Assessing ESA Climate Change Initiative data for the monitoring of phytoplankton abundance and phenology in deep lakes: Investigation on Lake Geneva

Mona Bonnier, Orlane Anneville, R. Iestyn Woolway, Stephen J Thackeray, Guillaume P Morin, Nathalie Reynaud, Frédéric Soulignac, Thierry Tormos, Tristan Harmel

Journal of Great Lakes Research 50 (2024) 102372

Contents lists available at ScienceDirect

Journal of Great Lakes Research

journal homepage: www.elsevier.com/locate/jglr

Assessing ESA Climate Change Initiative data for the monitoring of phytoplankton abundance and phenology in deep lakes: Investigation on Lake Geneva

Mona Bonnier [a,*], Orlane Anneville [a,1], R. Iestyn Woolway [b], Stephen J. Thackeray [c], Guillaume P. Morin [d], Nathalie Reynaud [e,f], Frédéric Soulignac [g], Thierry Tormos [h,g], Tristan Harmel [d,*]

[a] Université Savoie Mont Blanc - INRAE, UMR CARRTEL, pôle ECLA, Thonon-les-Bains, France
[b] School of Ocean Sciences, Bangor University, Menai Bridge, Anglesey, Wales
[c] UK Centre for Ecology & Hydrology, Library Avenue, Bailrigg, Lancaster, UK
[d] Magellium, 1 Rue Ariane, Ramonville-Saint-Agne, France
[e] INRAE, Aix Marseille Univ, RECOVER, Aix-en-Provence, France
[f] Pôle R&D « ECLA », Aix-en-Provence, France
[g] Commission Internationale Pour la Protection des eaux du Léman, Route de Duillier 50, 1260 Nyon, Switzerland
[h] OFB, Unité ECLA, Aix-en-Provence, France

ARTICLE INFO

Communicated by Caren Binding

Keywords:
Water quality
Earth observation
Peri-alpine Lake
Long-term trends
Remote sensing
Chlorophyll *a*

ABSTRACT

Lake water quality assessment requires quantification of phytoplankton abundance. Optical satellite imagery allows us to map this information within the entire lake area. The ESA Climate Change Initiative (ESA-CCI) estimates Chl-*a* concentrations, based on medium resolution satellite data, on a global scale. Chl-*a* concentrations provided by the ESA-CCI consortium were analyzed to assess their representativeness for water quality monitoring and subsequent phenology studies in Lake Geneva. Based on vertically resolved *in-situ* data, those datasets were evaluated through match-up comparisons. Because the underlying algorithms do not take into account the vertical distribution of phytoplankton, a specific analysis was performed to evaluate any potential biases in remote sensing estimation, and consequences for observed phenological trends. Different approaches to data averaging were performed to reconstruct Chl-*a* estimates provided by the remote sensing algorithms. Strong correlation (R-value > 0.89) and acceptable discrepancies (*rmse* ~ 1.4 mg.m^{-3}) were observed for the ESA-CCI data. This approach permitted recalibration of the ESA CCI data for Lake Geneva. Finally, merging satellite and *in-situ* data provided a consistent time series for long term analysis of phytoplankton phenology and its inter-annual variability since 2002. This combination of *in-situ* and satellite data improved the temporal resolution of the time series, enabling a more accurate identification of the timing of specific spring events characterising phytoplankton phenology.

1. Introduction

Pressures arising from demographic and economic growth, and associated global environmental change, challenge the use and preservation of inland waters. Increasing human pressures are resulting in a rapid degradation of lake ecosystems worldwide (Jenny et al., 2020). Appropriate management practices and environmental measures can limit the negative impacts of human induced changes such as eutrophication or global warming (Jacquet et al., 2014; Kao et al., 2020). However, to establish effective practices, it is essential to have a sound understanding of how ecosystems function, based in particular on accurate observations of the biotic and abiotic elements that influence them. Accurate monitoring of biotic components of lacustrine ecosystems, such as phytoplankton abundance and composition, is required by

* Corresponding authors.
E-mail addresses: mona.odile.bonnier@gmail.fr (M. Bonnier), tristan.harmel@magellium.fr (T. Harmel).
[1] Given their role as Editor, Orlane Anneville had no involvement in the peer-review of this article and has no access to information regarding its peer-review. Full responsibility for the editorial process for this article was delegated to Caren Binding.

https://doi.org/10.1016/j.jglr.2024.102372
Received 10 July 2023; Accepted 13 May 2024
Available online 8 June 2024

various legislations to provide essential evidence on water quality, trophic status, and ecological status (e.g., the European Water Framework Directives (WFD), the post-2020 Global Biodiversity Framework).

Amongst a widening array of monitoring technologies, remote sensing is increasingly being used to monitor water quality and the ecological status of aquatic environments. Satellite Earth Observation can provide repeated synoptic coverage of ecosystems at multiple scales (Kavanaugh et al., 2021; Maberly et al., 2020; Tormos et al., 2021), with spatially resolved data better capturing horizontal heterogeneities in the various parameters monitored. In addition, the ever-improving spatial, spectral, and temporal resolutions of satellite Earth Observation underscore the need to improve the quality and robustness of the various estimation algorithms for chlorophyll *a* (Chl-*a*) concentration, colored dissolved organic matter (CDOM) concentration, surface water temperature, and cloud cover (Ansper and Alikas, 2018; Topp et al., 2020).

A key metric used to evaluate water quality and ecosystem state in lakes is the abundance of phytoplankton. A number of environmental factors, linked to local pollution or climate change, can influence phytoplankton biomass, its vertical distribution in the water column, and its phenology (Ho et al., 2019; Kraemer et al., 2017; Thackeray et al., 2013). As a result, phytoplankton biomass and temporal dynamics are often considered key indicators of environmental and climate change (Adrian et al., 2009). This potential can be realized through the accurate monitoring of its seasonal dynamics. Indeed, even subtle changes in the timing of seasonal phytoplankton biomass can have numerous knock-on effects on aquatic ecosystems, such as a desynchronization of ecological interactions and unbalanced carbon fluxes in aquatic food webs, leading to detrimental consequences for secondary consumers (Winder and Schindler, 2004).

Phytoplankton abundance is monitored by measuring several proxies, including the concentration of Chl-*a*, an indicator widely used to estimate the primary production of an ecosystem (Chang et al., 2022). However, different approaches to measuring this proxy vary in their spatial and temporal coverage and resolution. *In-situ* measurements of Chl-*a* might provide accurate depth-resolved estimates of phytoplankton abundance, but only at specific point locations across the surface of a lake. Current remote sensing techniques provide greater (horizontal) spatial coverage, but do not account for the vertical dimension related to phytoplankton distribution (e.g., the presence of deep chlorophyll maxima). Moreover, the information derived by remote sensing depends upon light reflected back from the water column, which in turn depends on the transparency of the environment. The estimation of Chl-*a*, inherent optical properties (IOP), and associated biogeochemical parameters is therefore potentially impacted by the vertical distribution of water column components, such as total suspended matter (TSM) or colored dissolved organic matter (CDOM) (Hu et al., 2021; Nouchi et al., 2018).

To quantify the effect of vertical (water column) heterogeneity on satellite data, we exploited a long-term dataset that contains vertical profiles of Chl-*a* concentration in Lake Geneva, a large and deep lake located at the border between Switzerland and France. Lake Geneva is particularly appropriate for this topic with its comprehensive *in-situ* monitoring since the late 1950 s, a period of increasing eutrophication due to a growing human population and poor wastewater collection and treatment systems. Since the 1990′s, the phytoplankton community of Lake Geneva has undergone major changes due to the re-oligotrophication of its waters. This re-oligotrophication of the lake favored the growth of species indicative of oligotrophic lakes and resulted in a deepening of phytoplankton activity linked to modifications of the vertical profiles in phosphorus concentration (Anneville et al., 2002, 2019). The latter showed extremely low concentrations in the euphotic zone along with a deepening of the phosphocline. In turn, the vertical distribution of phytoplankton was significantly impacted with a decrease in surface concentrations and a deepening of the Chl-*a* maxima (Anneville et al., 2013). These changes in the state of Lake Geneva might bias the interpretation of satellite imagery from the lake (Nouchi et al., 2018; Pitarch et al., 2014) and, as such, should be investigated to inform the exploitation of remote sensing data from different missions and different processing algorithms.

In this study, we evaluated moderate spatial resolution satellite data to quantify their accuracy in representing the seasonal dynamics and phenology of deep lake phytoplankton. We hypothesized that, despite being less sensitive to deep phytoplankton growth, satellite information may be appropriate for analyzing long-term changes such as phenological variability of subsurface primary production. We addressed this hypothesis by analyzing and comparing *in-situ* measurements from Lake Geneva, and Sentinel-3 datasets provided by the European Space Agency Climate Change Initiative (ESA CCI, https://climate.esa.int/fr/), integrating the vertical dimension from the *in-situ* datasets. The aim of this study was to provide an unbiased assessment of Chl-*a* concentrations in Lake Geneva, and to evaluate the proportion of water column phytoplankton biomass potentially detected by satellite Earth Observation. Based on this approach, we created a long-term time series of Chl-*a* concentrations by combining satellite and *in-situ* measurements to obtain a higher temporal resolution data set before analyzing phenological change. Finally, we discuss these results as part of a more general framework to guide future assessments of the utility of satellite data for long-term monitoring, to provide the evidence needed for lake management (Feng et al., 2021).

2. Materials and methods

2.1. Lake Geneva and data acquisition

The *in-situ* data used in this study were collected as part of the lake environmental monitoring coordinated by the French and Swiss commission for the protection of Lake Geneva water (CIPEL). These data were made available through the OLA database (Rimet et al., 2020). To evaluate the long-term variability of the phytoplankton abundance, in this study we focused on data collected from sampling station SHL2, as used for the WFD monitoring of the lake. This station is located upon the deepest part of the lake (45.453°N, 6.594°E, Fig. 1). The sampling frequency is monthly in winter and fortnightly for the rest of the year. Water samples were collected at ten different depths (0, 1, 2.5, 3.5, 5, 7.5, 10, 15, 20, 30 m) and then analyzed at the laboratory on the day of sampling. In addition, a multi-parameter probe (Probe reference: CTD90M, serial no 1079 and manufacturer: Sea & Sun Marine Tech (Germany)) was used to quantify the vertical profiles of the parameters of interest at a finer vertical resolution (on average 2.7 measurements per meter). The period analyzed in this study ranged from 2003 to 2021, which included synchronous Chl-*a* and radiation measurements.

