

## FOSS4G BASED HIGH FREQUENCY AND INTEROPERABLE LAKE WATER-QUALITY MONITORING SYSTEM

M. Cannata \* [massimiliano.cannata@supsi.ch](mailto:massimiliano.cannata@supsi.ch), D. Strigaro \* [daniele.strigaro@supsi.ch](mailto:daniele.strigaro@supsi.ch), F. Lepori \* [fabio.lepori@supsi.ch](mailto:fabio.lepori@supsi.ch), C. Capelli \* [camilla.capelli@supsi.ch](mailto:camilla.capelli@supsi.ch), M. Rogora \*\* [michela.rogora@cnr.it](mailto:michela.rogora@cnr.it), D. Manca \*\* [dario.manca@cnr.it](mailto:dario.manca@cnr.it)

\* Istituto scienze della Terra, Dipartimento Ambiente Costruzione e Design Scuola Universitaria Professionale della Svizzera Italiana – SUPSI, Campus Mendrisio,  
Via Francesco Catenazzi 23 CH – 6850 Mendrisio, Switzerland.

\*\* CNR Istituto di ricerca sulle Acque (IRSA), Largo Tonolli 50, 28922 Verbania, Italy.

**KEY WORDS:** lake, monitoring, IoT, istSOS, SOS, open

### ABSTRACT:

Climate, together with human activities, is changing the natural dynamics in lake ecosystems and adding new challenges to the management of water resources. Recent studies on Lake Lugano, in Switzerland, showed for instance that the increased water temperature influence other processes such as lake stratification and mixing dynamics, algal blooms, colonisation by alien species, affecting the lake ecosystem as a whole. In such situation, real-time systems with high frequency measurements, together with the traditional discrete monitoring, can help in understanding dynamics and processes occurring on short time scales. To this aim, an open monitoring system largely composed by open source components is being developed for the high frequency monitoring of Lake Lugano. The system relies on the open source software istSOS either on the server and node sides applying the *edge computing* paradigm which is more and more adopted in the Internet of Things field. The implementation collects temperature and dissolved oxygen data from sensors positioned at six different depths of the lake and transmits them using the LoRa radio frequency to a data warehouse. At server side, the software architecture adopts the evolving technology based on containers where services can be grouped in a compose and easily deployed on a server. This paper aims to describe the adopted open source technology and demonstrate that it can be successfully used also in environmental monitoring where the accessibility is limited and the weather conditions can be unpredictable.

## 1. INTRODUCTION

### 1.1 Lake ecosystems

Lake ecosystems are exposed to growing threats due to climate change and other anthropic pressures. For example, water warming is predicted to favour harmful algal blooms (HABs) that are toxic to people and animals (Paerl et al., 2019; Chapra et al., 2017). In addition, warming tends to increase the thermal stratification of lakes and reduce turnovers, which can lead to oxygen depletion in deep layers (Rogora et al., 2018) and release of toxic gases (methane, hydrogen sulphide) from sediments (Woolway et al., 2021). Similarly, the increased use of plastics has produced nano- and micro-plastics pollution which, together with anthropogenic micro pollutants, is posing a new emerging risk factor to lake biota (Sighicelli et al., 2018). Historically, to effectively study and manage those issues, researchers and managers use monitoring data (observations) to derive effective data-driven management policies. Observations have traditionally been gathered from limnological vessels through periodic (often monthly) monitoring campaigns, during which water samples are collected for further analyses in the laboratory and various measurements are performed using on-board instruments (e.g. CTD probe measuring conductivity and water temperature or Secchi disk to measure transparency) (CIPAIS, 2020). However, environmental issues including HABs and changes in lake stratification due to warming, call for a shift towards monitoring approaches that allow higher-frequency (e.g. hourly or sub-hourly) automatic collection of key water-quality properties (e.g. algal pigments, pH, temperature, dissolved oxygen (Salmaso et al., 2020)). Therefore, to meet current challenges, gain better phenomena understanding and activate proactive measures, monitoring systems should be updated to provide a better temporal and

spatial resolution. At the same time, this development should not increase the cost of monitoring, which is very often a limiting factor in lake management.

