An open early-warning system prototype for managing and studying algal blooms in Lake Lugano

Daniele Strigaro[1], Camilla Capelli[1], Massimiliano Cannata[1]

[1] Institute of Earth Sciences, DACD, University of Applied Sciences of Southern Switzerland (SUPSI)
Via Flora Ruchat-Roncati 15, CH-6850 Mendrisio, Switzerland
(daniele.strigaro, camilla.capelli, massimiliano.cannata)@supsi.ch

Abstract

Lake Lugano is increasingly experiencing new processes and dynamics due to the combined effects of climate change and human activities. In recent years, abnormal phenomena of algal proliferation, known as Harmful Algal Blooms (HABs), have appeared on the lake. Particularly in 2023, after several critical events, public awareness of this issue has risen, as it can potentially impact socio-economic aspects related to the use of this locally important water resource. In this study, we aim to share the development process of a monitoring system designed to promptly inform about the increase in concentration levels, thereby initiating a protocol to support decision-makers. This system, based on open-source and cost-effective devices, is designed to gather information on these blooms, which pose potential toxicity risks to humans and animals. The system uses a Raspberry Pi connected to a Trilux fluorimeter, measuring algal pigments and transmitting data via NB-IoT. The data is processed using istSOS software, stored locally, and transmitted to a data warehouse. Alerts are set up based on phycocyanin (PC) levels, indicating cyanobacterial biomass and potential microcystin toxicity. The open-source nature of the system allows for easy replication and expansion, aiding decision-makers and researchers, and increasing citizen awareness.

1. Introduction

The effects of climate change, together with human activities, are stressing many natural resources. Such effects are altering distribution patterns, such as precipitation, and known dynamics in all natural spheres (Hydrosphere, Biosphere, Lithosphere, and Atmosphere). The monitoring of environmental parameters is becoming of primary importance to better understand the changes that we need to address. Satellite images, laboratory analysis of samples, and high-end real-time monitoring systems offer solutions to this problem. However, often such solutions require proprietary tools to better exploit data and interact with them. The open science paradigm fosters accessibility to data, scientific results, and tools at all levels of society. Hence, in this project, we aimed to apply such an approach to aid in managing a new phenomenon affecting Lake Lugano, primarily caused by the increase in water temperatures and the high load of nutrients from human activities. In fact, over the past years and particularly in 2023, distributed Harmful Algal Blooms (HABs) appeared on the lake, raising awareness of this phenomenon that can be dangerous for human and animal health. Since HABs are distributed on the water lake surface, an open source cost-effective solution based on open hardware, software and standards can potentially helps in building and reproduce devices to increase the spatial resolution to collect more dense measurements. The excessive algae growth could be composed by Cyanobacteria which can produce a wide range of toxic metabolities, including microcystins (MCs). These cyanotoxins, whose negative effect can be both acute at high concentrations and at low doses, are produced by common species in Lake Lugano. Among these, the most problematic is Microcystis, as it can give rise to blooms during the summer period that accumulate along the shores due to wind and currents. In these areas, the risk of exposure to people and animals is higher, especially in bathing areas. Considering the potential risks to human and an-

imal health, in this project an open early warning monitoring system has been designed and built upon previous experiences in water lake monitoring (Strigaro et al., 2022b, Tiberti et al., 2021) by leveraging the benefits derived from the application of open science principles. Most monitoring plans use microscopic counts of cyanobacteria as an indicator of toxicity risk. However, these analyses are time-consuming, therefore, in addition to or as an alternative to classical methods, sensors capable of measuring algal pigments are increasingly being used. In particular, phycocyanin (PC), characteristic of cyanobacteria, can be adopted as an indicator of cyanobacterial biomass, thus estimating the potential exceedance of critical levels of microcystins. Based on previous studies, this project aimed to develop a high-frequency sensor-based early warning system for real-time detection of phycocyanin in surface waters for bathing use. In particular, the study aimed to i) develop a pilot system for real-time phycocyanin surveillance, using a high-frequency fluorimeter positioned below the surface near a bathing beach; ii) develop a data management software that automatically notifies the exceeding of predicted phycocyanin risk thresholds; iii) test the system during cyanobacterial blooms, comparing the measured phycocyanin values with microcystin concentrations.

