

ESA Climate Space Tipping Elements Project

RESET lakes

Inventory

Version 1.0

29 September 2025

Authors

Michael Brechbühler

Iestyn Woolway

Elisa Calamita

Caren Binding

Daniel Odermatt

Content

1	Datasets	2
1.1	<i>Data overview</i>	3
1.2	<i>Post processing of EO</i>	4
2	Models	5
2.1	<i>Modelling Lake Biogeochemical Regimes with LAKE 2.0</i>	5
2.2	<i>Comparative Assessment of ALBM and LAKE 2.0 for Tipping Point Detection</i>	6
2.3	<i>Generic Indicators</i>	7
3	References	7

1 Datasets

A comprehensive inventory of Earth Observation (EO) datasets for water resources monitoring was compiled by Eawag as deliverable D3.1 in the SOS-Water project¹. The inventory comprises not only lakes datasets, but also land and catchment variables that influence lake dynamics and can hence act as potential triggers for tipping points. It was updated for RESETlakes, and the full Excel-sheet of data entries is available on demand. Here, we report on some of its metadata and key coverage attributes.

After identifying the relevant data sources, we systematically collected key metadata attributes for each dataset. Most of the entries are multivariate spatio-temporal EO products, which results in a diverse and heterogeneous collection. To make this collection usable across the project, each dataset was described using a consistent set of attributes (Table 1). These attributes capture essential information such as observed main variables and their units, spatial and temporal resolution and extent. This structured approach ensures that the content, limitations, and potential applicability of each dataset can be rapidly assessed by all project partners.

Table 1: 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)	Messenger, 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
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.

¹ <https://www.sos-water.eu/publications-and-press#deliverables>

The data inventory was developed to adhere to the FAIR principles (Findable, Accessible, Interoperable, Reusable). Applying FAIR standards ensures that datasets are well-documented, labelled, and easy to access, while also promoting interoperability and reusability across different applications. This is essential for building a robust, transparent, and effective data resource that supports the project’s modelling, analysis, and knowledge exchange activities.

1.1 Data overview

In total, 141 project-relevant EO datasets were reviewed and compiled into a comprehensive data inventory. The data inventory builds on an earlier compilation developed in the SOS-Water project, from which 87 dataset entries were initially reused (Brechtbühler, 2024). These entries were subsequently filtered, updated, and expanded with additional datasets to meet the specific scope of this study. The majority of entries originate from the ESA Climate Change Initiative (CCI) Open Data Portal (70 datasets, 50%) and the Copernicus Global Land Service (CGLS, 43 datasets, 31%), complemented by additional contributions from NASA, USGS, EUMETSAT, and other established providers. Figure 1 summarizes the variables covered in the inventory, grouped into six thematic categories such as surface and subsurface hydrology. This overview highlights the diversity of EO products available to study both direct lake processes and external catchment influences that may contribute to tipping behavior.