2.2. PAR and transparency

The photosynthetically active radiation (PAR) is the radiative energy contained in the visible light spectrum (i.e., wavelengths from 400 to 700 nm) which is potentially used during photosynthesis by primary producers. In this study, PAR data were used to estimate the daily diffuse attenuation coefficient (K_d). The K_d defined the rate of decrease in downward irradiance at a depth z, i.e., K_d is high when water transparency is low. The K_d was expressed in units of m^{-1} and, in the case of PAR attenuation with depth z, is defined as follows (Mueller et al., 2003):

$$PAR(z) = PAR(z=0)\exp\left(-\int_z K_d(z')dz'\right) \quad (1)$$

In this study, a least-squares optimization was used to fit this function to the actual PAR data taken from 0.8 to 30 m depth. Note that the fit was performed in the logarithmic space (i.e., taking the logarithm of Eq. (1) and assuming a constant value to minimize the influence of potential outliers related to data acquisition particularly close to the surface, due

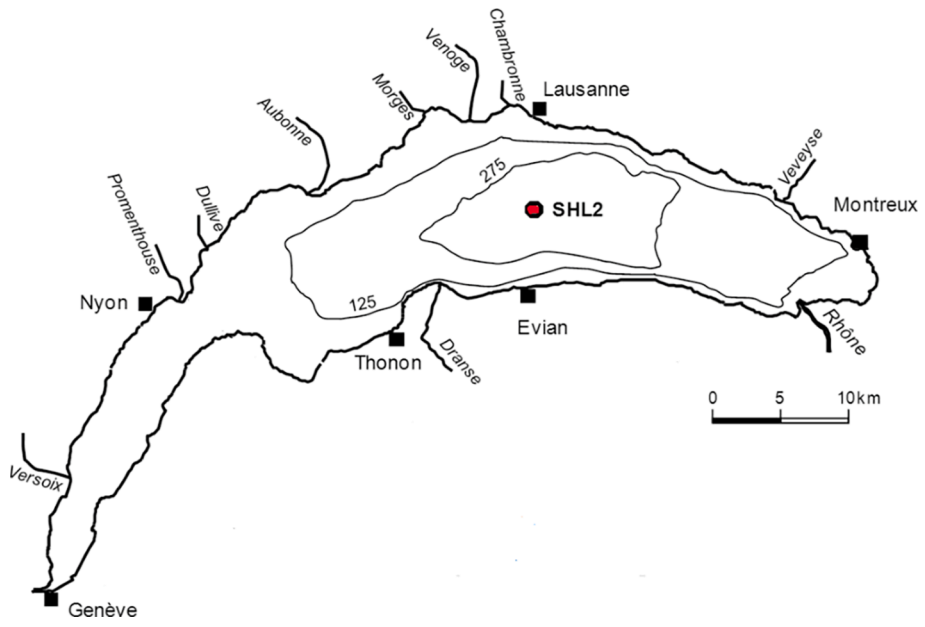

Fig. 1. Location of the sampling station (SHL2, 46.453°N, 6.594°E) for long-term monitoring of Lake Geneva.

to imprecise depth measurements and wave focusing (Gernez and Antoine, 2009). This method gave a value of K_d per sampling acquisition. Nevertheless, inconsistent K_d values were obtained for a few dates between 2003 and 2017 (uncalibrated probe period). To remove those outliers and to maximize the number of exploitable dates, we calculated an average K_d over the entire time series (2003–2021) which was applied to days with a K_d not included in the following interval: 0.12—–0.34. This interval was defined using the maximum and minimum K_d value during the calibrated period (2017–2021). This method was validated through Pearson correlation of the K_d-derived transparency with *in-situ* 30 cm white disk and Secchi disk data (Quétin et al., 2020) (cf: 3.4 Subsurface layers). Pearson's coefficients obtained were greater than 0.5 with p-values less than 0.05.

2.3. Chlorophyll-a concentrations

Chl-*a* concentrations were estimated following two different approaches: (i) laboratory analysis of collected water samples, (ii) from the multi-parameter probe via fluorescence measurements directly within the water column. The first approach was based on spectrophotometer measurements following the Strickland-Parsons protocol before 2012, and then by SCOR-UNESCO (Rimet et al., 2020). Unlike the first method, the fluorescence approach might be highly biased due to non-photochemical quenching (NPQ) effects on the intrinsic fluorescence of the phytoplankton cells (Gupana et al., 2022; Muller et al., 2001). The NPQ effect is well-known but highly challenging to correct (Huot and Babin, 2010). It can result in a significant underestimation of concentrations at the lake surface. Here, the methodology used to correct for NPQ was based on a modeling procedure for fitting the fluorometer digital counts and the unbiased laboratory measurements over the entire vertical profile. In addition, a specific method was used to account for the uncertainty attached to the depth $\underline{z}$ of each sample used for the laboratory measurements. Note that the depth of the fluorometer probe is well quantified based upon concurrent pressure measurements. Following this scheme, the calibration coefficients $C(z)$ (Eq. (2)) were obtained for each discrete water sample taking into account the sampling depth uncertainty $\alpha(z)$ corresponding to a normal distribution with a standard deviation σ_z of 1 m:

$$C(\overline{z}) = \frac{Chla_{spectro}(\overline{z})}{Chla_{fluo}(\overline{z})} \tag{2}$$

with

$$Chla_{fluo}(\overline{z}) = \frac{\sum_z \alpha(z) Chla_{fluo}(z)}{\sum_z \alpha(z)} \tag{3}$$

and

$$\alpha(z) = \exp\left(-\frac{1}{2}\left(\frac{z-\overline{z}}{\sigma_z}\right)^2\right) \tag{4}$$

where $Chla_{spectro}$ is the Chl-*a* concentration measured at the laboratory, and $Chla_{fluo}$ is the Chl-*a* concentration measured by the fluorometer probe. The uncertainty (i.e. standard deviation) σ_z was set to one meter as indicated in the OLA database. The fluorometer signal was thus averaged for the depth z based on equations (2) and (3). Finally, the coefficient $C(z)$ was linearly interpolated from 0.5 to 30 m at 0.3 m intervals to recover the fine vertical resolution of the probe profile.

To properly compare *in-situ* Chl-*a* data with that estimated by satellite Earth Observation, the *in-situ* data have to be integrated over depth, e.g., considering heterogeneity of the vertical profile. Several practical methods can be used. The simplest approach is to take the mean value within a given water layer. Nevertheless, this method is prone to errors due to the presence of subsurface Chl-*a* maxima (Pitarch et al., 2014). Typically, deep Chl-*a* maxima will have less impact on satellite measurements than the same maxima near the surface. To analyze the impact of variable vertical distributions of Chl-*a*, we averaged for each day (81 days) the *in-situ* data between 0.5 m and a limit depth (tested depths 2, 5 and 8 m) (**cf.** Appendix S1, **Fig. S1.2**). Another approach is to consider an optically weighted integration of the Chl-*a* over depth. Several methods were proposed based on full or approximated radiative transfer computations (Gordon and Clark, 1980; Zaneveld et al., 2005). Note that more elaborate approaches should consider full radiative transfer simulations (Piskozub et al., 2008) but are not applicable to our data set due to the lack of vertically-resolved measurements of the inherent optical properties.

Here, the two datasets of depth-resolved Chl-*a* and of the diffuse attenuation coefficient were exploited to properly compare the *in-situ* Chl-*a* with estimates obtained from satellite optical data. Thus, the weighted Chl-*a* was computed from values obtained through Eq. (3) and the following vertical integration:

$$Chla_{weighted} = \frac{\sum_z Chla(z)e^{-2K_d z}}{\sum_z e^{-2K_d z}} \quad (5)$$

This approach accounts for Chl-*a* concentrations throughout the water column, while applying higher weighting to surface concentrations (i.e., the concentrations that are potentially visible to the satellite).

2.4. Subsurface layers

We determined the depths at which the PAR values were equal to 5 %, 1 %, 0.1 % of the PAR values measured at the surface, using the calculated daily K_d obtained with the method introduced earlier. For this step, several formulae were tested: $1/K_d$ which is similar to the first optical depth, $1.7/K_d$ which is close to the Secchi disk depth, and finally $4.605/K_d$ for the euphotic depth (Roy and Das, 2022). These depths were used to estimate the depth limit of satellite visibility, over which the fluorometric measurements were averaged.

Finally, we estimated the total Chl-*a* that may be observed by satellite Earth Observation as the proportion (in %) of the depth-integrated Chl-*a* concentration measured *in-situ*, and the number of Chl-*a* maxima shallower than the visibility layers calculated with the boundary depths. These proportions were derived from the estimated total Chl-*a* in the water column (0.5 to 30 m) per day. Daily total Chl-*a* was calculated using the trapezoidal integration method applied to the *in-situ* vertical Chl-*a* profile. This method was based on the numerical calculation of the amount of Chl-*a* in the water column. This relies on the numerical calculation of an integral based on a linear interval interpolation, i.e., the interpolated depth values.

2.5. Satellite observations

2.5.1. CCI-lake data

The satellite data used in this study originate from the CCI-lake processing chain applied to the MERIS and the two Sentinel-3 OLCI satellite missions as summarized in Carrea et al., (2023). With: (i) the remote sensing reflectance (R_{rs}, in sr^{-1}) obtained after application of pixel masking based on Idepix software and the atmospheric correction processor POLYMER (Steinmetz et al., 2011), (ii) the concentration of Chl-*a*. This processing chain was based on a set of Chl-*a* algorithms applied per pixel according to prior recognition of the optical water types (OWT). For the typical OWT of Lake Geneva, the standard OC2 algorithm described below (O'Reilly et al., 1998, 2000) is mainly used (see Table 8 in (Carrea et al., 2023) for details).

2.5.2. Retrieval of the optical water types

Following the CCI-Lake approach, the optical water types described in Spyrakos et al. (2018) were used to map the OWT for each pixel from the spectral R_{rs} values. The normalized reflectance and the definition of those OWT are shown in Fig. S5. The OWT classification followed the method used by **Uudeberg et al. (2019)** based on the maximum likelihoods computed from individual spectra based on the combination of two robust metrics: spectral correlation similarity (SCS) and modified spectral angle similarity (MSAS). The Pearson statistical correlation can be used as a similarity measure through the SCS value defined for the OWT number *j* as:

$$SCS_j = \frac{1}{n-1}\left(\frac{\sum_{i=1}^{n}(R_{rs}(\lambda_i) - \overline{R_{rs}})\left(nR_{OWTj}(\lambda_i) - \overline{nR_{OWTj}}\right)}{\sigma_{R_{rs}}\sigma_{OWTj}}\right) \quad (6)$$

Where *n* is the number of spectral bands used, nR_{OWTj} is the normalized reflectance attached to the OWT number *j*, the overlying bar and σ symbols stand for the average and standard deviation (Table 1) computed over the spectral bands. The second metric MSAS provides the spectral angle between this pair of spectral vectors and can be formulated as follows:

$$\alpha_j = \arccos\left(\frac{\sum_{i=1}^{n} R_{rs}(\lambda_i) nR_{OWTj}(\lambda_i)}{\sqrt{\sum_{i=1}^{n} R_{rs}(\lambda_i)^2}\sqrt{\sum_{i=1}^{n} nR_{OWTj}(\lambda_i)^2}}\right) \quad (7)$$

For normalization purposes, MSAS for the OWT number *j* is given by:

$$MSAS_j = \frac{2\alpha_j}{\pi} \quad (8)$$

Finally, the OWT that maximizes the following expression δ is attributed to the pixel:

$$\delta_j = 10\left(SCS_j + \frac{1 - MSAS_j}{2}\right) \quad (9)$$

2.5.3. Algorithm for chlorophyll-a estimation

Numerous algorithms were developed to convert (or "inverse") the remote sensing reflectance in terms of Chl-a concentration (Neil et al., 2019). Among those algorithms, the "OCx" algorithms are based on a non-linear relationship between Chl-*a* and the ratio of the blue over green reflectance (O'Reilly et al., 1998, 2000). The OCx formulation can be written from the measured remote sensing reflectance (R_{rs}) as a fourth degree polynomial of the logarithm of the blue-green ratio:

$$\log_{10}(Chl\text{ - }a) = \sum_{i=0}^{4} a_i \log_{10}\left(\frac{R_{rs}(\lambda_{blue})}{R_{rs}(\lambda_{green})}\right)^i \quad (10)$$

The spectral bands used for the blue-green ratio might differ between satellites. In addition, several formulations were developed to make use of several spectral bands for the blue term. The OC2 formulation is based on the ratio defined by the bands centered on 490 nm (blue) and 560 nm (green); OC3 takes the maximum R_{rs} between 443 nm and 490 nm for blue reference. Similarly, OC4 takes the maximum R_{rs} between 443, 490 and 510 nm, and the maximum between 412, 443, 490 and 510 nm for OC5. The coefficients a_i were computed based on large-scale matchup comparison with *in situ* data. Typically, the current values for MERIS and OLCI are based on the OC4 formulation obtained after calibration on oceanic and coastal measurements with $a = (0.4249, -3.2097, 2.8972,$

Table 1
Results of phenology comparisons between dates retrieved from ESA CCI calibrated data, $Chla_{weighted}$, $Chl\text{-}a_{0\text{-}5m}$ and $Chla_{Merged.}$. Shown are the Pearson correlations between dates estimated from $Chla_{ESACCI}$ and $Chla_{weighted}$ (Cor. cci ~ wg). Percentage of dates when time-lags between events are higher than 21 days (% Large mismatch cci ~ wg), linear trend over a common period (2003–2019) for dates from $Chla_{ESACCI}$ and $Chla_{weighted}$ (Corcoeff cci / wg); Pearson correlation between dates from $Chla_{Merged}$ and $Chl\text{-}a_{0\text{-}5m}$ (Cor. merged ~ 0–5 m), percentage of dates when time-lags between events are higher than 21 days (% Large mismatch merged ~ 05 m), trend in the dates retrieved from $Chla_{Merged}$ over the period 2003–2021 (Sen's slope_merged).

