Remote Sensing of Environment 287 (2023) 113464

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Spectral features of ocean colour radiometric products in the presence of cyanobacteria blooms in the Baltic Sea

Ilaria Cazzaniga *, Giuseppe Zibordi, Frédéric Mélin

European Commission, Joint Research Centre, Ispra, Italy

ARTICLE INFO

Edited by Dr. Menghua Wang

Keywords:
Ocean colour
Remote sensing reflectance
Harmful algal blooms
Cyanobacteria
Baltic Sea

ABSTRACT

Cyanobacteria blooms are recurrent in the Baltic Sea with frequency and intensity increasing with temperature. By relying on autonomous multispectral measurements from the Ocean Color component of the Aerosol Robotic Network (AERONET-OC), this study exploited an unprecedented dataset of in situ remote sensing reflectance $R_{RS}(\lambda)$ spectra (with wavelength λ in the interval 400–667 nm) acquired during filamentous cyanobacteria blooms in the Baltic Sea. The study investigated the temporal evolution of the in situ $R_{RS}(\lambda)$ during these blooms with particular emphases on those spectral features that may show potential to identify cyanobacteria and their development stages. Additionally, it assessed operational satellite Ocean Colour $R_{RS}(\lambda)$ products from the Ocean and Land Colour Instrument (OLCI) and Moderate Resolution Imaging Spectroradiometer (MODIS) in the presence of cyanobacteria, for which only qualitative evaluations are available from previous studies. To ascertain to what extent satellite operational data products could be used for cyanobacteria detection in the Baltic Sea, the comparison of in situ and satellite derived $R_{RS}(\lambda)$ showed poor agreement with differences particularly pronounced at the blue centre-wavelengths. Nevertheless, band-differences in the green-red spectral region for OLCI and for MODIS exhibited less dependence on atmospheric correction issues with mean absolute relative differences between 8.3% and 9.6% for OLCI and between 12.6% and 12.9% for MODIS in the presence of cyanobacteria. Additionally, they showed potential to indicate the presence and development stage of cyanobacteria blooms in Baltic Sea waters while not being sensitive to other algal blooms.

1. Introduction

Cyanobacteria are prokaryotic organisms with some species causing harmful phytoplankton blooms due to secondary toxic metabolites (Carmichael, 1992; Huisman et al., 2018). These blooms, which are of concern for riparian communities (Paerl and Paul, 2012) and exhibit higher occurrence and intensity with increasing temperatures (Munkes et al., 2021), are recurrent during summer months in the Baltic Sea where they can affect areas up to 200,000 km^2 (Kahru and Elmgren, 2014).

Cyanobacteria blooms usually show patchy distributions with concentrations characterized by small-scale horizontal and vertical heterogeneity. In calm sea conditions, they tend to accumulate at the surface as swirls, eddies or even dense, thick, opaque scums when loosing buoyancy ability. During this stage, cyanobacteria blooms can be easily observed through remote sensing optical sensors (Kutser, 2004). In fact, the gaseous vacuoles providing buoyancy aptitude to cyanobacteria cells, enhance their scattering properties. Cyanobacteria can also be identified by the absorption features of accessory pigments such as phycocyanin, unique to cyanobacteria, and carotenoids such as zeaxanthin. Nevertheless, the accurate quantification of cyanobacteria through satellite imagery is challenged by:

i. The capability of these organisms to regulate their buoyancy, which affects the hypothesis of homogenous distribution within the water volume (Metsamaa et al., 2006). Specifically, Kutser et al. (2008) showed that the amplitude and shape of remote sensing reflectance $R_{RS}(\lambda)$ are largely impacted by the vertical distribution of cyanobacterial biomass.
ii. The spatial resolution of satellite sensors, which is sometimes too coarse to capture the small-scale variability in spatial distribution characterizing these blooms. In fact, the chlorophyll-*a* concentration (*Chl-a*) may change significantly at a scale of a few tens of meters across cyanobacteria blooms (Kutser, 2004).
iii. The potential failure of the atmospheric correction process or of the accurate determination of *Chl-a* in the presence of high

* Corresponding author.
E-mail address: ilaria.cazzaniga@ec.europa.eu (I. Cazzaniga).

https://doi.org/10.1016/j.rse.2023.113464
Received 28 July 2022; Received in revised form 12 January 2023; Accepted 16 January 2023
Available online 6 February 2023

phytoplankton concentrations or scums (Kutser et al., 2018; Reinart and Kutser, 2006). Additionally, under these conditions, satellite image pixels could be flagged as bright targets due to their high radiance, and likely not properly handled by standard processors (Banks and Mélin, 2015).

iv. The spectral resolution of multispectral sensors limiting the capability of distinguishing phycocyanin in the presence of other photosynthetic pigments as shown by Zolfaghari et al. (2022).

Additionally, the execution of in situ measurements and, consequently, the validation of satellite products are difficult tasks in the presence of cyanobacteria. For instance, ship movement itself may modify the surface pattern of the blooms (Reinart and Kutser, 2006) and therefore affect the spatial representativeness of ship-based $R_{RS}(\lambda)$ measurements. Finally, focusing on the Baltic Sea, satellite ocean colour products are affected by large uncertainties and biases even in the absence of blooms due to the high concentration of chromophoric dissolved organic matter (CDOM) and the low sun elevations challenging the atmospheric correction algorithms (Zibordi et al., 2009a; Kutser et al., 2018).

