand	Evaporation and evapotranspiration	Evapotranspiration monitoring
More frequent and intense extreme events	All EWVs	Land surface temperature, soil moisture, precipitation monitoring, vegetation productivity
Reduced environmental flows	Land cover, vegetation and land use, Water quality	Surface water extent, water quality monitoring
Eutrophication	Water Quality	Trophic state, chlorophyll-a
Less snow and earlier snow melt	Snow cover (including snow water equivalent, depth, freeze thaw margins)	Snow monitoring
Changing seasonality	Precipitation, Land cover, vegetation and land use	Vegetation phenology, vegetation productivity
Hydropower expansion	Water use/demand (agriculture, hydrology, energy, urbanization)	-
Sedimentation	Water quality	Turbidity
Sediment loss	Water quality	Turbidity
Biodiversity loss	Land cover, vegetation and land use	Land use/cover, surface water extent, vegetation productivity
Rising sea level	-	Satellite altimetry, water extent
Low access to clean water	Water Quality	Total Suspended Matter (TSM), Cholorphyll-a, Colored Dissolved Organic Mater (CDOM)
Salinisation	Water Quality	-

[1] JAXA, and GEO. (2014) The Geoss Water Strategy from Observations to Decisions.

transboundary water conflicts	Water use/demand	Land cover, vegetation and land use
water pricing/allocation conflicts	Water use/demand	Land cover, vegetation and land use

2.2.2 EO data needs and mismatches

Today, hydrological models strongly rely on EO derived global measurements of rainfall, soil moisture, snow cover, groundwater storage change, and surface water elevation to name a few. Table 6 shows a summary of all hydrologic variables currently available. Even though EO products have seen astounding advances in the past decades, there are still many knowledge gaps which are denoted in column three of Table 6.

Table 6 – Hydrological variables, associated EO methods and knowledge gaps.

Hydrological Variable	EO Methods	Knowledge Gaps (McCabe et al., 2017)[2]
Precipitation	Active and passive microwave RS	Spatio-temporal gaps introduce large uncertainties, limited spatial resolution, limited rainfall retrieval at high latitudes, lack in detection of snowfall and extreme events
Evaporation	Optical RS, thermal RS	Complex and site-specific retrieval algorithms, too coarse spatial resolution for critical applications (e.g., drought assessment)
Soil moisture	Active and passive microwave RS	Shallow retrieval depths
Runoff	Altimetry	Extensive calibration needed (reliant on algorithms that account for the unknown channel depths and flow velocities), generally limited to one-dimensional height retrieval, limited temporal resolution
Groundwater and terrestrial water storage	Gravimetric RS	Very coarse spatial resolution (> 300 km), can only infer changes in storage
Vegetation	Optical RS	Too much emphasis on simplistic broadband vegetation indices such as NDVI
Snow, ice and permafrost	Optical RS, active and passive microwave RS	Retrieval of snow depth, density or water equivalent (SWE) still very limited, no dedicated permafrost missions
Water vapour	Radar, Lidar	Lack of high spatio-temporal resolution data
Water quality	Optical RS, thermal RS	Limited data availability, lack of dedicated EO missions for inland waters

[2] McCabe et al. (2017) The Future of Earth Observation in Hydrology. Hydrol Earth Syst Sci. 21(7), 3879-3914

Bridging many of the identified knowledge gaps in Table 6 relies on the development of new and improved sensor technology and elaborate algorithms. For example, gravimetric measurements from space by the Gravity Recovery and Climate Experiment (GRACE) mission using the GRACE and GRACE-FO platforms can extract information on changes in groundwater storage capacity. This is a crucial input to calibrate and validate current hydrological models. However, the existing sensors can only provide information in a very coarse spatial resolution (> 300 km), while most global hydrologic models are striving for a finer spatial resolution (~ 1 km^2).

