Integration of novel EO products into basin-scale water resources modelling

D3.3 Basin-scale EO data integration

Lead beneficiary	6 – FUTUREWATER SL	Due Date	30 September 2025
WP no	3	New due date (if delay)	
Task no	3	Actual Delivery Date	30 September 2025
Dissemination level	PU – Public	Status	FINAL

Authors

Authors	Partner no	Partner organisation	Name of author
Main author	6 6.1	FUTUREWATER SL FUTUREWATER-NL	Gijs Simons Sergio Contreras Tijmen Schults Corjan Nolet
Contributing author(s)	11	EAWAG	Daniel Odermatt Michael Brechbühler

Review

Authors	Partner no	Partner organisation	Name of author
Technical review	2	UU	Niko Wanders
Formatting review	1 9	IIASA EUT-RS	Silvia Artuso Mikael Muegge

Document history

Date	Version	Chapters affected	Description of change	Author	Document status
27/08/2025	1.0	All		Gijs Simons Sergio Contreras Tijmen Schults Corjan Nolet Michael Brechbühler	Draft for review by UU
09/09/2025	1.1	All	Minor modifications / comments accepted, executive summary	Gijs Simons Daniel Odermatt	Draft final version for clearance by internal coordination team
29/09/2025	1.2	All	Formatting review	Silvia Artuso	Final
21/01/2026	2.0	All	Minor editing improvements, high-resolution Figure 1 and 3, fixed references to figures	Gijs Simons	Final (revised), incorporating EC feedback

Publishable Executive Summary

This report documents the integration of novel Earth Observation (EO) products into basin-scale water resources modelling within the SOS-Water project. The objective is to explore how EO can enhance the Integrated Water Modelling System (IWMS) that underpins the project's Safe Operating Space (SOS) framework, improving the accuracy, robustness, and transparency of water resources assessments across multiple case studies.

Six EO prototypes were developed to address specific gaps in monitoring and modelling:

- Suspended Particulate Matter Mapper: generated spatially detailed sediment concentration datasets for the Mekong Basin, filling observational gaps and enabling basin-scale sediment flux estimates.
- Lake and River Ice Mapper: produced accurate, high-resolution ice cover and phenology records for Alpine regions, meeting key requirements for cryospheric monitoring and climate studies.
- Satellite-based altimetry: operationalized SWOT data for water level monitoring, supporting improved model calibration while highlighting site-specific limitations in complex environments.
- Surface Water Temperature Mapper: built a multi-decadal thermal archive from Landsat data, providing benchmarks for water temperature simulations and revealing biases in hydrological and water quality models.
- Crop Water Productivity Mapper: tailored WaPOR methodologies to the Júcar Basin, demonstrating the feasibility of locally adapted EO-based productivity assessments to support hydro-economic modelling.
- Snow Water Equivalent (SWE) Downscaling: developed a machine-learning-based approach to enhance SWE representation, supporting model calibration and benchmarking in snow-dominated basins.

Collectively, these prototypes demonstrate how EO can serve both as input data and as independent benchmarks for hydrological modelling. Their integration has improved model realism, constrained uncertainties, and supported the evaluation of sustainability indicators within the SOS framework. At the same time, the work underscores the need for recalibration and methodological adaptation when embedding EO into model workflows.

The outcomes of Task 3.3 highlight the immediate benefits of EO integration for the SOS-Water case studies and lay the groundwork for longer-term applications. The prototypes and methodological frameworks developed here are openly available and designed to support the wider hydrological community, contributing to more reliable and observation-driven water resources management in the face of climate and societal change.

Table of Contents

1 Background 8
1.1 Objective and scope 8
1.2 Task management 8
2 Requirements for integrating EO products for water monitoring into model systems 10
2.1 Introduction 10
2.2 EO for model input 11
2.2.1 Meteorological Forcing 11
2.2.2 Parameterization of Spatial Heterogeneity 11
2.2.3 States (initial and boundary conditions of the system) 12
2.3 EO for model benchmarking 13
2.3.1 Benchmarking fluxes 13
2.3.2 Benchmarking storage components 14
3 EO integration in SOS-Water 15
3.1 EO and the SOS 15
3.2 Overview of SOS-Water EO prototypes 17
4 EO prototypes and their integration 18
4.1 EOP1: SPM Mapper 18
4.1.1 Synergies with models 18
4.1.2 Activities 18
4.1.3 Results 19
4.1.4 Prospects 20
4.2 EOP2: Lake and river ice mapper 20
4.2.1 Synergies with models 20
4.2.2 Activities 21
4.2.3 Results 21
4.2.4 Prospects 22
4.3 EOP3: Satellite-based altimetry 23
4.3.1 Synergies with models 23
4.3.2 Activities 23
4.3.3 Results 24

4.3.4 Prospects....25
4.4 EOP4: Surface water temperature mapper....25
4.4.1 Synergies with models....25
4.4.2 Activities....26
4.4.3 Results....27
4.4.4 Prospects....28
4.5 EOP5: Crop water productivity mapper....29
4.5.1 Synergies with models....29
4.5.2 Activities....29
4.5.3 Results....30
4.5.4 Prospects....31
4.6 EOP6: Snow-water equivalent downscaling....32
4.6.1 Synergies with models....32
4.6.2 Activities....32
4.6.3 Results....33
4.6.4 Prospects....35
5 Conclusions....36

List of Tables and Figures

Table 1. Task 3.3 Task management overview. 9
Table 2. List of SOS-Water indicators, their relevance to each river basin, and the suitability of using EO to quantify them. 15
Table 3.Overview of Earth Observation Prototypes (EOPs) developed and tested under the SOS-Water project. 17
Table 4. Averaged pyWaPOR results for different crop types 31
Table 5. Performance metrics of different runs of direct forcing CwatM with downscaled SWE dataset 33

Figure 1. Validation panel for satellite-derived sediment flux estimates at Stung Treng station (014501). 19
Figure 2. Comparison of aggregated lake ice cover timeseries from EOP2 against the ARLIE product at the Silersee, Switzerland test site. 22
Figure 3. SWOT Interferometry provides high-resolution, two-dimensional measurements 23
Figure 4. SWOT data operationalization workflow 24
Figure 5. Time Series and Scatter Plot with Regression Line for Lobith, NE: Comparison of in-situ and observed data for Reach (top) and Node (bottom) 25
Figure 6. Comparison of observed remote-sensing and simulated DynQual stream temperatures (MPI). 27
Figure 7. Seasonal residual analysis between observed remote-sensing and simulated DynQual stream temperatures (MPI, SSP126, historical and projected). 28
Figure 8. Interannual Variability of NBWP computed with the adapted pyWaPOR 30
Figure 9. Calibration results of CWatM snowmelt using the downscaled SWE dataset 34
Figure 10. Average time series of SWE test predictions against observations at 383 stations 35

Abbreviations

ARLIE	Aggregated River and Lake Ice Extent
ASCAT	Advanced Scatterometer
CWATM	Community Water Model
DEAP	Distributed Evolutionary Algorithms in Python
DSMP	Discharge Sediment Monitoring Project
DYNQUAL	Dynamic Water Quality Model
ECV	Essential Climate Variable
EO	Earth Observation
EOP	Earth Observation Prototype
ERA5	ECMWF Reanalysis v5

ET	Evapotranspiration
FAO	Food and Agriculture Organization
GCOS	Global Climate Observing System
GCM	Global Climate Model
GRACE	Gravity Recovery and Climate Experiment
GRACE-FO	GRACE Follow-On
GRU-D	Gated Recurrent Unit – Decay
HR-S&I	High-Resolution Snow and Ice
IWMS	Integrated Water Modelling System
KGE	Kling-Gupta Efficiency
LSTM	Long Short-Term Memory
LSWT	Lake Surface Water Temperature
MERRA	Modern-Era Retrospective analysis for Research and Applications
MODIS	Moderate Resolution Imaging Spectroradiometer
NBWP	Net Biomass Water Productivity
NPP	Net Primary Productivity
NSE	Nash-Sutcliffe Efficiency
PCR-GLOBWB2	PCRaster Global Water Balance, version 2
PLD	Prior Lake Database
RLIE	River and Lake Ice Extent
SAR	Synthetic Aperture Radar
SIGPAC	Sistema de Información Geográfica de Parcelas Agrícolas
SOS	Safe Operating Space
SPM	Suspended Particulate Matter
SSP	Shared Socioeconomic Pathway
SWE	Snow Water Equivalent
SWOT	Surface Water and Ocean Topography
SWORD	SWOT River Database
SWT	Surface Water Temperature
VMD	Vietnamese Mekong Delta
WaPOR	Water Productivity through Open-access of Remotely sensed data
WP	Work Package
WUE	Water Use Efficiency
WMO	World Meteorological Organization
XGBOOST	Extreme Gradient Boosting

1 Background

1.1 Objective and scope

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 and impact models with a focus on agricultural and energy production, drinking water supply and biodiversity.

Earth Observation (EO) is considered an important source of information to feed and enhance the IWMS, by providing spatial information and time series of critical environmental variables. Within the project, a comprehensive inventory was compiled of current state-of-the-art EO data products that satisfy hydrological modelling requirements (SOS-Water D3.1)[1]. Although a wealth of information is available and continuous developments and improvements are being made by the global EO community, the inventory also informed the identification of existing gaps and research opportunities to further leverage the potential of EO to serve the modelling community. This finding, together with the specific requirements of the IWMS in the context of the SOS-Water case study basins, prompted the development of a set of EO-based Prototypes (EOPs) which together comprise SOS-Water D3.2.[2]

In SOS-Water Task 3.3, the developed EOPs were linked to the IWMS and integrated into basin-scale water resources modelling. The nature, and extent, of this integration differs per EOP, model and case study, and is described in this report.

In this way, the purpose of this document (D3.3) is:

- to provide an overview of how EO is able to support and enhance water resources modeling efforts;
- to explore and discuss the potential of each SOS-Water EOP to serve the wider hydrological modelling community;
- to demonstrate how each SOS-Water EOP links to the IWMS and describe the activities that were undertaken.

1.2 Task management

Table 1 outlines the distribution of work across the project partners involved in Task 3.3.

[1] Data inventory and EO data needs for water resources monitoring, SOS-Water Deliverable Report, 2023.

[2] EOPs are publicly available through the SOS-Water Work Package 3 GitHub: https://github.com/mibrechb/SOSW_WP3.

Table 1. Task 3.3 Task management overview.

