Scuola universitaria professionale
della Svizzera italiana

SUPSI

UNIVERSITÉ DE GENÈVE

Final project for complementary certification in geomatics

Automation of Satellite Image Downloading and Processing for Lake Temperature and Algae Monitoring

Realised by

Yara Leone dos Santos

Under the direction of

Daniele Strigaro
(SUPSI)
&
Gregory Giuliani
(UNIGE)

December 2025 – January 2026

Abstract

This work addresses the challenge of accessing and processing satellite data for environmental monitoring activities conducted at the Institute of Earth Sciences (IST) of the University of Applied Sciences and Arts of Southern Switzerland (SUPSI), in the Geomatics sector. Two main projects were studied: the analysis of the land's surface temperature for the assessment of urban heat islands within the framework of a cantonal mandate, and the monitoring of the surface temperature of the lake and of algae concentration of Lake Lugano within the framework of the International Commission for the Protection of Italian Swiss Waters (CIPAIS) mandate.

Automated workflows were developed to download, process and organize satellite imagery from Sentinel-3 OLCI and Landsat-8/9 Collection 2 Level-2 products using open-source tools and designed to be reproducible by non-geomatics experts.

The implemented scripts for the Lake Lugano monitoring reproduce, as closely as possible, the manual processing chains previously applied by experts following the SIMILE project protocols, while improving reproducibility, traceability and processing efficiency.

The resulting products consist of documented inventories of satellite scenes and georeferenced images directly usable in GIS software. Comparison with manually processed reference products from the 2024 CIPAIS report shows strong agreement in spatial patterns and statistical distributions, confirming the reliability of the automated approach. The remaining differences can be attributed to temporal variability between the processed and reference images, or to shifts due to automated processing. However, they remain similar and could be used to analyze lake quality indicators such as water surface temperature and algae concentration.

Overall, this work demonstrates that automation can significantly reduce satellite image processing time while maintaining scientific rigor. This automation can be implemented by individuals without programming expertise, provided they use the appropriate tools. It lays the groundwork for future extensions such as continuous, long-term monitoring, near real-time processing, and the integration of additional satellite missions for improved data comparison. Ultimately, this approach can facilitate the sharing of information with a wider audience.

The transmission of the complete project was highlighted so that it could be reviewed and improved before the actual use of the processed products ("Yara-Leone_SUPSI_project" on GitHub)

Table of contents

Table of figures

List of acronyms

API = Application Programming Interface
OAuth = Open Authorization
Chla = Chlorophyll-a
CIPAIS = International Commission for the Protection of Italian Swiss Waters
CSDE = Copernicus Data Ecosystem
C2RCC = Case-2 Regional CoastColour
DACD = Department of Environment, Construction and Design
EFR = Extended Full Resolution
EPSG = European Petroleum Survey Group
ESA = European Space Agency
GEE = Google Earth Engine
GeoTIFF = Georeferenced Tagged Image File Format
GIS = Geographic Information System
IST = Institute of Earth Sciences
OLCI = Ocean and Land Colour Instrument
OLI = Operational Land Imager
LST = Land/ Lake surface Temperature
MODIS = Moderate Resolution Imaging Spectroradiometer
NTC = Non-Time-Critical
NASA = National Aeronautics and Space Administration
PNG = Portable Network Graphic
TIRS = Thermal Infrared Sensor
SIMILE = Informative System for the Integrated Monitoring of the Insubrian Lakes and their Ecosystems
SLSTR = Sea and Land Surface Temperature Radiometer
SNAP = Sentinel Application Platform
STAC = SpatioTemporal Asset Catalog
SUPSI = University of Applied Sciences and Arts of Southern Switzerland
UNIGE = University of Geneva
URL = Uniform Resource Locator
USGS = United States Geological Survey
UTM = Universal Transverse Mercator
VIIRS = Visible Infrared Imaging Radiometer Suite
WFR = Wide Full Resolution
WKT = Well-Known Text
(ZIP = Compressed archive file format)

1. Introduction

Satellite remote sensing has become a central component of environmental observation. Its ability to provide frequent, wide-area measurements offers significant advantages over traditional field-based monitoring, particularly as ecosystems are increasingly affected by climatic variability, extreme weather events, and human-driven transformations (Parevez Mosharof, 2025). Parameters such as land surface temperature, vegetation dynamics, and water quality indicators, including surface temperature and chlorophyll-a concentration, can now be assessed consistently over time through multispectral and thermal satellite missions (Gholizadeh et al., 2016; ESA, 2023). These advances have substantially expanded the possibilities for analyzing environmental processes at regional to global scales.

At the same time, the growing availability of Earth-observation data has introduced new challenges. Many datasets are distributed as large archives requiring careful selection, downloading, and preprocessing before they can be integrated into scientific or operational workflows (Sudmanns, 2018). Differences in spatial resolution, revisit frequency, radiometric characteristics, and product formats across satellite missions further complicate their use but can also be complementary. Turning raw satellite acquisitions into meaningful environmental information often demands substantial technical effort, specialized knowledge, and considerable time. These constraints affect researchers, public institutions, and monitoring programs that rely on regular and comparable observations.

Recent developments in open-source tools and automated processing pipelines offer opportunities to overcome these limitations. By streamlining the retrieval, organization and treatment of satellite images, such approaches can improve reproducibility, reduce manual workload and make remote-sensing data more accessible to a broader range of users. This is particularly relevant in applied contexts where environmental assessments must be carried out rapidly or periodically, and where the consistent use of satellite observations contributes to long-term understanding of environmental change.

1.1. Host Organization Overview

The Geomatics sector of SUPSI is part of the Institute of Earth Sciences, which is dedicated to sustainable management and protection of land, environment, and natural resources (SUPSI, 2025). Beyond pure research, IST conducts applied research and technological development, provides services to institutions and industries, and offers education (SUPSI, 2025). It thereby contributes directly to regional and national sustainability efforts.
The Geomatics team specializes in acquiring, managing, processing, and sharing georeferenced data. Through these activities, it generates evidence-based knowledge to support decision-making for land use planning and infrastructure management. The sector collaborates extensively with the other sectors of the IST in environmental monitoring.
Many of IST's projects are commissioned by public authorities, notably by the authorities of the Canton of Ticino or by international organizations, underlining the strong applied dimension of the work and its relevance for policy, land management and environmental governance. This close collaboration ensures that research outputs are tightly linked to real-world needs.

1.2. Project Overview

The internship carried out with the Geomatics sector during the last quarter of 2025 focused on this collaboration of the IST's sectors by dealing with two projects mandated by international public institutions.

Lugano lake monitoring

An important mandate of the SUPSI is given by CIPAIS, a cross-border cooperation program, which is responsible for monitoring, assessing, and protecting shared transboundary waters. Its mission includes identifying sources of pollution, evaluating the ecological condition of lakes and rivers, and guiding coordinated management actions between Switzerland and Italy (DACD SUPSI, 2023). Within this framework, the aquatic ecology sector of the IST conducts the limnological research program for Lake Lugano, a lake historically affected by eutrophication. The work focuses on monitoring water quality (temperature, mixing, oxygen, nutrient inputs) in lakes and rivers by in situ measurement and by integrating satellite imagery to track surface temperature and algae dynamics. These results support the CIPAIS Control Panel and Action Plan, guiding restoration efforts (DACD SUPSI, 2023). The monitoring of

Lake Lugano is a big part of the work of the aquatic ecology sector and the comparison between the in situ measurements and the telemetric images of surface temperature and chlorophyll-a concentration is currently performed using a manual process, which takes about one week of processing per year.

The SIMILE project (Informative System for the Integrated Monitoring of the Insubrian Lakes and their Ecosystems) was part of a collaborative program between Switzerland and Italy aimed at establishing an information system for the monitoring of the insubrian lakes (Lake Lugano, Lake Maggiore, Lake Como, and Lake Varese). It was mandated by the Interreg Italy-Switzerland (insubrilakes.eu, n.d.). The objective was to integrate data from various sources such as in situ measurements, satellite imagery and citizen data, and make it available at high spatial and temporal resolution to a wide audience. This has enabled a greater number of stakeholders to quickly detect critical situations and facilitates the integrated management of these lakes. The Geomatics sector worked on this project in collaboration with the Polytechnic University of Milan (Polimi), focusing on the processing of satellite images. The mandate ended in 2023, but the information remains accessible on the website "insubrilakes.eu".

Heat islands

In parallel with the main subject of this work, a few hours of geomatics processing were dedicated to a project of the "Area and territorial development" sector. As part of a mandate issued by the Canton of Ticino concerning urban heat island assessment in 2025, a dataset of satellite imagery related to land surface temperature was requested. The purpose was to determine which and how many images can be found in the canicular period, from June to September.

Staff working in other IST sectors are not geomatics experts so, as for the annual report of the CIPAIS, a lot of time is taken to find and process good images. The main limit of these monitoring projects is the time spent during image processing.

1.3. Research question

In this context, the Geomatics sector of SUPSI is involved in facilitating access to satellite images for the various collaborators of the institution.

The central issue of this work can be summarized as follows:

How can the downloading and processing of satellite images for temperature and algae monitoring be facilitated?

Of course, process automation is the key, and a series of scripts accompanied by explanations for collaborators is the central point of this work. The need lies in downloading, scientific processing, and viewing satellite images.
Throughout the work, the intention was to use open-source tools as much as possible.

2. Method and work steps

To carry out this project properly, it is divided into 3 distinct stages:

1) Review of existing work and scientific literature

The first phase was devoted to studying current knowledge on satellite monitoring of lakes and surface temperature, as well as previous studies conducted on small alpine lakes. This literature review provided the conceptual framework and ensured continuity with existing research practices (SIMILE project). A summary of important theoretical points can be found in chapter 2.1.
The different image processing methods used in the SIMILE protocols were tested before their automation, and the automation project of a previous intern was used as a basis for the work.

2) Selection and adoption of open-source software

Following various tests, suitable software for data extraction, processing, and analysis was identified. The selection of these open-source tools is presented visually at the end of Chapter 2.2.

3) Division of the project into different scripts

The project has been organized into independent but interconnected components:

- Downloading: automated acquisition of satellite imagery platforms
- Processing: implementation of the necessary models and masks, reprojections and geometry extraction
- Visualization: display of maps to support interpretation and reporting

Chapters 2.4 and 2.5 expose the logic and some explanation of the scripts used for downloads and image processing.

2.1. Satellite-Based Environmental Monitoring

This chapter presents the key theoretical concepts necessary to understand the role of different Earth observation satellites and the types of images they provide in environmental monitoring.
Many environmental variables, such as surface water temperature or chlorophyll-a concentration, can be estimated from satellite observations because land and water reflect solar and thermal radiation in ways that can be detected from space (Gholizadeh et al., 2016).
To perform these measurements, satellites are equipped with sensors, which can be broadly grouped into five main categories: thermal sensors, optical sensors, radar sensors, hyperspectral sensors, and lidar sensors (NASA, 2019). For analyzing lake surface temperature and algae concentration, estimated through chlorophyll-a concentration, thermal and optical sensors are used. The application of established physical models is necessary for processing satellite images (Siddique-E-Akbor et al., 2011).
The heat emitted by water surface is measured by its thermal infrared radiation (figure 1), and this signal is converted into lake surface temperature (LST, ESA n.d.). However, the same principle can be applied to land surface, such as studying urban heat islands.

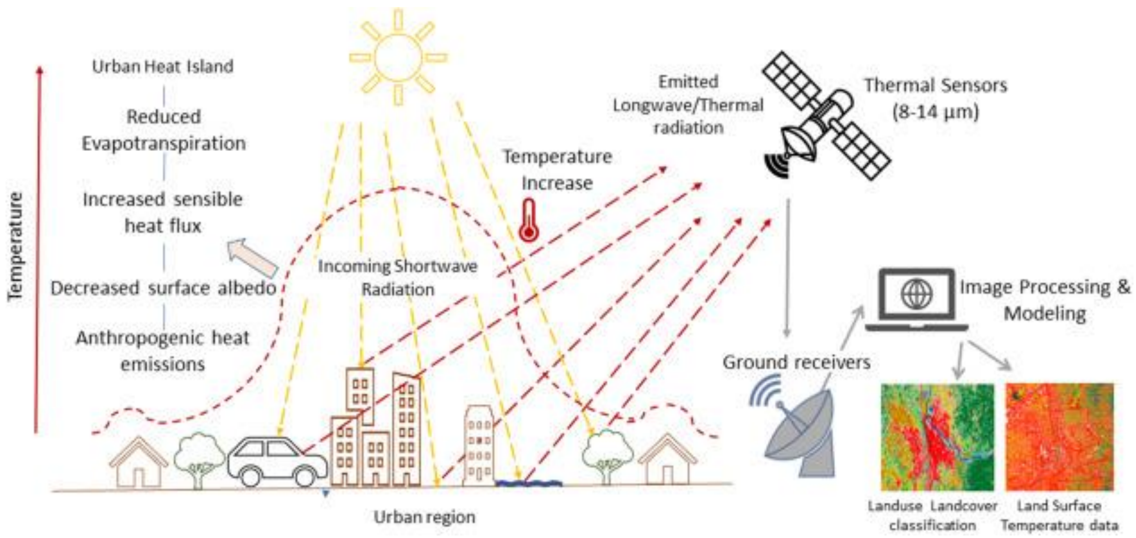


Figure 1 : Thermal sensor (sciencedirect.com)

Chlorophyll-a, which is abundant in lake phytoplankton and algae, absorbs sunlight and reflects it (Figure 2). Its concentration is used for detect presence of algae because is the primary photosynthetic pigment in algae and is therefore a key indicator of algal biomass.

