The Observatory on LAKes (OLA) database: Sixty years of environmental data accessible to the public
Frédéric Rimet, Orlane Anneville, Denis Barbet, Cécile Chardon, Laura Crepin, Isabelle Domaizon, Jean-Marcel Dorioz, Laurent Espinat, Victor Frossard, Jean Guillard, et al.

To cite this version:
Frédéric Rimet, Orlane Anneville, Denis Barbet, Cécile Chardon, Laura Crepin, et al.. The Observatory on LAKes (OLA) database: Sixty years of environmental data accessible to the public. Annales de Limnologie - International Journal of Limnology, EDP sciences, 2015, 10.4081/jlimnol.2020.1944. hal-02916312

HAL Id: hal-02916312
https://hal.inrae.fr/hal-02916312
Submitted on 17 Aug 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
The Observatory on LAkes (OLA) database: Sixty years of environmental data accessible to the public

Frédéric Rimet,¹ Orlane Anneville,¹ Denis Barbet,¹ Cécile Chardon,¹ Laura Crépin,¹ Isabelle Domaizon,¹ Jean-Marcel Dorioz,¹ Laurent Espinat,¹ Victor Frossard,¹ Jean Guillard,¹ Chloé Goulon,¹ Valérie Hamelet,¹ Jean-Christophe Hustache,¹ Stéphan Jacquet,¹ Leslie Lainé,¹ Bernard Montuelle,¹ Pascal Perney,¹ Philippe Quetin,¹ Serena Rasconi,¹ Antoine Schellenberger,² Viet Tran-Khac,² Ghislaine Monet¹

¹UMR CARRTEL, INRAE, Université de Savoie Mont-Blanc, 75 bis Avenue de Corzent, CS 50511, 74203 Thonon-les-Bains Cedex; ²US 1106 Infosol, INRAE, 2163 Avenue de la Pomme de Pin, Ardon CS 40001, 45075 Orléans Cedex 2, France

ABSTRACT
Lakes are essential ecosystems that provide a large number of ecosystem services whose quality is strongly impacted by human pressures. Optimal uses of lakes require adapted management practices which in turn rely on physico-chemical and biological monitoring. Long-term ecological monitoring provides large sets of environmental data. When such data are available, they have to be associated to metadata and to be stored properly to be accessible and usable by the scientific community. We present a data informatics system accessible to anyone who requests it. Maintained online since 2014 (https://si-ola.inrae.fr), it is originated from the Observatory on LAkes (OLA). It contains long-term data from 4 peri-alpine lakes (Lakes Aiguebelette, Annecy, Bourget, Geneva/Léman) and 24 high-altitude lakes of the northern French Alps. We describe the generated long-term data series, the data type, the methodologies and quality control procedures, and the information system where data are made accessible. Data use is allowed under the condition of providing reference to the original source. We show here how such a platform clearly enhances data sharing and scientific collaboration. Various studies referring to these data are regularly published in peer-reviewed journals; providing in fine a better understanding of lakes’ ecosystems functioning under local and global pressures.

INTRODUCTION
While lake water represents less than 0.01% of the total water on Earth (Shiklomanov, 1993), it provides a disproportionately high number of ecosystem services, such as water provisioning, fish resources, biodiversity support, climate buffering, scenic and cultural services. Any alteration of lake’s water quality has wide-ranging ecological and societal implications. In many ways, human societies depend on healthy aquatic ecosystems, and conversely, ecosystems depend on human pressures.

Humans have influenced lake ecosystems for thousands of years, both locally and regionally, and impacts on the functioning of lakes often lead to alterations of water quality and modifications of ecosystem services. The most striking example is eutrophication: after industrialization and introduction of fertilizers and domestic pollutants, major changes in nutrient loads occurred in lakes, leading to increased primary production and frequent associated dysfunctions such as hypoxia and Harmful Algal Blooms (Le Moal et al., 2019). When the first symptoms of eutrophication have appeared, scientists and stakeholders promptly highlighted the need for the implementation of systematic water quality monitoring, and, at best, this type of survey began in the late 1950s.

Data collected for water quality monitoring can be used for scientific purposes and contribute to the global data collection needed for better evaluating the impacts of long-lasting changes such as climate change. Long-term data that allow to detect changes in lake ecosystems are also essential for achieving an integrated understanding of how aquatic biological populations and communities interact and to develop and test ecological theories. Long-term monitoring also help establishing reference conditions in order to identify the magnitude of changes between past and present conditions (Kaiblinger et al., 2009). Long term ecological research (LTER) programs have been established in several regions such as North America and Europe over the last decades to facilitate adapting environmental policies in watersheds and protect terrestrial and aquatic resources for future generations.
Such programs are either national (e.g., in France: Zone Ateliers and Analysis and Experimentation on Ecosystems [AnaEE France]; in the USA: Long Term Ecological Research [LTER]) or involved in international networks (e.g., AnaEE Europe, Europe LTER, International LTER [ILTER] see Mirtl et al. 2019). These programs aim to facilitate interactions and collaborations within the multidisciplinary community of researchers such as the Global Lake Ecological Observatory Network -GLEON- (Rose et al., 2016).