### 1.2 Open Monitoring System

Environmental Monitoring Systems (EMS) are composed in the following six different domains (Cannata et al., 2017) : (1) sensors, measuring and delivery observations to a local storage; (2) acquisition, collecting data from sensors and transmitting them to a data warehouse; (3) data-service, organizing and structuring data so that they are accessible for quality control and distribution; (4) application, making information accessible through a user interface (i.e.: maps, diagrams, reports); (5) processing, elaborating raw data to derive complex indicators and information; (6) users, mediating information with human understanding toward a system governance or management. Advances in open technologies including hardware and geospatial software and standards together with the massification of sensing device production pushed by the Internet of Things (IoT) have enabled the development of cost-effective and fully open monitoring systems (Cannata et al., 2018; Strigaro et al., 2019).

Such a kind of systems make use of open technologies in every domain of the EMS so that the system is replicable, cost-effective and not constrained by licenses and copyrights. In addition, Open Monitoring Systems allow scientists to control, verify and eventually adapt every step of the data workflow from sensing to consuming making the system highly flexible.

### 1.3 Research scope

This specific work aims at design, implement and test a fully open monitoring system for the high-frequency observation of a

series of limnological key parameters. If validated, this solution will provide a cost effective monitoring approach that could be replicated to other sites within the lake (e.g. secondary gulfs) and other water bodies.

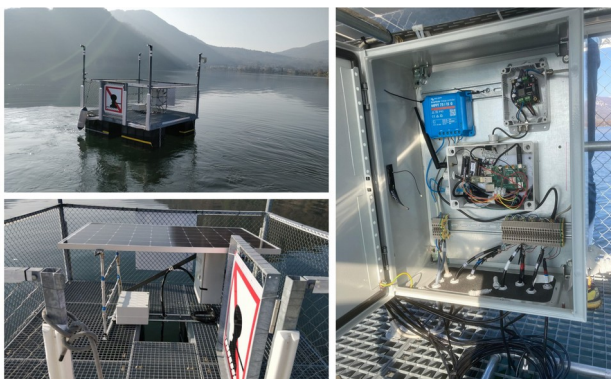
## 2. METHODOLOGY

We adopted the living lab approach based on the involvement of multiple stakeholders for the co-creation of open innovation in a real-world context generating sustainable value for the end-users. The design is the result of surveys, meetings and agile iterations with stakeholders, including administrations as end users of the monitoring systems, limnologists as scientific experts of the ecosystem issues/dynamics and local lake's eco-services users for example from the fishing, water sports, and tourism sectors. The main system requirements that the system should handle were identified in:

- ability to collect data in near real time
- ability to integrate observations from different sources and formats
- analytical capabilities to extrapolate lake water quality
- user friendly Web interface to evaluate biological, chemical and physical dynamics
- enough flexibility to accommodate additional sensors or protocols
- replicability to enable a standardization among different lakes

### 2.1 Sensors domain

The sensor domain is constituted by the components interacting with the data flux from sensing instruments to local data storage. This includes the support structure, sensor devices, the processing unit (with processor, memory and I/O), the power unit and the transceiver. The identified solution of the sensing structure, as depicted in figure 1, was conceived as a square mooring platform anchored to the lake bottom at the four edges. A central hole permits the immersion of sensors in the lake while a fixed structure supports a 300W solar panel and a control unit composed by a Raspberry Pi, a backup battery of 110Ah 12V (estimated to supply the system for more than 9 days in absence of solar recharge) and the LoRa communication unit.



**Figure 1.** Open Lake monitoring system, sensor domain composed by mooring platform, power components (solar panel and backup battery) and control unit (processor and communication).

Installed sensors consist in a chain of thermistors and optical dissolved oxygen sensors (Optod by Ponsel) with resolution 0.01 and accuracy +/- 0.1 mg/L and a standard signal interface

of type Modbus RS-485. Sensors are located at depths of 0.4m, 2.5m, 5m, 8m, 12.5m and 20m.

Edge computing capabilities was implemented at the processing unit to reduce computational effort at data center and bandwidth usage in data transmission. In particular, at the edge two major processes are executed: data standardization and data quality assessment. Thanks to its atomicity and light weight, the FOSS4G istSOS software (Cannata et al., 2019) has been installed at the platform to replace simple data logging. This software offers a standardized web service to collect and dispatch sensor data by implementing the Sensor Observation Service (SOS) standard from the Open Geospatial Consortium (OGC).