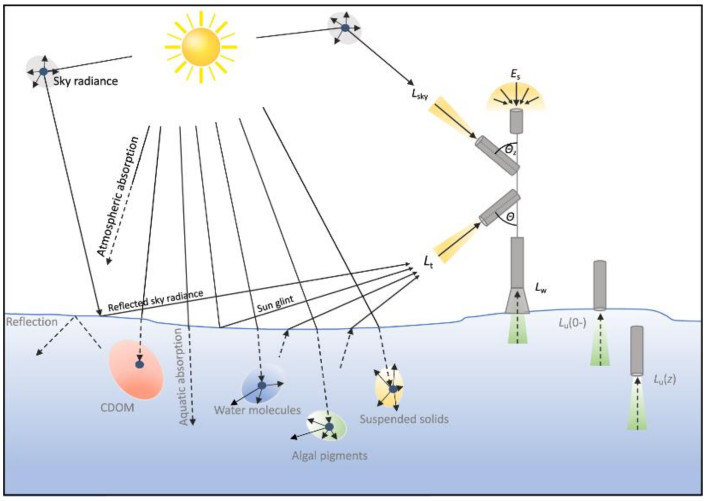


Figure 2 : Optical sensor (nature.com)

Optical sensors record this sunlight emission, and since chlorophyll-a has its own optical spectrum (Schriber, 2013), illustrated on figure 3, it can be detected by multispectral instruments (Bresciani et al., 2011). Chlorophyll-a strongly absorbs light in the blue–violet (approximately 400–450 nm) and red (approximately 600–700 nm) regions of the visible spectrum, while it reflects green wavelengths (approximately 500–600 nm). In this logic, it absorbs sunlight in the blue and red wavelengths and reflects green light.

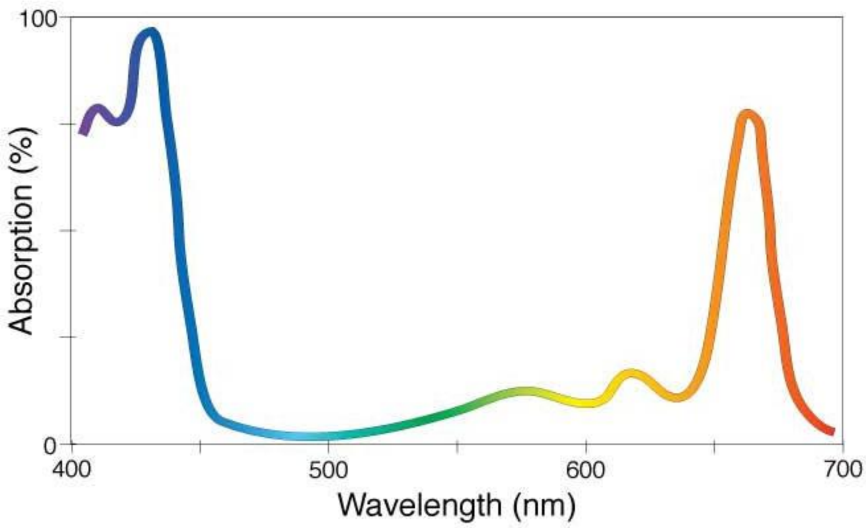


Figure 3 : Typical absorption spectrum of chlorophyll-a in the visible domain (adapted from Mobley, 1994)

The choice of satellite depends on the type of sensor and its bands, resolution and frequency of passage and the launch year.

The satellites most commonly used in the studies reviewed are the following:

- Sentinel-3 (DACD SUPSI, 2023)
- Sentinel-2 (Ansper, 2018)
- Landsat 8 and 9 (DACD SUPSI, 2023)
- MODIS and VIIRS (Bresciani et al., 2011)

In this work, the Sentinel-3 and Landsat 8 and 9 satellites have been chosen because they are the reference satellites used in the annual reports of CIPAIS. Their resolution and pass frequency complement each other well, and their sensors are sufficient to obtain the desired information on surface temperature and chlorophyll-a concentration. Both Sentinel and Landsat missions use different instruments.

The Sentinel-3 satellite, developed by the European Space Agency (ESA) for the Copernicus program, carries two main sensors relevant to this work:

- the Ocean and Land Color Instrument (OLCI), an optical spectrometer with 21 spectral bands, designed specifically to detect water optical properties, algae, suspended sediments, and other bio-optical properties (ESA Sentinel-3 Team, 2017)
- the Sea and Land Surface Temperature Radiometer (SLSTR), a thermal instrument that measures surface temperature over land and water with high radiometric accuracy (ESA, 2023; ESA Sentinel-3 Team, 2017).

Its pass frequency is daily, and the resulting image resolution is 300 meters. The satellites were launched in 2010 (3A) and 2018 (3B).

The example shown in Figure 4 illustrates the spatial distribution of luminance in the Oa08 band (665 nm) from the OLCI sensor onboard Sentinel-3, in the southern part of the canton of Ticino. The image is displayed using the "derived from cc_chl" color scale in the SNAP software, used here solely for visualization purposes.
The Oa08 band, centered at 665 nm, corresponds to a wavelength strongly absorbed by chlorophyll-a. Thus, the luminance variations observed at this wavelength are sensitive to the presence of phytoplankton in surface waters, without directly representing its concentration. Terrestrial areas exhibit high luminance (in red) due to their high reflectance in the visible spectrum, while aquatic areas appear with lower luminance (in blue), as the presence of chlorophyll-a results in increased absorption of radiation.
Although the luminance of the Oa08 band is not a direct measure of chlorophyll-a concentration, it represents an essential input variable for bio-optical algorithms used to estimate this concentration. It thus allows for an initial visualization of the satellite signal used in chlorophyll-a calculations.

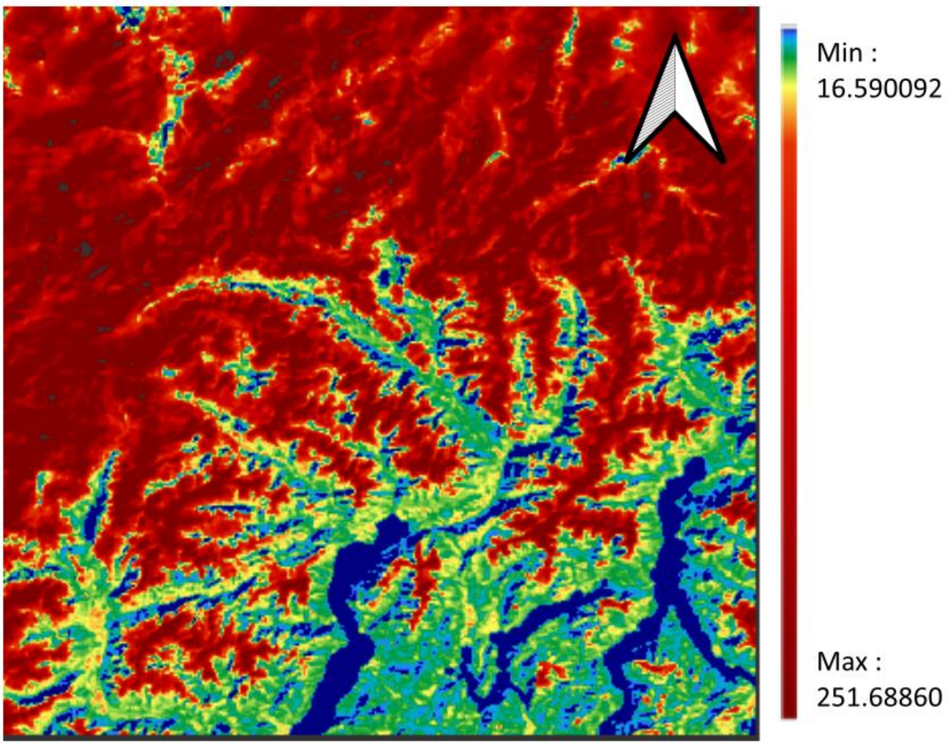


Figure 4 : band Oa08 radiance of Sentinel-3 OLCI

The Landsat 8 and 9 missions, jointly operated by NASA (National Aeronautics and Space Administration) and USGS (United States Geological Survey), carry two complementary sensors:

- the Thermal Infrared Sensor (TIRS), which measures thermal radiation and is used to retrieve surface temperature (NASA, 2019)
- The Operational Terrestrial Imaging Imager (OLI), which is a multispectral optical estimation instrument usable for clear water systems thanks to its visible and near-infrared bands. Although it can be used for chlorophyll-a analysis, this sensor and its bands were not used in this work.

The pass frequency is weekly, but the resolution is 100 meters. The satellites were launched in 2013 (Landsat 8) and 2021 (Landsat 9). An example of Land Surface Temperature (LST) derived from Landsat 9 TIRS Band 10 (Level-2 product), displayed using a continuous pseudo-colour scale is shown in figure 5. The ST_B10 band represents the land surface temperature retrieved from thermal infrared observations of the Landsat TIRS sensor. It provides an estimate of the physical temperature of the land surface, rather than air temperature, and reflects the combined effects of surface properties, solar heating and land cover.

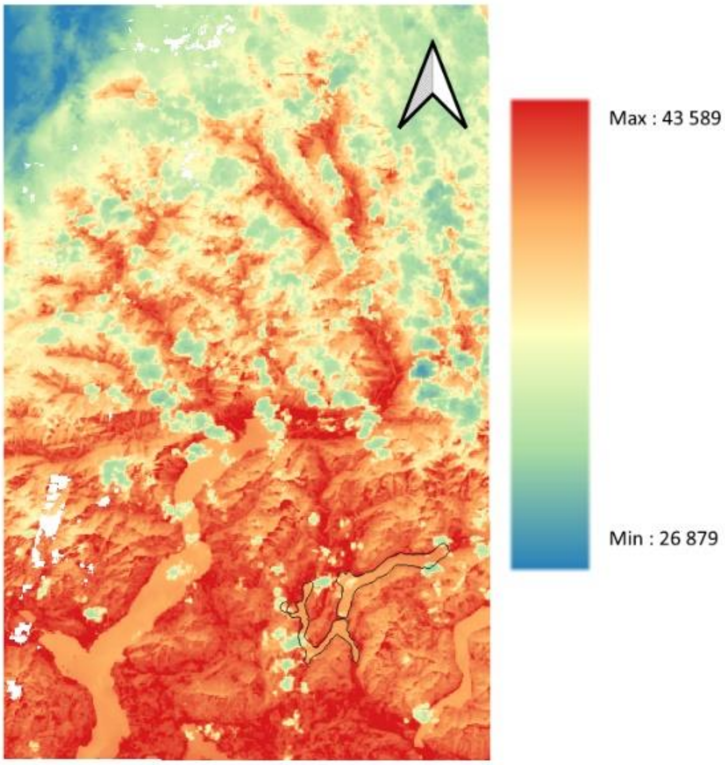


Figure 5 : Band ST_B10 of Landsat-9 TIRS

2.2. Platform and software selection

A series of exploration tests were conducted to identify the most suitable platforms and software tools for satellite data retrieval and analysis. These tests, combined with a review of existing procedures, helped to determine which solutions were the most reliable and efficient for the workflows required in this project. The selected tools and platforms have been consolidated and are summarized in the schematic overview presented in figure 6. The satellite images come from two platforms: USGS and Copernicus. They provide open-source images of Sentinel-3 and Landsat 8 and 9 satellites. Proficiency in Visual Studio Code with Python programming is essential for automation. SNAP software was used to process chlorophyll-a concentration images, and finally, QGIS software was used to visualize maps of the processed data. Artificial intelligence tools were used as coding assistance in this work. Their proper use relies on the ability to formulate precise and unambiguous instructions, such as avoiding conflicting requests, isolating a single code block to be modified without altering the rest of the script, and explicitly preserving the original structure and formatting. The quality and reliability of the generated code therefore depend largely on the user's capacity to define constraints, validate outputs, and maintain full control over the development process.

