POLITECNICO DI MILANO
DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING
DOCTORAL PROGRAMME IN ENVIRONMENTAL AND INFRASTRUCTURE ENGINEERING

Lake Water Quality Maps Monitoring and Dissemination – Insubric Lakes Case Study

Doctoral Dissertation of:
Juan Fernando Toro Herrera

Supervisor:
Prof. Daniela Carrion

Tutor:
Prof. Federica Migliaccio

The Chair of the Doctoral Program:
Prof. Riccardo Barzaghi

35 Cycle – 2023

Acknowledgements

Abstract

The objective of this work is to assist public administrations in conserving water resources by monitoring various parameters of water quality to evaluate its overall status. The research focuses on addressing a specific area within geo-visualization, namely geographic visualization, and aims to develop efficient tools and web-based applications for the visualization, sharing, and automated processing of water quality parameters in inland water bodies. The web-based applications, or Collaborative Data Sharing Platform and WebGIS, for sharing and monitoring the raster data have been implemented for a study case and are available on the Web.

The Collaborative Data Sharing Platform enables the multitemporal visualisation of the water quality parameters maps for the study case. The platform benefits from open standards, free and open-source software, and open data, most notably to ensure interoperability and reduce the barriers to access it. On the other hand, the necessary tools and workflow for the semi-automated production of the water quality parameters maps are available inside a virtual environment for further development and applications to other study cases. A processing chain was implemented and integrated to generate maps of water quality parameters, in the framework of SIMILE Interreg Italy-Switzerland project (Brovelli et al., 2019), which defined the algorithms to be used. This involved determining the most appropriate workflow and implementing a series of procedures to address issues encountered during map production. These identified issues, such as misaligned maps and the presence of anomalous values, had the potential to cause misinterpretation of parameter trends over time. Therefore, the processing chain integrated a co-registration algorithm and statistical inference methods for addressing the issues. The primary goal of this research is to bridge the divide between scientific knowledge and society by effectively sharing research findings and enabling the use of web-based tools and

containerised solutions for replication the work. By doing so, it aims to facilitate informed decision-making concerning different water quality phenomena.

This dissertation's literature review focused on monitoring water quality assessment through remote sensing and the different WebGIS infrastructures to design and implement a data-sharing application using open standards and free and open-source software. SIMILE project served as the study case for the implementation of the data-sharing platform and production of water quality parameters maps. Next, the document provides a brief background on the data production workflow and outputs of the data processing of the water quality indicators. The dissertation concludes with an overview and the future directions of the work.

Table of Contents

List of Figures

List of Tables

Nomenclature and Abbreviations

Chapter - 1.   INTRODUCTION

Water resources, and thus water quality, have been compromised due to the increasing pressures on the ecosystems such as anthropic activities and climate change (Foley et al., 2005). Therefore, sustainable management of the quality of water has become a major requirement. For example, eutrophication, mostly caused by growing agricultural and industrial activities, poses serious threats to the ecosystems' health, aquaculture and fisheries, recreation, and tourism. In this context, geographical and temporary data on water quality metrics are critical for enhancing our understanding of water bodies, as well as assessing the impact of human pressures (Vavassori et al., 2022).

Currently, there are available three different means for water quality assessment: (a) in-situ sampling; (b) mathematical modelling; and (c) remote sensing. In inland and coastal areas, traditionally, field-based measurements of water quality indicators have been an important source for monitoring water characteristics. However, these measurements are constrained both spatially and temporally, significantly limiting their use for recording spatiotemporal dynamics of constituents such as Chlorophyll-a (CHL-a), Total Suspended Matter (TSM) and Colour Dissolved Organic Matter (CDOM). Furthermore, field sampling is costly and time-consuming. For this reason, remote sensing techniques are being widely employed for complementing field observations to address its challenges. Remote sensing provides a synoptic view of the water quality status, providing a lower-cost solution to analyse water bodies with respect to in-situ sampling and mathematical modelling. While remote sensing of inland water bodies could classify as both satellite and airborne remote sensing, there are benefits and challenges concerning their use. Airborne products can provide a higher image resolution, but the acquisition frequency is limited to surveying campaigns. On the other hand, satellite images

can provide a continuous and much more frequent tool for acquiring images at the cost of a lower resolution than airborne images.

In this study, time series of multi-sensor satellite images are utilized to monitor the water quality of inland water bodies using optical remote sensing techniques. The image processing is introduced as a semi-automated procedure that enables the ongoing generation of maps over time. Additionally, the findings and data are shared through a collaborative platform dedicated to data sharing. In addition, to encourage the constant production of Water Quality Parameter (WQP) maps and their publication, semi-automated procedures and web-based applications, have been developed, to enable a broader audience to produce and access water quality estimates with limited expertise in the field.

The United Nations (UN) define water resources as part of one of the Sustainable Development Goals (SDGs). In 2021, the coordinating mechanism of the UN under the umbrella of SDG 6, UN-Water, delivered the "Summary Progress Update 2021: SDG 6 — Water and Sanitation for All" for ensuring a coordinated response to water-related challenges. The document details the UN's agenda concerning the status, challenges and goals for the conservation and management of water resources. In particular, target 6.3 focuses on water quality improvement by reducing pollution. Furthermore, target 6.5 remarks on the importance to strengthen integrated water resource management, including transboundary cooperation (Bresciani et al., 2018). UN-Water (2021) reported that 60% of the world's monitored water bodies have good ambient water quality, yet it is challenging to assess global status and trends. Warner et al. (2020) define good ambient water quality as the positive assessment of water to consider it as not harmful to the ecosystem or human health, accounting for the combination of effects of natural and anthropic pressures. However, UN-Water (2021) estimates of the ambient water quality variables are based on data from less than half of all countries (96 countries with data), of which many rely on relatively few measurements. Out of 49 countries reporting 2017 and 2020, 19 are

on track to improve water quality. Urgent action is necessary to improve surface and groundwater monitoring systems and define water quality standards. Ambient water quality needs improving regardless of national socioeconomic status, including transboundary cooperation. According to UN-Water (2021) there is a medium to low implementation of Integrated Water Resource Management (IWRM), with only 54% degree of implementation in 186 countries by 2020. The IWRM framework presents some key aspects addressing the drivers concerning the pressure agents on water resources, which affect the options and requirements for water management (UNEP-DHI Centre for Water and Environment, 2009). These components focus mainly on socioeconomic aspects. However, some less controllable factors, such as climate change, are included.

The present work was developed under the framework of the SIMILE Interreg project, between Italy and Switzerland, for the conservation of the water resources in the Insubric area (Brovelli et al., 2019). The project aims at providing innovative tools for water quality assessment, which favours some of the critical components defined within SDGs targets, such as: Collaborating in the management of freshwater and coastal water: Advances in the management and the coastal zone as a continuum; Ensuring sustainable water infrastructure: Assists in giving a cross-sectoral view of the water: development and a multipurpose infrastructure. Adequate cross-sectoral coordination, a sound scientific knowledge base, capacity assessment, water allocation mechanisms, environmental regulations, etc., prevent investment failures or sub-optimal economic and financial cost recovery. Planning transboundary collaboration: Assists water management with the catchment as the management unit, irrespective of whether it is within national boundaries or shared between two or more countries.

Understanding the world through science-based knowledge can improve the quality of life by managing necessary natural resources for commodities and services. However, science-based output

dissemination for the efficient management of resources has several challenges. An important example of this is the publication of a water management investigation, where users, researchers and managers hold different interests and perspectives (Hopfensperger et al., 2021; McKinnon & Vos, 2015; Weingart et al., 2021). For instance, many users are unaware of the water bodies' origins and factors influencing the resource availability and quality. Likewise, stakeholders may be uninformed of the consequences of the anthropic activities' effects on the waters. The lack of awareness usually leads to difficulties in reaching a consensus regarding the tradeoffs related to water management strategies (Golladay et al., 2021). On the other hand, researchers, often in academic environments isolated from the external public, fail to deliver the outcomes of their work efficiently. Public Engagement with Science (PES) defines itself as a strategy for "creating intentional, meaningful interaction that provides opportunities for mutual learning between researchers, water managers, and the general public water users"(Golladay et al., 2021). PES is not a unique solution; it changes according to their settings in terms of context or audience size. Yet, it should go beyond being unidirectional concerning knowledge transfer through exchanging views on the topic of interest.

In this research, the project SIMILE ("Integrated monitoring system for knowledge, protection and valorisation of the subalpine lakes and their ecosystems", Brovelli et al. (2019), served as the study case for the study and tools implemented for the monitoring of lake WQPs. The project recognizes the great availability of physical, chemical, and biological data in time, but these are not sufficient for the description for the limnological processes in the lakes in a short time scale. SIMILE aims at extending spatial and temporal scale investigations, through an innovative approach by integrating satellite data, in-situ sensor data, and user-contributed georeferenced data. SIMILE is a project financed by the Cooperation Programme INTERREG Italy-Switzerland 2014-2020 (*Sistema Informativo per il Monitoraggio Integrato dei Laghi insubrici e dei loro Ecosistemi*, 2022). The project concerns the lakes in the subalpine region of Canton Ticino and the Lombardy Region (specifically, Lake Maggiore, Lake

Lugano and Lake Como). The project sees Politecnico di Milano and the SUPSI (*Scuola Universitaria della Svizzera Italiana*) as the project leaders and CNR (*Consiglio Nazionale delle Ricerche*), Fondazione Politecnico, Lombardy Region and Canton Ticino as partners. The project's goal is to support the decision-making system and the definition of politics of the region of interest through an advanced information system based on innovative technologies for lake monitoring. The project encouraged an ecosystem perspective, where the aquifers in the cross-border area between Italy and Switzerland should be protected and restored, involving the local communities for improving the water quality of the regional water bodies.

The primary focus of this work is to develop and implement the essential tools for generating maps of water quality parameters. Additionally, it involves designing and implementing a WebGIS platform for effective sharing of these maps. The outcomes of the work provide a set of tools for monitoring WQPs, from satellite data acquisition to the publication of the WQPs maps which in this case aligns with the goals of project SIMILE. Its main goal is to enhance decision-making and data access through a strategic water quality monitoring tool, on the one hand, by implementing an environment using the appropriate set of algorithms and technologies that suit the study case for WQPs maps production.

This work introduces a semi-automated procedure to produce water quality parameter (WQP) maps using optical satellite images. The procedure incorporates two crucial components: co-registration techniques for rectifying misalignments and statistical inference tests for extracting outliers. The integration of these methods aim at improving the reliability of the WQP maps, ensuring the provision of high-quality data for water quality assessment and analysis.

On the other hand, thanks to the collaborative data-sharing platform for publishing the produced maps. The platform makes use of free and open-source infrastructure and open standards. Free and Open-Source Software (FOSS) technologies for the platform guarantee access to information without

concerns of licenses or purchase constraints. The collaborative data-sharing platform includes two web-based applications targeting different end-users. The first corresponds to a customized implementation of the GeoNode docker-composition of FOSS that allows the publication of geospatial data. The customized GeoNode targets the partners involved in producing the water quality parameters maps. The second is a WebGIS developed for exposing the WQPs maps uploaded through Open Geospatial Consortium (OGC) protocols provided by the first application. The WebGIS is dedicated both to stakeholders and citizens. Moreover, the collaborative platform looks forward to reducing the technological barriers to sharing and visualizing the WQPs maps monitoring.

The outline of this work describes the following points. 1) Introduction and relevance of the monitoring of water quality status for the management of water resources. 2) State-of-art of the use of remote sensing for the study of water bodies, in particular inland water bodies, using optical satellite images. The section presents the relevance of remote sensing in water quality assessment and current advances in sensors and algorithms devoted to inland water bodies monitoring. 3) Introduction of the inland water bodies used as study cases and the workflow used to produce the WQPs maps. 4) Assessment and additions to the presented workflow for detecting outliers in the WQP maps, including challenges and additional considerations for their production. 5) Design of a web-based application to disseminate geospatial data. In addition, the design and implementation of the platform for the publication and visualization of the WQPs maps through web-based applications. 6) Final remarks and conclusions. In the appendix sections, extensions of the work to benefit and improve the implemented tools for water quality monitoring can be found.

The solutions implemented in this study are intended to be easily replicable, allowing others to benefit from them through the sharing of methodologies and tools. By promoting replicability, these solutions can be applied to various other study cases beyond the initial scope of this research. The design and

implementation of the platform play a significant role in fostering the sharing of geospatial data. By providing a user-friendly interface and accommodating both expert and non-expert users, the platform encourages active participation and contribution from a diverse range of stakeholders. This collective sharing of data not only enhances the monitoring of dynamic geospatial variables but also facilitates knowledge-based support for public administrations. Public administrations can leverage the platform's capabilities to access valuable geospatial data, enabling them to make informed decisions concerning various environmental aspects, such as water quality, land use, and natural resource management. This knowledge-based approach empowers public administrations to develop more effective policies and strategies, leading to better protection and conservation of natural resources. The openness and replicability of the implemented solutions, combined with the data-sharing features of the platform, create a collaborative environment where researchers, experts, public administrations, and citizens can work together towards common environmental goals. This approach fosters a more sustainable and data-driven approach to managing our natural resources and promotes the overall well-being of our ecosystems.

Chapter - 2. REMOTE SENSING OF INLAND WATER BODIES

Inland water bodies act as sentinels when it comes to environmental changes such as climate change (Erler et al., 2019), population pressure (Pokhrel et al., 2017), and land use and land cover change (Mellander & Jordan, 2021). Rapid and uncontrolled environmental change, including nutrient pollution, drought, urbanization, and engineered changes to watersheds, frequently have adverse effects, including accelerated eutrophication, the growth of toxic blue-green algae, extreme turbidity and deterioration of water transparency, loss of aquatic benthos, and detrimental impact on human and animal health (Groom et al., 2019). Water quality management must be a major priority for environmental regulatory authorities worldwide, given the crucial uses of inland waters. "The Drama of the Commons"(*The Drama of the Commons*, 2002), presented by the National Research Council, outlined seven major difficulties in managing environmental resources. These underlined the necessity for low-cost environmental monitoring programs employing remote sensing technology, including "low-cost enforcement rules" and "monitoring the resource and users" compliance with the rules". To effectively control water quality, monitoring surveys employing remote sensing should be carried out frequently alongside less frequent field-based techniques (Ostrom et al., 2003). Remote sensing offers a low-cost, high-frequency, wide-coverage, and valuable alternative to conventional field-based water quality monitoring techniques, which are typically expensive and labour-intensive (Duan et al., 2010; Hadjimitsis & Clayton, 2008). To accurately monitor inland water resources and isolate the natural and anthropogenic stressors, new techniques must be developed, or bio-optical models and algorithms must be fine-tuned. Remote sensing's main advantage is its ability to perform large-scale synoptic monitoring of water resources.

2.1 OCEAN LAND COLOUR INSTRUMENTS FOR WATER QUALITY MONITORING

Research on remote sensing of water bodies has been performed to better understand the global cycle of carbon and how marine ecosystems respond to climate change, biochemistry variability, and feedback processes (Ogashawara et al., 2017). Since the 1970s, remote-sensing techniques for water quality monitoring applications using ocean colour began assuming the possibility of deriving water parameters such as chlorophyll-a (Chl-a, as a proxy for phytoplankton biomass) and surface water temperature (Gordon et al., 1988; A. Y. Morel & Gordon, 1980). The initial studies' assumptions led oceanographers to remotely monitor the optical properties of water constituents, such as phytoplankton, coloured dissolved organic matter, and total suspended solids. Due to these studies, hydrologic optics and radiative transfer theory concepts have been inherited for the currently known bio-optical modelling.

The analysis of complex waters are dependent on the characteristics of each ecosystem for retrieving the water constituents. In this work is presented the study case of subalpine lakes following the indications presented by Luciani et al. (2021). The study describes the lakes of Maggiore, Lugano, and Como, located in the Italy-Swiss cross border area, and introduces a workflow to produce WQPs maps. Here, the workflow has been implemented as a semi-automatic procedure to derive the WQP maps of chlorophyll-a, total suspended matter, and lake surface water temperature. The processing chain considers two optical satellites sensors, Sentinel-3 OLCI and Landsat-8 TIRS, for the production of the WQP maps. In the following sections are presented the evolution of the Ocean Land Colour Instruments for water quality monitoring, principles of optical remote sensing, and satellite images and methods used for the WQP processing (i.e. C2RCC processor, for Chl-a and TSM, and Barsi Method, for LSWT).

Water quality assessment is possible through remote sensing approaches based on the water constituents that affect the absorption and scattering from the incoming visible light spectrum (i.e. the visible fraction of the electromagnetic spectrum, ranging between $0.4-0.7\mu$). These are also called optical WQP and are the following: (i) Concentrations of chlorophyll-a pigments and cyano phycocyanin (CPC), suspended matter and aquatic humus; (ii) presence or absence of algae blooms; (iii) floating submerged macrophytes; (iv) vertical attenuation coefficients of downwelling and upwelling light; (v) water transparency; and(vi) point source pollution, e.g. oil spills (Ogashawara et al., 2017). It is possible to monitor the optical water constituents as a result of hydrologic optics and radiative transfer theory which led to the development of the bio-optical models (A. Morel, 2001).

2.1.1 Bio-optical properties and reflectance of the waters

Remote sensing of inland waters utilises the optical properties of water constituents. Two categories define the optical properties (Preisendorfer et al., 1976):

1. Properties that depend on the medium and directional structure, known as apparent optical properties (AOPs), and

2. Those which depend only on the medium and are independent of the ambient light field, known as inherent optical properties (IOPs)

Optical properties, such as irradiance reflectance (R), above and below water remote sensing reflectance (R_{RS} or r_{rs} respectively), and various diffuse attenuation functions (K) are AOPs since they vary according to the composition of the medium and the light field (C. Mobley, 2020). Properties such as absorption (a) and scattering (b) coefficients, which vary based on just the composition of the medium or constituents, are IOPs (C. Mobley, 2021).

For remote sensing of inland waters, the most commonly used AOP is R_{rs}, known as remote sensing reflectance, which is defined as the ratio of water-leaving reflectance radiance (L_w) to downwelling irradiance (E_d).

$$R_{rs} = \frac{L_w}{E_d} \ [sr^{-1}]$$
<div align="right">[Eq. 2-1]</div>

Since instruments for measuring scattering in inland waters are not widely used, absorption properties of water constituents are the most commonly used parameters for IOPs. More research is needed to evaluate their performance in these optically complex waters. The size and shape of R_{rs} in an aquatic ecosystem can be affected by varying compositions of constituents and IOPs that are associated with them.

At each study site, the variations in R_{rs} are connected to the variations in IOPs like a_{phy}, a_{NAP}, and a_{CDOM}. Ternary plots of the absorption coefficients at three distinct characteristic wavelengths – 440, 560, and 665 nm – can illustrate the compositional variability of the IOPs. The plots exemplify the optical complexities of inland waters and how multiple constituents influence reflectance at a specific spectral region. Because of that, bio-optical models for inland waters are increasingly focusing on red-NIR bands (Ogashawara et al., 2017).

Radiometric quantities, or IOPs and AOPs, such as downwelling spectral solar and sky radiation and the absorption and scattering properties of water column constituents, are the foundation of bio-optical models. There are two possible ways to define bio-optical models. The first definition describes the aquatic system's "bio-optical state" in various ways (A. Morel, 2001), indicating that the biological and geomorphological processes in the water body influence the optical properties. By establishing relationships between radiometric measurements and optically active constituents, bio-optical models frequently aim to obtain information about the biological and physical processes in the water body.

Based on statistical relationships between radiometric quantities, these models are typically categorized as empirical and descriptive, as stated by Morel (A. Morel, 2001). Bio-optical models have a second definition, which involves utilizing the radiative transfer theory to obtain optical characteristics of substances present in the water column (C. D. Mobley, 2001). Quantification of IOPs is the foundation of these bio-optical models, and a ratio of their absorption coefficients (a) to specific absorption coefficients (a^*) and backscattering coefficients (b_b) to specific backscattering coefficients (b_b^*) can be used to determine the constituents' composition. Analytical models are the common name for these bio-optical models. The utilization of bio-optical models has facilitated the monitoring of inland water quality, primarily through the estimation of chl-a or TSS concentrations, serving as indicators of the water body's trophic or turbidity condition. When coupled with multiplatform observations (such as in situ, airborne, and space-borne), these models can prove to be highly valuable in comprehending the interplay between the water body and its surrounding landscape.

Due to a lack of consistency in generally accepted terminology, various terms have been used to classify bio-optical models according to their formulation and objectives (Odermatt et al., 2012). Based on the formulation and final goals of the bio-optical models found in the literature, they may be classified into five broad categories (Niroumand-Jadidi et al., 2019; Ogashawara, 2015): 1) empirical, 2) semi-empirical, 3) semi-analytical, 4) quasi-analytical, and 5) analytical models. The current study focuses on empirical and physical modelling, which are applied by the processor used inside the processing chain defined by Luciani et al. (2021) deemed to be suitable for the lakes of interest. Empirical models use statistical relationships between in-situ measurements of water constituents and radiometric data from satellite sensors or proximal remote sensing devices. The models base on a combination of R_{rs} at different wavelengths, which provide the best correlation between reflectance and concentration of optically active water constituents.

2.2 EXISTING PRODUCTS FOR WATER QUALITY MONITORING - EUMETSAT

The European operational satellite agency for monitoring weather, climate, and the environment from space (EUMETSAT) is an intergovernmental organisation, whose activities refer to the operation of a system of meteorological satellites for observing the atmosphere, oceans, and land. In addition, EUMETSAT provides several services of satellite data and products publication useful for weather forecasting, the monitoring of the environment and climate change. EUMETSAT delivers products related to the Sentinel missions devoted to monitoring the atmosphere, ocean and climate in cooperation with the European Space Agency and the Sentinel-3 Marine mission.

EUMETSAT uses the Sentinel-3 Ocean and Land Colour Instrument (OLCI) to produce marine Level 1B and Level 2 products (Lavender, 2021). The Level-2 products delivered include pre-processing outputs such as pixel classification, gaseous correction, computation of the Integrated Water Vapour product (IWV) and the smile correction (Top of Atmosphere – TOA reflectance correction after the gaseous absorption correction), adjustment for white caps, sun-glint, and application of System Vicarious Calibration (SVC) gains. The default two atmospheric corrections (AC) methods are run in parallel: Baseline Atmospheric Correction (BAC) and Alternative Atmospheric Correction (AAC). The BAC is used for the Ocean Colour for MERIS (OC4ME) chlorophyll products (*OC4Me Chlorophyll - Sentinel Online*, 2023), while the AAC feeds the Neural Network (NN) of the C2RCC processing and outputs of CHL and TSM (among others). OC4ME and NN WQP maps for CDOM, CHL-a and TSM are the standard products delivered by EUMETSAT.

The OC4ME Maximum Band Ratio (MBR) semi-analytical algorithm, created by Morel et al. (A. Morel et al., 2007), defines chlorophyll concentration, Chl (OReilly et al. (O'Reilly et al., 1998), for a more general description of such algorithms). The MERIS pigment index algorithm, which is fully described

in Morel et al. (A. Morel & Maritorena, 2001) and the MERIS ATBD 2.9, is at its most recent iteration. The polynomial OC4Me is based on the semi-analytical model, which was, in turn, based on the analysis of AOPs measured in situ over the previous decades in various oceanic regions (Bricaud et al., 1998) (A. Morel & Maritorena, 2001) (algorithms which were developed for case 1 waters).

$$\log_{10}[Chl] = \sum_{x=0}^{4} \left(A_x \left(\log_{10}\left(R_j^i\right)^x \right) \right)$$

[Eq. 2-2]

R_j^i is the proportion of band i's irradiance-reflectance at 443, 490, and 510 nm to band j's at 560 nm. The band selection aims at maximizing the ratio.

The output for the AAC processing is water reflectance, which is an internal processor product. The estimates are based on a NN, considering the correction over complex waters with significant CDOM and/or TSM contributions. The TOA reflectances in 15 bands (400-753nm, 778nm, 865nm and 1020nm), corrected for absorbing gases and smile effect, are used together with wind speed, salinity, temperature, altitude of the water surface and observation geometry to estimate the water reflectance and atmospheric parameters. These results are not provided with the standard products. The second NN retrieves Inherent Optical Properties from the internal water reflectance, including total back-scattering coefficients b_{bp}, phytoplankton absorption coefficient $a_{(pig)}$ and coloured Detrital and Dissolved Material absorption coefficient $a_{detritus+ag}$, all at 442.5nm (Doerffer and Schiller (Doerffer & Schiller, 2007) presented the alternative approach developed for MERIS). The derived products from the IOPs are the following:

$$ADG443_{NN} = ADG443 \left[\frac{1}{m}\right]$$

[Eq. 2-3]

$$CHL_{NN} = 21a_{pig@442.5}^{1.04} \left[\frac{mg}{m^3}\right]$$

[Eq. 2-4]

$$TSM_{NN} = 1.06 b_{bp@442.5}^{1.04} \left[\frac{g}{m^3}\right]$$

The NNv2 processor reports similar results to the C2RCC implemented within SNAP. However, some differences may concern pixels classification, gaseous correction, and SVC applied to the operational processor but not automatically included in SNAP processing. The Algorithms Theoretical Basis Document can be found (ATBD) are publicly accessible (*Ocean Colour Services | EUMETSAT*, 2018).

This works aims at delivering WQPs maps targeting the specific ecosystems under the study case. The analysis on previous investigations devoted to the characterisation of the lakes. By lakes characteristics, it refers to the bio-optical model parameters necessary for the modelling of the WQPs. While for the WQPs modelling, the same processor has been used, it is expected that the use of these modelling parameters will lead to improving the WQPs maps estimated to those presented by EUMETSAT.

2.3 WQP PRODUCTION CHALLENGES

Various compounds contribute to the optical signature of the water body, i.e. the concentration of the phytoplankton pigment Chlorophyll-a, inorganic suspended sediments, and coloured dissolved organic matter govern the light field reflected from the water surface of coastal waters, estuaries, rivers, and inland waters, and recorded by an optical satellite Brockmann et al. (2016). The retrieval of concentrations of these elements requires bio-optical models that contain the constituents' inherent optical properties (absorption and backscattering coefficients), concentration ranges and covariations, and a mechanism for inverting the water leaving reflectance spectrum. Even though satellites measure the light field emerging from the top of the atmosphere, an atmospheric adjustment must be performed as a part of the data processing of water. Because of the low reflectance's on the ground, the atmospheric route radiance accounts for more than 90% of the signal seen by a satellite. As a result, atmospheric adjustment is critical (Brockmann et al., 2016).

For waters with CDOM, typical values of above-water remote sensing reflectance (R_{rs}) can be in the range of 1% in peak reflectance bands, and much lower in the NIR areas with efficient absorption by water soluble-compounds. As a result, achieving a good optical signal-to-noise ratio is critical whe inferring optical and biophysical properties of the observed water body from a remote sensor. In the visible spectrum, atmospheric path radiance surpasses water leaving radiance by at least 80%-90%, and this is primarily due to molecule and aerosol scattering in the atmosphere, which decreases with rising wavelengths (Warren et al., 2019). These signal components must be eliminated so that the residual signal may be assigned to the water's surface reflectance. Additional impacts at the water's surface, such as whitecaps, sun-glint, adjacency, and proximity to land, may also be considered. Their effects result in additional reflectance to that of the water column, posing problems for atmospheric correction. Validation of current procedures in a variety of water bodies and atmospheric

circumstances is necessary to aid in the improvement of atmospheric correction routines (Gholizadeh et al., 2016).

For the spectrally based retrieval of constituents, a number of strategies have been developed, which can be divided into three categories (Niroumand-Jadidi et al., 2019): empirical methods, semi-empirical methods and physically based methods. This research follows the third strategy, which is a physically based approach that relies on inverting the radiative transfer model to obtain water body elements. The considered method entails utilizing a neural network that have been trained on a large database of radiative transfer simulations.

Nevertheless, WQP maps production relying on remote sensing products come with limitations and challenges that must be taken into account. The temporal resolution and number of WQPs produced maps are dependent on the water conditions at the time of the image acquisition. The defined workflow benefits from optical satellite images, then it must deal with cloud coverage that might obstruct the sensor coverage for days. As a result, the dataset's temporal resolution is lower than the satellite revisit time (Hestir et al., 2015). Sun glint is a significant problem when observing water colour from space, resulting in anomalously bright pixels that should be treated (Overstreet & Legleiter, 2017). In addition, standard satellite data processing techniques, which are mainly designed for oceans, presume an unbounded water surface and, as a result, ignore the presence of neighbouring land. As an outcome, radiance reflected by the land and subsequently scattered by the atmosphere in the field of view of a satellite sensor observing the target water is a source of perturbations, resulting in water reflectance-based product errors. De Keukelaere et al. (2018) defines this phenomenon as adjacency effects, and it always happens when a scattering medium is placed in the proximity of the surface.

The optical characteristics of the lake waters under study, categorized as case 2 (to distinguish them from case 1 waters, corresponding to the oceans), are heavily influenced by inorganic and/or

sediments, necessitating high accuracy in atmospheric correction methods to properly retrieve water parameters. The problem of atmospheric corrections in case 2 waters has yet to be overcome. As a result, much effort has gone into developing atmospheric correction processors that span a wide range of methodologies (Warren et al., 2019, 2021). However, the performance of the processors varies depending on the scenario (sun observation geometry, atmospheric, optical, and site-specific conditions), and there is currently no standard approach; however, atmospheric correction processors continue to evolve as new methods and data become available. This brings the need to continue to test various atmospheric correction procedures as well as water quality retrieval methos using in situ data accounting for various water types and environmental circumstances (Soriano-González et al., 2022; Spyrakos et al., 2018).

Electromagnetic radiation's ability to spread through water is severely constrained by absorption. In most spectral regions, the penetration is relatively low and exceeds just 1 m in the visible; the constituents found in natural water further reduce it. Therefore, only this limited range is pertinent for biological activities using it (photosynthesis and vision), as well as for remote sensing of objects below the surface of the water. The physical method to quantitatively characterize these pathways and processes is Radiative Transfer theory. The terms electromagnetic radiation and light are used interchangeably in the discussion of radiative transfer theory and material properties, which is limited to the visible spectrum and only slightly expanded to the ultraviolet and infrared.

According to the RT theory, the density and physical characteristics of the insulating materials determine the wavelength-dependent changes in the direction and intensity of electromagnetic radiation. A mathematical formula plus a set of parameters makes up the theory. The radiative transfer equation is a mathematical formulation that links the physical processes of attenuation, scattering, and luminescence with the characteristics of matter.

The IOPs, absorption and backscattering, cannot be detected due to the low light penetration along the water column and the signal returned to the remote sensor is too weak to identify the presence of an algal bloom (Bresciani et al., 2018).

Figure 2-1 introduces a C2RCC processor sample output of a Chl-a map within SNAP's GUI. In this case, the products viewers display two different maps for a single output. On the left-hand side of the figure presents the initial output without applying the flags cited in Figure 2-1, where the white filled areas correspond to the estimates considered as "valid" after applying the filters. The right-hand side of the map presents the final output from the processing chain defined by the initial workflow. However, from the initial processing, it was noticed the presence of anomalous values which may indicate issues in the estimates in vicinity of the lake boundaries (see Figure 2-2), due to sun-glint effects, proximity to cloud coverages, etc.

A similar view from the outputs obtained from the computation of a LSWT map is shown in Figure 2-2. In the product viewer it is possible to review the original images in real colours, or false colours, for identifying possible disturbances that take place inside of the water body (as seen in the window on the top-left of Figure 2-2). A remarkable point from the figure is that after applying the masks devoted for the analysis of the LSWT, it is noticed the presence of darker blue colours around areas which have been masked by cloud coverages.
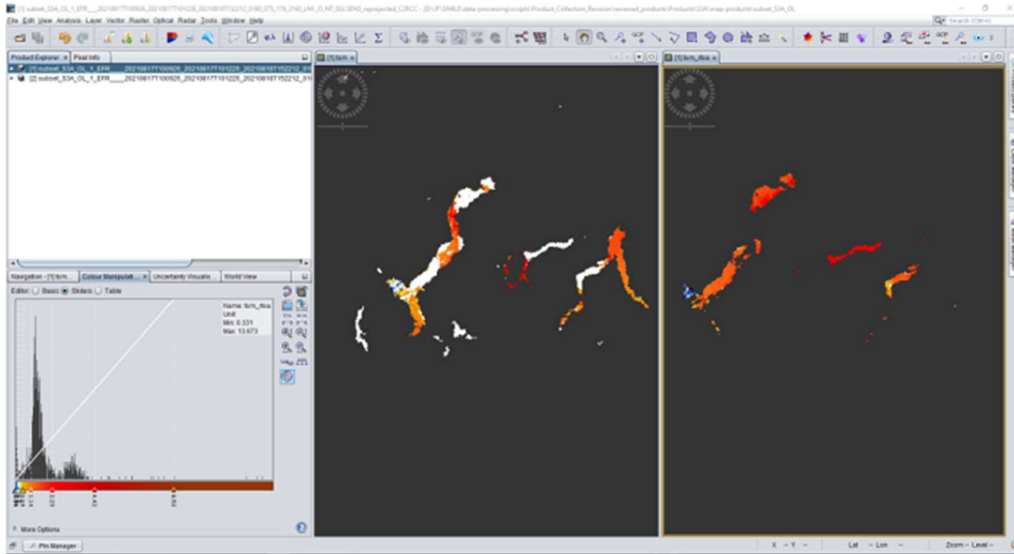
Figure 2-1 – SNAP example chlorophyll-a output from the C2RCC processor applied to a Sentinel 3 image
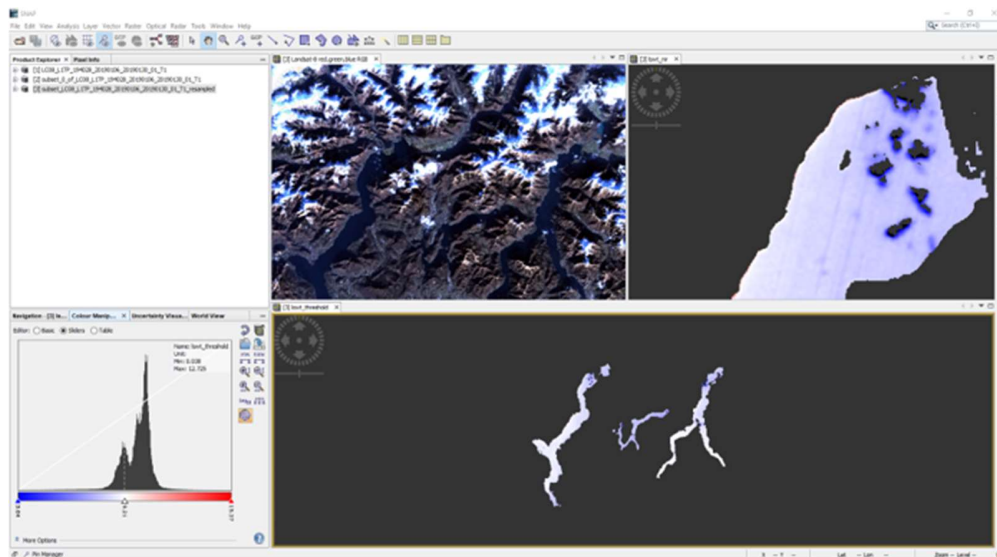
Figure 2-2 – SNAP example LSWT map output using the Barsi et al (2005) method from a Landsat-a image

Chapter - 3. DATASETS AND METHODS

In contrast to in-situ data, remote sensing technologies can monitor large areas with significant temporal coverage, capturing optically active water quality parameters (WQP)'s geographical and temporal variability (Bresciani et al., 2020; Free et al., 2021; Giardino et al., 2013, 2015). Inland water quality studies and monitoring applications based on optical satellites have increased dramatically in recent years as a result of the general ongoing improvement of satellite sensor design and resolutions (Bresciani et al., 2020; Topp et al., 2020). Earth observation has emerged as a well-established and proven method for lake monitoring, providing a comprehensive picture of the state of the water's quality. This is in addition to the conventional point-based sampling methods.