Storage of long-term environmental raw data is crucial and it is essential to associate to them meta-information about data traceability. Indeed, protocols are highly susceptible to change over time because of technological improvements or natural adaptation of ecosystems. A clear description of those raw data and how they have been collected is crucial to interpret and re-use them. Consequently, that information on data acquisition have to be documented in metadata which must be associated to raw data. Moreover, it is necessary to develop platforms that provide a persistent access to this information. The observatory OLA (Observatory on LAKes) offers such a structure.

The observatory OLA (https://www6.inrae.fr/soere-ola_eng/) was created by the research unit CARRTEL (INRAE-USMB), an alpine research center devoted to the study of limnetic systems and associated food webs. It focuses on lakes of the northern French Alps and covers the region between 45°32’ and 46°26’north. This region has been particularly affected by growing human populations and activities during the last decades. Human populations were concentrated primarily around the largest lakes (Aiguebelette, Bourget, Annecy and Geneva), where densities were particularly elevated — more than 1000 inhabitants per km² in many places. In the 20th century, the need to protect these precious aquatic resources became crucial because of the ecological, social and economic importance of these four lakes. Therefore, an ecological monitoring started in the 1960’s as soon as the first signs of important eutrophication occurred (Jacquet et al., 2014a; Montuelle and Clemens, 2015; Tab. 1). These monitorings are still ongoing today and are funded through various administrative entities, representing several municipalities neighboring the lakes (see Supplemental data S1 for details).

OLA provides high-quality scientific data for understanding and modeling the ecological functioning of lakes. Data are publicly available and are regularly used in scientific studies. An information system (IS) accessible at https://si-ola.inrae.fr gives access to the data. In this paper, we briefly describe the analytical methods for the main parameters, the database content, rules for using the data, IS data storage and the website interface.

**Data available in the database**

The efficient partnership between scientists and local stakeholders allowed the collection of physical, chemical and biological data from 1957 on Lake Geneva (also called Lake “Léman”), 1966 on Lake Annecy, 1974 on Lake Aiguebelette and 1987 on Lake Bourget. More recently, since the late 90s, 24 small high-altitude lakes situated on a latitudinal gradient in the French Alps were included in the observatory, and are since then monitored. Studying these lakes, considered as ‘pristine’ and highly sensitive to climate change, opens perspectives on the impacts of global change on lake functioning. The location and intrinsic physical characteristics of the largest lakes monitored in the framework of OLA are shown in Tab. 1 and Fig. 1. Samples are collected from the middle of the lake above its deepest point. Detailed description of all sampling stations can be found at https://si-ola.inrae.fr and in Tab. 2.

**ACQUISITION OF BIOLOGICAL DATA**

**Taxonomic composition of phytoplankton community**

The water sample for phytoplankton taxonomic identification and counting is collected within the first 20 m (0-10 or 0-18 m, depending on the lake and year) using an integrated sampling bottle. Samples are preserved with Lugol and then are kept in a dark and cool place until microscopic observations are performed. The Utermöhl technique (Utermöhl, 1958), standardized at the European level (CEN, 2006), is used. Briefly, after the settling of a known volume of sample in a microscopic chamber, algal counts and determinations are carried out under an inverted microscope at x64 magnification using specialized literature (for instance, the various volumes of the *Süsswasserflora von Mitteleuropa* edited by B. Büdel, G. Gärtnner, L. Krienitz and M. Schagerl). Results are expressed in biovolume per volume of water (μL.m⁻³) or in biomass.

**Tab. 1.** Limnological characteristics of the main lakes monitored in the OLA observatory.

<table>
<thead>
<tr>
<th>Lake</th>
<th>Volume (km²)</th>
<th>Renewal time (years)</th>
<th>Max depth (m)</th>
<th>Watershed surface (km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geneva</td>
<td>89</td>
<td>11.3</td>
<td>310</td>
<td>7395</td>
</tr>
<tr>
<td>Bourget</td>
<td>3.6</td>
<td>9.0</td>
<td>145</td>
<td>560</td>
</tr>
<tr>
<td>Annecy</td>
<td>1.12</td>
<td>3.8</td>
<td>82</td>
<td>251</td>
</tr>
<tr>
<td>Aiguebelette</td>
<td>0.116</td>
<td>3.1</td>
<td>70</td>
<td>59</td>
</tr>
</tbody>
</table>
per volume of water considering that phytoplankton density is 1 kg.L⁻¹ (µg.L⁻¹). A database of species biovolumes (Rimet and Druart, 2018) is used to obtained these results. Over more than 50 years, this biovolumes’ database has been regularly updated to include species observed in the alpine lakes in France (Lakes Geneva, Annecy, Bourget and Aiguebelette) and also in high-altitude lakes in the French Alps (Feret et al., 2017) and in various other French lakes.

Simultaneously, picocyanobacteria are examined in Lakes Annecy, Bourget and Geneva from samples taken at 6 discrete depths since 2003 (see details in supplemental data S2); these data will be included in the IS in the forthcoming years.

