



Guidelines for Retrieving Data from the Alplakes Platform

D2.1.1 Alpine meta-data set and correlation report with multiparameter table

Name of Author(s)	Mario Hess
Department/Organization/Institution	Institute for Limnology, Mondsee
Date of the document	04/08/2026
Version	1.1
Project	DiMark



Introduction

The Alplakes platform provides satellite-derived water quality data for alpine lakes. These datasets include parameters such as lake water surface temperature (LWST), secchi depth, chlorophyll-a concentration, turbidity and more.

The objective of this report is to describe the process of retrieving data from the platform and to compare the satellite-derived results with in-situ measurements. This guideline was written based on the training provided in the PEAK course A52/26 Satellite data as a supplement to lake monitoring on 19 March 2026 by partner EAWAG.

The atmospheric correction algorithm is a current area of active development. Ongoing refinements are aimed at further improving the consistency and accuracy of satellite-derived data, especially chlorophyll-a concentration. The guidelines described in this document should remain applicable as these improvements are released.

Overview

Alplakes enables retrieval of satellite-derived datasets in different formats and spatial representations. These main types of data can be accessed:

- Lake-wide averages, min and max values of a specific parameter
- Pixel-based values of a specific parameter within a window (specific geographic coordinates within a radius)

Available outputs include:

- **.csv files:**
CSV (Comma-Separated Values) is a plain text file format used for storing tabular data, such as spreadsheets or databases, where each line represents a record and commas separate fields. Usually, most users will use Microsoft Excel for further processing.
- **.tif files:**
TIF (Tagged Image File Format) is a raster file format used for storing raster graphics images. It can be used in GIS software such as QGIS or ESA SNAP to process and represent image-based information.

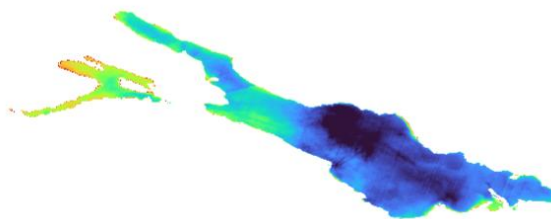


Figure 1: TIF file of lake surface temperature in QGIS with a blue-to-red ramp showing interpolated temperatures.