Partner organization	Partner no	Description of work
FutureWater	7	Task lead and coordination of D3.3, development of EO prototypes, supporting their implementation and connection to the IWMS
EAWAG	6	Development of EO prototypes, supporting their implementation and connection to the IWMS
IIASA	1	Modeling to explore and test implementation of EO prototypes
UU	2	Modeling to explore and test implementation of EO prototypes

2 Requirements for integrating EO products for water monitoring into model systems

2.1 Introduction

The integration of Earth Observation (EO) data into hydrological and water resources models has emerged as a critical approach to improving the accuracy and efficiency of water resource management, particularly in regions where traditional ground-based data collection methods are limited or costly. With the growing demand for sustainable water management and climate adaptation strategies, the need for high-resolution, comprehensive, and real-time data is more pressing than ever. EO products, derived from satellites, provide invaluable insights into various hydrological variables such as precipitation, soil moisture, evapotranspiration, and snow cover, which are fundamental to water resources modelling.

Hydrological models rely on accurate data to simulate water flow, storage, and distribution across river basins. However, many regions face challenges in obtaining detailed in-situ data due to factors such as remote locations, political barriers, or financial constraints. EO data, with its global coverage and regular updates, offers a cost-effective solution to these data gaps. Moreover, satellite-based products can provide continuous monitoring over large areas, enabling the assessment of spatial and temporal variations in water resources that are crucial for effective management.

The rationale for linking EO and hydrological models is therefore grounded in the need for robust, data-driven tools that can effectively address the complex challenges faced by water resources management in a changing climate. EO data provides a rich source of information that, when integrated with hydrological models, can significantly improve the understanding of hydrological processes, enhance model predictions, and support more informed decision-making in water resources planning and management. The research domain of EO to support hydrological modeling is very much an evolving field, as many challenges exist when it comes to streamlining spatial and temporal resolutions. Cutting-edge applications being explored include the integration of EO into Digital Twin Earth models[3], and supporting hyper-resolution global modelling[4], among many others. A comprehensive overview of existing EO products for the hydrological modeling community, organized by variable of interest, has previously been compiled under the SOS-Water project[5]

[3] Brocca L, Barbetta S, Camici S, Ciabatta L, Dari J, Filippucci P, Massari C, Modanesi S, Tarpanelli A, Bonaccorsi B, Mosaffa H, Wagner W, Vreugdenhil M, Quast R, Alfieri L, Gabellani S, Avanzi F, Rains D, Miralles DG, Mantovani S, Briese C, Domeneghetti A, Jacob A, Castelli M, Camps-Valls G, Volden E and Fernandez D. A Digital Twin of the terrestrial water cycle: a glimpse into the future through high-resolution Earth observations. Front Sci (2024) 1:1190191. doi: 10.3389/fsci.2023.1190191

[4] van Jaarsveld, B., Wanders, N., Sutanudjaja, E. H., Hoch, J., Droppers, B., Janzing, J., van Beek, R. L. P. H., and Bierkens, M. F. P.: A first attempt to model global hydrology at hyper-resolution, Earth Syst. Dynam., 16, 29–54, https://doi.org/10.5194/esd-16-29-2025, 2025

[5] Data inventory and EO data needs for water resources monitoring, SOS-Water Deliverable Report, 2023

2.2 EO for model input

Hydrological models rely on a combination of time-varying climate inputs, spatially distributed catchment characteristics, and information on current water storage and fluxes. Traditionally, these inputs have come from in-situ networks, which remain sparse or absent in many regions. Satellites now routinely deliver globally consistent information on the land surface, atmosphere, cryosphere, and water bodies at spatial and temporal resolutions that make them highly valuable for driving, parameterizing, and constraining hydrological simulations. EO products can serve as input to model simulations, provide input on key model parameters, improving predictive capabilities for water availability, drought monitoring, flood forecasting, and climate change impact assessments. EO data can be directly ingested by hydrological models or used to improve model parameterization and initialization. Although technically a step beyond "inputs," EO products can also be continuously assimilated into model states (e.g., soil moisture, snowpack, water levels) to constrain model trajectories and reduce errors from uncertain forcing data.

EO inputs can be grouped into three main categories: (i) **forcings, (ii) parameters, and (iii) states**.

2.2.1 Meteorological Forcing

The term forcing relates to the time-varying drivers of the water cycle. Hydrological models require meteorological variables as continuous inputs. EO products fill this role by providing near-real-time, spatially distributed estimates of the main fluxes driving the water cycle:

- Precipitation: Satellite-based rainfall products such as TRMM, GPM, and CMORPH supply rainfall intensity and distribution at regional to global scales. This is particularly valuable in regions without dense ground-based rain gauge networks, where precipitation uncertainty is often the dominant source of hydrological model simulation error.
- Radiation and temperature: EO products (e.g., MODIS, AIRS) and reanalysis products that blend EO with modelling provide inputs for evapotranspiration and snowmelt calculations. These forcings influence key processes such as potential evapotranspiration, crop water use, and snowpack dynamics.
- Evapotranspiration (ET) estimates: Satellite-derived evapotranspiration from MODIS, Landsat, or Sentinel can act as a substitute where meteorological data are lacking, or as a complementary forcing dataset to constrain energy balance models. Satellite-derived ET can serve either as direct forcing or as benchmark for model calibration (see Section 1.3).

2.2.2 Parameterization of Spatial Heterogeneity

Model parameters are the properties of the landscape that define hydrological behavior. Whereas forcings drive hydrological processes, parameters define the physical template of the basin. EO datasets offer spatially continuous maps of the characteristics that shape processes such as runoff generation, infiltration, and evapotranspiration. Major ones include:

- Land cover and vegetation: Products from MODIS, Sentinel-2, and Landsat classify vegetation type, canopy density, and phenology. These attributes affect interception, rooting depth, and transpiration, and are particularly important in ecohydrological and land surface models.

- Topography: Digital elevation models (SRTM, TanDEM-X, ASTER) capture slope, flow direction, and catchment boundaries. These parameters define runoff routing, drainage density, and orographic influences on precipitation.
- Soil properties: Global soil maps informed by EO (e.g. SoilGrids[6], HiHydroSoil[7]) provide estimates of soil texture, porosity, and hydraulic conductivity, which control infiltration capacity and soil-water retention.
- Cryospheric features: EO products mapping glaciers, snow cover, and ice thickness (from MODIS, Sentinel-1, ICESat-2) supply critical parameters for high-altitude or polar catchments where seasonal snow and ice melt dominate streamflow.

Together, these datasets help capture (sub-)basin heterogeneity that strongly affects hydrological response, especially in distributed or hyper-resolution models.

2.2.3 States (initial and boundary conditions of the system)

Hydrological models also require information about the current condition of the system. EO data provide direct observations of key state variables, which can initialize models, reduce spin-up time, or serve as boundary conditions. Using EO-derived states helps ensure that hydrological models begin simulations from realistic initial hydrological conditions, reducing uncertainty in streamflow forecasts and water balance assessments.

Examples include:

- Soil moisture: Microwave missions such as SMAP, SMOS, Sentinel-1 and ASCAT measure near-surface soil moisture at global scales. These datasets are frequently used to initialize soil water profiles, improving model stability and predictive skill.
- Snow and ice storage: MODIS snow cover, Sentinel-1 snow water equivalent, and ICESat-2 elevation change products characterize snowpack and glacier storage. Accurate representation of these states is vital for snowmelt-dominated catchments.
- Terrestrial water storage: GRACE and GRACE-FO satellites provide basin-scale estimates of groundwater and total water storage anomalies. These products offer unique insights into groundwater depletion and seasonal storage changes, which are otherwise difficult to observe.
- Surface water dynamics: Altimetry (Sentinel-3), radar (Sentinel-1), and SWOT mission data provide water level and surface extent of rivers, lakes, and reservoirs. These observations can define inflow/outflow conditions and capture human-water interactions such as dam operations.

[6] https://www.isric.org/explore/soilgrids

[7] Simons, G.W.H., R. Koster, P. Droogers. 2020. HiHydroSoil v2.0 - A high resolution soil map of global hydraulic properties. FutureWater Report 213.

2.3 EO for model benchmarking

The synergy between EO and hydrological models extends beyond data input. Since EO captures variables that are typically also simulated by water resources models, such as actual evapotranspiration/latent heat, water levels, snow cover and water equivalent (SWE) or soil moisture, these datasets can be consulted for benchmarking of (specific modules of) water resources models. This can be done with the purpose of quantitatively assessing model performance, and/or performing a calibration process where certain model parameters are tuned to achieve a better alignment between model output and EO data. Here, the uncertainties associated with EO data should always be considered.

Most hydrological models are traditionally calibrated and validated against river discharge time series. However, discharge alone does not uniquely constrain all processes— equifinality can occur, meaning that different parameter sets can produce similar streamflow. EO products provide additional, independent benchmarks that help distinguish between competing model structures or parameterizations by testing whether fluxes and storage components simulated by the model are realistic.

By providing independent benchmarks for fluxes and storage, EO data enable multi-variable evaluation of hydrological models. This improves diagnostic power: for example, a model that reproduces discharge but fails to match observed soil moisture or evapotranspiration may be "right for the wrong reasons." EO also supports multi-scale benchmarking, from local field scales (e.g., Sentinel-2 ET) to continental or global scales (e.g., GRACE storage anomalies). Using EO data for benchmarking encourages the development of models that capture the internal water balance more correctly, helping to improve model transferability to ungauged basins and increase confidence in projections under climate and land-use change.

2.3.1 Benchmarking fluxes

- Evapotranspiration: ET can be inferred from thermal imagery, using algorithms based on surface energy balance theory. Benchmarking against satellite-derived ET products (e.g., MODIS, Landsat, WaPOR[8]) allow modelled evaporation and transpiration to be compared against independent estimates. This helps assess whether models partition water correctly between evaporation and runoff. Often, models make assumptions with regard to water use parameters per land cover type or sector (e.g. irrigation efficiency or crop water requirements). Such assumptions can be tested by comparing with EO-derived ET information.
- Snow dynamics: snow cover/SWE estimates (MODIS, Sentinel-1, ICESat-2) provide checks on the snow dynamics (build-up and melting of the snowpack) simulated by the model.
- Surface Water Fluxes: Flood extent and inundation dynamics, for example from SAR (Sentinel-1) or altimetry (SWOT), can be compared with modelled floodplain storage and model routing performance.