**Primary production and Chl a**

Primary production and Chl a are measured at 9 discrete depths (0, 1, 2.5, 3.5, 5, 7.5, 10, 15 and 20 m) in Lake Geneva (two times a month the entire year except once a month from December to February). Primary production has been measured using the ¹⁴C method (Søemm-Nielsen, 1952) with the incorporation of NaH¹⁴CO₃.
from 1971 to 2012 (Pelletier, 1983). Since 2012, measurements of primary production are carried out using an isotopic method (incubation with $^{13}$C sodium bicarbonate whose incorporation into the biomass is measured by isotope mass spectrometry). Measures of primary production are expressed in $\mu$g C/L/h. Chlorophyll $a$ is analyzed using spectrophotometry after extraction into 90% acetone (ISO 10260: 1992). Chlorophyll $a$ discrete measurements are completed with vertical profiles measures obtained with multiparameter probes.

**Zooplankton**

Zooplankton is sampled with a vertical tow from 50 m to the surface using a 200-µm plankton net for crustaceans and a 64-µm plankton net for rotifers. Samples are fixed in the field using 5% buffered formaldehyde.

Crustacean identification and counts are carried out in a subsample of a known volume of the net sample using a light microscope at x10 magnification. The identification books of Amoros (Amoros, 1984) and Dussart (Dussart, 1967, 1969) are used to distinguish species. Larger organisms (*Leptodora kindtii* Focke and *Bythotrephes longimanus* Leydig) are identified using the same procedure at x4 magnification. Final results are given in number of individuals per square meter.

Samples dedicated for observing rotifers are processed following the Utermöhl technique (Utermöhl, 1958). A subsample of known volume is allowed to settle in a microscopic chamber, and counts and determinations are carried out using an inverted microscope at x16 magnification. The identification books of Stemberger (Stemberger, 1979) and Voigt and Koste (Voigt and Koste, 1978a, 1978b) are used to distinguish species. Results are given in number of individuals per square meter.

---

**Tab. 2.** Complete list of the lakes and sampling stations integrated in the OLA.

<table>
<thead>
<tr>
<th>Type</th>
<th>Name of lake</th>
<th>Name of sampling station</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Altitude (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large lake</td>
<td>Aiguebelette</td>
<td>Point a</td>
<td>45.55066</td>
<td>5.00900</td>
<td>390</td>
</tr>
<tr>
<td>Large lake</td>
<td>Ancey</td>
<td>Grand lac</td>
<td>45.87270</td>
<td>6.16453</td>
<td>447</td>
</tr>
<tr>
<td>Large lake</td>
<td>Ancey</td>
<td>Petit lac</td>
<td>45.81748</td>
<td>6.22140</td>
<td>447</td>
</tr>
<tr>
<td>Large lake</td>
<td>Bourget</td>
<td>Point B</td>
<td>45.75583</td>
<td>5.86000</td>
<td>231</td>
</tr>
<tr>
<td>Large lake</td>
<td>Léman</td>
<td>Stl2</td>
<td>46.45345</td>
<td>6.59423</td>
<td>372</td>
</tr>
<tr>
<td>High-lake</td>
<td>Anterne</td>
<td>Sampling point 1</td>
<td>45.99063</td>
<td>6.79889</td>
<td>2063</td>
</tr>
<tr>
<td>High-lake</td>
<td>Arpont</td>
<td>Sampling point 1</td>
<td>45.31580</td>
<td>6.77683</td>
<td>2672</td>
</tr>
<tr>
<td>High-lake</td>
<td>Aumar</td>
<td>Aumar : sampling point 1</td>
<td>42.84250</td>
<td>0.14972</td>
<td>2192</td>
</tr>
<tr>
<td>High-lake</td>
<td>Blanc du Carro</td>
<td>Sampling point 1</td>
<td>45.42000</td>
<td>7.12668</td>
<td>2753</td>
</tr>
<tr>
<td>High-lake</td>
<td>Bresses inférieur</td>
<td>Centroid</td>
<td>44.15812</td>
<td>7.24204</td>
<td>2458</td>
</tr>
<tr>
<td>High-lake</td>
<td>Bresses supérieur</td>
<td>Centroid</td>
<td>44.15932</td>
<td>7.24406</td>
<td>2501</td>
</tr>
<tr>
<td>High-lake</td>
<td>Brévent</td>
<td>Sampling point 1</td>
<td>45.92906</td>
<td>6.82739</td>
<td>2127</td>
</tr>
<tr>
<td>High-lake</td>
<td>Corne</td>
<td>Sampling point 1</td>
<td>45.22490</td>
<td>6.07870</td>
<td>2098</td>
</tr>
<tr>
<td>High-lake</td>
<td>Cornu</td>
<td>Sampling point 1</td>
<td>45.95837</td>
<td>6.84954</td>
<td>2275</td>
</tr>
<tr>
<td>High-lake</td>
<td>Cos</td>
<td>Sampling point 1</td>
<td>45.23250</td>
<td>6.07890</td>
<td>2183</td>
</tr>
<tr>
<td>High-lake</td>
<td>Izourt</td>
<td>Izourt : sampling point 1</td>
<td>42.68805</td>
<td>1.49777</td>
<td>1645</td>
</tr>
<tr>
<td>High-lake</td>
<td>Jovet</td>
<td>Sampling point 1</td>
<td>45.75697</td>
<td>6.73199</td>
<td>2173</td>
</tr>
<tr>
<td>High-lake</td>
<td>Lauzanier</td>
<td>Sampling point 1</td>
<td>44.3791</td>
<td>6.87234</td>
<td>2280</td>
</tr>
<tr>
<td>High-lake</td>
<td>Malrif</td>
<td>Centroid</td>
<td>44.82292</td>
<td>6.86922</td>
<td>2580</td>
</tr>
<tr>
<td>High-lake</td>
<td>Merlet Supérieur</td>
<td>Sampling point 1</td>
<td>45.36040</td>
<td>6.64023</td>
<td>2447</td>
</tr>
<tr>
<td>High-lake</td>
<td>Mont Coua</td>
<td>Sampling point 1</td>
<td>45.31740</td>
<td>6.63899</td>
<td>2666</td>
</tr>
<tr>
<td>High-lake</td>
<td>Muzelle</td>
<td>Sampling point 1</td>
<td>44.95050</td>
<td>6.09722</td>
<td>2099</td>
</tr>
<tr>
<td>High-lake</td>
<td>Noir du Carro</td>
<td>Sampling point 1</td>
<td>45.42140</td>
<td>7.12491</td>
<td>2750</td>
</tr>
<tr>
<td>High-lake</td>
<td>Pavé</td>
<td>Centroid</td>
<td>44.98770</td>
<td>6.32588</td>
<td>2846</td>
</tr>
<tr>
<td>High-lake</td>
<td>Pétarel</td>
<td>Centroid</td>
<td>44.80214</td>
<td>6.16875</td>
<td>2090</td>
</tr>
<tr>
<td>High-lake</td>
<td>Pisses</td>
<td>Centroid</td>
<td>44.72245</td>
<td>6.37941</td>
<td>2515</td>
</tr>
<tr>
<td>High-lake</td>
<td>Plan Vianney</td>
<td>Sampling point Ebytem</td>
<td>44.95750</td>
<td>6.04333</td>
<td>2250</td>
</tr>
<tr>
<td>High-lake</td>
<td>Pormenaz</td>
<td>Sampling point 1</td>
<td>45.96254</td>
<td>6.79460</td>
<td>1947</td>
</tr>
<tr>
<td>High-lake</td>
<td>Rabuons</td>
<td>Sampling point 1</td>
<td>44.26970</td>
<td>6.98250</td>
<td>2501</td>
</tr>
<tr>
<td>River</td>
<td>Dranse</td>
<td>Defaut dranse – main tributary of Lake Léman</td>
<td>46.38350</td>
<td>6.50744</td>
<td>388</td>
</tr>
</tbody>
</table>
PHYSICAL AND CHEMICAL DATA ACQUISITION