GUI Details

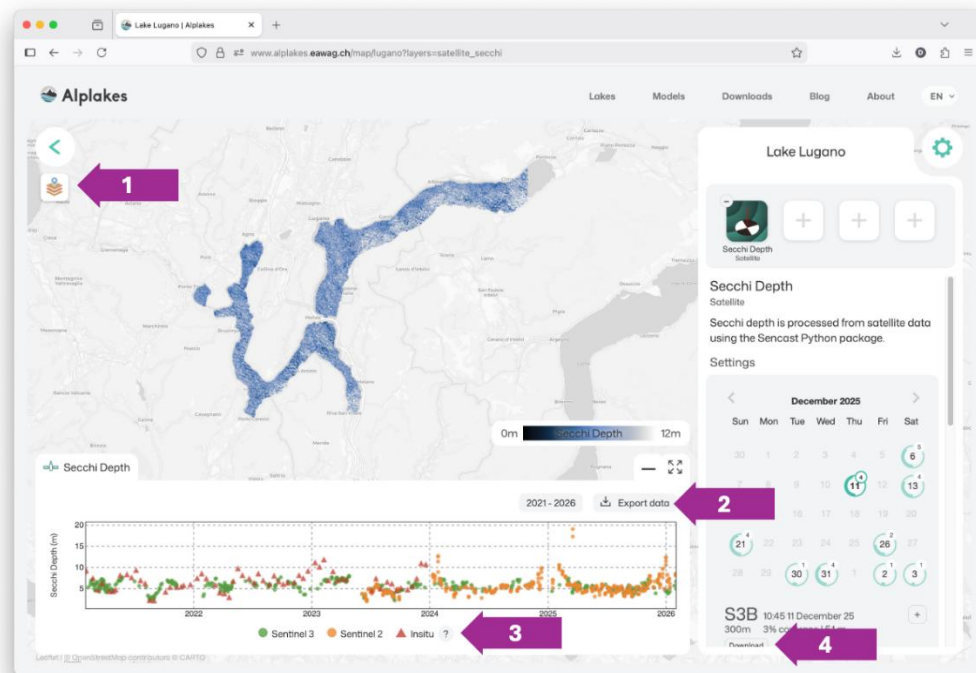


Figure 2: Screenshot of Lake Lugano's landing page on www.alplakes.eawag.ch. (1) is the button for extracting EO time series at virtual stations, (2) is the button to extract the time series visualised in the plot, (3) provides information about the in situ measurements, and (4) allows downloading individual EO raster products.

EO timeseries of lake-wide spatial averages

The initial “Sentinel-2” and “Sentinel-3” data shown in the time series plot in Figure 2 represent the spatial average of each day with valid pixels. Such spatial aggregation allows a better spatial representation and an improved temporal sampling for water quality variables that tend to be relatively homogeneous across the lake. Contrariwise, in case of heterogeneous variables, it can create artificial variations in the time series.

For example, if one half of the lake is clouded during one week, and the other half during the next week, an extraction of homogeneous lake-averaged surface water temperature can help to collect observations from both weeks. However, if one half comprises a turbid estuary, the extraction of spatially averaged turbidity will show an artificial decrease for the week in which the estuary is clouded.

Clicking the button “Export data” next to **arrow (2)** in Figure 2 downloads csv tables for the in-situ measurements and EO lake averages (and min, max) for the selected lake and variable. Note that EO data from all selected satellite missions are included in the table, regardless of whether they were selected and displayed in the time series plot.

EO timeseries at a virtual station

The slicing tool icon in the top left of the lake map (**arrow (1)** in Figure 2) allows extracting all valid observations of the selected variables at any location in the lake. Clicking the slicing tool will shade the icon change the mouse cursor to crosslines. When clicking at the lake with this crossline cursor, a satellite timeseries dialog box will open (Figure 3):

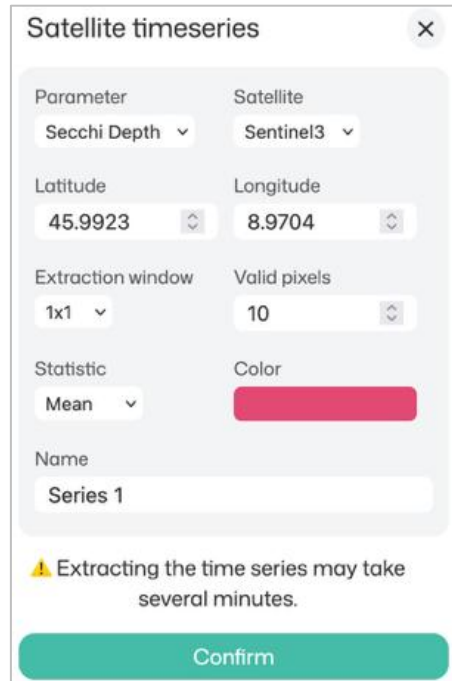


Figure 3: Satellite time series dialog box for the extraction and visualisation of EO data at virtual sampling stations. The value “Valid pixels” defines the minimum number of pixels available in a satellite image for the selected lake to be considered. In this way, cloudy images that yield notoriously poorer retrievals can be removed. Note that small lakes may contain only a couple of Sentinel-3 (300 m) pixels even when it’s cloud free, while Sentinel-2 may still observe thousands of pixels of a large lake that is mostly cloudy.

The **Latitude** and **Longitude** set in this box correspond to the location chosen with the mouse cursor. It can manually change them to an actual sampling location. Sampling location coordinates of the Alplakes in situ data can be displayed using the “?” button next to the time series labels (**arrow (3)** in Figure 2), and entered here.

The menu item **Extraction Window** sets whether a single pixel location is extracted, or whether a selectable average (Statistic: Mean) of a surrounding window is extracted.