2. Materials & methods

2.1 Study area

The selected site for the installation of the monitoring system is situated at the Lido of Riva San Vitale, at the bottom of the North Basin of Lake Lugano (Figure 1). This lake is a glacial body of water located on the border between southern Switzerland and northern Italy (Barbieri and Polli, 1992). This is a public area frequented by people for recreational activities. Additionally, it has peculiar microclimatic conditions governed by wind currents that cause the transport and subsequent accumu-

lation of materials on the lido shores. This process aids in the accumulation of algae, making it one of the most affected places on the lake, subject to the harmful phenomena of HABs.
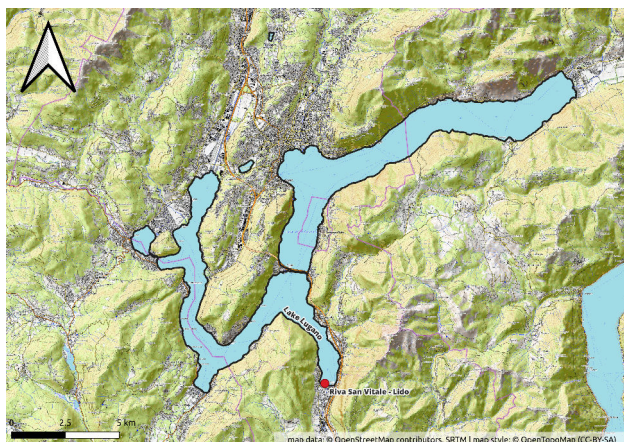

Figure 1. Study area.

2.2 Data collection

In April 2023, near the Lido of Riva San Vitale (Figure. 1), an automatic system for real-time high-frequency detection of PC was installed. The system, anchored to a pole of the pier about 6 meters from the shore (Figure 2), consisted of:

- a *sensor unit* composed by a TriLux fluorometer (Chelsea Technologies), for high-frequency analysis of chlorophyll-a (Chl-a) and PC, positioned just below the surface (20-30 cm). The sensor was equipped with: i) a copper adhesive tape to slow down and prevent the formation of algal biofilm and manually cleaned every week; ii) a 3D printed plastic cap to minimize the errors caused by the sunlight, since it is positioned very near to the surface;

- a *logger unit* composed by a datalogger developed (i) to manage the reading of data from the sensor every minute, (ii) to quality control the values with basic checks on the raw values collected and (iii) to aggregate the verified data based on the configured frequency. In addition, the system allows the transmission of such aggregated data via NB-IoT to the SUPSI datacenter establishing a connection based on the MQTT protocol.

The logger unit utilizes an open-source software developed within the SIMILE project, which has been further enhanced to meet the specific needs of this application. This software (Strigaro and Cannata, 2024), licensed under the GNU GPL v3, is a Python implementation capable of reading data from sensors produced by various companies (e.g., Lufft, Ponsel, Chelsea Technologies) and storing them in a locally installed istSOS service. This approach enables the device to comply with the Sensor Observation Service (SOS) Standard of the Open Geospatial Consortium (OGC), aiming to facilitate the edge computing paradigm that seeks to shift some computational efforts to the node.

Since fluorimetric measurements are easily affected by artifacts caused by the accidental passing of obstacles and sunlight reflection in the water, a data quality test has been implemented for when raw data are collected. This range control ensures that
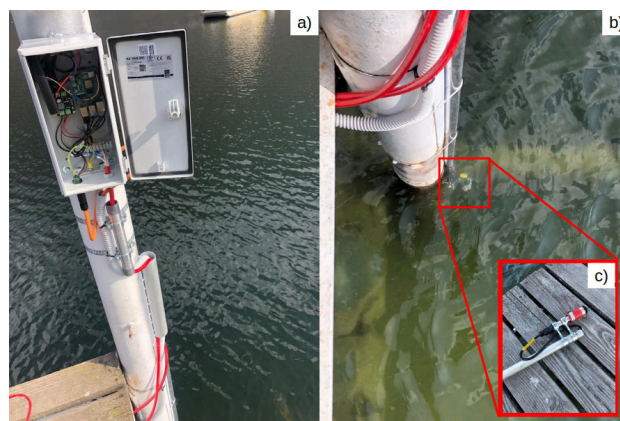