In situ measurements

Transparency, using a Secchi disk, is systematically measured following the international standard ISO 7027. Probes are used to measure depth (pressure), pH, conductivity, oxygen concentration, turbidity, in situ chlorophyll a fluorescence and photosynthetic active radiation. In recent years, the most commonly used probes were CTM214 FO and CTD009 (Sea and Sun Technology©).

Water samples are collected at discrete depths for subsequent chemical and physical laboratory analyses. Sampling is performed using bottles such as the Freeflow bottle (Hydrosbios©) or the integrated sampling bottle (Hydrosbios © IWS bottle). In past years, other kinds of bottles were used, including the Niskin bottle for collecting integrated depth samples (Pelletier and Orand, 1978).

Chemical and physical laboratory data

In the laboratory, water samples are filtered immediately through Whatman© GF/F filters (0.7-µm nominal pore size) and then stored at 4°C in high-density polyethylene plastic bottles or amber glass bottles until analyses, which are performed within 48 h after collection. A large range of physical and chemical analyses are carried out following international or national standards. Chemical procedures follow rigorous quality assurance and control to ensure the quality of obtained analytical data. Tab. 3 shows the major analyses performed.

Total suspended matter is determined using an ultrasonic bath (XP2U; Mettler Toledo©). Major cations (Ca2+, Mg2+, K+ and Na+) are determined using flame (acetylene/air) atomic absorption spectrophotometry (AA240FS; Varian©) following NF 90-020. Lanthanum is added as a matrix modifier to minimize interference. Major anions (Cl−, NO3− and SO42−) are quantified using ion exchange (861 Advanced Compact ion chromato-graph; Metrohm©) with chemical suppression following ISO 10304-1.

An automatic titrator (Basic Titrino 794, Metrohm©) with a glass electrode is used to perform successive pH measurements. Temperature corrections are systematically applied when performing lab measurements of the conductivity (ISO 10523:2008). Total alkalinity is obtained using potentiometric titration (ISO 9963-1).

A UV-vis spectrophotometer (Cary 50 scan; Varian©) is used to determine total phosphorus and orthophosphate based on sulfuric acid digestion and a molybdenum blue method (EPA 365.3). Ammonium is determined based on an indophenol blue method (ISO 5664:1984), and nitrite ions are determined using a diazotization method (ISO 6777:1984).