In such cases, the improvement of available EO data is out of the scope for the SOS-Water project. However, the understanding of other hydrological variables can be greatly improved by deriving meaningful spatio-temporal aggregates and value-added derivatives from existing EO datasets. Here we list feasible development efforts as part of the upcoming Task 3.2, "Improved EO applications". These development efforts have the potential to leverage existing EO data to bridge some of the identified knowledge gaps and a key set of data is Space altimetry

Space altimetry can give accurate measurements of water surface elevation in rivers, lakes, and oceans. This information can further be used to derive storage volumes and river flow dynamics, including the speed and volume of water moving through a river channel. The currently active Sentinel-3 and Sentinel-6 missions, launched as part of the Copernicus program, both carry along-track altimetry sensors and can provide water level measurements from inland waters. Sentinel-3 can retrieve inland water levels down to small water bodies (300 meters at < 27-day frequency), whereas Sentinel-6 has a higher temporal resolution (300 meters at < 10-day frequency)[3]. Jason-2 and Jason-3 are other water level sensors used in the hydrological community, and harmonized Jason and Sentinel-3 time series are disseminated through the Copernicus Land platform[4]. Figure 6 gives an overview of the timeline of relevant past, current and future altimetry missions.

Potentially, altimetry-derived water levels hold great added value for the hydrological modelling community. Altimetry comes close to direct satellite observations of river discharge, which is a crucial variable in most hydrological model calibration efforts. In addition, for reservoirs and other water bodies managed by humans, altimetry provides an opportunity for detecting human influence on the water body elevations. However, limitations in spatial resolution, temporal resolution, as well as coverage (e.g., water bodies located in intertrack areas are not covered) have led to limited uptake in the modelling community compared to other EO data sources. Integration of different altimetry data sources can also improve its usefulness for hydrological monitoring.[5]

Improving on the current state of art regarding satellite altimetry, the Surface Water and Ocean Topography (**SWOT**) mission, which is a collaboration between NASA and the French space agency CNES, will measure the height and width of water bodies at an unprecedented level of detail: 100 meters at < 21-day frequency, covering most parts of the globe every ~10 days. In contrast to conventional along-track altimetry missions, which provide data at a one-dimensional spatial level,

[3] Donlon et al. (2021), The Copernicus Sentinel-6 mission: Enhanced continuity of satellite sea level measurements from space, Remote Sensing of Environment 258

[4] https://land.copernicus.eu/global/products/wl

[5] Li et al. (2022, Monitoring of hydrological resources in surface water change by satellite altimetry, Remote Sensing 14, 4904.

SWOT will be able to provide data from hundreds of thousands of lakes as well as the discharge volumes of medium-to-large rivers at a two-dimensional spatial level. With its successful launch in December 2022, the first calibrated Level-2 altimetry data are expected to be released from October 2023 onwards.

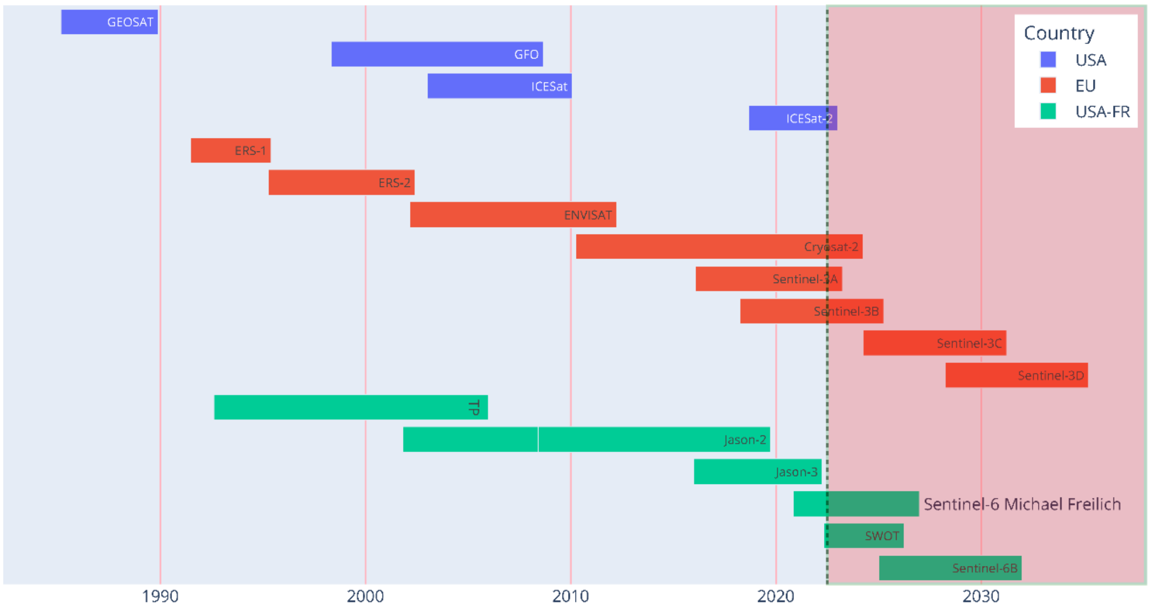

Figure 6 – Timeline overview of relevant past, current and future altimetry missions. Adapted from Yang et al., 2022[6].