Figure 2. The monitoring system installed near the Lido of Riva San Vitale. a) the datalogger unit; b) and c) the sensor unit.

the measured value is between 0 and 1000 μg/L for both the Chl-a and PC parameters. The istSOS software automatically checks the values during the insertObservation process at the device level. If the test is passed, it assigns a good quality flag; otherwise, the value is flagged as suspicious (Figure 3).

The monitoring system was developed with an emphasis on keeping implementation costs low. The following (Table 1) is a list of components and their associated costs:

Name	Description	Cost (CHF)
Trilux	Fluorimeter from Chelsea Technologies to measure algal pigments (chlorophyll-a and phycocyanin)	~2600
Raspberry pi 4	MicroController unit	~60
Cable	Unterminated, 5 meters length, 11.5 mm dimater	~400
Power module DFROBOT	Voltage converter from 12 to 5V	~7
Mikroe LTE IoT 6 Click	NB-IoT module for data transmission	~70
Miscellaneous	Metal box, cables, teminators, etc.	~350
	Total	~3487

Table 1. List of materials and costs.

The data collected by the logger is sent to the server every 15 minutes using the MQTT protocol. Each set of observations is sent to a topic composed as follows:

```
{base_path}/{istsos_service}/{assignedid}
```

```
Example:
istsos/ceresioremote/810eeb33153c4533a1d4d681acf51079
```

where *base_path* is the path that points to the istsos, *istsos_service* is the service name, a PostgreSQL database schema where all the procedures and observations are stored and, finally, the *assignedid* which is a unique identifier of a procedure. It can be considered as a password to access and communicate new observations to a specific procedure.

This communication uses Quality of Service 1, which guarantees that data has been correctly received by the server since it waits for an acknowledgment in order to end the communication. Otherwise, the data is considered not sent and it will be transmitted the next time.
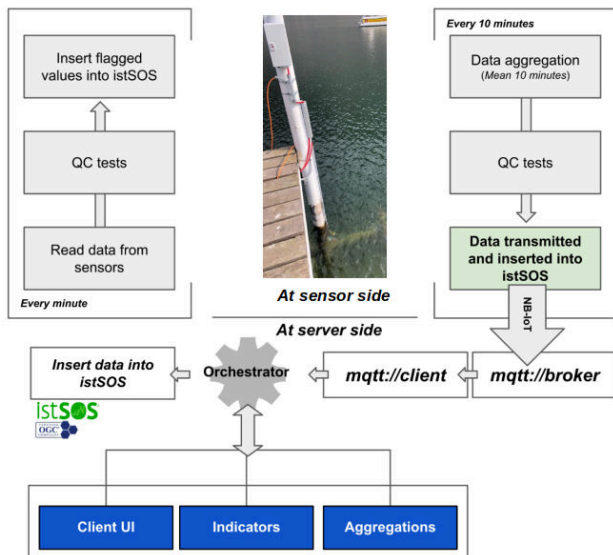

Figure 3. The data flow schema from sensor to data management system.

2.3 Data management architecture

The data management system is composed by a set of services which perform different tasks:

- Keycloak is used to manage authorizations and user profiles;

- VerneMQ is a MQTT broker which collects the messages sent by the monitoring system device;

- istSOS is the Python implementation of the SOS of the OGC and it is used to store and share the data;

- grafana is an open software which has been connected to the database and to the keycloak application in order to provide use-firendly dashboard to the user. It has been also adopted to perform the notification based on pre-defined rules;

- orchestrator is a python implementation to manage the income requests and distributed them over the different services.