[8] https://data.apps.fao.org/wapor/

2.3.2 Benchmarking storage components

- Soil Moisture: Global soil moisture datasets from SMAP, SMOS, and ASCAT provide direct benchmarks for modelled root-zone soil water dynamics. Agreement with these observations is an indicator of process realism in infiltration and evapotranspiration.
- Snowpack and Glaciers: Remote sensing of snow cover duration, glacier mass balance, and elevation changes provide benchmarks for cryosphere-influenced basins, reducing uncertainty in snowmelt timing and glacier runoff contributions.
- Groundwater and Total Water Storage: GRACE[9] and GRACE-FO terrestrial water storage anomalies enable basin-scale benchmarking of groundwater and long-term storage changes—variables that are rarely monitored in situ but are critical for model realism. These data are relatively coarse but can be useful for large basins.
- Surface Water Levels and Volumes: Altimetry and SAR-derived water levels (e.g., Sentinel-3, SWOT) allow models to be benchmarked against observed lake, reservoir, and river storage. This can help for example to derive reservoir operational rules, which often need to be assumed in hydrological models due to a lack of information.

[9] Sutanudjaja, E. H., van Beek, R., Wanders, N., Wada, Y., Bosmans, J. H. C., Drost, N., van der Ent, R. J., de Graaf, I. E. M., Hoch, J. M., de Jong, K., Karssenberg, D., López López, P., Peßenteiner, S., Schmitz, O., Straatsma, M. W., Vannametee, E., Wisser, D., and Bierkens, M. F. P.: PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model, Geosci. Model Dev., 11, 2429–2453, https://doi.org/10.5194/gmd-11-2429-2018, 2018.

3 EO integration in SOS-Water

3.1 EO and the SOS

Implementation of the framework for Safe Operating Space for water systems (SOS-Water) requires information on a range of indicators, together describing the physical and socio-economic processes that define sustainability of water resources management in a given river basin. The SOS-Water concept relies on models to evaluate these indicators across various scenarios and management pathways. SOS-Water uses an integrated water modelling system (IWMS) to represent all natural and socio-economic processes of the water system, including ecological, environmental, agriculture, and energy aspects. the IWMS developed in the SOS-Water project consists of two components: 1) global hydrology models (i.e., model that assess the impacts of climate variability and change and human activities on water resources availability) and 2) the local impact models (i.e., models that assess the impact of water availability constraints on the economy, society, and environment).

Although EO by definition provides historical information, it plays an important role in the diagnostics and improvement of the performance of models to assess the required indicators. SOS-Water Deliverable 5.2[10] reports on the evaluation of the SOS-Water in the project case study areas and defines the indicators that have been identified as relevant in each particular context, through a comprehensive assessment involving stakeholder consultations, literature review, and modeling efforts. Table 2 categorizes these indicators, highlights their relevance in each case study and presents an assessment of the extent to which they can be quantified with EO data.

Table 2. List of SOS-Water indicators, their relevance to each river basin, and the suitability of using EO to quantify them. ++ = suitable (can be reliably derived or monitored with satellite data), + = partially suitable (can be approximated or supported with satellite-derived proxies, but needs in-situ data too), - = not suitable (satellites cannot directly quantify; requires ground-based monitoring or modeling)

Indicator	Júcar	Danube	Rhine	Upstream Mekong	Mekong Delta	EO suitability
Groundwater levels in key aquifers	✔					-
Streamflow at basin control points	✔	✔	✔	✔	✔	-
Reservoir storage levels	✔	✔	✔	✔	✔	++ (altimetry, imagery)
Ecological status of water bodies (WFD criteria)	✔	✔	✔			+ (chlorophyll, turbidity, extent)
Salinity intrusion in delta areas					✔	+ (indirect proxies color, temperature)

[10] SOS-Water Deliverable Report 5.2, Final Evaluation of Safe Operating Space, 2025

Indicator	Júcar	Danube	Rhine	Upstream Mekong	Mekong Delta	EO suitability
Sediment transport at key stations		✔	✔	✔	✔	++ (turbidity, suspended sediment concentration)
Flood frequency and extent	✔	✔	✔	✔	✔	++ (flood mapping from SAR/optical)
Drought frequency and severity	✔	✔	✔	✔	✔	++ (soil moisture, vegetation stress, precipitation proxies)
Nutrient concentrations (N, P)	✔	✔	✔	✔	✔	+ (chlorophyll /trophic state proxies, but not direct)
Pollutant loads (organic, heavy metals, pesticides)	✔	✔	✔	✔	✔	-
Land use / land cover change (remote sensing)	✔	✔	✔	✔	✔	++ (optical & radar land cover monitoring)
Groundwater abstraction volumes	✔	✔			✔	-
Water demand (irrigation, urban, industry)	✔	✔	✔	✔	✔	+ (crop area from satellites, but withdrawals require models)
Hydropower generation		✔	✔	✔		-
Fish catch / fisheries productivity				✔	✔	- (can map habitat change, not catch)
Glacier melt contribution to river flow				✔		++ (glacier area, snow cover, mass balance)
Riverine and delta erosion / subsidence rates					✔	++ (shoreline change, subsidence via InSAR)

3.2 Overview of SOS-Water EO prototypes

Based on the requirements of the models employed in the IWMS, a number of "Earth Observation Prototypes" (EOPs) have been developed as part of SOS-Water. These prototypes were selected based on their expected added value in the process of model benchmarking and calibration, thus supporting the models' performance in assessing the indicators listed in Table 2. Moreover, they are also expected to serve a broader purpose beyond the SOS-Water project and have been made available to the hydrological community as a whole[11]. Table 3 lists all EOPs developed as part of SOS-Water.

Table 3.Overview of Earth Observation Prototypes (EOPs) developed and tested under the SOS-Water project.

ID	Title	Description	Case study
EOP1	Suspended Particulate Matter (SPM) Mapper	EO tool for suspended sediment estimation	Mekong
EOP2	Lake and River ice mapper	EO tool for extraction of ice coverage and phenology	Rhine, Danube
EOP3	Satellite-based altimetry	EO tool for water level estimation	Rhine, Mekong
EOP4	Surface water temperature mapper	EO tool for extraction of surface water temperature	Rhine
EOP5	Crop water productivity mapper	EO tool for estimation of crop water productivity	Jucar
EOP6	Snow-water-equivalent (SWE) downscaling	EO tool for downscaling of SWE	Danube, Rhine

[11] https://github.com/mibrechb/SOSW_WP3

4 EO prototypes and their integration

4.1 EOP1: SPM Mapper

4.1.1 Synergies with models

Suspended particulate matter (SPM), referring to the main sediment load - fine particles (clay, silt, fine sand) carried within the water column, is a key variable for understanding sediment transport, nutrient delivery, and morphological change in river basins. In the Mekong River Basin, sediment transport is essential for sustaining the fertility of floodplains and the stability of the delta. Over recent decades, the construction of numerous reservoirs has significantly altered the natural sediment regime, reducing downstream sediment supply and impacting agricultural and ecological systems. In D3.1 we highlighted the lack of spatially extensive and temporally resolved sediment observations as a main limitation for modelling sediment dynamics and assessing the impacts of infrastructure development on sediment budgets. The existing monitoring network is limited in coverage and frequency, leaving large spatial and temporal gaps in the sediment record, which in turn makes it difficult to directly quantify the contributions of sub basins and effects of reservoirs.

An EO-based SPM mapper addresses these shortcomings by providing fine-resolution, spatially extensive sediment concentration estimates that complement and extend existing monitoring data. When combined with hydrological simulations of river discharge, these EO observations can be used to derive sediment-discharge relationships and estimate sediment fluxes across the basin. Beyond filling observational gaps, the EO-based approach also provides an alternative to, and potential improvement of, simulation-driven sediment modelling frameworks such as CASCADE[12], by offering direct observational constraints on suspended sediment concentrations.

4.1.2 Activities

The SPM prototype developed in Task 3.2 uses optical satellite imagery to retrieve suspended particulate matter concentrations at high spatial resolution. Multi-sensor datasets from Landsat-5 TM, Landsat-7 ETM+, Landsat-8/9 OLI, and Sentinel-2 MSI are atmospherically corrected and processed to remove clouds, shadows, and mixed shoreline pixels, with dynamic water detection applied to delineate water surfaces. Suspended sediment concentrations are then estimated using a retrieval algorithm calibrated for the Mekong River Basin with in-situ measurements from the Mekong River Commission covering the full basin and from project partner SWIRP (partner No. 8) for the Vietnamese Mekong Delta (VMD). This calibration ensures that the retrieval accounts for the specific optical and sediment characteristics of both upstream and deltaic waters. The result is a time series of SPM concentration maps at up to 30 m resolution, with increasing temporal coverage in recent years due to the addition of Sentinel-2 observations.

One of the main challenges in applying the EO-based sediment flux framework is the limited temporal density of optical satellite observations, particularly during the cloudy wet season when sediment fluxes are highest. As investigated during prototype development, the current multi-sensor

[12] Tangi, M., Schmitt, R. J. P., Bizzi, S. & Castelletti, A. (2019). The CASCADE toolbox for analyzing river sediment connectivity and management. Environmental Modelling & Software. [119. 10.1016/j.envsoft.2019.07.008.]

configuration achieves a combined revisit frequency of about 6.6 days in the dry season and 10.4 days in the wet season. To address these temporal gaps, the integration framework employs sediment rating curves that link remotely sensed concentrations with simulated discharge, enabling daily discharge-driven interpolation of SPM values. This approach reduces the risk of missing large fluxes during periods of poor optical coverage and ensures that sediment transport estimates better capture seasonal dynamics.

During the reporting period, this prototype was applied to the Mekong River Basin to generate spatially extensive SPM concentration datasets for integration with hydrological model outputs. The satellite-derived observations were combined with daily discharge simulations from the VIC-Res[13] hydrological water management model to establish sediment–discharge rating curves for key stations along the Mekong mainstem. These data were then used to produce continuous daily suspended sediment concentration and flux timeseries. This integration step, carried out jointly with POLIMI (partner No. 4), marks the first operational use of the sediment flux framework within SOS-Water.

4.1.3 Results

An initial benchmarking of the EO-based sediment flux framework was carried out at the Stung Treng station, Cambodia (ID 014501), located at the confluence of the 3S Basin and the Upper Mekong. Here, sediment fluxes were estimated by combining VIC-Res simulated discharge with EO-derived SPM concentrations and were compared against interpolated monthly in-situ measurements from the Sediment Monitoring and Discharge Measurement (DSMP) program. The validation, illustrated in Figure 1, shows reasonable agreement between the EO-based flux estimates and the ground-based observations, both in terms of magnitude and seasonal variability. This demonstrates that integrating EO-derived SPM concentrations with hydrological model outputs can bridge existing observational gaps and support basin-scale sediment budget assessments.