Previous studies in the area with the combined use of L8, S2A and S2B provided a collection of medium resolution optical sensors with an average revisit time of 2-3 days. Whereas, S3-A/B OLCI provides a revisit time of 1-2 days, with a lower spatial resolution of 300m, but offers 21 spectral bands inside the visible-NIR regions which allows for the distinction of phytoplankton families (while sensors such as S2 are not able to distinguish between them). Bresciani et al. (2018) mentions that a multi-source approach is the most suitable operational and full monitoring of lake quality. For this reason, this study takes advantage of the capabilities of the S3-A/B OLCI images and Landsat-8 TIRS. In addition, the following sections introduce the methods integrated within the processing workflow of the WQP maps. The extensions to the workflow aim to align the WQPs maps relative shift due to the Sentinel-3 images geolocation and to identify outliers in the WQP maps. The first considers the use of the GeFolki algorithm for the alignment of the WQP maps. The second relies on statistical methods based on the WQP maps to determine the presence of anomalous.

3.1 SATELLITE SENSORS FOR WQPS MONITORING STUDY CASE

Sentinel-3 A/B Ocean Land Colour Instrument

In general, a conditioning factor for the monitoring of different processes is the frequency of available data. This is not different for following the dynamics of the WQPs. For this reason, this work adopted the use of the Sentinel 3 A/B OLCI sensors. First, the revisit time for the sensors can be approximately a day. Second, the sensors provide a rich number of bands within the visible and near infrared zone of the electromagnetic spectrum. While the spatial resolution is the main limitation, when compared to other ocean colour sensors such as MODIS (250m), it became a widely used source for ocean colour research and monitoring.

In Europe, the Copernicus programme plans four Sentinel 3 spacecraft carrying the Ocean Land Colour Instrument (OLCI) with Sentinel 3A and 3B launched in 2016 and 2018. The motivation for Copernicus is ensuring continuity and consistency for the observations enabling the construction of services utilising the Earth Observation (EO) data. Copernicus has an Open Data policy and supports upstream data processing; for OLCI, processing data from L1 and L2 will be undertaken by European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) and are data available via File Transfer Protocol – FTP, and satellite digital broadcast (Eumetcast). Level 3 and regional (European) products are provided through the Copernicus Marine Environment Monitoring Service, while the CDRs are available from the Copernicus Climate Change Service.

OLCI observation is performed simultaneously in 21 spectral bands, listed in the table below, ranging from the visible to the near-infra-red (400nm to 1020nm). Each of these bands is programmable in position and width (*User Guides - Sentinel-3 OLCI - Radiometric Resolution - Sentinel Online - Sentinel Online*, 2023).

Band	Function	Spatial Resolution [km]	LB [μm]	UB [μm]
Oa01	Aerosol correction improved water constituent retrieval	0.3	392.5	407.5
Oa02	Yellow substance and detrital pigments (turbidity)	0.3	407.5	417.5
Oa03	Chlorophyll absorption maximum, biogeochemistry, vegetation	0.3	437.5	447.5
Oa04	High Chlorophyll	0.3	485	495
Oa05	Chlorophyll, sediment, turbidity, red tide	0.3	505	515
Oa06	Chlorophyll reference (Chlorophyll minimum)	0.3	555	565
Oa07	Sediment loading	0.3	615	625
Oa08	Chlorophyll (2nd Chlorophyll absorption maximum), sediment, yellow substance/vegetation	0.3	660	670
Oa09	For improved fluorescence retrieval and to better account for smile together with the bands 665 and 680 nm	0.3	670	677.5
Oa10	Chlorophyll fluorescence peak, red edge	1.2	677.5	685
Oa11	Chlorophyll fluorescence baseline, red edge transition	1.2	703.75	713.75
Oa12	O2 absorption/clouds, vegetation	1.2	750	757.5
Oa13	O2 absorption band/aerosol correction	1.2	760	762.5
Oa14	Atmospheric correction	1.2	762.5	766.25
Oa15	O2 used for cloud top pressure, fluorescence over land	1.2	766.25	768.25
Oa16	Atmos. corr./aerosol corr.	1.2	771.25	786.25
Oa17	Atmospheric correction/aerosol correction, clouds, pixel co-registration	1.2	855	875
Oa18	Water vapour absorption reference band. Common reference band with SLSTR instrument. Vegetation monitoring	1.2	880	890
Oa19	Water vapour absorption/vegetation monitoring (maximum reflectance)	1.2	895	905

Band	Function	Spatial Resolution [km]	LB [μm]	UB [μm]
Oa20	Water vapour absorption, Atmospheric correction/aerosol correction	1.2	930	950
Oa21	Atmospheric correction/aerosol correction	1.2	1000	1040

Table 3-1 - Sentinel-3 Ocean Land Colour Instrument (OLCI) band characteristics. Radiometric resolution - 21 bands in VIS/SWIR

This work benefit from the OLCI Level-1 Earth observation processing mode. The mode is available in two resolutions (Full and Reduced Resolution). The processing chain used the non-time critical OL_1_EFR products (i.e., output during EO processing mode for Full Resolution). Where, non-time critical refers to the timeliness of the delivery of the products (in this case, within one month after the acquisition of the image). The individual products can be accessed through the Copernicus Open Access Hub[1].

The WQPs maps processing workflow described by Luciani et al. (2021) and detailed in section 3.6 has been automated to provide tools for ensuring their continuous production for monitoring the water resources by the public administration, even with limited expertise in remote sensing. The image processing and filtering are implemented in a set of scripts that benefit from the SNAP-Python API (Zuhlke et al., 2015) to obtain the WQPs maps. The WQPs' automated process involves the retrieval of the input values for various parameters and executing processing tools within SNAP, depending on the workflow followed by each WQP. For example, for the estimates of Chl-a and TSM, the parameters required by the C2RCC processor to describe the water ecosystem (such as CHL factor, CHL exponent, TSM factor, TSM exponent, temperature, among others) are retrieved from datasets with time series for the parameters matching the date of acquisition of the Sentinel-3 imagery. In this

---

[1] https://scihub.copernicus.eu/dhus/#/home

case, the only variable parameter is the temperature retrieved through the Lombardy Region Open Data Catalogue Socrata API[2].

Landsat-8 – OLI and TIRS

The Landsat 8 satellite consists of two instruments —the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). These two sensors provide seasonal coverage at a spatial resolution of 30 meters (5 visible and 1 near infrared bands); 100 meters (thermal; 2 short wave infrared); and 15 meters (1 panchromatic band). Landsat 8 was developed as a collaboration between the National Aeronautics and Space Administration (NASA) and the United States Geological Survey (USGS)(*Landsat 8 | Landsat Science*, 2021, p. 8). The revisit time for both sensor is of 8 days (16 days each). Landsat-8 products can be retrieved from the United States Geological Survey (USGS) Earth Explorer platform[3].

In this case, the Landsat-8 Collection 1 (before 2022-01-01) and 2 Level 1 Tier 1 have been used for modelling the LSWT by applying the Barsi method. Where, the measured radiance at the top of the atmosphere, L_{TOA}, correspond to the signal estimates of band-10 (TIR-1) from the Landsat-8 images. Then, the spatial resolution of the corresponding WQP maps is of 100m. The image acquisition of Landsat-8 products has been carried out individually for its processing inside the semi-automated procedure. Additionally, the LSWT maps were masked accounting for the presence of clouds, shadows, and unexpected values. The first two conditions are available in the quality assessment band of the Landsat-8 products. The pixels values in the quality assessment band provide information on the conditions (*Landsat 8 (L8) Data Users Handbook*, 2019). Here, the LSWT maps processing

---

[2] https://dati.lombardia.it/
[3] https://earthexplorer.usgs.gov/

accounted for the mask concerning "clouds", "Cloud Shadow – High" and "Cirrus confidence – High". A further step taken into the filtering of unexpected values, by discarding values below 0°C.

Landsat-8 OLI and TIRS	Band Detail	Spatial Resolution [m]	UB[μm]	LB[μm]
Band 1	Coastal/Aerosol	30	0.435	0.451
Band 2	Blue	30	0.452	0.512
Band 3	Green	30	0.533	0.59
Band 4	Red	30	0.636	0.673
Band 5	NIR	30	0.851	0.879
Band 6	SWIR-1	30	1.566	1.651
Band 7	SWIR-2	30	2.107	2.294
Band 8	Pan	15	0.503	0.676
Band 9	Cirrus	30	1.363	1.384
Band 10	TIR-1	100	10.6	11.19
Band 11	TIR-2	100	11.5	12.51

Table 3-2 - Landsat-8 band characteristics. Radiometric resolution - 11 bands in VIS, NIR, SWIR

3.2 WQP MAPS PROCESSING WORKFLOW

The multiple process and specific lakes characteristics can have a large effect in the selection of algorithms for the water quality assessment using remotely sensed data. In this case, the study case selection accounted for the replicability of the estimation of the production of WQP maps, then a wide part of the work consisted in identifying a tool which performed well in the analysis of the deep subalpine lakes. Luciani et al. (2021), following validation of results exposed by Giardino et al. (2013) and Bresciani et al. (2018) for the parametrisation of the algorithm, presented a workflow devoted to the analysis of a set of subalpine lakes in the Insubric region showing that the C2RCC processor displayed good results for monitoring these lakes. Figure 3-1, describes the processing chain for retrieving a set of WQPs, Chlorophyll-a and Total Suspended Matter, taking advantage of the C2RCC processor.

Figure 3-1 present a set of additional steps to the original processing chain highlighted with green background. These steps consider a set of procedures address different items identified on the WQP maps to enhance the WQP maps' cross-temporal monitoring. The procedures integrated correspond to the image co-registration and statistical analysis of the WQPs maps. These actions have been considered to fix relative shift in between the WQPs maps derived from the sentinel-3 OLCI images and to detect and remove anomalous values which were not targeted by the masks of the C2RCC processor.
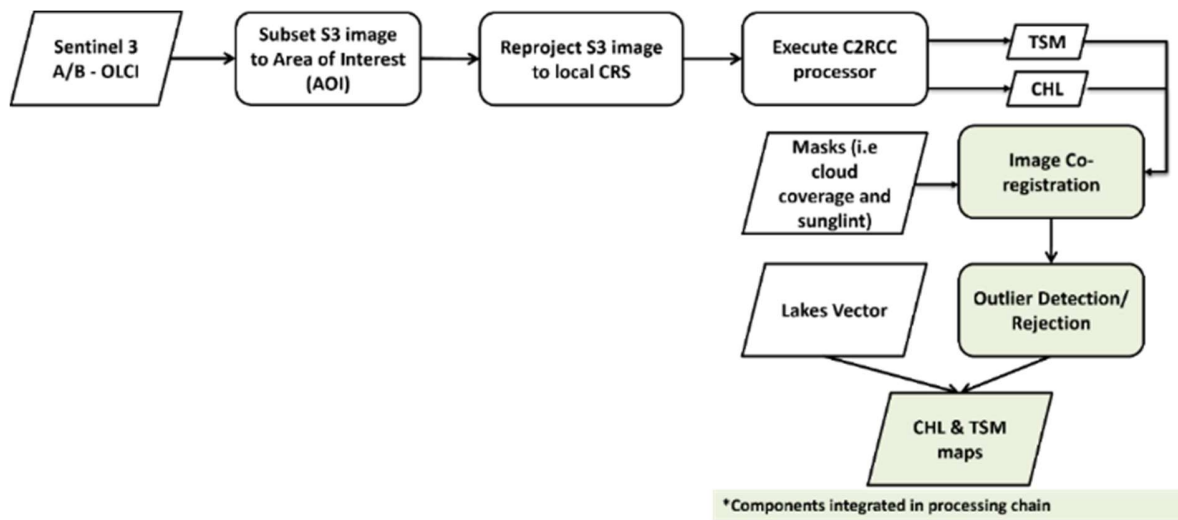
Figure 3-1 – Sentinel-3 images processing workflow for WQP maps

The TIRS sensor's images were radiometrically calibrated, and the Barsi method was used to determine the lakes' surface temperatures (Barsi et al., 2005). The radiative transfer equation is used to calculate the BOA radiances, and the inversion of the Planck equation is used to approximate the water's surface temperature using specific Landsat 8 constants. Additionally, in this instance, the Landsat cloud masks on some images were insufficient to exclude pixels affected by cloud shadows or thin clouds, necessitating additional masking in SNAP based on user selected NIR band thresholds. Furthermore, the statistical outlier rejection method for anomalous values that were not extracted by the initial processing chain masks is incorporated into the processing of the LSWT maps (see Figure 3-2).
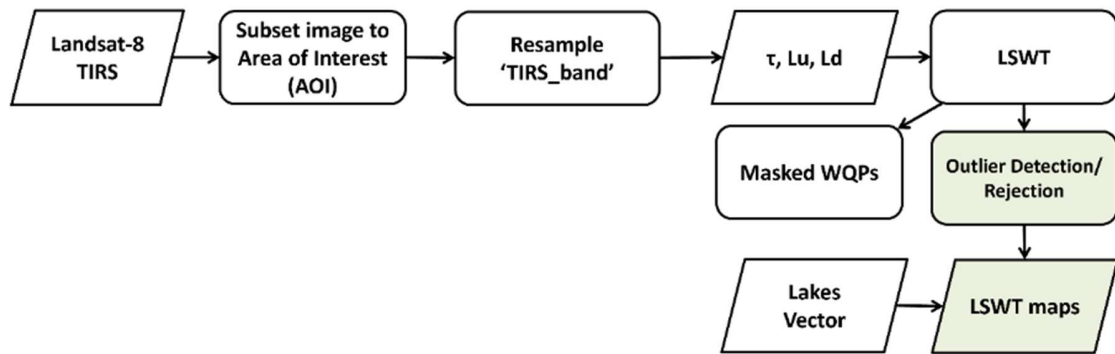
Figure 3-2 – Landsat-8 images processing workflow for WQP maps

Images are processed using opensource and free technologies (Luciani et al., 2021), which have been validated for inland aquatic environments (Free et al., 2021), and are implemented in the European Space Agency's (ESA's) free software Sentinel Application Platform (SNAP) (Zuhlke et al., 2015). Case 2 Regional Coast Colour (C2RCC) (Brockmann et al., 2016) employs a neural network to perform radiometric and atmospheric corrections and compute Chlorophyll-a concentrations using Sentinel-3 data. Instead, LSWT is calculated using the Barsi Method (Barsi et al., 2005), a methodology validated and used in different case studies (Amadori et al., 2021; Sharaf et al., 2019). Furthermore, the implementation of the processing environment integrates a Python-based library for the processing of the satellite imagery. The developed library includes a class for the WQP processing used for applying the image coregistration algorithm and outliers detection analysis. Figure 3-3 displays the WQP parameters maps outputs from the processing chain for the corresponding sensors.
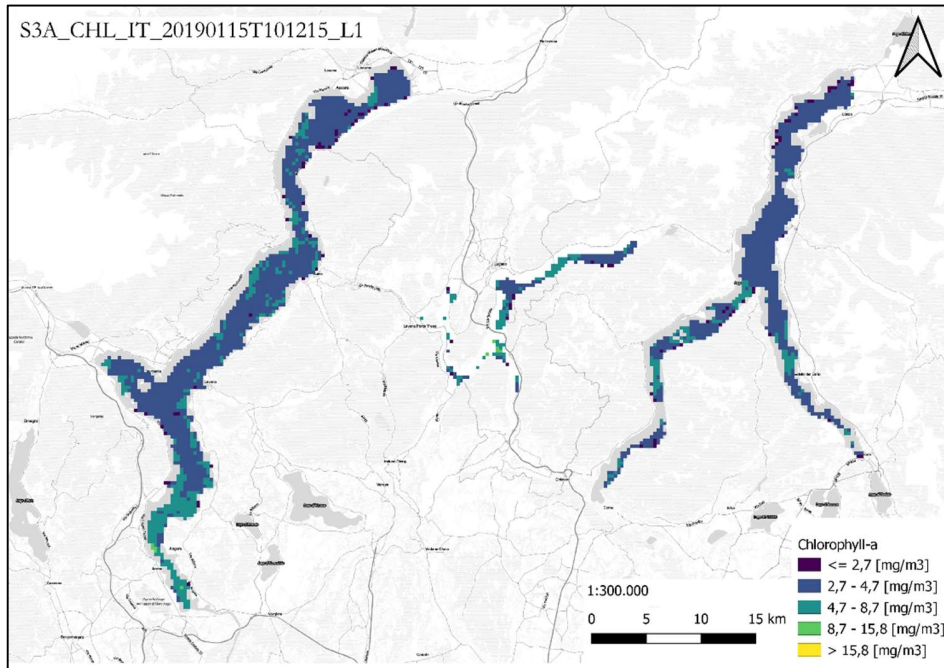
(a)
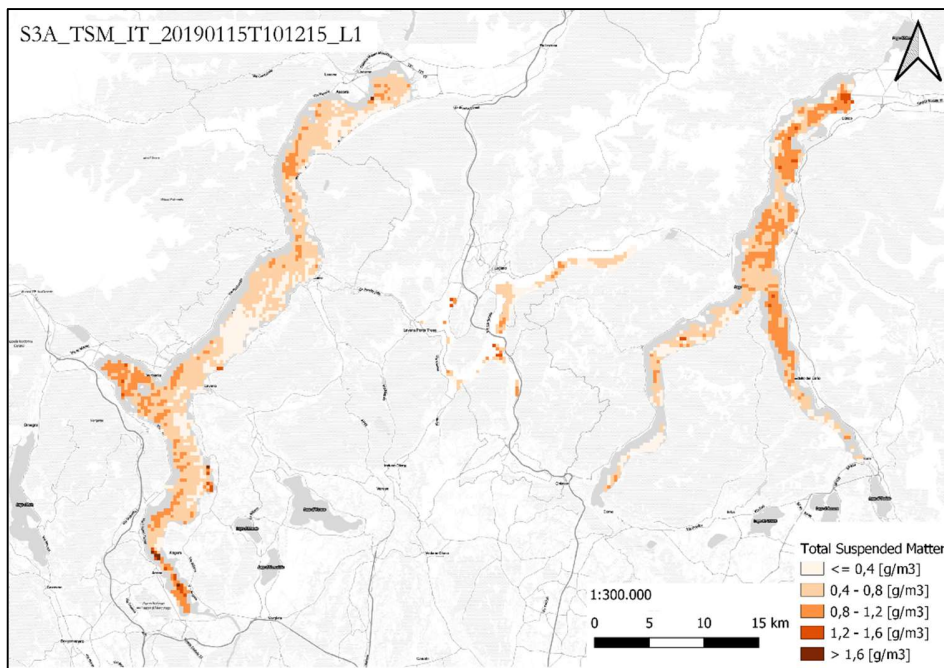
(b)
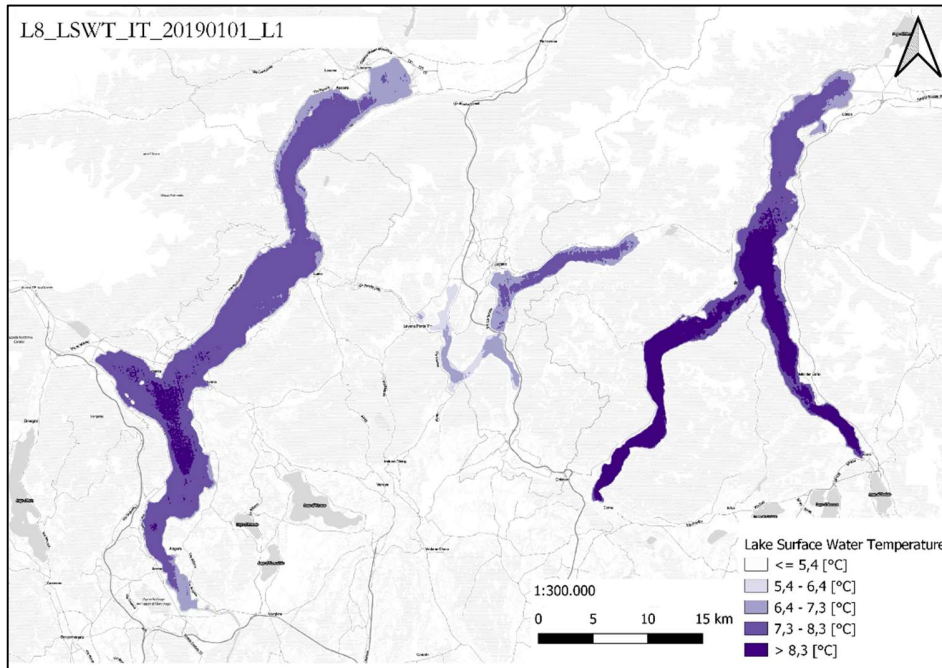
(c)

Figure 3-3 - Sample outputs of the processing workflows of WQPs maps. (a) Chlorophyll-a map on 15-01-2019 derived from Sentinel-3 OLCI; (b) Total Suspended Matter on 15-01-2019 derived from Sentinel-3 OLCI; (c) Lake Surface Water Temperature on 01-01-2019 derived from Landsat-8

3.2.1 ESA Sentinel Application Platform (SNAP)

The Sentinel Application platform, or SNAP, is a common architecture for all Sentinel toolboxes developed by Brockmann Consult, SkyWatch, and C-S. It is a Java-based software that is ideal for Earth Observation. It has been developed with an emphasis on Extensibility, Portability, Modular Rich Content Platform, Generic EO Data Abstraction, Tiled Memory Management, and Graphic Processing Framework. The software lets the user work with various geospatial data sources in text, vector, and raster formats[4]. It is feasible to read and analyse various information sources and their metadata giving different modules and plugins. The software also supports multithreading and multi-core processing and incorporates a product library for quickly scanning and cataloguing large archives. The Graphical User Interface (GUI) integrates WorldWind for data visualization. Some of the core

---

[4] http://step.esa.int/main/doc/tutorials/

technologies and frameworks implemented for the platform are NetBeans platform [5] (desktop application framework), Install4J [6] (multi-platform installer builder), GeoTools [7] (geospatial tools library), GDAL[8] (reading/writing raster vector geospatial data formats), Jira[9] (issue tracker) and Git[10] (version control). Alongside the installation procedure it is possible selecting specific toolboxes for the analyses of Sentinel-1, Sentinel-2, Sentinel-3, within others for specific satellites. The source code of the project is available on GitHub[11].

In this work, the processing chain benefits from two tools inside of SNAP. The first corresponds to the C2RCC processor plugin within the platform, and the second being the band calculator for implementing the Barsi algorithm (Barsi et al., 2005).

3.2.2 Case 2 Regional Colour Coast (C2RCC) Processor

The C2RCC processor (Coastal and Continental Shelf Reflectance Correction Algorithm) is a computer program used to correct atmospheric and adjacency effects in satellite images of coastal and continental shelf waters. Adjacency effects refers to the influence of the reflectance of surrounding features in the signal detected by a remote sensing instrument of the observed target. The phenomenon can have a significant impact on measurements' accuracy of water quality and quantity. An example of this is the influence of land reflectance on water reflectance measurements. When land surface light is reflected into the water, it can change the spectral signature of the water, making it difficult to measure WQPs. Also, it is the case for the influence of clouds and atmospheric conditions, where aerosols can scatter and absorb light that can lead to errors in the WQPs estimations. The

---

[5] https://netbeans.org/features/platform/
[6] http://www.ej-technologies.com/products/install4j/overview.html
[7] http://www.geotools.org/
[8] http://gdal.org/
[9] https://www.atlassian.com/software/jira
[10] https://git-scm.com/
[11] https://github.com/senbox-org

C2RCC processor is designed to improve the accuracy of water quality measurements derived from satellite images by removing the influence of the atmosphere and surrounding pixels on the reflectance values of a particular pixel[12]. The algorithm is specifically designed to process satellite data from ocean colour missions, such as Sentinel-3, MODIS Aqua and Terra, VIIRS, MERIS, and SeaWiFS.

The C2RCC processor is commonly used in water quality monitoring applications, as it can improve the accuracy of measurements of key water quality parameters, such as chlorophyll concentration, suspended sediment concentration, and water clarity. The C2RCC processor can be used with other tools and algorithms to support a range of water quality monitoring efforts, including monitoring the health of coastal ecosystems, assessing the impacts of human activities on water quality, and supporting the management of water resources. C2RCC is a widely used water quality processor for complex and inland waters tested in various studies (Kyryliuk & Kratzer, 2019; Niroumand-Jadidi et al., 2021; Schütt et al., 2022; Soomets et al., 2020; Soriano-González et al., 2022; Toming et al., 2017).

The atmospheric correction is performed by the C2RCC processor, provided an accurate remote sensing reflectance (R_{rs}), as assessed in previous studies over the analysed water bodies (Luciani et al., 2021). It should be noted that the publicly available version of the C2RCC works only in the built-in atmospheric correction, which could be a limitation considering the sensitivity of the inversion to the quality of atmospheric correction.

The specific algorithms used by the C2RCC processor are proprietary and are not publicly disclosed. However, the C2RCC processor is based on a combination of empirical and physical algorithms, which consider the scattering of light by the atmosphere, the presence of clouds or aerosols, and the characteristics of the water surface itself.

---

[12] https://c2rcc.org/

In comparison to other processors, the C2RCC processor is generally considered to be more accurate and reliable than empirical algorithms, which are based on statistical relationships between the reflectance values of the water body in the satellite image and the inherent optical properties of the water (Hafeez et al., 2019). The C2RCC processor for atmospheric correction has been updated: the current C2RCC is a modified version of the original Case 2 Regional processor, which has been adapted to many multispectral satellites (e.g., Sentinel-2, Sentinel-3, Landsat-8). Now, the C2RCC comprises a set of three processors (i.e., C2-Nets: C2RCC, C2X and C2X-COMPLEX) (Soriano-González et al., 2022).

The C2RCC plugin within SNAP can be setup defining the parameters presented in Table 3-3. Note that the factor coefficient and exponents for the calculating the CHL-a and TSM differ from those proposed within the EUMETSAT model (see Equations 2-6 & 2-7). The difference in the parameters for the processing rely on the studies presented in (Luciani et al., 2021) devoted to the study for the specific inland water bodies analysed. In this case, the only time dependent variable that modified for each retrieved image was the temperature acquired from a local meteorological sensor (see Section 3.2.2).

OLCI Processing Parameters			
Parameter	**Description**	**Units**	**Input**
Valid-pixel expression	The arithmetic expression defines the pixels which are valid for processing. Pixels which are not valid will be marked as no-data in the target product.	[-]	
Salinity	The value used as water salinity for the scene.	[PSU]	0.5
Temperature	The value used as water temperature for the scene.	[°C]	
Ozone	The value used as ozone if not provided by auxiliary data.	hPa	330
Air pressure	The surface air pressure at sea level if not provided by auxiliary data.	[-]	1000

OLCI Processing Parameters			
Parameter	**Description**	**Units**	**Input**
TSM factor	TSM conversion factor ($TSM = TSMfac * iop_{btot}^{TSMexp}$). **Note:** This parameter was formerly named TSMfakBpart. Also the equation to derive TSM from the IOPs has been changed in version 7.0.1 of S3TBX.	[-]	1.06
TSM exponent	TSM conversion exponent ($TSM = TSMfac * iop_{btot}^{TSMexp}$). **Note:** This parameter was formerly named TSMfakBwit. Also the equation to derive TSM from the IOPs has been changed in version 7.0.1 of S3TBX.	[-]	0.942
CHL exponent	Chlorophyll exponent ($CHL = iop_{apig}^{CHLexp} * CHLfac$).	[-]	0.65
CHL factor	Chlorophyll factor ($CHL = iop - a_{pig}^{CHLexp} * CHLfac$).	[-]	19.8
Threshold rtosa OOS	Threshold for out of scope of NN training dataset flag for gas corrected top-of-atmosphere reflectances.	[-]	0.01
Threshold AAC reflectaances OOS	Threshold for out of scope of NN training dataset flag for atmospherically corrected reflectances.	[-]	0.15
Threshold for clod flag on down transmittance	Threshold for cloud test based on downwelling transmittance @865.	[-]	0.955
Atmospheric aux data path	Path to the atmospheric auxiliary data directory. Use either this or the specified products on the I/O Parameters (ozone, air pressure) tab. If the auxiliary data is not available at this path, the data will automatically be downloaded.	[-]	
Alternative NN path	Path to an alternative set of neuronal nets. Use this to replace the standard set of neuronal nets. For MSI there exits a neural net where ranges of parameters was extended to moderate-to-extreme cases (C2X-Nets). However, this is a parameter which should only used during the development of new neural nets.	[-]	
Output AC reflectances as rrs instead of rhow	Write remote sensing reflectances instead of water leaving reflectances.	[-]	
Derive water reflectance from path radiance and transmittance	Alternative way of calculating water reflectance. Still experimental.	[-]	

OLCI Processing Parameters			
Parameter	**Description**	**Units**	**Input**
Use ECMWF aux data of source product	Use ECMWF auxiliary data (total_ozone, sea_level_pressure) from the source product.	[-]	☑
Output TOA reflectances	Add TOA reflectances to the target product.	[-]	
Output gas corrected TOSA reflectances	Add TOSA reflectances to the target product.	[-]	
Output gas corrected TOSA reflectances of auto nn	Add TOSA reflectances of the autoassociative neural net to the target product.	[-]	
Output path radiance reflectances	Add path radiance reflectances to the target product	[-]	
Output downward transmittance	Add downward transmittance to the target product	[-]	
Output upward transmittance	Add upward transmittance to the target product	[-]	
Output atmospherically corrected angular dependent reflectances	Add atmospherically corrected angular dependent reflectances to the target product	[-]	☑
Output normalized water leaving reflectances	Add normalized water leaving reflectances to the target product	[-]	☑
Output of out of scope values	Add out of scope to the target product	[-]	☑
Output of irradiance attenuation coefficients	Add irradiance attenuation coefficients to the target product	[-]	☑
Output uncertainties	Add uncertainties to the target product	[-]	☑

Table 3-3 - C2RCC processor parameters for computing the CHL-a and TSM maps. e.g. Settings for the study case of the inland water bodie– - Maggiore, Lugano and Como Lakes (*SNAP Data Processors - C2RCC OLCI Processing Parameters*, n.d.)

For each image, the C2RCC was parameterized with temperature data at the time of the sensor's acquisition, and they were adopted site-specific coefficients for the inherent optical properties (IOP) of the lakes. The processor allows to modify the default setting of some processing parameters to

better adapt to local regional characteristics. Regarding OLCI, the flags generated by the C2RCC were used to mask WQP products. The processor's neural nets detected potential cloud risk and anomalies in the water spectra, such as adjacency effects along lake borders. To achieve this goal, the C2RCC flags Cloud_risk, Rtosa_OOS, Rtosa_OOR, Rhow_OOS, and Rhow_OOR were deemed to be the most appropriate (see Table 3-3). These flags indicated that the input spectrum to the atmospheric correction neural net or the input spectrum to derive the WQPs was neither within the training range of the respective neural nets nor outside the range of the expected results.

Flag	Description
Rtosa_OOS	The input spectrum to the atmospheric correction neural net was out of the scope of the training range and the inversion is likely to be wrong
Rtosa_OOR	The input spectrum to the atmospheric correction neural net out of training range
Rhow_OOR	One of the inputs to the IOP retrieval neural net is out of training range
Rhow_OOS	The Rhow input spectrum to IOP neural net is probably not within the training range of the neural net, and the inversion is likely to be wrong
Cloud_risk	High downwelling transmission is indicating cloudy conditions

Table 3-4 - C2RCC quality flags (*Sentinel-3 Toolbox (S3TBX)*, 2014/2023)

Temperature Time Series – ARPA Lombardia

A relevant part in the production of the WQP maps relies in the acquisition of temperature data to ingest data in the C2RCC processor. In this case, the research relied in a local environmental agency for the acquisition of the temperature time series in the vicinity of the lakes of interest. The Regional Environmental Protection Agency of Lombardy for Lombardy [13], known as ARPA Lombardia

---

[13] https://www.arpalombardia.it/

("*Agenzia regionale per la protezione dell'ambiente della Lombardia*"), takes care of the prevention and protection of the environment, collaborating with local and regional institutions in different activities, from the study of atmospheric and acoustic pollution, to the intervention of superficial and underground water resources, to the monitoring of soil pollution through the use of the electromagnetic field. The data acquisition from the sensors network is possible in two ways. The first being on-demand request directly on the ARPA Lombardia platform specifying the sensor and period for retrieving the data[14]. The second benefits from ARPA Lombardia and Lombardy Region collaboration to make the meteorological network data available through the native API of Lombardy's Region data catalogue (Socrata API[15]). This option allows obtaining a specified sensor time series only for the current month of the request ("*Dati sensori meteo*" – Meteorological Sensors Data). However, the Regione Lombardia data catalogue provides a collection of CSV files including the yearly time series for all sensors (*Dati sensori meteo 2020-2022* – Meteorological Sensors Data 2020-2022). For this work, the second method was implemented to ease the data retrieval.

IdSensore	Tipologia	Unità DiMisura	NomeStazione	IdStazione	Provincia	DataStart	UTM_Nord	UTM_Est
14606	Temperatura	°C	Tremezzo	1380	CO	18/04/2013	5092705	516580
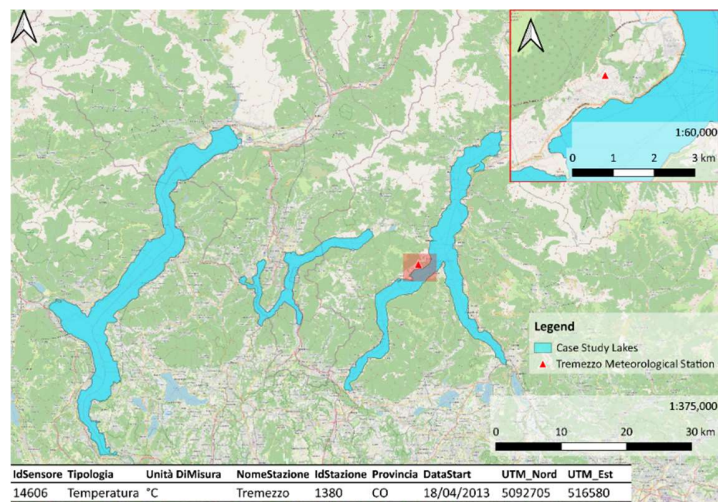
Figure 3-4 – ARPA Lombardia Tremezzo Meteorological station

---

[14] https://www.arpalombardia.it/Pages/Meteorologia/Richiesta-dati-misurati.aspx#
[15] https://dev.socrata.com/foundry/www.dati.lombardia.it/647i-nhxk

Figure 3-4 indicates the location of the reference station and specification for the selected sensor. The station is situated in Tremezzo, Como Province, and is active since April 18th of 2013. As expected, as seen in Figure 3-5 (a), the temperature behaviour in the area follows a seasonal pattern with a lower peak during the winter season (December to February), and a higher peak during the summer season (June to August). In general, over the years the temperature trends remain constant. Nevertheless, it is observed that for 2022 there is an increase in the average temperatures between May and August (Figure 3-5 (b)).