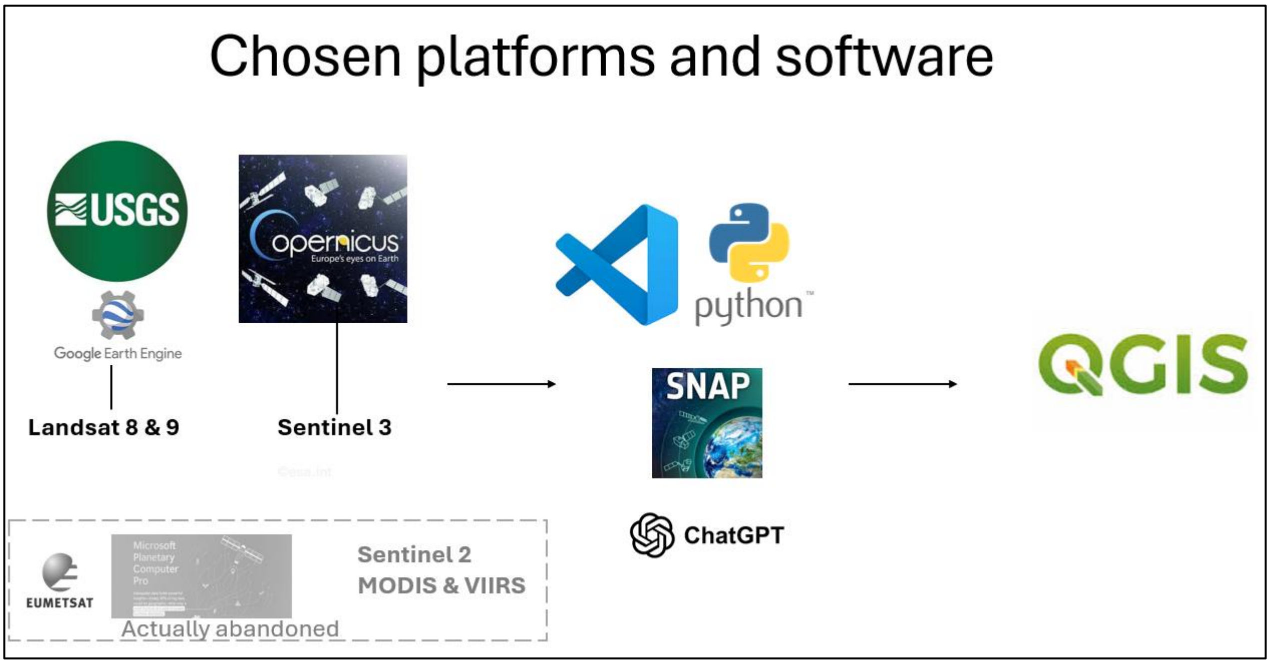


Figure 6 : Visual diagram of the selected platforms and software

2.3. Transmission logic

The submission of the final work is a crucial element to consider throughout the process. Bugs, limitations, and specific workarounds must be documented to facilitate adoption by collaborators. In addition to scripts created and uploaded to a platform to make them accessible (GitHub), an explanation of the process is indispensable. A few files needed to get started with the scripts are included with the raw scripts. The "README.md" file, which presents and explains the project, and the input data are required. The "requirements.txt" file specifies the Python dependencies necessary to run the scripts, and the ".gitignore" file defines the files ignored by Git and not included in the repository.

2.4. Satellite Image Downloading

Authentication is the process used to verify the identity of a user before granting access to a platform, an API, or a service. Both USGS and Copernicus platforms require authentication to control access to satellite data services, ensure proper usage, and monitor downloads. Platforms commonly rely on account-based authentication and API keys or session tokens for automated access.

Copernicus platform, particularly through the Copernicus Data Space Ecosystem, uses OAuth-based authentication, providing access tokens that can be reused across multiple Copernicus services.
Direct download from the USGS platform could not be established during the tests. Instead, data access was achieved through Google Earth Engine (GEE), which provides integrated access to USGS datasets. This approach required authentication via a Google account.

A STAC (SpatioTemporal Asset Catalog) is a standardized way to describe and organize geospatial datasets, particularly satellite imagery. In the context of image download, STAC allows users to search, identify, and access datasets based on spatial extent, acquisition date, sensor, or other metadata. To successfully retrieve data using STAC, it is essential to provide the exact name of the collection being queried, as STAC catalogs rely on precise identifiers. STAC does not handle authentication because it only defines how data collections and assets are described and discovered.

The flowcharts summarizing the download steps are in Annex 1 and 2, taken the example of the Heat islands project, and the scripts are available on GitHub (“Yara-Leone_SUPSI_project).

2.4.1 Heat islands monitoring

The characteristics of the Landsat 8 and Landsat 9 satellites complement those of the Sentinel-3 satellites well. This combination makes them particularly suitable for local studies, such as the detection of urban heat islands in the study of daily dynamics, but also for a more precise study of the effect in cities.

The download scripts used for Landsat and Sentinel-3 not only retrieve the images but also generate Excel files summarizing the available satellite products, as shown in the examples in Annexes 1 and 2. This approach directly addresses the canton's request for comprehensive documentation of existing imagery. For Landsat imagery, two informative Excel files are generated. The first contains images available directly on the USGS platform, but whose download script does not allow it, and the second contains information on downloads made directly from GEE. The difference in the number of images found is significant, so it is worthwhile to keep both files for quality reasons.

Integrating the shapefile of the Canton of Ticino into the user settings streamlines processing and facilitates interpretation. An additional column has been added to the Excel files to indicate the percentage of the canton covered by each image. This allows for an initial quality assessment before more in-depth analysis.

Since the canton’s primary objective was to determine how many images exist for the heatwave period, image processing was not performed. However, this would be a logical next step should SUPSI receive a mandate to further analyze these data. Notably, all necessary bands have already been downloaded.

2.4.2. Lake Lugano monitoring

In the annual report of CIPAIS, Landsat 8 and 9 images are used for surface temperature estimation, and Sentinel-3 OLCI EFR images are used for chlorophyll-a. The necessary image collections have been therefore selected for automated download within the framework of this project so that SUPSI limnologists could continue to compare indicators derived from satellite data with in-situ measurements.

The 30% cloud cover limit, request in scientific analyses, is intentionally removed because it applies to the entire image and sometimes risks excluding an image where the lake would still be visible despite some indicated cloud cover.

All available images are downloaded, but it is important to keep in mind that the images used in CIPAIS reports must be within a 5-day period before or after the in situ sampling used for comparison in the report. If the image quality is insufficient, a longer time may be used.

Sentinel-3 OLCI images

The retrieval of Sentinel-3 OLCI Level 1 EFR products are automatically downloaded using the Copernicus Data Ecosystem (CDSE). The script begins by defining key initialization parameters, such as the lake shapefile import, output directory, date range, and Copernicus authentication credentials. After a STAC search using the exact name of the collection ("sentinel-3-olci-1-efr-ntc"), a summary table is generated using pandas and exported to an Excel file.

Before downloading, the script requests confirmation from the user, then extracts the correct ZIP URL for each product. This is important because CDSE STAC resources sometimes contain incorrect value fields, which cause errors during initial tests. The script avoids overwriting existing files and introduces a short pause between downloads to prevent overloading the server. All valid ZIP files are saved to the designated output folder. Currently, downloading a full year's worth of content is not possible because it represents more than 500 images, a volume too large for the STAC request, which appears to stop after about fifteen downloads, at least with the computer used during this work.

Landsat images

As with the urban heat island project, Landsat images are downloaded from Google Earth Engine. Several time periods have been tested, and every year appears to be available in the selected collections (“LANDSAT/LC08/C02/T1_L2” and “LANDSAT/LC09/C02/T1_L2”).
After initialization and the STAC search, the list of available images is generated in an Excel file just before download confirmation.

2.5. Image Processing for Lake Lugano monitoring

For the CIPAIS project, the next step in image processing is essential for visualizing and analyzing the images on a map and making an interpretation of the situation. Indeed, the directly downloaded images are not yet visible in QGIS, it is therefore essential to modify the existing bands, particularly by adding masks and constants. The main processing steps required for valid processing, derived from the SIMILE protocols, are available in Annex 3. The objective is to reproduce, as closely as possible, the manual processing chain described in the SIMILE documentation while ensuring consistency, reproducibility and efficiency of the future comparisons. For complete verification of the processing, the full SIMILE manual treatment protocols are available from Camilla Capelli and Aron Castelli of the SUPSI aquatic ecology sector.

2.5.1. Sentinel-3 OLCI images

As summarized in the processing flowchart in Annex 4, each ZIP archive is extracted, and the script identifies the root folder by locating the xfdumanifest.xml file. This file contains the product metadata. It describes the product structure, acquisition parameters, and references all the files necessary for processing. It is essential for opening and using the data in remote sensing software such as SNAP.
The lake geometry is then converted into a bounding rectangle containing WKT coordinates for use with SNAP's Subset operator, thus restricting the processing to Lake Lugano. The product is then reprojected in UTM 32N (ESPG: 32632) at a resolution of 300 meters.

The C2RCC (Case-2 Regional CoastColour) processor is used to derive chlorophyll-a concentration from OLCI reflectances. This bio-optical model performs atmospheric correction and isolates the water-leaving reflectance signal by excluding suspended solids, and naturally occurring dissolved substances (ESA, n.d.). The workflow follows SIMILE project conventions, including specific parameter settings such as salinity, temperature, CHLexp, CHLfac, and useNNValues=True, which improves robustness in optically complex waters.
The chlorophyll-a band (conc_chl) is extracted and converted into a numpy array. Invalid or non-positive values are removed. An external Excel table of observed water monthly measured temperatures is loaded to calibrate the model. It concerns application of cloud masks and out-of-scope value masks to filter unreliable pixels. Masked pixels are replaced by NaN values.

Depending on image quality, applying masks may lead to loss of valid pixels. To handle this, the script includes three mask modes:

- STRICT : applies all cloud and Out-Of-Scope masks
- LITE: applies only essential masks, and is the one used for the CIPAIS report
- RAW: no masks (to be used only if strict filtering eliminates all valid pixels)

After masking with “LITE” mode, summary statistics such as minimum, maximum, and median chlorophyll-a values are printed in the terminal for quality control.

To reduce the influence of outliers and obtain more robust estimates of chlorophyll-a concentration, statistical filtering is applied. Values outside the range defined by the 2.5th and 97.5th percentiles are removed to eliminate extreme observations while preserving the central distribution of the data.

The filtered chlorophyll-a concentration raster is exported to GeoTIFF format using SNAP's GeoTIFF writing tool. It is then clipped to the lake polygon using a raster mask (rasterio library) and saved with a filename indicating the masking method.

2.5.2. Landsat images

As for Sentinel-3 imagery, the SIMILE project's protocol is applied and is translated into an automated workflow to process Landsat images for the CIPAIS annual report.
Lake surface temperature is derived from Landsat-8 and Landsat-9 Collection 2 Level-2 thermal products following the SIMILE manual protocol based on the atmospheric correction Barsi method (Barsi et al., 2005). The workflow relies on the use of thermal top-of-atmosphere radiance and intermediate atmospheric parameters provided in the Landsat Level-2 products (Annex 5).
For each Landsat scene, the thermal radiance band (ST_TRAD) and the intermediate atmospheric bands ATRAN, DRAD and URAD are extracted and are clipped to the lake mask. This step reproduces the manual GIS-based clipping applied in the SIMILE protocol. Pixels affected by clouds, cloud shadows, snow or radiometric saturation are excluded using the official Landsat QA_PIXEL and QA_RADSAT quality assurance bands. In addition, a near-infrared threshold is applied to further reduce residual cloud and shoreline contamination over the water surface. Atmospheric parameters are computed on a scene-by-scene basis. For each scene, mean digital number values of the ATRAN, DRAD and URAD bands are calculated over the lake surface and are converted to physical units using the official Landsat scale factors. These mean values are used as scalar atmospheric inputs representing atmospheric transmittance (τ), downwelling radiance (L_d) and upwelling radiance (L_u), consistently with the manual procedure.
Surface-leaving radiance is then obtained by applying the radiative transfer equation of Barsi, using per-pixel thermal radiance (ST_TRAD) and scene-specific atmospheric parameters. A constant water emissivity value (EMIS) of 0.98 is assumed. Land surface temperature is retrieved by inverting Planck's law using Landsat-specific thermal constants and is subsequently converted from Kelvin to degrees Celsius.
An additional calibration step is applied to ensure consistency between satellite-derived lake surface temperature and in situ observations. This step relies on a linear regression between Landsat-derived temperature estimates and concurrent in situ water temperature measurements collected within the SIMILE monitoring framework. The regression parameters are derived from a reference dataset containing paired estimated and observed temperature values, originally stored in an internal Excel file used in the manual SIMILE workflow. This dataset is not anonymized for the moment because the dataset is publicly available. The resulting linear model is recalculated for each processing period and is applied to all Landsat scenes prior to the production of the final temperature maps.
The Landsat thermal band ST_B10 is also extracted and is converted to Kelvin using USGS scale factors for diagnostic purposes only. This step mirrors the manual workflow and allows consistency checks between brightness temperature and atmospherically corrected surface temperature, without being used in the final temperature computation.
Finally, temperature maps are statistically trimmed by removing extreme values outside the 2.5–97.5 percentile range, as performed in the SIMILE processing chain. The resulting lake surface temperature products are saved as GeoTIFF files, fully compatible with standard GIS software such as QGIS to be read as rasters.