Such a software architecture has been firstly implemented during the SIMILE interreg project (Strigaro et al., 2022a) and then re-use to host other new dataset that in the future will be implemented on the lake. It is based on different service in order to provide a complex and complete tool where each task is performed at the best by the assigned service. The "compose"

technology has been used for the automatic generation of the services whose have been implemented as individual docker container. Compose is a tool for defining and starting multi-container docker applications which are defined in a YAML file where the various configurations can be set. The configuration file of the software in production mode is showed in the following code block. Various parameters have been defined and configured to allow the application, which is otherwise very complex to manage and launch, to be simplified, while maintaining the possibility of customising ports, domains and access credentials. These parameters are present in an environmental file containing environment variables that the user can modify before performing the first initialisation phase of the application. Once this phase has been started, the application will be available and accessible with all its functionalities.

2.4 Thresholds definition

Based on the data collected between 2020 and 2022 as part of the monitoring of Lake Lugano (IST-SUPSI 2021, 2022), threshold values of PC were obtained, using a regression model to interpolate the ninetieth percentile (Q90) of the distribution of MCs for each observed PC value. With this model, three PC thresholds were defined within which the risk of exceeding a certain value of MCs is less than 10%. The thresholds were set based on the EPA (United States Environmental Protection Agency) and World Heath Organization (WHO) guidelines for bathing waters (risk range between 8 and 24 μg/L), considering the WHO limit of 24 μg/L of MCs as the term to ban bathing (Chorus and Welker, 2021, EPA, 2019). Each threshold was developed including a hypothetical bathing water management plan, and were thus defined:

1. Monitoring - PC threshold of 3.4 Chl-a eq μg/L, corresponding to a value of 5 μg/L of MCs (with PC greater than Chl-a). This threshold defines an abundant growth of phytoplankton with dominance of cyanobacteria. Upon exceeding this threshold, the risk of exceeding 5 μg/L of MCs increases, and frequent monitoring of the situation and identification of the dominant genus to predict its potential toxicity is recommended;

2. Alert - PC threshold of 6.7 Chl-a eq μg/L, corresponding to a value of 10 μg/L of MCs. This threshold defines an abundant growth of cyanobacteria and the potential start of a bloom. Upon exceeding this threshold, the risk of exceeding 10 μg/L of MCs increases, and a site inspection, identification of the dominant genus, and analysis of cyanotoxins are recommended;

3. Ban - PC threshold of 13.4 Chl-a eq μg/L, corresponding to a value of 20 μg/L of MCs. This threshold defines an ongoing cyanobacteria bloom. Upon exceeding this threshold the toxic risk is maximum, as the values are close to the maximum limits imposed for the bathing ban (24 μg/L of MCs, WHO). Therefore, a temporary bathing ban is recommended until the confirmation of bloom toxicity with verification of any exceeding of the WHO limit.

3. Results

The system developed and installed for monitoring the concentration of algal pigments such as chlorophyll-a and phycocyanin has operated without particular problems from the 6th of May 2023 to the 29th of February 2024. In the graph in Figure