(a)
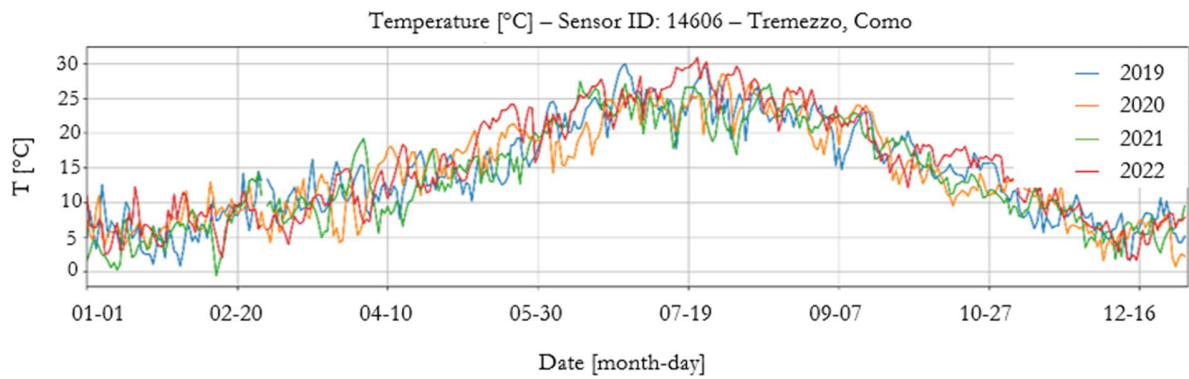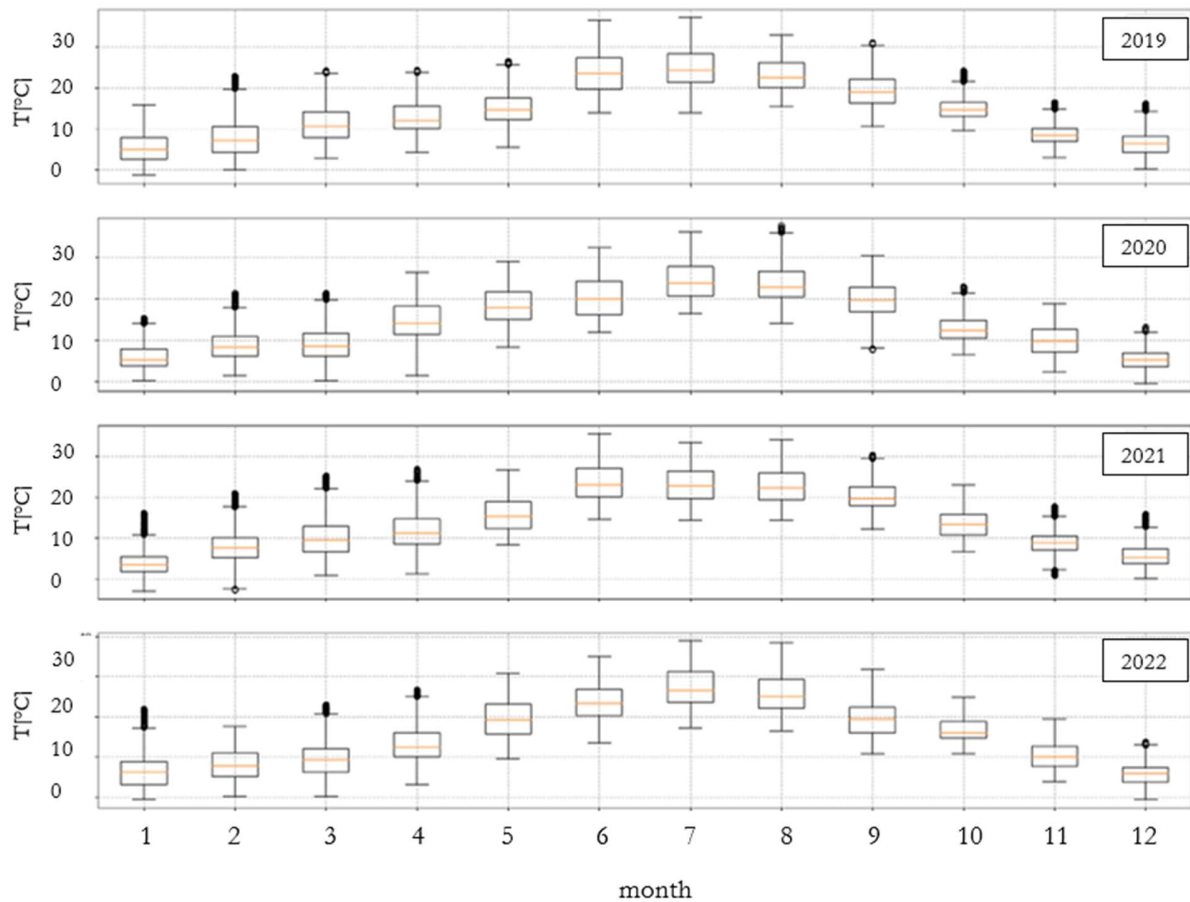
Figure 3-5 - Temperature time series (2019-2022). Tremezzo, Como, Italy (see Figure 3-4)

3.2.3 Barsi Method

Barsi et al. (2005) suggest that it is possible to derive the Lake Surface Water Temperature (LSWT) temperature by inverting the Planck Law with some approximation. To invert the Planck Law, it is necessary to remove the presence of atmospheric gases from the measured radiance at the top of the atmosphere (L_{TOA}), as these effects can increase the emitted radiance in the thermic band of a surface. The method has been implemented in the processing chain applied to Landsat-8 TIRS-1 band to produce the LWST maps.

The atmospheric correction parameters must be defined with respect to the acquisition date. The parameters are the following: specific data on the coefficient of atmospheric radiance (τ), and the

ascending and descending solar radiance (L_u and L_d, respectively). Then, it is computed the radiance at the bottom of the atmosphere (L_{BOA}, i.e. corrected radiances from atmospheric disturbances).

$$L_{BOA} = \frac{L_{TOA} - \tau(1 - \varepsilon)L_d - L_u}{\tau\varepsilon}$$

[Eq. 3-1]

In the expression, the surface's thermal emissivity coefficient ε is assumed constant and equal to 0.98.
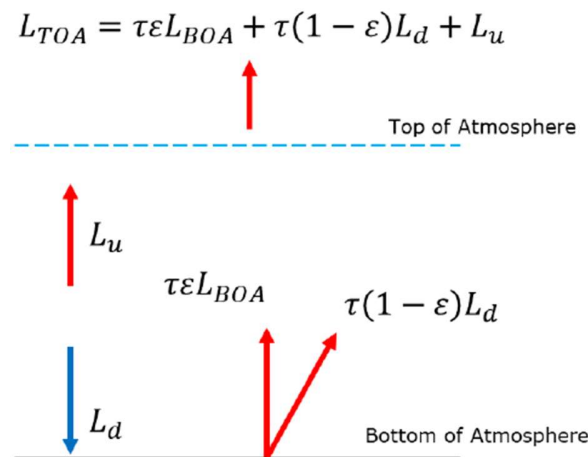
Figure 3-6 - Radiative transfer equation

From the L_{BOA}, it is obtained the lake surface water temperature by inverting the Planck's equation, with some approximations specifical for the Landsat 8 products.

$$T = \frac{k_2}{\ln\left(\frac{k_1}{L_{BOA}} + 1\right)}$$

[Eq. 3-2]

Where T is the LSWT, k_1, k_2 are thermic constants for Landsat 8 ($k_1 = 1321.08, k_2 = 774.89$).The atmospheric correction parameters are retrieved through the web-based application presented by Barsi et al. (2005) (see Figure 3-6). The application employs local interpolation models of atmospheric profiles.

(a)

(b)

Figure 3-7 - Atmospheric correction parameters calculator[16]

---

[16] https://atmcorr.gsfc.nasa.gov/

3.3 SENTINEL-3 A/B GEOLOCATION – IMAGE CO-REGISTRATION

One of the challenges for analysing the WQPs involved the relative shift between the satellite images retrieved from the Sentinel 3 – OLCI sensor.

> "The georeferencing accuracy of OLCI-A and OLCI-B is validated using the visible data from channel Oa17 correlated with ground control points. Current georeferencing shows a global accuracy below about 0.3 pixels (90m) RMS".[17]

Considering that the analysis relies on the temporal monitoring of the different WQPs, it is important to have a regular grid to analyse the evolution of the different parameters. For this reason, the need for implementing a co-registration algorithm for the alignment of the WQPs maps was considered. Co-registration in image processing is a procedure that minimises the shift between the images at a pixel scale. Leprince et al. (2007) mentions that the reason why it is fundamental for remote sensing applications such as environmental mapping, change detection, mosaicking, or image fusion is that by minimising the displacement between image datasets, the co-registration process avoids errors at further analysis.

---

[17] https://sentinel.esa.int/documents/247904/4598066/Sentinel-3-OLCI-Land-Handbook.
pdf
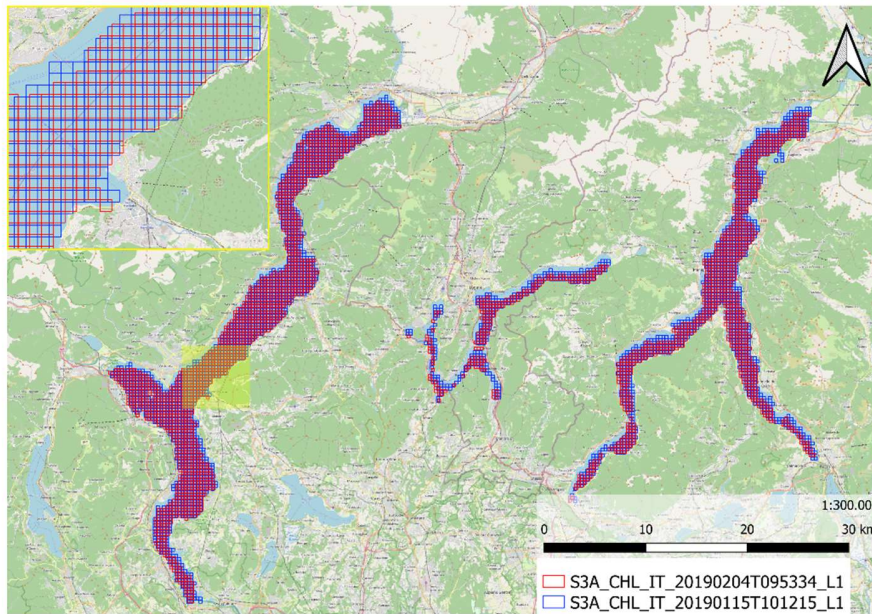
Figure 3-8 – Example Chlorophyll-a map relative shift.
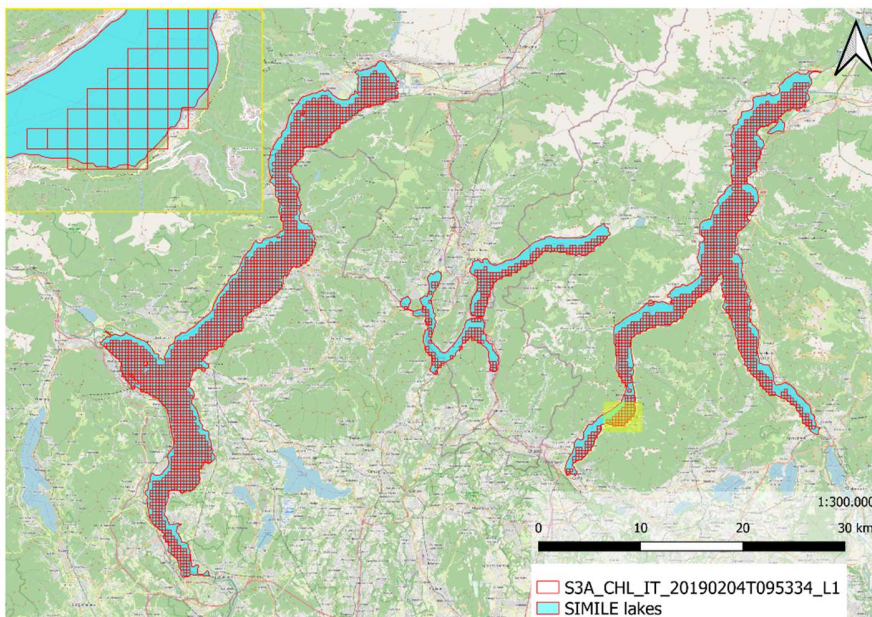
Figure 3-9 - Example Chlorophyll-a out-of-lake estimates due to

The co-registration process aims at fixing deformations and distortions with respect to a reference image. In general, a reference image is a georeferenced image with ground control points. Within the co-registration process, the images are projected and resampled onto the same reference system. The

procedure by taking both "projected" and "resampled" images, computes the measurement of displacement in the plain and corrects the displacements in the image.

The displacements on the images get calculated by using image registration techniques, which could be classified into two types: intensity-based and feature-based.

- Intensity-based: Similarity of the pixel values that appears on the target and reference image.

- Feature-based: Detect the position of a ground object distinct in both target and reference images.

The image co-registration procedure performs the correction of the target image by warping it (i.e. applying a mathematical transformation based on displacement values). Most methods of remote sensing image co-registration are based either on geocoding, or on nonrigid image registration methods that use only the images as input (Brigot et al., 2016). In the case of geocoding, the accuracy of co-registration will be highly dependent on the availability and precision of both a digital terrain model and the orbit parameters.

On the other hand, nonrigid image registration without geocoding is widely investigated in various fields beyond the scope of remote sensing, for example, computer vision and medical imaging. In computer vision, video image coregistration must meet constraints of robustness and speed of the execution but often focuses on images taken from the same sensor with little delay in time. In medical imaging or remote sensing, difficulty lies in the different nature of the images to compare. Moreover, the context of remote sensing is also changing today with larger quantities of time series data, and some time-sensitive applications require fast processing. This is the case for example for near real-time change detection for rapid post-disaster assessment (Guida et al., 2018; Stumpf et al., 2018), wildlife tracking (Hyun et al., 2020), and surveillance across broad areas(Kristollari & Karathanassi, 2022).

Most nonrigid registration methods are parametric methods, meaning that an assumption is made about a parametrized model that constrains the form of the expected deformations between processed images. Then, a similarity function is optimized to find an approximation of a real underlying deformation. Among them, feature-based approaches establish a correspondence between a number of especially distinct points in images.

The choice of an image similarity measure is a key point. One of the most widespread used for the registration of multimodality images is mutual information. Already used in remote sensing image co-registration, the main drawback of the similarity measure is that it is quite time-consuming.

Another family of nonrigid co-registration methods is nonparametric. Among them, dense methods compute a displacement for every pixel in the image. They are particularly interested in the case of very local deformation due for example to terrain elevation that has a lot of influence on high-resolution images. Most of dense and nonparametric methods belong to optical flow estimation. Optical flow is the pattern of apparent motion of objects in a visual scene caused by the relative motion between the sensor and the scene. Optical flow methods have been developed in a context where the constraints of speed and robustness to environmental effects have led to intensive efforts in producing algorithms that combine robustness, precision, and high computing speed.

The focus of this work presents the implementation of an optical flow algorithm, GeFolki (Geoscience Extended Flow Lukas—Kanade Iterative, (Brigot et al., 2016), that computes the displacement between two images. Flow algorithms classical methods benefit from partial derivatives over the image intensity concerning the spatial and temporal coordinates. These algorithms rely on two points:

1. An observation model where intensity remains the same from one image to another (i.e. brightness constancy constraint).

2. Linearization step for solving the system, which enables estimating the intensity variation as a local gradient.

GeFolki

Considering the registration of two images I1 and I2, defined on a 2-D support S ∈ R2. The dense optical flow in computer vision is the pattern of the apparent motion, that corresponds to the displacements found between two images. The displacements are defined as u : x→u(x) ∈ R2.

The initial assumption in most optical flow models is the assumption that the pixel intensity remains constant from one frame to another.

$$I_1(\mathbf{x}) = I_2(\mathbf{x} + \mathbf{u}(\mathbf{x}))$$

[Eq. 3-3]

[Eq. 3-3], is known as the brightness constancy. The equation accounts for the intensity variations between the images by introducing the displacement variable in the parameters for the second frame.

The numerical solution for [Eq. 3-3], needs a linearizing step. And the change of intensity is commonly rewritten as:

$$I_1(\mathbf{x} + \mathbf{u}(\mathbf{x})) = I_2(\mathbf{x}) + \nabla I_2 \cdot \mathbf{u}$$

[Eq. 3-4]

[Eq. 3-4] is called the linearized optical flow constraint. The constraint assumes small displacements.

However, in remote sensing, the assumptions of frame proximity and controlled acquisition conditions cannot be guaranteed. Then, unfortunately, classical algorithms are not suitable to model or compensate for these effects. The system presented in equations 3.2 and 3.3 is undetermined. For this reason, optical-flow methods introduce spatial conditions on the distribution of the flow, local or global, assuming the spatial smoothness in the flow field. GeFolki implements a local window-based optical-flow method based on the work presented by Lucas and Kanade(1981).

Lucas—Kanade algorithms are divided into two different approaches, classical and modern. The first group assumes that the flow is constant around the neighbourhood, solving the optical flow equations by the least squares criterion using a Gauss—Newton strategy with a first-order Taylor expansion. The second group, is iterative and multiresolution, benefitting of pyramids to compute the optical flow at different scales (coarse-to-fine strategy); the multiresolution strategy allows estimating large displacement besides the local constraint. GeFolki, belongs to a Lucas—Kanade the method mentioned last.

GeFolki (Terradue, 2019/2022) requires two input images resampled at the same rate and recreates the components of the flow at every pixel position.

$$J(\mathbf{u}; \mathbf{x}) = \sum_{\mathbf{x}' \in S} \omega(\mathbf{x}' - \mathbf{x}) \left(f_1\big(I_1(\mathbf{x}')\big) - f_2\left(I_2(\mathbf{x}' + \mathbf{u}(\mathbf{x}))\right) \right)^2 \qquad \text{[Eq. 3-5]}$$

Where,

- ω defines a local window of radius r and size $(2r + 1) \times (2r + 1)$, thus $\omega(x) = 1 \Leftrightarrow |x|\infty \le r$, otherwise 0.

- f_1 is a function applied to the reference (master) image I_1, and f_2 is a function designed to project the images in a space where $f_1(I_1)$ and $f_2(I_2)$ are similar enough to validate the brightness constancy model.

- The first step is creating a pyramidal structure of down-sampled images to find large displacements. L affects the amplitude of the desired movements.

It is possible defining the pyramid levels L to determine the amplitude of the movements for the window of analysis. Then, the algorithm proceeds to solve the minimisation presented in [Eq. 3-5] iteratively for each scale, refining the solutions from coarser to finer.

The algorithm benefits from the multiscale to initialise the flow from coarser scales. Nevertheless, it must be highlighted that large down-sampling can remove relevant information useful for finding matches.

In case of heterogeneous images co-registration, it is possible defining a different rank filter to make the images more "similar". A function of local contrast inversion is defined for all cases where the contrast between different sensors is not the same. This is the case for the co-registration of images acquired in different fields of the electromagnetic spectrum.

$$f_1 = R \diamond g \qquad \text{[Eq. 3-6]}$$

$$f_2 = C \diamond R \diamond g \qquad \text{[Eq. 3-7]}$$

Where,

- g being the rolling guidance filter (presented in Zhang et al. [56]), which concerns about the texture and resolution of the images. The filter removes small scal features using a Gaussian filter focusing on the retrieval of edge information;
- R the rank filter;
- g the local inversion decision function.

The Rank function $R(I)$ consists in replacing the intensity value of $I(x)$, and is applied to the image I (thus, the rank transform depends on the relative order of the local neighbourhood), is expressed as:

$$R\big(I(\mathrm{x})\big) = \#\mathrm{x}' \in S_R(\mathrm{x}) \ with \ |I(\mathrm{x})| > |I(\mathrm{x}')| \qquad \text{[Eq. 3-8]}$$

with S_R being a neighbour of pixel x.

The Local Contrast Inversion decision function, C, addresses the contrast differences between images. To determine the need of inverting the value of a pixel x_0 the algorithm follows the criteria presented next.

$$C_1(\mathrm{x}_0) = \sum_{x_i \in W(\mathbf{x}_0)} \left| I_1\left(\mathrm{x}_i - I_2\left(\mathrm{x}_i + u(\mathrm{x}_i)\right)\right)\right|$$

$$C_2(\mathrm{x}_0) = \sum_{x_i \in W(\mathbf{x}_0)} \left| 1 - I_1\left(\mathrm{x}_i - I_2\left(\mathrm{x}_i + u(\mathrm{x}_i)\right)\right)\right|$$

If $C_1(\mathbf{x}_0) < C2(\mathbf{x}_0)$ it is not required the inversion of the target image, otherwise for image 2 $I_2(\mathbf{x}_0 + \mathbf{u}(\mathbf{x}_0))$ by $1 - I_2(\mathbf{x}_0 + \mathbf{u}(\mathbf{x}_0))$.

There are multiple reasons behind the need for implementing the co-registration algorithm for the WQP maps. First, the production of maps under the same grid limits further errors on the cross-temporal analysis of the WQPs. In addition, the co-registration of the WQPs maps helps avoiding the removal of pixels which are not located within the lakes delineated by the vector used during the statistical analysis (which, otherwise, could be cropped).

The co-registration algorithm has been applied to the complete set of WQPs maps. For the selection of the reference image, it was performed a visual inspection from the set of products to determine the image which overlapped the lakes shape in the "best" way. Once identified the reference image, the transformation W obtained by the method would be applied to the corresponding WQPs maps obtained from the target image. The transformation benefits from the band Oa17 from the original products, which is employed for the initial geolocation of the satellite images.
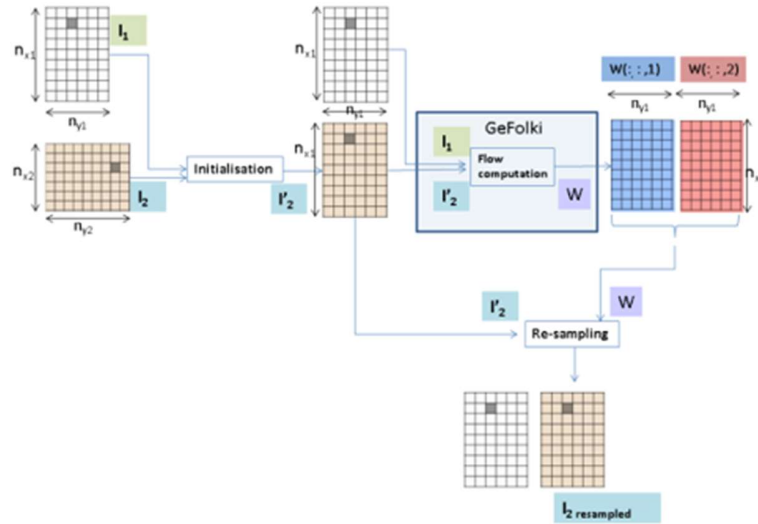
Figure 3-10 exemplifies the workflow for the execution of the co-registration algorithm. The algorithm decomposes into three steps:

1. Initialization step. Input the reference (master) and target(slave) images. Where, the target image is assigned the same dimensions as the reference image (n_{x1} and n_{y1}).

2. Flow calculation step (GeFolki). With two images of equal dimensions, the output W is a matrix of size $(n_{x1} \times n_{y1} \times 2)$. With $W(:,:,1)$ being the flow component on the horizontal axis and W(:, :,2) the flow component on the vertical axis.

3. Resampling step. In this step, the resampling of the slave image is performed using the flow matrix. The output matrix is denoted as $I_{2resample}$ that is the co-registered image to the master image.

4. In this case, the processing chain benefits from the computed flow matrix for the coregistration of the WQPs maps. Notice from Figure 3-11 the misalignment of the products on two different dates.
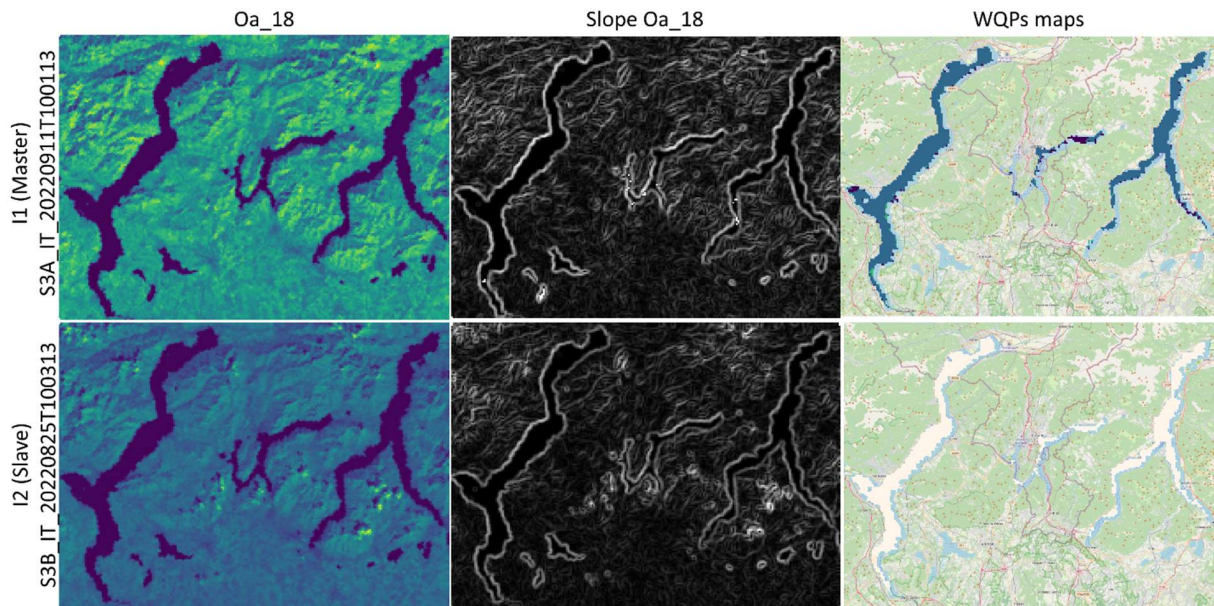
Figure 3-11 - Application of the image coregistration transformation to the WQPs maps

3.4 OUTLIERS DETECTION/REJECTION

The comparison with the in-situ measures provided by the project partners and the inspection of the WQP maps' statistics showed the need for filtering data to be marked as outliers and excluded from the maps. However, in the case of WQPs we are referring to values which could pertain to specific local behaviours, which implies that the filtering must avoid possible reasonable anomalies due to physical reasons (for example, negative estimates or close to zero estimates for the LSWT maps, or anomalies in the expected R_{rs} values in the Sentinel 3 products). Thus, the time series of in-situ data have been considered, and the limnologists who study the waters of the lakes under study have been interviewed, in order to select the plausible ranges of values.

The outliers' detection in WQP maps considered three statistical methods within the processing chain. It is worth mentioning that the detection of outliers, while effective in some cases related to mixed pixels, still requires some expertise in understanding the WQP maps for identifying the abnormal

values within the estimates. For this reason, the statistical analysis in the processing chain is referred to as a semi-automated method, where an experienced researcher should get insights from the identified outliers. The processing chain considered three different techniques concerning the statistical dispersion of the data to identify outliers presented in Table 3-5.

Method

Interquartile Range Test (IQR): Estimate the outliers using the 1st and 3rd quartile from the median of a data set.

$$IQR = Q3 - Q1$$

$$LB = Q1 - 1.5 \cdot IQR$$

$$UP = Q3 + 1.5 \cdot IQR$$
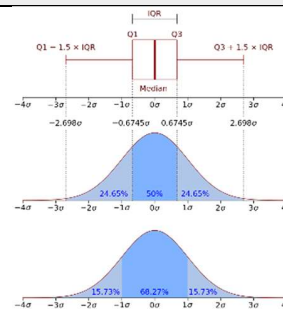
Figure 3-12 - Boxplot (with an interquartile range) and a probability density function (pdf) of a Normal N(0,σ2) Population

Empirical rule: 2σ and 3σ. For an approximately normal data set, the distribution must be within the percentages presented in Figure 3-13.

$$LB = \mu - k\sigma$$

$$UB = \mu + k\sigma$$

Where k is the number of deviations from the mean.
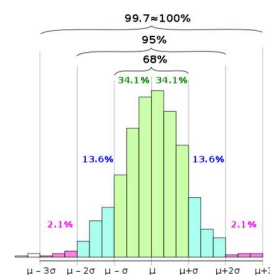
Figure 3-13 - Empirical rule, expected distribution from the mean for σ, 2σ and 3σ

Table 3-5 - Outliers' detection methods implemented in the WQP processing chain.

Notice that two of the presented methods rely on the assumption of normality of the distribution of the analyzed variables, where it is required to devote attention to the percentage of anomalous values

indicated by the tests. To check if the values are suitable for applying the outlier rejection filter, the Chebyshev inequality was revised to avoid over-extracting feature values from the WQPs maps.

In probability theory, Chebyshev's inequality (also called Chebyshev Theorem), for a broad class of probability distributions, guarantees that a certain fraction of the values can be no more than a certain distance from the mean. Specifically, the values cannot be further than $1/k^2$ of the distribution's values can be k or more standard deviations away from the mean.

$$\Pr(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$
<div align="right">[Eq. 3-11]</div>

k	Min. % within k standard deviations of mean	Max. % beyond k standard deviations from mean
1	0.00%	100.00%
√2	50.00%	50.00%
1.5	55.56%	44.44%
2	75.00%	25.00%
2√2	87.50%	12.50%
3	88.89%	11.11%
4	93.75%	6.25%
5	96.00%	4.00%
6	97.22%	2.78%
7	97.96%	2.04%
8	98.44%	1.56%
9	98.77%	1.23%
10	99.00%	1.00%

Table 3-6 – Chebyshev inequality outlier detection based on the number of standard deviations

Following Chebyshev's theorem, it was observed that the 3σ filter satisfied the inequality for the maps produced within the workflow. In addition, it has been noted that the filtering method did not remove some of the estimates which were considered as outliers by the IQR method and were not abnormal values. In Figure 4-7, it is observed that, in some cases, the IQR removed a significant set of pixels from the maps, which after further inspection did not display anomalies to consider them outliers. While, the IQR method can be useful when having skewed distributions, as it is less affected from

extreme values, in cases such as the one of Figure 4-7 a more uniform distribution of the values in the dataset can precede the extraction of many pixels.
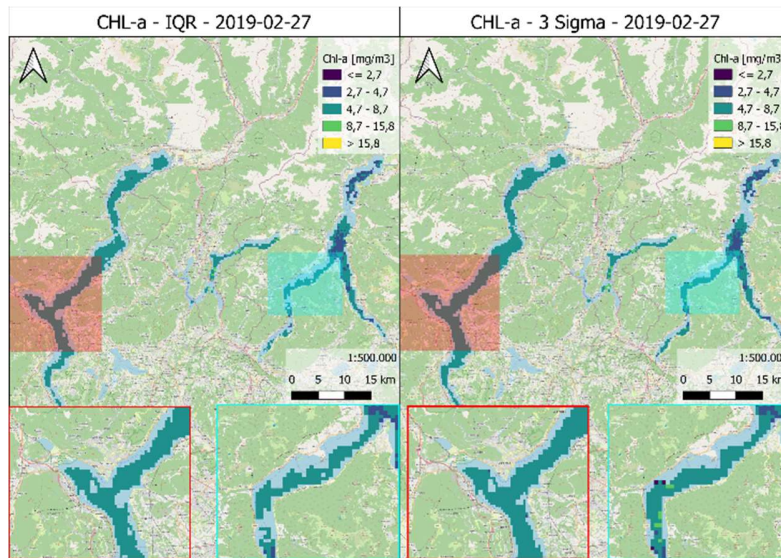
(a)
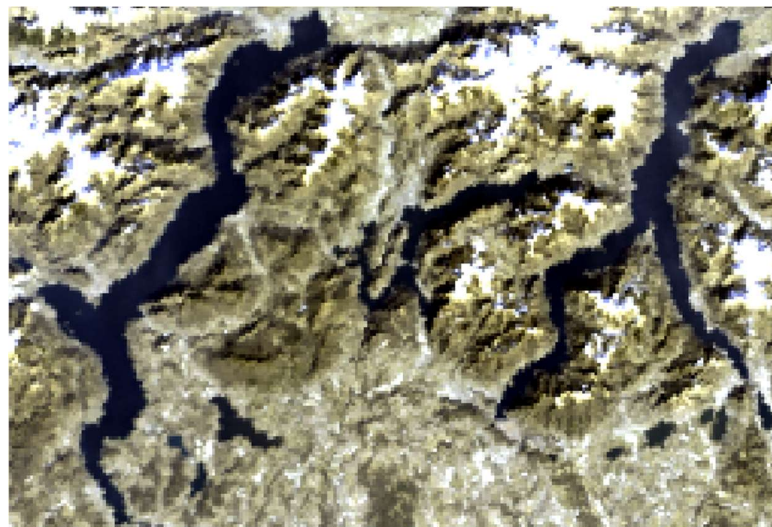
(b)

Figure 3-14 – IQR vs 3-Sigma outlier detection/rejection methods comparison. (a) Comparison IQR (Left) vs. 3-Sigma (Right) for the Chl-a map for 27th of February 2019. (b) S3A-OLCI real color image over the area of interest on the 27th of February 2019.