The great combined potential of these three missions, as well as their coinciding timelines with the SOS-Water project, give an opportunity to harness their data shown in Table 7. Through extracting and pre-processing data from these three missions for the case studies at hand, we will be able to facilitate inclusion of this novel data for the project partners in an early stage. In fact, inclusion of altimetry data will ensure better calibration of the large-scale hydrologic models at the 1km spatial resolution.

Table 7 – Key properties of three satellite altimetry missions.

Satellite, *sensor*	Data provider	Spatial coverage	Spatial resolution	Revisit time	Temporal extent
Sentinel-3, *SRAL*	ESA	Global, with 52 km intertrack gaps (65,000 data points)	Up to 300 m	27 days	2016 - present
Sentinel-6, *Poseidon-4*	ESA	Global, with 20 km intertrack gaps (>30,000 data points at the time of launch)	Up to 300 m	10 days	Nov 2020 - present

[6] Yang et al. (2022) Satellite Altimetry: Achievements and Future Trends by a Scientometrics Analysis. Remote Sens 14(14), 3332.

SWOT, *KaRin*	NASA, CNES, CSA, UKSA	Global	250 m (lakes) 100 m (rivers)	< 21 days (10 days for most regions)	2023 -

Water quality retrieval in lakes and reservoirs is key in maintaining aquatic ecosystem services such as safe drinking water, recreation, fishing and aquatic biodiversity. EO-based water quality retrieval is limited to a handful of water variables including clarity, turbidity, water color, and the concentrations of optically active constituents. Despite its limitations, the use of EO data to obtain these variables allows us to achieve an extensive spatial coverage, which traditional and generally sparsely distributed in-situ stations cannot produce. Recent efforts at Eawag have made use of the in-house developed SenCast[7] software to put forward datalakes (https://www.datalakes-eawag.ch/), an operational near-real-time water quality monitoring system for all large Swiss lakes.

SenCast is an open-source software package that enables automatic processing of data from the Sentinel-2 and Sentinel-3 satellites. It is implemented in Python and mainly utilises processors from the Sentinel Application Platform (https://step.esa.int), an open-source toolbox from ESA (e.g., Idepix, C2RCC, MPH and Sen2cor). With SenCast the decisive water parameters chlorophyll-a, Secchi depth (water clarity), turbidity and particle content can be determined. Additionally, **phytoplankton phenology** can be further inferred from retrieved chlorophyll-a concentrations and is important to understand associated changes in water quality and lake productivity. Development efforts to apply this methodology to the SOS-Water case studies to retrieve water quality parameters or phytoplankton phenology are conceivable.

Spatio-temporal gaps in time-series acquired from EO sensors are a major problem for most applications. Hydrological modelling often requires inputs with gap-free spatial and temporal coverage. As a result, a multitude of gap-filling methods to deal with this problem have been developed and used (e.g., DINEOF, DINCAE, CLIMFILL). Despite this, the application of gap-filling methods is often omitted in EO datasets provided by official sources and the efforts associated with gap-filling are left to the data user.

A reliable gap-filling method applicable to EO data is **DINEOF**[8] (Data Interpolating Empirical Orthogonal Functions). DINEOF is a technique to fill in missing data in geophysical datasets and is based on an empirical orthogonal function decomposition. Thus, it uses the spatio-temporal coherence present in the data to calculate missing values. Even though this method has been developed and improved for nearly two decades and is widely adopted for gap-filling of oceanic datasets, it has seen very low uptake in freshwater use cases. Previous work done by Eawag using the ESA CCI Lakes dataset and DINEOF interpolation methods has shown, that reliable and gap-free outputs of Lake Surface Water Temperature can be achieved. Alternatively, this methodology could

[7] Odermatt et al. (2020) SenCast: Copernicus Satellitendaten auf Knopfdruck. Geomatik Schweiz, Géomatique Suisse, Geomatica Svizzera 118(9), 12-16.

[8] Alvera Azcarate, A., Barth, A., Rixen, M., & Beckers, J.-M. (2005) Reconstruction of incomplete oceanographic data sets using empirical orthogonal functions: application to the Adriatic Sea surface temperature. Ocean Modelling, 9 (4), 325-346.

be expanded and applied to other water variables to provide gap-free data for the SOS-Water project.