2.6. Validation methodology: comparison with reference products

A visual comparison between the GeoTIFFs generated by the automated workflow and the reference GeoTIFFs obtained by the manual processing chain used for the CIPAIS 2024 report allows for verification of the automation results.
This comparison is based on five complementary visual and statistical outputs that together provide a comprehensive understanding of the level of agreement between the two datasets. PNGs are generated for this visual analysis.
The flowchart of the script providing this analysis, for the example of Landsat images, is presented in Annex 6.
Before performing the comparison, the images already reprojected are systematically resampled to match the grid of the monthly reference raster. It is important to note that, for each month, all automated image scenes were compared

to a single monthly reference image. These reference images are based on a single date per month, while the automated workflow processes all available scenes from that same month. Consequently, the natural temporal variability of lake surface temperature and chlorophyll-a concentration can lead to significant differences between the processed scenes and the reference images, even if they differ by only a few days. The examples presented in the Results chapter account for this variability, and the compared processed image is as close as possible to the date of the reference image.

The side-by-side display offers an initial qualitative impression of the overall similarity between the images by presenting them with a consistent color scale. It allows the reader to visually assess whether spatial patterns, temperature gradients, and regional trends appear consistently across the two sources, or whether the processed image introduces noticeable artefacts, smoothing effects, or localized deviations.

The difference map, representing the pixel-wise subtraction of the processed raster from the reference, highlights where and by how much the two datasets diverge. Areas with strong positive or negative deviations reveal spatially structured biases, such as systematic warm or cold shifts, processing artefacts, or inconsistencies related to land cover, terrain, or atmospheric conditions.

The relative difference map, expressing these deviations as percentages, is particularly useful when absolute values vary significantly across the year. It helps identify regions where seemingly small absolute differences correspond to large proportional departures, providing insight into the robustness of the processing method under varying seasonal and environmental conditions.

The pixel-to-pixel scatter plot provides a quantitative evaluation of the relationship between the two rasters. When points cluster tightly around the 1:1 diagonal, the datasets show strong agreement. A consistent offset from this diagonal indicates systematic bias, while increased dispersion suggests weaker correlation or higher noise levels. Outliers may reveal issues such as residual cloud contamination, misclassification, or localized processing anomalies.

Finally, the histogram comparison examines the overall distribution of pixel values in each raster. Overlapping histograms indicate similar global ranges and variability. Shifts in position or differences in spread may reveal systematic biases, changes in dynamic range, or modifications affecting extremes (e.g. smoothing of hot or cold outliers).

Together, these visualizations allow for an integrated assessment of global similarity, spatial coherence, systematic bias, and the presence of local anomalies. A high-quality processed raster should exhibit coherent spatial patterns in the side-by-side view, minimal deviations in the difference maps, a tight, near-diagonal relationship in the scatter plot, and overlapping value distributions in the histograms. This combined approach provides a robust and transparent framework for evaluating the reliability of processed raster data relative to a trusted reference.

3. Results

3.1. Output products

The automated processing workflow produced two main categories of outputs: Excel files summarizing the Landsat and Sentinel-3 scenes downloaded and GeoTIFFs processed for the study period.
The tables present, for each scene, the acquisition date and time, the satellite name or product identifier and the associated cloud cover or coverage of canton for the heat islands project. These summary files ensure full traceability of the input data used in the automated workflow. Screenshots of example Excel results are shown in Annex 1 and 2.
The main products generated by the automated processing are georeferenced lake surface temperature and chlorophyll-a concentration raster exported as GeoTIFF files with the projection EPSG:32632, which is used in Switzerland and in the annual report of the CIPAIS. Each GeoTIFF corresponds to a single image scene processed according to the SIMILE atmospheric correction protocols and clipped to the lake surface. These raster products are the main output of the automation and are directly compatible with standard GIS software such as QGIS. Examples of the resulting maps are shown in Annex 7 and 8. Initial visual verification shows that the generated GeoTIFFs are similar to the reference image.

3.2. Comparison with reference SIMILE products

The more in-depth comparison confirms that the automated implementation reproduces the main results of the manual protocol as we can see in the comparison histogram of the figure 7. However, there are some differences between the processed images and the reference images, as it's shown in figure 8, suggesting that the scripts can still be improved. Complete examples of these comparisons are presented in Annex 9.

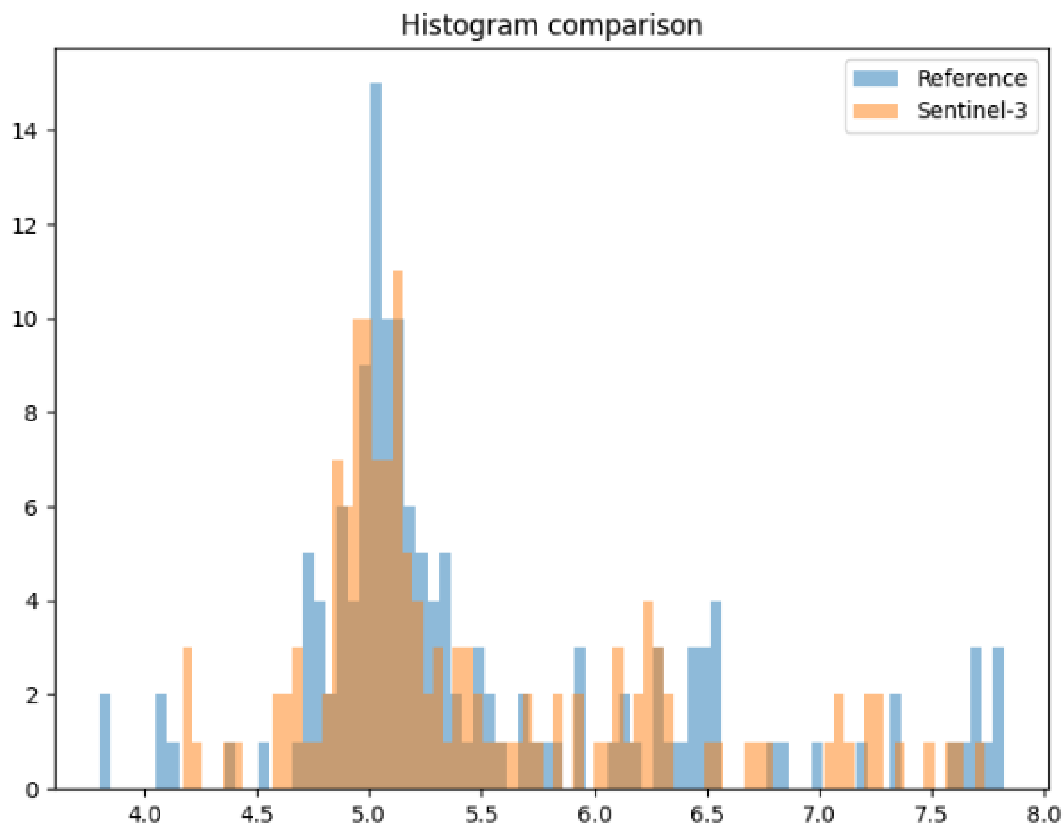


Figure 7 : Comparison of the 16.03.2024 for chlorophyll-a concentration with histogram method

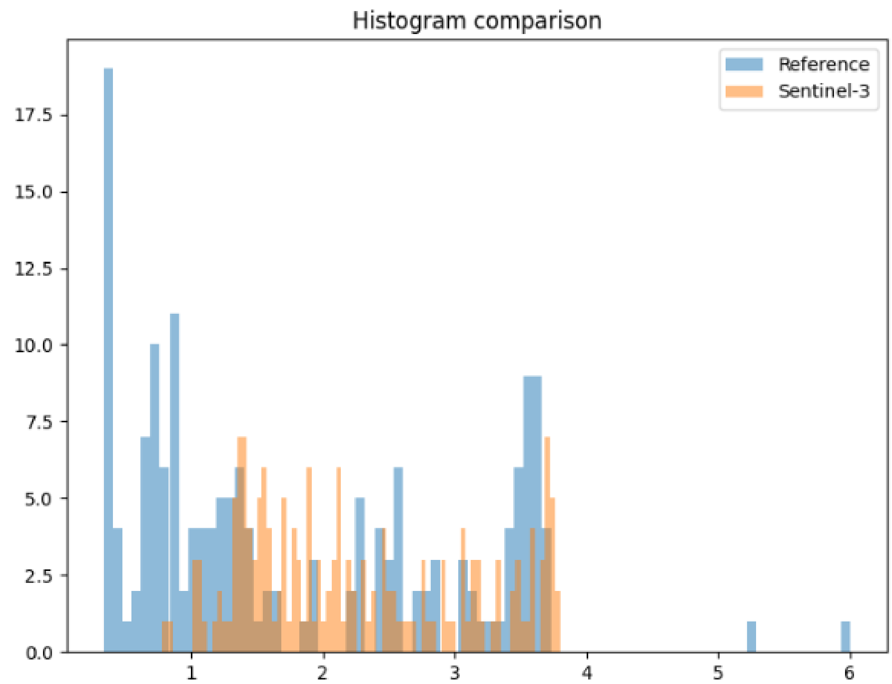
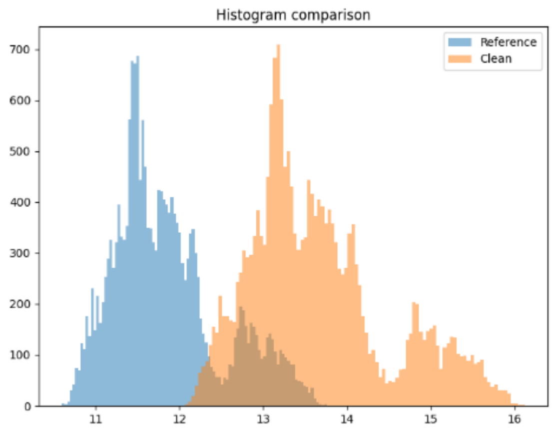


Figure 8 : Comparison of the 23.07.2024 for chlorophyll-a concentration with histogram method



Visual and statistical inspection of the Landsat images shows that the lake surface temperature values are consistent between the automated products and the SIMILE reference maps, proving that the atmospheric correction and masking steps were correctly applied. However, as illustrated in figure 9, some differences take the form of a systematic temperature offset rather than spatial inconsistencies. This difference could not yet be explained due to time constraints, but it could be a little persistent processing error.

Figure 9 : Comparison of the 17.04.2024 for surface temperature with histogram method

This analysis could, however, have been based on a more scientific foundation because, despite the fact that it was invented at the end of the work, a validated scientific study could have supported its justification.

3.1. Scripts presentation

The various flowcharts created using the open-source JavaScript library "Mermaid" provide an initial visualization of the script automation steps. This step is important because raw scripts are not very informative for a presentation of the work.

The complete scripts have been stored on the GitHub platform to ensure their accessibility and the login URL is: Yara-Leone_SUPSI_project . This reduces the size of this report while making the entire automation process accessible not only to report reviewers but also to colleagues in the SUPSI geomatics sector who will take over the project. The project on GitHub includes the automation scripts and other documents required for processing.

4. Learnings and limits

This work provided valuable insights into both the technical and methodological challenges associated with satellite data processing and automation in an operational context.

4.1. Platform-related limitations

One of the main limitations encountered during this project concerns the constraints imposed by satellite data platforms themselves. While several platforms offer powerful tools for accessing and processing Earth observation data, each comes with inherent spatial, technical or administrative limitations. For instance, the Swiss Data Cube would have been a highly relevant resource for this study. However, this platform only provides data within Swiss national boundaries, which represents a significant limitation for transboundary water bodies such as Lake Lugano. This spatial constraint prevents its direct application to lakes or other objects shared between multiple countries and highlights the importance of platform coverage when selecting data sources.

Other platforms, such as USGS, EUMETSAT or Microsoft Planetary Computer portals, were explored but ultimately not used due to time constraints or the complexity of identifying and extracting the appropriate products and bands. These platforms remain promising options for future work, as they offer open access and do not necessarily require user registration.

4.2. Automation and workflow complexity

Methodological reflection

Although automation significantly improves efficiency and reproducibility, the development of even relatively small automated workflows proved to be time-consuming. Designing robust scripts requires careful planning, clear structuring of the processing chain, and a well-defined strategy before implementation. Researching existing work and taking it into account when writing one's own scripts is a valuable tool. This project highlighted the importance of logical code design, explicit parameter definitions and systematic testing at each processing step.
In addition, documenting the workflow proved to be as important as coding itself. Producing clear workflows, schematic illustrations and commented scripts greatly facilitates understanding of the process logic and ensures that the methodology can be communicated effectively to collaborators.