Here, it is presented the case of the Chl-a maps outliers detection. Figure 4-3 to Figure 4-5 provide an overview of the valid pixels within the Chl-a maps accounted by the different methods. The boxplots display distribution of the valid pixels present in the maps, on a monthly basis, for each of the methods.
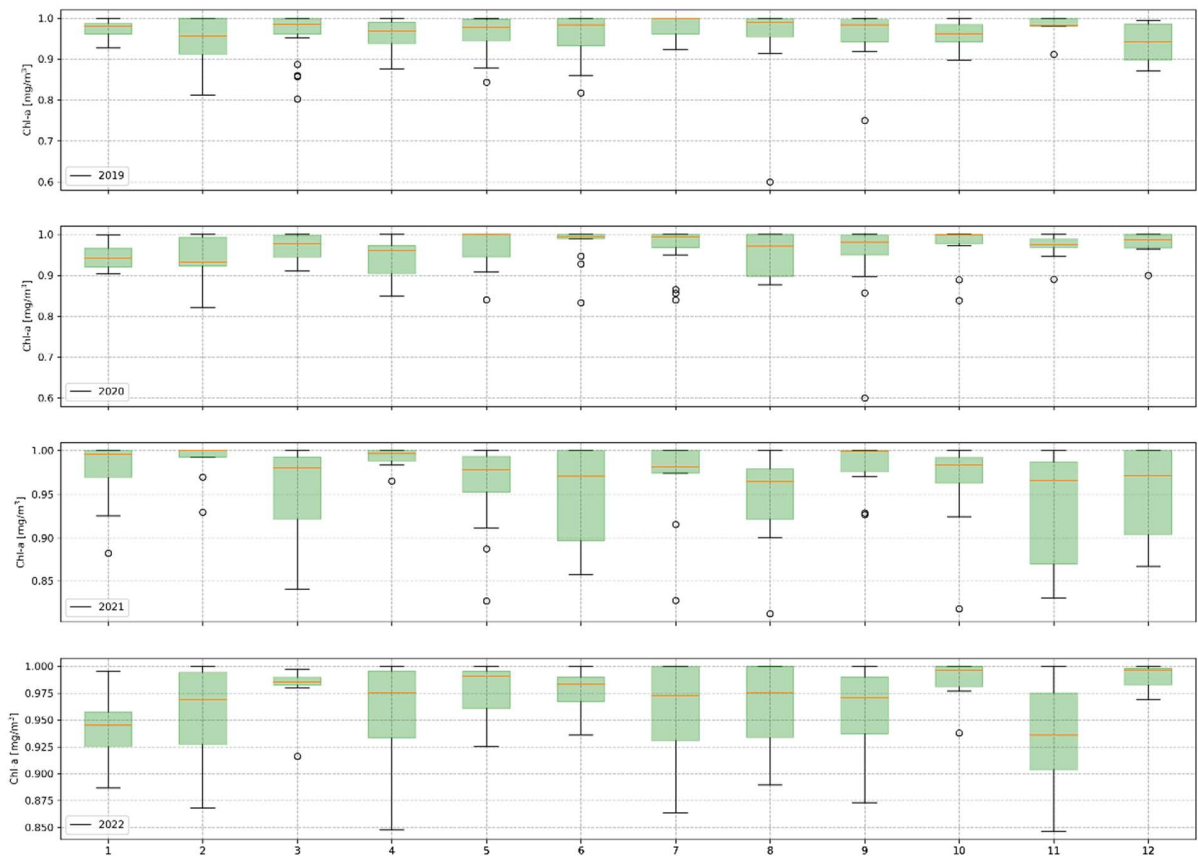
Figure 3-15 - IQR Chl-a valid estimates boxplot per month

First, it is possible to identify for all filtering methods, that the detection of outliers was low for the first two years of the analysis, and that for the first and last months of the maps Chl-a maps the estimates variability led to the extraction of many pixels (pattern which maintained across the different methods). Concerning the k-Sigma methods, as expected, the number of outliers is larger for the 2-Sigma filter when compared to the 3-Sigma. On the other hand, the assessment with respect to the IQR filter is more complex as the amount of identified outliers rely in the distribution of the estimates in each of the maps. For this reason, there are cases in which the IQR will identify more or less outliers compared to the k-Sigma filtering. Then, the 3σ methods was selected to proceed with the analysis of the WQPs maps. However, it is relevant to remark that the filtering alone is not an absolute tool for determining the presence of outliers within the maps. The filtering method removes estimates which

55

are extremely large with respect to the trend inside of the maps, as the spatial distribution of the parameters is not expected to vary greatly.
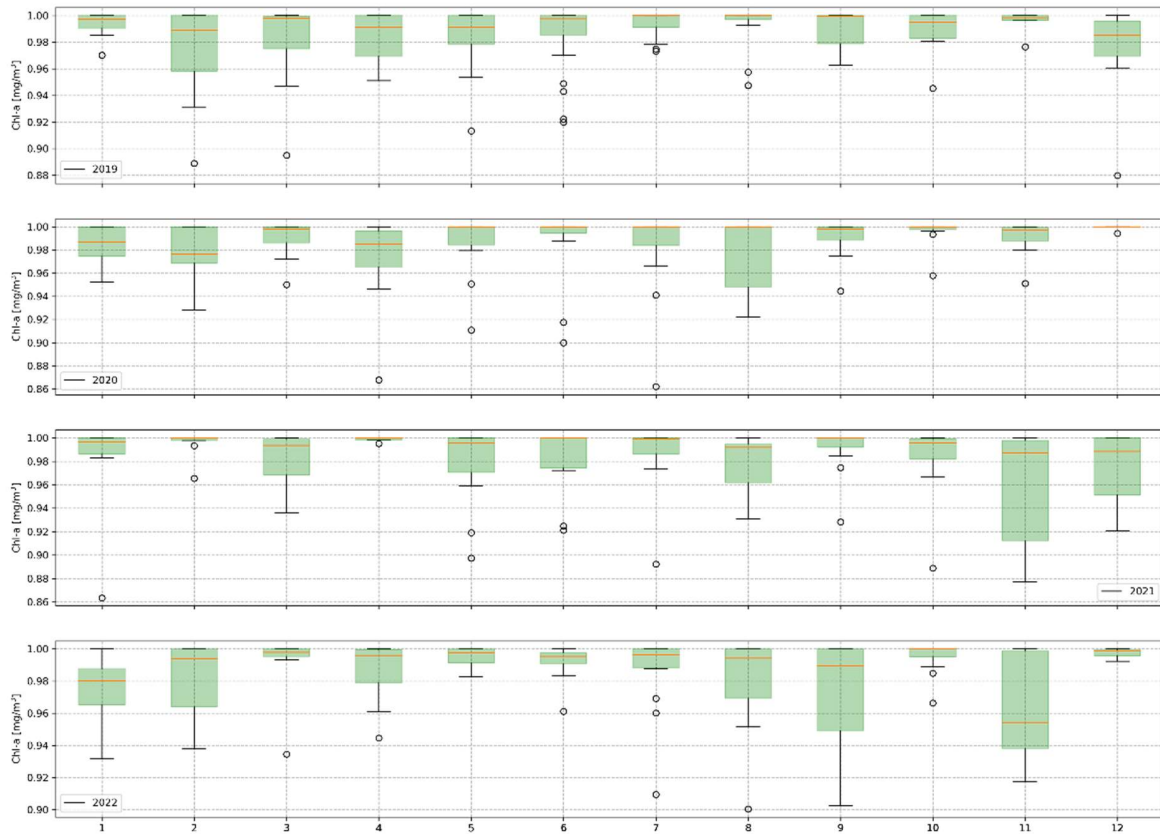
Figure 3-16 - 2-Sigma Chl-a valid estimates boxplot per month

To perform the outlier rejection, a 3σ filtering has been applied in order to single out data which showed a different behaviour from the one of the lake population. Then, the out-of-range values detected with the 3σ filtering were explored on the map to interpret the reason for the anomaly in terms of geographical location (e.g., lower surface temperature values could be reasonable when detected in an area affected by inflowing waters, higher surface temperature due to shallow water, etc.).
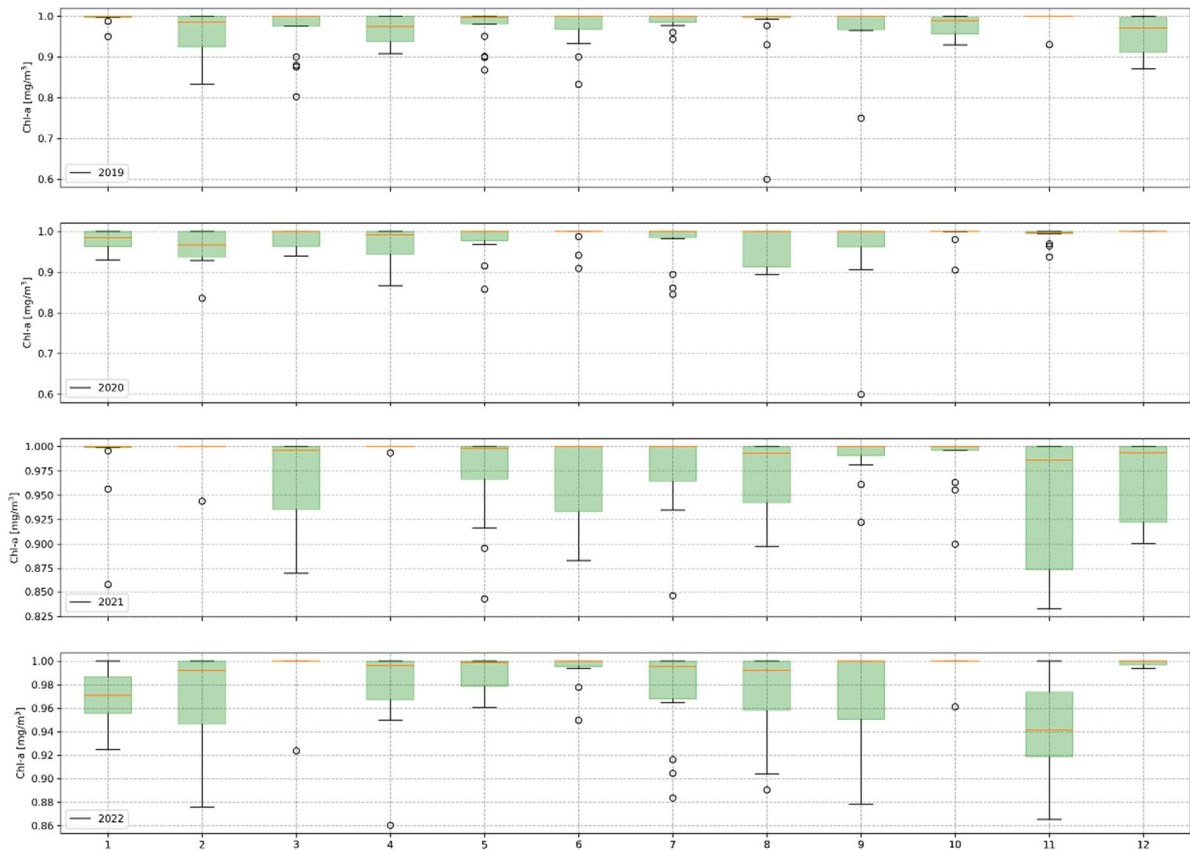
Figure 3-17 - 3-Sigma Chl-a valid estimates boxplot per month

Chapter - 4. WQP MONITORING – SIMILE STUDY CASE

In this chapter, we present the outcome of the implemented processing chain used to obtain the WQPs maps. Throughout the processing of the satellite images some of the main challenges concerning the expected values for the analysed water of the WQP, and the presence of anomalous values and the relative shift in between the WQP maps. It is possible tackling the limited information concerning the data availability of the water quality parameters through the regular annual reports of an environmental agency for the protection of the water bodies in the region of interest. The available reports provide in-situ measurements of chlorophyll-a for the complete set of lakes. Then, the scope is introducing the methods implemented to tackling these challenges. On the one hand, by presenting a statistical analysis over the WQPs estimates descriptive statistics to identify values which may be regarded as abnormal values. The statistical inference methods used for this analysis considered the Interquantile Range, and two and three sigma methods. On the other hand, the relative shift between images has been addressed through implementing an optical flow co-registration algorithm known as GeFolki. In a subsequent step, each image has been analysed over global descriptive statistics for assessing different trends over the WQP. The chapter presents the tools, both desktop software and data acquisition APIs, that were utilised for generating the maps. Lastly, we propose potential future directions for analysing these WQP maps to monitor water resources.

4.1 STUDY AREA AND DATASETS

The research has been developed under the framework of project SIMILE (Brovelli et al., 2019). The focus of the project is the conservation of the subalpine lake in the cross-border area of Italy and Switzerland. More specifically, the study addresses lakes Maggiore, Lugano and Como. For this reason, this introductory section will focus in providing an overview of the trends of the analysed water bodies

WQP's as a reference for the outputs of the maps processing chain. In this case, the entities working on behalf for the conservation of the lakes of interest the Italy-Switzerland International Commission for the Water Protection (CIPAIS; Commissione Internazionale per la Protezione delle Acque Italo-Svizzere) and ARPA Lombardia. CIPAIS investigates problems with pollution and other anomalies in the Lakes Maggiore and Lugano waters. The Commission is made up of numerous delegations that serve as the central administrative representatives for the cantons of Ticino, Vallese, and Grigioni in Switzerland, as well as the regions of Lombardy and Piedmont in Italy. The study aims to offer recommendations to local governments for managing their water resources. ARPA Lombardia on the other hand delivers reports and data for the Lake of Como as part of the available network of stations for monitoring the lakes. Here, we characterise the lakes morphology and target WQP, indicating the reference values for the analysis considering that these are in-situ measurements which are more reliable. However, it must be noted that the measurements are made in a column depth from surface to 20m below surface which makes them also limiting for making a direct comparison between the WQP maps which are estimated from the water surface leaving reflectance.
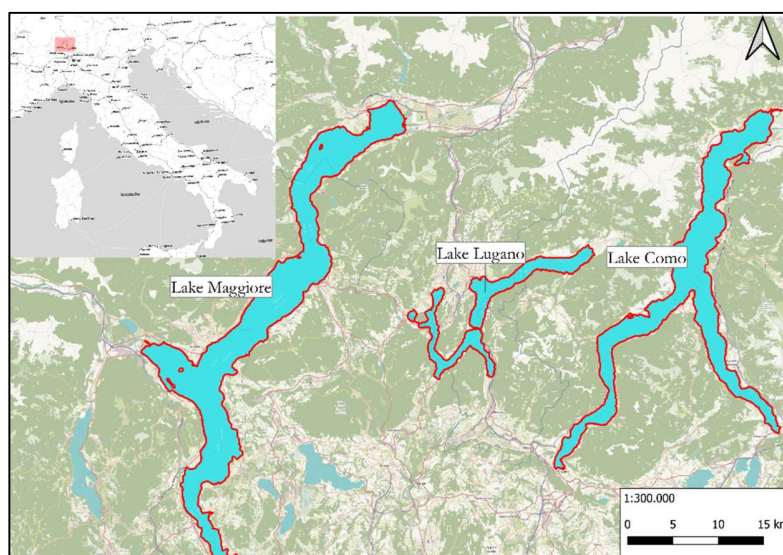
Figure 4-1 - Study case lakes of interest. Insubric Lakes (Italy-Swiss cross-border area)

Lake Maggiore

When it comes to surface area and volume, Lake Maggiore is the second-largest Italian basin (212km^2 and 37.5km^3). The lake, one of the biggest in the subalpine region, was created by two glaciers that carved their way out of the Alps through the valleys of the rivers Ticino and Toce.

	Units	Italian Basin	Swiss Basin	Maggiore Basin
Basin Area	[km2]	3370	3229	6599
Lake Area	[km2]	42.6	169.4	212
Volume	[km3]	7.5	30	37.5
Mean Depth	[m]	-	-	176
Max Depth	[m]	-	-	370
Theoretical Average Time of Change	[years]	1.6	2.4	2.4

Table 4-1 – Lake Maggiore basin characteristics

At an elevation of 1,283 m.a.s.l., the lake's catchment area extends over 6,599 km^2. The catchment is administratively located in Switzerland and the regions of Piedmont and Lombardy in Italy. The waters of Lugano enter the Maggiore Lake through the river Tresa in Luino, which is part of the Verbano basin catchment. The Ticino and Toce rivers serve as lake Maggiore's main tributaries in the hydrographic basin, while the Ticino River serves as its primary emissary. There are numerous uses for Lake Maggiore that benefit the local economy and citizens. For industrial purposes, for example, the outgoing rivers are especially important. hydroelectric power plants and extensive Po valley irrigation. The Miorina dam regulates the lake's main inlet in the Golasecca area of Sesto Calende (VA).

In 2019, the mean value of chlorophyll-a was 3.95µg/l, which aligns with the decrease reported in the previous year with respect to the value of 2017 of 4.8µg/l. The maximum reported value was registered in April (9.4µg/l), in correspondence to the algae bloom of Diatoms (March presented values of 9.2µg/l).

Analogously to the patterns in previous years, the increase in cyanobacteria was focused on the late summer period (starting in August). In 2018, both, the mean yearly value of the biovolume and the concentration of the mean chlorophyll-a decreased, with values of 0.75cm3/3 and 3.6mg/m3.

In 2020, as in the preceding years, is highly dominated by Diatoms species. The peaks in the biomass are prevalently due to Diatoms species between the periods of March-April and June-July. The mean value for chlorophyll-a 3.9μg/L, is similar to the values reported in 2019, and so displays the decreasing trend from 2017. The maximum value of 8.1μg/L on the 13th of July. Again, the cyanobacteria population increased in the late period of the summer season from the start of August.

In 2021, the development of the algal population of Diatoms, while dominant, presented a different behaviour from previous years. The population of Diatoms appeared in January with low biomass, yet it anticipates their regular pattern, probably due to the favourable meteorological conditions, with atmospheric temperatures higher than normal. The first months of the year (from January to April) were characterized by a lower biovolume. It has been hypothesized that the diminishment in biomass, and subsequent slowdown in algal growth, are a consequence of biotic and abiotic factors.

Lake Lugano

Lake Lugano locates in the cross-border area between Switzerland (Canton Ticino) and Italy (Regione Lombardia) 271m.a.s.l. The lake locates in a valley, originated from fluvial erosion in the tertiary period (Messian), shaped by the glaciers of the Adda and Ticino during the last glacial period (Pleistocene). Lake Lugano divides into three different basins: the northern basin (between Melide and Porlezza), the southern basin (between Capolago and Agno) separated by the Melide dam-bridge built in the past on a sub-lake moraine and the small basin of Ponte Tresa located near river Tresa. The three basins have different morphological and hydrological characteristics. The northern basin is very deep (288m) with a catchment basin of 270km2 and a volume of 4.68km3. The eutrophication process of the lake

began in the 1950s due to the disappearance of oxygen in the deep waters in the northern basin and the increase in the saline density. The stratified waters during this period led to a significant increase in the permanence of waters over the 100m depth.

	Units	North Basin	South Basin	Ponte Tresa
Basin Catchment Area	[km2]	269.7	290.3	5.6
Lake Basin Area	[km2]	27.5	20.3	1.1
Basin Mountain Area	[km2]	-	297.2	607.8
Total Basin Catchment Area	[km2]	297.2	607.8	614.5
Lake Basin Volume	[km3]	4.69	1.14	0.03
Theoretical Lake Retention Time	[Years]	12.3	1.4	0.04
Mean Depth	[m]	171	55	33
Maximum Depth	[m]	288	95	50

Table 4-2 – Lake Lugano basin characteristics

In 2019, the yearly mean values for chlorophyll-a concentration in the productive stratum (0-20m) was 6.4mg/m3 in the north basin and 6.7mg/m3 in the south basin. Concerning 2018, where the values approached the objective, the chlorophyll-a concentration grew, as observed for the phytoplankton biomass.

In the south basin, the Spring peaks constitute the annual maximum (Melide: 13.6mg/g3; Figino: 12.2mg/m3), coinciding with the initial growth phase of Diatoms. In the north basin, regardless of the peak reached in February of Diatoms, the maximum levels of chlorophyll were achieved in October due to the Cyanophyceae (10.7mg/m3). The values reported in Summer (June-August), commonly treated as an indicator of the trophic state, followed the past year's trend (Gandria: 5.5mg/m3, Melide: 6.6mg/m3, Figino: 6.4mg/m3).

The spatio-temporal distribution of chlorophyll-a evidenced the beginning of the vegetative phase (February-April) due to the Diatoms and the Summer growth of phytoplankton, initially higher in the superficial stratum and then concentrated in the lower stratum where other cyanobacteria develop.

Lake Como

Lake Como is a natural-regulated lake located in the northern part of the Lombardy Region. It is the deepest, with a maximum depth of 425m, and third largest Italian lake with a surface area of 145.5km^2.

	Units	Lake Como Basin
Basin Catchment Area	[km2]	4524
Lake Basin Area	[km2]	145.5
Lake Basin Volume	[km3]	22,500
Maximum Depth	[m]	425

Figure 4-2 – Lake Como basins characteristics (Copetti et al., 2020)

Lake Como encompasses three main sub-basins: northern, south-western, and south-eastern. The northern basin. The northern basin is fed by two tributaries, Adda and Mera Rivers (with a mean annual flow of 88m^3/s and 22 m^3/s).

4.2 WQP DESCRIPTIVE STATISTICS TIME SERIES

Here, it is presented the case of the Chl-a maps outliers' detection. Figure 4-3 to Figure 4-5 provide an overview of the valid pixels within the Chl-a maps accounted by the different methods. The boxplots display distribution of the valid pixels present in the maps, monthly, for each method. As the boxplots distribution separates from one, it represents that the set of outlier pixels is larger.
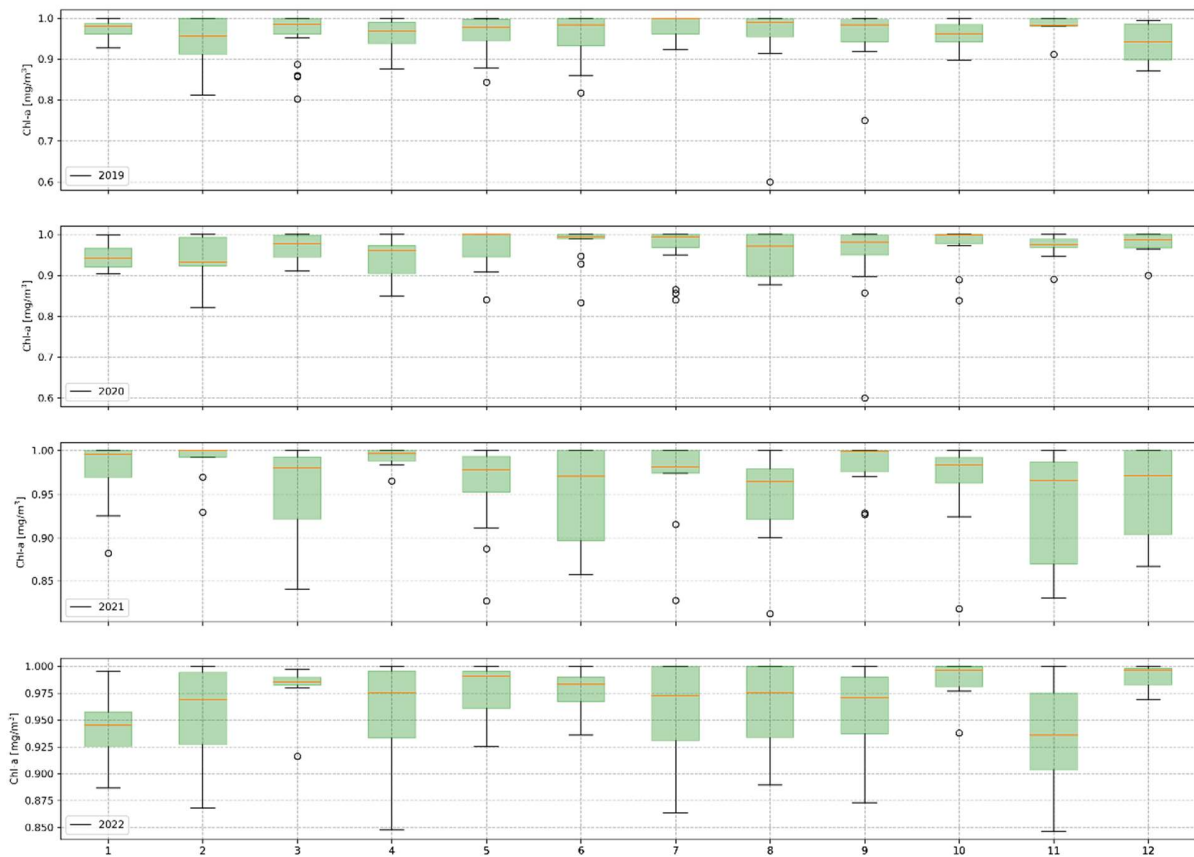
Figure 4-3 - IQR Chl-a valid estimates boxplot per month

First, it is possible to identify for all filtering methods, that the detection of outliers was low for the first two years of the analysis, and that for the first and last months of the maps Chl-a maps the estimates variability led to the extraction of many pixels (pattern which maintained across the different methods). Concerning the k-Sigma methods, as expected, the number of outliers is larger for the 2-Sigma filter when compared to the 3-Sigma. On the other hand, the assessment with respect to the IQR filter is more complex as the number of identified outliers rely in the distribution of the estimates in each of the maps. For this reason, there are cases in which the IQR will identify outliers compared to the k-Sigma filtering. Then, the 3σ method was selected to proceed with the analysis of the WQPs maps. However, it is relevant to remark that the filtering alone is not an absolute tool for determining the presence of outliers within the maps. The filtering method removes estimates which are extremely

large with respect to the trend inside of the maps, as the spatial distribution of the parameters is not expected to vary greatly.
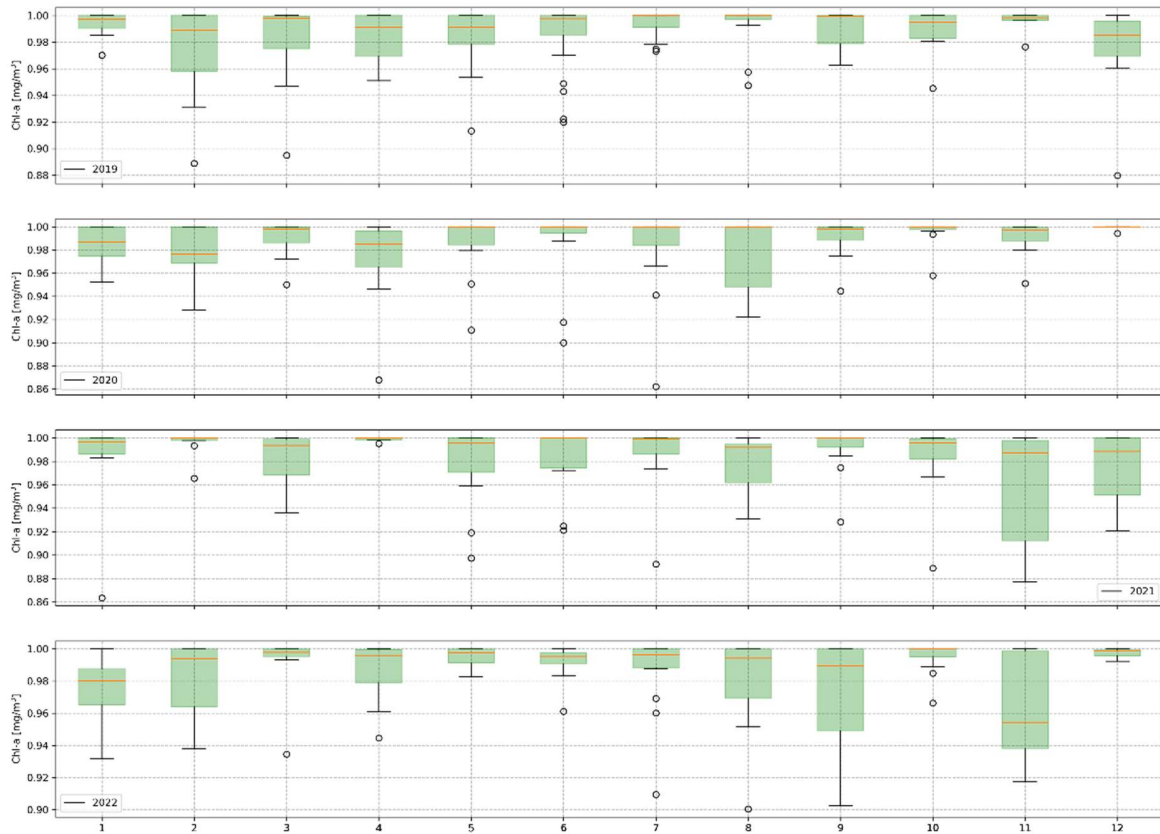
Figure 4-4 - 2-Sigma Chl-a valid estimates boxplot per month

To perform the outlier rejection, a 3σ filtering has been applied in order to single out anomalous WQP estimates for a lake. Then, the out-of-range values detected with the 3σ filtering were explored on the map to interpret the reason for the anomaly in terms of geographical location (e.g., lower surface temperature values could be reasonable when detected in an area affected by inflowing waters, higher surface temperature due to shallow water, etc.).
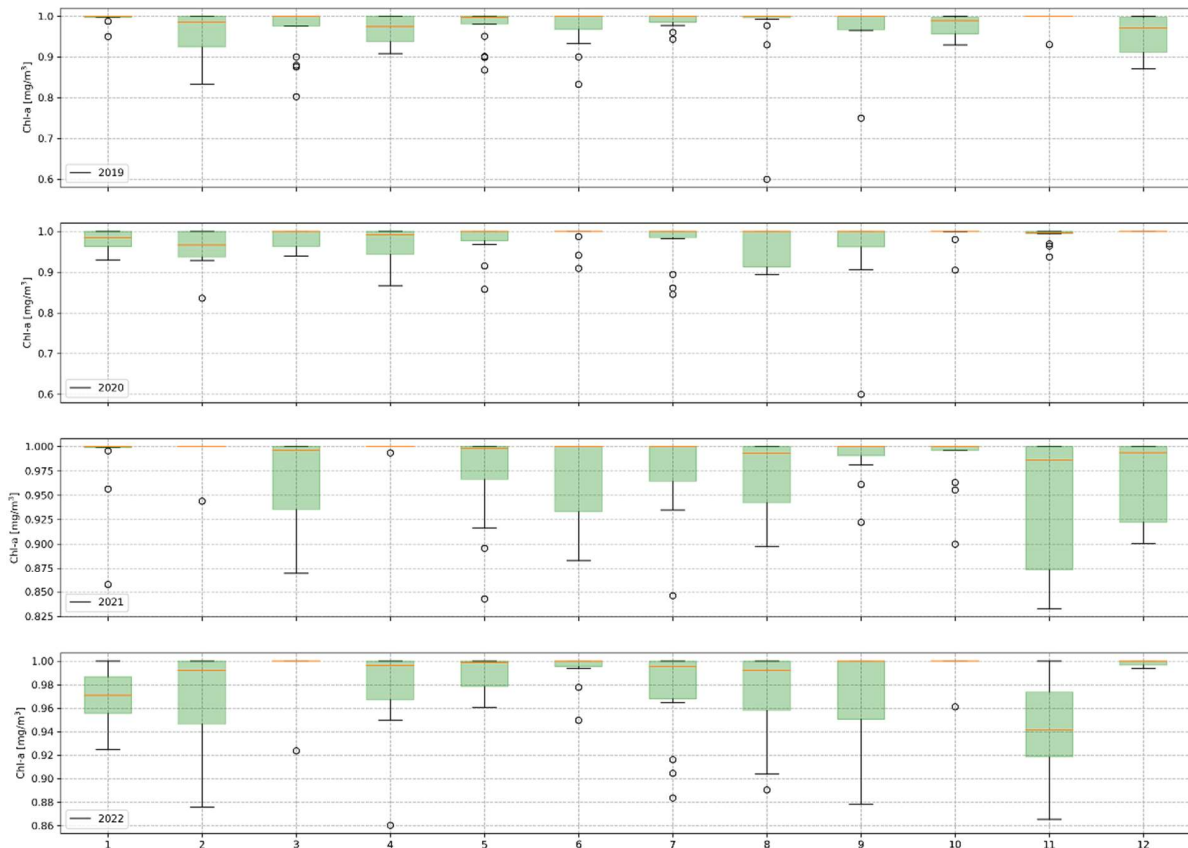
Figure 4-5 - 3-Sigma Chl-a valid estimates boxplot per month

4.3 COMPARISON WITH REFERENCE PRODUCTS (SIMILE & EUMETSAT)

The produced WPQ have been compared to the EUMESAT products to check the differences in the products, for inland water bodies; the comparison was of relevance due to the unavailability of ground truth measurements. EUMETSAT delivers a set of products including the Chlorophyll-a and Total Suspended Matter, and in one of the outputs it uses the C2RCC processor to compute the WQP. These products are identified by the name of the WQP accompanied by the suffix "_CNN" (e.g., CHL_CNN, TSM_NN). However, the outputs can vary due to the characteristics of the water body used for the bio-optical model. Here, we consider that the result of the current work uses lake specific parameters for modelling the subalpine lakes of interest. The comparison between the maps,

EUMETSAT and SIMILE, concerned a set of a 100 randomly distributed sample point over the lakes (see Figure 4-6).
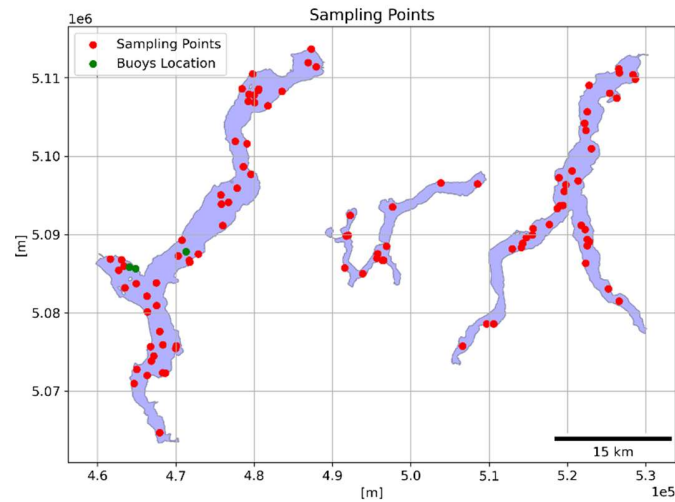
Figure 4-6 - Sampling points for the comparison of the WQP maps

When comparing the products there were three noteworthy points, and which could be of relevance for further investigation of the water bodies of interest. First, while using the same processor for modelling the WQP, the results would vary in at least one order of magnitude between the maps. The variations remark the relevance for setting the proper bio-optical properties for the model. On the one hand, for the CHL-a, SIMILE's estimates tend to be higher with respect to those presented by EUMETSAT. On the contrary, for the TSM, presented the opposite behaviour, with EUMETSAT's estimates having around one, up to two, orders of magnitude greater than those of SIMILE. In addition, there is the presence of greater variations in some specific months of the year, which match in both analysed years. For instance, for the month of March, in both 2021 and 2022, displayed larger differences between the estimates which after further review highlighted the methods in which cloud masking of the products was performed. In some cases, the EUMETSAT outputs was more lenient in the extraction of features as outliers and kept extreme values which may be affected by clouds or cloud shadows.
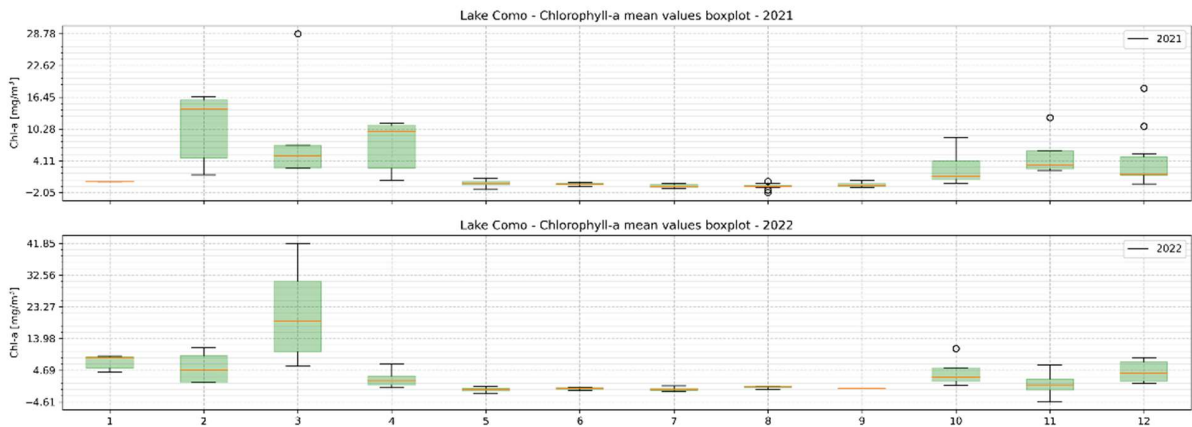
(a)
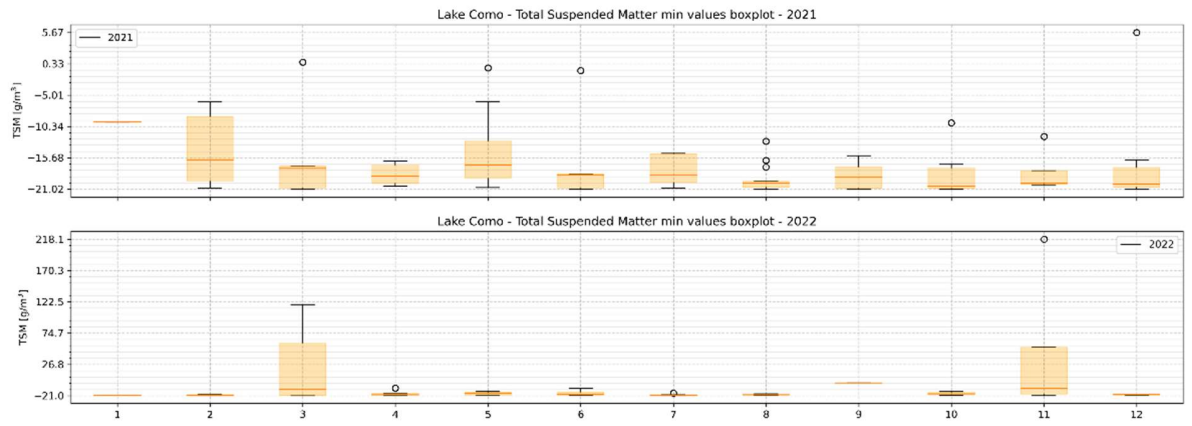
(b)

Figure 4-7 - Monthly boxplots of the difference between SIMILE and EUMETSAT maps. (a) CHL SIMILE & CHL_NN EUMETSAT. (b) TSM SIMILE & TSM_NN EUMETSAT.

Here, the comparison between the products does not expect to provide validation on the WQP maps produced. Nevertheless, it provides important information on the need for providing suitable parameters to the model to obtain more precise and accurate estimates.

4.4 FUTURE DEVELOPMENTS

In the realm of water quality monitoring, ongoing research seeks to explore future developments on Water Quality Parameter (WQP) maps. Here is presented a few exploratory analyses into the potential advancements that could significantly favour the monitoring of the WQP through satellite imagery derived products. Among these prospective improvements, special attention is given to studying the near border areas of lakes, where the interaction between temperature and bathymetry. Additionally, the investigation of water surface reflectance clustering analysis offers an exciting opportunity for identifying outliers, which could lead to more comprehensive water quality assessment. There is potential for gaining valuable insights that may contribute to a more sustainable and resilient future for our aquatic ecosystems.