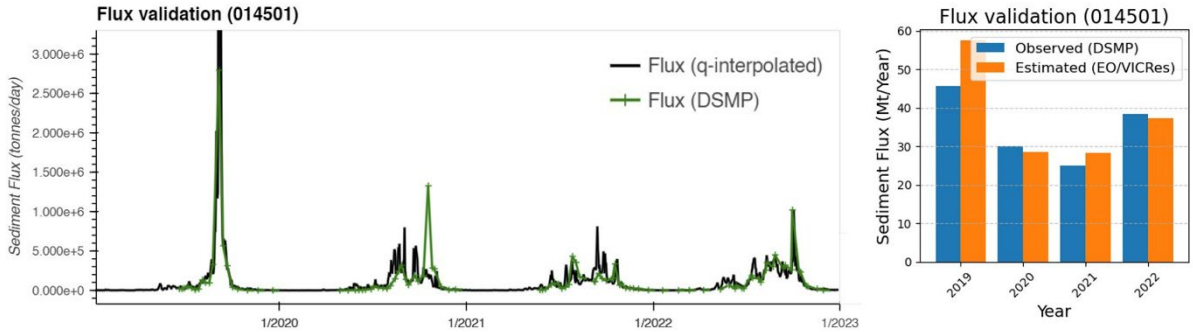

Figure 1. Validation panel for satellite-derived sediment flux estimates at Stung Treng station (014501). Fluxes are calculated as product of VICRes discharge and stallite-derived suspended sediment concentrations. Validation is based on interpolated monthly insitu measurements of the Sediment Monitoring and Discharge Measurement (DSMP).

[13] Eldardiry, H., Mahto, S. S., Fatichi, S., & Galelli, S. (2025). VIC-Res Mekong: Hydrological Model for the Mekong River basin (2.0). Zenodo. https://doi.org/10.5281/zenodo.15635798

4.1.4 Prospects

Looking ahead, further improvements of both the SPM mapper and the VIC-Res discharge simulations are underway, with the objective of increasing the accuracy and robustness of sediment flux estimates. Together with POLIMI (partner No. 4), work has started on extending the integration framework to reservoir-scale applications. Building on an inventory of reservoirs across the Mekong, a representative subset of sites has been identified that captures the diversity of storage capacity, inflow regimes, and geographic settings. This will allow us to test the feasibility of estimating sediment trapping efficiency in a systematic way, ultimately moving from single-station benchmarking towards basin-wide reservoir impact assessments.

As these developments progress, the combined EO-model framework is expected to provide an increasingly powerful tool for understanding sediment transport processes under regulated and naturalized flow conditions. This integration offers the potential to improve model-based sediment budget assessments, assess the impacts of sediment trapping in reservoirs, and ultimately support the development of management strategies for sediment-sensitive river systems like the Mekong.

Future applications could also explore replacing modelled discharge inputs with satellite-derived discharge estimates from missions such as the Surface Water and Ocean Topography (SWOT) mission. Integrated with sediment concentration retrieved from optical sensors, such as the SPM mapper, this could provide a fully EO-based framework for estimating sediment fluxes. However, during the reporting period, postponed data delivery meant that no feasible SWOT discharge products were yet available for testing. In addition, the limited temporal density of SWOT acquisitions may significantly constrain its usefulness for sediment flux applications. It should also be noted that SWOT-derived discharge is not entirely independent of modelling, as it relies on external hydraulic geometry models and ancillary ground data to convert satellite-measured elevations, widths, and slopes into discharge estimates.

4.2 EOP2: Lake and river ice mapper

4.2.1 Synergies with models

Lake and river ice dynamics are important variables for both hydrological and ecological systems. Ice cover is directly linked to the timing of seasonal hydrological flows and influences thermal stability and mixing regimes. In rivers, ice formation can alter flow conditions and cause ice jams that affect discharge dynamics, while in lakes, freeze cycles stabilize winter inverse stratification and regulate the length of the stratification period. In Deliverable D3.1 we identified the lack of spatially extensive and long-term ice cover and phenology records as a significant data gap. Existing in-situ datasets are scarce and fragmented, particularly in Alpine regions. An EO-based ice mapping approach can fill these gaps and provide extensive, consistent measurements of ice cover and ice phenology that can be used to gain insights into ice dynamics against their climatic drivers or to benchmark hydrological models.

A key challenge is that the large-scale hydrological models used in the SOS-Water project do not simulate water ice conditions explicitly, which limits the direct possibilities for assimilation or benchmarking. For the time being, the produced EO tool and dataset therefore serve primarily as an independent observation product. Due to its critical linkage to the Earth's climate, lake ice cover has

been introduced by the Global Climate Observing System (GCOS) as part of the Essential Climate Variables (ECV). In the requirement report of 2022, GCOS redefined the goal requirements for lake ice coverage measurements to daily observations at 50 m resolution and 1% uncertainty (WMO, 2022). The developed prototype can help move towards meeting these requirements.

4.2.2 Activities

The ice mapping prototype developed in D3.2 integrates Sentinel-1 SAR, Sentinel-2 MSI, and Landsat-7/8/9 imagery into a multi-sensor machine learning framework, trained and validated with over 40,000 webcam images from 21 lakes across the European Alps. This approach produces temporally dense ice cover maps, as well as aggregated phenological indicators such as freeze-up and break-up dates. Transferability to river ice has also been demonstrated, showing the potential for broader cryospheric monitoring across basins.

During the D3.3 reporting period, the prototype was further enhanced and evaluated. Methodological improvements focused on extending the temporal coverage to the 2015–2025 period, thereby capturing a full decade of ice conditions. Additional refinements strengthened the robustness of the classification and improved its generalization across lakes and seasons. Alongside the XGBoost framework, exploratory work was carried out with deep learning architectures such as Long Short-Term Memory (LSTM) and GRU-D networks to test their potential for improving temporal modelling and sensor harmonization. Together, these developments move the prototype closer to the GCOS goal requirements. The methodological advances and key outputs are being documented in a forthcoming manuscript, “High-resolution lake ice monitoring using multiple Earth observation missions” (Brechbühler, Wulf, Tom, Odermatt, in preparation).

4.2.3 Results

The EO-based ice mapper demonstrated high accuracy in retrieving lake ice coverage and associated phenological events. Validation against webcam observations across 21 Alpine lakes showed mean deviations of only –0.17 ± 4.3 days for freeze and break-up events. Classification accuracy, measured as mean intersection over union – binary Jaccard score, reached 98% for Landsat and Sentinel-2, and 85% for Sentinel-1, with good generalization to unseen lakes confirmed by leave-one-lake-out cross-validation. Validation has shown clear improvements over existing multi-sensor datasets, such as the Copernicus River and Lake Ice Extent (RLIE), particularly in terms of signal quality, marking a clear improvement and a step towards the GCOS goals.

Figure 2 shows a comparison between aggregated lake ice cover from the developed prototype and the Copernicus pan-European High-Resolution Snow and Ice service (HR-S&I) Aggregated River and Lake Ice Extent (ARLIE) product for Silsersee, Switzerland. The ARLIE is a comparable product that relies on Sentinel-1 C-SAR and Sentinel-2 MSI data processed with backscatter threshold approach and minimal distance classifier, respectively. The two datasets achieve a comparable temporal density, with a mean revisit frequency of 1.7 days for our prototype and 1.4 days for ARLIE. However, the prototype strongly outperforms ARLIE in signal fidelity, reaching an R^2 of 0.91 against fully frozen and unfrozen ground truth states, compared to only 0.03 for ARLIE. This demonstrates that, while temporal

density is comparable, the prototype delivers a higher-quality signal, representing a clear step towards reliable lake ice monitoring.

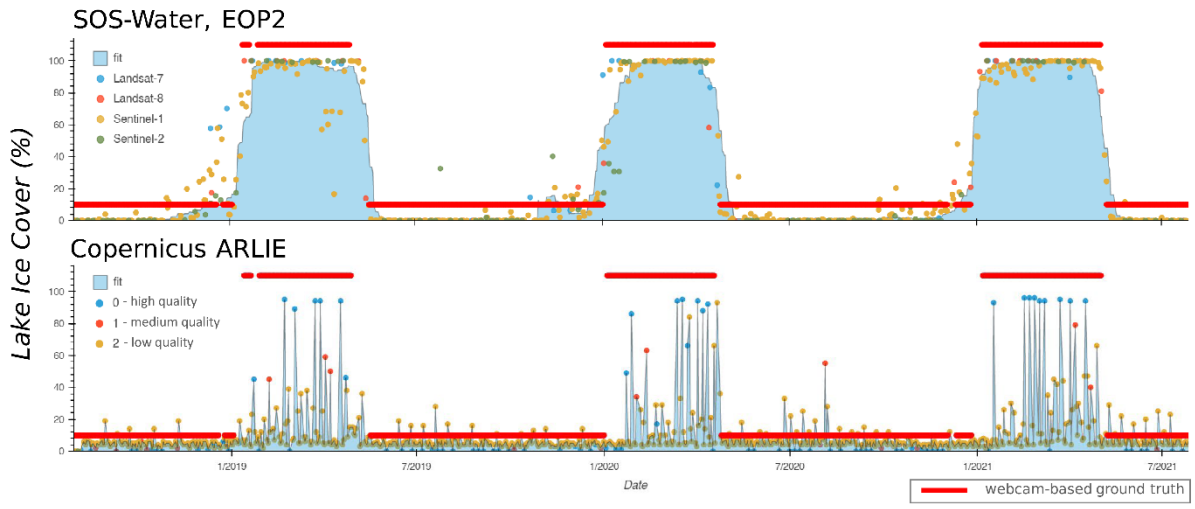

Figure 2. Comparison of aggregated lake ice cover timeseries from EOP2 against the ARLIE product at the Silersee, Switzerland test site. The red line indicated the ground truth states (offset for visibility) for fully frozen (100%) and fully unfrozen (0%) lake states.

4.2.4 Prospects

The EO-based ice mapper provides a consistent, observation-based dataset that can be used to benchmark modelled surface water temperatures, since the presence of ice corresponds to near-0 °C conditions and defines key seasonal transitions. While large-scale hydrological models in SOS-Water do not yet simulate ice cover explicitly, the dataset already supports thermal benchmarking and helps identify biases in seasonal timing.

Relative to existing datasets such as RLIE and ARLIE, the prototype demonstrates clear improvements in the consistency and fidelity of lake ice cover time series. Unlike these products, it also provides derived phenological indicators such as freeze-up and break-up dates, which are essential for linking observed ice dynamics to climatic and ecological processes. In terms of GCOS requirements, the prototype already meets the goal resolution of 50 m and is close to the suggested accuracy for phenology detection of ±2 days[14]. Over the European Alps it achieves ~40% daily coverage, corresponding to an effective observation frequency of ~2.5 days, bringing the dataset close to the GCOS goal of daily to sub-daily observations. The resulting historical dataset is therefore well suited for assessing climate-driven changes in lake ice cover and for supporting ecological studies. In particular, the reduction or loss of winter ice cover has direct implications for lake mixing regimes, which in turn affect oxygen availability and ecosystem functioning.