Porting istSOS standard service at the edge brings several benefits. First, data are managed within databases instead of files: this expose its intrinsic advantages of data consistency, concurrency, integrity, recovery and security. Additionally, it brings flexibility in adding new sensors, because SOS allows to register new sensors to the system using a standardized sensor description (including their metadata like technical specifications and observed properties formalized using the *SensorML* format) that enable the flux of data to the system. Finally, since SOS offers a standard interface to explore sensors and retrieve data with filter capabilities, the process of data elaboration, enrichment and quality assessment can also be conceptualized and applied reacting to sensor devices modifications. For example, gross error detection can be programmed for the different observed properties and executed on each discovered sensor registered to the service.

Currently, at the platform a preliminary data quality assessment is performed taking advantage of the istSOS features that associates a quality flag, named *qualityIndex*, to each single observation. In this process, data are tested with progressive data quality checks that, if passed, associate a higher quality index to the observation. At sensor specific sampling interval (e.g.: 1 minute for DO), raw data are collected from sensors and while registered in the local istSOS they are in real-time checked with a soundness test and a gross-error test. At specific time intervals data are retrieved, aggregated and associated with a new quality index resulting from raw data quality and a number of data checks, related for example to time consistency or step test. At the end of this quality assessment process 10 minutes aggregated data with *qualityIndex* are stored in the local istSOS and ready to be dispatched.

### 2.2 Acquisition domain

The acquisition domain is the part of the system that is responsible to collect data from the sensor units and register them to a data warehouse. As mentioned above, data are collected every minute and archived in a service of istSOS. A specific service is dedicated to the raw data collected from each sensor installed on the platform. In fact, istSOS permits to create different services which can be defined as an independent SOS instance composed by different constellation of sensors (or procedures). The raw data are subsequently aggregated at pre-configured times, using a script that collects the raw data from istSOS using a *GetObservation*. The data are quality checked using the *qualityIndex* flag and are archived in another service of istSOS which only stores already aggregated data. In this way, raw data are stored and could be retrieved if anything wrong happens. A third script is run at the time to send data. Since at this development state we are using the LoRaWAN protocol to send data through LoRa radio frequency, we need to respect some limitations in terms of bandwidth usage. Therefore, only aggregated data (30 minutes) are sent to the data warehouse. The raw data are collected from the data logger

more or less every one/two months, when the platform is visited for maintenance activities.

### 2.3 Data-service domain

The data-service domain is composed of a number of containerized Web services dedicated to data collection, protection and serving (Figure 2). All the web services are based on open source software. In particular, Keycloak is used for authentication and authorization, istSOS for standard data management according to the OGC Sensor Observation Service, Grafana for time series data plotting, PostgreSQL as database for istSOS and Keycloak.

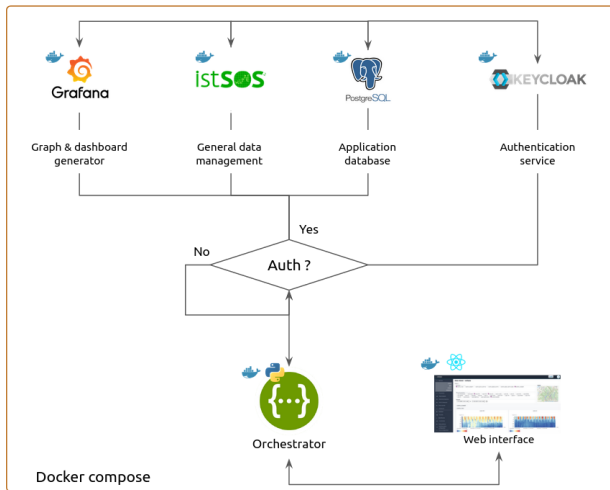


Figure 2. The software architecture.

Specifically, for this project, istSOS software has been expanded to support two specific data types: profile of observations and specimens. Specimens (Figure 3) have been implemented on the base of the OGC samplingSpecimen while profiles (Figure 4) have been considered as a collection of sensors located at different heights and characterized by identical sampling time. These enhancements have been included in the latest istSOS release 2.4-RC2 accessible on github (<https://github.com/istSOS/istsos2>).

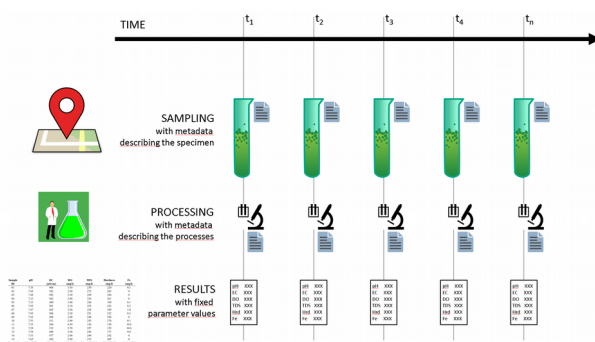


Figure 3. The Specimen observation type.