Temperature-Depth analysis

Dyba et al. (2022), emphasizes that for specific study cases the water temperature variability in the near-surface water layer should be regarded as inconsiderable for depths above 1m. For example, Figure 4-8 (b) shows the Maggiore Lake bathymetry (on the right) and the LSWT SIMILE map before (top right) and after (bottom right) the 3σ outlier rejection for the map produced for 20th August 2020. It can be observed that the LSWT values have been filtered by the outlier rejection within the Angera gulf. However, the rejected values exceed of 1-2 °C the neighbouring unfiltered data and we can see that they correspond to very shallow water in a small gulf, which are conditions that could justify the increase in LSWT. Figure 4-8 (a) shows the correlation between water depth and LSWT). For this reason, the procedure foresees the inspection of filtered data to prevent as much as possible the exclusion of acceptable values.
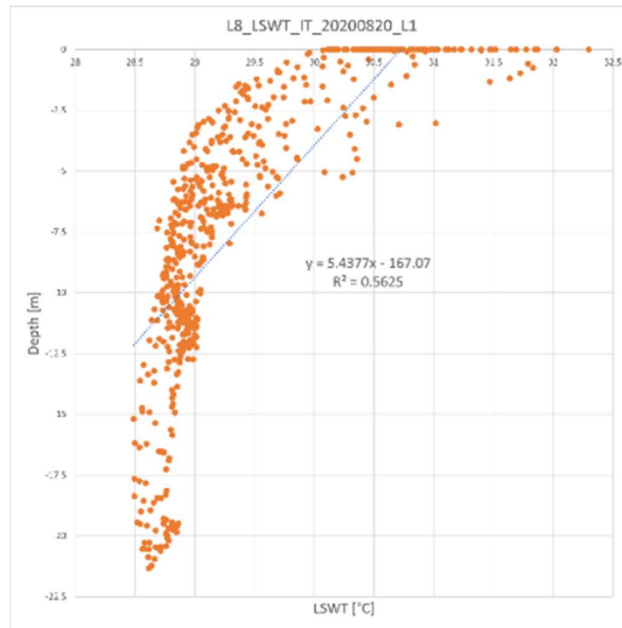
(a)
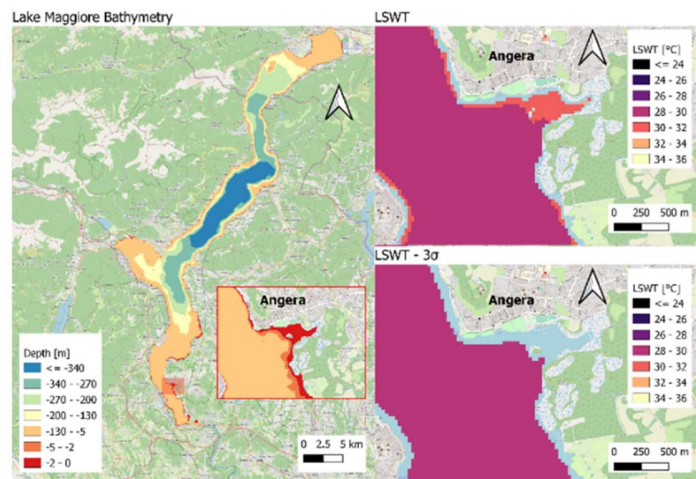
(b)

Figure 4-8 –(a) Correlation between water depth and LSWT for Maggiore lake in Angera gulf. (b) Lake Maggiore bathymetry (on the left) and SIMILE WQP LSWT map of 20th August 2020 before (top right) and after (bottom right) the 3σ outlier rejection.

Clustering analysis WQP's R_{rs}

Here, it is presented an exploratory analysis of the R_{rs} for identifying the presence of anomalous values in the estimates of the WQP maps. The method follows a data analysis to identify statistically

robust groups for the R_{rs} spectra over the map pixels, commonly employed for the distinction of OWT (Spyrakos et al., 2018). The classification of the radiometric quantities consider the scaled values to account for moderating the variation that is attributed to the change of concentrations of the optically active constituents (Ficek et al., 2012). In this case, the clustering considers the normalized values of the $R_{rs}(\lambda)$ to reduce the mean spectral reflectance on the separation of the clusters. Furthermore, Craig et al. (2006) mentions that uncertainties in the $R_{rs}(\lambda)$ reflect on the amplitude rather than the shape of the spectra. For this reason, the standardization served on preserving the shape of the R_{rs} across the spectrum (Vantrepotte et al., 2012).

In order to explore the reasons for anomalous values of Chl-a, a step back with respect to the output of the processor has been made, checking directly the reflectance. In-situ measures of reflectance, collected during dedicated field campaign with handheld WISP-3 spectroradiometers have been taken into account and considered as reference (Hommersom et al., 2012). It was possible to single out the anomalous response (see Figure 4-9b) of point number 1, which corresponds to the buoy position and to an outlier for the map produced upon the satellite image of 1st September 2020 (see also Figure 4-9a).
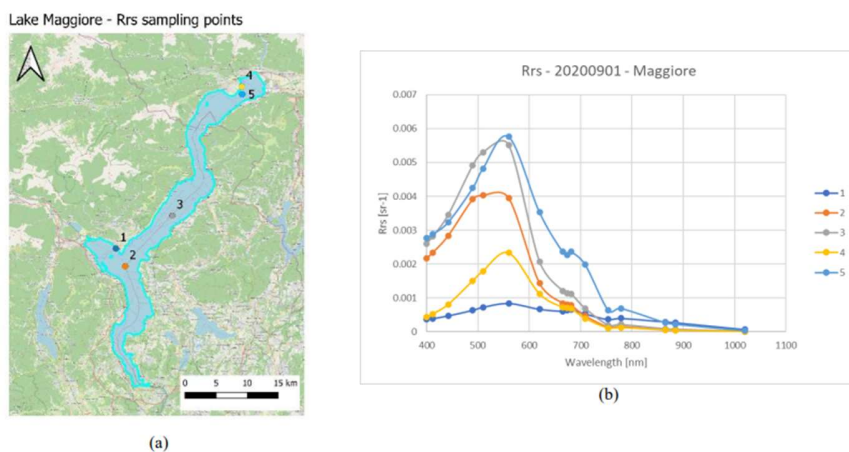
Figure 4-9 – Graph (b) shows the spectral signatures of the sampling points distributed on the lake (a), for the 1st of September 2020 Sentinel-3 image
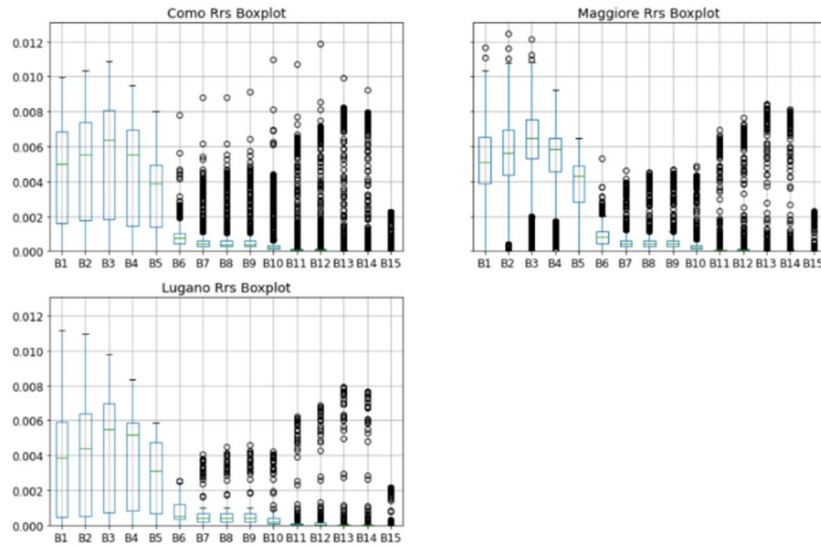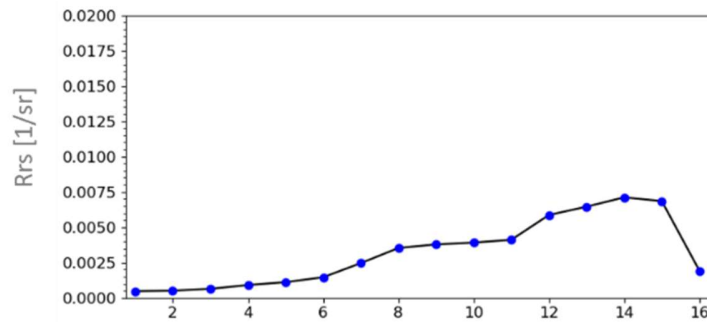
Figure 4-10 – Sample R_{rs} response for March 15th, 2019 (a) $R_{rs}(\lambda)$ estimates across the spectrum. (b) Example of a coast pixel $R_{rs}(\lambda)$ response.

Figure 4-10 (a) summarizes the distribution of the $R_{rs}(\lambda)$ estimates along the spectrum for the 15th of March 2019. The figure presents a set of boxplots for the corresponding $R_{rs}(\lambda)$ considering the complete set of pixels available in the dataset (i.e., pixels which have been clipped within the shape of each lake). For this, the expected behaviour of a water pixel corresponds to the median values found in a boxplot, while the circles display the reflectance of those pixels with different coverages or anomalies (e.g., see Figure 4-11 (b), which corresponds to a pixel affected by adjacency effects, and whose response is that of a near shore pixel). Figure 4-10 presents the outputs for the K-means clustering methods for a set of n=[3, 4, 5], which distinguishes the different patterns for the $R_{rs}(\lambda)$

in the dataset. Observing the first three clusters it is already possible identifying the likely water cover pixels (in green), and a large set of disturbed pixels (in blue and red) by adjacency effects due image's coarse spatial resolution in the borders of the maps or by clouds coverage in the northern part of the lake.

(a)
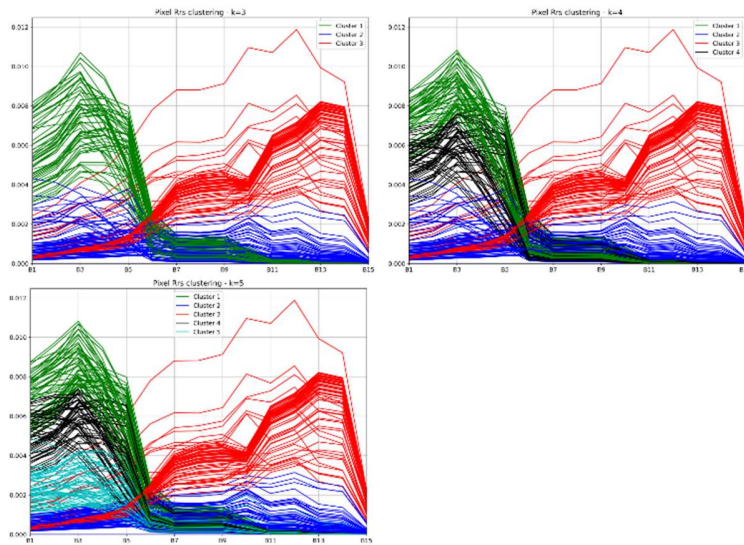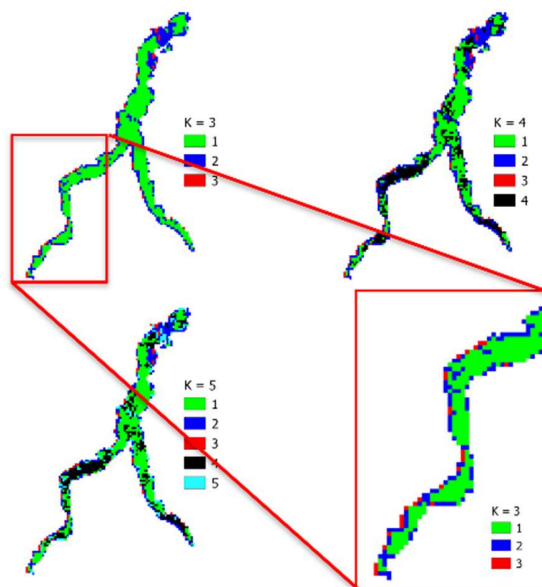
Figure 4-11 - K-means clustering of R_{rs} for the March 15th, 2019, at Lake Maggiore. (a) K- means clustering for n=[3,4,5]. (b) Representation of the clusters in the WQPs map.

In this study, the clustering analysis aims at providing an additional indicator for the detection of anomalous values inside the datasets. The detection methods search not only observing at the WPQs estimates, but also the spectral behavior of the R_{rs} which can rely on the expected behavior of water pixels.

4.5 DESIGN OF TOOLS FOR THE CONTINUOUS LAKE MONITORING

An important part of this work is ensuring the replicability and continuous monitoring of the water bodies. For this reason, the processing chain for the production of the Water Quality Parameters maps, their statistical analysis and outlier rejection methods have been implemented and made available in a virtual machine[18].

The virtual environment contains all necessary software and scripts for the computation of the water quality estimates (Toro Herrera et al., 2022). The environment benefits from docker for providing a portable solution available for different platforms. Docker is an open-source platform designed to streamline the deployment, scaling, and management of applications by utilizing containers. These containers are self-contained units, lightweight and portable, bundling an application alongside its dependencies, libraries, and configuration files. This approach ensures applications run consistently across diverse environments, enhancing reproducibility and maintaining consistency. In the context of research and sharing script tools, Docker serves as a comprehensive solution, providing various advantages for replicability, such as isolation, portability, reproducibility, version control, easy sharing, and simplified setup. When researchers utilize Docker to publish their script tools, they actively

---

[18] https://hub.docker.com/repository/docker/jftoro/wqp_vm

contribute to the broader objective of fostering transparency, replicability, and robustness within scientific research (Giuliani et al., 2020; Graf et al., 2022; Trifonov et al., 2018). The docker composition for building, or importing, the image can be found in GitHub[19]. The docker image used for implementing the virtual machine is built upon a Linux-based environment using Python 3.6, which is the compatible version of Python for using the snappy library to execute the methods available within SNAP software and extending the capabilities for the analysis of the data via plugins. SNAP is a Java-based software and therefore its "native" API is based in Java, snappy is a wrapper for implementing python code using the Java SNAP engine API.

The virtual machine aims at simplifying the workflow on the production of the Water Quality Parameters maps. Here, the implementation will not make use of the SNAP interface to use existing functionalities, but they will be reachable through a Jupyter Notebook with a ready-to-go setup for maps production. Some of the considerations for the implementation of the building the virtual environment concerned:

---

[19] https://github.com/JFToro192/jpnbWQPsimile
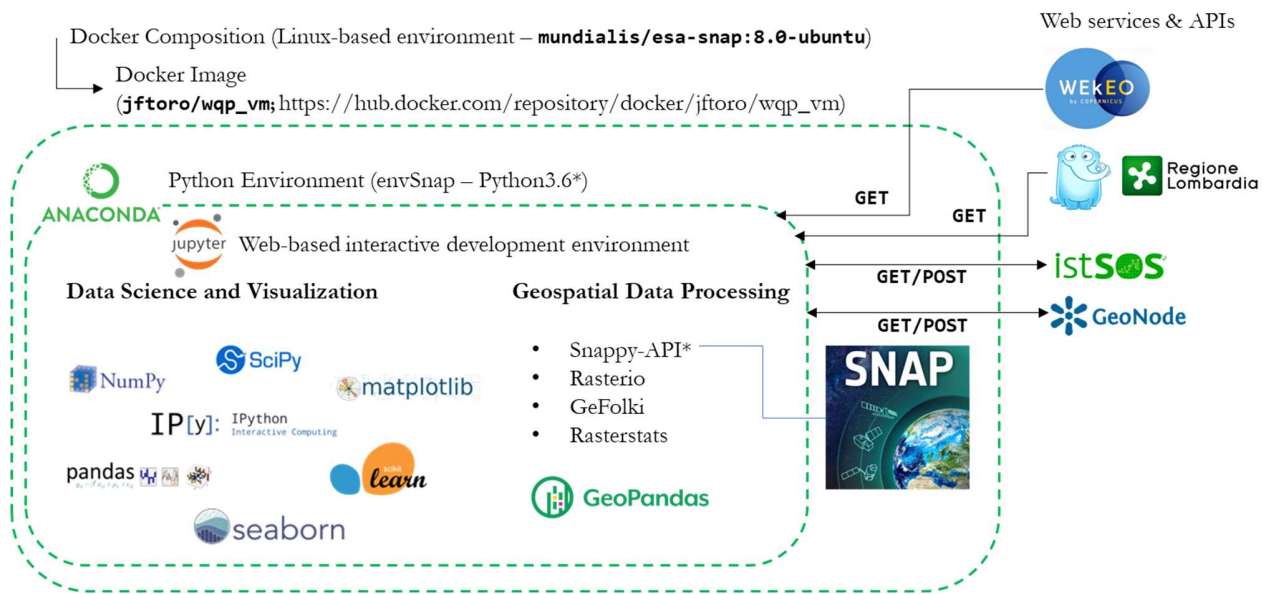
Figure 4-12 - Virtual machine with tools for the processing and publication of WQP maps and statistics

- Use of python scientific and data visualisation libraries such as numpy[20], pandas[21], scipy[22], matplotlib[23], scikit-learn[24], etc.

- Use of libraries for executing HTTP protocols to retrieve and submit data through multiple APIs (e.g., Download Sentinel-3 OLCI images, retrieve meteorological data and post descriptive statistics into a istSOS instance, and the upload of the WQP metadata)

- Use of libraries for treating geospatial data such as rasterio[25] and geopandas[26].

- Implement a Graphical User Interface to perform analysis over the satellite images and Water Quality Parameters maps, which enables users to implement extended functionalities to the existing workflow based in Jupyter[27].

---

[20] https://numpy.org/
[21] https://pandas.pydata.org/
[22] https://scipy.org/
[23] https://matplotlib.org/
[24] https://scikit-learn.org/stable/
[25] https://rasterio.readthedocs.io/en/latest/
[26] https://geopandas.org/en/stable/
[27] https://jupyter.org/

- The use of Python 3.6 compatible libraries is required as this is the base Linux image python compiler.

The image installs two of the modules for the SNAP software required for using the C2RCC processor and reading the Landsat images. Limiting the number of SNAP modules goal is avoiding a large size for the working image. However, within the docker composition, it is possible to set up and install other modules if needed by the user[28].

Toolbox	Module Name
Sentinel-3 Toolbox - Landsat products reader	org.esa.s3tbx.s3tbx.landsat.reader
Sentinel-3 Toolbox - C2RCC processor function	org.esa.s3tbx.s3tbx.c2rcc

The GUI for the data processing benefits from the python JupyterLab development environment. The interface enables the use of the different python scripts required for following the different passages for the WQPs maps processing. The standard workflow for the processing of the satellite image to obtain the WQPs estimates divides into severalnotebooks. The order for replicating the workflow is as shown next:

1. downloadMeteoARPA.ipynb: Retrieval and storage of the temperature measurements from an in-situ sensor of the meteorological network of ARPA Lombardia

2. downloadWEkEO: Request the Sentinel-3 images through the Harmonized Data Access API from WEkEO

3. downloadMeteoARPA: Retrieval and storage of the temperature measurements from an in-situ sensor of the meteorological network of ARPA Lombardia

---

[28] SNAP modules: https://senbox.atlassian.net/wiki/spaces/SNAP/pages/15269950/SNAP+Configuration

4. downloadWEkEO: Request the Sentinel-3 images through the Harmonized Data AccessAPI from WEkEO

5. WQP_Production: Production of the WQP maps

6. WQP_Coregistration: Coregistration of the WQP maps from the Sentinel-3 satellite images

7. WQP_Statistics: Computation of the descriptive statistics of the WQPs for each lake

8. WQP_istSOS.ipynb: Upload the retrieved descriptive statistics into the istSOS platform

In addition, the JupyterLab interface allows the integration of new notebooks, outside of the processing chain defined for the production of the WQP maps for their analysis. For example, the WQP_istSOS.ipynb is a notebook devoted to the computation of the lakes basin level descriptive statistics, which are later uploaded into the Sensor Observation Platform presented in section 4.6. Furthermore, it includes the integration of procedures for uploading water quality parameter (WQP) metadata, which enhances the dataset information by providing users with details such as the sensor used during map acquisition, date of image acquisition, typology, and coordinate reference system. The virtual environment is designed to offer flexibility in map analysis, serving not only as a tool for continuous production of water quality parameter maps but also as a platform for implementing new methods of map analysis and sharing processed results.

4.6   ISTSOS PLATFORM

istSOS[29] is a Free and Open Source Sensor Observation Service Data Management System. The platform implements OGC Sensor Observation Service standards (SOS [30]) used for data interoperability of retrieved from sensors. istSOS is released with the GNU General Public License.

The OGC SOS server, istSOS, is implemented in Python, allowing the management and dispatch of observations from monitoring sensors. istSOS is provided with a Graphical User Interface and RESTfull Web API for the automation of administrative procedures (e.g. such as data upload).
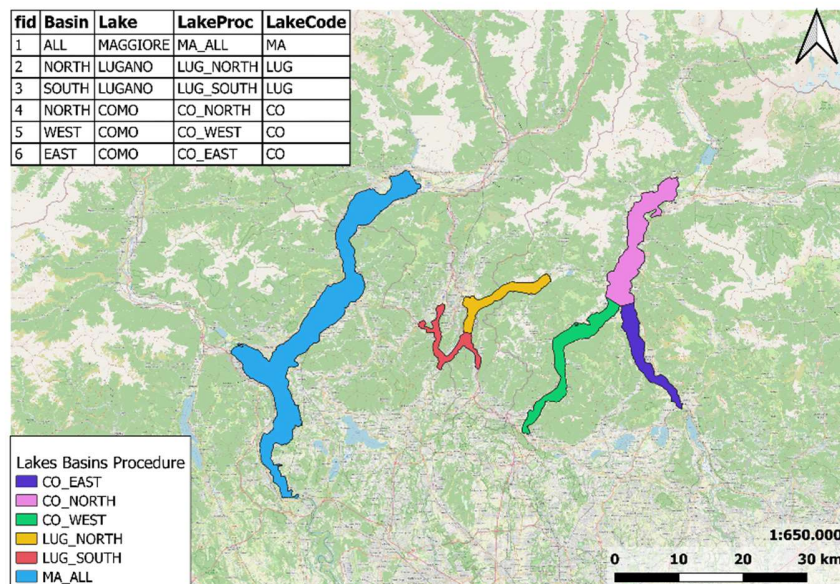
Figure 4-13 - Lakes basin level division

---

[29] http://istsos.org/en/latest/doc/
[30] https://www.ogc.org/standards/sos

Procedures

```
SATELLITE_{WQP}_{Basin}_{Stat}
```

The combinations for naming the procedures are defined as follows:

Parameter	Description	Values
'WQP'	Water Quality Parameters	['CHL_TURB', 'TEMP']

'Basin'	Lakes Basins	{
		 'Como':['CO_E', 'CO_N', 'CO_W'],
		 'Lugano':['LUG_N', 'LUG_S'],
		 'Maggiore':['MA']
		 }

'Stat'	Statistics	{
		 'Mean':[''],
		 '1st Quantile':['1Q'],
		 '3rd Quantile':['3Q'],
		 'Standard Deviation':['SD']
		 }

Sample output of a getObservation request from istSOS.

```
{'name': 'SATELLITE_CHL_TURB_LUG_S',
'samplingTime': {'beginPosition': '2019-01-15T10:12:15+00:00',
'endPosition': '2022-10-07T09:48:10+00:00',
'duration': 'P1360DT23H35M55S'},
'procedure': 'urn:ogc:def:procedure:x-istsos:1.0:SATELLITE_CHL_TURB_LUG_S',
```

```
'observedProperty': {'CompositePhenomenon': {'id': 'comp_110',
'dimension': '5',
'name': 'timeSeriesOfObservations'},
'component': ['urn:ogc:def:parameter:x-istsos:1.0:time:iso8601',
'urn:ogc:def:parameter:x-istsos:1.0:water:Chl:a',
'urn:ogc:def:parameter:x-istsos:1.0:water:Chl:a:qualityIndex',
'urn:ogc:def:parameter:x-istsos:1.0:water:TSS',
'urn:ogc:def:parameter:x-istsos:1.0:water:TSS:qualityIndex']},
'featureOfInterest':           {'name':           'urn:ogc:def:feature:x-
istsos:1.0:Point:lugano_basin_south',
'geom':                                          "<gml:Point
srsName='EPSG:4326'><gml:coordinates>8.90780249008498,45.9177648931359,270.5</
'result': {'DataArray': {'elementCount': '5',
'field': [{'name': 'Time',
'definition': 'urn:ogc:def:parameter:x-istsos:1.0:time:iso8601'},
{'name': 'water-Chl-a',
'definition': 'urn:ogc:def:parameter:x-istsos:1.0:water:Chl:a',
'uom': 'mg/m3'},
{'name': 'water-Chl-a:qualityIndex',
'definition': 'urn:ogc:def:parameter:x-istsos:1.0:water:Chl:a:qualityIndex',
'uom': '-'},
{'name': 'water-TSS',
'definition': 'urn:ogc:def:parameter:x-istsos:1.0:water:TSS',
'uom': 'g/m3'},
{'name': 'water-TSS:qualityIndex',
'definition': 'urn:ogc:def:parameter:x-istsos:1.0:water:TSS:qualityIndex',
'uom': '-'}],
'values': [['2022-10-07T09:48:10+00:00', 2.99, 100, 0.41, 100]]}}}}
```

The uploaded water quality parameters estimates can be visualised in the Business Intelligence Platform of SIMILE project[31]. Figures A.1a and A.1b present an example on exploiting the data stored

---

[31] https://simile.ddns.net/

and dispatched through the istSOS platforms using the RESTful API to visualise the evolution of the

WQPs maps on a basin level.
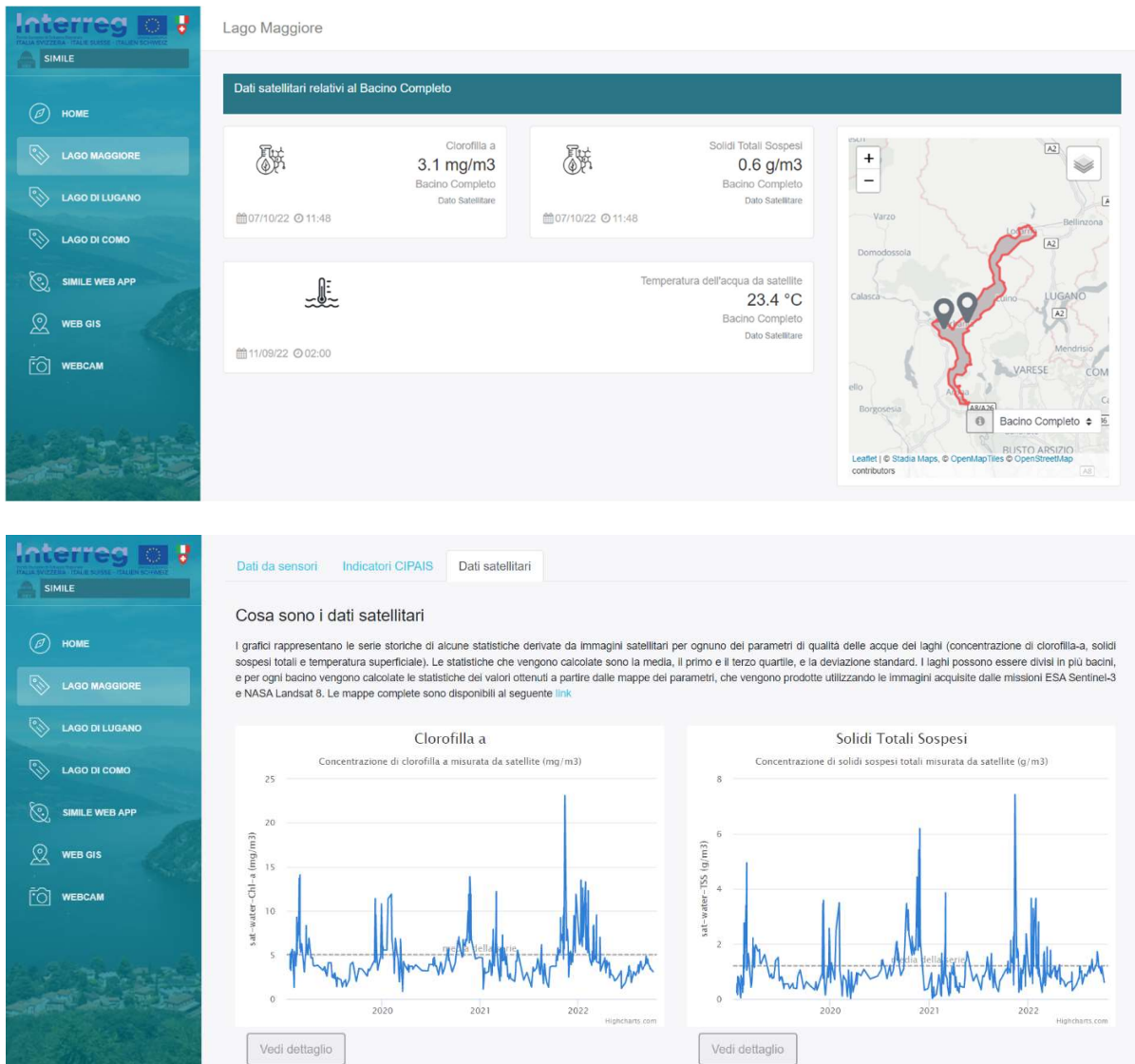
Figure 4-14 - Business Intelligence Platform. Satellite data. (a) Basin level coverage and last uploaded measurements of the WQP maps; (b) Chlorophyl and Total Suspended Matter time series for the selected basin

Chapter - 5. WATER QUALITY PARAMETERS DISSEMINATION

Due to the widespread use of geo-referenced data sets in numerous fields of science and business, such as earth observation, environmental sciences, city planning, BIM, real-time processing, and analytics for geospatial data, geospatial data management is becoming an increasingly important step in the workflow of geospatial data processing (Breunig et al., 2020). Open data platforms and containers are being considered by current technologies for effectively managing large amounts of geographic raster and vector data. The topic of interest is the strong connection between geospatial data management and other geo-information science research fields like data modelling, visualization, and analysis. The scope of this work is implementing a Content Management System, for the upload of the water quality status parameters derived from the processing chain, and WebGIS, for the temporal visualisation and querying of WQP maps, for the monitoring of inland water bodies.

Here is presented the set of technologies employed in implementing a Collaborative Data Sharing platform and WebGIS for the dissemination and monitoring of the WQPs maps. The first relies on a customized version of the GeoNode project to publish and share the WQP maps. GeoNode is a flexible collection of FOSS and frameworks for the sharing of geospatial data with proven success among different fields such as disaster and emergency management, e.g. floods (Bhattacharya et al., 2014; Cristofori et al., 2015) and earthquakes (Boccardo & Pasquali, 2012; Liberzon, 2019), socioeconomic and sustainable development studies (Bhattacharya & Painho, 2017; Guan et al., 2012; Rudiastuti et al., 2020; Tulokhonov et al., 2020), management and analysis of satellite-derived products (Buonanno et al., 2019a; Menegon et al., 2018), and many others. The second is a WebGIS addressing the WQPs maps shared in the Collaborative-Data Sharing Platform. WebGISs has proven to be an effective tool for water resources management and monitoring, providing access and interactive visualization of the water quality indicators and models through the internet to engage the public and

promote public awareness (Arias Muñoz et al., 2017; Delipetrev et al., 2014; Kulawiak & Lubniewski, 2014).

5.1 GEODATA PUBLICATION AND OPEN STANDARDS

Creating a comprehensive system that encompasses the necessary tools for visualization and analytics, supporting spatial decision-making, poses a considerable difficulty due to the intricate nature of spatial decision problems. As the utilization of diverse tools from various origins is unavoidable, ensuring their interoperability becomes crucial. Interoperability enables technology developers to work together and benefit from one another's efforts. In addition, traditional desktop data processing and visualization systems are being superseded by web-based platforms. Interoperable technologies are even more crucial on online platforms than desktop systems. The establishment of standards is fundamental in achieving interoperability. The Open Geospatial Consortium (OGC) defines web standards for geospatial data and services.

OGC is an international consortium with the goal of making geospatial data and services FAIR (i.e., findable, accessible, interoperable, and reusable) by establishing web standards. Software developers use them to incorporate open interfaces and encodings into their services and products. The primary "products" of the Open Geospatial Consortium (OGC) are membership-created standards that address specific interoperability issues like the online publication of map content, the exchange of crucial location data during disaster recovery and response, and the facilitation of the fusion of data from various IoT devices[32].

The progress in this field empowers individuals to accomplish various tasks such as processing and analysing intricate spatial data, as well as visualizing information. In today's interconnected world, data

---

[32] https://www.ogc.org/standards

is accessible as a combination of spatial information from diverse sources. Geographic Information Systems (GIS) and related disciplines have witnessed significant advancements through extensive research in recent decades. GIS has been specifically developed to effectively manage large volumes of geographical data. By incorporating computer and network technologies into spatial data, GIS has enhanced its capabilities, enabling end users to perform complex data processing, spatial analysis, and data visualization tasks.

In the realm of webGIS applications, a web browser acts as the client, communicating requests to a web server, which in turn responds to those requests. Unlike non-spatial web applications that typically rely solely on a web server, webGIS applications involve a data or map server that handles spatial data. This server can execute GIS functions and manage geospatial data, providing seamless integration and exchange of geospatial data across diverse systems and applications through web services. (Agrawal & Gupta, 2017). Within the numerous Community provided by the OGC, the Collaborative Data Sharing Platform in this work takes particular interest in the Web Map Service (WMS). The incorporation of the Collaborative Data Sharing Platform and WebGIS in this endeavor was guided by the utilization of OGC Community Standards for the publication of geospatial data.

By employing a simple HTTP interface, clients can acquire geo-registered map images from distributed geospatial databases using the standard Web Map Service (WMS) interface. A WMS request defines the desired processing area and the corresponding geographic layer(s). Furthermore, for certain geo-registered map images that are returned in formats such as JPEG or PNG, the response can be visualized within a web application. Additionally, users can indicate in the interface whether the retrieved images should possess translucency, enabling the combination of layers from different servers.[33].

---

[33] https://www.ogc.org/standards/wms

5.2 SERVICE-ORIENTED ARCHITECTURE

Client-server architecture performance, scalability, maintenance, and other quality parameters are major architectural issues. Meanwhile, small-scale web services are vast and cover specific client needs. According to Harrison and Reichardt (2001) (Golladay et al., 2021), web services are the new model for internet-based applications without the need for complex and bundled software packages. Furthermore, geospatial web services differ from traditional ones due to the variety of spatial data models, data formats, data semantics, and relationships. This variety leads to concerns regarding data interoperability, which the OGC standards have addressed. The OGC standards provide an interface specification of several services (e.g., WMS, WFS – Web Feature Service, WCS – Web Coverage Service, etc.).

On their own, these web services work as a stand-alone solution, where their integration is complex. Service Oriented Architectures, known as SOA, emerged as a solution to harmonize disparate systems and data by merging applications and information of different natures. SOA architectural style concerns the temporary relationship between service providers and consumers, runtime issues from the service provider, and expectations of the service consumer. OASIS (Organization for the Advancement of Structured Information Standards) defined the reference model for organizing and using distributed capabilities under the control of different ownership domains (OASIS SOA Reference Model TC | OASIS, n.d.).

Web services in SOA have three roles: provider, broker (catalogue), and requestor (end-user), each with its implementation and interface. The internal specification of the web services is known as the implementation. In contrast, the external behaviour and interaction with other web services is the interface. The interfaces can divide into two types, buy and sell. Buy/Sell interfaces defines the required/provided services and plays the requestor/producer role.