Reproducibility and software dependencies

Reproducibility emerged as a key challenge. For another collaborator to reproduce or continue this work, it is necessary to provide a complete set of requirements, including Python libraries, software versions used, and input data such as shapefiles and Excel files with input parameters. File paths must also be adapted to local environments, which can represent an additional barrier if not clearly documented.
As introduced, software compatibility issues can further complicate reproducibility. A notable example encountered during this project is the Python library snappy, used to interface with ESA's SNAP software. Unlike standard Python packages, snappy cannot be installed directly via common package managers. Its installation requires a previous version of Python (version 3.9.12) and SNAP (version 9) software, and correct configuration of environment variables and directory paths in SNAP installation. This step is not straightforward and remains insufficiently documented, which can represent a significant obstacle for new users.

4.1. Data quality and processing assumptions

The quality of satellite-derived products is strongly dependent on external factors such as cloud cover, atmospheric conditions and acquisition geometry. Even with a correct processing chain, some images may remain unusable. Furthermore, the use of fixed constants and default values, such as right masks and scale factors, can introduce uncertainty if not carefully selected. In cases where in-situ observations are unavailable, the reliance on default calibration values may propagate uncertainty into derived products, such as chlorophyll-a concentration estimation.

Platform-imposed limitations on data access also represent a constraint. For example, restrictions on the number of images that can be queried or downloaded in a single request may require compromises between completeness and processing time. Choosing inappropriate limits or pause time between image downloads may result either in missing relevant images or in excessively long processing times.

4.2. Perspectives for future work

Several extensions could improve and complement the current workflow. Integrating additional satellites, such as Sentinel-2 MSI, MODIS, or VIIRS, would allow for satellite comparisons and offer higher temporal resolution. Although the Sentinel-3 OLCI Level 2 WFR products include chlorophyll-related bands, these products have not been fully utilized due to difficulties in visualizing scientifically relevant concentration values in GIS software. Removing download limits would enable near-instantaneous image processing. With further methodological developments, these datasets could significantly enrich future analyses. For example, the comparison between satellite imagery and in-situ measurements taken on the lake could be resumed, and the discrepancies between the two measurement methods could be analyzed.

5. Conclusion

This work demonstrates that the processing of satellite images for lake monitoring and urban heat island study can be efficiently streamlined through automated workflows, while remaining compliant with established scientific protocols. This work can also be carried out by someone with limited programming knowledge if the available support tools are well used. By translating SIMILE's manual procedures into reproducible scripts, the developed workflows significantly reduce the time required for image selection, downloading, and preprocessing, while maintaining compatibility with the methodologies used manually.

The comparison with manually produced reference products confirms that the automated results are scientifically reliable. The spatial patterns and statistical distributions are preserved, and the observed discrepancies are mainly explained by the temporal variability between the acquisition dates and by the choice of application of processing parameters. These differences do not undermine the validity of the approach but rather highlight the intrinsic sensitivity of satellite-derived products to processing choices. A few additional modifications could, however, strengthen the certainty of the automated results.

Beyond operational gains, this work emphasizes the importance of rigorous documentation, clear workflow design and transparent parameter choices to ensure reproducibility and knowledge transfer. The scripts, accompanied by inventories and flowcharts, provide a concrete basis for future reuse by SUPSI collaborators.

In the future, once the image processing is fully validated and the accuracy of the reference images confirmed, the aquatic ecology sector will be able to use the processed images in the CIPAIS report. The geomatics sector plans to implement daily automation to access the data as soon as the current scenes are available on download platforms. This will allow the insubrilakes.eu website to be updated and these water quality parameters to be made accessible to a wider audience. This automation could also be applied to other lakes in the region. It should be noted, however, that the alplakes.eawag.ch website already offers a series of parameters estimated from satellite imagery for various lakes in Switzerland.

Finally, while automation offers undeniable advantages, it also creates new dependencies regarding platform availability, software compatibility, and data quality. Addressing these constraints will be essential for future developments, particularly if workflows are extended to daily monitoring, historical analysis using past images, and additional satellite missions. Overall, this project establishes a robust and transferable framework for automated satellite-based environmental monitoring in applied institutional contexts.

Bibliography

Mentions in the text

Ansper, A., & Alikas, K., (2018). Retrieval of Chlorophyll a from Sentinel-2 MSI Data for the European Union Water Framework Directive Reporting Purposes. *Remote Sensing, 11(1),* 64. https://doi.org/10.3390/rs11010064

Bresciani, M., & Stroppiana, D., & Odermatt, D., & Morabito, G., & Giardino, C. (2011). Assessing remotely sensed chlorophyll-a for the implementation of the Water Framework Directive in European perialpine lakes. *Science of the Total Environment, 409,* 3083-3091. https://doi.org/10.1016/j.scitotenv.2011.05.001

DACD SUPSI. (2023). *Ricerche sull'evoluzione del Lago di Lugano. Aspetti limnologici. Programma triennale 2022-2024. Campagna 2024 e rapporto triennale.* Commissione Internazionale per la Protezione delle Acque Italo-Svizzere. 100pp. https://www.cipais.org/web/lago-di-lugano/rapporti

European Space Agency (ESA). (2023). *Sentinel-3 SLSTR instrument overview.* Accessed 7 December 2025, on https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-3-slstr

European Space Agency (ESA). (n.d.-a). *Case-2 Regional CoastColour (C2RCC) processor.* Earth Online – EO Gateway. Accessed 7 December 2025, on https://earth.esa.int/

European Space Agency (ESA). (n.d.-b). *Landsat-8 and 9 OLI/TIRS worldwide data products.* Earth Online – EO Gateway. Accessed 7 December 2025, on https://earth.esa.int/eogateway/catalog/landsat-8-9-oli-tirs-worldwide-data-products

ESA Sentinel-3 Team. (2017). *Sentinel-3 user handbook.* Accessed 9 December 2025, on https://filetransfer.itc.nl/pub/dragon4/Optical-Thermal/D2OTP1-Hyperspectral-DOdermatt/references/Sentinel-3_User_Handbook-iss1_v1_20170113.pdf

Gholizadeh, M. H., & Melesse, A. M. & Reddi L. (2016). A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors, 16(8),* 1298. https://doi.org/10.3390/s16081298

Gitelson, A. A., & Gurlin, D., Moses, W. J., & Barrow, T. (2008). A bio-optical algorithm for the remote estimation of the chlorophyll-a concentration in case 2 waters. *Environmental Research Letters 4 045003.* DOI 10.1088/1748-9326/4/4/045003

Insubri Lakes Project. (n.d.). *Insubrilakes platform.* Accessed 5 December 2025, on https://insubrilakes.eu/#/admin/home

Mobley, C. D. (1994). Light and water: Radiative transfer in natural waters*. Academic Press.*

NASA. (2019). Landsat 8–9 Thermal Infrared Sensor (TIRS) user guide*.* Accessed 7 December 2025, on https://landsat.gsfc.nasa.gov

Parvez Mosharof md. (2025). The Importance of Remote Sensing in Environmental Science. https://www.researchgate.net/publication/393977426_The_Importance_of_Remote_Sensing_in_Environmental_Science

Schriber, M. (2013). The Full Palette of Photosynthesis. *NASA.* Accessed 10 December 2025, on https://www.giss.nasa.gov//research/features/archive/201311_kiang/

Siddique-E-Akbor, A. H. M., & Hossain, F., & Lee, H., Shum, C. K. (2011). Inter-comparison study of water level estimates derived from hydrodynamic–hydrologic model and satellite altimetry for a complex deltaic environment. *Remote Sensing of Environment,* 115, 1522-1531 . https://doi.org/10.1016/j.rse.2011.02.011

Sudmanns, M., & Tiede, D., & Lang, S., & Bergstedt, H., & Trost, G.,& Augustin, H. (2020). Big Earth data: disruptive changes in Earth observation data management and analysis? *International Journal of Digital Earth*, *13*(7), 832–850. https://doi.org/10.1080/17538947.2019.1585976

SUPSI. (2025). *Istituto Scienze della Terra (IST).* Accessed 9 December 2025, on https://www.supsi.ch/ist

Images

Lehmann, M. et al. (2023). GLORIA - A globally representative hyperspectral *in situ* dataset for optical sensing of water quality. Accessed 15 December 2025, on https://www.nature.com/articles/s41597-023-01973-y

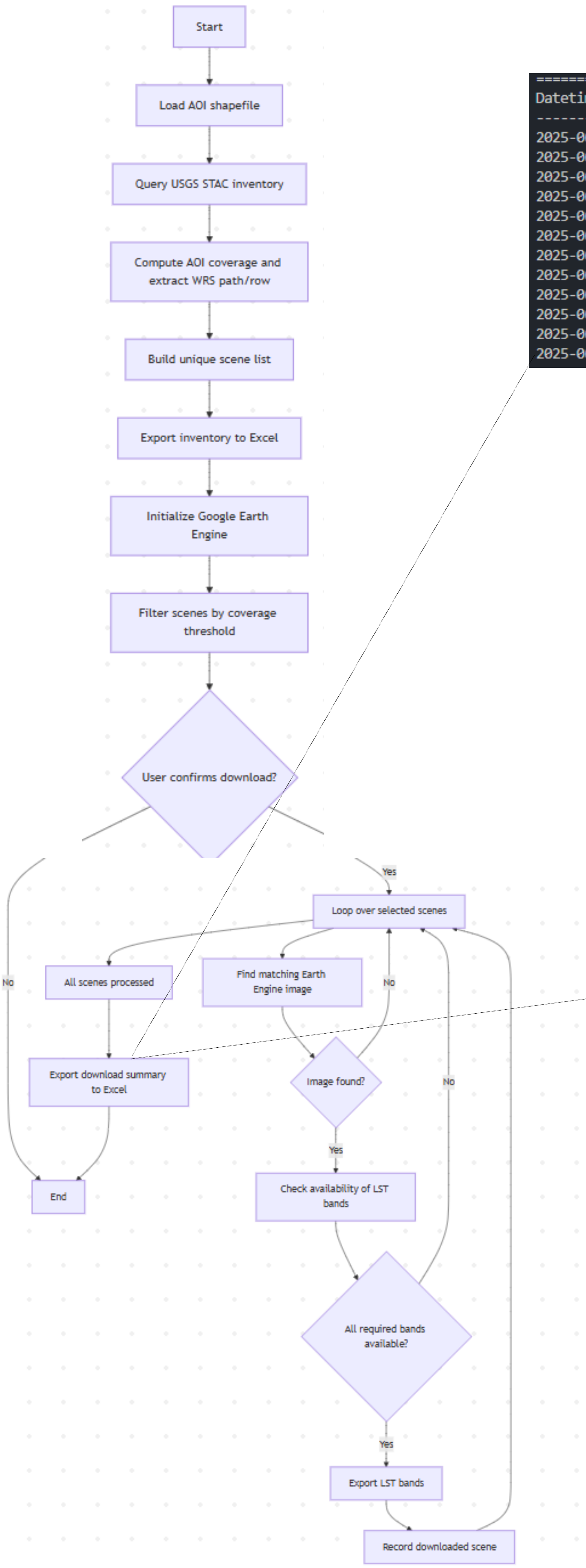
(n.d). Thermal Infrared Remote Sensing. Accessed 15 December 2025, on https://www.sciencedirect.com/topics/earth-and-planetary-sciences/thermal-infrared-remote-sensing

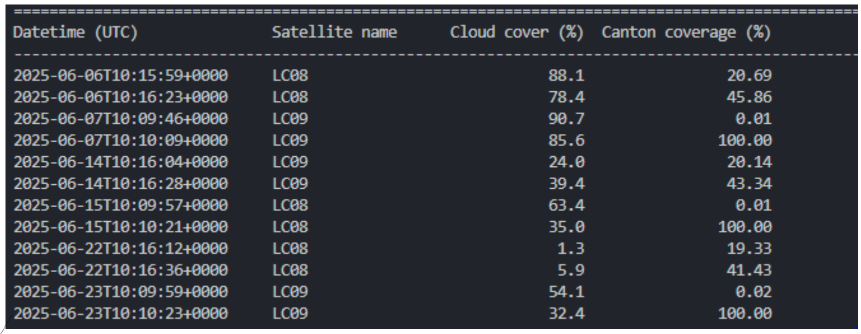
(n.d.). Python Programming Language: Everything You Need to Know. Accessed 15 December 2025, on https://loudbench.com/python/

Schriber, M. (2013). The Full Palette of Photosynthesis. *NASA.* Accessed 10 December 2025, on https://www.giss.nasa.gov//research/features/archive/201311_kiang/

(n.d.). Google Earth Engine Frequently Asked Questions. Accessed 15 December 2025, on https://sanborn.com/blog/google-earth-engine-frequently-asked-questions/