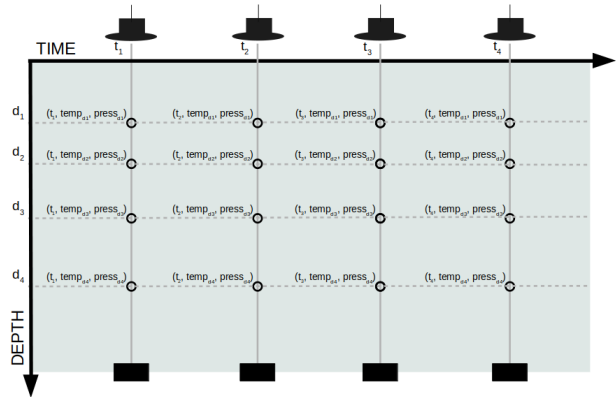


Figure 4. The Profile observation type.

### 2.4 Application domain

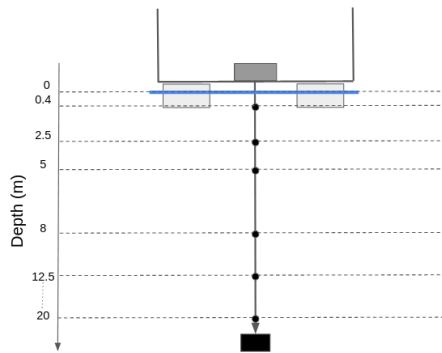
The application domain concerns all the tools developed to make the information available to other services or to users. In Figure 2, there are two other components, developed within the project activity. The first is a Python Orchestrator that controls the user permission and allows the access to the APIs of the different components (istSOS, Keycloak, Grafana). The second is a web user interface which is implemented using the last open source technologies and libraries (ReactJS, NextJS, Node, etc.) to offer an easy way for users to import data, visualize profiles and time series and modify sensor metadata. This component always refers to the orchestrator that sorts the different kind of requests towards the corresponding service.

### 2.5 Processing domain

The processing domain groups all the possible analyses that can be performed on the data. In this context, a plugin system is developed in order to be easily extend the software with specific algorithm not only to check the quality of data, but also to calculate automatically indicators and other post-produced parameters. At this state, we are developing Python scripts to calculate some relevant lake indicators such as water density, metalimnion depths, Schmidt stability and others.

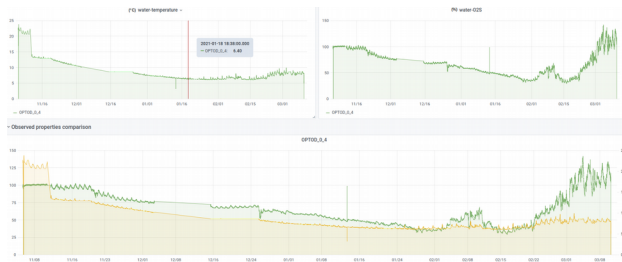
## 3. RESULTS

The preliminary results of the described activities are promising with regard to both the monitoring system installed and the server side software system. A system to monitor the DO and water temperature of the lake is implemented and deployed on a floating platform in the central part of Lake Lugano. The system (Figure 5) is composed by a chain of six sensors positioned at different depths. because the collected data will be used for the calculation of whole-lake metabolism (gross primary productivity and net ecosystem production). The productivity of a lake is usually concentrated in the photic area since it is based on the algae photosynthesis process. Since the photic area in Lake Lugano is generally located in the first 5-10 meters, we decided to position more sensors within this layer.



**Figure 5.** Scheme of the position of the DO sensors.

From the 5<sup>th</sup> of November 2020 to the 10<sup>th</sup> of March 2021, 161346 raw data (1-minute frequency) were collected for each parameter sensed and for each of the six sensors installed on the system. In Figure 6. Graphs produced by the web user interface showing water temperature (°C) and DO data (as % saturation) at 0.4 meters., a screenshot of the data graphs built using the user interface is shown. The data shown are those collected by the sensor located at 0.4 meter depth. The station worked well for the whole period, except during 3-15 December, when the monitoring system stopped working due to a power issued which was then solved.



**Figure 6.** Graphs produced by the web user interface showing water temperature (°C) and DO data (as % saturation) at 0.4 meters.