Software and procedures from intelligent structures will be imported more frequently to create novel GIS services (Breunig et al., 2020). Nowadays, GIS services rely on a service-oriented architecture (SOA) approach, which enables support sharing and recycling through services-oriented characteristics. Furthermore, with the possibility of building blocks of infrastructure as containers, it is possible to modify, edit, update, and migrate existing SOA. The containerized solutions allow to monitor resource allocation, are easy to update, are reusable and carry less overhead (run virtual services instead of a full virtual machine). The current work presents two sets of containerized solutions enabling the publication, management, and production of the WQP maps. Each container in the collection, detailed later in the chapter, is a software-defined environment for running scripts and is not dependent on the hosting environment.

5.3 WEBGIS – CLIENT SERVER ARCHITECTURE

WebGISs follow a client-server architecture, where the client represents an interface of interaction on a web-based application and the server handles the processing, storage, and management of geospatial data. The breakdown of a client-server WebGIS architecture is of a client, that is a web or desktop application that interacts with a GIS system, and the server which processes and serves data to a client. The communication between client and server is achievable through requests and response these two components by means of web services (e.g., maps visualisation, querying data, analyse spatial data, and other GIS capabilities provided by the server). This architecture follows a traditional network architecture. It has different approaches: thin client, thick client, and hybrid architecture (Agrawal & Gupta, 2017). In this work we focus using a hybrid client architecture, where the client has minimal processing capabilities and relies on the GIS functionalities provided by the server. However, WebGIS is client-side rendered which benefit from technologies enabling its functionality. The use of this type of architecture is preferred as it reduces the client-side requirements but also relies on server-side

operations, or APIs, for accessing the information of the WQP maps. Hybrid client web GIS applications offer a balance between rich client-side functionality and the ability to leverage server-side resources.

Thin client architecture

The client-side resource requirements are low for this architecture. On the server side, the bulk of the processing takes place. The server responds to a client's request by creating a response, which in its most basic form can be a map created from a database. This spatial response has been formatted for the web so a web browser can display it. The Common Gateway Interface (CGI) or a gateway script must be used because the client cannot call the GIS server directly. The Java program known as a servlet is another choice for server-side technology. Due to its ability to handle multiple client requests and lack of need to start, load, and stop for each request, the servlet is more efficient than the CGI. The Application Programming Interface (API), Active Server Pages (ASP), and Java Server Pages (JSP) are some additional technologies.

The advantage of this architecture is that the client side has no responsibilities. In this case, the resources are available through a server HTTP request and the display of the processed result sent by the server. The results are mainly in image formats such as jpeg, gif, or png. Therefore, raster data results are easier to display; on the other hand, vector data display is impossible unless rendered in raster format. The server centralizes the data processing control, which eases the data updates and maintenance activities. A drawback in a production environment is the capacity to deal with individual client needs and requirements. The architecture is a low-cost solution from the client-side viewpoint. In contrast, the server receives the entire load and limits its operation on bandwidth and other issues. However, while benefitting the client on the processing workload, the problem relies on the limited functionalities on the client side. For example, basic operations such as zooming, querying and other

basic operations result in a new request to the server. In many cases, this increases the number of interactions and underutilises the client side.

Thick client architecture

Thick client architecture takes advantage of the browsers' capabilities by using plugins, applets, or ActiveX. For this, the processing is no longer entirely supported on the server; now, the client has a role in the processing. The plugins are executable on the client side. They must be pre-installed in the client machine, allowing the client to work on the native data types, specialist viewing and functions manipulation. An alternative to the plugins is using applets, or Java executables, which do not require previous installation. Applets can support visualization processing on the client side. Also, the implementation enables the manipulation of raw data and its processing by the client, e.g., filtering, mapping to a geographic representation, rendering using opacity, etc.

Running a client-side system is advantageous when there is low connectivity with the server, as the client receives the raw data for its local processing. This also implies that the number of requests to the server is reduced for minor actions, lessening the server load.

Hybrid architecture

A hybrid client architecture is a combination of both thin and thick architectures. In this work, the WebGIS benefits from both sides' technologies, client, and server. A hybrid client architecture in web GIS combines the features of both thin and thick client architectures. It allows the client-side to handle some processing and rendering tasks while also relying on server-side operations (Agrawal & Gupta, 2017). The client can cache data and perform local processing for improved efficiency, but it still maintains interaction with the server for data updates and more complex operations. Hybrid client web GIS applications strike a balance by providing rich client-side functionality while leveraging

server-side resources. Frameworks like OpenLayers or Leaflet exemplify hybrid client web GIS, where data rendering and basic analysis take place on the client, while advanced functionality depends on server-side APIs.

5.3.1   Collaborative Data Sharing Platform

GeoNode is a web-based platform composed of open-source frameworks to implement a GIS (Geographic Information System) and an SDI (Spatial Data Infrastructure) (Buonanno et al., 2019b; Corti et al., 2019). Interoperability is one of the key drivers for implementing an SDI; GeoNode complies with the OGC standards. Furthermore, GeoNode environment provides functionalities such as:spatial search of data and metadata;

- management and sharing of raster, vector and metadata;

- management of the security policies on data sharing;

- data visualization and integration from different sources, both stored on infrastructures and from services supplied from outside through WMS, by using an integrated web development environment to build interactive maps

GeoNode leverages on the Python programming language for its development environment using a Python-based web-framework. The framework supports the orchestration of different components of the Content Management System (CMS). GeoNode's collection of frameworks and software can summarise in the following:

- Docker: a platform to help developers build, share, and run modern applications. The platform uses OS-level virtualization for the delivery of software in packages called containers;

- Django: a high-level Python Web framework;

- GeoServer: an open-source software written in Java that allows users to share and edit geospatial data. The software implements several standards, including OGC standards, that enable the creation of maps and sharing of data inside GeoNode;

- GeoWebCache: a Java web application used to cache maps, accelerating, and optimising the image delivery. It aims to eliminate redundant request processing and save processing time. Typically, GeoNode mapping client request tiles from the WMS-C GeoServer endpoint. If a given tile has already been generated and is not expired, it returns from GeoWebCache to the mapping client. Otherwise, it generates a WMS request to GeoServer.

- PostgreSQL: an object-relational database that makes possible the storage, query, and analysis of spatial information;

- PostGIS: a spatial database extension for PostgreSQL database. The extension adds support for geographic objects. It enables spatial queries in Structured Query Language (SQL). SQL is a standard for "Create" "Read" "Update" and "Delete" operations in the database.

GeoNode enables the management of geospatial datasets, either vector or raster. Vector data is stored in a PostgreSQL/PostGIS geodatabase, while a file system hosts raster data as GeoTIFFs. After storing the datasets, GeoNode exports the datasets to GeoExplorer (WebGIS component) through GeoServer by using GeoWebCache to accelerate and optimize image delivery. Moreover, GeoServer allows accessing the DBMS through GeoExplorer and other third-party OGC-compliant software to interact and visualize data. Users can fill in metadata on uploaded data to ease the searching and retrieval through the CSW (Catalogue Service for the Web) utilizing pyCSW that is is an open-source, standards-compliant catalog server for publishing and discovering geospatial metadata.

GeoNode follows a resource-based model as the baseline for defining the core properties in its information model. Properties such as a Universal Unique Identifier (UUID), a bounding box, a title,

other metadata properties and publishing workflow statuses serve as a standard to manage relevant resources (layers, maps, documents) by inheriting them from the model. In GeoNode, a layer represents a spatial dataset, which can be of vector or raster type or a remote web service. Each layer can be associated with multiple styles and, in the case of vector datasets, can be composed of several layer attributes.

GeoNode implements Django templates and many JavaScript libraries and frameworks to build the client's user interface. Other web frameworks, such as AngularJS[34], let users access the Django Tastypie API. JQuery[35] framework is used widely in the application context, for example, on the autocompletion search boxes. GeoNode implements a RESTful API (Tastypie API) allows users to interact with geospatial data stored on the platform. The API is built using the Django REST framework, which provides a set of standard views for common operations such as GET, POST, PUT, and DELETE.

Some of the key features of the Geonode API include:

- Access to geospatial data: The Geonode API allows users to retrieve metadata, browse and download data files, and search for specific datasets based on various criteria such as keywords, location, and time.

- Authentication and authorization: The API supports user authentication and authorization using OAuth2 or Token authentication, allowing users to access private data and perform actions such as uploading and editing datasets.

- Integration with other applications: The API can be used to integrate Geonode with other applications or services, such as web mapping applications or data processing workflows.

---

[34] https://angularjs.org/
[35] https://jquery.com/

- Customization and extension: The Django REST framework used by Geonode provides a flexible and extensible architecture, allowing developers to customize or extend the API to meet specific needs.

Overall, the Geonode API provides a powerful and flexible way to interact with geospatial data stored on the platform, and enables users to integrate Geonode with other applications and workflows.

The interface integrates a mapping client based on OpenLayers[36] or Leaflet[37]. Some of the most renowned mapping clients in the platform are:

- GeoExplorer[38], based on OpenLayers, GeoExt[39], and ExtJS[40] JavaScript frameworks, which is the default and possibly the most common choice for GeoNode deployments;

- MapLoom[41], which is built on top of Angular and OpenLayers. An interesting feature of MapLoom is its GeoGig[42] integration, which enables geospatial feature versioning;

- MapStore[43], based on the React JavaScript framework. It is not tied to a specific Javascript mapping library, but it can indifferently use OpenLayers, Leaflet or CesiumJS[44];

- WorldMap[45], developed by Harvard University Center of Geographic Analysis, a forked version of GeoExplorer, adds features such as attribute-based classifications style editor, layer categories in the table of content, a note editor and a gazetteer.

---

[36] https://openlayers.org/
[37] https://leafletjs.com/
[38] https://connect.boundlessgeo.com/docs/suite/4.6/geoexplorer/
[39] http://geoext.org/
[40] https://www.sencha.com/products/extjs/
[41] https://github.com/ROGUE-JCTD/MapLoom
[42] https://geogig.org/
[43] https://mapstore.geo-solutions.it
[44] https://cesiumjs.org/
[45] https://github.com/cga-harvard/geoexplorer-worldmap-client

GeoServer is the selected spatial data server in the GeoNode collection of technologies. The client can reach GeoServer from its web browser or an external GIS application for rendering map tiles or stream/edit vector geometries (WMS, WFS and WFS-T).

5.3.2 Map service

GeoServer is a Java-based FOSS implementing open standards for sharing geospatial data. It can run using a servlet engine such as Tomcat[46] or Jetty[47] and benefits from the GeoTools Java geospatial library. OGC standards, such as WMS, WFS, WFS-T, and WCS, are integrated within the spatial server and, in some cases, it is possible to enrich the standards catalogue through extensions (for example, for WMTS – Web Map Tile Service). One more extension refers to the CSW. OGC CSW is a standard, first published in 2010. Its latest version is 3.0, published in 2016, that divides in two parts: the general model and the HTTP protocol binding, also known as CSW which builds upon the general mode. The general mode abstractly specifies clients and catalogue services interfaces. The CSW enables publishing and searching metadata for geospatial data and services related to resources.

Typically, GeoServer clients request that from web browsers and desktop GIS software. GeoServer has a web-based management interface connected to the data sources on the back-end. GeoTools is part of the core technologies of GeoServer, released with version 2.1 of the GNU Lesser General Public License (LGPL). The source code of GeoTools is available on GitHub[48]. GeoTools and GeoServer are OSGeo projects. GeoServer is a community-driven project developed, tested and supported by individuals and organizations worldwide. Its latest release is version 2.20, and it works on Microsoft Windows, Linux and Mac OS X operating systems with an appropriate Java Runtime

---

[46] https://tomcat.apache.org/
[47] https://www.eclipse.org/jetty/
[48] https://github.com/geotools/geotools

Environment (JRE). GeoServer latest releases use the GNU General Public License (GPL). GeoServer's source code is available on GitHub[49].

GeoServer supports vector and raster geospatial data. Concerning vector data, the supported formats are shapefile, GeoPackage (GPKG), and GML. On the other hand, the supported raster data formats are GeoTIFF, ImageMosaics, GPKG, and ArcGrid. In addition, GeoServer can connect with different remote data sources, such as databases, including PostgreSQL with PostGIS, MySQL, and Oracle, and remote web services, such as WMS, WMTS, and WFS servers. GeoServer extensions and community modules may support other data formats, databases and OGC web services.

GeoServer primarily enables the viewing and editing of geospatial data. Viewing and downloading the geospatial data in several vector and raster formats is available by default to any user accessing a GeosServer instance. Some available formats for downloading geospatial data are KML, PDF, PNG, JPEG, GeoTIFF, SVG, GIF, GML, GeoJSON, JSONP, CSV, and shapefile, among many others. GeoServer implements a viewer based on the Free and Open Source JavaScript library OpenLayers to visualize raster data. OpenLayers, a web mapping library, allows displaying of dynamic and interactive 2D maps using various vector and raster formats on a web application. The visualization benefits from the response made to a WMS, WMTS, or WFS server. The web mapping library is released under the Simplified Berkley Source Distribution (BSD) License, also known as FreeBSD License. The source code of OpenLayers is available on GitHub[50].

Another relevant component for GeoServer is GeoWebCache[51] (GWC). GWC is a Java-based FOSS for caching map tiles web application. It is also available as a standalone product used on other servers. It runs as a proxy between clients, servers and cached tiles on their request. When the maximum

---

[49] https://github.com/geoserver/geoserver
[50] https://github.com/openlayers/openlayers
[51] https://www.geowebcache.org/

cached size is reached, the least frequently or recently stored tiles are removed. The administrator of the GeoServer instance can set the maximum cache size and criteria for removing tiles. Moreover, it can generate all the tiles for specific zoom levels in EPSG:4236 (WGS 84) or EPSG:3857 (WGS 84 / Pseudo-Mercator), in PNG or JPEG format, with a user-defined style, bounding-box and store them permanently. Caching and storing tiles eliminates redundant requests and saves processing time. GWC implements different tiling protocols, such as Web Map Service Tile Caching (WMS-C), Tile Map Service specification (TMS), Web Map Tile Service (WMTS), XYZ, and Bing Maps Tile System, to accelerate and optimize the mp imagery delivery. WMS-C provides a method for clients to fetch tiles from tiles servers and cache them on the server, an intermediate location, or pre-generate them completely. Furthermore, it offers a method for identifying that a tileset for a particular layer from a WMS or WMS proxy is available. In this way, it leverages the existing resource in the WMS server. OSGeo is the developer of WMS-C, and while TMS and WMTS have superseded its use, GeoServer still supports it. TMS is a specification that standardizes how map tiles are requested by clients and the ways that server describe their holdings. It has been developed by OSGeo. TMS served as the basis for WMTS. Following the XYZ scheme, images are served through a web server, with a URL similar to

```
http://...//Z/X/Y.png
```

, where Z is is the zoom level, and X and Y identify the tile. GWC is released with version 3 of the GNU LGPL. The source code of GWC is available on GitHub[52]. GWC is an OSGeo community project.

The pre-production and caching of map tiles has become a standard practice in online map services since these processes use significantly fewer server resources than maps rendered on demand. This

---

[52] https://www.geowebcache.org/

practice has grown in popularity with web-mapping applications and the rapid growth of map data availability.

5.4 COLLABORATIVE DATA SHARING PLATFORM AND WEBGIS DESIGN AND IMPLEMENTATION

This section presents the design and implementation of the collaborative web platform for sharing the WQP maps. SIMILE project (Toro et al., 2019) served as the study case for testing the platform and WebGIS. Figure 5-1 presents the current implementation of the Collaborative Data-Sharing Platform and WebGIS. As mentioned, the architecture of the tools divides into two components working within an independent Docker environment. The components for the GeoNode platform dedicated to sharing the WQP maps of the lakes are shown on the right-hand side of the figure. Then, on the left-hand side of the figure, there is the WebGIS VueJS application. Both applications provide a client web-based environment which can be accessed through a browser to upload, edit and visualize geospatial data, more specifically, WQP maps. The use of open standards, WMS, supported achieving the communication in between the applications for the interactive visualization of the layers on the WebGIS.
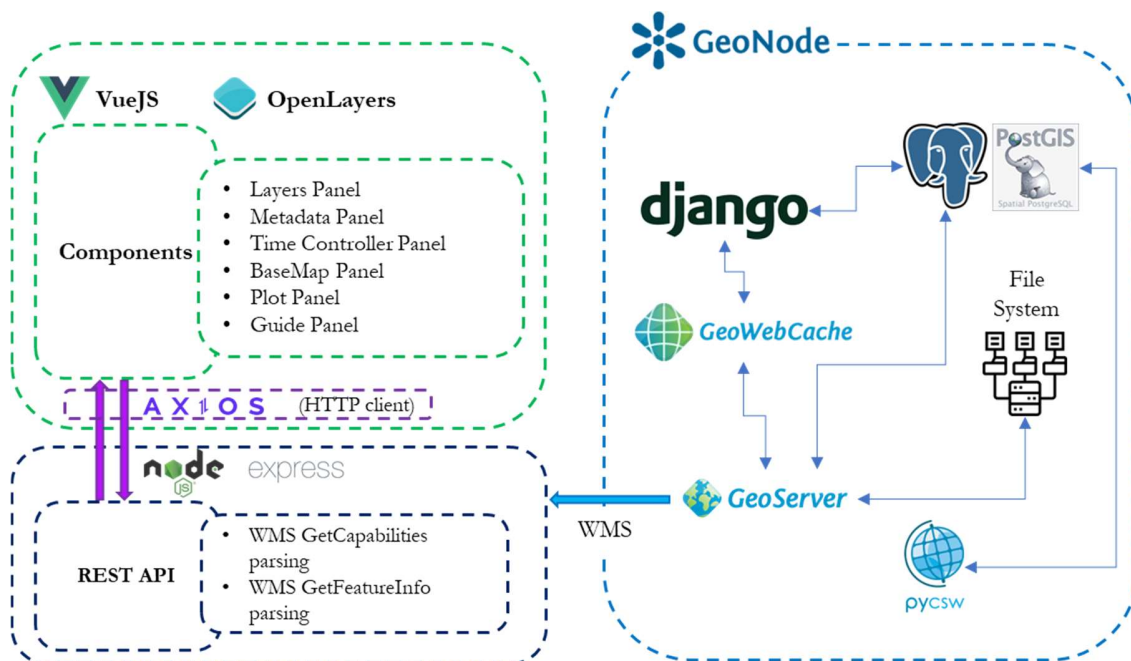
Figure 5-1 - Collaborative data-sharing platform architecture. WebGIS (left-side). GeoNode (right-side). SIMILE project

GeoNode is developed with scalability in mind. For this reason, the recommended installation method is through Docker[53] containers and the docker-compose[54] tool. For the purpose of the WQP maps data sharing, the solution benefits from this composition. In this work, the docker-compose implements different services, one for each component of the GeoNode stack, to deploy it into the following containers:

- nginx[55], on which the webserver runs;

- django[56], in which the Geonode Django web process runs;

- postgreSQL[57], on which PostgreSQL and its spatial extension, PostGIS are running;

- geoserver[58], on which GeoServer runs;

[53] https://www.docker.com/
[54] https://docs.docker.com/compose/
[55] https://www.nginx.com/
[56] https://www.djangoproject.com/
[57] https://www.postgresql.org/
[58] https://geoserver.org/

- celery[59], on which Celery runs. Celery is an open-source distributed task queue framework written in Python. It used for building distributed and asynchronous systems, primarily for handling background tasks or executing tasks in a distributed environment;

- rabbitmq[60], on which RabbitMQ runs. RabbitMQ is a widely used open-source message broker that facilitates communication between different systems or components by enabling asynchronous message passing;

- elasticsearch[61], optional, where the search engine runs.

On the other hand, the WebGIS, focuses on map visualization allowing exploration and navigability of the datasets available within the GeoNode map server. It aims to be a user-friendly environment for users to explore the products of water quality parameters published on the GeoNode application. When designing and implementing the WebGIS, the pattern followed a client-server architecture pattern. The design considers popular JavaScript based web application technologies, which provides benefits such as the use of the same language over the entire stack, simplifying development, code consistency and a rich ecosystem due to active communities.

First, the client-side component used VueJS[62], that an open-source model-view-viewmodel front end JavaScript framework for building user interfaces and single-page applications. The application is client-side rendered. It is released under the MIT License. The WebGIS implements OpenLayers, which is a popular web mapping framework. OpenLayers[63] is an open-source JavaScript library for displaying map data in web browsers. It provides a rich API for building rich web-based geographic applications similar to Google maps and Bing maps. Other relevant libraries for the development on

---

[59] https://github.com/celery/celery
[60] https://www.rabbitmq.com/
[61] https://www.elastic.co/
[62] https://vuejs.org/
[63] https://openlayers.org/

the client-side were Axios[64] (a promise-based HTTP client for Node.js and the browser) and Sass[65] (a pre-processor scripting language that is interpreted or compiled into CSS).

Second, the server-side components incorporate Node.js and Express.js. Node.js is a cross-platform open-source server and runtime environment for JavaScript. It serves as the backend server for the application, handling incoming requests and generating responses. It runs on the V8 JavaScript Engine, enabling the execution of JavaScript on the back end instead of a browser. On the other hand, Express.js[66] is a back-end application framework for building RESTful APIs and web applications with Node.JS released as free and open-source software under the MIT License. It simplifies the development of web applications by providing a robust set of features for routing, middleware, and server-side logic. An example of the middleware functions are the requests for the layers information available inside the server were handled by Express.js providing and serving the parsed data as an API endpoint to the VueJS app. In addition, Express.js enabled serving the static assets from the Vue.js application build process. The build process compiles the Vue.js components, transpiles the JavaScript code, and bundles the application's assets (CSS, images, fonts, etc.). Node.js and Express offer a lightweight and flexible server-side framework, allowing for rapid development and easy integration with other systems.

The interactive map uses OGC standards, such as WMS, to display maps published by map providers. The publication of datasets via OGC standard is available thanks to GeoServer on the backend of the GeoNode project. The WebGIS was implemented using a node.js runtime environment, which allows for creating server-side applications using JavaScript. The application was implemented using Vue.js, an open-source modelview-viewmodel front-end JavaScript framework for building user interfaces

---

[64] https://axios-http.com/
[65] https://sass-lang.com/
[66] https://expressjs.com/

and single-page applications. The use of Vue.JS aimed to make it easier for the website to render new layers as they are loaded into the platform . For example, an update of the time series list as the *getCapabilities* WMS request responds with a new set of layers after their upload. The distribution version is a build of the Vue.js app provided through Express.js. Express is a minimal and flexible Node.js web application framework that provides robust features for web and mobile applications. Express allows for the creation of RESTful APIs, which, in this case, allows importing the *getCapabilities* response from the map server and handling layer metadata on the server side and providing the data required to expose the time series on the client side. The design of the WebGIS benefits from the OpenLayers JavaScript web mapping library. The application includes features that allow users to control layer visibility (i.e. display and opacity) and navigate the map. The WebGIS retrieves metadata for each layer in two ways. The first relates to the metadata from the *getCapabilities* requests from GeoServer (such as the abstract and style). The second concerns the information deduced from the naming convention (i.e. sensor data, image acquisition timestamp, map type and native projection). The analysis of the response of the *getCapabilities* functions will allow the construction of the different time series based on the types of corresponding layers and the timestamp of the images. Furthermore, the WebGIS dynamically updates the set of layers via the *GetCapabilities* URL provided by GeoServer.

The deployment of the WebGIS involves the creation of a corresponding Docker composition. In this case, the docker composition accounts for a single container running an ExpressJS server for executing the WebGIS web application. The WMS server at hand for the project can be reached using the endpoint: https://www.geonode.eo.simile.polimi.it/geoserver/geonode/wms. An example of the WMS request from GeoServer can be defined as follows:

```
https://www.geonode.eo.simile.polimi.it/geoserver/geonode/wms?service=WMS&vers
ion=1.1.0&request=GetMap&layers=geonode:L8_LSWT_IT_20190101_L1&bbox=459765.0,5
```

```
063385.0,530325.0,5114025.0&width=768&height=551&srs=EPSG:32632&styles=&format
=application/openlayers
```

The example represents a GetMap request, from which obtains several pieces of information regarding the layer and the corresponding representation of it inside the WebGIS through the parameters:

- Layers: LSWT map (2019-01-01), found in the workspace named geonode;

- Bbox: a set of South, East, North, and West coordinates defining the bounding box for the requested layer;

- Width: width of the requested image in terms of pixels;

- Height: width of the requested image in terms of pixels;

- Srs: selected projection for representing the image;

- Format: application selected for representing the image through a web mapping library.

The WebGIS benefits from different methods to request the layers in the GeoNode platform using the OpenLayers library through its API:

- ol/layer/Image[67]

- ol/layer/Tile[68]

- ol/layer/Group[69]

The created map viewer in OpenLayers is a map object with extended functionalities for the widgets.

```
const map = new Map({
        target: 'map',
        controls: defaultControls()
        .extend([
                new ScaleLine(),
```

---

[67] https://openlayers.org/en/latest/apidoc/module-ol_layer_Image-ImageLayer.html
[68] https://openlayers.org/en/latest/apidoc/module-ol_layer_Tile-TileLayer.html
[69] https://openlayers.org/en/latest/apidoc/module-ol_layer_Group-LayerGroup

```
            overviewMapControl,
            new MenuControl(),
            new LayersControl(),
            new MetadataControl(),
            new TimeControl(),
            new BaseMapControl(),
            new PlotControl(),
            new SlideControl(),
            new ProjectLogo(),
            new GeonodeControl(),
    ]),
    view: new View({
    center: [994904.360160,5775887.605341],
    zoom: 10
    })
});
```

The design of the Collaborative Data-Sharing platform considered two critical challenges in the dissemination of scientific-based outputs from water quality monitoring. First, provide a tool to publish WQPs maps to monitor the lake's status, enabling the management of the uploaded datasets inside the platform. For management, it refers to the capability of uploading/deleting datasets, adding/editing/deleting the dataset's metadata and styles, and their publication through web services. The platform implementation took advantage of GeoNode as a tool to cover the data publication demands. Second, identify the platform users, either data providers or general users, who can have a different experience depending on their assigned privileges (see Table 5-1). Each user's privileges will determine the actions they can perform inside the platform. For example, only data providers, e.g. researchers, are in charge of the data upload of WQP maps, and the public can only access the final products.

User Role	WQP maps Upload	Create Groups	Create Users	Layers Management	Styles Management	Define Privileges	View Layers	Download Layers
Administrator	✓	✓	✓	✓	✓	✓	✓	✓
Data Provider	✓		✓	✓	✓	✓	✓	✓
General User							✓	✓

Table 5-1 - User's privileges in the platform according to their role.

Table 5-1 defines the expected roles inside the platform (i.e., administrator, data provider and general user). The administrator is in charge of the platform's maintenance; for this purpose, it has access to the complete set of functionalities existing inside the platform, including assigning privileges to other users. The Data Providers, or map producers, have a dedicated workspace and groups for the maps upload and the provision of styles and metadata. A group will enable identifying the platform source of the data and additional contact information of the map producers. In case more than one group exists in the platform, the Administrator can restrict the interaction to the uploaded data to the corresponding Data Provider. The Data Provider can define the way in which users can interact with the WQPs maps by defining the privileges of layers to the General Public, allowing them to visualize them or download them. The General Public will have access to all maps which have been made enabled by the data providers.

The GeoNode customized application for project SIMILE devotes to the producers of the WQP maps. The platform permits uploading layers, following the stated naming convention, along with additional metadata provided by filling out a form. Depending on the role of the providers, they have assigned privileges allowing them to modify or not the datasets. Also, the users can take advantage of introducing the preferred styling for a map typology, and the user must use the same nomenclature of table 5.3 for the new style definition.

WQP	Time Span	Number of Layers	File Format
CHL	Jan-01-2019 Dec-31-2022	587	GeoTiFF
TSM	Jan-01-2019 Dec-31-2022	587	GeoTiFF
LSWT	Jan-01-2019 Dec-31-2022	131	GeoTiFF

Table 5-2 – Datasets currently ingested into the platform

WQP	Projection	EPSG	Spatial Resolution	Pre-processing source
CHL	CH1903+/LV95 – Swiss CH1903+/LV95	2056	300m	Sentinel-3 A/B OLCI
	WGS84/UTM zone 32N	32362		
TSM	CH1903+/LV95 – Swiss CH1903+/LV95	2056	300m	Sentinel-3 A/B OLCI
	WGS84/UTM zone 32N	32362		
LSWT	CH1903+/LV95 – Swiss CH1903+/LV95	2056	100m	Landsat-8 TIRS
	WGS84/UTM zone 32N	32362		

Table 5-3 – Processed WQPs maps currently ingested into the platform

Layer name	Sensor	WQP
S3A_CHL_CH	Sentinel-3 A OLCI	CHL
S3B_CHL_IT	Sentinel-3 B OLCI	TSM
S3A_TSM_IT	Sentinel-3 A OLCI	CHL
S3B_TSM_IT	Sentinel-3 B OLCI	TSM
L8_LSWT_CH	Landsat-8 TIRS	LSWT
L8_LSWT_IT	Landsat-8 TIRS	LSWT

Table 5-4 – Processed WQPs maps nomenclature

The second application corresponds to a WebGIS. This application focuses on the visualization of the water quality parameters maps. The WebGIS aims at enabling the exploration and navigability of the datasets available within the map server (i.e. the uploaded water quality maps in the GeoServer instance of GeoNode). The WebGIS builds upon a Node.js runtime environment, using Express.js

middleware for serving the single-page application implemented on Vue.js. The WebGIS benefits from the OpenLayers web mapping library, Axios and Jquery framework to execute the WMS requests (i.e. *GetCapabilities*, *GetMap*, *GetFeatureInfo*, *GetLegend*) to the map server for the visualization and data acquisition of the georeferenced images.

The two applications have been composed independently in separate docker compositions. However, the use of open standards ensures communication between the docker containers. The WebGIS benefits from these standards to request the map server the information of the available layers uploaded into GeoNode.

(a)

(b)
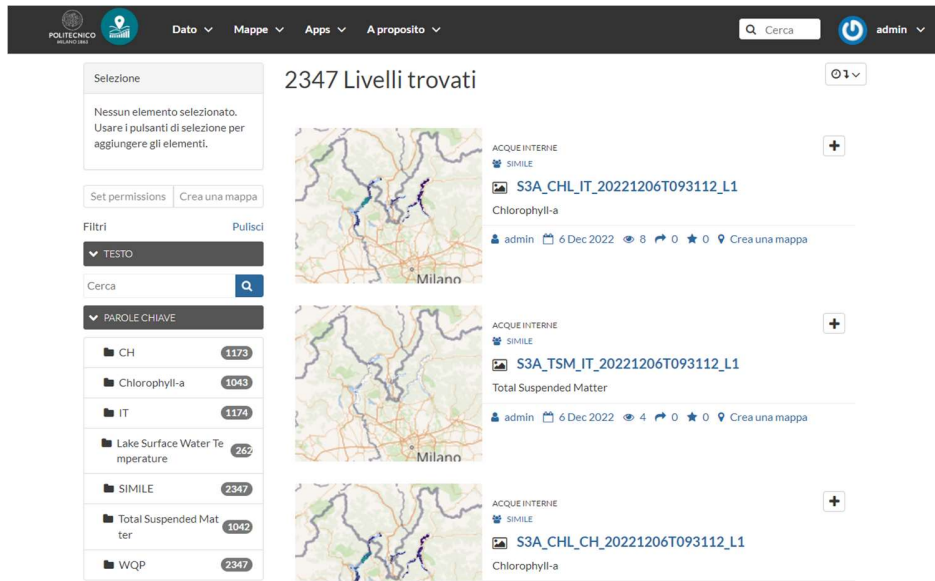
(c)

Figure 5-2 – Custom GeoNode platform[70]; (a) Welcome page GeoNode SIMILE; (b) Layers page GeoNode SIMILE; (c) Example: Layer page and metadata, December 06, 2022.

The WebGIS aims at being a user-friendly environment for the users to explore the water quality parameters products published on the GeoNode application. The interactive map uses OGS standards, such as Web Mapping Services (WMS), to display the published maps. The publication of the datasets through OGC standards is available thanks to GeoServer on the backend of the GeoNode

---

[70] https://www.geonode.eo.simile.polimi.it/

project. The WebGIS includes capabilities that allow the users to control the visibility of the layers (i.e. display and opacity), explore the metadata, download the datasets and navigate on the map. The WebGIS retrieves the metadata for each layer in two ways. The first is the one relative to the metadata available from the getCapabilities WMS request from GeoServer (such as the abstract and the style), and the second, regards the information which can be inferred from the naming convention detailed in Table 5-2 to Table 5-4 (i.e., sensor data, timestamp of image acquisition, map typology, and native projection). The parsing of the information of the *getCapabilities* functionalities takes place in the middleware, Express.js, which provides the data to the Vue.js application for building the different OpenLayers layers collection to recreate the time series for the matching layer typologies and the imagery timestamps. Furthermore, the WebGIS dynamically updates the set of layers through the getCapabilities URL provided by Geoserver.

In addition, in GeoNode allows developers or users to access specific resources or perform actions through the GeoNode API. The authentication to the API is via key is a security token used to authorize API requests. The key provides a secure mechanism to control access to its API and protect the resources and functionalities exposed through the API. It allows developers to integrate with GeoNode and build applications that interact with its data and services while enforcing proper authentication and authorization protocols. The API key determines the level of access and the operations that can be performed, such as retrieving geospatial data, updating metadata, uploading new resources, or managing user accounts. The API provides an alternative to ease the automatic batch update from existing layers. Here, an example of how it is possible to use a GET request the dataset available in the platform through the API's endpoint `https://www.geonode.eo.simile.polimi.it/api/v2/layers.` The response is a JSON formatted text with the information of the layers page inside the platform, including details on the metadata for the layers.

```
{
  "links": {
    "next": "http://www.geonode.eo.simile.polimi.it/api/v2/layers?page=2",
    "previous": null
  },
  "total": 2347,
  "page": 1,
  "page_size": 10,
  "layers": [
    {
      "pk": "2335",
      "uuid": "509bc946-cc11-11ed-91cf-0242ac180008",
      "name": "S3B_CHL_IT_20190818T095857_L1",
      "workspace": "geonode",
      "store": "S3B_CHL_IT_20190818T095857_L1",
      "storeType": "coverageStore",
      "charset": "UTF-8",
      "is_mosaic": false,
      "has_time": false,
      "has_elevation": false,
      "time_regex": "",
      "elevation_regex": null,
      "use_featureinfo_custom_template": false,
      "featureinfo_custom_template": null,
      "default_style": {
        "pk": 6755,
        "name": "chl",
        "workspace": "geonode",
        "sld_title": "chl",
        "sld_url":
"https://www.geonode.eo.simile.polimi.it/geoserver/rest/workspaces/geonode/styles/chl.sld"
      },
…
}
```

In this case, GeoNode's API key served for updating the uploaded layer's metadata. In order to update the layer's metadata the procedure implement a GET request to access the layers names, and a POST request to update the layer's metadata. The updated metadata included information inferred by their naming (i.e. CRS, date of acquisition, WQP map typology), layer category (e.g. inland waters) and keywords. Here, the metadata not only works to provide information about the layer, but also to ease the querying of specific products by the user inside the platform.