	Cor. cci ~ wg	% Large mismatch cci ~ wg	Corcoeff cci / wg	Cor. merged ~ 0–5 *m*	% Large mismatch merged ~ 05 *m*	Sen's slope_merged
SOS	0.75***	8	−0.38/-0.25	0.93***	6	−1
DOM	0.67**	8	−0.16/-0.07	0.88***	6	−0.47
COG	0.51*	0	−0.08/-0.12	0.89***	0	−0.17

Significant: * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

−0.7526, −0.9826) for MERIS, and a = (0.4254, −3.2168, 2.8691, −0.62623—1.0933) for OLCI. Note that the CCI-lake processing chain is based on the OC2 formulation which has been recalibrated on inland Chl-*a* measurements with a = (0.1731, −3.9630, −0.5620, 4.5008, 3.0020).

2.6. Data comparison strategy

The quality of the CCI products was assessed based on match-up comparison between satellite and *in situ* measurements. To assess the sensitivity of the dataset comparison to spatial heterogeneities or temporal dynamics several modalities were tested: (i) megapixel size windows with 3x3 or 5x5 pixels extracted around the SHL2 location, (ii) time windows of 1, 2, 3 and 4 days between the sampling date and the satellite overpass, (iii) changing the vertical integration method of the *in situ* Chl-*a*. Statistical indicators were computed within those regions of interest such as the mean, median and standard deviation. Classically, satellite images are excluded from the comparison if the standard deviation is greater than 20 % of the average value (Hlaing et al., 2013) (**cf. Suppl. Mat, Fig. S1**). Following other studies (Pahlevan et al., 2021; Seegers et al., 2018) the statistical metrics were computed both in the linear-space and the log-space. Those statistical indicators were used to quantify the performance of the Chl-*a* satellite retrievals (noted *y* below) against the reference *in situ* data (noted *x*) and to be comparable with other studies from the literature. A linear fit was first applied to provide the equation of the regression line, and the Pearson correlation (*r-value*) was computed. In the linear space, the bias, the Root Mean Squared Error (*rmse*) and the mean absolute percentage error (*mape*) were computed as follows:

$$bias = \frac{1}{N}\sum_{i=1}^{N} y_i - x_i \tag{11}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (y_i - x_i)^2} \tag{12}$$

$$mape = \frac{1}{N}\sum_{i=1}^{N} \left|\frac{y_i - x_i}{x_i}\right| \tag{13}$$

In the log-space, the symmetric signed percentage bias, β, and the median symmetric accuracy, ε, were computed as follows:

$$\begin{aligned} \beta &= 10^{Z} - 1 \\ Z &= median\left(\log_{10}(y/x)\right) \end{aligned} \tag{14}$$

$$\begin{aligned} \varepsilon &= 10^{Y} - 1 \\ Y &= median\left(\left|\log_{10}(y/x)\right|\right) \end{aligned} \tag{15}$$

2.7. Phenology analysis

The ESA CCI data availability ranged from 2003 to 2021 which corresponded to 1035 exploitable dates for Lake Geneva. This comprehensive time series is an asset to analyze seasonal, annual, and even decadal changes in surface water biomass through the Chl-*a* estimates. On the other hand, it is also beneficial to merge remote sensing and *in-situ* data to increase the temporal intensity (resolution) of monitoring and obtain information under conditions that challenge remote sensing alone (e.g., haze or clouds over the lake). The merged dataset was established based on the following steps. First, recalibration of the $Chla_{ESACCI}$ data was performed by correcting the Chl-*a* value for the slope and offset observed during the match-up(s) comparison (i.e., harmonization based on the absolute *in-situ* Chl-*a* values). Then, the calibrated $Chla_{ESACCI}$ were merged with $Chla_{weighted}$ values by averaging the daily values when both satellite and *in-situ* data were available.

Specific seasonal events in the phytoplankton time series were quantified using *in-situ* only, satellite only, and merged Chl-*a* data. We adopted a multi-metric approach to our phenology analysis (Thackeray et al., 2012; Thackeray et al., 2013). Specifically, we calculated proxies for the Start-Of-Season (SOS) (Park et al., 2016), Date-Of-Maximum concentration (DOM), and the Center-Of-Gravity (COG) (Edwards et Richardson, 2004) of the spring bloom. The proxies were computed throughout the entire spring growing season from February to the beginning of June. SOS is the first date when Chl-*a* reached a concentration above the 25 % amplitude computed over the spring growing season. The DOM stands for the date of the seasonal Chl-*a* maxima. The COG was the date corresponding to the central point (T) of the area under the monthly means graph.

$$T = \frac{\sum_{i=2}^{6} i \cdot Chl\text{ - }a_i}{\sum_{i=2}^{6} Chl\text{ - }a_i} \tag{16}$$

where i is the month number and Chl-a_i the mean Chl-*a* observed over the month *i*.

These metrics, for each year, were computed for $Chla_{ESACCI}$ calibrated, $Chla_{weighted}$, $Chla_{Merged}$ and $Chla_{fluo}$ (corrected for NPQ) averaged over 0–5 m ($Chla_{0\text{-}5m}$). Dates of biological events were then compared between $Chla_{ESACCI}$ calibrated and $Chla_{weighted}$ to test the differences between *in-situ* and satellite measurements. Comparisons between dates for $Chla_{ESACCI}$ calibrated and $Chla_{weighted}$ were processed for years with one or more observations per month. Temporal trends in the timing of spring biological events were estimated using a linear regression. Phenology comparisons between $Chla_{Merged}$ and $Chla_{0\text{-}5m}$ were performed to evaluate the benefit of merging different sources of data. Comparisons between $Chla_{Merged}$ and $Chla_{0\text{-}5m}$ were processed over the entire period (2003–2021) and temporal trends were assessed using Sen's slope (O'Reilly et al., 2015; Sen, 1968).

3. Results

3.1. In-situ data

Based on the *in-situ* profile data, three main patterns of Chl-*a* vertical distributions frequently occur at the central location of Lake Geneva (SHL2 site). The first is characterized by a maximum of Chl-*a* at the surface (i.e., in the first 5 m) and is potentially detectable by the satellite Earth Observation, during early spring (Fig. 2). The second pattern is characterized by a deep maximum Chl-*a* below 5 m depth (Fig. 2) during summer and autumn. The third pattern presents a vertically homogeneous distribution (e.g. Fig. 2, during winter). The data presented are those obtained after the application and validation of NPQ correction (cf. Suppl. Mat, **Fig. S2, S3**).

Those vertical distributions might in turn impact the estimation of Chl-*a* from satellite optical measurements. To delineate this potential impact, several "optical depth" metrics were used. The range of depths reached by underwater light were on average 4.5 m, 7.7 m and 21 m, estimated using $1/K_d$, $1.7/K_d$ and $4.605/K_d$ respectively, which correspond to the so called first optical depth, the estimate of the Secchi disk depth, and the 1 % euphotic zone (cf. 3.4 Subsurface layers). Between 2003 and 2021, 26.3 % of Chl-*a* maxima were located above the $1/K_d$ layer, 13.6 % and 50.1 % were observed in the depth intervals between the $1/K_d$ and $1.7/K_d$, and $1.7/K_d$ and $4.605/K_d$ thresholds, respectively, and only 10.3 % occurred below the $4.605/K_d$ threshold. Most maxima (60.4 %) were observed below depths determined using $1.7/K_d$ (Fig. 3, note that a maximum is not necessarily a peak in this figure). The water-leaving radiance is highly influenced by optically active constituents within the first optical depth $1/K_d$, whereas 23.4 % of the maxima are within this layer between 2017 and 2021 (calibration probe period).

We also examined whether trends in water column integrated Chl-*a*, and in visibility layers, could affect the trends observed from optical remote sensing. Obviously, the presence of the phytoplankton itself might modify the optical depth but the presence of non-algal particles

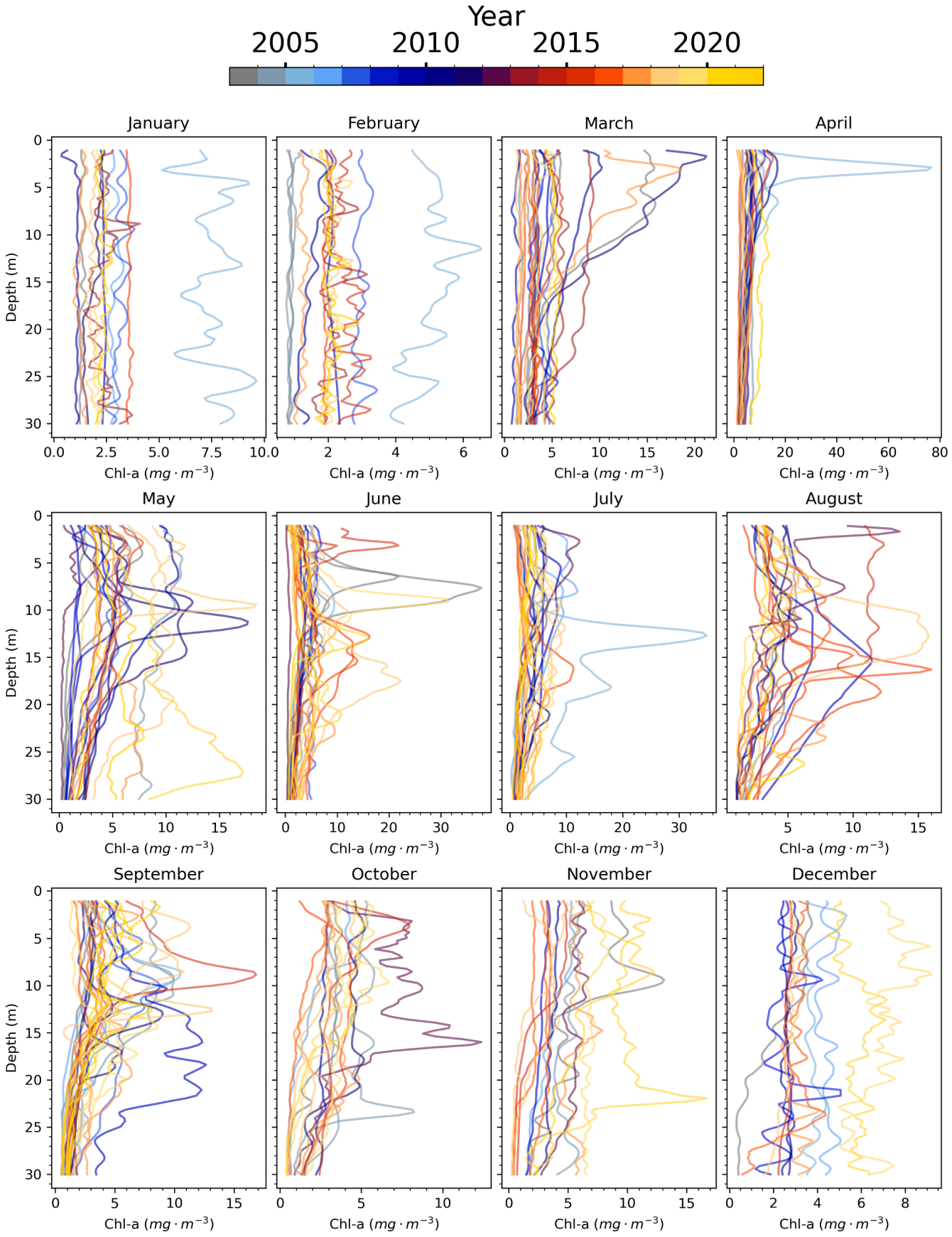

Fig. 2. Time series from 2003 to 2021 of the vertical profiles of Chl-*a* concentration obtained after merging the spectrometric and NPQ-corrected fluorometric measurements at SHL2.