Total nitrogen is determined using chemiluminescence after high-temperature digestion and catalytic postcombustion using an ElementarVario© TOC/TNb coupled with a chemiluminescence detector (APNA-370; Horiba©). Total inorganic nitrogen (ammonium and nitrate ions) is calculated as the difference between total nitrogen and or-

Tab. 3. Physical and chemical analyses carried out in the laboratory. References to the standard methods and to the main principles are given.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Standard</th>
<th>Principle</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH, Conductivity</td>
<td>ISO 10523:2008</td>
<td>Potentiometric and conductometric measure</td>
</tr>
<tr>
<td>Alkalinity</td>
<td>ISO 9963-1</td>
<td>Potentiometric determination</td>
</tr>
<tr>
<td>Total suspended matter</td>
<td>EN 872:2005</td>
<td>Filtration</td>
</tr>
<tr>
<td>Dissolved oxygen</td>
<td>ISO 5813:1983</td>
<td>Iodimetric method</td>
</tr>
<tr>
<td>Total and dissolved organic carbon</td>
<td>EN 1484:1997</td>
<td>Heated-Persulfate oxidation, Non-dispersive infrared measure</td>
</tr>
<tr>
<td>Particulate carbone and nitrogen</td>
<td>ASTM D2579-93e1</td>
<td>CHN elemental analyzer (High-temperature combustion method, gas chromatography, thermal conductivity detector)</td>
</tr>
<tr>
<td>Ammonium</td>
<td>ISO 5664:1984</td>
<td>UV-Visible spectrometer</td>
</tr>
<tr>
<td>Nitrite</td>
<td>ISO 6777:1984</td>
<td>UV-Visible spectrometer</td>
</tr>
<tr>
<td>Nitrate, chloride, sulfate</td>
<td>ISO 10304-1</td>
<td>Ion chromatography</td>
</tr>
<tr>
<td>Total nitrogen</td>
<td>EN 12260:2003</td>
<td>High temperature digestion and catalytic post combustion, chemiluminescence measure</td>
</tr>
<tr>
<td>Orthophosphate, total phosphorus, particulate phosphorus</td>
<td>EPA 365.3</td>
<td>UV-Visible spectrometer</td>
</tr>
<tr>
<td>Silica</td>
<td>NF T 90-007</td>
<td>Discret Chemistry Analyzer</td>
</tr>
<tr>
<td>Majors cations (sodium, potassium, calcium, magnesium)</td>
<td>NF 90-020</td>
<td>Atomic Absorption Spectrometry</td>
</tr>
<tr>
<td>Chlorophyll-a</td>
<td>ISO 10260 : 1992</td>
<td>UV-Visible spectrometer</td>
</tr>
</tbody>
</table>
ganic nitrogen (EN 12260:2003). Total and dissolved organic carbon are measured using a non-dispersive infrared detector coupled to an Aurora 1030W carbon analyzer, applying a heated persulfate oxidation method (EN 1484:1997). Dissolved organic carbon is determined from filtered (muffled Whatman© GF/F filter) samples. Particulate organic carbon and nitrogen are analyzed by applying a high-temperature combustion method (ASTM D2579-93e1) with a gas chromatograph equipped with a thermal conductivity detector (Flash 2000; Thermoscientific©). Biogenic silica is analyzed using a Smartchem 200 Discrete Analyzer (WESTCO Scientific Instruments©), applying NF T90-007.

**DATA STORED IN THE IS: FOCUS ON A SELECTION OF LONG-TERM LIMNOLOGICAL DATA**

Tab. 4 gives an overview of the data available in the IS. The longest series of data are available for Lake Geneva (going back to 1957 for chemical parameters).

The most recent series are available for the high-altitude lakes. A detailed description of the data is available at https://si-ola.inrae.fr/

From such data, the responses of lakes to local and global pressures have been described, as for instance the effect of global warming on water temperature (O’Reilly et al., 2015), thermal structures (Kraemer et al., 2015), lake mixing, oxygen availability (Foley et al., 2012), plankton communities and fisheries (O’Reilly et al., 2003). Local and global pressures deeply modify alpine lake functioning and biodiversity, as testified by recent paleo-reconstructions showing multi-level responses in biodiversity facing eutrophication and warming (Perga et al., 2015).

Here we selected some of these long-term data to illustrate some of the scientific advances they recently provided. Supplemental data 2 and 3 show complementary data collected for other biological compartments (fish and picophytoplankton) that will be included soon in the OLA-IS.

**Tab. 4.** Temporal distribution of the data available in the information system (white: no data, grey: data available in the information system).