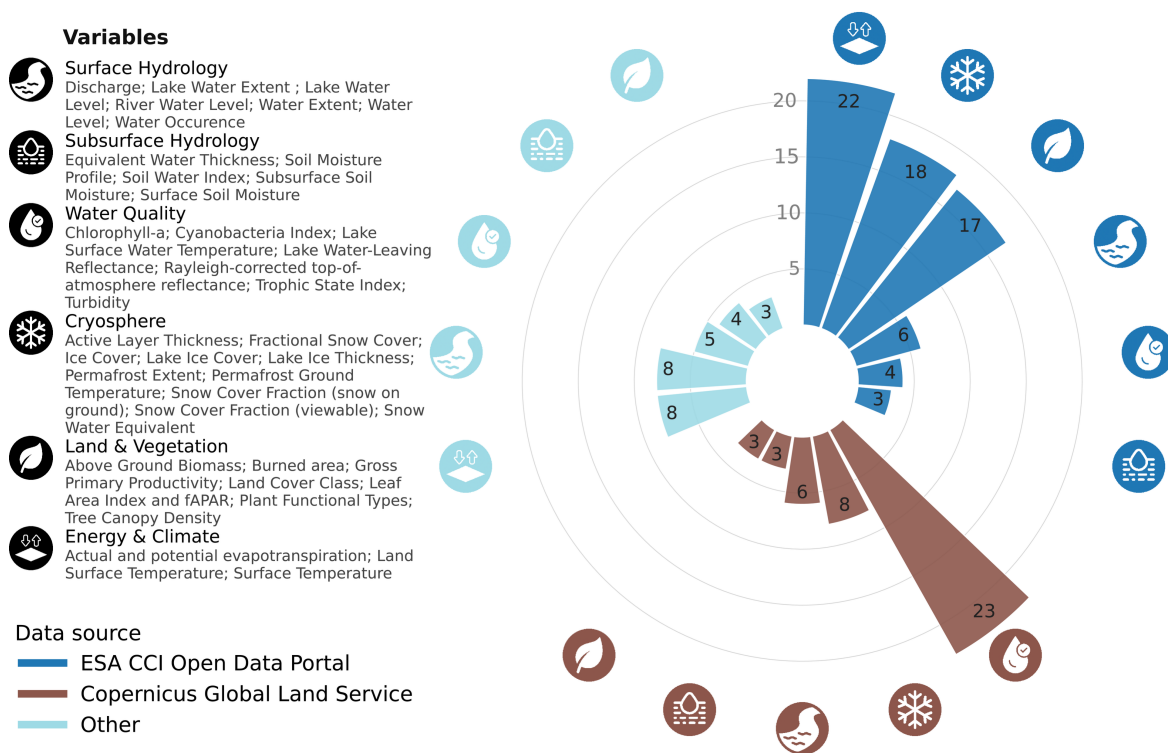


Figure 1: Number of datasets compiled in the inventory, grouped by thematic variable categories. The figure highlights the main contributing sources.

A central consideration when applying EO data to monitor lakes and identifying tipping behavior is the interplay between spatial and temporal resolution. Spatial resolution determines the level of detail observable, ranging from a few meters (e.g. Sentinel-2/MSI optical data) to several hundred kilometers (e.g. GRACE-FO gravimetric data). Temporal resolution reflects how frequently observations are available over the same location, varying from sub-hourly sampling on geostationary platforms to weekly, monthly, or longer intervals. No single spatio-temporal scale is universally optimal; rather, requirements depend strongly on the process under investigation. High-resolution imagery (10 to 300 meters) is critical for capturing dynamics in small and medium-sized lakes, such as algal blooms,

shoreline changes, or ice cover dynamics. In contrast, coarser-resolution products at kilometer scales are most suitable for catchment-scale drivers including land use, soil moisture, snow, and permafrost dynamics. Similarly, high-frequency time series (daily to weekly) are needed to detect abrupt transitions and short-term events, while long and continuous records spanning decades are indispensable for recognizing gradual trends and potential early-warning signals.

Figure 2 provides an overview of the compiled datasets in terms of spatial and temporal resolution as well as temporal extent. While many products are nominally distributed as “daily” (61, 48%), it is crucial to consider that the effective observation frequency is often lower due to sensor limitations, orbital characteristics, or cloud cover interference. This means that actual usable data can be available at weekly or even coarser intervals depending on location and season. The temporal extent of the datasets is highly variable: several records reach back to the 1980s, while others represent short operational data streams that are still ongoing. Importantly, many datasets now provide two decades or more of consistent observations, forming a robust basis for long-term trend analyses and tipping-point detection.