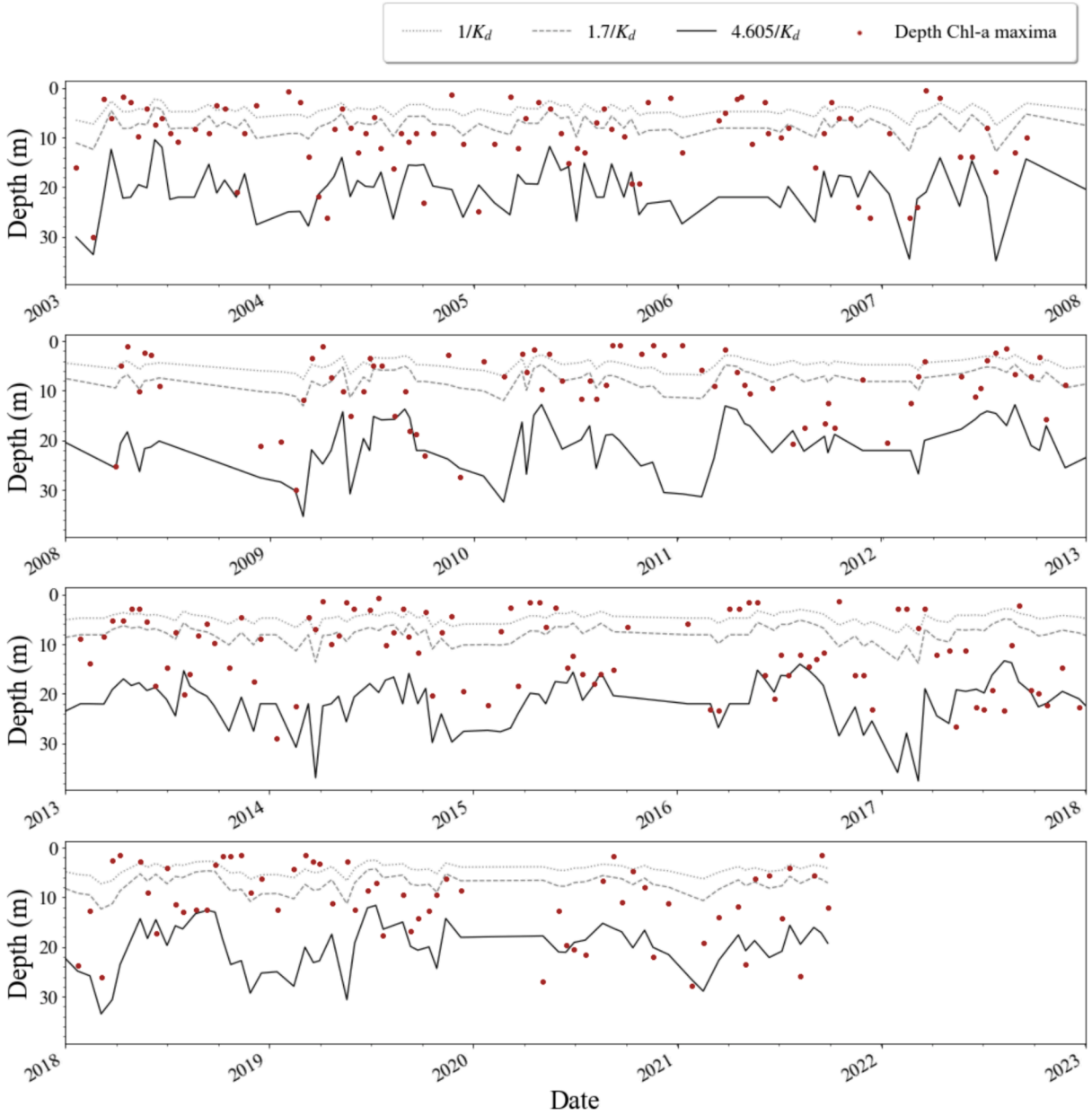

Fig. 3. Depths of Chl-*a* maxima (red dots) determined by Chl-a_{fluo}, and estimates of subsurface layers calculated as 1/ K_d, 1.7/ K_d and 4.605/ K_d (cf. 3.4 Subsurface layers).

may also induce spurious trends in the observed Chl-*a* through remote sensing. The euphotic zone, estimated by 4.605/K_d would result in approximately 70——90 % of Chl-*a* content being included in concentration estimates (Fig. 4), but this proportion would decrease down to 50——70 % in summer. The two other indicators (for Secchi depth and first optical depth) exhibit similar, but less marked, trends between winter and summer with approximately no more than 50 % of the total, whole water column Chl-*a* within those layers. Assuming that the optical depths are robust indicators of the impact of the Chl-*a* content on satellite estimations, these biases would have to be considered in temporal and phenology analyses of satellite time series.

3.2. Optical water types

The OWT classification was applied to all the valid pixels of the CCI images based on the common wavelengths between MERIS and OLCI, namely, the bands centered on 412, 443, 490, 510, 560, 620, 665, 681, 754, 779 nm. Based on 965 images, only three distinctive OWT were retrieved with the majority (879) of the reflectance spectra being representative of the OWT-3 as well as 53 dates of OWT-9 and 33 of OWT-13 (Fig. 5). OWT-3 and OWT-13 are defined as clear water and very clear blue water, respectively. OWT-9 is an intermediate case defined as a common inland water type with higher values in the shorter wavelengths. For these three OWT, the observed spectral shapes indicate very low water-leaving signals beyond 650 nm. These low values preclude the use of retrieval algorithms based on the fluorescence line height or those based on empirical or semi-analytical relationships using the red and near-infrared bands (Gilerson et al., 2010; Gons et al., 2005). For those reasons, the CCI-lake Chl-*a* product is based on the OC2 relationship using the blue-green band ratio for those three OWT. This classification also indicates that in the central part of the lake at the SHL2 point the water column is mostly impacted by phytoplankton and colored dissolved organic matter without strong impact of potential sediment load.

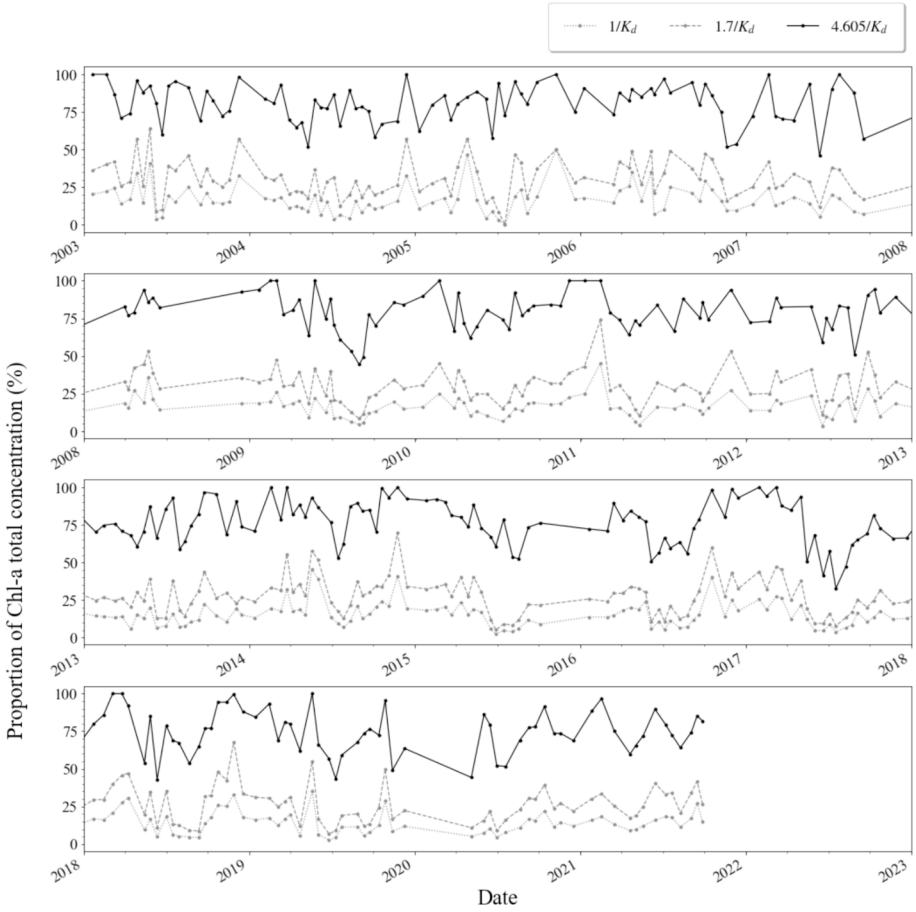

Fig. 4. Time series of the proportion of the total Chl-*a* located within the subsurface layer defined by the bottom limit corresponding to the first optical depth (1/ K_d), the Secchi depth (1.7/ K_d) and the euphotic zone (4.605/ K_d). Average proportion of 1/ K_d layer: 15.8 %, 1.7/ K_d: 28.9 % and 78.1% for 4.605/ K_d.

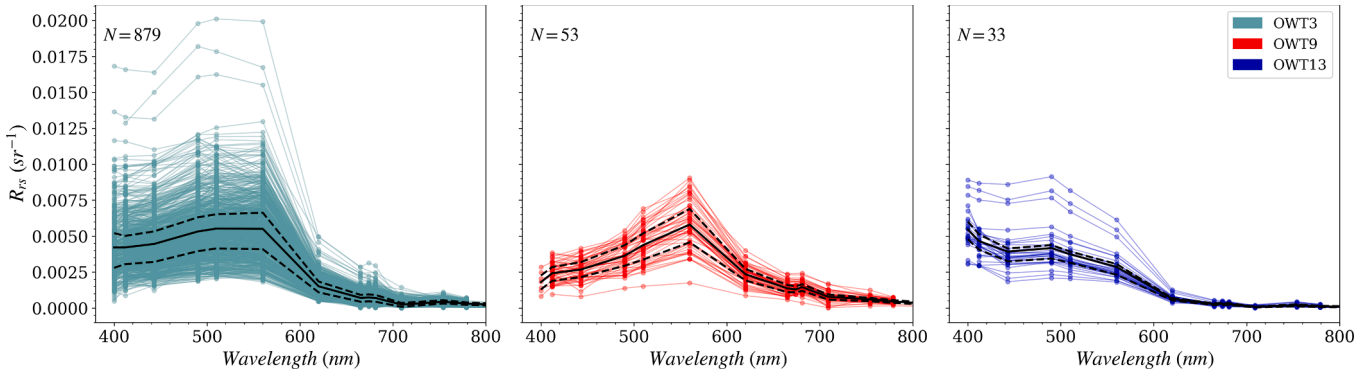

Fig. 5. CCI-lake remote sensing reflectance (R_{rs}) averaged over the region of interest (3x3 pixels centered on SHL2 location) after classification based on the thirteen optical water types defined by Spyrakos et al., 2018. The solid line is the mean reflectance for each OWT and the dashed lines represent the first and third quartiles.

3.3. Performance of CCI-lake Chl-a retrieval

First, a sensitivity analysis was applied to evaluate the most representative matchup settings for further evaluation of the performance of the Chl-*a* retrieval by the CCI processing chain. Based on several modalities of the spatial and time windows, the best comparisons were obtained for a time window of one day; for larger windows correlations between satellite and *in-situ* data were weaker, with higher *rmse(s)* (cf. Suppl. Mat, **Table S1, S2**). These observations are related to the fact that, at any point in the lake, the Chl-*a* concentrations can vary from one day to another, as well as during the day (Minaudo et al., 2021). Further, the best correlations between *in-situ* and satellite data were obtained with either the mean or median value computed over the 3x3-pixel window. While the correlations between *in-situ* and satellite data were similar for both means and medians of the values from the individual pixels, the number of pixel(s) accounted for was highly influential to the agreement between data sets. The comparisons indicated strong and significant correlations, with low *rmse*(s) and *mse*(s) for all depth intervals, but the best results were obtained by averaging Chl-*a* *in-situ* over 0–2 and 0–5 m. Finally, it is worth highlighting that more matchups are available for the spring-summer period (68 % for a 1-day matchup window). This could be due to failure of the atmospheric correction for the very high Sun zenith angles in winter. Nevertheless, this is not thought to critically affect the findings due to the low concentration and

variation of Chl-*a* over this period of the year.