<table>
<thead>
<tr>
<th>Lake</th>
<th>Measures</th>
<th>50ies</th>
<th>60ies</th>
<th>70ies</th>
<th>80ies</th>
<th>90ies</th>
<th>2000</th>
<th>2010</th>
<th>Present</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aiguelebette</td>
<td>Chemical par.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multipar. probe</td>
<td>1974</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phytoplankton</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chlorophyll</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1989</td>
</tr>
<tr>
<td>Annecy</td>
<td>Chemical par.</td>
<td>1966</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multipar. probe</td>
<td>1966</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zooplankton</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phytoplankton</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chlorophyll</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1990</td>
</tr>
<tr>
<td>Bourget</td>
<td>Chemical par.</td>
<td>1999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multipar. probe</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2003</td>
</tr>
<tr>
<td></td>
<td>Zooplankton</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phytoplankton</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chlorophyll</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1987</td>
</tr>
<tr>
<td>Geneva</td>
<td>Chemical par.</td>
<td>1957</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multipar. probe</td>
<td>1957</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zooplankton</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phytoplankton</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1974</td>
</tr>
<tr>
<td></td>
<td>Chlorophyll</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1976</td>
</tr>
<tr>
<td>High</td>
<td>Chemical par.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1992</td>
</tr>
<tr>
<td>Altitude</td>
<td>Multipar. probe</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2014</td>
</tr>
<tr>
<td>Lakes*</td>
<td>Zooplankton</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2015</td>
</tr>
<tr>
<td></td>
<td>Phytoplankton</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1998</td>
</tr>
<tr>
<td></td>
<td>Chlorophyll</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2015</td>
</tr>
</tbody>
</table>

Interannual dynamics of phosphorus in deep peri-alpine lakes

Lakes Geneva, Bourget, Annecy and Aiguebelette are typologically similar but they have been impacted differently by local anthropogenic pressures. Restoration of these lakes, after phosphorus pollution several decades ago, provides nowadays exemplary cases of successful restoration of trophic status in a world where freshwater eutrophication is increasing globally (Le Moal et al., 2019). Lakes Geneva and Bourget, and to a lesser degree, Annecy and Aiguebelette, underwent rapid eutrophication during the 60s and 70s as a result of development of nutrient point sources, mostly untreated domestic effluent discharges. The temporal series of total phosphorus concentrations in Lake Geneva from 1957 to 2017, in Lake Bourget from 1999-2017, in Lake Annecy from 1966 to 2017 and in Lake Aiguebelette from 1974 to 2017 are shown in Fig. 2.

Phosphorus maxima observed in the data were reached during the late 70s in Lake Geneva and during the 90s in Lakes Annecy and Aiguebelette. The maximum phosphorus concentration in Lake Bourget during the 80s is also well known but not included in the database (Perga et al., 2015). Effective measures for reducing the phosphorus load in the 70s resulted in strong decreases in phosphorus concentrations in the four large alpine lakes. Re-oligotrophication processes continue today (Jacquet et al., 2014a).

Long-term chloride concentration in Lake Geneva

Chloride ion concentration in Lake Geneva increased continuously from 1973 to 2017 (Fig.3). The annual average concentration reached 10.31 mg Cl⁻/L. The contributions of various tributaries to the lake largely explain this trend (Klein, 2016). Chloride concentration reflects a growing utilization of chemical chlorine-based products in the industrial sector (metallurgy and pharmaceutical sectors) and the application of deicing salt during winter (Gumy and De Alencastro, 2001). These data have been used to consolidate a database of 529 lakes in Europe and North America (Dugan et al., 2017b); the associated study predicts that many lakes could exceed the aquatic life threshold criterion for chronic chloride exposure (230 mg L⁻¹), stipulated by the US Environmental Protection Agency (EPA), in the next 50 years if current trends do continue (Dugan et al., 2017a).

Temporal dynamics of phytoplankton and zooplankton in Lake Geneva

Over the last 50 years, the abundance, taxonomic structure and phenology of phytoplankton communities have changed (Anneville and Leboulanger, 2001; Anneville et al., 2005, 2010, 2018; Rimet et al., 2009; Jacquet et al., 2014c) because of decrease in phosphorus concentrations, warming in water temperature and change

---

in the thermal stratification dynamic. Except in Lake Geneva (Fig. 4), phytoplankton biomass has systematically decreased with phosphorus concentration. The unexpected trend in phytoplankton biomass observed in Lake Geneva is explained by a decrease in zooplankton grazing (Tadonleke et al., 2009), a change in phytoplankton taxonomic composition and an increase in phytoplankton biomass during winter and spring as a result of temperature increases during these seasons (Anneville et al., 2018). Moreover, the abundance and taxonomic composition in zooplankton gradually changed with time because of decreases in food quality (Anneville et al., 2007) and trophic status (Molinero et al., 2006), and increased fish predation (Nõges et al., 2018).