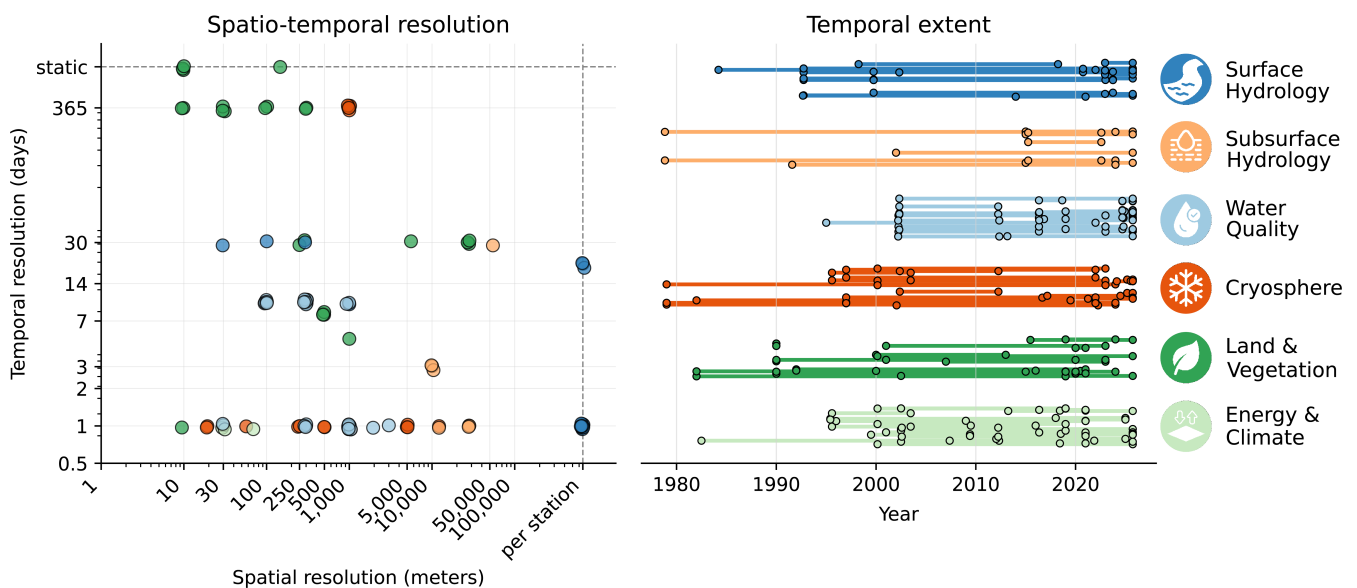


Figure 2: Spatio-temporal resolution of datasets compiled in the inventory (left) and temporal extent (right), grouped by thematic variable categories.

Most of the collected datasets (106, 75%) offer global coverage, which ensures their applicability across diverse lake regions. Nevertheless, global spatial extent is often referring to globally distributed lakes or stations and limitations often remain when monitoring smaller sites. For instance, the “ESA Lakes Climate Change Initiative (Lakes_cci): Lake products, Version 2.1” dataset delivers a consistent multivariate time series of essential lake variables on a global scale, but its coverage is restricted to approximately 2,000 of the world’s largest lakes. Such constraints underline the importance of carefully matching dataset characteristics to the scale and type of processes under investigation.

1.2 Post processing of EO

To ensure robust detection of ecosystem shifts and tipping points, gap filling is essential to reduce spatial and temporal bias in EO datasets. As highlighted in (Bathiany et al., 2024), missing observations, caused by cloud cover, seasonal effects, or sensor limitations, can distort early-warning signals such as rising variance or autocorrelation. Without adequate gap filling or bias correction, these methods may yield misleading results or fail to detect critical transitions. For lakes, where both short-term dynamics and long-term trends are crucial, careful pre-processing of EO data, including interpolation, smoothing, and uncertainty estimation, is a prerequisite for reliable resilience assessment and trend analysis.