Based on the best time and space windows, that is 1-day maximal time difference and 3x3 pixels region, the CCI Chl-*a* products were compared against the *in-situ Chl-a*$_{0\text{-}5m}$ averaged over the five top meters of the water column (Fig. 6a) and the *in-situ Chl-a*$_{weighted}$ weighted by the diffuse attenuation coefficient (see Eq. (6), Fig. 6b). It is readily visible from the two matchup modalities that the CCI Chl-*a* are underestimated with a slope of the regression line of 0.35 and 0.47, respectively. Nevertheless, the matchup comparisons show a robust linear relationship with the *in-situ* data with Pearson correlation coefficients of 0.80 and 0.85 for *Chl-a*$_{0\text{-}5m}$ and *Chl-a*$_{weighted}$, respectively. It is worth noting that the second modality using *Chl-a*$_{weighted}$, as reference provided better results compared to *Chl-a*$_{0\text{-}5m}$ with lower values of the statistical indicators (i.e. *rmse*, *bias*). This better agreement with the optically weighted Chl-a$_{weighted}$ demonstrates the necessity to properly integrate the vertically sampled *in situ* concentrations to obtain meaningful comparison with satellite retrievals.

Even if the integration of Chl-*a* vertical heterogeneity through the computation of *Chl-a*$_{weighted}$ improved the correlation between *in-situ* and satellite, the slope between the two data sets is still unsatisfactory. To provide more realistic values from the satellite estimations a straightforward solution is to apply a detrending routine from the matchup regression line. This operation was applied to all the satellite data by removing the offset (−0.149) and then dividing by the slope (0.47) the different Chl-*a* values. Thus, the matchup comparison shown in Fig. 6c provides the unbiased performances of the satellite retrievals with *mse* around 1.5 mg m^{-3} and *mape* of 25 %. Finally, we find that highest Chl-*a* values were retrieved for the OWT-9 (intermediate water) and the majority of the matchup dates are representative of the OWT-3 (clear waters).

3.4. Recalibrated OCx retrieval

The matchup results showed the need to recalibrate the CCI Chl-*a* for Lake Geneva. The linear adjustment previously performed is an easy and direct means to provide unbiased and scaled data from the range of values observed in the matchup comparison, that is between 1 and 12 mg m^{-3}. Nevertheless, this simple recalibration might provide unexpected results for values beyond this range. In order to provide long-term satellite data, a proper recalibration of the OCx relationships was performed based on the R_{rs} retrievals using the OC2, OC3, OC4 and OC5 formulations, see Fig. S6 and Table S3 in the Supplementary Material. The matchup results based on these recalibrations are shown in Fig. 7. Interestingly, the Pearson correlation coefficient was significantly improved in comparison to the previous matchup comparison whatever the OCx formulation considered with values ranging from 0.891 (OC2 and OC3) to 0.899 (OC4). Moreover, the *rmse*, *mape* and *bias*, as well as β and ε, are significantly reduced in comparison to the CCI Chl-*a* comparison. The best performances were achieved for the OC4 formulation with a regression slope of 0.98 and *rmse* of around 1.05 mg m^{-3} and a mean absolute percentage error of 21 %. Based on these results, the entire satellite time series was processed using the recalibrated OC4 formulation to provide the final satellite Chl-*a* values.

3.5. Long term change

The *in-situ* and satellite values (Fig. 8) were used to investigate long-term trends in phytoplankton biomass at the SHL2 site. The merged dataset was obtained after averaging the *Chla*$_{weighted}$ and recalibrated *Chla*$_{ESACCI}$. Addition of the satellite data to the *in-situ* data set greatly enhanced the time series, showing a greater variability in Chl-*a* concentrations with Chl-*a* concentrations ranging from 0-30 mg m^{-3}.

Dates of phenological events identified by satellite and also using the *in-situ Chla*$_{weighted}$ data were significantly correlated for SOS, DOM (p-value < 0.05) and COG (p-value < 0.1). No exact matches between estimates from *in-situ* and satellite data were observed. Half of the SOS indicator values (50 %) occurred earlier based on satellite measurements than when based on *in-situ* data, and the temporal gaps between dates for *in-situ* and satellite data did not exceed 6 days in 67 % of the observations, indicating a relatively good match between the two datasets (Fig. 9). Temporal differences in phenological indicators derived from satellite and *in-situ* data exceeded 7 days more often, for both the DOM (67 %) and COG (58 %). Large inconsistencies between phenological dates estimated from satellite and *in-situ* data (more than 3 weeks apart) occurred only once for both DOM and SOS. No large inconsistencies were recorded for COG. Finally, over the studied period (2003–2019), neither the satellite nor *in-situ* data indicated significant (p-value < 0.05) trends in the seasonal timing of phenological indicators.

Chla$_{Merged}$ allowed analyses for the 2013 to 2016 period, when satellite data were not available. Therefore, the merged dataset contained 19 consecutive years that could be compared to the *in-situ* dataset from 2003 to 2021. Temporal changes in the dates of seasonal events retrieved from *Chla*$_{Merged}$ and *Chla*$_{0\text{-}5m}$ data were strongly correlated (Fig. 9, p-value < 0.01).

Furthermore, there were good matches between dates of seasonal events (Fig. 9), especially for DOM. For this seasonal event we recorded 12 exact matches (67 %) and only one observation presented a difference higher than 21 days. SOS also matched relatively well between merged and *in-situ*; 56 % of pairings were exact matches and only 1 date

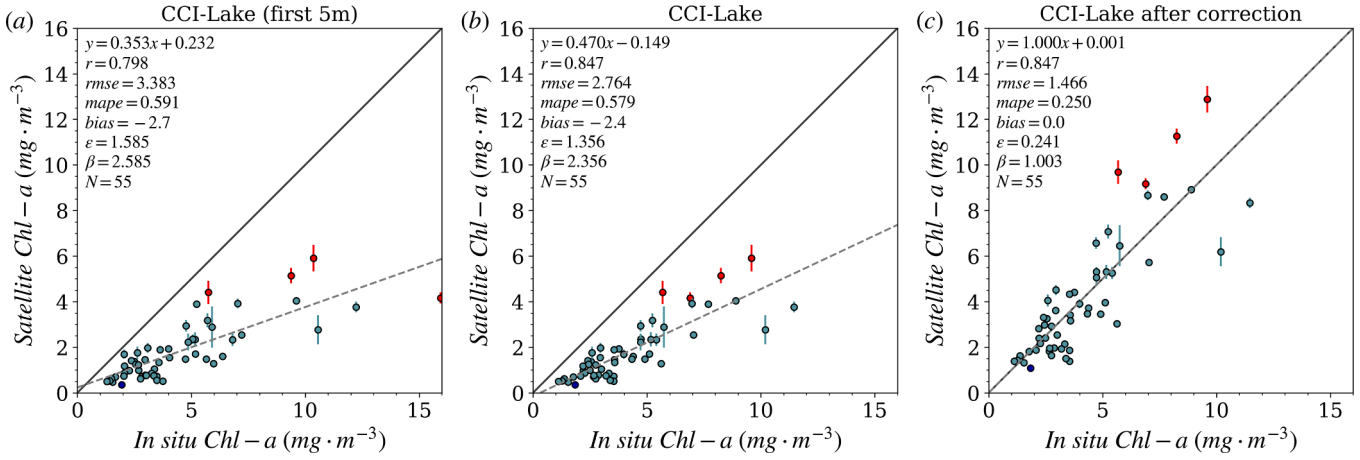

Fig. 6. Matchup comparison between (a) CCI-lake Chl-*a* product versus *in situ* Chl-*a* averaged over the 0 to 5 m top layer, (b) similar but using the optically weighted *in situ* Chl-*a*, and (c) CCI-lake Chl-*a* after recalibration based on the regression curve found in (b). The colors represent the OWT retrieved of each matchup date (grey: OWT-3, red: OWT-9, blue: OWT-13).

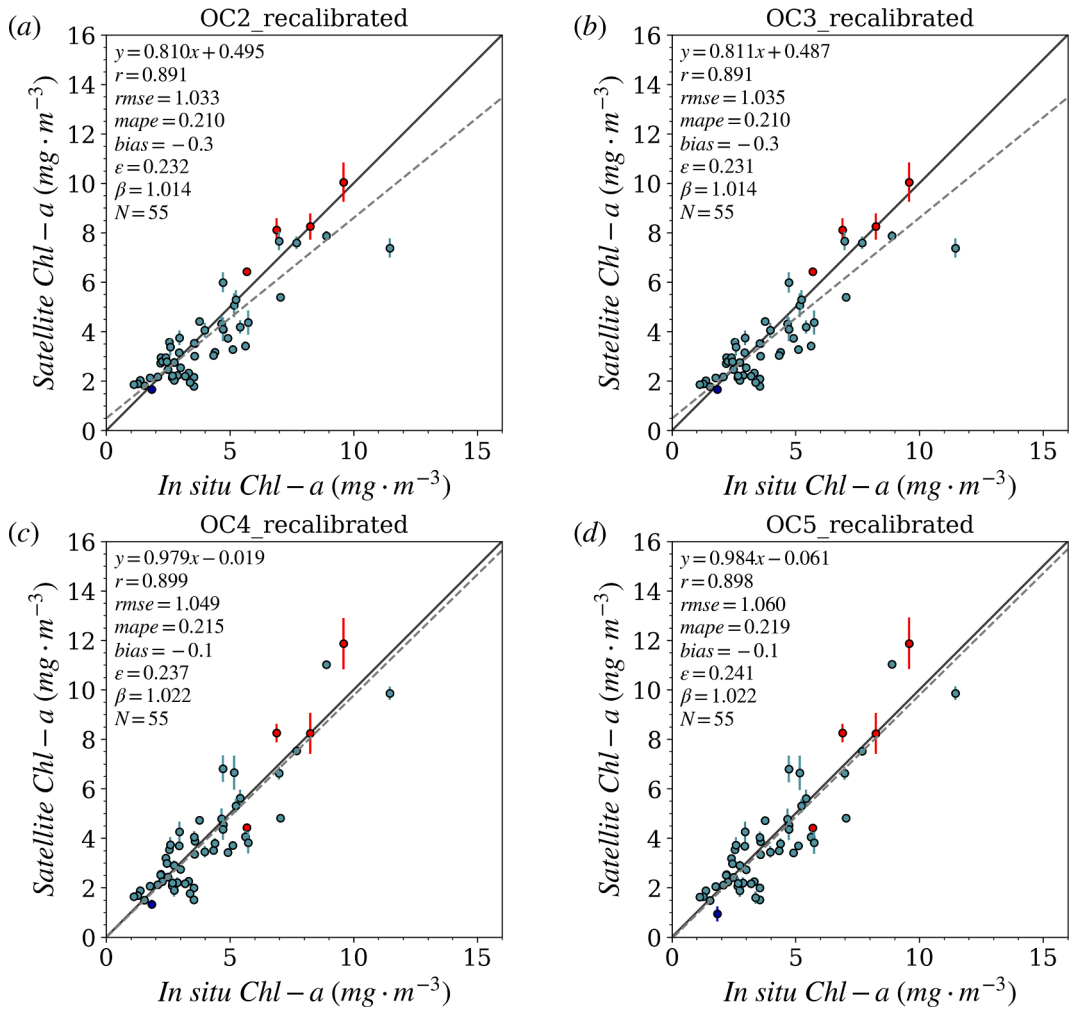

Fig. 7. Matchup comparisons between *in situ* data and the satellite product computed after recalibration of the OCx relationships. The colors indicate the OWT retrieved for each date, see Fig. 5.

had a difference higher than 21 days. An exact match was recorded 3 times for COG and no date with more than 21-day difference was recorded. Finally, phenological events computed with $Chla_{Merged}$ did not show any significant (p-value < 0.05) trend over the studied period (2003–2021).