[14] Tom, M., Suetterlin, M., Bouffard, D., Rothermel, M., Wunderle, S., & Baltsavias, E. (2019). Integrated monitoring of ice in selected Swiss lakes Final Report - Final Report. https://doi.org/10.25678/0000JJ

Importantly, all of the satellite missions used in the prototype (Sentinel-1, Sentinel-2, and Landsat-7/8/9, with successors planned) are expected to continue into the future, ensuring the long-term viability of the framework. With further development, the approach could also be extended towards near-real-time monitoring, providing an additional tool for tracking ice conditions in support of water resources management.

4.3 EOP3: Satellite-based altimetry

4.3.1 Synergies with models

The integration of satellite-based altimetry in the SOS-Water project responds to persistent gaps in the availability of consistent water level observations across the project's basins. As identified in D3.1, hydrological models require temporally resolved datasets of river and lake water levels to calibrate routing and storage parameters effectively. In many transboundary regions, gauge networks remain sparse or fragmented, limiting the capacity to validate model performance and assess seasonal dynamics. The Surface Water and Ocean Topography (SWOT) mission offers a unique contribution in this context. Unlike earlier nadir-only altimetry missions, SWOT provides high-resolution, two-dimensional measurements of water surface elevation, width, and slope over wide swaths (Figure 3). These observations can be retrieved at both reach (approximately 10 km) and node (~200 m) scales, offering enhanced spatial detail and more frequent coverage (21-day repeat cycle). The expectation was that SWOT data would supplement in situ observations, improve the characterization of storage variability, and support calibration workflows in the IWMS framework.

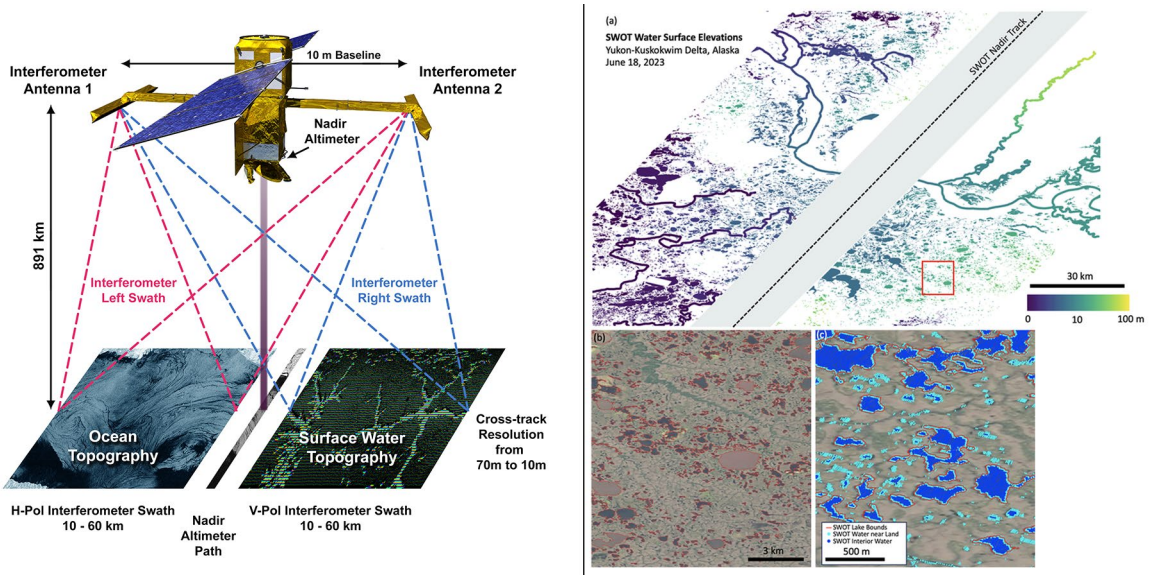

Figure 3. SWOT Interferometry provides high-resolution, two-dimensional measurements

4.3.2 Activities

To operationalize SWOT data for model integration, a programmatic workflow was developed during the reporting period (Figure 4). The approach began with a detailed review of available pre-release and science data products, associated uncertainties, and relevant processing methods. The workflow was implemented in Python within a Jupyter Notebook environment and relied primarily on the

Hydrocron API to retrieve and repackage SWOT Level 2 hydrology products into analysis-ready time series. Spatial reference features were defined using the SWORD (Prior River Database) and the Prior Lake Database.



Figure 4. SWOT data operationalization workflow

Data extraction focused on five locations spanning the Upper Danube, Danube Delta, and Rhine case studies. For each site, node-level and reach-level water surface elevation time series were compiled. In situ observations were collated from national hydrological services and harmonized to ensure alignment in temporal resolution and vertical datum. Validation procedures included bias correction, filtering of outliers, and calculation of Pearson's correlation coefficients and unbiased RMSE. The activities also documented retrieval limitations associated with specific site characteristics, notably in urban areas where signal degradation and misclassification were observed.

4.3.3 Results

The validation results demonstrated that SWOT data can substantially enhance the capacity to benchmark hydrological model outputs. Node-level observations consistently yielded higher agreement with in-situ measurements compared to reach-level aggregates (Figure 5), reflecting the advantages of finer spatial sampling. In the Danube Delta and Rhine reaches, correlation coefficients at node level approached or exceeded 0.9, with low error magnitudes after bias adjustment. The ability to resolve seasonal fluctuations in water level with this precision provides a robust basis for constraining routing parameters and validating simulated discharge dynamics.

In urbanized sections, such as the Linz (Rhine) reach, accuracy declined, with node-level correlations remaining moderate and reach-level correlations becoming negative in some cases. These discrepancies are likely linked to infrastructure interference and local topographic complexity, as noted in SWOT documentation. The findings also confirmed the importance of careful alignment of vertical datums and consistent filtering of low-quality observations prior to model integration.

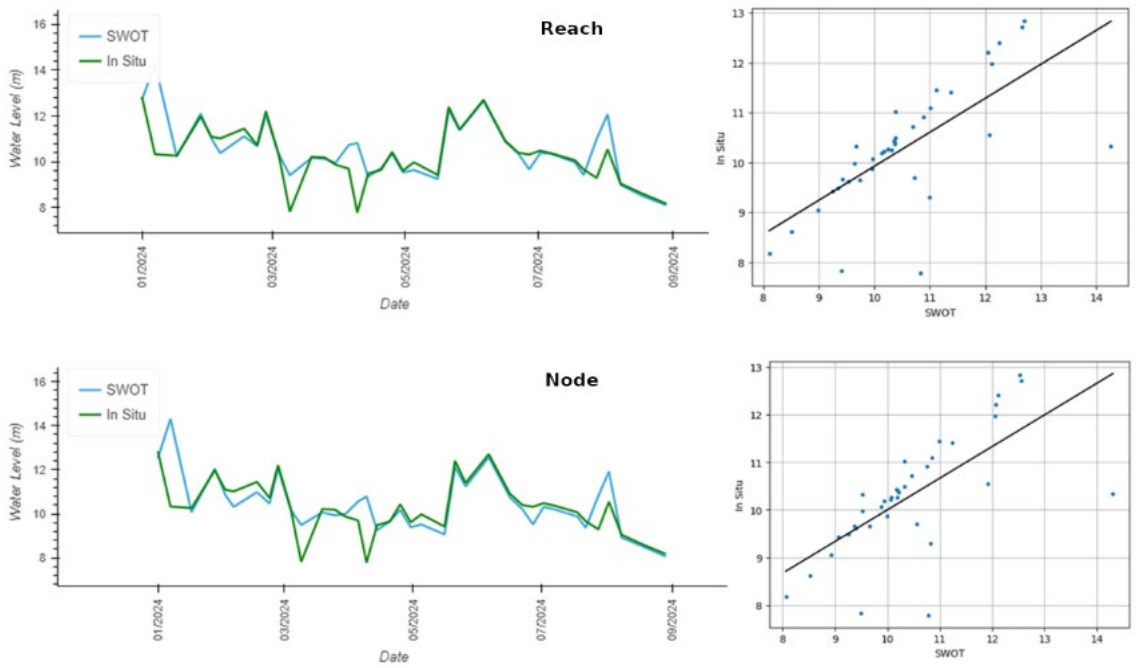

Figure 5. Time Series and Scatter Plot with Regression Line for Lobith, NE: Comparison of in-situ and observed data for Reach (top) and Node (bottom)

4.3.4 Prospects

The SWOT prototype illustrates clear potential for broader application in hydrological modelling and operational water monitoring. The combination of high spatial resolution, frequent revisit time, and global coverage represents a step change relative to previous altimetry missions. As future SWOT data releases become available, including discharge and volume estimates, the scope of integration will expand further and allows a more direct comparison with hydrological models. Incorporating SWOT-derived water levels into model calibration and assimilation workflows can improve the representation of storage processes, floodplain dynamics, and seasonal variability in river systems.

Operational adoption will require continued development of automated processing pipelines, robust correction procedures for known retrieval issues, and standardization of workflows across different modelling platforms. The experiences gained during this prototype phase provide a foundation for these next steps and demonstrate that satellite altimetry can be integrated effectively to strengthen model-based water resources assessments.

4.4 EOP4: Surface water temperature mapper

4.4.1 Synergies with models

Surface Water Temperature (SWT), as derived from satellite observations, serves as a proxy for water temperature, which regulates habitat suitability, governs biogeochemical reaction rates, and influences the overall ecological status of rivers and lakes. In SOS-Water, the DynQual surface water quality model, coupled to the PCR-GLOBWB2 global hydrological model, has been used to produce river and lake temperature outputs as part of broader water quality projections. These simulations are

used in biodiversity modelling workflows in WP4, where projected changes in thermal habitat under different climate scenarios are assessed through species distribution modelling (SDM).

Deliverable D3.1 identified a critical gap in the availability of spatially explicit, long-term SWT observations suitable for benchmarking large-scale hydrological and water quality models. Existing in-situ monitoring networks are spatially sparse and biased towards a limited number of sites, which restricts the ability to validate simulated water temperatures across large and hydrologically diverse river networks. An EO-derived SWT product addresses this limitation by providing spatially extensive observations at fine spatial granularity, which can be aggregated to reach and lake scale, enabling systematic validation of distributed model outputs.