Annex 1: Landsat images downloading flowchart (heat islands project)

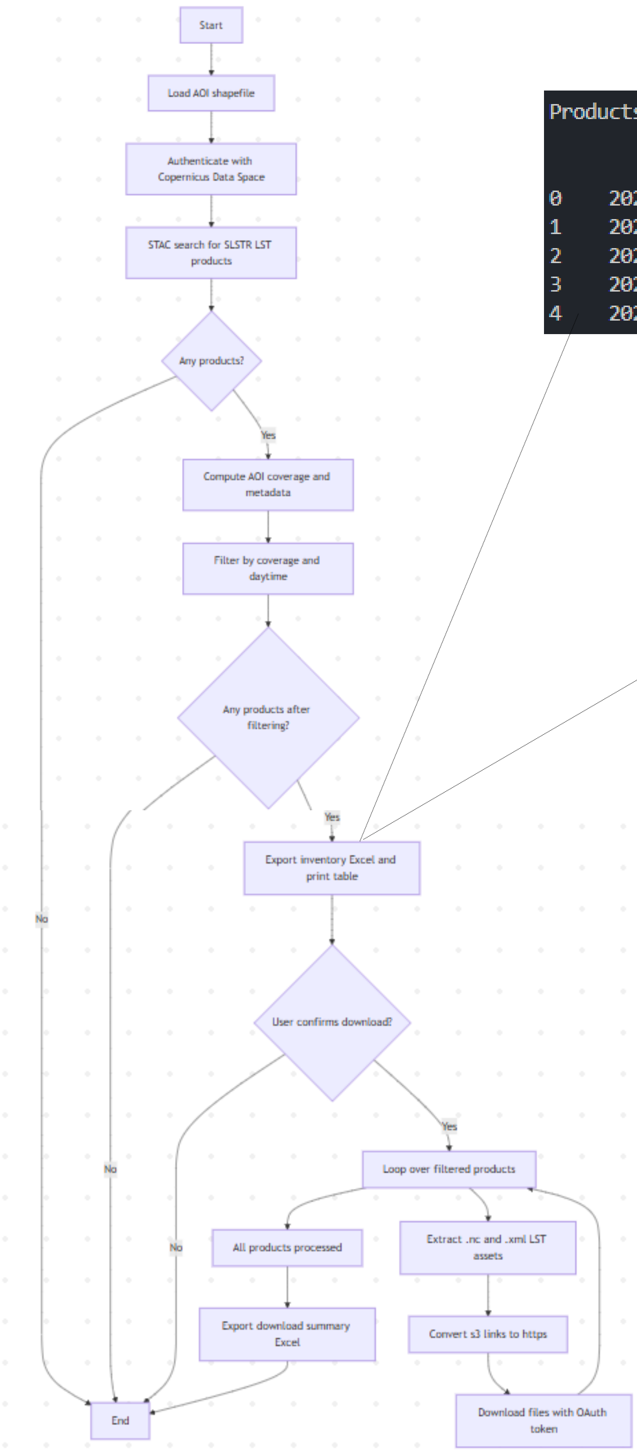




Datetime (UTC)	Satellite name	Cloud cover (%)	Canton coverage (%)
2025-06-06T10:15:59+0000	LC08	88.1	20.69
2025-06-06T10:16:23+0000	LC08	78.4	45.86
2025-06-07T10:09:46+0000	LC09	90.7	0.01
2025-06-07T10:10:09+0000	LC09	85.6	100.00
2025-06-14T10:16:04+0000	LC09	24.0	20.14
2025-06-14T10:16:28+0000	LC09	39.4	43.34
2025-06-15T10:09:57+0000	LC08	63.4	0.01
2025-06-15T10:10:21+0000	LC08	35.0	100.00
2025-06-22T10:16:12+0000	LC08	1.3	19.33
2025-06-22T10:16:36+0000	LC08	5.9	41.43
2025-06-23T10:09:59+0000	LC09	54.1	0.02
2025-06-23T10:10:23+0000	LC09	32.4	100.00

Datetime (UTC)	Satellite name	Cloud cover (%)	Canton coverage (%)
2025-06-06T10:15:59+0000	LC08	88,07	20,69020979
2025-06-06T10:16:23+0000	LC08	78,42	45,86033077
2025-06-07T10:09:46+0000	LC09	90,7	0,005478003
2025-06-07T10:10:09+0000	LC09	85,57	100
2025-06-14T10:16:04+0000	LC09	24,02	20,14330122
2025-06-14T10:16:28+0000	LC09	39,41	43,34088301
2025-06-15T10:09:57+0000	LC08	63,35	0,007824273
2025-06-15T10:10:21+0000	LC08	35,02	100
2025-06-22T10:16:12+0000	LC08	1,29	19,33289839
2025-06-22T10:16:36+0000	LC08	5,87	41,42601518
2025-06-23T10:09:59+0000	LC09	54,08	0,01514747
2025-06-23T10:10:23+0000	LC09	32,44	100
2025-06-30T10:16:13+0000	LC09	3,7	20,04101011
2025-06-30T10:16:37+0000	LC09	29,79	43,14504662
2025-07-01T10:10:05+0000	LC08	43,32	0,007788468
2025-07-01T10:10:29+0000	LC08	16,02	100
2025-07-08T10:16:16+0000	LC08	89,18	20,30752063
2025-07-08T10:16:40+0000	LC08	59,96	44,5569966
2025-07-09T10:10:02+0000	LC09	89,07	0,005068561
2025-07-09T10:10:26+0000	LC09	21,27	100
2025-07-16T10:16:18+0000	LC09	89,66	20,55077295
2025-07-16T10:16:42+0000	LC09	50,02	45,01303072
2025-07-17T10:10:10+0000	LC08	42,52	0,004064631
2025-07-17T10:10:34+0000	LC08	13,9	100
2025-07-24T10:16:23+0000	LC08	88,22	20,01520027
2025-07-24T10:16:47+0000	LC08	64,34	43,7092578
2025-07-25T10:10:11+0000	LC09	94,36	0,00836445
2025-07-25T10:10:35+0000	LC09	85,3	100
2025-08-01T10:16:23+0000	LC09	63,99	20,88397899
2025-08-01T10:16:47+0000	LC09	53,49	45,93667709
2025-08-02T10:10:15+0000	LC08	99,21	0,003034772
2025-08-02T10:10:38+0000	LC08	66,49	100
2025-08-09T10:16:30+0000	LC08	0,58	19,88448658
2025-08-09T10:16:54+0000	LC08	2,1	43,10617064

Annex 2: Sentinel-3 images downloading flowchart (Heat islands project)




```
Products found:

                    Date & Time (UTC)   Satellite ProductType
0   2024-12-31 09:29:43.684703+00:00  sentinel-3a    OL_1_EFR
1   2024-12-30 09:55:54.665392+00:00  sentinel-3a    OL_1_EFR
2   2024-12-30 09:17:10.169227+00:00  sentinel-3b    OL_1_EFR
3   2024-12-29 09:43:21.315687+00:00  sentinel-3b    OL_1_EFR
4   2024-12-28 10:09:32.411383+00:00  sentinel-3b    OL_1_EFR
```

Datetime (UTC)	Satellite name	Canton coverage (%)
2025-06-01T09:49:27Z	SENTINEL-3B	100
2025-06-01T10:28:15Z	SENTINEL-3A	48,71590214
2025-06-02T09:23:16Z	SENTINEL-3B	100
2025-06-02T10:02:04Z	SENTINEL-3A	100
2025-06-03T09:35:53Z	SENTINEL-3A	100
2025-06-04T09:09:42Z	SENTINEL-3A	100
2025-06-04T10:11:53Z	SENTINEL-3B	100
2025-06-05T09:45:42Z	SENTINEL-3B	100
2025-06-05T10:24:30Z	SENTINEL-3A	100
2025-06-06T09:19:31Z	SENTINEL-3B	100
2025-06-06T09:58:19Z	SENTINEL-3A	100
2025-06-07T09:32:08Z	SENTINEL-3A	100
2025-06-08T10:08:09Z	SENTINEL-3B	100
2025-06-09T09:41:59Z	SENTINEL-3B	100
2025-06-09T10:20:44Z	SENTINEL-3A	100
2025-06-10T09:15:48Z	SENTINEL-3B	100
2025-06-10T09:54:33Z	SENTINEL-3A	100
2025-06-11T09:28:22Z	SENTINEL-3A	100
2025-06-12T10:04:26Z	SENTINEL-3B	100
2025-06-13T09:38:16Z	SENTINEL-3B	100
2025-06-13T10:17:00Z	SENTINEL-3A	100
2025-06-14T09:12:05Z	SENTINEL-3B	100
2025-06-14T09:50:49Z	SENTINEL-3A	100
2025-06-15T09:24:39Z	SENTINEL-3A	100

Annex 3 : Steps of the SIMILE protocol

The following tables summarize the correspondences of the processing steps between the manual processing protocols and their automation:

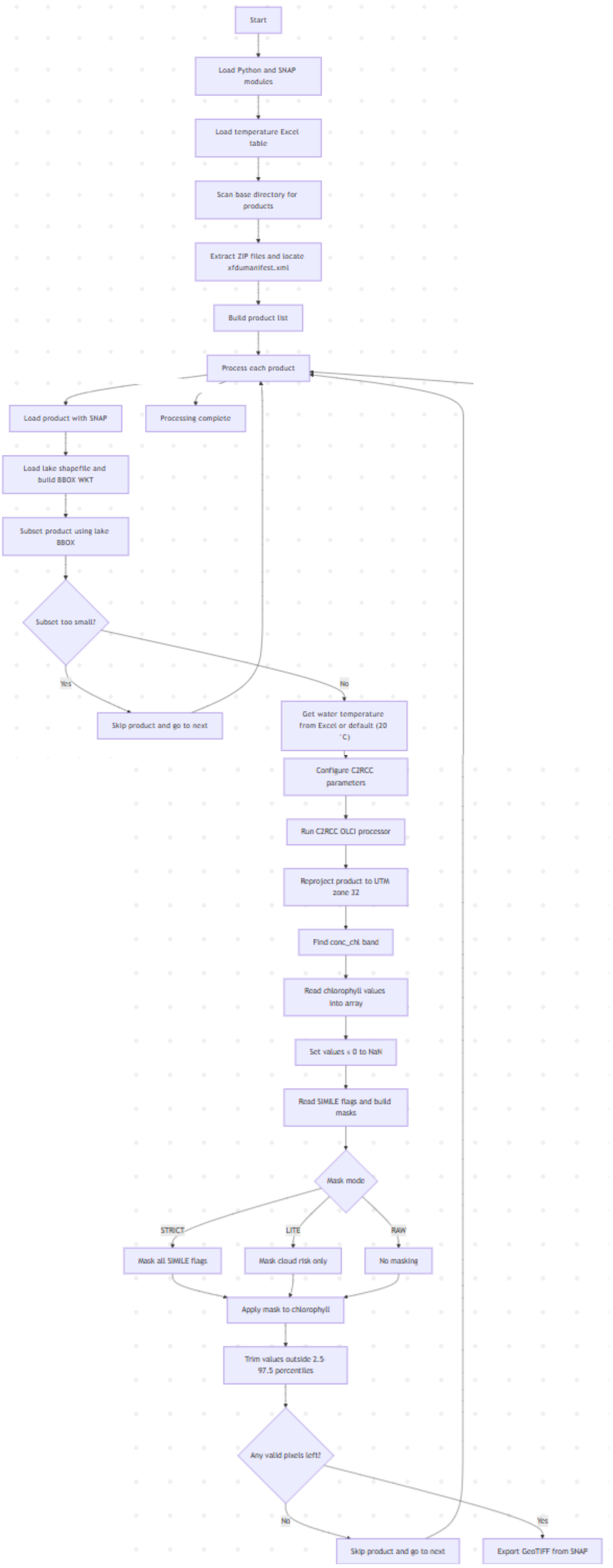
Sentinel-3 OLCI images

SIMILE Processing Step	Manual Protocol	Python Automated Script
Data acquisition	Download Sentinel-3 OLCI L1B ZIP from Copernicus.	Automatically detects ZIP files and extracts products.
Load product	Open *xfdumanifest.xml* manually in SNAP.	ProductIO.readProduct() loads the product automatically.
Subset lake area	Manual "Subset" tool after zooming on the lake.	Shapefile → WKT bounding box → automatic SNAP Subset.
Reprojection	Reproject to UTM 32N, set pixel size = 300 m manually.	Reproject operator with EPSG:32632 and 300 m resolution.
C2RCC processing	Run C2RCC OLCI with SIMILE parameters (salinity, CHL exp/factor, temperature).	SIMILE parameters implemented; temperature retrieved automatically from Excel.
Extract conc_chl	Convert the *conc_chl* band to raster.	Reads conc_chl directly into arrays (read_band).
Masking	Apply masks: Cloud_risk, Rtosa_OOS, Rtosa_OOR, Rhow_OOR + lake boundary mask.	Automatic detection of SIMILE flags + masking + shapefile clip.
Outlier removal	Not included in protocol (was said orally)	Removes extreme values using 2.5–97.5 percentile trimming.
Export outputs	Export GeoTIFF with the Chl-a band.	Automatic GeoTIFF export + clipping to lake.