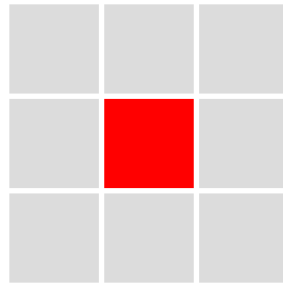


Figure 4: 3x3 pixel extraction window.

Through this step, referred to as spatial binning, very noisy EO data can be smoothed to obtain more reasonable spatial gradients. For example, most Sentinel-2 products are very noisy, such as the Secchi map of Lake Lugano in December 2025 in native 20 m resolution. By using a 5x5 averaging window, this noise is removed. Contrariwise, spatial smoothing is not necessary for naturally smooth products such as most Landsat surface water temperature data, or Sentinel-3 products that already come at coarse resolution that would decrease even more through the smoothing. Note that the effect of such spatial binning can also be assessed using the “Smoothing” slider (Figure 5):



Figure 3: Spatial binning using the ‘Smoothing’ Slider in the Alplakes map viewer, for the same Sentinel-2 Secchi depth map without (left) and with 5x5 pixel binning (right).

After clicking “Confirm” in the satellite timeseries dialog, the time series at your selected location and using your configuration will show up in the time series plot, with the extraction parameters described under the “?” next to its label item. When you click “Export data” (arrow (2) in Figure 2) now, this time series will be extracted along with the in situ data and lakewide average described in the previous section.

Comparing in situ and EO time series

Different use cases require different comparisons of in situ data and EO time series. If the user wants to derive phenological metrics, or calculate annual or monthly means, then the two timeseries are processed individually and derivatives are compared at the very end (e.g., annual EO

average vs. annual in situ average). If matching measurements within a certain time window are needed, e.g., for validation are needed, then the time window should be as small as possible, yet large enough to yield a sufficient number of so-called matchups. For example, if measurements are acquired specifically for the purpose of validating EO data, and considering dynamic areas like estuaries, then time windows may be as short as 1-2 hours.

Generally speaking, for every year of monthly in situ measurements on a 50% cloudy lake, we expect about three 24h matchups with Sentinel-3A (revisit 2-3 days) in the years 2016-2019, or six matchups with Sentinel-3A/B after Sentinel-3B's launch in 2019. In case of Sentinel-2, whose individual satellites have 10-day revisit, we expect only about one matchup with such in situ data per year in 2015-2018, and two in the years after 2019. It may hence be necessary e.g., to increase the matchup to 48 hours if the lake is extraordinarily cloudy, or if the in-situ data isn't continuous monthly since 2015.

This time window is thus an important attribute that can be set by the user for every matchup comparison, and it must be reported accordingly. The time window is calculated based on the satellite image acquisition time included in the Alplakes time series download (in datetime string format YYYY-MM-DDThh:mm:ss). Likewise, a datetime string can be extracted from the in situ timeseries, but note that acquisition time is often unknown and hence set to 12:00. Depending on your preferred data analysis language, any Large Language Model (chatGPT, Gemini, Claude etc.) can be asked to write code that "merges columns a from table x and column b from table y when their shared datetime in format YYYY-MM-DDThh:mm:ss is within T".

Retrieving Data Step-by-Step Guidelines

Data can be accessed either by downloading it from the website or through the API. The following sections serve as step-by-step guides for both methods.

Retrieving Data via the Website

Most users will opt for this method when retrieving data. The Website provides an intuitive graphical user interface with several options to specify the data needed.

1. Open the Alplakes platform in a web browser: <https://www.alplakes.eawag.ch/>
2. Search for and select the desired lake.
3. Choose the parameter of interest (e.g., chl-a, LWST, secchi depth) and click on **"See more"**.