The multi-decadal Landsat thermal archive, extending from 1982 to the present, provides a unique basis for deriving long-term SWT records. These EO-based observations offer an independent measurement source for validating model performance over extended time periods. By delivering observation-based benchmarks for simulated water temperature from DynQual and PCR-GLOBWB2, the EO product supports the identification of temporal, seasonal, and spatial biases in the simulations and provides information to modelers to guide model refinement.

4.4.2 Activities

The EOP4 Surface Water Temperature mapper builds on the complete Landsat thermal archive to generate 30-meter resolution SWT estimates at up to approximately 8-day temporal resolution. The workflow integrates thermal imagery from Landsat-4/5 TM, Landsat-7 ETM+, and Landsat-8/9 TIRS, applying consistent cloud and shadow masking, refined water masking, and pure water pixel selection to minimize mixed-pixel effects in narrow channels. This pure water pixel selection step is particularly important given the lower nominal thermal sensor resolution of 70 to 120 meters, which can otherwise lead to land contamination in narrow water bodies. Temperatures can further be aggregated spatially using the SWOT Prior Lake Database (PLD) and the SWORD river reach database, providing a continuous record of satellite-derived surface water temperature from 1982 to the present. Validation against Rhine in-situ records as part of T2.2 has shown high accuracy (five stations, R^2 from 0.921 to 0.981, mean absolute error 1.12 ± 1.02 °C), confirming the product's suitability for integration with basin-scale modelling frameworks.

During the reporting period, these EO data were processed and linked with outputs from the DynQual water quality model coupled to PCR-GLOBWB2. The WP2 partner UU (No. 2) provided monthly stream water temperature simulations for the Rhine basin covering 2005 to 2100, based on five Global Climate Models (GCMs; GFDL, IPSL, MPI, MRI, UKESM) and three SSP scenarios (SSP126, SSP370, SSP585). A benchmarking workflow was developed that spatially and temporally matches EO observations aggregated to reach-scale with modelled 1 km river cells using the SWOT SWORD river database. EO observations were further subset to produce two overlapping benchmarking datasets: one for the historical period 2005 to 2014 to assess baseline model performance, and another for the projection overlap period 2015 to present to evaluate agreement during the start of the projection horizon. This workflow forms the basis for a structured benchmarking process designed to quantify seasonal, temporal, and spatial simulation biases and to provide targeted feedback for model refinement.

4.4.3 Results

A first benchmarking exercise was carried out for the Rhine basin using a linear regression analysis between EO-derived and modelled SWT for both the historical (2005–2014) and the projection overlap period (2015–present). Figure 6 illustrates the results for the MPI-driven simulation, which achieved the best overall agreement among the five GCMs. For the historical period, R^2 values ranged from 0.68 to 0.75, mean absolute error (MAE) from 2.87 to 3.33 °C, and bias from –1.72 to –1.31 °C. For the projection period, R^2 values ranged from 0.59 to 0.72, MAE from 3.55 to 4.14 °C, and bias from –2.95 to –2.34 °C. These results indicate generally reasonable agreement but also highlight a consistent underestimation of SWT by all GCM-driven simulations, with larger discrepancies observed in the projection period compared to the historical baseline.

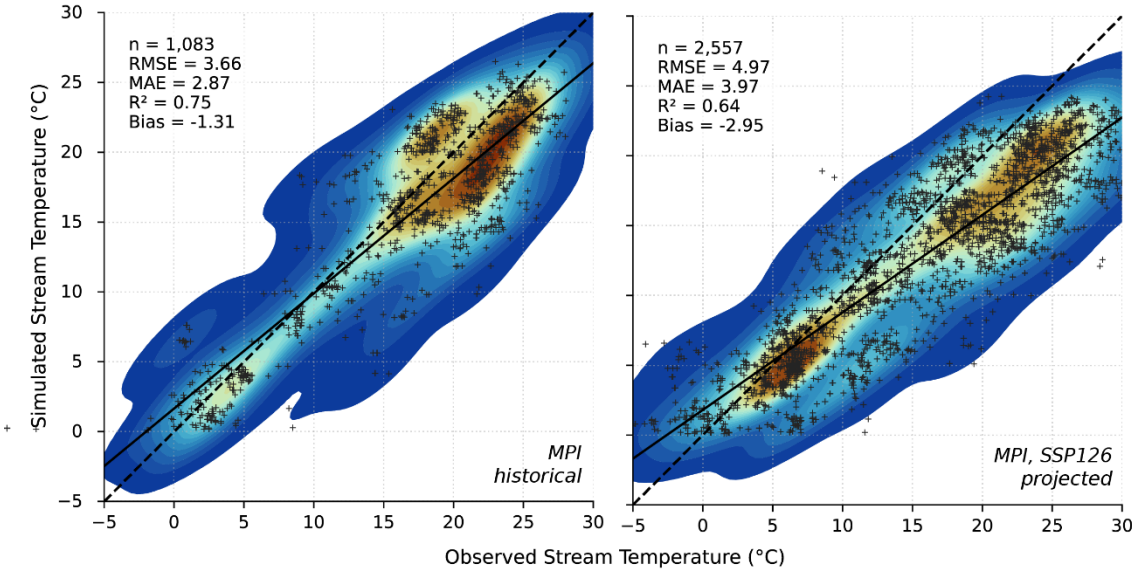

Figure 6. Comparison of observed remote-sensing and simulated DynQual stream temperatures (MPI). The left panel shows historical benchmarks and the right panel projected SSP126 benchmarks after the 2015 cutoff. Black dashed lines indicate the 1:1 relationship, and solid lines show the linear regression fit.

Seasonal residual analysis showed that positive residuals, where simulated SWT was lower than EO observations, occurred most frequently from late spring to early summer. This period coincides with snowmelt-driven high flows in the Rhine, suggesting that model biases may be linked to the representation of heat exchange processes during these conditions. Figure 7 illustrates the MPI-driven results for the entire overlap period. The integration of EOP4 demonstrates that EO-derived SWT can reveal both systematic biases and finer-scale patterns, such as reach-specific and season-specific discrepancies, that would be difficult to detect using the sparse and uneven coverage of existing in-situ monitoring networks.

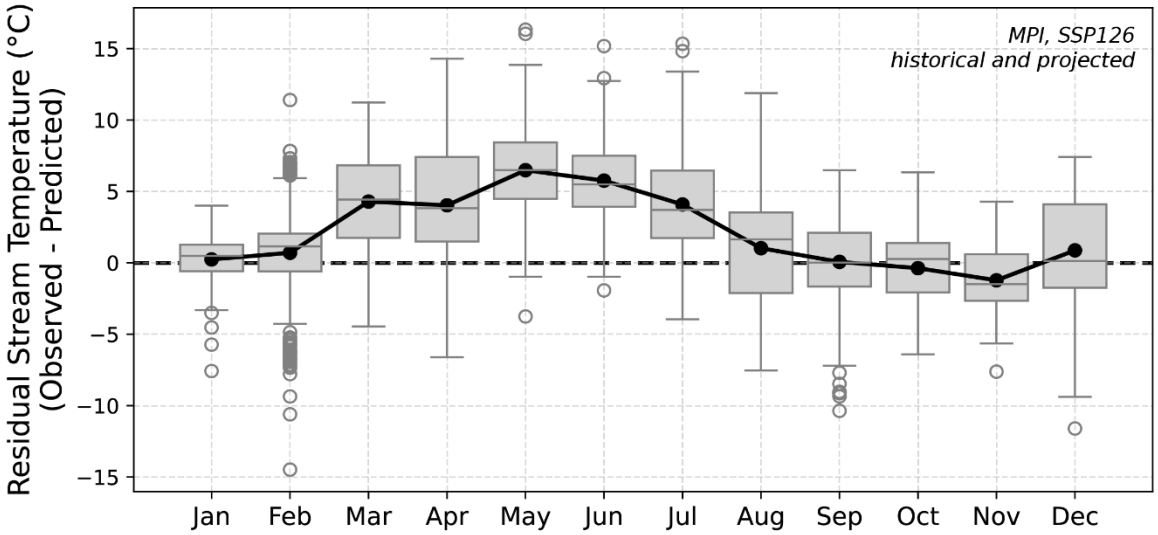

Figure 7. Seasonal residual analysis between observed remote-sensing and simulated DynQual stream temperatures (MPI, SSP126, historical and projected). High positive residuals, indicating underestimation of simulated values, can be seen in late spring to early summer.

Further potential of multi-year satellite-based SWT analysis for targeted applications has been illustrated in a study currently under review in the Journal of Hydrology: Regional Studies (Woszczyk and Brechbühler, 2025) examining thermal pollution in Lake Licheńskie, central Poland. This case study demonstrates how long-term EO temperature records, extracted using the EOP4, can be applied to assess thermal impacts from anthropogenic activities and to inform management decisions.

4.4.4 Prospects

Expanding this benchmarking framework across the SOS-Water case studies will enable a more comprehensive evaluation of DynQual water temperature simulations. The resulting benchmarks can be exploited across different temporal and spatial scales, such as reach- or season-specific scales. This allows modelers to pinpoint parameterization issues and to identify when and where discrepancies between simulated and observed water temperatures arise. Future inclusion of lakes will offer a more complete assessment of thermal model performance across both riverine and lacustrine systems. In the longer term, the workflow could mature into a standard EO-model benchmarking toolbox for large-scale hydrology and water quality models.

In a broader context, ground-based stations will remain essential for accurate calibration and validation due to their higher measurement accuracy, regular sampling, and much finer temporal resolution. EO-derived SWT and LSWT from Landsat should therefore be viewed as a complementary, independent validation source, offering a promising option for spatially extensive assessments that cannot be achieved through in-situ networks.

4.5 EOP5: Crop water productivity mapper

4.5.1 Synergies with models

Water productivity (WP) in crops, also known as water use efficiency (WUE), is a measure of the amount of biomass or yield produced per unit of water consumed by a crop (kg/m^3). It is a key performance indicator of agricultural productivity and sustainability, particularly in regions where water resources are limited or where efficient water management is essential. Thus, it is also an important indicator for definition of the Safe Operating Space.

Improving crop productivity is a key mechanism for maximizing agricultural production while minimizing water consumption and environmental impacts. Strategies for enhancing water productivity in crops may include adopting more efficient irrigation systems, optimizing planting densities, selecting drought-tolerant crop varieties, improving soil water retention, and implementing conservation practices (Pereira et al., 2012; Tuberosa et al., 2007). Agrohydrological models can support assessment of such management scenarios.

Water Productivity is an important input in hydro-economic models employed in SOS-Water to define the SOS of agroecosystems. Particularly a clear synergy has been identified for supporting the hydro-economic model applied in the Jucar River basin (a model description is provided in D2.2, and preliminary results reported in D5.1 and D5.2). The current version of the hydro economic model uses as input estimates of max. agricultural production under a full-irrigation scheme - without crop water deficits. Actual values of crop production, considering deficits in irrigation, are computed following the FAO66 method. The original configuration requires yearly values of agricultural production, but static-average values can be also included.