SIMILE dedicated WebGIS focuses on publishing raster datasets of water quality parameters maps as time series. GeoNode supports uploading raster datasets in GeoTIFF format, by taking advantage of the data storage system implemented by GeoServer. Notice that GeoServer provides data formats dedicated to multidimensional raster datasets (e.g., image mosaics and NetCDF). However, the direct interaction with the server contemplates a barrier in the data-sharing workflow (due to the need for FTP access to the server). To overcome the challenge of single image layers without a timestamp attribute, it was possible to use the timestamps available in the layer's name to execute the temporal queries. Then, for matching layer typologies, unique dates were extracted to build the group of layers. The groups of layers are produced as an OpenLayers LayerGroup object. In this way, it was possible to enable the time series visualization for the water quality parameters in the WebGIS and maintain GeoNode as a suitable tool for easing the publication of raster temporal data (See Figure 5-3).
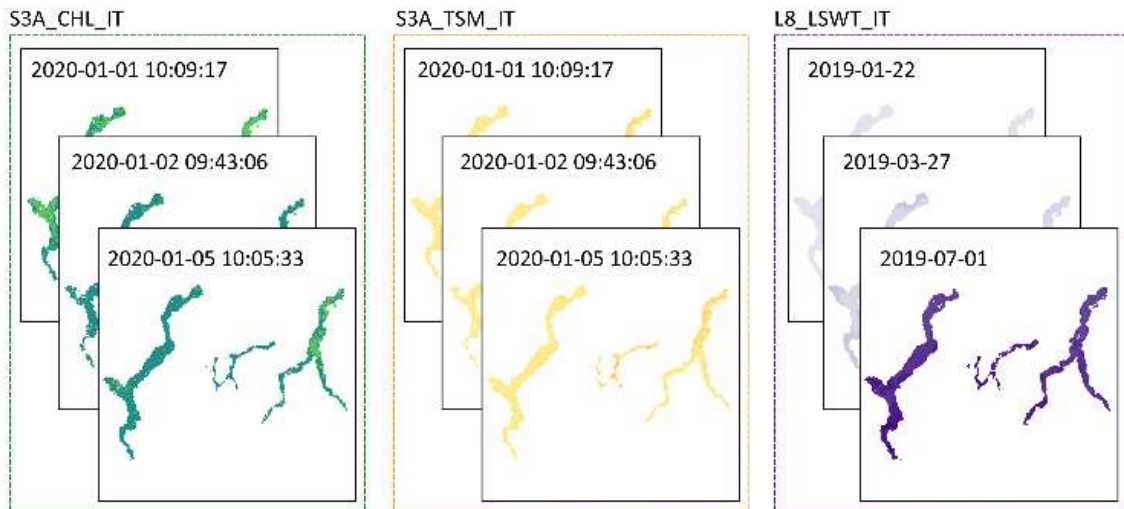
Figure 5-3 - Water quality parameters maps time series examples; OpenLayers LayerGroup

Figure 5-4 presents the different functionalities/widgets available inside the WebGIS which enable monitoring of the time series of the WQPs maps. The Welcome page in Figure 5-4(a) introduces the Menu bar to activate/deactivate the multiple functionalities available in the web application: Layer's Panel, Metadata Panel, Time Panel, Basemaps Panel, Plot Panel and Guide Panel. The Layer's Panels present the layers groups, or time series, for each WQP depending on the sensor used on the image acquisition (see Figure 5-4(b)). The Layer's Panel comprises different functions for the user to interact with the WQP maps, such as, (2) setting the layer visibility, (3) downloading the current layer, (4) requesting information on the WQP in the Metadata Panel and activating the Time Panel, (5) defining the layer opacity in the map viewer. Figure 5-4(c) introduces the Metadata Panel (1) and Time Panel (2). The Metadata Panel provides information on the most recently activated WQP map (i.e. description of the WQP, layer typology, WQP units, the period covered by the time series and legend). The Time Panel provides controls to navigate the time series and activate the layers by selecting one from the list of available dates, a slider Figure 5-4€, or step forwards/backwards buttons. Users can customize the maps viewer by clicking on the desired base map options (see Figure 5-4(f)). In addition, users can benefit from the implementation of pop-ups providing the estimates of the WQPs maps at
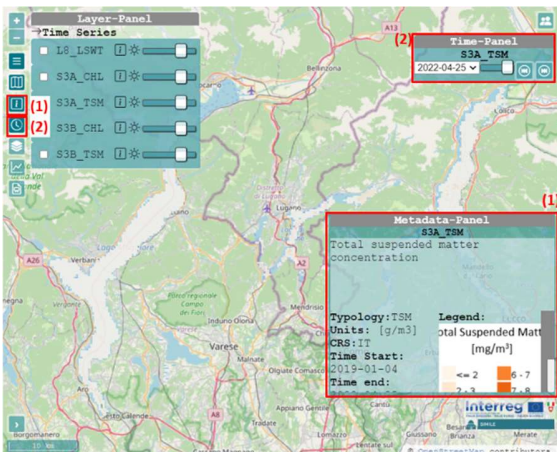
111

a specific location by clicking on them and explore the complete time series at the location if the Plot panel is active (see Figure 5-4 (g-h)). The last component of the WebGIS displays the Guide Panel, which hints the user on how to move around the web application.

(a)

(b)

(c)

(d)

(e)
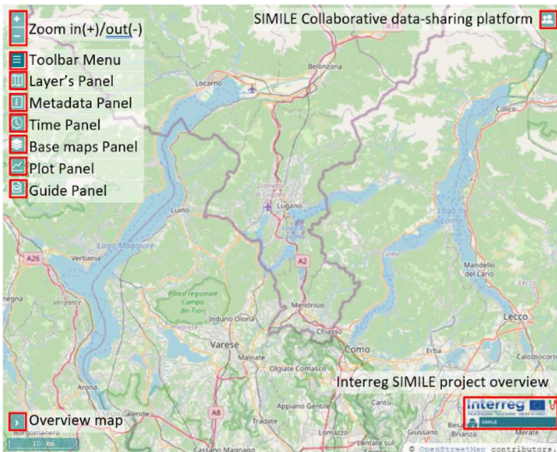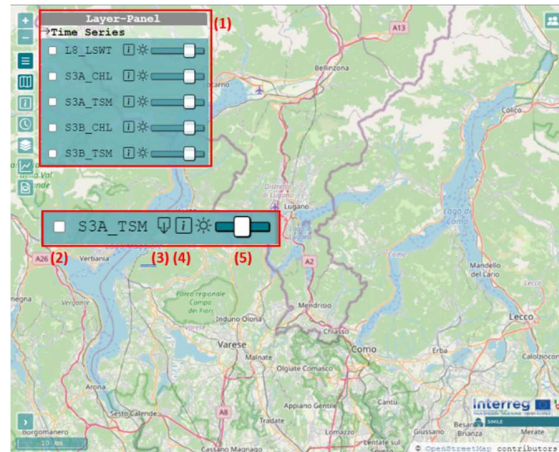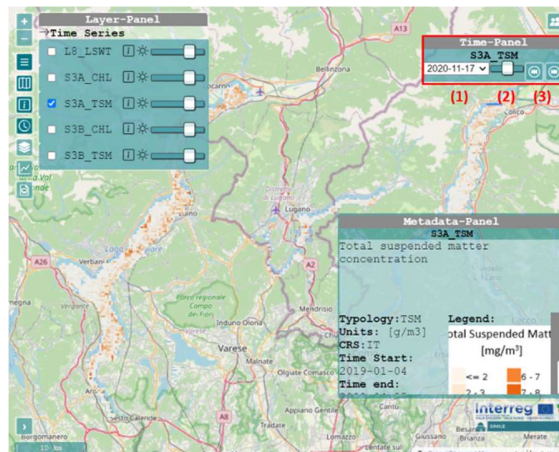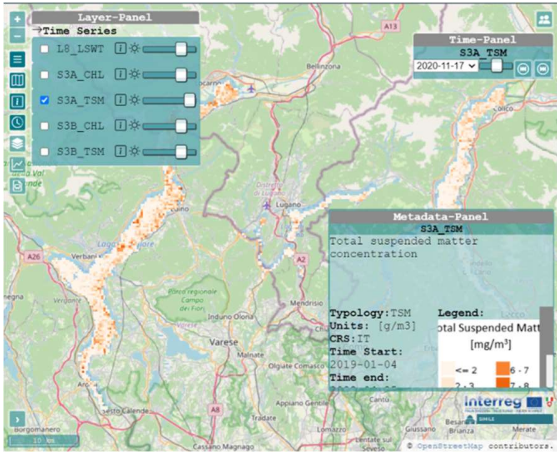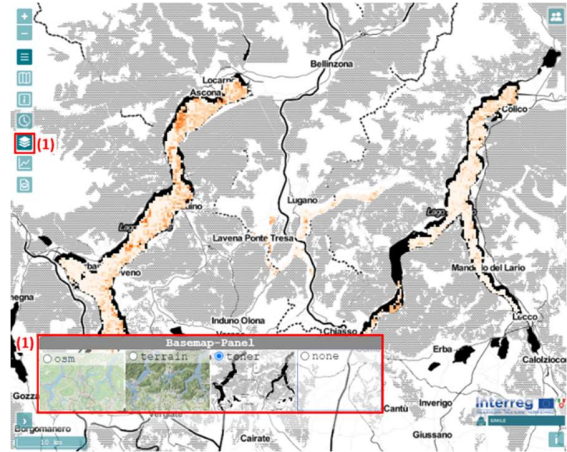
(f)
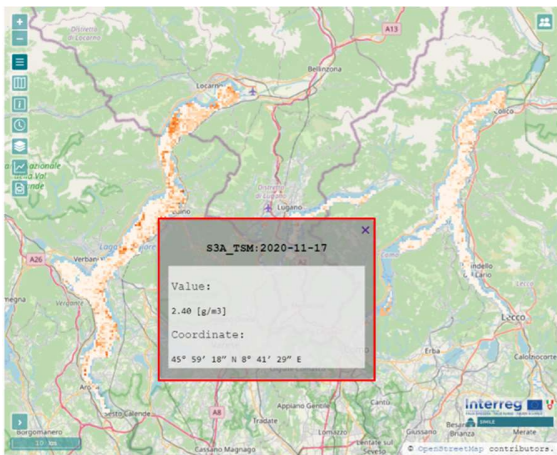
(g)
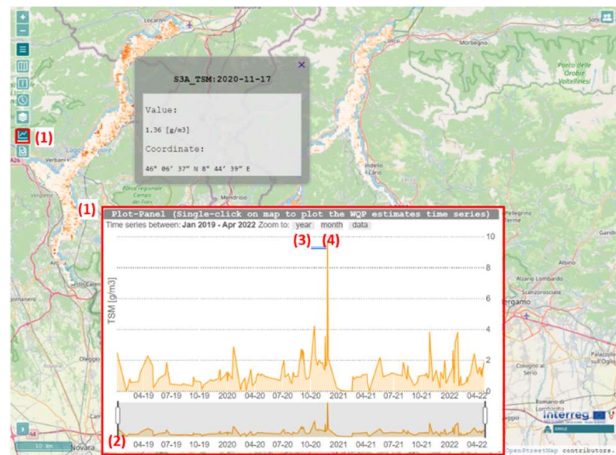
(h)
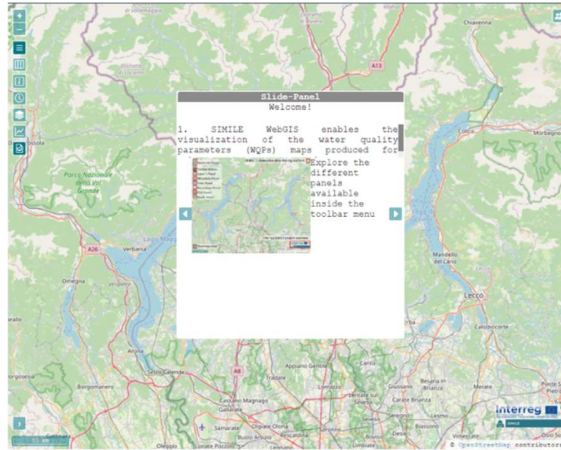
(i)

Figure 5-4 - WebGIS Panels; (a) WebGIS controllers; (b) Layers Panel; (c) 1. Metadata Panel, 2. Time Panel; (d) Time panel controls. 1. List of dates for the selected time series, 2. Time slider of the time series, 3. Step backward/forward controls; (e) Layer visualization; (f) Basemap Panel; (g) Pop-up. GetFeature info for the active layer; (h) Plot Panel; (i) Guide Panel.

Chapter - 6.   DISCUSSION AND CONCLUSIONS

Designing and implementing a system to share water quality maps is crucial for effective monitoring and management of water resources. This system aims to provide a comprehensive platform where users can access and visualize up-to-date information about water quality in different locations. The design process involves considering various factors such as data collection methods, data storage and management, visualization techniques, and user interface design. Implementing this system requires integrating data from multiple sources, developing data processing algorithms, and creating an intuitive web-based interface for users to explore and interact with the water quality maps. The successful design and implementation of such a system will contribute to better decision-making, resource planning, and environmental conservation efforts related to water quality.

The implemented collaborative platform provides a strategic tool for the dissemination and monitoring of water quality parameters maps. The platform presents two separate applications devoted to two different target audiences. First, there is the geospatial data-sharing platform, directed at the map producers. Here, the users will be able to upload the water quality maps, accompanied by their metadata and style. The uploaded dataset will follow a specific naming convention which opens the doors for the enriched temporal visualization on the second web-based application. On the other hand, the WebGIS provides a user-friendly environment for the visualisation of lake water quality maps produced from satellite imagery granting access to everyone. The application includes different modules such as the table of contents, the metadata panel and the time widget. Then, the users are capable to visualize the different sets of water quality parameters maps by controlling their visibility through the table of contents, explore the characteristics of a specific layer through the metadata panel and, most importantly, displaying the evolution of the different variables in time. The temporal visualization promotes the monitoring of different processes described by the water quality maps. For

example, the eutrophication level evolution in a lake characterised by the presence of chlorophyll-a, or the influence of climate change in the variation of the lake surface water temperature over the same seasons on different years. The monitoring and accessibility to the datasets will provide the users' information for assessing the health of the water bodies and their aquatic life. Moreover, the thorough monitoring of the water resources is a must to ensure the livelihood of the communities surrounding the lakes which depend on the consumption of the resource and other anthropic related activities.

In this study, we emphasize the utilization of a hybrid client architecture, where the client possesses limited processing capabilities and relies on the server's GIS functionalities. However, it is important to note that WebGIS is predominantly rendered on the client-side, benefiting from technologies that enable its functionality. The adoption of this architecture is favoured as it reduces the requirements on the client-side, while still relying on server-side operations or APIs to access the information pertaining to the Water Quality Parameter (WQP) maps. Hybrid client web GIS applications provide a harmonious equilibrium between extensive client-side functionality and the ability to leverage the resources offered by the server-side.

The platform aims at becoming a powerful tool for collaboration and publication of water quality parameters monitoring data. These benefits will enable not only the project partners and stakeholders, but also the community of citizens, to have a comprehensive understanding of the scope of the project that is the cross-border management of water resources.

The design and implementation of the platform has been tested under the framework of project SIMILE. The project addresses the use of multiple resources for water quality monitoring of the subalpine lakes in the cross-border area between Italy and Switzerland (specifically, Lakes Como, Lugano and Maggiore). The data sources analysed for the project derive from high-frequency in-situ sensor, citizen science and remote sensing. In this work, we present the approach adopted for

monitoring the lakes using remote sensing data. The WQP maps processing considers previous studies on the processing of optical satellite imagery for monitoring the lakes of interest. The presented processing chain has been built into a virtual environment with the different technologies and scripts required to replicate the procedure and ensure the continuous production of the maps. The sensors selection accounted for the desire of having a higher frequency monitoring of the lake with a robust methodology for estimating the WQPs maps. In this work, the processing chain considered the use of Sentinel 3 OLCI, for chlorophyll-a and total suspended matter (derived at least on a weekly basis), and Landsat 8 TIRS, for lake surface water temperatures (derived at least on a monthly basis).

In addition, the existing processing chain has integrated a set of algorithms to target two different challenges in the processing concerning the relative shift between the WQP maps derived from Sentinel 3 images and the identification of anomalous estimate after the processing for each WQP map. The discrepancy in image alignment can result in erroneous analysis when monitoring changes over time. This can hinder the identification of specific areas on a grid where estimates exhibit abnormal values. Consequently, resources may be misallocated, as efforts to manage water resources may be focused on incorrect areas. For instance, addressing high concentrations of chlorophyll associated with algae blooms may divert resources away from where they are truly needed. Furthermore, within the processing chain, the maps were clipped to the boundaries of the lakes. However, if there is misalignment between the maps and the actual positions of the lakes, this could result in the inadvertent removal of estimates from the generated maps. For this reason, it was implemented the co-registration algorithm into the processing chain known as GeFolki. Then, this study employs a statistical analysis of the Water Quality Parameter (WQP) estimates, utilizing descriptive statistics to identify values that can be considered abnormal. The statistical inference method employed in this analysis include the three-sigma method. The performance of the method was tested using in-situ measurement provided by a local environmental agency network of buoys

117

located in one of the lakes. Although, few matches between the dates of the measurements of the in-situ data and maps were found, it was possible identifying some cases in which the filter was capable of extracting pixels in the maps with large values.

Local measures, provided by SIMILE project partners, have been considered as reference as well as in situ reflectance measures. Ancillary data have been taken into account for complementing the metadata and support the analysis of the final WQPs maps such as: the sun light incidence to account for non-desirable effects in the processing such as glint, other adjacency effects due to the topographic configuration around the lake, visual examples of the previously mentioned disturbances to help the operator in the interpretation.

Outlier rejection has been applied to remove unacceptable values, trying to interpret, when possible, the reason for the anomalies in order to decide when the out-of-range values had to be included because due to peculiar conditions, such as high temperatures in small gulfs, or if they had to be discarded.

As further analysis, the spectral signatures rather than the final WQPs estimates are being considered to look for outliers. The principal component analysis is being exploited, to isolate disturbed pixels according to the spectral signature. The first results are encouraging, singling out the same outliers as the processors' flags with some useful addition. This will be investigated in the next future.

According to the literature (Bresciani et al., 2020; Free et al., 2021; Niroumand-Jadidi et al., 2021; Schütt et al., 2022; Soriano-González et al., 2022), the replicability of research, especially concerning case 2 water types bodies and inland water, can greatly support the enhancement of processing algorithms for satellite imagery. In the context of processing algorithms for satellite imagery, replicability allows other researchers and experts to test, verify, and validate the effectiveness and accuracy of these algorithms on different datasets and scenarios. In the case of case 2 water types of

bodies and inland water, which often present specific challenges due to their varying optical properties and inherent complexity, replicability is crucial. When multiple researchers and groups can replicate and validate the processing algorithms on different water bodies, it serves several important purposes such as benchmarking, identification of limitations, generalization, consistency in results and collaboration. Overall, replicability in research is fundamental for advancing the field of processing algorithms for satellite imagery, particularly when dealing with complex water types. It promotes transparency, rigor, and the refinement of algorithms, ultimately leading to more accurate and efficient methods for monitoring and analysing inland water bodies.

The work introduces a collection of open-source solutions designed for processing satellite images and facilitating the sharing of finalized map products. These solutions cater to both expert and non-expert users. For expert users, the implementation workflow involves a virtual machine that incorporates essential software and scripts required for satellite image processing. On the other hand, the project focuses on creating a platform to share geospatial data, enabling the time monitoring of various parameters. This platform offers diverse capabilities, allowing data producers and citizens alike to monitor dynamic geospatial variables. The resulting web applications provide a user-friendly interface for sharing geospatial data. The significance of this platform lies in its ability to accommodate both non-expert and expert users. As a case study, it has been tested for the project SIMILE, demonstrating its potential in monitoring geospatial variables of interest. Moreover, the provision of such tools is highly relevant to public administration, as it empowers decision-makers to make informed choices for the protection of water resources (Bresciani et al., 2018; Brovelli et al., 2019). The platform equips them with knowledge-based insights, aiding in better management and conservation of water resources.

REFERENCES

Agrawal, S., & Gupta, R. D. (2017). Web GIS and its architecture: A review. *Arabian Journal of Geosciences*, *10*(23), 518. https://doi.org/10.1007/s12517-017-3296-2

Amadori, M., Giovannini, L., Toffolon, M., Piccolroaz, S., Zardi, D., Bresciani, M., Giardino, C., Luciani, G., Kliphuis, M., van Haren, H., & Dijkstra, H. A. (2021). Multi-scale evaluation of a 3D lake model forced by an atmospheric model against standard monitoring data. *Environmental Modelling & Software*, *139*, 105017. https://doi.org/10.1016/j.envsoft.2021.105017

Arias Muñoz, C., Brovelli, M. A., Kilsedar, C. E., Moreno-Sanchez, R., & Oxoli, D. (2017). WEB MAPPING ARCHITECTURES BASED ON OPEN SPECIFICATIONS AND FREE AND OPEN SOURCE SOFTWARE IN THE WATER DOMAIN. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *IV-2-W4*, 23–30. https://doi.org/10.5194/isprs-annals-IV-2-W4-23-2017

Barsi, J. A., Schott, J. R., Palluconi, F. D., & Hook, S. J. (2005). *Validation of a web-based atmospheric correction tool for single thermal band instruments* (J. J. Butler, Ed.; p. 58820E). https://doi.org/10.1117/12.619990

Bhattacharya, D., Ghosh, J. K., Komarkova, J., Banerjee, S., & Kutoglu, H. (2014). Distributed GIS for automated natural hazard zonation mapping internet-SMS warning towards sustainable society. *Cogent Engineering*, *1*(1). https://doi.org/10.1080/23311916.2014.961346

Bhattacharya, D., & Painho, M. (2017). *SMART CITIES INTELLIGENCE SYSTEM (SMACiSYS) INTEGRATING SENSOR WEB with SPATIAL DATA INFRASTRUCTURES (SENSDI).* *4*(4W3), 21–28. Scopus. https://doi.org/10.5194/isprs-annals-IV-4-W3-21-2017

Boccardo, P., & Pasquali, P. (2012). *Web mapping services in a crowdsource environment for disaster management: State-of-the-art and further development. 39*, 543–548. Scopus.

Bresciani, M., Cazzaniga, I., Austoni, M., Sforzi, T., Buzzi, F., Morabito, G., & Giardino, C. (2018). Mapping phytoplankton blooms in deep subalpine lakes from Sentinel-2A and Landsat-8. *Hydrobiologia*, *824*(1), 197–214. https://doi.org/10.1007/s10750-017-3462-2

Bresciani, M., Pinardi, M., Free, G., Luciani, G., Ghebrehiwot, S., Laanen, M., Peters, S., Della Bella, V., Padula, R., & Giardino, C. (2020). The Use of Multisource Optical Sensors to Study Phytoplankton Spatio-Temporal Variation in a Shallow Turbid Lake. *Water*, *12*(1), Article 1. https://doi.org/10.3390/w12010284

Breunig, M., Bradley, P. E., Jahn, M., Kuper, P., Mazroob, N., Rösch, N., Al-Doori, M., Stefanakis, E., & Jadidi, M. (2020). Geospatial Data Management Research: Progress and Future Directions. *ISPRS International Journal of Geo-Information*, *9*(2), Article 2. https://doi.org/10.3390/ijgi9020095

Bricaud, A., Morel, A., Babin, M., Allali, K., & Claustre, H. (1998). Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters: Analysis and implications for bio-optical models. *Journal of Geophysical Research: Oceans*, *103*(C13), 31033–31044. https://doi.org/10.1029/98JC02712

Brigot, G., Colin-Koeniguer, E., Plyer, A., & Janez, F. (2016). Adaptation and evaluation of an optical flow method applied to coregistration of forest remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(7), 2923–2939. https://doi.org/10.1109/JSTARS.2016.2578362

Brockmann, C., Doerffer, R., Peters, M., Kerstin, S., Embacher, S., & Ruescas, A. (2016). Evolution of the C2RCC Neural Network for Sentinel 2 and 3 for the Retrieval of Ocean Colour Products in Normal and Extreme Optically Complex Waters. *Living Planet Symposium*, *740*, 54.

Brovelli, M. A., Cannata, M., & Rogora, M. (2019). SIMILE, A GEOSPATIAL ENABLER OF THE MONITORING OF SUSTAINABLE DEVELOPMENT GOAL 6 (ENSURE AVAILABILITY AND SUSTAINABILITY OF WATER FOR ALL). *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-4-W20*, 3–10. https://doi.org/10.5194/isprs-archives-XLII-4-W20-3-2019

Buonanno, S., Zeni, G., Fusco, A., Manunta, M., Marsella, M., Carrara, P., & Lanari, R. (2019a). A GeoNode-Based Platform for an Effective Exploitation of Advanced DInSAR Measurements. *Remote Sensing*, *11*(18), Article 18. https://doi.org/10.3390/rs11182133

Buonanno, S., Zeni, G., Fusco, A., Manunta, M., Marsella, M., Carrara, P., & Lanari, R. (2019b). A GeoNode-based platform for an effective exploitation of advanced DInSAR measurements. *Remote Sensing*, *11*(18). Scopus. https://doi.org/10.3390/rs11182133

Copetti, D., Guyennon, N., & Buzzi, F. (2020). Generation and dispersion of chemical and biological gradients in a large-deep multi-basin lake (Lake Como, north Italy): The joint effect of external drivers and internal wave motions. *Science of The Total Environment*, *749*, 141587. https://doi.org/10.1016/j.scitotenv.2020.141587

Corti, P., Bartoli, F., Fabiani, A., Giovando, C., Kralidis, A. T., & Tzotsos, A. (2019). *GeoNode: An open source framework to build spatial data infrastructures* (e27534v1). PeerJ Inc. https://doi.org/10.7287/peerj.preprints.27534v1

Cristofori, E. I., Balbo, S., Camaro, W., Pasquali, P., Boccardo, P., & Demarchi, A. (2015). *Flood risk web-mapping for decision makers: A service proposal based on satellite-derived precipitation analysis and geonode*. *2015-November*, 1389–1392. Scopus. https://doi.org/10.1109/IGARSS.2015.7326036

De Keukelaere, L., Sterckx, S., Adriaensen, S., Knaeps, E., Reusen, I., Giardino, C., Bresciani, M., Hunter, P., Neil, C., Van der Zande, D., & Vaiciute, D. (2018). Atmospheric correction of Landsat-8/OLI and Sentinel-2/MSI data using iCOR algorithm: Validation for coastal and

inland waters. *European Journal of Remote Sensing*, *51*(1), 525–542. https://doi.org/10.1080/22797254.2018.1457937

Delipetrev, B., Jonoski, A., & Solomatine, D. P. (2014). Development of a web application for water resources based on open source software. *Computers and Geosciences*, *62*, 35–42. Scopus. https://doi.org/10.1016/j.cageo.2013.09.012

Doerffer, R., & Schiller, H. (2007). The MERIS Case 2 water algorithm. *International Journal of Remote Sensing*, *28*(3–4), 517–535. https://doi.org/10.1080/01431160600821127

Duan, H., Ma, R., Xu, J., Zhang, Y., & Zhang, B. (2010). Comparison of different semi-empirical algorithms to estimate chlorophyll-a concentration in inland lake water. *Environmental Monitoring and Assessment*, *170*(1), 231–244. https://doi.org/10.1007/s10661-009-1228-7

Dyba, K., Ermida, S., Ptak, M., Piekarczyk, J., & Sojka, M. (2022). Evaluation of Methods for Estimating Lake Surface Water Temperature Using Landsat 8. *Remote Sensing*, *14*(15), Article 15. https://doi.org/10.3390/rs14153839

Erler, A. R., Frey, S. K., Khader, O., d'Orgeville, M., Park, Y.-J., Hwang, H.-T., Lapen, D. R., Peltier, W. R., & Sudicky, E. A. (2019). Evaluating Climate Change Impacts on Soil Moisture and Groundwater Resources Within a Lake-Affected Region. *Water Resources Research*, *55*(10), 8142–8163. https://doi.org/10.1029/2018WR023822

Ficek, D., Meler, J., Zapadka, T., Woźniak, B., & Dera, J. (2012). Inherent optical properties and remote sensing reflectance of Pomeranian lakes (Poland)**This paper was carried out within the framework of the SatBałtyk project funded by the European Union through European Regional Development Fund (contract No. POIG.01.01.02-22-011/09 entitled 'The Satellite Monitoring of the Baltic Sea Environment'). The partial support for this study was also provided by the MNiSW (Ministry of Science Higher Education) as a research project N N306 066434 in the years 2008–2011 also as a part of Pomeranian University IO PAS's statutory

research.The results published in this paper were presented at the international conference Ocean Optics XXI, Glasgow, Scotland, 8–12 October 2012. *Oceanologia*, *54*(4), 611–630. https://doi.org/10.5697/oc.54-4.611

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, I. C., Ramankutty, N., & Snyder, P. K. (2005). Global Consequences of Land Use. *Science*, *309*(5734), 570–574. https://doi.org/10.1126/science.1111772

Free, G., Bresciani, M., Pinardi, M., Ghirardi, N., Luciani, G., Caroni, R., & Giardino, C. (2021). Detecting Climate Driven Changes in Chlorophyll-a in Deep Subalpine Lakes Using Long Term Satellite Data. *Water*, *13*(6), Article 6. https://doi.org/10.3390/w13060866

Gholizadeh, M. H., Melesse, A. M., & Reddi, L. (2016). A Comprehensive Review on Water Quality Parameters Estimation Using Remote Sensing Techniques. *Sensors*, *16*(8), Article 8. https://doi.org/10.3390/s16081298

Giardino, C., Bresciani, M., Stroppiana, D., Oggioni, A., & Morabito, G. (2013). Optical remote sensing of lakes: An overview on Lake Maggiore. *Journal of Limnology*, *73*(s1). https://doi.org/10.4081/jlimnol.2014.817

Giardino, C., Bresciani, M., Valentini, E., Gasperini, L., Bolpagni, R., & Brando, V. E. (2015). Airborne hyperspectral data to assess suspended particulate matter and aquatic vegetation in a shallow and turbid lake. *Remote Sensing of Environment*, *157*, 48–57. https://doi.org/10.1016/j.rse.2014.04.034

Giuliani, G., Chatenoux, B., Piller, T., Moser, F., & Lacroix, P. (2020). Data Cube on Demand (DCoD): Generating an earth observation Data Cube anywhere in the world. *International*

Journal of Applied Earth Observation and Geoinformation, *87*, 102035. https://doi.org/10.1016/j.jag.2019.102035

Golladay, S. W., Craig, L. S., DePalma-Dow, A. D., Emanuel, B. N., & Rogers, S. G. (2021). Building resilience into water management through public engagement. *Freshwater Science*, *40*(1), 238–244. https://doi.org/10.1086/712514

Gordon, H. R., Brown, O. B., Evans, R. H., Brown, J. W., Smith, R. C., Baker, K. S., & Clark, D. K. (1988). A semianalytic radiance model of ocean color. *Journal of Geophysical Research: Atmospheres*, *93*(D9), 10909–10924. https://doi.org/10.1029/JD093iD09p10909

Graf, L. V., Perich, G., & Aasen, H. (2022). EOdal: An open-source Python package for large-scale agroecological research using Earth Observation and gridded environmental data. *Computers and Electronics in Agriculture*, *203*, 107487. https://doi.org/10.1016/j.compag.2022.107487

Groom, S., Sathyendranath, S., Ban, Y., Bernard, S., Brewin, R., Brotas, V., Brockmann, C., Chauhan, P., Choi, J., Chuprin, A., Ciavatta, S., Cipollini, P., Donlon, C., Franz, B., He, X., Hirata, T., Jackson, T., Kampel, M., Krasemann, H., … Wang, M. (2019). Satellite Ocean Colour: Current Status and Future Perspective. *Frontiers in Marine Science*, *6*. https://doi.org/10.3389/fmars.2019.00485

Guan, W. W., Bol, P. K., Lewis, B. G., Bertrand, M., Berman, M. L., & Blossom, J. C. (2012). WorldMap—A geospatial framework for collaborative research. *Annals of GIS*, *18*(2), 121–134. Scopus. https://doi.org/10.1080/19475683.2012.668559

Guida, L., Boccardo, P., Donevski, I., Lo Schiavo, L., Molinari, M. E., Monti-Guarnieri, A., Oxoli, D., & Brovelli, M. A. (2018). POST-DISASTER DAMAGE ASSESSMENT THROUGH COHERENT CHANGE DETECTION ON SAR IMAGERY. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII–3*, 431–436. https://doi.org/10.5194/isprs-archives-XLII-3-431-2018

Hadjimitsis, D. G., & Clayton, C. (2008). Assessment of temporal variations of water quality in inland water bodies using atmospheric corrected satellite remotely sensed image data. *Environmental Monitoring and Assessment*, *159*(1), 281. https://doi.org/10.1007/s10661-008-0629-3

Hafeez, S., Wong, M. S., Ho, H. C., Nazeer, M., Nichol, J., Abbas, S., Tang, D., Lee, K. H., & Pun, L. (2019). Comparison of Machine Learning Algorithms for Retrieval of Water Quality Indicators in Case-II Waters: A Case Study of Hong Kong. *Remote Sensing*, *11*(6), Article 6. https://doi.org/10.3390/rs11060617

Hestir, E. L., Brando, V. E., Bresciani, M., Giardino, C., Matta, E., Villa, P., & Dekker, A. G. (2015). Measuring freshwater aquatic ecosystems: The need for a hyperspectral global mapping satellite mission. *Remote Sensing of Environment*, *167*, 181–195. https://doi.org/10.1016/j.rse.2015.05.023

Hommersom, A., Kratzer, S., Laanen, M., Ansko, I., Ligi, M., Bresciani, M., Giardino, C., Beltran-Abaunza, J. M., Moore, G., Wernand, M. R., & Peters, S. W. (2012). Intercomparison in the field between the new WISP-3 and other radiometers (TriOS Ramses, ASD FieldSpec, and TACCS). *Journal of Applied Remote Sensing*, *6*(1), 063615. https://doi.org/10.1117/1.JRS.6.063615

Hopfensperger, K. N., Larson, E. I., Washko, S. E., & Moody, E. K. (2021). Elevate your work through incorporation of public engagement. *Freshwater Science*, *40*(1), 221–227. https://doi.org/10.1086/712607

Hyun, C.-U., Park, M., & Lee, W. Y. (2020). Remotely Piloted Aircraft System (RPAS)-Based Wildlife Detection: A Review and Case Studies in Maritime Antarctica. *Animals : An Open Access Journal from MDPI*, *10*(12), 2387. https://doi.org/10.3390/ani10122387

Kristollari, V., & Karathanassi, V. (2022). Change Detection in VHR Imagery With Severe Co-Registration Errors Using Deep Learning: A Comparative Study. *IEEE Access*, *10*, 33723–33741. https://doi.org/10.1109/ACCESS.2022.3161978

Kulawiak, M., & Lubniewski, Z. (2014). SafeCity—A GIS-based tool profiled for supporting decision making in urban development and infrastructure protection. *Technological Forecasting and Social Change*, *89*, 174–187. Scopus. https://doi.org/10.1016/j.techfore.2013.08.031

Kyryliuk, D., & Kratzer, S. (2019). Evaluation of Sentinel-3A OLCI Products Derived Using the Case-2 Regional CoastColour Processor over the Baltic Sea. *Sensors*, *19*(16). https://doi.org/10.3390/s19163609

Landsat 8 | Landsat Science. (2021, November 30). https://landsat.gsfc.nasa.gov/satellites/landsat-8/

Landsat 8 (L8) Data Users Handbook. (2019).

Lavender, S. (2021). *Sentinel-3 OLCI Marine User Handbook*.

Leprince, S., Barbot, S., Ayoub, F., & Avouac, J.-P. (2007). Automatic and Precise Orthorectification, Coregistration, and Subpixel Correlation of Satellite Images, Application to Ground Deformation Measurements. *IEEE Transactions on Geoscience and Remote Sensing*, *45*(6), 1529–1558. https://doi.org/10.1109/TGRS.2006.888937

Liberzon, M. R. (2019). Technology geonod: Status and prospects of automation and control. *Mekhatronika, Avtomatizatsiya, Upravlenie*, *20*(11), 696–701. Scopus. https://doi.org/10.17587/mau.20.696-701

Lucas, B. D., & Kanade, T. (1981). *An Iterative Image Registration Technique with an Application to Stereo Vision*.