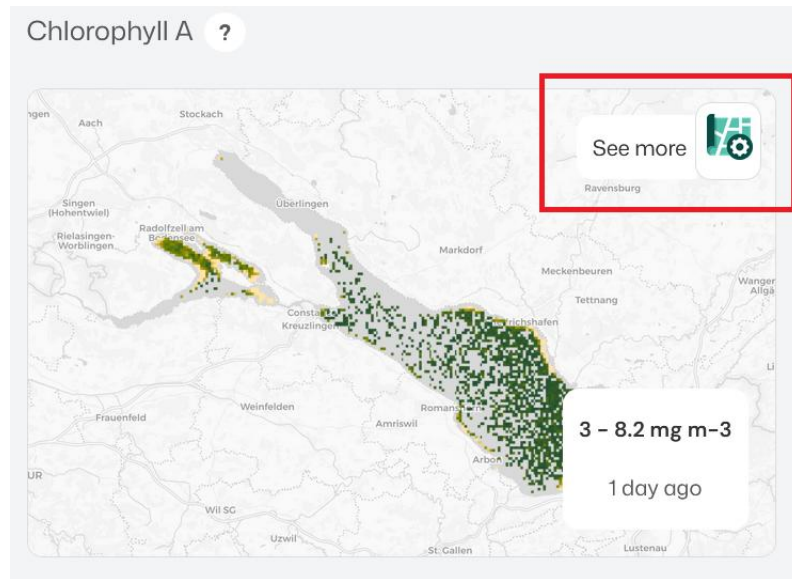


Figure 4: GUI Element for the Chlorophyll-A parameter.

4. Select the desired date:
 - (Optional) Right panel: Export of a .tif raster file for a specific date. Select the desired date and click on “**Download**” to download the file.
 - Left panel: Adjust the desired date range of the timeseries.
5. (Optional) Select a specific location by clicking on “**Satellite timeseries**” and placing a marker (as described in Figure 3):

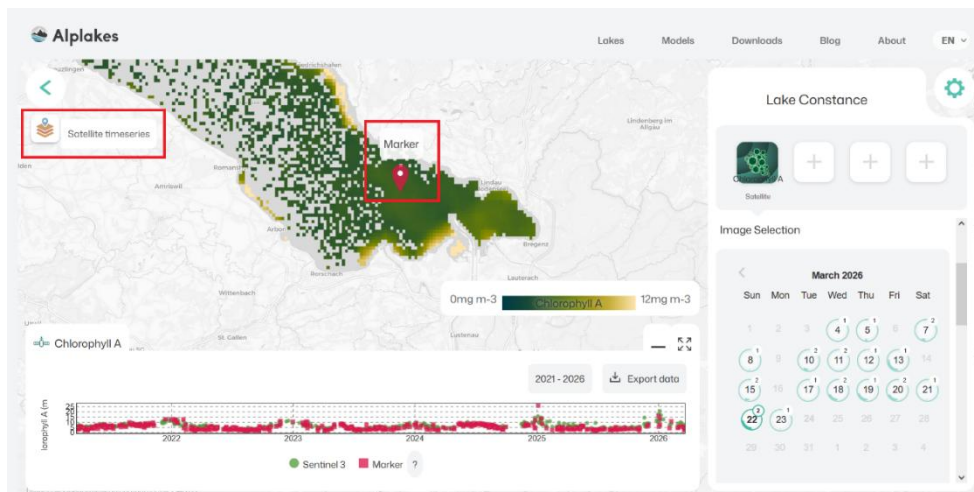


Figure 5: Placing a marker at a specific location.

- If no marker is placed, the platform returns timeseries data with lake-wide data (average, mean, min, max).
 - If a marker is placed, the platform returns timeseries data of a specific location (e.g., geographic coordinates within the extraction window) in addition to lake-wide data.
6. Click “**Export data**” to download the dataset.