WaPOR (Water Productivity Open-access of Remote sensing data)[15] is the FAO's portal to assist countries in monitoring water productivity, identifying water productivity gaps, and proposing solutions to reduce these gaps (FAO, 2023). WaPOR rests on the EO-based ETLook model for the high-resolution (30m) quantification of water productivity in croplands. Based on this methodology, PyWaPOR[16]is an open-source Python package that provides users with the ability to generate water productivity and related outputs at any region and period of interest by using open-access satellite imagery in combination with local data. EOP5 was developed based on the pyWaPOR algorithm, under the premise that WP values can be used as direct inputs into the hydro-economic model, rather than the traditional FAO66 approach.

4.5.2 Activities

The Crop Water Productivity Mapper of SOS rests on the WaPOR model developed by FAO. that has been tested as a complementary tool in the frame of the SOS-Water project, and particularly in the pilot case of the Jucar River Basin (Spain). For the purposes of SOS-Water, WaPOR was enhanced from its original configuration and properly adapted for increasing the spatial resolution and the accuracy

[15] https://data.apps.fao.org/wapor/?lang=en

[16] https://www.fao.org/aquastat/py-wapor/index.html

of the outputs. The testing of new model configurations and the ingestion of a high-resolution, local landuse-landcover dataset (SIGPAC) were addressed. The original climate forcing layers from the MERRA (0.5deg.) and GEOS5 (0.25deg.) datasets were replaced by the ERA5-Agro (0.1 deg) and ERA5-reanalysis (0.25 deg.). Values of Net Biomass Water Productivity (NBWP) - as surrogate of crop productivity - for different cropland categories were generated for the Júcar Water Resource System. one of the three exploitation units of the Júcar River Basin.

4.5.3 Results

The EO prototype, tailoring the generic pyWaPOR code to the context of a small Mediterranean catchment, has been locally implemented in a first assessment which aimed to provide representative values of crop water productivity for the most dominant crops in the Jucar River Basin (Figure 8)[17]. Average evaporative fluxes and NBWP values retrieved from WaPOR – spatially and timely averaged per cropland category and irrigation scheme – are shown in Table 4 for the most dominant croplands in the region. These values capture the annual variability and the spatial variability of the NBWP, so we consider may be adopted as representative values for each crop type

Results are currently being cross validated with field measurements of ET and crop productivity retrieved from lysimeters. It is currently undergoing a critical revision of preliminary WaPOR outputs, fine-tuning of internal algorithms, and optimization of the computational workflow to reduce processing requirements.

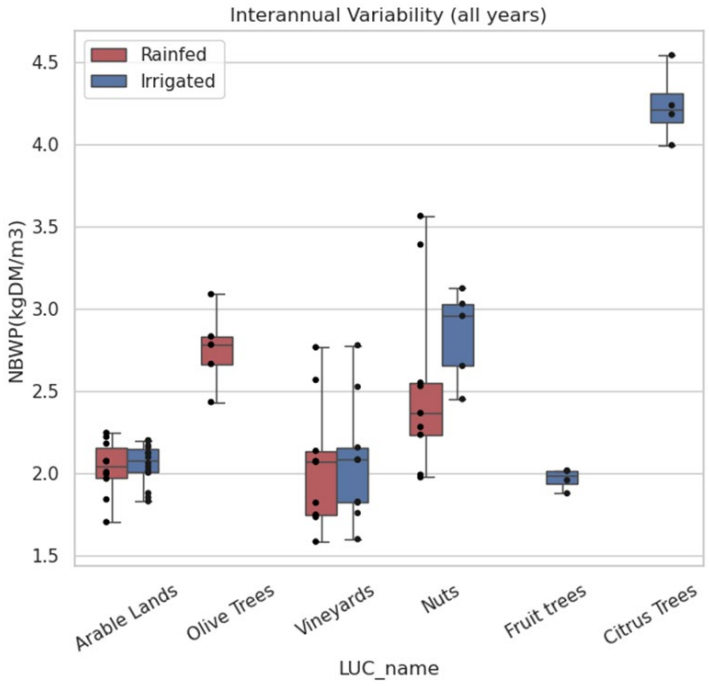

Figure 8. Interannual Variability of NBWP computed with the adapted pyWaPOR

[17] Férnandez-Rodriguez, Contreras & Simons, 2024. FutureWater Report 250

Table 4. Averaged pyWaPOR results for different crop types

Crop	Irrigation scheme	T (mm/y)	ET (mm/y)	NPP[1)] (gCm^{-2}/y)	TBP ($kgDM \cdot ha^{-1}/y$)	NBWP ($kgDM/m^3$)
TA - Arable land	Rainfed	164	269	150	3344	2.03
	Irrigated	353	483	328	7295	2.05
FS – Nuts	Rainfed	121	600	131	2930	2.54
	Irrigated	314	939	412	9173	2.84
FY – Fruit trees	Rainfed					
	Irrigated	460	570	411	9142	1.96
OV – Olive trees	Rainfed	89	182	111	2478	2.76
	Irrigated					
VI - Vineyards	Rainfed	157	230	137	3053	2.05
	Irrigated	183	256	160	3577	2.07
CI – Citrus trees	Rainfed					
	Irrigated	278	381	529	11761	4.23

4.5.4 Prospects

Water productivity plays a crucial role in understanding links between agricultural practices and water resources. Accurate assessment for water productivity across both space and time has significant importance for applications in water resources management, especially within a context of changing climate and society. EOP5 achieved the high-resolution quantification of water productivity per cropland type and irrigation scheme in order to contribute to the definition of a safe operating system (SOS) for water resources.

WaPOR technical feasibility to be implemented in new pilot sites has been proved by setting up the algorithm to cover the Júcar River Basin area. The official WaPOR database does not report data at Level 3 (30-meter resolution) for this area, so this exercise demonstrates the adaptability of WaPOR for being applied across different geographical settings. The implementation of the algorithm allows the user to choose the native sources from which remote sensing data are downloaded. Sideloading is available, so the user can even generate its own remote sensing datasets to be ingested prior to the model run.

The flexibility of the algorithm for the assimilation of high-resolution local datasets has been tested by the ingestion of a local land use/land cover dataset, tailoring the analysis to the study area. Due to the complexity of the model and the multiple interdependencies between parameters, the ingestion of local datasets can be challenging in the sense that extra parametrization might be needed. Some advanced knowledge on the algorithm is needed in order to identify how the workflow may be affected

by the replacement of a default dataset (in this case, the ingestion of a local land use dataset required the parametrization of three additional variables).

In the current version, there are two main limitations that impact prospects for upscaling EOP5 to other regions:

(i) to run pyWaPOR is computationally expensive, especially during the pre-processing module. Upscaling at the river basin scale requires powerful computing infrastructure and considerable runtimes

(ii) the need for local calibration and validation was demonstrated; however, the intrinsic complexity of the model behind the algorithm is very high, which makes it difficult to calibrate and validate the outputs.

4.6 EOP6: Snow-water equivalent downscaling

4.6.1 Synergies with models

Snow Water Equivalent (SWE) is one of the most important variables in mountain hydrology, yet at the same time one of the least understood. At the same time, its complex to monitor due to its high horizontal and vertical heterogeneity. As SWE controls the timing and amount of runoff, it strongly affects water availability and the risk of flooding. Most hydrological models rely on simplified estimates derived from temperature and precipitation, often failing to capture local snow processes governing snow accumulation and melt. Comparing snow outputs of hydrological models with improved SWE products can bridge this gap and aid in the calibration of these complex snow processes. Using datasets that combine observations with modelling, such as ERA5-Land, can improve the accuracy of hydrological model simulations by improving the uncertain snow storage term with data that better reflect actual conditions.

Earth Observation Prototype 6 (EOP6), the algorithm created in SOS-Water for downscaling of ERA5-Land Snow Water Equivalent (SWE), and the SWE dataset it generated, supported model calibration and benchmarking of hydrological models within the SOS-Water project: the CWatM model for the Upper Danube and the PCR-GLOBWB model for the Rhine-Meuse basin. The added value and synergy with the models lies in its ability to reduce spatial bias and better capture temporal dynamics in the snow modules of these models.

4.6.2 Activities

The integration of EOP6 into hydrological modelling was pursued along two major activities:

1) Direct forcing of hydrological models with SWE

The downscaled SWE dataset was directly ingested into the CWatM model for the Melach catchment in the Upper Danube and calibrated against discharge observations at In der Au. In this configuration, the downscaled SWE dataset replaced the CWatM default temperature-precipitation driven snow module, thereby directly forcing the model with observation-informed SWE rather than parameterized snow storage dynamics. Performance was tested both without

additional calibration and with recalibration using the DEAP algorithm. Similar work is in progress for the PCRGLOB-WB simulation for the Rhine-Meuse.

2) Snow module calibration against downscaled SWE

Secondly, the downscaled SWE dataset was used as a benchmark for recalibrating the snow modules of both CWatM and PCR-GLOBWB. Key snow parameters present in these models, including the degree-day melt factors, accumulation thresholds, and snow albedo decay rates, were reviewed and where possible adjusted to reproduce the spatial and temporal SWE patterns captured by the downscaled dataset. The goal to realign model snow process representation with the observation-based SWE dataset, rather than relying solely on discharge-based tuning.

4.6.3 Results

1) Direct forcing of hydrological models with snow melt

The produced downscaled SWE dataset was directly assimilated into the CWatM model for the Melach catchment in the Upper Danube and calibrated on river discharge observations at In der Au. The downscaled dataset replaced the CwatM default temperature-precipitation based snow simulation. This approach forced the model with melt derived from observation-informed SWE, replacing parameterized snow storage dynamics with data-driven representation.

Table 5. Performance metrics of different runs of direct forcing CwatM with downscaled SWE dataset

Run	Period	KGE	NSE	R^2
Standard CWatM (calibrated)	Calibration	0.76	0.51	0.85
	Validation	0.77	0.51	0.82
CWatM with downscaled SWE (uncalibrated)	Calibration	0.35	-0.12	0.61
	Validation	0.33	-0.08	0.64
CWatM with downscaled SWE (DEAP calibrated)	Calibration	0.66	0.40	0.68
	Validation	0.64	0.38	0.67

The performance metrics of the direct forcing runs of CwatM with the downscaled dataset show mixed results for the different experiments (Table 5). When the CwatM snow module was replaced and not calibrated, the model performance dropped substantially from ~0.76 KGE to ~0.34 KGE, NSE even dropping to negative values, indicating a poor agreement with the physically more realistic SWE input. Once the model was calibrated a second time using the DEAP algorithm, the performance recovered substantially, but never overcoming the initial simulations of the standard calibrated CWatM run with temperature-precipitation snow module. These results highlight that direct forcing with EO-based observations are difficult as a drop-in replacement for conventional parameterizations and model structures but likely require dedicated recalibration and revision of formulas.