4. Discussion

Representativeness is defined as how well a sample corresponds to the population from which it is drawn. Here, we interpret the representativeness of satellite data as how well satellite-derived estimates of absolute Chl-*a* concentrations, and temporal dynamics in Chl-*a*, match with those determined from depth-resolved *in-situ* data. Specifically, we used *in-situ* data to estimate the impact of the vertical distribution of phytoplankton on trends in Chl-*a* observed by satellite Earth Observation. The performance of the ESA-CCI was assessed based on both regional satellite algorithms and following match-up comparison with the available *in-situ* data.

The vertical distribution of the phytoplankton may induce significant changes in the spectral water-leaving radiance and in turn might modify the estimation of the Chl-*a* concentration where homogeneity of the water column is often assumed (Nouchi et al., 2018; Xue et al., 2017). In this study, we tested a practical approach to account for this potential impact on Chl-*a* estimation and check if it may induce spurious trends when not corrected for. The rationale of the method (see section 2.2 and 3.1) was to exploit the *in-situ* data available within the long-term OLA dataset including vertically resolved Chl-*a* and PAR measurements. The latter was used to assess water transparency through the computation of the diffuse attenuation coefficient. The various matchup modalities we used revealed that the satellite signal was mainly driven by the optically active material of the first optical depth layer (i.e., $1/K_d$) ranging from 2.5 to 8.2 m deep in the case of Lake Geneva. This explains the good quality of the matches obtained between satellite data and *in-situ* data from 2, 5 and 8 m deep. Despite this, satellite data showed the strongest agreement with the $Chla_{weighted}$ data. This leads us to strongly advocate for highly resolved profiling within the first optical depth layer. Nonetheless, the most appropriate devices to do so rely on fluorometric acquisition with the known impact of NPQ within this layer. As a result, both fine scale vertical measurements and accurate NPQ correction are essential to properly compare *in-situ* Chl-*a* with their satellite estimate counterpart.

The accuracy of satellite estimates of Chl-*a* in the lake strongly depends upon the depth of the Chl-*a* maximum relative to the water layers observable by the satellite. Here, we found that only about 15.8 % of the total (depth-integrated) Chl-*a* was likely to be observable by remote sensing if we assume a visibility limit based upon $1/K_d$. This suggests that vertical heterogeneity in phytoplankton populations in deep lakes could introduce important uncertainties into satellite-derived estimates of primary production.

This limitation might be more frequent in deep lakes where deep Chl-*a* are common (Leach et al., 2018) or where metalimnetic species like the toxic cyanobacterium *P. rubescens* can be abundant (Soulignac et al.,

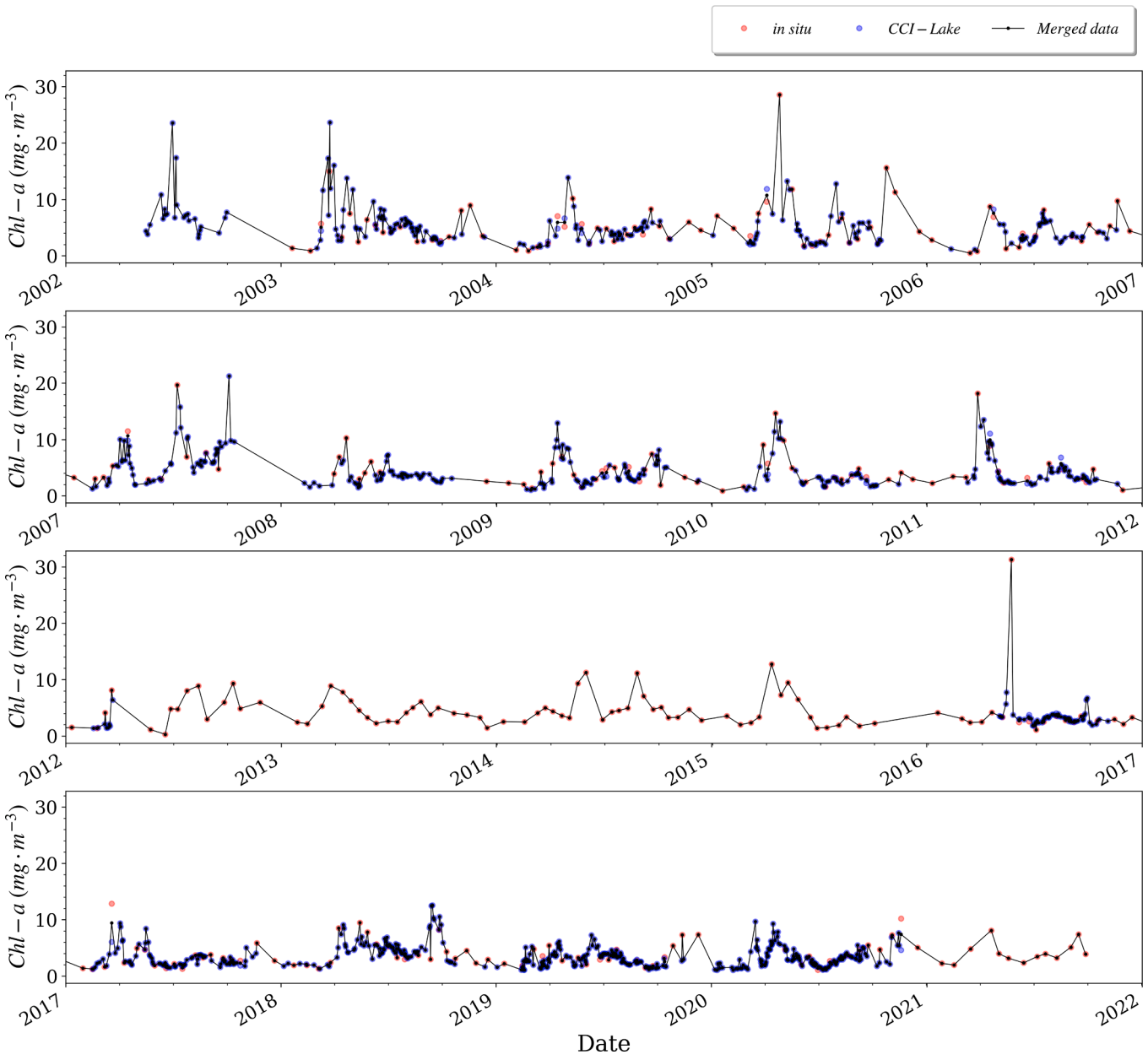

Fig. 8. Time series of the Chl-*a* concentration obtained from recalibrated ESA CCI (blue) and *in situ* measurements (red), the continuous line represents the merged data set after averaging values from CCI and *in situ* measurements.

2018). For Lake Geneva, the long-term deepening of phytoplankton maxima in the water column (Anneville et al., 2013) poses a considerable challenge to the efficiency of satellite-based remote sensing. Such results suggest that a better understanding of the link between deep Chl-*a* maxima and trophic state, and the extent to which such lake ecosystem behaviors can be observed from space, is needed (Anneville et al., 2002; Scofield et al., 2020).

Further research is necessary to understand the conditions under which deep Chl-*a* blooms occur and whether they are associated with specific trajectories of ecological succession. In Lake Geneva, it is unlikely to observe a deep Chl-*a* event in winter and early-spring when the water column is weakly stratified, making satellite-observed blooms more reliable. In contrast, summer is more conducive to the development of deeper Chl-*a* maxima (Anneville and Leboulanger, 2001) when the phytoplankton community is dominated by low light adapted species like *Mougeotia gracillima* and *Planktothrix rubescens*, which develop near the thermocline (Jacquet et al., 2005, 2014; Tapolczai et al., 2015). In these cases, relying solely on satellite data for bloom detection carries a high degree of uncertainty. Therefore, both satellite and *in situ* monitoring are essential for accurate bloom detection.

The use of remote sensing as a tool to better understand the dynamics of continental water quality has developed significantly over the last 50 years. However, it is important to recognize the benefits of merging the resulting data with *in-situ* field data, provided that appropriate pre-processing and quality control is conducted, in order to meet the expectations of political decision-makers and water managers (Gholizadeh et al., 2016; Topp et al., 2020). Inter-dataset comparisons, and efforts towards *in-situ* and satellite data integration, are highly dependent on the range of the observed datasets. Here, the merged dataset obtained by aggregating (averaging) *in-situ* data ($Chla_{weighted}$) and calibrated ESA CCI data, allowed the analysis of dynamics at a finer temporal resolution. This dataset is thus particularly appropriate for the analysis of phytoplankton phenology. Phytoplankton phenology is sensitive to changes in meteorological conditions and can thus be a good indicator of the impact

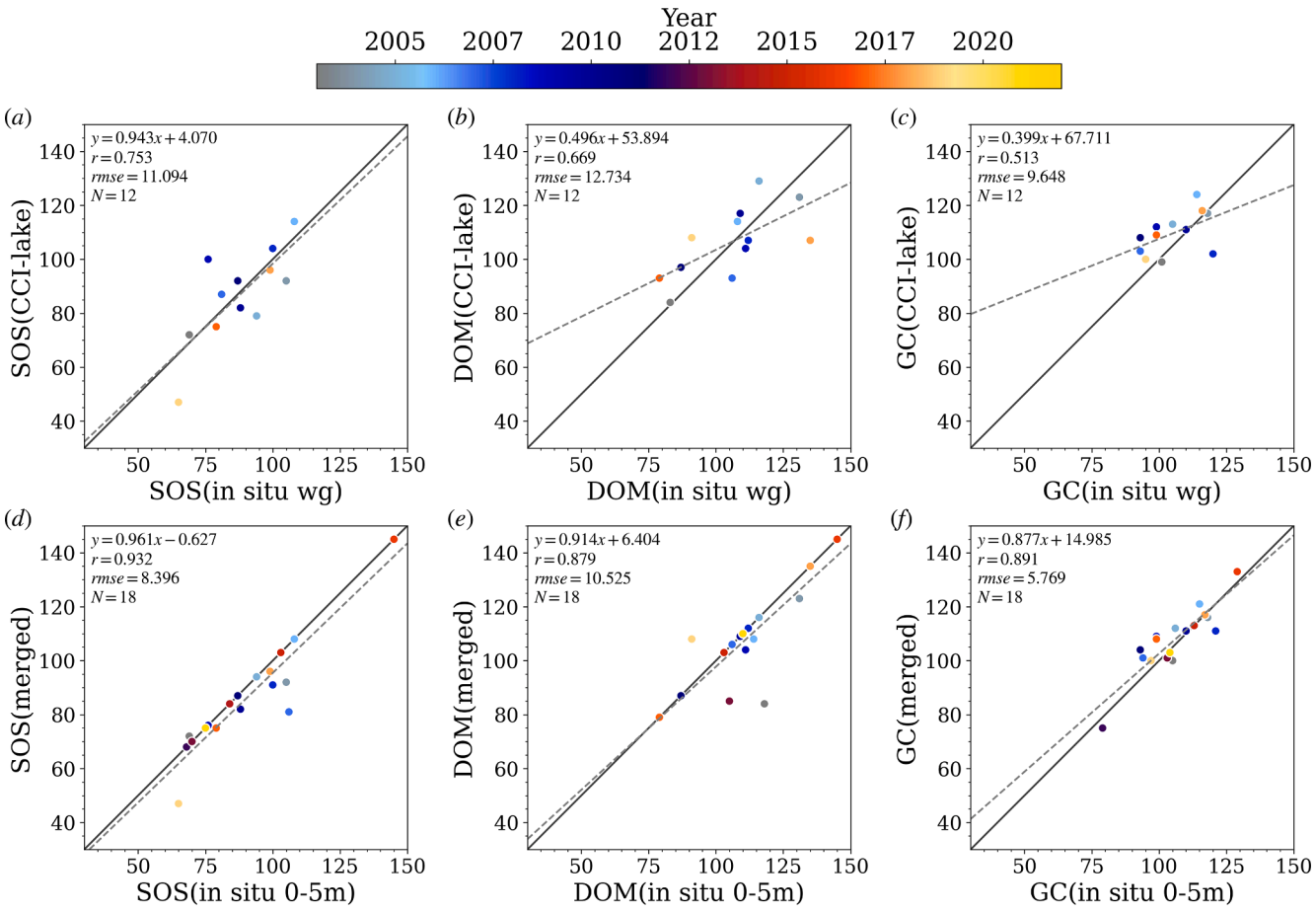

Fig. 9. Comparisons of inter-annual variations in phenological indicators among datasets.