4, the trend of concentrations throughout the entire time frame taken into consideration is present. In total, 51889 observations were expected for each observed property (chlorophyll-a and phycocyanin), but only 40859 have been flagged as good data, which constitutes 78.46% of the expected dataset. Although this result might be seen as unsatisfactory, it's important to consider that the measurement of such parameters can easily be biased due to various factors, such as quenching (Sackmann et al., 2008, Mitchell et al., 2024), fouling, and interference from fish. Furthermore, given that the sensor is located near the surface, it is more susceptible to the effects of light reflection, wave turbulence, and water movements caused by anthropogenic or natural factors (e.g., swimming, boats). In fact, there were instances when spikes appeared, reaching the upper end of the scale, leading the automatic data validation algorithm to flag these data points as erroneous. In Figure 5, the quenching effect between day and night is visible. Essentially, it involves a decrease in the concentration of chlorophyll-a during the day due to the effects of sunlight on the algal pigment and an increment in the total concentration increases during the night. In the literature, there are some works that have attempted to correct such effects using post-processing procedures. These will be evaluated in the future to provide corrected measurements. As described in paragraph 2.2, to counteract the growth of the biofilm on the sensor, it has been equipped with a copper cup. However, even though this system helps in delaying such phenomena, the sensor is cleaned weekly by a technician.
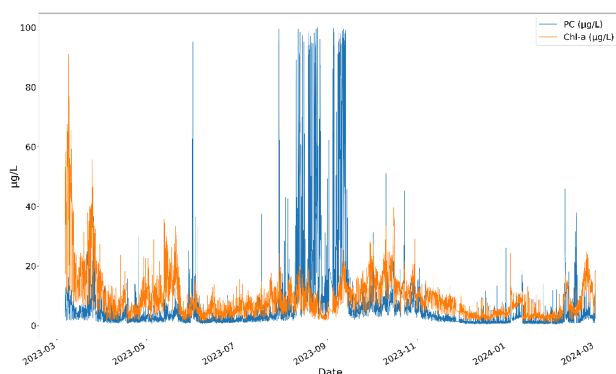

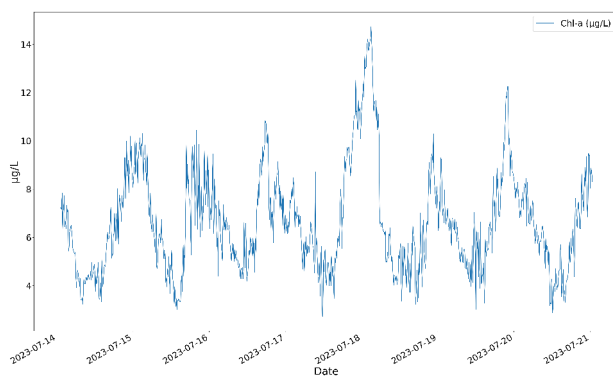Figure 4. Data collected from 06-03-2023 to 29-02-2024.

Figure 5. The quenching effect between day and night.

In Table 2, the results of the analysis performed on the time series regarding the number of times that the measured values adhered to the defined rules are shown. In this analysis, the values have been aggregated hourly to provide more robust data and avoid considering spikes that can distort the data. The data

in the table show that: the "Monitoring" rule have been respected 115 times; the "Alert" level has been triggered 967 times; and finally, the "Ban" level limit has been exceeded 620 times by the data coming from the fluorimeter. This indicates that when there was an algal bloom caused by cyanobacteria, most of the time it was a phenomenon that needed to be monitored in depth and potentially dangerous, since about 87% of the values are in the "Alert" and "Ban" classes.

Threshold	Rule	N°
Monitoring	$3.4 >= PC < 6.7 \ \& \ PC > Chl(a)$	115
Alert	$6.7 >= PC < 13.4$	967
Ban	$PC >= 13.4$	620

Table 2. Number of times that phycocyanin concentration respected the designed rule. Values are expressed in μg/L.
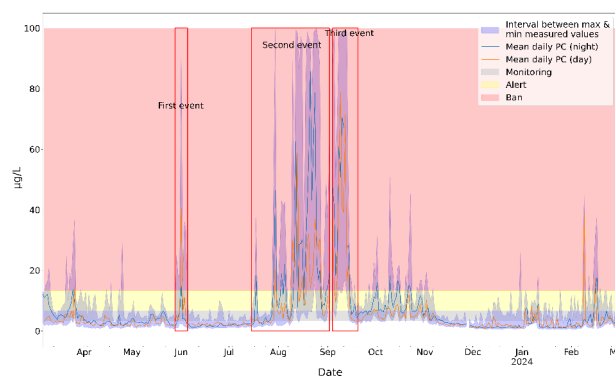

Figure 6. The quenching effect between day and night.