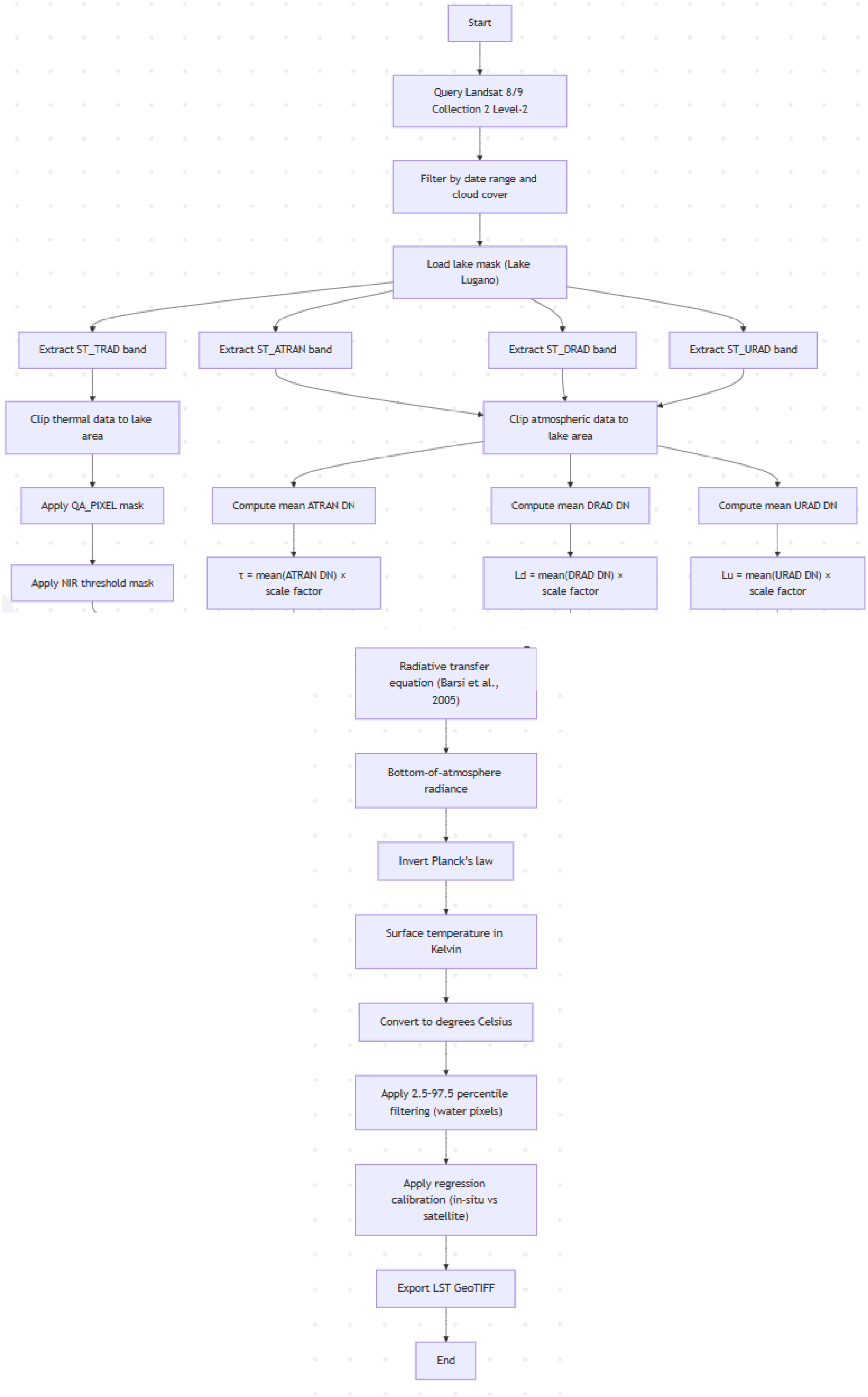
Landsat images

SIMILE Processing Step	Manual Protocol	Python Automated Script
Satellite data	Use of Landsat 8 thermal imagery	Image collections LANDSAT/LC08/C02/T1_L2 and LANDSAT/LC09/C02/T1_L2, processing level
Thermal band	Use of TIRS thermal infrared data	Use of ST_TRAD (TOA thermal radiance)
Atmospheric parameters	Use of τ (atmospheric transmissivity), Ld (downwelling radiance), Lu (upwelling radiance)	Use of ST_ATRAN, ST_DRAD, ST_URAD bands (Collection 2 Level-2 scaling applied)
Water emissivity	Constant emissivity value $\varepsilon = 0.98$	Constant EMIS = 0.98
Atmospheric correction (Barsi et al., 2005)	($L_{BOA} = \frac{L_{TOA} - \tau (1-\varepsilon)L_d - L_u}{\tau \varepsilon}$)	Direct implementation of the radiative transfer equation
Planck equation inversion	Conversion from radiance to surface temperature using Landsat-specific constants	Constants K1 = 1321.08, K2 = 774.89
Conversion to °C	Conversion from Kelvin to Celsius	Subtraction of 273.15
Cloud masking (Landsat QA)	Removal of cloud, cloud shadow and cirrus-affected pixels	Bitwise masking using QA_PIXEL and QA_RADSAT
Additional NIR filtering	Optional NIR threshold to remove residual cloud contamination	Automatic NIR threshold using the 97.5th percentile of SR_B5
Lake boundary masking	Exclusion of areas outside lake surfaces using lake shapefile	Clipping with SIMILE lake geometry and negative buffer to remove shoreline effects
Spatial resampling	Homogenization of spatial resolution prior to processing	Native 30 m resolution preserved
Outlier removal	Removal of anomalous temperature values based on histogram inspection	Statistical trimming using 2.5–97.5 percentile range
Satellite–in situ calibration	Optional empirical correction using in-situ measurements	Annual linear regression based on Excel (estimated vs observed temperatures)
Final map projection	Final products in metric projection for cartographic use	Reprojection to EPSG:32632 (UTM Zone 32N)
Output format	Raster map of lake surface temperature (°C)	GeoTIFF, float32

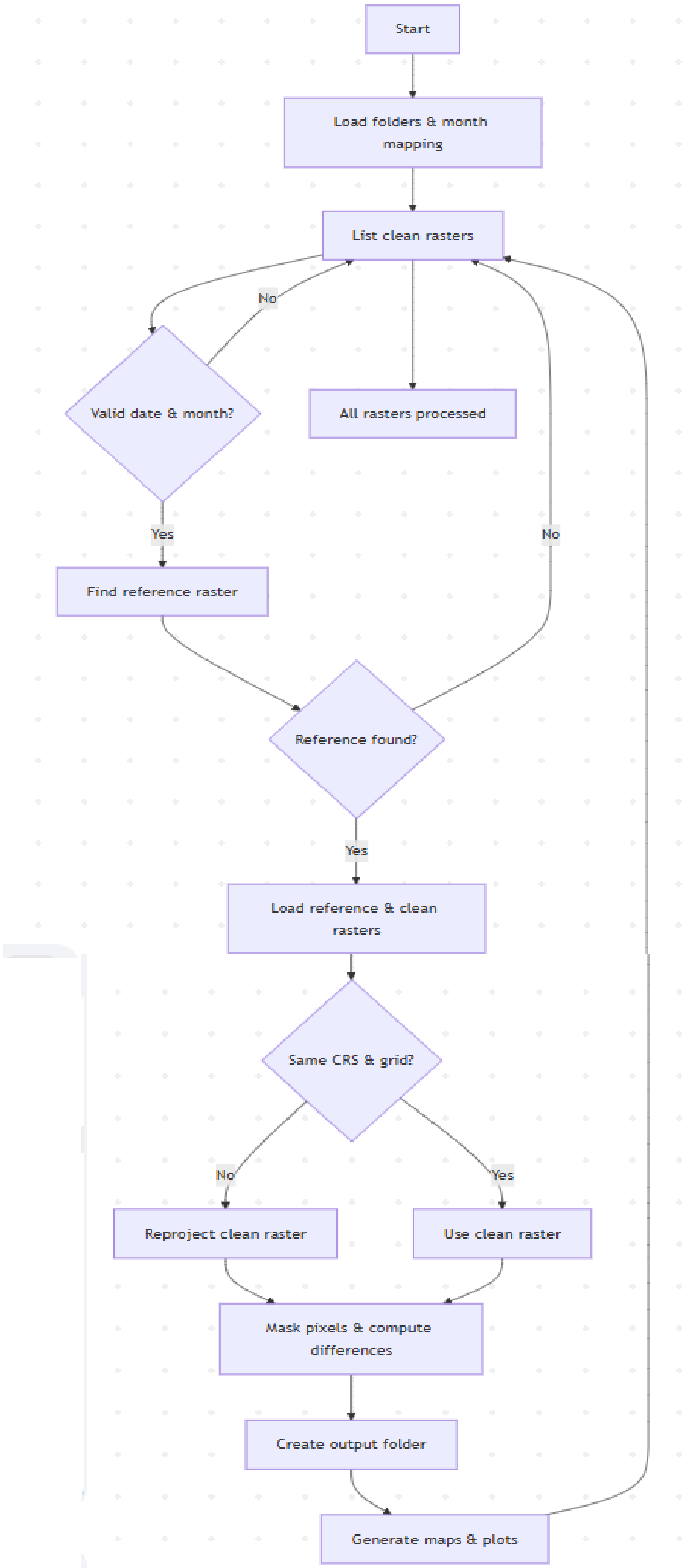
Annex 4: Sentinel-3 images processing flowchart (CIPAIS project)



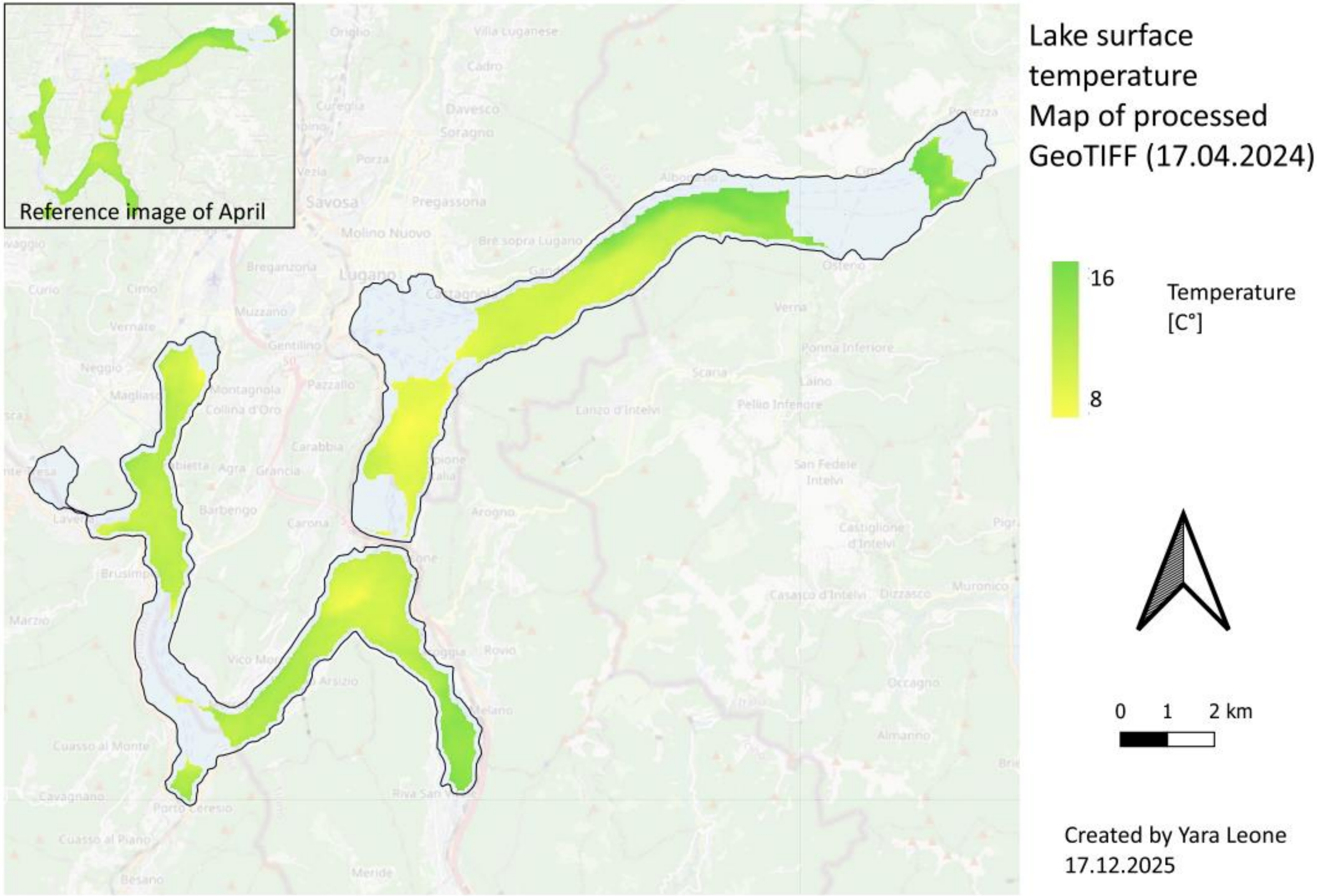
Annex 5: Landsat images processing flowchart (CIPAIS project)