of climate change. Warmer springs induce earlier disappearance of the spring phytoplankton community due to earlier zooplankton grazing (Straile et al., 2003). Similar change in spring phenology was also observed in Lake Geneva, where warmer springs also indirectly induced an earlier development of summer assemblages (Anneville et al., 2002). In this paper, we tested the merged dataset on the spring phenology. Unlike phytoplankton summer assemblages, which are largely comprised of low-light adapted species thriving deeper in the water column where nutrients persist, spring phytoplankton communities are typically dominated by species that require higher light availability flourishing in the upper part of the euphotic layer (Anneville et al., 2002). This phytoplankton biomass is thus more likely to be detected by satellite Earth Observation. In Lake Geneva, the start of spring phytoplankton growth is driven by the onset of thermal stratification (Sommer et al., 2012; Soulignac et al., 2018). With climate warming, we expect an earlier onset of thermal stratification and as a consequence an earlier phytoplankton growth. However, our results do not support this hypothesis for the considered time period (2003–2019). They align with those of other studies that focused on phytoplankton biomass and did not report marked changes after 2000 (Anneville et al., 2019). Our analyses were conducted over a relatively short time period (19 years), which is not sufficient to encompass changes in relation to global warming that occur over a temporal scale larger than 30 years. Furthermore, the considered period coincides with the global warming hiatus which resulted in lower warming rates of lake surface temperature (Winslow et al., 2018). Nevertheless, it is interesting to notice the strong inter-annual variations in phenology, with some years exhibiting the development of phytoplankton at an early date in the season. Further analysis is required to evaluate whether these fluctuations in the phenology are due to inter-annual fluctuations in weather conditions and/or thermal stratification. If this is the case, our findings suggest that climate-induced earlier stratification would likely result in earlier development of phytoplankton in the future (Thackeray et al., 2008).

Finally, despite the enhanced temporal resolution afforded by the integration of *in-situ* and satellite data, the OC model series algorithms (O'Reilly et al., 2000) do not estimate Chl-*a* concentrations effectively in optically complex waters above 10 mg.m^{-3} (Ogashawara et al., 2017). As a result, the match-up(s) and merged dataset derived here cannot be extrapolated to concentrations above 10 mg.m^{-3}.

There is no doubt that satellite observations offer new perspectives and scales of investigation that will enhance our knowledge of patterns and drivers of phytoplankton abundance and productivity. Indeed, empirical models based on R_{rs} can enhance knowledge derived from *in-situ* measurements e.g., through combination, as illustrated here (*Chla-Merged*). On the other hand, the use of satellite data for management and risk assessment purposes still requires *in-situ* measurements to advantageously complement remote-sensing-based information to provide integrative and synoptic observations of large lakes.

The potential for high frequency data acquisition means that satellites could identify short term events, such as dates of maximum concentrations, more effectively than manual *in-situ* low frequency sampling. Our observed differences between dates of phenological indicators retrieved from satellite and *in-situ* sampling are suggestive of these potential advantages of satellite monitoring.

The combination of satellite and *in-situ* data yields information at finer temporal resolution and, in our case, allows the inclusion of years for which satellite data were not available. Phenological indicators estimated using the merged data yielded long-term trends similar to those derived from indicators based upon *in-situ* data, and dates of events were well correlated with significant match-up(s). Combining the

two sources allowed a consistent increase in the number of observations and thus the robustness of the results, together with an enhanced capability to observe the occurrence of short-term events.

5. Conclusion

Understanding of the state and dynamics of complex systems, such as large *peri*-alpine lakes, requires complementary approaches, such as satellite and *in-situ* monitoring, to provide a synoptic and/or punctual vision of the ecosystem and its temporal evolution. Satellite data alone cannot be used to observe the vertical distribution of phytoplankton biomass and fails to properly quantify the totality of this biomass. Satellite information on phytoplankton biomass is limited to surface water layers and is affected by water clarity. Based upon *in-situ* data, the proportion of the water column integrated Chl-*a* observed by Sentinel-3-like satellites rarely exceeds 30 %. At present, it is not clear whether satellite-observable surface Chl-*a* concentrations can provide useful indirect information on likely sub-surface growth, and this would be a valuable area of future study.

With the ongoing re-oligotrophication of Lake Geneva, accompanied by a deepening of phytoplankton niches, satellite observation will face further challenges in observing phytoplankton biomass representatively. *In-situ* data on Chl-*a* concentrations often have limited temporal resolution and spatial coverage. As a result, the integration of satellite observations and *in-situ* measurements in the water column has the potential to accurately describe an important ecological state in these stratified ecosystems. The combination of these two sources of information provides a finer temporal resolution of Chl-*a* concentrations enabling more accurate analysis of trends and phenology. Finally, this study demonstrated the representativeness and robustness of the CCI satellite data base in the central region of the lake opening potential ways to decipher spatial–temporal patterns of the phenology over the full extent of Lake Geneva.

CRediT authorship contribution statement

Mona Bonnier: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Orlane Anneville:** Conceptualization, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing, Data curation, Funding acquisition. **R. Iestyn Woolway:** Formal analysis, Investigation, Methodology, Resources, Writing – review & editing. **Stephen J. Thackeray:** Methodology, Writing – review & editing. **Guillaume P. Morin:** Investigation, Methodology, Writing – review & editing. **Nathalie Reynaud:** Resources. **Frédéric Soulignac:** Project administration, Resources. **Thierry Tormos:** Resources. **Tristan Harmel:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Authors thank CIPEL and pole Ecla (OFB- INRAE) for funding. In-situ data for Lake Geneva are contributed by the Observatory of alpine LAkes (OLA), © SOERE OLA-IS, AnaEE-France, INRAE at Thonon-les-Bains, CIPEL. We would like to thank the operational technicians of the OLA PRECCI platform at INRAE Thonon-les-Bains for their work, their involvement, and their support. We would also like to thank pole ECLA (INRAE-OFB) for its financial support and for the active participation of its members in this study. Finally, we would like to acknowledge the financial support provided by CNES through the TOSCA-OBS2MOD project. RIW was supported by the UKRI Natural Environment Research Council (NERC): Independent Research Fellowship [grant number NE/T011246/1]. OA, RIW and SJT acknowledge support of the AFTER-GEISHA group funded by the INRAE and the synthesis center - CESAB of the French Foundation for Research on Biodiversity (FRB; www.fondationbiodiversite.fr). The authors would like to thank the three reviewers for their constructive feedback and Irani Rahaghi Abolfazl and Daniel Odermatt for their help to improve this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jglr.2024.102372.

References

Adrian, R., O'Reilly, C.M., Zagarese, H., Baines, S.B., Hessen, D.O., Keller, W., Livingstone, D.M., Sommaruga, R., Straile, D., Van Donk, E., 2009. Lakes as sentinels of climate change. Limnol. Oceanogr. 54 (6part2), 2283–2297.

Anneville, O.; Leboulanger, C. 2001. Long-term changes in the vertical distribution of phytoplankton in the deep Alpine Lake Geneva: a response to the reoligotrophication. vol. 14. Atti Associazione italiana di Oceanologia e Limnologia: 25-35.

Anneville, O., Souissi, S., Ibanez, F., Ginot, V., Druart, J.C., Angeli, N., 2002. Temporal mapping of phytoplankton assemblages in Lake Geneva: annual and interannual changes in their patterns of succession. Limnol. Oceanogr. 47 (5), 1355–1366.

Anneville, O., Beniston, M., Gallina, N., Gillet, C., Jacquet, S., Lazzarotto, J., 2013. L'empreinte du changement climatique sur le Léman. Arch. Sci. 16.

Anneville, O., Chang, C.-W., Dur, G., Souissi, S., Rimet, F., Hsieh, C., 2019. The paradox of re-oligotrophication: the role of bottom–up versus top–down controls on the phytoplankton community. Oikos 128 (11), 1666–1677.

Ansper, A., Alikas, K., 2018. Retrieval of chlorophyll a from Sentinel-2 MSI data for the European Union water framework directive reporting purposes. Remote Sens. (Basel) 11 (1), 64.

Carrea, L., Crétaux, J.-F., Liu, X., Wu, Y., Calmettes, B., Duguay, C.R., Merchant, C.J., Selmes, N., Simis, S.G., Warren, M., 2023. Satellite-derived multivariate world-wide lake physical variable timeseries for climate studies. Sci. Data 10 (1), 30.

Chang, C.-W., Miki, T., Ye, H., Souissi, S., Adrian, R., Anneville, O., Agasild, H., Ban, S., Be'eri-Shlevin, Y., Chiang, Y.-R., 2022. Causal networks of phytoplankton diversity and biomass are modulated by environmental context. Nat. Commun. 13 (1), 1–11.

Feng, L., Dai, Y., Hou, X., Xu, Y., Liu, J., Zheng, C., 2021. Concerns about phytoplankton bloom trends in global lakes. Nature 590 (7846), E35–E47.

Gernez, P., Antoine, D., 2009. Field characterization of wave-induced underwater light field fluctuations. J. Geophys. Res. 114.

Gholizadeh, M.H., Melesse, A.M., Reddi, L., 2016. A comprehensive review on water quality parameters estimation using remote sensing techniques. Sensors 16 (8), 1298.

Gilerson, A.A., Gitelson, A.A., Zhou, J., Gurlin, D., Moses, W., Ioannou, I., Ahmed, S.A., 2010. Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. Opt. Express 18 (23), 24109.

Gons, H.J., Rijkeboer, M., Ruddick, K.G., 2005. Effect of a waveband shift on chlorophyll retrieval from MERIS imagery of inland and coastal waters. J. Plankton Res. 27 (1), 125–127.

Gordon, H.R., Clark, D.K., 1980. Remote sensing optical properties of a stratified ocean: an improved interpretation. Appl. Opt. 19 (20), 3428–3430.

Gupana, R.S., Damm, A., Rahaghi, A.I., Minaudo, C., Odermatt, D., 2022. Non-photochemical quenching estimates from in situ spectroradiometer measurements: implications on remote sensing of sun-induced chlorophyll fluorescence in lakes. Opt. Express 30 (26), 46762–46781.

Hlaing, S., Harmel, T., Gilerson, A., Foster, R., Weidemann, A., Arnone, R., Wang, M., Ahmed, S., 2013. Evaluation of the VIIRS ocean color monitoring performance in coastal regions. Remote Sens. Environ. 139, 398–414.

Ho, J.C., Michalak, A.M., Pahlevan, N., 2019. Widespread global increase in intense lake phytoplankton blooms since the 1980s. Nature 574 (7780), 667–670.

Hu, M., Zhang, Y., Ma, R., Xue, K., Cao, Z., Chu, Q., Jing, Y., 2021. Optimized remote sensing estimation of the lake algal biomass by considering the vertically heterogeneous chlorophyll distribution: Study case in Lake Chaohu of China. Sci. Total Environ. 771, 144811.

Huot, Y., Babin, M., 2010. Overview of fluorescence protocols: theory, basic concepts, and practice. In: Chlorophyll a Fluorescence in Aquatic Sciences: Methods and Applications. Springer, pp. 31–74.

Jacquet, S., Briand, J.-F., Leboulanger, C., Avois-Jacquet, C., Oberhaus, L., Tassin, B., Vinçon-Leite, B., Paolini, G., Druart, J.-C., Anneville, O., Humbert, J.-F., 2005. The proliferation of the toxic cyanobacterium Planktothrix rubescens following restoration of the largest natural French lake (Lac du Bourget). Harmful Algae 4 (4), 651–672.