The results of the latest analysis, conducted in this study, are summarized in Figure 4. The plot is constructed using the measured values of phycocyanin from March 6, 2023, to February 29, 2024, and it consists of:

1. An orange line representing the daily average values between the hours of 10:00 and 15:00. This time interval has been chosen to capture the values most affected by the quenching effects, which primarily occur during the most intense sunlight of the day;

2. A blue line representing the daily average values between 23:00 and 03:00, ensuring that only nighttime values are selected to obtain values unaffected by the quenching effect;

3. A light blue area surrounding the blue and orange lines, representing the range between the daily maximum and minimum values. This allows for a better visualization of the range of data collected during the day;

4. Three areas colored in gray, yellow, and red to visualize the classes for monitoring the algal bloom (Monitoring, Alert, and Ban). It should be noted that for the "Monitoring" level, the value should follow the rule described in Table 2, which also considers the comparison with the value of the chlorophyll-a concentrations, which is not considered in this graph.

The trend of daytime and nighttime values clearly shows that the concentration of phycocyanin is almost always lower during the day than at night. This behavior changes when algal blooms are detected. In the graph, three main events are highlighted, and in all these events, the values, either of the night or of the day, are significantly above the upper limit threshold. This indicates that when there is a hazardous event, daytime values can also be considered since the amount of pigments present in the water is so high that it mitigates the quenching effect. In fact, considering that the sensor is below the water surface, the sensor is somewhat shaded by the surface film of algae that occurred during these events (Figure 7), although this hypothesis should be further explored. The first event occurred at the end of May and the beginning of June (Corriere del Ticino, 2023a). It was a bloom that peaked for just one day and then dissolved quite rapidly. The second major event occurred during August 2023. In this case, the highest values of the observations collected were reached. The bloom persisted for a relatively long period, raising awareness among the population and local administrations (Corriere del Ticino, 2023b). A third event followed at the beginning of September. Also in this case, values of algal pigment concentration were very high and in some cases reached the end of the sensor's scale.

Figure 7. Photo about the cyanoHABs events during the summer on Lake Lugano, Riva San Vitale.

A final consideration pertains to how this system communicates data to the end users, who are currently limited to a small group of individuals such as researchers and stakeholders from the local administration, given that this was just a pilot study. The communication strategy involves the use of a dashboard (refer to Figure 8), where data aggregated on an hourly basis are updated in real-time. It is also possible to visualize three classes to determine whether the values have exceeded the limits or not. A second method is the triggering of notifications via email when a limit is reached. This allows researchers to react promptly and monitor the situation in order to decide if a collection of a water sample is needed to check for the presence of dangerous species of cyanobacteria. When a bloom like the one that occurred last August is present, the system checks every 5 minutes which threshold has been exceeded. If the limit exceeded has changed, then users are re-notified with the change in class for both the worsening and improving one. On the contrary, when the value fell in the same class as before, only after 6 hours have passed does the system send a summary notification via email. Finally, when the event concludes, users will be notified with an email reporting the cessation of the phenomenon and the return to normal levels of pigment concentrations. The emails are formatted with the following information:

- The name of the threshold category;

- The value of phycocyanin concentration in Chl-a equivalent (in µg/L);

- The date and time when the event started;

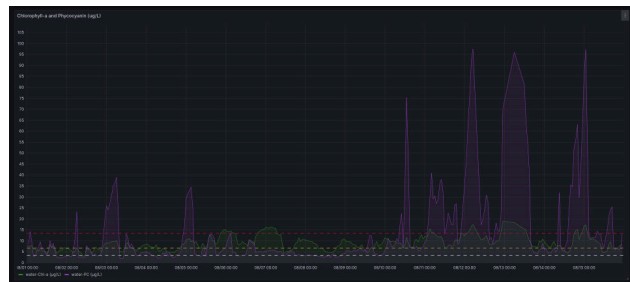- A link to the online dashboard for visualizing data (Figure 8).

Figure 8. The web dashboard accessible by stakeholder involved in the pilot study.

4. Conclusions

The adoption of open hardware, software, and standards enables the implementation of a toolchain that can be easily replicated. The promising results and transparency of the solution will allow for further expansion of the network, assisting decision-makers and researchers in better managing and studying this phenomenon using sensor data. The solution can also effectively raise citizen awareness by implementing kits that local stakeholders can use to monitor the status of the lake water, thereby providing additional data.