Addressing data gaps and bias in Earth Observation records represents a critical knowledge gap that must be prioritized. Reliable detection of ecosystem shifts, especially in lake systems, depends on the availability of temporally and spatially continuous data. To support robust analysis and early warning efforts, we must focus on developing and validating gap filling methodologies that minimize bias while preserving meaningful variability in the signal. Producing user-ready, gap-filled datasets will be essential for enabling confident application of resilience indicators and supporting operational monitoring systems.

2 Models

In this section we address the question whether the Advanced Lake Biogeochemistry Model (ALBM), combined with machine learning algorithms, can be used to identify tipping points in lake thermal and biogeochemical regimes. The Advanced Lake Biogeochemistry Model (ALBM) (Tan et al., 2015, 2018) is a one-dimensional, vertically resolved model that simulates lake thermodynamics, ice and snow cover, sediment heat exchange, and coupled carbon–methane cycling. Its modular architecture explicitly represents nonlinear feedbacks between lake thermal structure, redox-sensitive biogeochemical processes, and greenhouse gas emissions, all of which are central to identifying tipping points in lake systems.

ALBM represents heat diffusion, convective overturn, radiation transfer, and snow–ice dynamics in a simplified yet physically consistent way. It couples these processes with detailed modules for organic carbon cycling, methane production and oxidation, oxic methane production, and bubble-mediated fluxes. By distinguishing between labile and recalcitrant organic carbon pools, ALBM also captures sediment “memory” effects that can delay or amplify biogeochemical responses.

The model simulates a wide range of feedbacks: stronger stratification suppresses oxygen renewal in deep waters, leading to anoxia and enhanced methane release; dissolved organic carbon (DOC) alters heat absorption, stratification, and greenhouse gas fluxes; and long-term organic matter accumulation primes lakes for abrupt regime shifts when thresholds are crossed. These dynamics make ALBM well suited for exploring abrupt transitions such as the onset of persistent hypoxia or a switch to continuous methane emissions.

Machine learning (ML) enhances ALBM by detecting early-warning signals of tipping behaviour, constraining uncertain parameters with Earth Observation and in situ data, uncovering nonlinear thresholds, and emulating large ensembles of simulations to rapidly explore climate and catchment scenarios. Together, ALBM and ML provide a powerful hybrid framework to identify and anticipate tipping points in lake thermal and biogeochemical regimes.

2.1 Modelling Lake Biogeochemical Regimes with LAKE 2.0

While ALBM emphasizes computational efficiency and detailed biogeochemistry, LAKE 2.0 is a more mechanistically complex one-dimensional model designed to represent turbulence, ice–snow processes, and energy balance in greater detail (Stepanenko et al., 2011). It has been widely applied in climate and ecosystem studies, particularly in Arctic and sub-Arctic environments.

LAKE 2.0 employs a turbulence closure scheme that dynamically links mixing intensity to stratification and wind forcing. It includes a multilayer snow and ice model, accounting for black ice, white ice, compaction, and refreezing events. These features make it particularly suited for high-latitude lakes where ice-cover dynamics and snow–radiation interactions dominate.

The model also resolves sediment heat conduction and benthic interactions, linking bottom-water temperature with methane production. Shortwave radiation absorption is dynamic, depending on water turbidity, snow, and ice. Methane cycling includes production in sediments, oxidation in the water column, and bubble-mediated fluxes that bypass oxidation to the atmosphere.

LAKE 2.0 captures feedbacks between stratification and turbulence, snow–ice cover and radiation, and nonlinear methane oxidation. Its detailed physical schemes make it highly suitable for process studies

and regional applications, though its higher computational cost limits its use in global-scale ensemble modelling compared to ALBM.

2.2 Comparative Assessment of ALBM and LAKE 2.0 for Tipping Point Detection

Both ALBM and LAKE 2.0 provide valuable but complementary insights for RESETLakes. ALBM is efficient and includes a wide range of methane pathways and sediment memory effects, making it ideal for ensemble modelling and integration with machine learning. LAKE 2.0, by contrast, provides a more detailed treatment of turbulence, snow and ice, and radiation feedbacks, making it particularly valuable for Arctic and high-latitude systems where ice phenology strongly influences tipping dynamics.