Lake ice cover provides a variety of ecosystem services. Therefore, the loss of ice cover is linked to the reduction in water quantity, degradation of water quality, reduction in dissolved oxygen and proliferation of algal blooms. Current efforts at Eawag focuses on the development of high-resolution lake ice monitoring based on the Copernicus Sentinel and NASA/USGS Landsat satellite suite. Making use of both optical and SAR-based EO data, this multi-sensor approach makes it possible to retrieve lake ice coverage and phenology with nearly daily resolution and down to lake sizes smaller than 10 ha. If lake ice is simulated as part of hydrological models in the SOS-Water project, EO-based lake ice coverage could be used as validation for simulated ice-on/off state.

Snow cover, accumulation and melt play an important role in the water exchange between the atmosphere and the soil. Current EO-based methods to retrieve the amount of water stored in snow estimate snow water equivalent (SWE) based on snow depth and density and are still limited in their capabilities and associated with high uncertainties. However, EO-based methods can be used to accurately map two-dimensional snow coverage with high spatial and temporal resolution. The extracted on/off state can be used to validate snow coverage simulated as part of hydrological models. Thus, existing snow cover products, such as MODIS-derived Snow Cover Fraction products from the Copernicus Land Monitoring Service, the ESA CCI Open Data Portal or the open-source tool SnowWarp[9], give an opportunity to validate the hydrological models used in this project.

2.2.3 In-situ data needs and mismatches

While EO data has many benefits in hydrological modelling, it is limited to the retrieval of a small number of water variables. In-situ data is generally considered more important as it provides direct and accurate measurements of a large number of hydrological variables directly at the location of interest. In comparison, remotely collected EO dataset are associated with higher uncertainties and retrieved measurements are generally available at much coarser temporal resolution. Thus, matching up in-situ data with EO-based estimates is a standard approach to calibrate and validate EO algorithms and products.

In the next subsections we want to explore the mismatches between in-situ data needs and current in-situ data availability. To do this we will focus on the Mekong Delta and the Danube Delta, two of the SOS-Water case studies.

Mekong Delta

The Mekong Delta has significantly faced salinity intrusion, floods, erosion, and sediment loss, water scarcity, land subsides, and biodiversity degradation due to anthropogenic drivers rather than climate change impact in the following three decades (Eslami et al., 2021). Significantly, compound effects driving climate change and man-made drivers could intensify consequences (e.g., water scarcity, flood hazards, crop damages). The hydrometeorological and environmental data captured from the in-situ monitoring networks play a crucial role in the water management. Notably, these data are inputs for assessing the region's status and investigating appropriate interventions to make

[9] Vaglio Laurin et al. (2022) SnowWarp: An open science and open data tool for daily monitoring of snow dynamics. Environmental Modelling & Software 156, 105477.

the region more climate-resilient. Here, we assess mismatches between the current in-situ data availability and the data demands of water management.

A Hydromet network for river monitoring, which is managed by the Mekong River Commission (MRC)[10], comprises 45 fully operational stations of hydrological cycle observing system (HYCOS), 63 water level stations and 127 rainfall stations in the whole lower Mekong River basin (stretching from the border of China and Lao PDR to the Vietnamese Mekong Delta). Discharge and sediment monitoring has occurred since 1957 and handed over to riparian countries (Thailand, Lao PDR, Cambodia, Vietnam) in the discharge and sediment monitoring program of MRC in 2009. The member countries are provided with equipment and skills to carry out high-quality sediment measurements. With 17 stations along the mainstream Mekong, and its main tributaries, some key parameters are collected: daily water level, weekly flow, suspended sediment, bed load, and sediment grain size. However, the monitoring activity was postponed from 2016 to mid-2018. Currently, this monitoring is done weekly from June to October on a yearly basis. Results are reported on a semi-annual or annual basis by line agencies of the member countries. A water quality monitoring network[11] has been running with 22 stations along the mainstream of the Mekong River. Water samples are collected at mid points of the Mekong mainstream every two months, starting from February, and collects about 21 parameters. River biomonitoring data was collected for a few years, such as in 2013, 2015, and 2017, according to the MRC[12] publications. For fisheries monitoring, a program has been supported by the MRC which monitors regional fisheries in three categories: (i) fish abundance and diversity monitoring (at 38 sites, since 2003), (ii) fish larvae drift monitoring (in Vietnam and Cambodia since 2000, and Lao PDR in 2020, Thailand in 2021), and (iii) bagnet (dai) fishery monitoring (on the Tonle Sap River in Cambodia, since 1995).