**Changes in phytoplankton composition**

Long-term monitoring of phytoplankton in Lake Geneva is one of the longest time series in the world (Fig. 5). It shows deep changes in its biomass and moreover on its taxonomic composition from the 70ies to nowadays. These changes were mostly controlled by bottom-up factors, such as nutrient level. For instance, the decrease of Dinophyceae from 1974 is mostly due to *Ceratium hirundinella*, a species preferring eutrophic epilimnia (Reynolds et al., 2002) and which is nowadays quite rare in Lake Geneva phytoplankton compared to 40 years before. On the other hand, some species known to prefer oligotrophic waters and which have mixotrophic capacities (Padisak et al., 2009), mostly belonging to the Chrysophyceae (e.g. *Dinobryon* spp., *Kephyrion* spp., *Chrysolykos planktonicus*) showed an increase in their biomass and are common in the summer epilimnion of the last decade, whereas they were absent in the 70ies. Inside the diatom class, some species have disappeared, like *Stephanodiscus binderanus*, a centric usually observed in eutrophic waters, and have been replaced by species like *Cyclotella costei*, characterizing re-oligotrophication stage in Lake Geneva (Rimet et al., 2009). Global warming has an impact on lake stratification and phytoplankton composition; it is the case with *Mougeotia gracillima*, a filamentous Zygophyceae which showed higher biomass in the metalimnion after year 2000, sometimes blooming when a thick layer of the epilimnion is stratified (Tapolczai et al., 2015).

Such changes in the phytoplanktonic structure and biomass have also been observed in Lake Bourget. During almost 20 years, the toxic and filamentous cyanobacteria *Planktothrix rubescens* developed and bloomed (representing between 20% and more than 50% of the algal biomass each year) despite efforts to reduce phos-
phorus loading since the early 1980’s. Between 1996 and 2009 blooms occurred and persisted while the lake was mesotrophic. By the end of 2009, however, a conjunction of events occurred so that the cyanobacterium disappeared. Interactions among nutrients, light availability, temperature and water column stability as well as zooplankton grazing could be identified as important factors for explaining bloom collapse (Jacquet et al., 2014b). In parallel, as for the Lake Geneva, mixotrophic algal species and picophytoplankton groups have significantly increased during the last decade in response to the re-oligotrophication and the disappearance of *P. rubescens*. It is noteworthy, however, that 2016 and 2017 constituted again two years where the species raised again despite nearly oligotrophic conditions (Jacquet et al., 2018).

**INFORMATION SYSTEM STRUCTURE**

The OLA-IS is built around generic components offering common functionalities to various observatories dedicated to a variety of environmental topics (lakes, grasslands, forests and crops) (Schellenberger et al., 2019b; https://doi.org/10.15454/QJJJZU). However, several features dedicated specifically to lakes were added (specific datatype modules), making the OLA-IS specific (Schellenberger et al., 2019a; https://doi.org/10.15454/VBWYWG). The OLA-IS was written in Java as a Maven project and is divided into several lake datatype modules. The OLA database was designed according to a relational model (Monet et al., 2020a, https://doi.org/10.15454/OHXVJY) managed by PostgreSQL Release 12. Technically, the database has been designed to make the modeling as generic as possible so that each type of data (for instance: phytoplankton data type, probes data type, etc.) is modeled in the same way. In addition, the metadata tables are the same regardless of the type of data studied.

The web interface is based on the JSF framework, and Primefaces, Spring and Hibernate graphics components libraries complement the technologies. The IS is accessed online through an apache server (https://si-ola.inrae.fr).

The interface provides information on the available data and states the Terms of Use (ToU) as stipulated in the data access license (Monet et al., 2020b, https://doi.org/10.15454/HHN2GA). Both the interface and the data content (variable names, units, etc.) are available in English and French.

The interface proposes extraction by data type. For each data type, the user can choose the lake, the time period and the desired variables. After extraction is complete, the user can download the data as text (.cvs) files. Online notifications show the user if the extraction is in progress or completed, and this does not imply the user

---

**Fig. 4.** Interannual dynamic of phytoplankton (µg.L⁻¹) and microcrustacean (ind.m²) biomasses in lake Geneva.
The Observatory on Lakes (OLA) database must remain logged in. Another interface, restricted to administrators, enables management of the users, the rights granted to them, and control, deposit and publication of validated data.

OLA is part of the French national research infrastructure Anaee-f. Anaee-f works to publish its metadata and its data through the use of the semantic web (https://www.anaee-france.fr/en/infrastructure-services/).

**Fig. 5.** Taxonomic composition of lake Geneva phytoplankton expressed in biomass (µg/L) from 1974 to 2018 (each value is an annual mean of the water column integrated phytoplankton samplings carried out 2 times a month except once a month between December and February).
modelling-and-data/semantic-reference-framework). OLA fits in with this perspective and uses the vocabulary of thesaurus Anaee-F, data are semantically annotated using the AnaEE-F (http://agroportal.lirmm.fr/ontologies/ANAEEFTHESES and OBOE ontologies (Madin et al., 2007). In this way, OLA is developed according to FAIR Data principles (Findable, Accessible, Interoperable, Re-usable) (Clobert et al., 2018; Pichot, 2017, Wohner et al., 2019).

DATA MANAGEMENT PLAN

Development and implementation of adapted policies and processes are essential to maintain the integrity, security and usability of the data. Data validation, data access and sharing policy are important component of the data management plan.

Quality control

The OLA-IS contains only curated data. All along the data collection, from field sampling to data archiving, rigorous quality controls are carried out (Fig. 6). These data follow a controlled vocabulary included in the AnaEE-France thesaurus (Clastre et al., 2018).