Table 1: Comparison of ALBM and LAKE 2.0 for modelling lake thermal–biogeochemical regimes.

Feature	ALBM	LAKE 2.0	Implications for Tipping Point Studies
Core design philosophy	Parsimonious 1D thermal model with detailed biogeochemistry	Mechanistic 1D turbulence-resolving model with intermediate biogeochemistry	ALBM better for ensembles/global runs; LAKE 2.0 for process-rich regional case studies
Thermal dynamics	Heat diffusion + implicit convective adjustment	k-ε turbulence closure; dynamic vertical mixing	LAKE 2.0 resolves turbulence feedbacks, ALBM robust for large-scale scenario exploration
Ice & snow dynamics	Prognostic snow–ice balance with simplified refreezing	Multilayer snow–ice energy–mass balance, includes black/white ice	LAKE 2.0 superior for Arctic/sub-Arctic tipping points linked to ice phenology
Sediment coupling	Simplified thermal conduction, two OC pools (labile/recalcitrant)	Thermal conduction; simpler OC representation	ALBM captures sediment “memory” (legacy OC); relevant for delayed CH ₄ tipping
Methane pathways	Ebullition, anoxic + oxic CH ₄ production, nonlinear oxidation	Ebullition, sedimentary CH ₄ production, Michaelis–Menten oxidation	ALBM captures wider range of CH ₄ feedbacks, including OMP; LAKE 2.0 more traditional
Feedback representation	Explicit DOC–heat–mixing, oxygen depletion, sediment memory	Turbulence–stratification, snow–ice–radiation, nonlinear CH ₄ oxidation	Both capture distinct feedbacks; combining them provides full coverage
Computational cost	Lightweight; suited to ensembles and climate sensitivity analysis	Higher; suited to detailed process studies but less scalable	Suggests a complementary two-tier strategy
Integration with EO/ML	Well suited for ML emulation, parameter calibration, ensemble detection of early warning signals	Provides high-fidelity process data for training/validating ML and benchmarking ALBM	Use ALBM+ML for large-scale EO integration; use LAKE 2.0 for calibration and validation

ALBM excels in scalability and its explicit treatment of organic carbon pools, methane oxidation pathways, and sediment memory effects, all of which are critical for understanding delayed and nonlinear tipping behaviours in lake carbon cycling. When combined with ML, it enables early-warning signal detection across large ensembles of lakes using EO data.

LAKE 2.0 provides superior mechanistic representation of turbulence, ice–snow dynamics, and energy balance, making it indispensable for high-latitude lakes where ice-cover feedbacks and snow–radiation interactions are central to tipping dynamics.

Together, the models support a two-tiered strategy: LAKE 2.0 can be applied in detailed case studies to capture mechanistic feedbacks, while ALBM can be scaled globally with ML emulators and EO

integration to detect early-warning signals and assess tipping-point risks across large ensembles of lakes.

2.3 Generic Indicators

The recent review “Ecosystem Resilience Monitoring and Early Warning Using Earth Observation Data: Challenges and Outlook” by (Bathiany et al., 2024) evaluates the potential, current challenges, and future directions of using EO data to monitor ecosystem resilience and provide early warning signals of ecological tipping points. The paper emphasizes that EO holds great potential for resilience monitoring and early-warning of ecosystem shifts, but current approaches face limitations related to data quality, theoretical clarity, and methodological rigor. It calls for cautious but ambitious development of new tools and a stronger theoretical–observational interface.