In the future, the application could be further improved by increasing the spatial distribution through the adoption of other devices to collect data on pigment concentrations at other lidos. This could also be beneficial in creating a network to increase the spatial monitoring resolution, providing more data and information to create materials that can be used to inform citizens and provide a more reliable system to manage algal blooms.

In addition, considering the number of potentially interested individuals from various groups could be useful in creating a more intelligent and reliable system for generating notifications and managing them in a smarter way (e.g., managing and grouping notifications by area or by the custom choice of clients).

Finally, considering the potential support by citizen the software for managing data can be improved by adopting international standards to handle citizen science data such as STAplus. STAplus aims to standardise citizen science data and make it accessible, interoperable and reusable among different citizen observatories (COs) and services. STAplus extend the already recognize SensorThings API model.

References

Barbieri, A., Polli, B., 1992. Description of lake Lugano. *Aquatic Sciences*, 54(3-4), 181–183. Publisher: Citeseer.

Chorus, I., Welker, M. (eds), 2021. *Toxic Cyanobacteria in Water: A Guide to Their Public Health Consequences, Monitoring and Management*. Taylor & Francis. Accepted: 2021-03-05T12:48:31Z.

Corriere del Ticino, 2023a. Alghe tra Agno e Caslano, sconsigliata la balneazione. https://www.cdt.ch/news/ticino/alghe-tra-agno-e-caslano-sconsigliata-la-balneazione-318693. Section: Ticino.

Corriere del Ticino, 2023b. I cianobatteri fanno preoccupare (e arrabbiare). https://www.cdt.ch/news/i-cianobatteri-fanno-preoccupare-e-arrabbiare-325155. Section: News.

EPA, U., 2019. Recommended human health recreational ambient water quality criteria or swimming advisories for microcystins and cylindrospermopsin. Technical report, EPA 822-R-19-001. Available online at: https://www.epa.gov/sites/default/files/2019-05/documents/hh-rec-criteria-habs-document-2019.pdf.

Mitchell, C., Drapeau, D., Pinkham, S., Balch, W. M., 2024. A chlorophyll a, non-photochemical fluorescence quenching correction method for autonomous underwater vehicles in shelf sea environments. *Limnology and Oceanography: Methods*, 22(3), 149–158. doi.org/10.1002/lom3.10597. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/lom3.10597.

Sackmann, B. S., Perry, M. J., Eriksen, C. C., 2008. Seaglider observations of variability in daytime fluorescence quenching of chlorophyll-*a* in Northeastern Pacific coastal waters. *Biogeosciences Discussions*, 5(4), 2839–2865. doi.org/10.5194/bgd-5-2839-2008. Publisher: Copernicus GmbH.

Strigaro, D., Cannata, M., 2024. A Python software to collect data from sensors. doi.org/10.5281/zenodo.11093851.

Strigaro, D., Cannata, M., Capelli, C., Lepori, F., 2022a. AUTOMATIC ASSESSMENT OF LAKE STATUS USING AN OPEN SOURCE APPROACH: LAKE LUGANO'S CASE STUDY. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-4-W1-2022, 457–463. doi.org/10.5194/isprs-archives-XLVIII-4-W1-2022-457-2022.

Strigaro, D., Cannata, M., Lepori, F., Capelli, C., Lami, A., Manca, D., Seno, S., 2022b. Open and Cost-Effective Digital Ecosystem for Lake Water Quality Monitoring. *Sensors*, 22(17), 6684. doi.org/10.3390/s22176684. Number: 17 Publisher: Multidisciplinary Digital Publishing Institute.

Tiberti, R., Caroni, R., Cannata, M., Lami, A., Manca, D., Strigaro, D., Rogora, M., 2021. Automated high frequency monitoring of Lake Maggiore through in situ sensors: system design, field test and data quality control. *Journal of Limnology*, 80(2). doi.org/10.4081/jlimnol.2021.2011. Number: 2.