In this example, since a marker is placed, the platform returns two datasets:

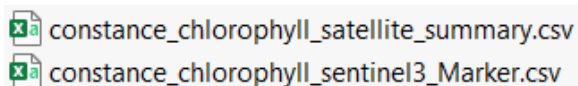


Figure 6: Retrieved Datasets

1. `constance_chlorophyll_satellite_summary.csv`:
This file contains timeseries data with the date, lake-wide average, min and max values. Download links for .tif raster files are also included:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Datetime	Satellite	Mean (mg	Min (mg m	Max (mg n	Pixel	Cove	URL												
2	2016-05-04T09:31:00.000+00:00	S3A	5.59208	3.05386	17.53888	31		https://eawags.s3.eu-central-1.amazonaws.com/alplakes/cropped/sentinel3/constance/OC3_chla_S3A_20160504T093142_sui_constance.tif												
3	2016-05-03T09:57:00.000+00:00	S3A	5.45566	2.57648	38.76119	14		https://eawags.s3.eu-central-1.amazonaws.com/alplakes/cropped/sentinel3/constance/OC3_chla_S3A_20160503T095753_sui_constance.tif												
4	2016-04-27T09:13:00.000+00:00	S3A	5.95105	2.77861	14.51166	11		https://eawags.s3.eu-central-1.amazonaws.com/alplakes/cropped/sentinel3/constance/OC3_chla_S3A_20160427T091300_sui_constance.tif												
5	2016-05-07T09:54:00.000+00:00	S3A	7.21561	3.25552	23.29832	80		https://eawags.s3.eu-central-1.amazonaws.com/alplakes/cropped/sentinel3/constance/OC3_chla_S3A_20160507T095408_sui_constance.tif												
6	2016-05-15T09:46:00.000+00:00	S3A	10.14316	5.52171	18.95463	1		https://eawags.s3.eu-central-1.amazonaws.com/alplakes/cropped/sentinel3/constance/OC3_chla_S3A_20160515T094640_sui_constance.tif												
7	2016-05-14T10:12:00.000+00:00	S3A	9.56912	8.58883	10.77247	1		https://eawags.s3.eu-central-1.amazonaws.com/alplakes/cropped/sentinel3/constance/OC3_chla_S3A_20160514T101251_sui_constance.tif												
8	2016-05-22T10:05:00.000+00:00	S3A	9.08801	4.65326	18.16886	68		https://eawags.s3.eu-central-1.amazonaws.com/alplakes/cropped/sentinel3/constance/OC3_chla_S3A_20160522T100522_sui_constance.tif												

Figure 7: Timeseries lake-wide.

2. `constance_chlorophyll_sentinel3_Marker.csv`:
This file contains the timeseries data with the date, average, median, min and max values at the given coordinates within the specified extraction window:

	A	B	C	D	E	F	G
1	Datetime	Mean (mg	Median (m	Min (mg m	Max (mg n	Count	Std
2	2026-03-2	4.129938	4.129938	4.129938	4.129938	1	0
3	2026-03-2	4.997629	4.997629	4.997629	4.997629	1	0
4	2026-03-2	4.551517	4.551517	4.551517	4.551517	1	0
5	2026-03-1	2.655519	2.655519	2.655519	2.655519	1	0
6	2026-03-1	3.629955	3.629955	3.629955	3.629955	1	0
7	2026-03-1	4.288596	4.288596	4.288596	4.288596	1	0
8	2026-03-1	5.793848	5.793848	5.793848	5.793848	1	0

Figure 8: Timeseries location-specific

Retrieving Data via the API

In addition to the graphical user interface, the platform provides an API that returns data in the JSON format, enabling programmatic retrieval or integration into custom data-processing pipelines. The documentation can be found at: <https://www.alplakes.eawag.ch/downloads>.

This section describes the process of constructing requests and inspecting responses in a web browser.

The root endpoint of the Remote Sensing API is <https://alplakes-api.eawag.ch/> and is defined as follows:

Remote Sensing		^
GET	/remotesensing/metadata Remote Sensing Metadata	∨
GET	/remotesensing/products/{lake}/{satellite}/{variable} Remote Sensing Products	∨
GET	/remotesensing/timeseries/{lake}/{satellite}/{variable}/{lat}/{lng}/{start_date}/{end_date} Remote Sensing Products	∨

Figure 9: Definition of the Remote Sensing API.