Figure 0 shows simulated versus observed daily discharge at the Melach River, In der Au station (basin area: 205 km²), using the downscaled SWE as forcing input. The simulated discharge (red) tracks the observed discharge (blue), however snow buildup simulated is too often late compared to the observations. The months of peak timing do not coincide, though the simulated peaks tend to be in a similar range than the observed peaks. Due to poor performance the direct forcing of SWE onto the CWatM model was halted and work on calibrating the snow module started (2).

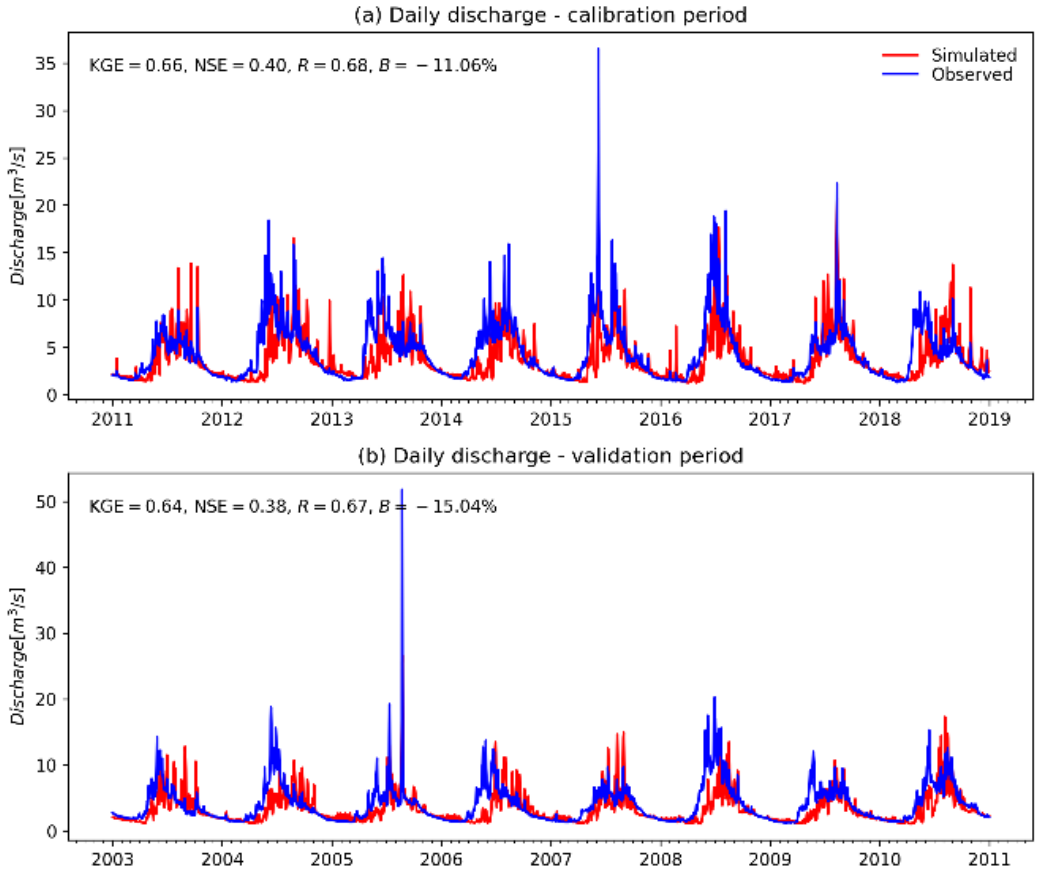

Figure 9. Calibration results of CWatM snowmelt using the downscaled SWE dataset

2) Snow module calibration against downscaled SWE

In addition to the direct forcing experiments, the downscaled SWE dataset is being prepared as a reference for recalibrating the snow modules of both CWatM and PCR-GLOBWB. This process involves adjusting snow-related parameters, such as the degree-day melt factors, accumulation thresholds, and albedo decay rates, to reproduce the spatial and temporal SWE patterns captured by the downscaled and bias-corrected SWE dataset. The objective is to realign model snow dynamics with observation-based snow storage dynamics rather than relying solely on discharge-based parameter tuning. Using the improved SWE dataset as a target dataset will provide an

independent constraint and reduce the risk of errors in calibration, or overcompensating parameters for the wrong reasons. In this way, model equifinality is managed and limited.

Benchmarking of this calibration pathway is planned to quantify the added value of the downscaled SWE in improving modelled snow dynamics and runoff simulation. The comparative experiments will assess performance under standard reanalysis forcing, direct forcing with the SWE dataset, and recalibrated snow modules informed by the downscaled SWE. While this work is still in progress, the preparatory steps have confirmed the feasibility of the approach and its potential to reduce bias in SWE representation.

4.6.4 Prospects

EOP6 demonstrates the potential of EO-based SWE downscaling to enhance hydrological modelling in snow-dominated basins. By constraining snow dynamics with observation-informed datasets, models can achieve a more physically consistent representation of snow storage and melt processes, thereby strengthening their reliability for historical benchmarking. Even while showing some pitfalls, in for instance the direct forcing, there is still potential for further improvement of this algorithm.

During the execution of this project, the prototype has evolved considerably, with methodological improvements such as the transition from Random Forest to XGBoost, enhanced training and validation strategies, and the integration of additional remotely sensed predictor datasets. These advances have in the meanwhile improved the predictive skill, and continued refinement is expected to further improve performance. As shown in Figure 10, the time series of SWE predictions averaged over all stations shows a good model performance. Continued research in the form of a scientific paper will further this EO prototype.

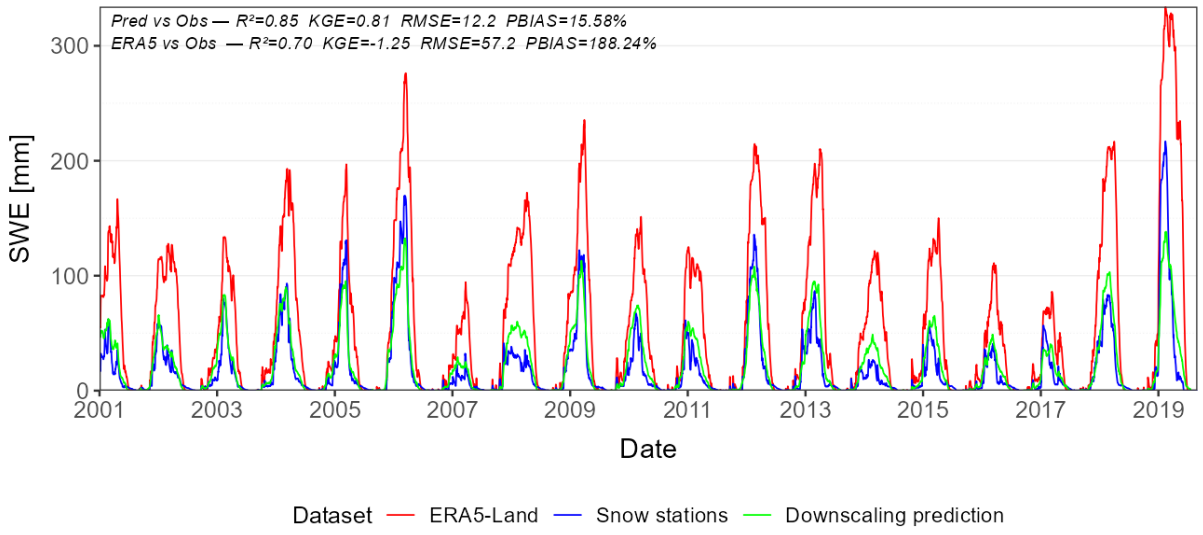

Figure 10. Average time series of SWE test predictions against observations at 383 stations

5 Conclusions

This report has demonstrated the integration of a suite of Earth Observation (EO) prototypes into basin-scale water resources modelling within the SOS-Water project. Each prototype addressed a specific gap identified in earlier project stages and was developed to enhance the accuracy, robustness, and applicability of the integrated water modelling system (IWMS):

- The Suspended Particulate Matter (SPM) Mapper advanced sediment flux monitoring in the Mekong by combining EO-derived sediment concentrations with modelled discharge, enabling basin-scale assessments of sediment dynamics in data-sparse conditions.
- The Lake and River Ice Mapper delivered accurate, high-resolution ice cover and phenology datasets, filling a long-standing gap in cryosphere monitoring and providing essential benchmarks for climate and ecological studies.
- Satellite-based altimetry, operationalized through SWOT data, proved valuable for water level benchmarking and has the potential to improve calibration of routing processes, while also highlighting challenges in urbanized or complex environments.
- The Surface Water Temperature Mapper established a long-term record of thermal conditions, enabling systematic benchmarking of simulated water temperatures in the Rhine basin and revealing model biases in seasonal and spatial dynamics.
- The Crop Water Productivity Mapper demonstrated the feasibility of tailoring EO-based productivity estimates to regional agricultural systems.
- Finally, the Snow Water Equivalent (SWE) downscaling prototype showed clear potential to improve the representation of snow processes in hydrological models, despite challenges in direct forcing, and has already evolved into a robust, observation-informed benchmark dataset.

Collectively, the integration of these prototypes illustrates the potential of EO to improve both the input data and the benchmarking capabilities of hydrological models. While the degree of integration varied across prototypes and case studies, each demonstrated tangible benefits for constraining uncertainty, enhancing model realism, and supporting water resources assessments within the Safe Operating Space framework. At the same time, the work has also underlined the importance of methodological adaptation, recalibration, and validation to ensure EO products can be meaningfully assimilated into model workflows.

The experiences gained through SOS-Water Task 3.3 highlight both the immediate and long-term opportunities for EO integration into water resource modelling. In the short term, the developed prototypes are already contributing to improved model performance in the SOS-Water case studies. In the longer term, methodological frameworks and datasets such as those established here can serve as building blocks for broader applications, supporting the modelling community in advancing towards more reliable, transparent, and observation-driven water resources management.

Disclaimer

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.

Acknowledgement of funding

Funded by the European Union

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

This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI).

Project funded by

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs,
Education and Research EAER
State Secretariat for Education,
Research and Innovation SERI