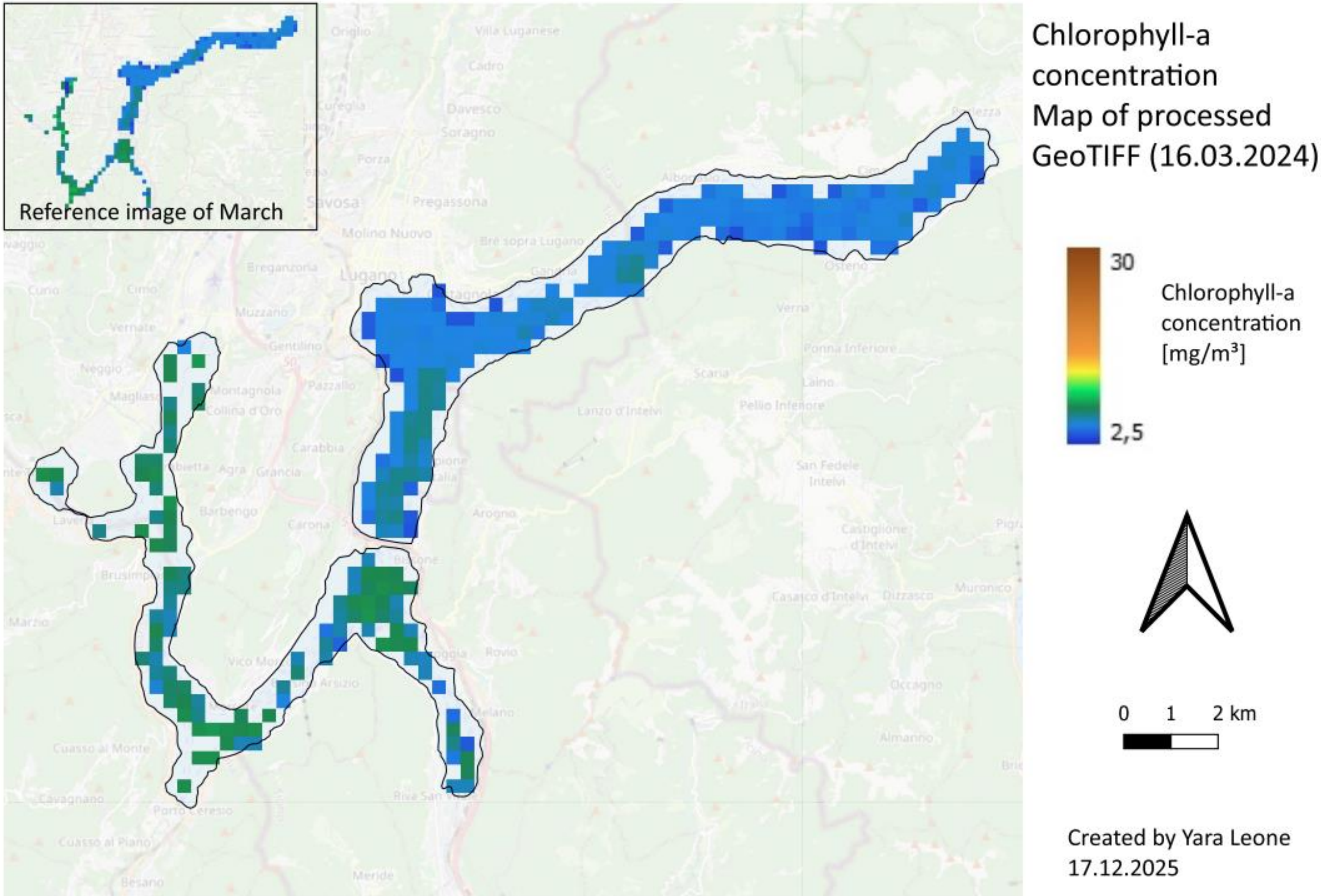
Annex 6 : Flowchart comparing processed image and reference image (Landsat example)



Annex 7: Final processed GeoTIFF (Landsat-9 17.04.2024)

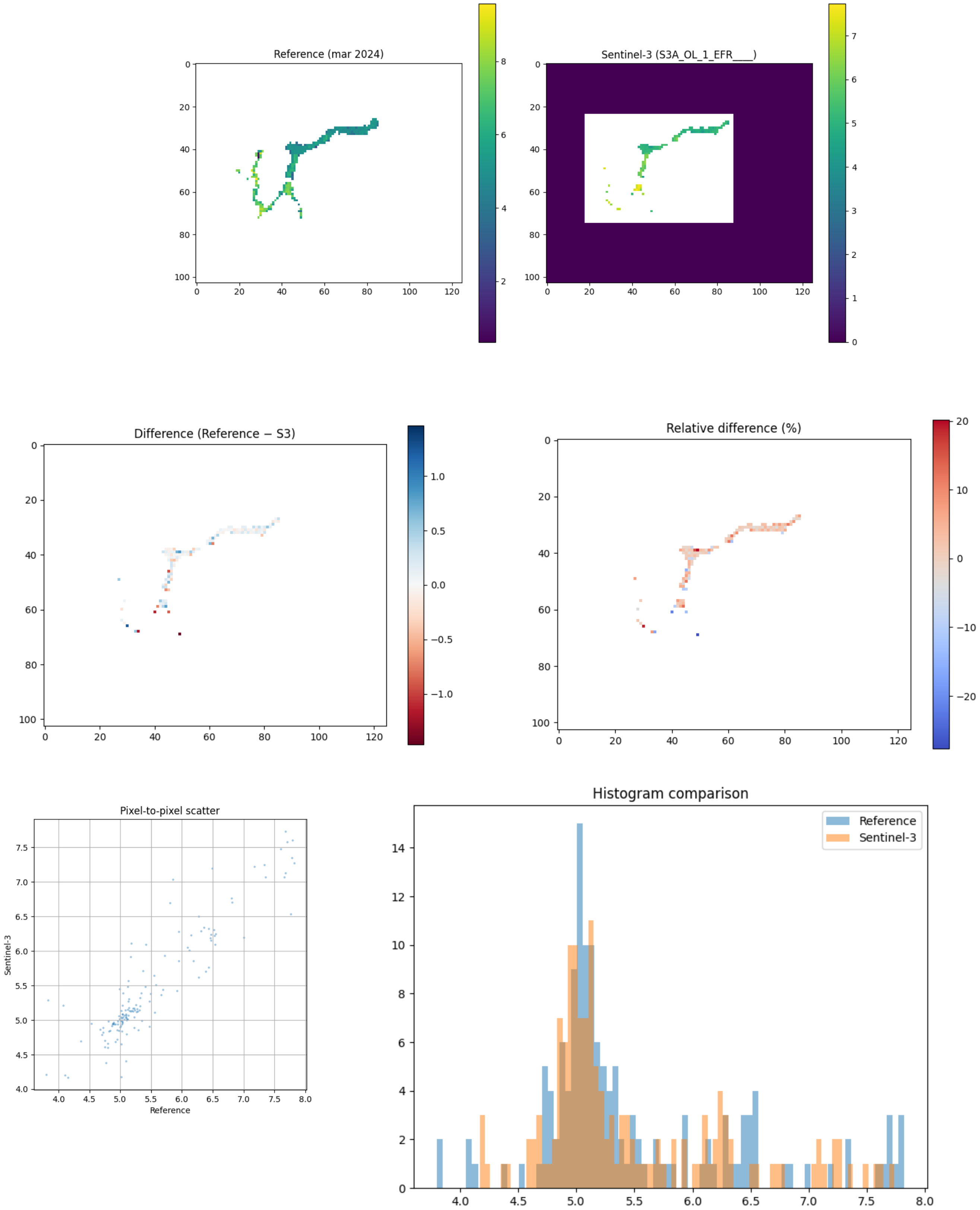


Annex 8: Final processed GeoTIFF (Sentinel-3 16.03.2024)

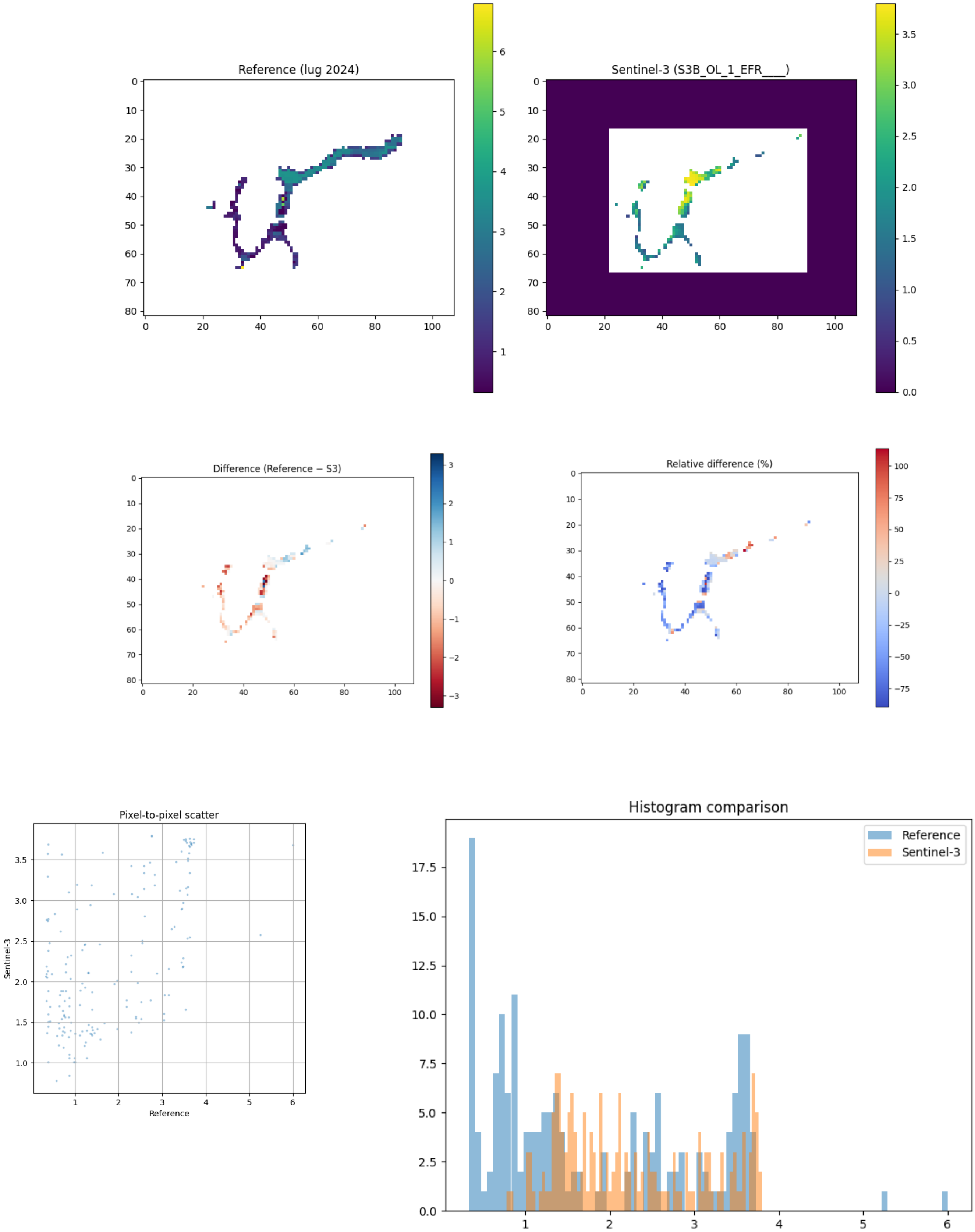


Annex 9: Examples of comparisons between processed images and references

Comparison between the processed Sentinel-3 image from 16 March 2024 and the Mars reference image from the CIPAIS 2024 report, chlorophyll-a concentration



Comparison between the processed Sentinel-3 image from 23 July 2024 and the July reference image from the CIPAIS 2024 report, chlorophyll-a concentration



Comparison between the processed Landsat-9 image from 17 April 2024 and the April reference image from the CIPAIS 2024 report, chlorophyll-a concentration