Metadata Endpoint

</remotesensing/metadata>

This endpoint provides information about available parameters, which is necessary to construct requests for the dynamic endpoints listed in Figure 13:

1. </remotesensing/products/...>
2. </remotesensing/timeseries/...>

By opening <https://alplakes-api.eawag.ch/remotesensing/metadata> in a web browser, the response returns metadata describing available lakes, satellites, variables and their corresponding identifiers:

```

constance:
  collection:
    0: "ST"
  sentinel2:
    0: "SPM_Nechad2016_665"
    1: "TUR_Nechad2016_665"
    2: "Z490"
    3: "forel_ule"
    4: "p3qaa_zSD"
    5: "tsm_dogliotti665"
  sentinel3:
    0: "Zsd_lee"
    1: "a_ph443"
    2: "area_bgr"
    3: "bgr_whit"
    4: "chla"
    5: "forel_ule"
    6: "pp_integral"
    7: "tsm_binding754"
  
```

Figure 10: Part of the response of the Metadata endpoint.

This metadata provides the necessary reference for constructing requests to the dynamic endpoints. It lists the identifiers for lakes, satellites, and variables, allowing users to specify which data they want to retrieve in API calls. In other words, the metadata serves as a catalog of available options and their corresponding IDs, ensuring requests are correctly formed and valid.

Products Endpoint

/remotesensing/products/{lake}/{satellite}/{variable}

This endpoint provides access to .tif raster files suitable for GIS processing.

There are two additional query options:

- `start_date`: Start of the date range
- `end_date`: End of the date range

Example: Retrieve a timeseries of raster files for **Lake Constance**, **Sentinel-3** satellite, and the **chl-a** variable for **2025**.

To achieve this, the placeholder variables have to be replaced by the target metadata values. To specify the date range, a **query string** has to be used at the end of the URL:

```
https://alplakes-api.eawag.ch/remotesensing/products/constance/sentinel3/chla?start_date=20250101&end_date=20251230
```

By opening this URL in a web browser, the response can be inspected. It contains a list of download links for .tif files for the specified lake, satellite, variable and date range:

```

▼ 0:
  time:           "2025-01-02T10:16:51+00:00"
  full_tile:      "https://eawags.s3.eu-central-1.amazonaws.com/dataLakes/sui/S3A_20250102T101651/OC3_chLa_S3A_20250102T101651.tif"
  lake_subset:    "https://eawags.s3.eu-central-1.amazonaws.com/alplakes/cropped/sentinel3/constance/OC3_chLa_S3A_20250102T101651_constance.tif"
  valid_pixels:   "1%"
▼ 1:
  time:           "2025-01-04T09:24:29+00:00"
  full_tile:      "https://eawags.s3.eu-central-1.amazonaws.com/dataLakes/sui/S3A_20250104T092429/OC3_chLa_S3A_20250104T092429.tif"
  lake_subset:    "https://eawags.s3.eu-central-1.amazonaws.com/alplakes/cropped/sentinel3/constance/OC3_chLa_S3A_20250104T092429_constance.tif"
  valid_pixels:   "17%"
    
```

Figure 11: Part of the response of the Products endpoint.

Timeseries Endpoint

/remotesensing/timeseries/{lake}/{satellite}/{variable}/{lat}/{lng}/{start_date}/{end_date}

This endpoint returns timeseries data of a specific variable, at a specific geographic location of a lake (within a specified extraction window) for a specific date range.

There are two additional query options:

- `window`: Extraction window radius (0=1px, 1=9px, 2=25px...)
- `valid_pixels`: Minimum percentage of valid pixels

Example: Retrieve a **chl-a** timeseries for **Lake Constance** at **latitude 47.5189** and **longitude 9.7386** for **2025**, with an **extraction window radius** of 3x3 pixels and a **minimum of 50%** of valid pixels.