Luciani, G., Bresciani, M., Biraghi, C. A., Ghirardi, N., Carrion, D., Rogora, M., & Brovelli, M. A. (2021). Satellite Monitoring system of Subalpine lakes with open source software: The case of

SIMILE Project. *Baltic Journal of Modern Computing*, *9*(1). https://doi.org/10.22364/bjmc.2021.9.1.08

McKinnon, M., & Vos, J. (2015). Engagement as a Threshold Concept for Science Education and Science Communication. *International Journal of Science Education, Part B*, *5*(4), 297–318. https://doi.org/10.1080/21548455.2014.986770

Mellander, P.-E., & Jordan, P. (2021). Charting a perfect storm of water quality pressures. *Science of The Total Environment*, *787*, 147576. https://doi.org/10.1016/j.scitotenv.2021.147576

Menegon, S., Sarretta, A., Depellegrin, D., Farella, G., Venier, C., & Barbanti, A. (2018). Tools4MSP: An open source software package to support Maritime Spatial Planning. *PeerJ Computer Science*, *2018*(10). Scopus. https://doi.org/10.7717/peerj-cs.165

Mobley, C. (2020). *Apparent Optical Properties: Ocean Optics Web Book.* https://www.oceanopticsbook.info/view/inherent-and-apparent-optical-properties/apparent-optical-properties

Mobley, C. (2021). *Inherent Optical Properties: Ocean Optics Web Book.* https://www.oceanopticsbook.info/view/inherent-and-apparent-optical-properties/inherent-optical-properties

Mobley, C. D. (2001). Radiative Transfer in the Ocean. In J. H. Steele (Ed.), *Encyclopedia of Ocean Sciences* (pp. 2321–2330). Academic Press. https://doi.org/10.1006/rwos.2001.0469

Morel, A. (2001). Bio-optical Models. In J. H. Steele (Ed.), *Encyclopedia of Ocean Sciences* (pp. 317–326). Academic Press. https://doi.org/10.1006/rwos.2001.0407

Morel, A., Gentili, B., Claustre, H., Babin, M., Bricaud, A., Ras, J., & Tièche, F. (2007). Optical properties of the "clearest" natural waters. *Limnology and Oceanography*, *52*(1), 217–229. https://doi.org/10.4319/lo.2007.52.1.0217

Morel, A., & Maritorena, S. (2001). Bio-optical properties of oceanic waters: A reappraisal. *Journal of Geophysical Research: Oceans*, *106*(C4), 7163–7180. https://doi.org/10.1029/2000JC000319

Morel, A. Y., & Gordon, H. R. (1980). Report of the working group on water color. *Boundary-Layer Meteorology*, *18*(3), 343–355. Scopus. https://doi.org/10.1007/BF00122030

Niroumand-Jadidi, M., Bovolo, F., & Bruzzone, L. (2019). Novel Spectra-Derived Features for Empirical Retrieval of Water Quality Parameters: Demonstrations for OLI, MSI, and OLCI Sensors. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(12), 10285–10300. https://doi.org/10.1109/TGRS.2019.2933251

Niroumand-Jadidi, M., Bovolo, F., Bruzzone, L., & Gege, P. (2021). Inter-Comparison of Methods for Chlorophyll-a Retrieval: Sentinel-2 Time-Series Analysis in Italian Lakes. *Remote Sensing*, *13*(12). https://doi.org/10.3390/rs13122381

OC4Me Chlorophyll—Sentinel Online. (2023, February 5). https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-3-olci/level-2/oc4me-chlorophyll

Ocean colour services | EUMETSAT. (2018, July 27). https://www.eumetsat.int/ocean-colour-services

Odermatt, D., Gitelson, A., Brando, V. E., & Schaepman, M. (2012). Review of constituent retrieval in optically deep and complex waters from satellite imagery. *Remote Sensing of Environment*, *118*, 116–126. https://doi.org/10.1016/j.rse.2011.11.013

Ogashawara, I. (2015). Terminology and classification of bio-optical algorithms. *Remote Sensing Letters*, *6*(8), 613–617. https://doi.org/10.1080/2150704X.2015.1066523

Ogashawara, I., Mishra, D. R., & Gitelson, A. A. (2017). Chapter 1 - Remote Sensing of Inland Waters: Background and Current State-of-the-Art. In D. R. Mishra, I. Ogashawara, & A. A. Gitelson (Eds.), *Bio-optical Modeling and Remote Sensing of Inland Waters* (pp. 1–24). Elsevier. https://doi.org/10.1016/B978-0-12-804644-9.00001-X

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., Kahru, M., & McClain, C. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research: Oceans*, *103*(C11), 24937–24953. https://doi.org/10.1029/98JC02160

Ostrom, E., Stern, P. C., & Dietz, T. (2003). Water Rights in the Commons. *Water Resources IMPACT*, *5*(2), 9–12.

Overstreet, B. T., & Legleiter, C. J. (2017). Removing sun glint from optical remote sensing images of shallow rivers. *Earth Surface Processes and Landforms*, *42*(2), 318–333. https://doi.org/10.1002/esp.4063

Pokhrel, Y. N., Felfelani, F., Shin, S., Yamada, T. J., & Satoh, Y. (2017). Modeling large-scale human alteration of land surface hydrology and climate. *Geoscience Letters*, *4*(1), 10. https://doi.org/10.1186/s40562-017-0076-5

Preisendorfer, R. W., Laboratory (U.S.), P. M. E., & Laboratories (U.S.), E. R. (1976). *Hydrologic Optics*. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Research Laboratories, Pacific Marine Environmental Laboratory. https://books.google.it/books?id=9B0jxAEACAAJ

Rudiastuti, A. W., Suryanegara, E., Wirawan, A., Purwanto, B., Gill, S., Prihanto, Y., Windiastuti, R., & Munawaroh. (2020). Design and implementation of a user-centered web-app using open source platform: Indonesia disaster data (InDITA). *International Journal on Informatics Visualization*, *4*(4), 243–249. Scopus. https://doi.org/10.30630/joiv.4.4.460

Schütt, E. M., Lehmann, M. K., Hieronymi, M., Dare, J., Krasemann, H., Hitchcock, D., Platt, A., Amai, K., & McKelvey, T. (2022). Dataset of five years of in situ and satellite derived chlorophyll a concentrations and its spatiotemporal variability in the Rotorua Te Arawa Lakes, New Zealand. *Data in Brief*, *40*, 107759. https://doi.org/10.1016/j.dib.2021.107759

Sentinel-3 Toolbox (S3TBX). (2023). [Java]. SNAP - ESA's SentiNel Application Platform. https://github.com/senbox-

org/s3tbx/blob/fa8ff45fe5f9c05621d35121e2d3c4e4e8a3d890/s3tbx-

c2rcc/src/main/java/org/esa/s3tbx/c2rcc/msi/C2rccMsiOperator.java (Original work published 2014)

Sharaf, N., Fadel, A., Bresciani, M., Giardino, C., Lemaire, B. J., Slim, K., Faour, G., & Vinçon-Leite, B. (2019). Lake surface temperature retrieval from Landsat-8 and retrospective analysis in Karaoun Reservoir, Lebanon. *Journal of Applied Remote Sensing, 13*(4), 044505. https://doi.org/10.1117/1.JRS.13.044505

Sistema Informativo per il Monitoraggio Integrato dei Laghi insubrici e dei loro Ecosistemi. (2022). https://progetti.interreg-

italiasvizzera.eu/it/b/78/sistemainformativoperilmonitoraggiointegratodeilaghiinsubriciedeil oroe

SNAP Data Processors—C2RCC OLCI Processing Parameters. (n.d.). Retrieved 30 January 2023, from https://seadas.gsfc.nasa.gov/help-8.2.0/c2rcc/C2RCC_OLCI_ProcParameters.html

Soomets, T., Uudeberg, K., Jakovels, D., Brauns, A., Zagars, M., & Kutser, T. (2020). Validation and Comparison of Water Quality Products in Baltic Lakes Using Sentinel-2 MSI and Sentinel-3 OLCI Data. *Sensors, 20*(3), Article 3. https://doi.org/10.3390/s20030742

Soriano-González, J., Urrego, E. P., Sòria-Perpinyà, X., Angelats, E., Alcaraz, C., Delegido, J., Ruíz-Verdú, A., Tenjo, C., Vicente, E., & Moreno, J. (2022). Towards the Combination of C2RCC Processors for Improving Water Quality Retrieval in Inland and Coastal Areas. *Remote Sensing, 14*(5). https://doi.org/10.3390/rs14051124

Spyrakos, E., O'Donnell, R., Hunter, P. D., Miller, C., Scott, M., Simis, S. G. H., Neil, C., Barbosa, C. C. F., Binding, C. E., Bradt, S., Bresciani, M., Dall'Olmo, G., Giardino, C., Gitelson, A. A.,

Kutser, T., Li, L., Matsushita, B., Martinez-Vicente, V., Matthews, M. W., … Tyler, A. N. (2018). Optical types of inland and coastal waters. *Limnology and Oceanography*, *63*(2), 846–870. https://doi.org/10.1002/lno.10674

Stumpf, A., Michéa, D., & Malet, J.-P. (2018). Improved Co-Registration of Sentinel-2 and Landsat-8 Imagery for Earth Surface Motion Measurements. *Remote Sensing*, *10*(2), Article 2. https://doi.org/10.3390/rs10020160

Terradue. (2022). *Gefolki anaconda package* [Python]. terradue. https://github.com/Terradue/gefolki/blob/5aa2cc03aaeb6dd159968bd52a812707e432ba6e/documents/manual_gefolki_english.pdf (Original work published 2019)

The Drama of the Commons. (2002). National Academies Press. https://doi.org/10.17226/10287

Toming, K., Kutser, T., Uiboupin, R., Arikas, A., Vahter, K., & Paavel, B. (2017). Mapping Water Quality Parameters with Sentinel-3 Ocean and Land Colour Instrument imagery in the Baltic Sea. *Remote Sensing*, *9*(10). https://doi.org/10.3390/rs9101070

Topp, S. N., Pavelsky, T. M., Jensen, D., Simard, M., & Ross, M. R. V. (2020). Research Trends in the Use of Remote Sensing for Inland Water Quality Science: Moving Towards Multidisciplinary Applications. *Water*, *12*(1). https://doi.org/10.3390/w12010169

Toro Herrera, J. F., Carrion, D., Bresciani, M., & Bratić, G. (2022). SEMI-AUTOMATED PRODUCTION AND FILTERING OF SATELLITE DERIVED WATER QUALITY PARAMETERS. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLIII-B3-2022*, 1019–1026. https://doi.org/10.5194/isprs-archives-XLIII-B3-2022-1019-2022

Toro, J. F., Carrion, D., Albertella, A., & Brovelli, M. A. (2019). CROSS-BORDER OPEN DATA SHARING: GIOCONDA PROJECT. *ISPRS - International Archives of the Photogrammetry,*
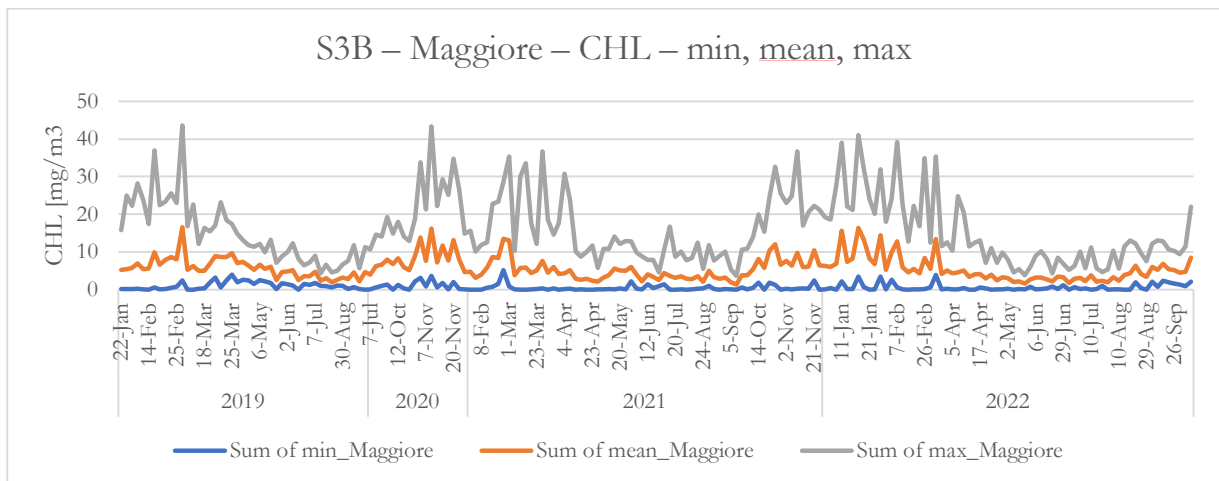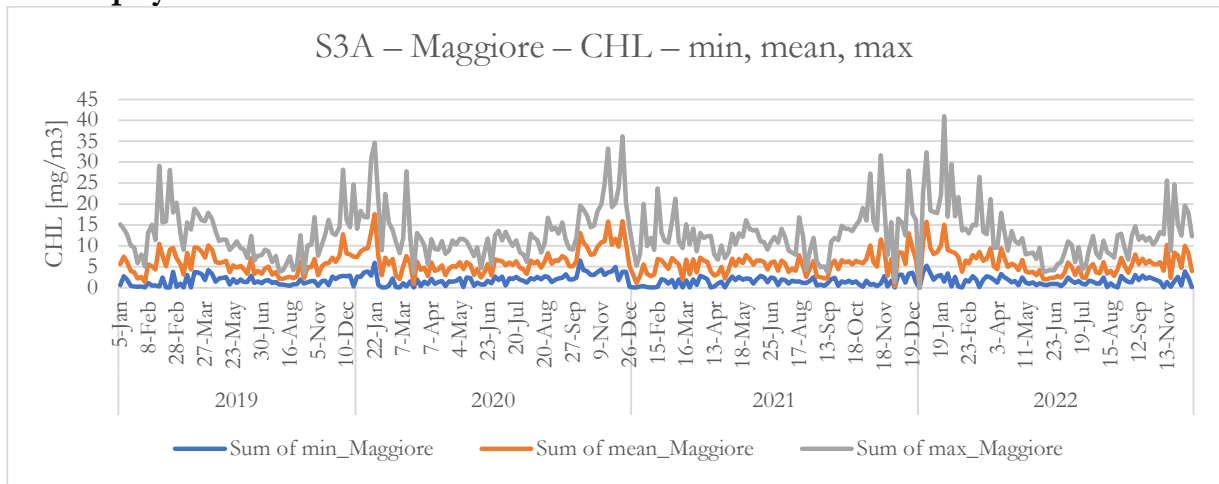
Remote Sensing and Spatial Information Sciences, *XLII-4-W14*, 233–238. https://doi.org/10.5194/isprs-archives-XLII-4-W14-233-2019

Trifonov, G. M., Poyda, A. A., Andreev, A. A., Prosvetov, A. V., & Zhizhin, M. N. (2018). Nocturnal Lights: Object-oriented catalog of the nighttime Earth light sources. *Procedia Computer Science*, *136*, 91–98. https://doi.org/10.1016/j.procs.2018.08.241

Tulokhonov, A. K., Suochen, D., Baklanov, P. Ya., Garmaev, E. Zh., Plyusnin, V. M., Vladimirov, I. N., Batuev, A. R., Voronov, B. A., & Beshentsev, A. N. (2020). *Atlas of sustainable development of north asia in the context of the project «one belt, one road»*. *26*, 352–360. Scopus. https://doi.org/10.35595/2414-9179-2020-1-26-352-360

UNEP-DHI Centre for Water and Environment. (2009). *Integrated water resources management in action*. https://unesdoc.unesco.org/ark:/48223/pf0000181891

UN-Water. (2021). *Summary Progress Update 2021: SDG 6 — water and sanitation for all*. UN-Water. https://www.unwater.org/publications/summary-progress-update-2021-sdg-6-water-and-sanitation-all

User Guides—Sentinel-3 OLCI - Radiometric Resolution—Sentinel Online—Sentinel Online. (2023, February 1). https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-3-olci/resolutions/radiometric

Vantrepotte, V., Loisel, H., Dessailly, D., & Mériaux, X. (2012). Optical classification of contrasted coastal waters. *Remote Sensing of Environment*, *123*, 306–323. https://doi.org/10.1016/j.rse.2012.03.004

Vavassori, A., Oxoli, D., & Brovelli, M. A. (2022). Population Space–Time Patterns Analysis and Anthropic Pressure Assessment of the Insubric Lakes Using User-Generated Geodata. *ISPRS International Journal of Geo-Information*, *11*(3), Article 3. https://doi.org/10.3390/ijgi11030206

Warner, S., Chapman, D., & Dickens, C. (2020). Good Ambient Water Quality. In W. Leal Filho, A. M. Azul, L. Brandli, A. Lange Salvia, & T. Wall (Eds.), *Clean Water and Sanitation* (pp. 1–11). Springer International Publishing. https://doi.org/10.1007/978-3-319-70061-8_83-1

Warren, M. A., Simis, S. G. H., Martinez-Vicente, V., Poser, K., Bresciani, M., Alikas, K., Spyrakos, E., Giardino, C., & Ansper, A. (2019). Assessment of atmospheric correction algorithms for the Sentinel-2A MultiSpectral Imager over coastal and inland waters. *Remote Sensing of Environment*, *225*, 267–289. https://doi.org/10.1016/j.rse.2019.03.018

Warren, M. A., Simis, S. G. H., & Selmes, N. (2021). Complementary water quality observations from high and medium resolution Sentinel sensors by aligning chlorophyll-a and turbidity algorithms. *Remote Sensing of Environment*, *265*, 112651. https://doi.org/10.1016/j.rse.2021.112651

Weingart, P., Joubert, M., & Connoway, K. (2021). Public engagement with science—Origins, motives and impact in academic literature and science policy. *PLOS ONE*, *16*(7), e0254201. https://doi.org/10.1371/journal.pone.0254201

Zuhlke, M., Fomferra, N., Brockmann, C., Peters, M., Veci, L., Malik, J., & Regner, P. (2015). *SNAP (Sentinel Application Platform) and the ESA Sentinel 3 Toolbox. 734*, 21.

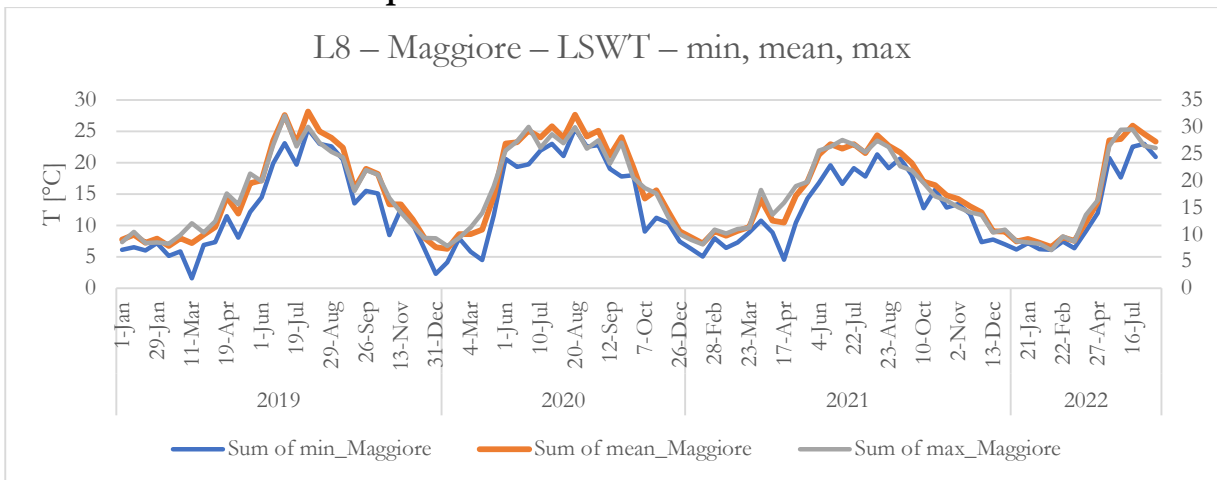# APPENDIX – WQP TIME SERIES DESCRIPTIVE STATISTICS

A.1.   LAKE MAGGIORE

Chlorophyll-a

Total Suspended Matter

S3A – Maggiore – TSM – min, mean, max

S3B – Maggiore – TSM – min, mean, max

Lake Surface Water Temperature

L8 – Maggiore – LSWT – min, mean, max
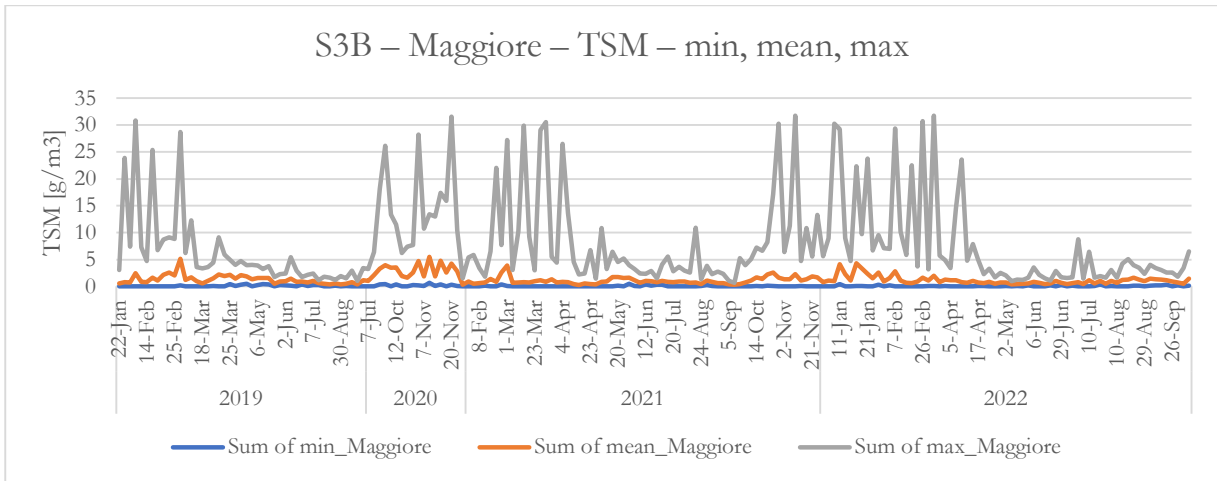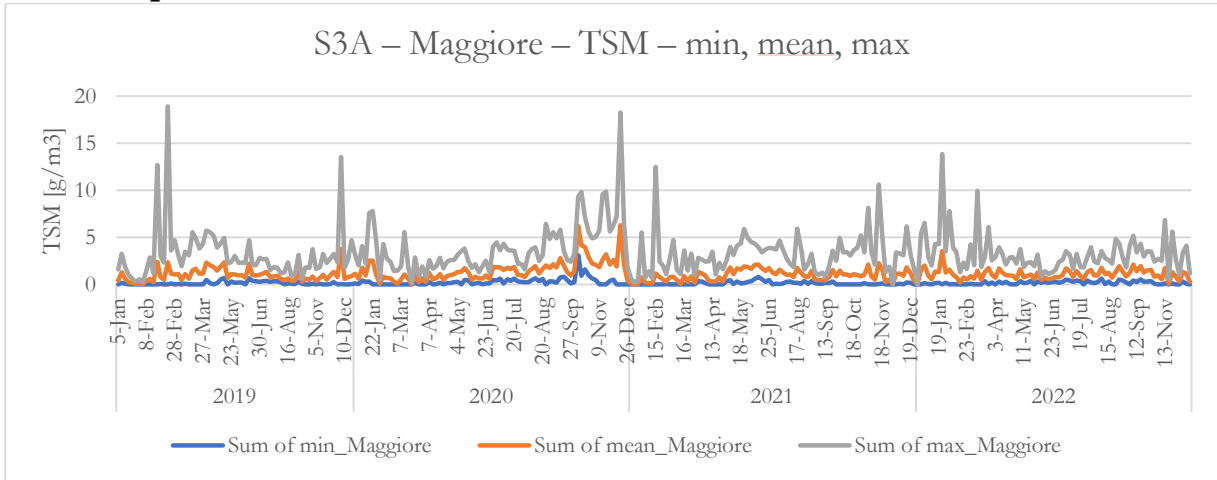
A.2.   LAKE COMO

Chlorophyll-a

S3A – Como – CHL – min, mean, max

S3B – Como – CHL – min, mean, max
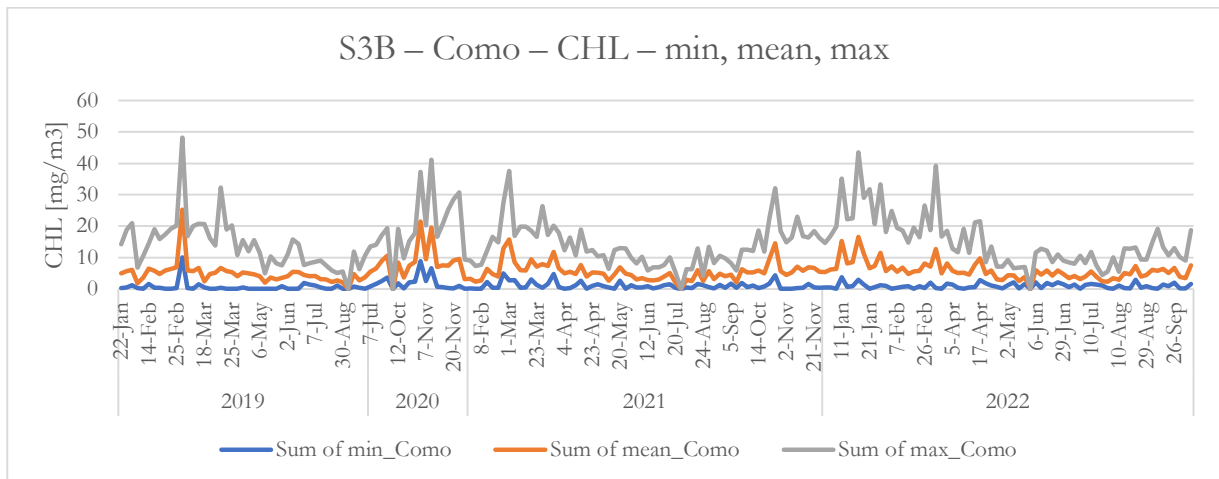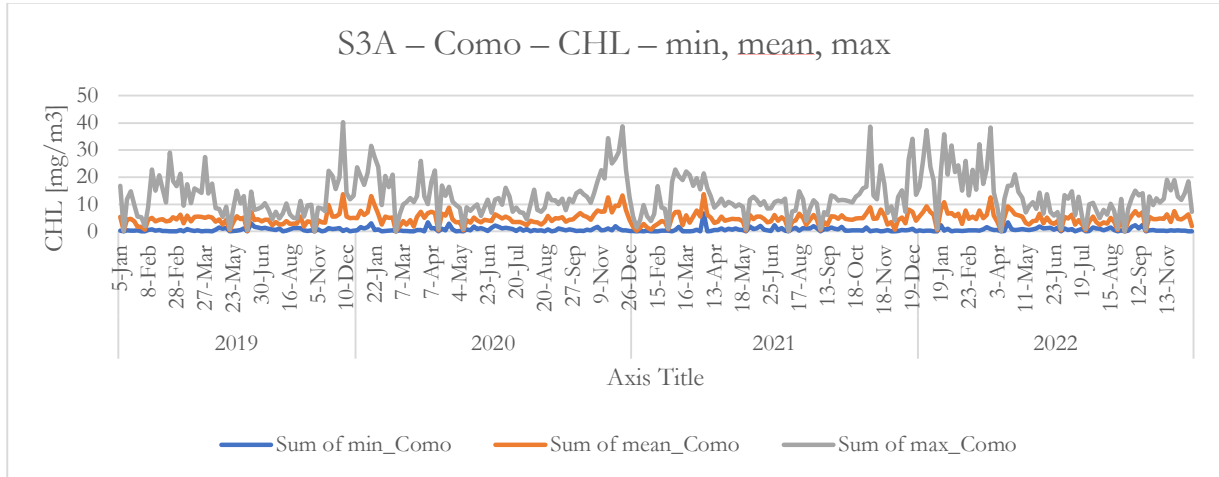
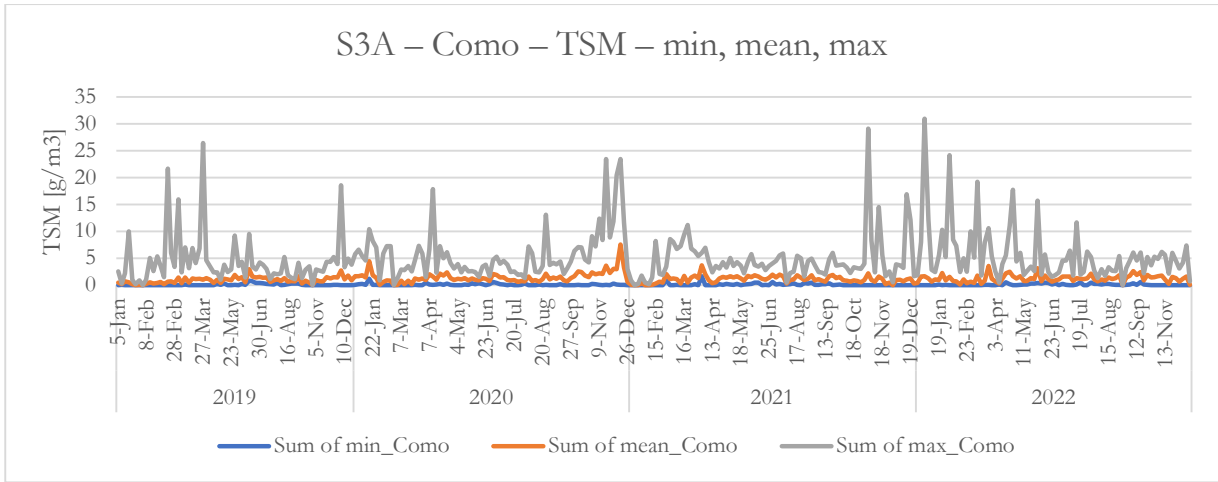# Total Suspended Matter
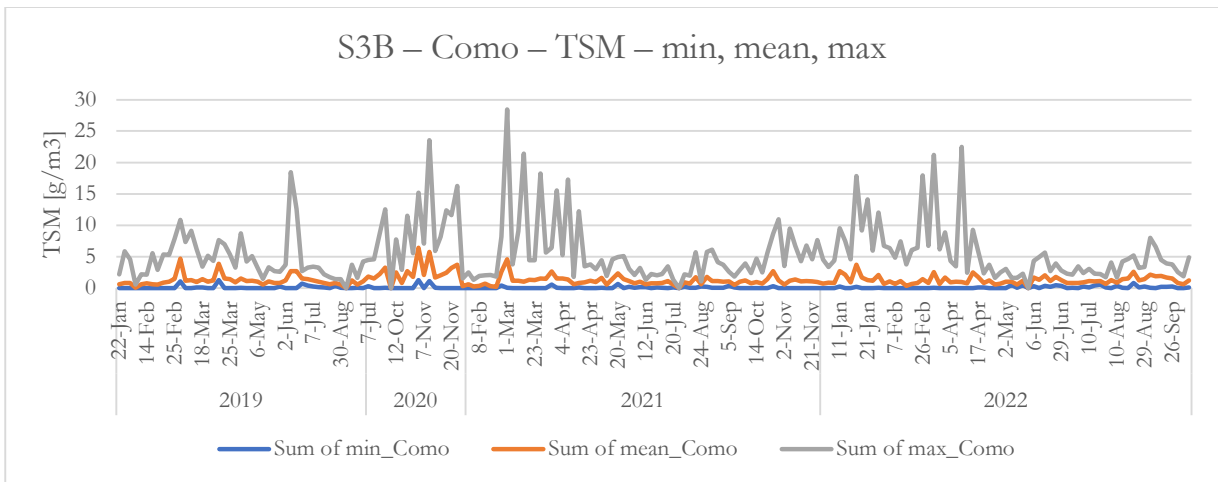
Lake Surface Water Temperature
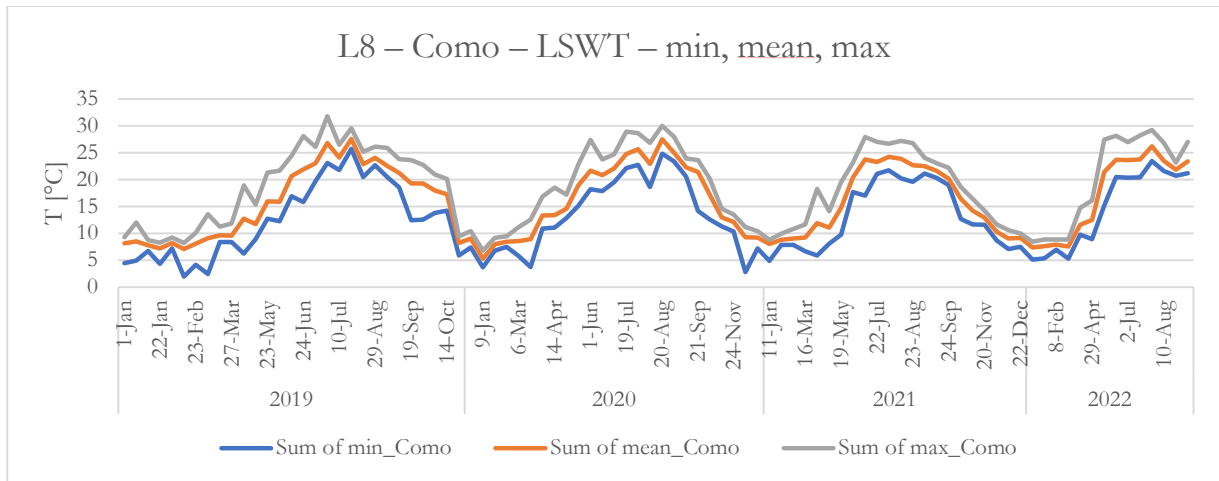
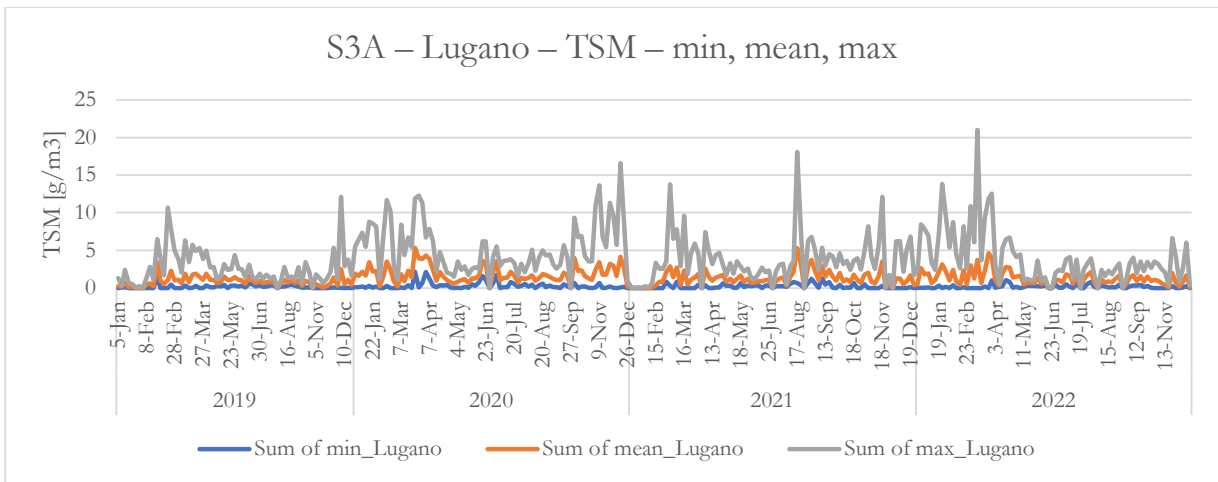## A.3.   LAKE LUGANO

Chlorophyll-a

S3B – Lugano – CHL – min, mean, max

Total Suspended Matter

S3A – Lugano – TSM – min, mean, max

S3B – Lugano – TSM – min, mean, max

Lake Surface Water Temperature

L8 – Como – LSWT – min, mean, max