With regards to in-situ data monitoring networks in the Vietnamese Mekong Delta (VMD) managed by Vietnam agencies there are 13 stations fully recording climatic parameters and 60 rain stations in the whole the VMD. For hydrological stations, the discharge and sediment are mainly measured at five stations on the mainstreams, of which four stations (Tan Chau, Chau Doc, Can Tho, My Thuan stations) are part of the MRC HYCOS. In contrast, the water level is monitored at 40 stations with long-term monitoring covering the VMD. There are 35 stations managed by provincial authorities for water quality monitoring, which collect 15 basic parameters (temperature, pH, electrical conductivity, alkalinity/acidity, dissolved oxygen, chemical oxygen demand, total phosphorous, total nitrogen, ammonium, total nitrite and nitrate, faecal coliform, total suspended solid, calcium, magnesium, sodium, potassium, sulphate, chloride, BOD_5) quarterly. Furthermore, 39 permanent salinity stations are installed along the mainstream and in canals covering the coastal region. There are also over 250 stations for groundwater monitoring spreading over the VMD with the observed frequency of two to four times per year.

[10] https://portal.mrcmekong.org/monitoring/river-monitoring-telemetry

[11] https://portal.mrcmekong.org/monitoring/water-quality

[12] https://mekongmrc.org/resource/ajhypg

In general, the monitoring stations partially meet the data requirement for water management. However, there are still some mismatches between water management activities and the available water and water-related in-situ data.

- The density of hydrologic stations (4,140 km^2/station) in VMD is quite low compared to the density target of the World Meteorological Organization (1,875 km^2/station)[13].
- For water management, the hydrology monitoring stations provide sufficient information for the water management tasks (flood forecast, flood preparedness, water supply, etc). However, the discharge and sediment data are still insufficient in station density and monitoring frequency, especially in the floodplains, and canals. Despite procedures for Data and Information Exchange and Sharing among member countries adopted in the Mekong lower basin in 2001, approaching the operation information of water infrastructures (e.g., flow released, dam operation regimes, monthly and weekly dam operation planning, water amounts for irrigation) is challenging for downstream countries. Furthermore, the lack of operation information on dams in the Upper Mekong River in China is also a significant challenge to the downstream countries (e.g., Vietnam).
- For water quality monitoring, the MRC network has provided practical, long-term data for assessing the surface water quality status and inputting in the water resources planning in the Mekong region; however, most stations are located on the mainstream and there is sparse frequency of the water sample collections (two times a month). In contrast, local water quality monitoring stations are located in canals, but the sampling frequency is once every three months.
- There are 39 salinity stations along the coastal area, but most observations are during the dry season, and salt stratification measurements are not captured.
- The missing information and inconsistent monitoring on groundwater in transboundary aquifers between Cambodia and Vietnam also yields gaps in the water management.
- Although biomonitoring data has been provided baseline data for decision-making in the management of regional natural resources in the Lower Mekong Basin, limitations include a lack of monitoring sites in the VMD (8) and inconsistent timing in samplings in discrete years. Similarly, the fish monitoring locations are mostly located in the mainstreams.
- Numerous studies have found that land subsidence is a driving force to intensify flooding, salt intrusion, and land loss impacts in the VMD. Yet, there has been a lack of stations monitoring the land subsidence in the VMD. Most studies have employed water level stations, groundwater level stations/wells, InSAR (Interferometric Synthetic Aperture Radar) imagery, and numeric modelling methods to assess land subsidence rates.
- Transboundary plastic pollution has recently been revealed by a recent MRC study on the status and trends of riverine plastic pollution in the Lower Mekong River Basin[14]. Mekong river was ranked as the 10th contributor to marine plastic pollution worldwide. This issue

[13] World Meteorological Organization (2008) Guide to Hydrological Practices, Volume I: Hydrology – From Measurement to Hydrological Information.

[14] MRC. (2022). The status and trends of riverine plastic pollution in the Lower Mekong River Basin. Vientiane: MRC Secretariat.

could also threaten the aquatic environment and living organisms. Although the riverine plastic pollution was recognized by MRC member countries, there are still gaps in riverine plastic waste monitoring and regulations for its monitoring implementation.