#### 4. CONCLUSION

The application of a cost-effective approach based as much as possible on open source components permitted the development and deployment of a high frequency monitoring system on Lake Lugano. Such system is installed on a floating platform appositely designed and built to host instruments to collect some limnological parameters. The monitoring system is composed by a Raspberry Pi data logger where all the sensors are connected. Data collected from sensors are stored inside an istSOS instance installed on the logger. This strategy allows to perform some computation at the edge lowering the load on the server side (edge computing paradigm). Also in the communication domain, we tried to use an open approach thanks to the adoption of the LoRa radio frequency for data transmission which offers some critical benefits such as long range coverage, long battery life and potential zero costs of communication.

One of the most important part of the activity concerns the development of a data warehouse composed by different services containerized in a single compose. This application is able not only to store all the data coming from the monitoring system, but also the time series of discrete data previously collected during the traditional monitoring campaigns performed on the lake. To this end, istSOS software is extended to support specimen and profile type of observations.

The whole system is working since November 2020 with promising results. The results demonstrate that an open source approach is a concrete and affordable alternative to more expensive system characterized by proprietary communication protocols and closed source software solutions.

#### REFERENCES

Cannata, M., Antonovic, M., Strigaro, D., & Cardoso, M., 2019: Performance Testing of istSOS under High Load Scenarios. *ISPRS International Journal of Geo-Information*, 8(11), 467.

Cannata, M., Strigaro, D., Shahid, I., Ratnayake, R., & Sudantha, B. H., 2018: The analysis and deployment of a fully open monitoring system in the Sri Lankan Deduroya basin: the 4onse project. In *Free and Open Source Software for Geospatial (FOSS4G) Conference Proceedings (Vol. 18, No. 1, p. 8)*.

Cannata Massimiliano, Strigaro Daniele, Sudantha BH, Rangajeewa Ratnayake, Warusavitharana Emeshi, & Bandara Prasad, 2017: 4onse D2.1 - System design report. Zenodo. <http://doi.org/10.5281/zenodo.1244162>

Chapra, S. C., Boehlert, B., Fant, C., Bierman Jr, V. J., Henderson, J., Mills, D., ... & Paerl, H. W., 2017: Climate change impacts on harmful algal blooms in US freshwaters: a screening-level assessment. *Environmental Science & Technology*, 51(16), 8933-8943.

CIP AIS, Istituto Scienze della terra (IST-SUPSI). 2020: Ricerche sull'evoluzione del Lago di Lugano. Aspetti limnologici. Programma triennale 2019-2021. Campagna 2019. Commissione Internazionale per la Protezione delle Acque Italo-Svizzere (Ed.); 80pp

Paerl, H.W., Huisman, J., 2009. Climate change: a catalyst for global expansion of harmful cyanobacterial blooms. *Environ. Microbiol. Rep.* 1, 27–37.

Pozzoni, Maurizio, Andrea Salvetti, and Massimiliano Cannata. "Retrospective and prospective of hydro-met monitoring system in the Canton Ticino, Switzerland." *Hydrological Sciences Journal* (2020): 1-15.

Rogora, Michela, Fabio Buzzi, Claudia Dresti, Barbara Leoni, Fabio Lepori, Rosario Mosello, Martina Patelli, and Nico Salmaso. "Climatic Effects on Vertical Mixing and Deep-Water Oxygen Content in the Subalpine Lakes in Italy." *Hydrobiologia* 824, no. 1 (November 1, 2018): 33–50.

Salmaso, N., Buzzi, F., Capelli, C., Cerasino, L., Leoni, B., Lepori, F., & Rogora, M., 2020: Responses to local and global stressors in the large southern perialpine lakes: present status and challenges for research and management. *Journal of Great Lakes Research*, 46(4), 752-766.

Sighicelli, M., Pietrelli, L., Lecce, F., Iannilli, V., Falconieri, M., Coscia, L., ... & Zampetti, G., 2018: Microplastic pollution in the surface waters of Italian Subalpine Lakes. *Environmental Pollution*, 236, 645-651.

Strigaro, D., Cannata, M., & Antonovic, M., 2019: Boosting a weather monitoring system in low income economies using open and non-conventional systems: Data quality analysis. *Sensors*, 19(5), 1185.

Woolway, R. I., Sharma, S., Weyhenmeyer, G. A., Debolskiy, A., Golub, M., Mercado-Bettin, D., ... & Jennings, E., 2021: Phenological shifts in lake stratification under climate change. *Nature communications*, 12(1), 1-11.