DATA ACCESS AND SHARING POLICY

Before downloading and using the data, a user must register on the SI, agree to the GCUs and formulate the data request. The data request is evaluated by the OLA scientific committee (scientists of the research unit CARTEL). Once the request is validated, the user is provided a login which allow to enter the OLA-IS for a period of three months and download the data. Access to the IS and the GCUs is available at https://si-ola.inrae.fr/. The GCUs describe the rights and obligations of the user when using the data provided by the OLA-IS. Data are accessible free of charge via registration on the website and compliance with the GCUs. The data are protected under copyright law, and compliance with the GCUs will not change this. All the data stored in the OLA-IS are usable by any applicant, under the condition of providing reference to the original source of the data in any communication by indicating the lake(s) studied, as given in the GCUs and by citing the present paper:

Rimet et al. (2020) The Observatory on Lakes (OLA) database: Sixty years of environmental data accessible to the public.

Furthermore, the user is asked to transmit to the OLA-IS manager a copy of the work wherein the downloaded data were analyzed and valorized. The GCUs expire after a period of five years after which the user must destroy the data. The user must formulate a new request in case the data are required for a new project. Data cannot be shared or stored in another public database.

CONCLUSIONS AND PERSPECTIVES

Since the OLA-IS was created in December 2014, an average of more than one peer-reviewed publication providing limnological data from the OLA has been accepted monthly. More than 55 publications have been referenced on Google Scholar from 2015 to 2018. This also corresponds to the monthly data demands received through the OLA-IS website. It is obvious that this website and IS fuel data-sharing and scientific collaborations between environmental scientists. Proof of this is found in large international co-authorships of several papers that include OLA data. Examples are Sharma et al. (2015), where several tens of scientists studied surface lake temperatures, Taranu et al. (2015), where cyanobacteria increases along with nutrient concentrations were studied on a large temporal and spatial scale, and Leach et al. (2018), where the relative importance of light and thermal stratification were found to be key drivers of deep chlorophyll maxima structure in a variety of lakes. Other examples include studies of the impact on stratification and response to climate change of lake morphometry (Kraemer et al., 2015), the rise in chloride concentrations in European and American lakes (Dugan et al., 2017b), or the widespread diminishing anthropogenic effects on calcium in freshwaters because of industrial acid depositions (Weyhemmeyer et al., 2019).

However, even if the OLA-IS is an excellent platform for data accessibility, the inclusion of new types of data must be scheduled. Some of these new types of data are related to technological evolutions that enable high-throughput data (DNA high-throughput sequencing and high-frequency sensors) to be accessed. Until now, biological analyses have been carried out under microscopes (e.g., for phytoplankton and zooplankton), requiring experts trained in taxonomical recognition. An alternative is DNA barcoding (Hebert et al., 2003) and high-throughput-sequencing, which enables rapid analyses of many environmental samples at lower costs compared to the involvement of specialized analysts. This method was tested on mock communities of algae (Kermarrec et al., 2013) and then on natural algal communities (Vasselon et al., 2017; Rivera et al., 2018). It also enables global assessments of protist diversity, including rare biosphere (Taib et al., 2013) and small eukaryotes whose taxonomy remains still not well-established (Lepere et al., 2008; Mangot et al., 2013). Therefore assessments of diversity are enlarged compared to the restricted diversities observable using optical devices. Such methodologies can also be applied to fish (Pont et al., 2018) and macroinvertebrates (Elbrecht et al., 2017), also providing promising results. The OLA lakes are already studied using metabar-
Fig. 6. Simplified flow chart for data validation and curation before integration in the OLA information system.
Recent technological developments have increased the number of variables monitored in lakes using automatic high-frequency sensors. Marcé et al. (2020) and Cantour, de la Vanoise and 3CMA are thanked for funding the relevant data production. The GIS Lac Sentinelle is also thanked.

High-throughput data storage (and accessibility) also presents an imminent challenge related to the development and storage of data from high-frequency sensors. Recent technological developments have increased the number of variables monitored in lakes using automatic high-frequency sensors (Marcé et al., 2016). To facilitate posterior data handling and interpretation, it is critical that an interface as the OLA-IS provides storage capacity and sharing access for this type of data, given it is becoming essential in near-future limnological surveys.

Monitoring additional matrices and habitats in lakes and their integration into the OLA-IS are already envisaged. For instance, data on fish populations, pico-cyanobacteria and new sampled habitats such as the littoral zone, are now routinely collected and are planned to be archived in the OLA-IS whose architecture allows to host new data. Therefore, integrating high-throughput data (DNA and sensors), new sampling sites and matrices into the OLA-IS would considerably enrich and diversify this database, offering thus new opportunities to enlarge scientific collaborations.

To conclude, the OLA-IS success was made possible by three items. First thanks to the quality of its data and the presence of experts in different fields (taxonomy, chemistry, physics). Second, thanks to the traceability of the data, since metadata are associated to each data, detailing the protocols of acquisition, the material used and the responsible persons. Finally, thanks to the permanence and accessibility of the IS made possible by the commitment of the regular staff dedicated to its maintenance and climate variability: mediators of synchrony in phytoplankton changes among European peri-alpine lakes. Freshwater Biol. 50:1731-1746.