Danube Delta

In the Danube Basin, upstream water management, energy production, and agricultural activities alter downstream flood regimes, sediment transport, and water quality, leading to a broad range of aggregated impacts in the delta region. For example, current efforts to mitigate sediment loss succeed by capturing sediments inside the delta through channelization. But conversely, they also favor increases in pollution and eutrophication levels and enhanced coastal erosion. This highlights the complex interactions in this region. Tensions especially exist between upstream users such as energy, and agricultural interests, and downstream interests such as environmental protection, tourism, and local communities. The National Management Plan for the Romanian part of the Danube International Basin is an important tool in managing the water resources in the Danube Delta. In order to better evaluate the tradeoffs in this basin, we assess the in-situ data availability versus the data needs for implementing the water management plan in the Danube delta.

Danube delta's water resources are managed by the National Administration of Romanian Waters (ANAR) through its regional authority (Dobrogea – Litoral Water Bazinal Administration – ABADL) together with other local authorities and research institutions. ANAR is also responsible for implementing the Water Framework Directive (WFD) 2000/60/CE in Romania through its National Management Plans. Therefore, water quality monitoring is done within the framework of the WFD to assess the ecological impact. Both Danube distributaries and lakes are monitored inside the delta, as well as the coastal waters (up to – 20 m isobath) in front of the delta. The monitoring program started in 2009 (with WFD implementation since 2015) and gradually developed over the years to, expand the number of parameters and monitoring stations. Currently the monitoring program reaches 15 river sections, 8 lakes and 5 coastal stations and publishes reports yearly[15]. Table 8 gives a comprehensive overview of the monitored parameters and sampling frequencies. These reports, however, lack any quantitative data on the measured parameters, and rather provide only qualitative data (e.g., percentage share of the five ecological status classes amongst the different water bodies).

[15] https://rowater.ro/despre-noi/descrierea-activitatii/managementul-european-integrat-resurse-de-apa/planurile-de-management-ale-bazinelor-hidrografice/planuri-de-management-ale-bazinelor-spatiilor-hidrografice-2016-2021/

Table 8 – Parameters and sampling frequencies for the Danube Delta

Quality elements		Parameters	Frequency		
			rivers	lakes	coastal
Biological elements	Phytoplankton	Species composition, density (cells/ml), biomass (mg/l)	2/year	4/year	4/year
	Phytobenthos	Species composition, density (cells/sample)	2/year	1/year	1/year
	Macrophytes	Species composition, Kohler abundance	1/3 years	1/3 years	NA
	Macro-invertebrates	Species composition, density (individuals/m^2)	2/year	1/year	1/year
	Fish	Species composition, density (individuals/m^2), length (mm/item)	1/3 years	1/3 years	1/3 years
Hydro-morpholo-gical elements	Hydrologic regime	Water level	2/day	1-30/month	1/day
		Discharge	20-60/year	NA	NA
		Wave exposure (for coastal waters)	NA	NA	4/year
	River continuity		1/6 years	NA	NA
	Morphological parameters	Depth and Width variation	1/year	1/6 years	1/ year
		Bed structure and texture	1/6 years	1/6 years	1/ year
		Riparian/bank zone structure	1/6 years	1/6 years	NA
Physico-chemical elements	Thermic conditions	Temperature	4/year	4/year	4/year
	Oxygen conditions	Dissolved oxygen - O_2, CCO-Cr, CBO_5	4/year	4/year	4/year
	Transparency	Secchi disk	4/year	4/year	4/year
	Salinity	Conductivity	4/year	4/year	4/year
	Acidification	pH	4/year	4/year	4/year
	Nutrients	N-NO_2, N-NO_3, N-NH_4, TN, TP, P-PO_4, chlorophyll-*a*	4/year	4/year	4/year
	Water specific pollutants	Heavy metals: Cu, Zn, As, Cr, Hg, Cd, Ni, Pb; Toluene, Naphthalene, Xylene, Phenol, PCBs (sum of 7), PAHs, Cyanides, Anionic-active detergents	4/year	4/year	4/year

Basic hydrological parameters: water level and discharge (daily measurements), suspended sediments, and navigable channel depth (1-4/year), are also monitored by the Lower Danube River Administration (AFDJ). From these measurements only the water level is reported on a daily basis to ensure safe navigation[16]. Despite of this, water level data is not archived. Water and sediment discharge were historically monitored in the Danube Delta before the establishment of ANAR and AFDJ. Annual means are available starting from 1840 at Ceatal Izmail (delta apex), 1921 at Sulina (river mouth) and 1960 at Sfantu Gheorghe (river mouth).