Bathiany et al., (2024) review the current state-of-the-art in methods and indicators used to detect resilience and tipping behavior in ecosystems using EO data. They assess the limitations of existing EO datasets and analytical approaches in reliably capturing early-warning signals. They also Outline methodological, technical, and conceptual challenges, such as noise, scale mismatches, and uncertainties in EO-derived indicators. Moreover, they identify opportunities for improvement through advanced modeling, machine learning, multi-sensor data fusion, and interdisciplinary collaboration. Finally, they propose a roadmap for better integration of EO in resilience science and practical early-warning applications (e.g., for forest collapse, desertification, or aquatic regime shifts).

Many of the methods reviewed in Bathiany et al. (2024), although primarily developed and applied for detecting tipping points in terrestrial ecosystems (such as the Amazon rainforest) are also highly relevant and adaptable to lake systems. Given that lakes are responsive to both climatic and catchment-scale changes, these techniques offer valuable frameworks for anticipating critical transitions in lake thermal structure, productivity, and biogeochemistry.

The following methods are particularly applicable to lake systems:

- 1) **Critical Slowing Down (CSD) indicators:** Detecting rising autocorrelation and variance in EO-derived lake variables (e.g., surface temperature, chlorophyll-a) to signal reduced resilience.
- 2) **Recovery rate analysis:** Measuring how quickly lake systems return to equilibrium after disturbances (e.g., storm events, heatwaves), using EO time series.
- 3) **Abrupt shift detection:** Identifying sudden, persistent changes in lake surface area, albedo, or optical properties, suggesting possible regime shifts.
- 4) **Spatial resilience metrics:** Using spatial patterns of turbidity, temperature, or ice cover heterogeneity as indicators of fragmentation or destabilization within large lakes.
- 5) **Multi-sensor data fusion:** Integrating optical, thermal, and radar data to track complex lake processes, especially under ice or cloud cover.
- 6) **Novel biophysical variables:** Leveraging EO products like lake surface water temperature (LSWT), floating vegetation cover, or reflectance-based proxies for dissolved organic matter.
- 7) **Machine learning and AI approaches:** Applying unsupervised learning or anomaly detection algorithms to long-term lake EO datasets to identify emerging patterns or tipping signals.
- 8) **Uncertainty-aware modeling:** Incorporating uncertainty explicitly in EO-derived indicators and model-data comparisons to ensure robustness in lake resilience assessments.

These methods, especially when integrated with in situ observations and lake modeling tools (e.g., ALBM, LAKE 2.0), can help develop early-warning systems tailored to aquatic ecosystems.

3 References

Bathiany, S., Bastiaansen, R., Bastos, A., Blaschke, L., Lever, J., Loriani, S., De Keersmaecker, W., Dorigo, W., Milenković, M., Senf, C., Smith, T., Verbesselt, J., & Boers, N. (2024). Ecosystem Resilience Monitoring and Early Warning Using Earth Observation Data: Challenges and Outlook. *Surveys in Geophysics*. <https://doi.org/10.1007/s10712-024-09833-z>

- Brechbühler, M. (2024). *SOS-Water D3.1 Data inventory*. Zenodo. <https://doi.org/10.5281/zenodo.12683903>
- Stepanenko, V. M., Machul'skaya, E. E., Glagolev, M. V., & Lykossov, V. N. (2011, April). *Numerical modeling of methane emissions from lakes in the permafrost zone*. *Izvestiya - Atmospheric and Ocean Physics*. <https://doi.org/10.1134/S0001433811020113>
- Tan, Z., Yao, H., & Zhuang, Q. (2018). A Small Temperate Lake in the 21st Century: Dynamics of Water Temperature, Ice Phenology, Dissolved Oxygen, and Chlorophyll a. *Water Resources Research*, 54(7), 4681–4699. <https://doi.org/10.1029/2017WR022334>
- Tan, Z., Zhuang, Q., & Walter Anthony, K. (2015). Modeling methane emissions from arctic lakes: Model development and site-level study. *Journal of Advances in Modeling Earth Systems*, 7(2), 459–483. <https://doi.org/10.1002/2014MS000344>