Jacquet, S., Domaizon, I., Anneville, O., 2014. The need for ecological monitoring of freshwaters in a changing world: a case study of Lakes Annecy, Bourget, and Geneva. Environ. Monit. Assess. 186 (6), 3455–3476.
Jenny, J.-P., Anneville, O., Arnaud, F., Baulaz, Y., Bouffard, D., Domaizon, I., Bocaniov, S.A., Chèvre, N., Dittrich, M., Dorioz, J.-M., Dunlop, E.S., Dur, G., Guillard, J., Guinaldo, T., Jacquet, S., Jamoneau, A., Jawed, Z., Jeppesen, E., Krantzberg, G., Lenters, J., Leoni, B., Meybeck, M., Nava, V., Nõges, T., Nõges, P., Patelli, M., Pebbles, V., Perga, M.-E., Rasconi, S., Ruetz, C.R., Rudstam, L., Salmaso, N., Sapna, S., Straile, D., Tammeorg, O., Twiss, M.R., Uzarski, D.G., Ventelä, A.-M., Vincent, W.F., Wilhelm, S.W., Wängberg, S.-Å., Weyhenmeyer, G.A., 2020. Scientists' Warning to Humanity: Rapid degradation of the world's large lakes. J. Great Lakes Res. 46 (4), 686–702.
Kao, Y.-C., Rogers, M.W., Bunnell, D.B., Cowx, I.G., Qian, S.S., Anneville, O., Beard Jr, T. D., Brinker, A., Britton, J.R., Chura-Cruz, R., 2020. Effects of climate and land-use changes on fish catches across lakes at a global scale. Nat. Commun. 11 (1), 2526.
Kavanaugh, M.T., Bell, T., Catlett, D., Cimino, M.A., Doney, S.C., Klajbor, W., Messié, M., Montes, E., Muller-Karger, F.E., Otis, D., 2021. Satellite Remote Sensing and the Marine Biodiversity Observation Network. Oceanography 34 (2), 62–79.
Kraemer, B.M., Mehner, T., Adrian, R., 2017. Reconciling the opposing effects of warming on phytoplankton biomass in 188 large lakes. Sci. Rep. 7 (1), 10762.
Leach, T.H., Beisner, B.E., Carey, C.C., Pernica, P., Rose, K.C., Huot, Y., Brentrup, J.A., Domaizon, I., Grossart, H.-P., Ibelings, B.W., 2018. Patterns and drivers of deep chlorophyll maxima structure in 100 lakes: The relative importance of light and thermal stratification. Limnol. Oceanogr. 63 (2), 628–646.
Maberly, S.C., O'Donnell, R.A., Woolway, R.I., Cutler, M.E., Gong, M., Jones, I.D., Merchant, C.J., Miller, C.A., Politi, E., Scott, E.M., 2020. Global lake thermal regions shift under climate change. Nat. Commun. 11 (1), 1232.
Minaudo, C., Odermatt, D., Bouffard, D., Rahaghi, A.I., Lavanchy, S., Wüest, A., 2021. The imprint of primary production on high-frequency profiles of lake optical properties. Environ. Sci. Tech. 55 (20), 14234–14244.
Mueller, J.L.; Morel, A.; Frouin, R.; Davis, C.; Arnone, R.; Carder, K.; Lee, Z.P.; Steward, R.G.; Hooker, S.; Mobley, C.D.; McLean, S.; Holben, B.; Miller, M.; Pietras, C.; Knobelspiesse, K.D.; Fargion, G.S.; Porter, J.; Voss, K. 2003. Ocean Optics Protocols For Satellite Ocean Color Sensor Validation, Revision 4. Volume III: Radiometric Measurements and Data Analysis Protocols.
Muller, P., Li, X.-P., Niyogi, K.K., 2001. Non-photochemical quenching. A response to excess light energy. Plant Physiol. 125 (4), 1558–1566.
Neil, C., Spyrakos, E., Hunter, P.D., Tyler, A.N., 2019. A global approach for chlorophyll-a retrieval across optically complex inland waters based on optical water types. Remote Sens. Environ. 229, 159–178.
Nouchi, V., Odermatt, D., Wüest, A., Bouffard, D., 2018. Effects of non-uniform vertical constituent profiles on remote sensing reflectance of oligo-to mesotrophic lakes. Europ. J. Rem. Sens. 51 (1), 808–821.
O'Reilly, J.E., Maritorena, S., Mitchell, B.G., Siegel, D.A., Carder, K.L., Garver, S.A., Kahru, M., McClain, C., 1998. Ocean color chlorophyll algorithms for SeaWiFS. J. Geophys. Res. Oceans 103 (C11), 24937–24953.
O'Reilly, J.E., Maritorena, S., Siegel, D.A., O'Brien, M.C., Toole, D., Mitchell, B.G., Kahru, M., Chavez, F.P., Strutton, P., Cota, G.F., 2000. Ocean color chlorophyll a algorithms for SeaWiFS, OC2, and OC4: Version 4: SeaWiFS Postlaunch. Calibrat. Validati. Anal. Part 3, 9–23.
O'Reilly, C.M., Sharma, S., Gray, D.K., Hampton, S.E., Read, J.S., Rowley, R.J., Schneider, P., Lenters, J.D., McIntyre, P.B., Kraemer, B.M., 2015. Rapid and highly variable warming of lake surface waters around the globe. Geophys. Res. Lett. 42 (24), 10–773.
Ogashawara, I., Mishra, D.R., Gitelson, A.A., 2017. Remote sensing of inland waters: background and current state-of-the-art. In: Bio-Optical Modeling and Remote Sensing of Inland Waters. Elsevier, pp. 1–24.
Pahlevan, N., Mangin, A., Balasubramanian, S.V., Smith, B., Alikas, K., Arai, K., Barbosa, C., Bélanger, S., Binding, C., Bresciani, M., 2021. ACIX-Aqua: A global assessment of atmospheric correction methods for Landsat-8 and Sentinel-2 over lakes, rivers, and coastal waters. Remote Sens. Environ. 258, 112366.
Park, T., Ganguly, S., Tømmervik, H., Euskirchen, E.S., Høgda, K.-A., Karlsen, S.R., Brovkin, V., Nemani, R.R., Myneni, R.B., 2016. Changes in growing season duration and productivity of northern vegetation inferred from long-term remote sensing data. Environ. Res. Lett. 11 (8), 084001.
Piskozub, J., Neumann, T., Woźniak, L., 2008. Ocean color remote sensing: choosing the correct depth weighting function. Opt. Express 16 (19), 14683–14688.
Pitarch, J., Odermatt, D., Kawka, M., Wüest, A., 2014. Retrieval of vertical particle concentration profiles by optical remote sensing: a model study. Opt. Express 22 (103), A947–A959.
Quétin, P.; Hustache, J.-C.; Perney, P.; Gérard, P. 2020. Analyse mathématique pour la comparaison des mesures de transparence de l'eau avec deux différents disques de Secchi. Le Cahier des Techniques de L'Inra 99.
Rimet, F., Anneville, O., Barbet, D., Chardon, C., Crepin, L., Domaizon, I., Dorioz, J.-M., Espinat, L., Frossard, V., Guillard, J., Goulon, C., Hamelet, V., Hustache, J.-C., Jacquet, S., Lainé, L., Montuelle, B., Perney, P., Quétin, P., Rasconi, S., Schellenberger, A., Tran-Khac, V., Monet, G., 2020. The Observatory on LAkes (OLA) database: Sixty years of environmental data accessible to the public. J. Limnol. 79 (2), 164.
Roy, S., Das, B.S., 2022. Estimation of Euphotic Zone Depth in Shallow Inland Water using Inherent Optical Properties and Multispectral Remote Sensing Imagery. J. Hydrol. 128293.
Scofield, A.E., Watkins, J.M., Osantowski, E., Rudstam, L.G., 2020. Deep chlorophyll maxima across a trophic state gradient: A case study in the Laurentian Great Lakes. Limnol. Oceanogr. 65 (10), 2460–2484.
Seegers, B.N., Stumpf, R.P., Schaeffer, B.A., Loftin, K.A., Werdell, P.J., 2018. Performance metrics for the assessment of satellite data products: an ocean color case study. Opt. Express 26 (6), 7404–7422.
Sen, P.K., 1968. Estimates of the Regression Coefficient Based on Kendall's Tau. J. Am. Stat. Assoc. 63 (324), 1379–1389.
Sommer, U., Adrian, R., De Senerpont Domis, L., Elser, J.J., Gaedke, U., Ibelings, B., Jeppesen, E., Lürling, M., Molinero, J.C., Mooij, W.M., Van Donk, E., Winder, M., 2012. Beyond the Plankton Ecology Group (PEG) Model: Mechanisms Driving Plankton Succession. Annu. Rev. Ecol. Evol. Syst. 43 (1), 429–448.
Soulignac, F., Danis, P.-A., Bouffard, D., Chanudet, V., Dambrine, E., Guénand, Y., Harmel, T., Ibelings, B.W., Trevisan, D., Uittenbogaard, R., 2018. Using 3D modeling and remote sensing capabilities for a better understanding of spatio-temporal heterogeneities of phytoplankton abundance in large lakes. J. Great Lakes Res. 44 (4), 756–764.
Steinmetz, F., Deschamps, P.-Y., Ramon, D., 2011. Atmospheric correction in presence of sun glint: application to MERIS. Opt. Express 19 (10), 9783–9800.
Straile, D.; Livingstone, D.M.; Weyhenmeyer, G.A.; George, D.G. 2003. The response of freshwater ecosystems to climate variability associated with the North Atlantic Oscillation.
Tapolczai, K., Anneville, O., Padisák, J., Salmaso, N., Morabito, G., Zohary, T., Tadonléké, R.D., Rimet, F., 2015. Occurrence and mass development of Mougeotia spp. (Zygnemataceae) in large, deep lakes. Hydrobiologia 745 (1), 17–29.
Thackeray, S.J., Jones, I.D., Maberly, S.C., 2008. Long-term change in the phenology of spring phytoplankton: species-specific responses to nutrient enrichment and climatic change. J. Ecol. 96 (3), 523–535.
Thackeray, S.J., Henrys, P.A., Jones, I.D., Feuchtmayr, H., 2012. Eight decades of phenological change for a freshwater cladoceran: what are the consequences of our definition of seasonal timing? Freshw. Biol. 57 (2), 345–359.
Thackeray, S.J., Henrys, P.A., Feuchtmayr, H., Jones, I.D., Maberly, S.C., Winfield, I.J., 2013. Food web de-synchronization in E ngland's largest lake: an assessment based on multiple phenological metrics. Glob. Chang. Biol. 19 (12), 3568–3580.
Topp, S.N., Pavelsky, T.M., Jensen, D., Simard, M., Ross, M.R., 2020. Research trends in the use of remote sensing for inland water quality science: Moving towards multidisciplinary applications. Water 12 (1), 169.
Tormos, T., Reynaud, N., Danis, P.-A., Harmel, T., Morin, G., Martinez, J.M., Andral, A., Coque, A., Peroux, T., Baudoin, J.-M., 2021. Quand la surveillance des plans d'eau prendra de la hauteur. Sci. Eaux Territoir. (37), 114–123.
Winder, M., Schindler, D.E., 2004. Climate change uncouples trophic interactions in an aquatic ecosystem. Ecology 85 (8), 2100–2106.
Winslow, L.A., Leach, T.H., Rose, K.C., 2018. Global lake response to the recent warming hiatus. Environ. Res. Lett. 13 (5), 054005.
Xue, K., Zhang, Y., Ma, R., Duan, H., 2017. An approach to correct the effects of phytoplankton vertical nonuniform distribution on remote sensing reflectance of cyanobacterial bloom waters: Phytoplankton vertical nonuniform distribution. Limnol. Oceanogr. Methods 15.
Zaneveld, J.R.V., Barnard, A.H., Boss, E., 2005. Theoretical derivation of the depth average of remotely sensed optical parameters. Opt. Express 13 (22), 9052–9061.