Fish resources are managed by the Danube Delta Biosphere Reserve Administration (ARBDD) and annual reports are published starting from 2016, however there are no quantitative data, nor a comprehensive monitoring approach being reported[17]. Since 2021, ARBDD is running a fish resource evaluation program where they are estimating the total catch allowed for the year to come for the Danube delta and the Black Sea[18]. ARBDD is monitoring tourism activities as well, and reports these activities and uses annually.

Water demand in the Danube delta is evaluated by ANAR through its regional authority ABADL while INHGA (National Institute for Hydrology and Water Management) forecasts the water demands for the different economic sectors (industry, irrigation, farming, aquaculture, and household use). The frequency of this evaluation is not clear. The last report was done in 2011 and based on data collected from 2007 – 2011. That said, water forecast demands were provided for 2020 and 2030, although it is unclear if these forecasted demands were accurate.

Climate data is provided by the National Administration of Meteorology (ANM) through its five monitoring stations within the delta on a daily (55 parameters) and monthly (376 parameters) basis. The oldest data record starts in 1878, while the latest starts in 1992[19].

In recent years Romania undertook important steps in implementing the WFD and developing a relevant assessment methodology for the Danube River Delta waters through new management plans. Nevertheless, mismatches between water in situ data and management actions still exist.

- Currently, there is no comprehensive, systematic, large scale monitoring scheme for the Danube Delta.
- Few institutes/institutions (ANAR, INHGA, ARBDD, etc.) collect data and report some basic parameters (such as water level, water temperature, tourism activities, and biodiversity) but most data are fragmented, inconsistent or not open-access. Moreover, there is no clear information on the monitoring location, timing, and sometimes methods or analytical procedures. The monitoring stations cover a minimum of river areas (only main distributaries), but are failing to address other wetland environment types such as channels, lakes or open-sea lagoons.

16 https://www.afdj.ro/ro/cotele-dunarii

17 https://ddbra.ro/rapoarte-anuale-de-activitate/

18 https://ddbra.ro/studii/

19 https://www.meteoromania.ro/catalog/

- There is no large-scale sediment management strategy. The Sulina river mouth is dredged by AFDJ, however the sediment is not reintroduced in the coastal system. Rather it is dropped at greater depths outside the wave redistribution zone which increases the erosional processes downdrift of the river mouth. Inside the delta, ARBDD is responsible for dredging the channel network, but the lack of strategy leads to clogged environmental fluxes and increased eutrophication in the delta lakes.
- Minimum to no collaboration between the above-mentioned institutions for a systematic in-situ measurement strategy along the Danube River and inside the Danube Delta.
- There is an urgent need of a continuous, high-frequency, large-scale monitoring program of the Danube Delta in terms of water quality parameters, both general (T, pH, O2, conductivity, alkalinity) and more specific (nutrients, CO2, CH4, N2O, TC, DIC, DOC, POC, PON, isotopic signature of C (d13C-DOC, d13C-POC, d13C-DIC), sediments concentrations and loads, heavy metals, etc.).

3 Conclusions

In this report, we provide a data inventory **with a total of 101 datasets** available for SOS-Water with a focus on project-relevant EO data. Each of the collected datasets were reviewed and complemented with key dataset attributes in terms of geographical and temporal scope, geographical and temporal granularity, data usage and accessibility. These attributes allow all project partners to quickly assess the potential usability that each of these datasets may have for the project.

From the gap analysis between EO data needs and data availability, we concluded the opportunities for development efforts for Task 3.2 listed in Table 9. Two key opportunities for the upcoming task and the SOS-Water project lie in the exploration of SWOT and Sentinel-3/6 altimetry for water level and runoff estimation and in Sentinel-based water quality retrieval.

Table 9 – List of hydrological variables and opportunities for development efforts as part of Task 3.2. For hydrological variables with no opportunities listed, data needs are mainly limited by currently available sensor technology.