To achieve this, the placeholder variables have to be replaced by the target metadata values. To specify the size of the pixel window and minimum percentage of valid pixels, a [query string](#) has to be used at the end of the URL:

https://alplakes-api.eawag.ch/remotesensing/timeseries/constance/sentinel3/chla/47.5189/9.7386/20250101/20251230?window=1&valid_pixels=50

By opening this URL in a web browser, the response can be inspected. It contains a list of average, median, min and max values for the specified lake, satellite, variable, latitude, longitude, date range, extraction window radius and minimum percentage of valid pixels:

```

  0:
    time:      "2025-12-12T09:19:41+00:00"
    value:
      mean:    21.75567626953125
      median:  21.75567626953125
      min:     21.75567626953125
      max:     21.75567626953125
      std:     0.0 JS: 0
      count:   1
  
```

Figure 12: Part of the response of the Timeseries endpoint.



Comparison Results

Lake Water Surface Temperature

Lake Surface Water Temperature - Lake Constance

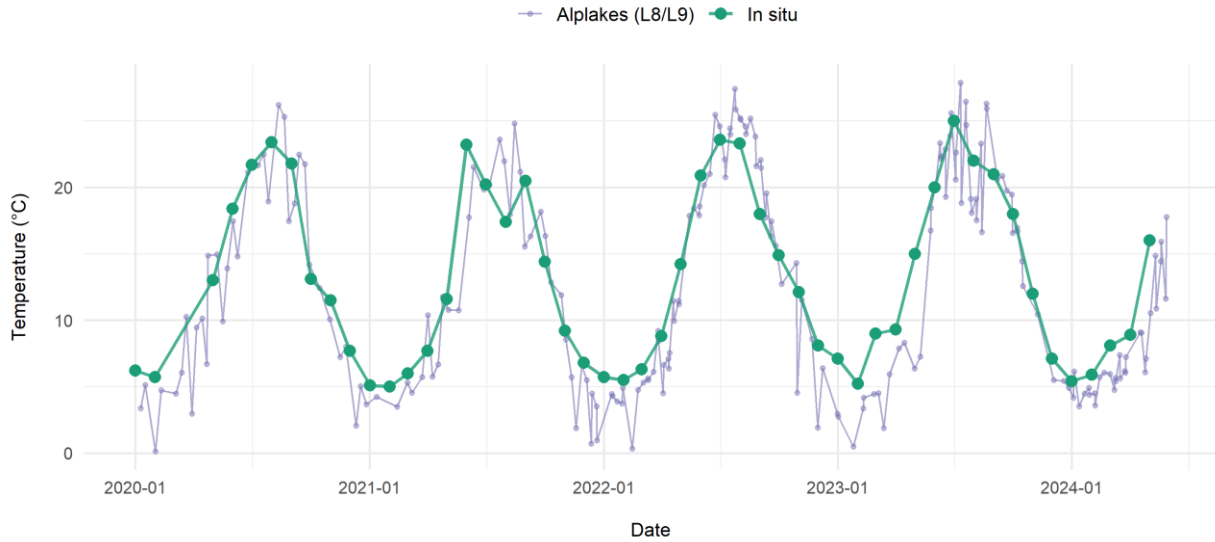


Figure 13: LWST comparison Lake Constance

Secchi Depth

Secchi Depth - Lake Constance

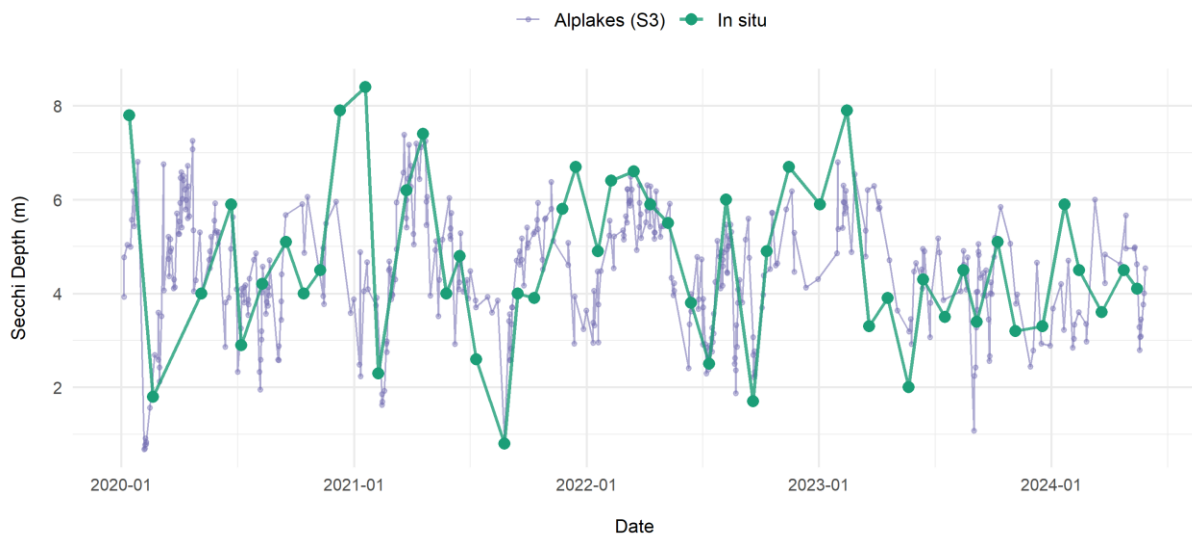


Figure 14: Secchi Depth comparison Lake Constance



Cholorphyll-A

Chl-a - Lake Constance

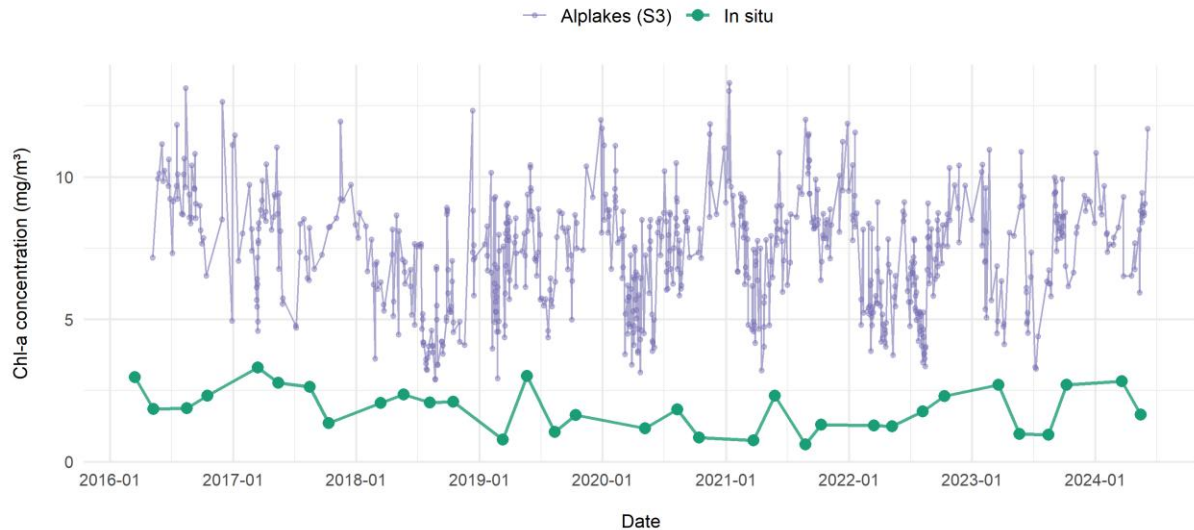


Figure 15: Chlorophyll-A comparison Lake Constance

Note: The Chl-a in-situ measurements are based on 20 m composite samples, so exact agreement with satellite-derived Chl-a estimates is highly unlikely.

Glossary

TIF (Tagged Image File Format)

Raster file format for GIS systems

CSV (Comma-Separated-Values)

Text file format used for storing tabular data

JSON (JavaScript Object Notation)

Data format used to store and transport data

API (Application Programming Interface)

Software communication

Endpoint

URL for accessing an API address



Query string

Part of a URL that passes parameters to an API

Request

A message sent to an API to retrieve or send data

Response

Data returned by an API after a request