Hydrological Variable	Opportunities for development efforts (Task 3.2)
Precipitation	• Exploration of spatial downscaling algorithms, • Improvement of coverage of high latitude/elevation areas • Harmonization with ground-based radar
Evaporation	• partitioning of evapotranspiration products to evaporation and transpiration
Soil moisture	
Runoff / Water level	• Exploration of SWOT and Sentinel-3/6 altimetry for water level and runoff estimation
Groundwater	
Vegetation	• Exploration of indices other than NDVI (e.g., SIF) • Retrieval of lake phytoplankton phenology
Snow, ice and permafrost	• Validation of hydrological models using snow cover and lake ice products
Water vapour	
Water quality	• Sentinel-based water quality retrieval • Gap-filling of existing products using DINEOF/DINCAE

Additionally, we explored and assessed key mismatches between the current in-situ data availability and the data demands of water management for the Mekong Delta (Table 10) and the Danube Delta (Table 11).

Table 10 – List of key mismatches between the current in-situ data availability and data demands of water management in the **Mekong Delta**.

Mismatch	Description
Low measuring station density	Low density (1,875 km^2/station) of hydrological stations compared to density target (4,140 km^2/station) of the World Meteorological Organization.
Lack of discharge and sediment data	Discharge and sediment data are insufficient in density and frequency, especially in the floodplains, and canals. Lack of operation information on upstream dams.
Limited measuring frequency	For water quality monitoring, measuring frequencies are generally limited to two-monthly in the mainstream and three-monthly in tributaries.
Limitations in salinity measurements	Most salinity observations are during the dry season, and salt stratification is not captured.
Lack of groundwater measurements	Missing information and inconsistent monitoring on groundwater in transboundary aquifers between Cambodia and Vietnam.
Lack of biomonitoring data	Lack of biomonitoring sites in the VMD and inconsistent timing in samplings in discrete years.
Lack of land subsidence data	Lack of stations for monitoring of land subsidence.
Lack of riverine plastic waste data	Gaps in riverine plastic waste monitoring and regulations for monitoring implementation.

Table 11– List of key mismatches between the current in-situ data availability and data demands of water management in the **Danube Delta**.

Mismatch	Description
No large scale monitoring system for Danube Delta	Currently, there is no comprehensive, systematic, large scale monitoring scheme in place for the Danube Delta.
Lack of collaboration	There is a general lack of collaboration between the involved institutions. As a result, no systematic in-situ measurement strategy along the Danube River and inside the Danube Delta is in place.
Data inconsistency and limited data accessibility	Most collected in-situ data are fragmented across institutions, are not open-access and have inconsistencies in associated methods, analytical procedures and reporting.

Coverage limited to main distributaries	Current measuring stations mainly cover main distributaries and fail to address wetland environment types such as channels, lakes or open-sea lagoons.
Lack of water quality data	There is an urgent need of a continuous, high-frequency, large-scale monitoring program of the Danube Delta in terms of water quality parameters.

4 Appendix A: Data Inventory

The Data Inventory is a table consisting of 101 project-relevant datasets compiled as part of Task 3.1 of the SOS-Water project. The datasets were reviewed and enriched with key dataset attributes in terms of geographical and temporal scope, geographical and temporal granularity, data usage and accessibility. Table A1 shows an example of one of the dataset entries.

Table A1 – Example entry from the Data Inventory for the ESA CCI Lakes v2.0.2 LSWT product

Dataset name	**ESA Lakes_cci: Lake products, Version 2.0.2, LSWT**
Variable	**Lake Surface Water Temperature**
Unit	Degrees Kelvin
Description	The LSWT product is part of the ESA CCI Lakes dataset and contains lake skin temperature. It is generated from ERS-2 (ATSR-2), Terra/Aqua (MODIS), Metop-A/B (ATSR-2) and Sentinel-3 (SLSTR) observations.
Spatial resolution	1 km at equator (1/120 arc-degrees)
Spatial extent	Global
Temporal resolution	Daily
Temporal extent	1995-04-21, 2020-12-31
URL	https://catalogue.ceda.ac.uk/uuid/a07deacaffb8453e93d57ee214676304
DOI	10.5285/a07deacaffb8453e93d57ee214676304
References	Carrea, L., Crétaux, JF., Liu, X. et al. Satellite-derived multivariate world-wide lake physical variable timeseries for climate studies. Sci Data 10, 30 (2023). 10.1038/s41597-022-01889-z
Data type	Raster
File format	NetCDF
Catchment	All
Type	Earth Observation
Data source	ESA CCI Open Data Portal
License	Free and open access (https://climate.esa.int/en/terms-and-conditions/)
Comments	Temporal coverage until 2006 relies on ERS-2 (ATSR-2) data only and is low as a result.

The full Data Inventory is **available upon request**.

Disclaimer

Acknowledgement of funding

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059264.