A number of investigations addressed the validation of radiometric ocean colour products in the Baltic Sea. In particular, $R_{RS}(\lambda)$ products from the National Aeronautics and Space Administration (NASA) 2005.1 reprocessing of the Moderate Resolution Imaging Spectroradiometer (MODIS) were evaluated by Zibordi et al. (2009a) using data from the Ocean Color component of the Aerosol Robotic Network (AERONET-OC) while $R_{RS}(\lambda)$ products from the MODIS 2009.0 reprocessing were later assessed by Zibordi et al. (2011) with ship-based radiometric data from the Bio-Optical mapping of Marine Properties (BiOMaP) field program. These analyses showed $R_{RS}(\lambda)$ overestimations in the blue spectral region and underestimations in the red, implying a positive bias for green to blue band-ratios and a negative bias for red to green ratios. Medium Resolution Imaging Spectrometer (MERIS) reflectance products were evaluated during the 2008 cyanobacteria bloom by Kratzer and Vinterhav (2010). MERIS reflectance products were obtained alternatively through the standard MERIS processor (MEGS 7.4.1), the Case 2 Regional processor (C2R), and the Case 2 Water Properties processor developed at the Freie Universität Berlin (FUB). With a small number of matchups, the analysis showed underestimations by all processors at all spectral bands, excepted at 412 nm. Atypical spectral shapes were observed for the C2R products in the blue. Additional assessments of $R_{RS}(\lambda)$ from the Ocean and Land Colour Instrument (OLCI) were performed by Kutser et al. (2018) and Toming et al. (2017) for data products generated with the Neural Network Case 2 Regional CoastColour (C2RCC) algorithm and the Baseline Atmospheric Correction (BAC) embedded in the Level-2 processor of the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). Kutser et al. (2018) reported that C2RCC-derived $R_{RS}(\lambda)$ showed "quite reasonable" spectra when excluding cases characterized by the presence of cyanobacteria. Conversely, BAC-derived $R_{RS}(\lambda)$ exhibited systematically negative $R_{RS}(\lambda)$ at the visible centre-wavelengths (Kutser et al., 2018). More recently, Mélin (2022) and Zibordi et al. (2022) reported low accuracies at the blue centre-wavelengths for operational ocean colour products including MODIS (R2018 reprocessing) and OLCI (Collection 3), using AERONET-OC Level 2 data from Baltic Sea sites. Still, these latter validation studies did not specifically address cases exhibiting the presence of cyanobacteria. Because of this, quantitative assessments of ocean colour products for cyanobacteria blooms are currently limited for the Baltic Sea.

This notwithstanding, several studies applied $R_{RS}(\lambda)$, or the equivalent normalized-water leaving radiance $L_{WN}(\lambda)$, from satellite ocean colour sensors to quantitatively investigate cyanobacteria blooms in the Baltic Sea. For instance, Kahru et al. (2007) and Kahru and Elmgren (2014) proposed a single band approach relying on the MODIS turbid water flag (called 'TURBIDW' activated by $R_{RS}(667) > 0.0012$ sr^{-1}) to identify areas affected by cyanobacteria accumulation. Neural Networks (Riha and Krawczyk, 2011) were specifically applied to determine phycocyanin absorption. Beltrán-Abaunza et al. (2014) suggested to use the high values at 709 nm in conjunction with a dip in the blue reflectance to identify surface cyanobacteria accumulation. Exploiting in situ $R_{RS}(\lambda)$ measurements acquired during cyanobacteria blooms in the Gulf of Gdansk (southern Baltic Sea), Woźniak et al. (2016) proposed empirical algorithms relying on the band-ratios $R_{RS}(595)/R_{RS}(660)$, $R_{RS}(625)/R_{RS}(650)$ and $R_{RS}(620)/R_{RS}(710)$, best correlating with phycocyanin concentration. The same authors proposed an alternative algorithm relying on the two band-ratios $R_{RS}(620)/R_{RS}(665)$ and $R_{RS}(620)/R_{RS}(709)$. However, these latter band-ratios showed significant dependence on the spatial distribution of *Chl-a*, not necessarily linked to the presence of cyanobacteria. Karabashev (2021), solely relying on MODIS-derived $R_{RS}(\lambda)$, proposed the use of local minima of $R_{RS}(\lambda)$ at the 488 and 443 nm centre-wavelengths to detect cyanobacteria blooms. Finally, Kutser et al. (2006) used laboratory measurements of cyanobacteria optical properties combined with a bio-optical model to simulate $R_{RS}(\lambda)$ spectra.

Following a previous attempt to exploit in situ radiometric data from AERONET-OC acquired during a cyanobacteria event (Zibordi et al., 2006), which however did not propose radiometric solutions for the detection of cyanobacteria in the Baltic Sea, this study further investigates time-series of AERONET-OC $R_{RS}(\lambda)$ embracing multiple cyanobacteria blooms occurring in the northern Baltic Proper and the Gulf of Finland. Aiming at contributing to quantitative ocean colour applications, with focus restricted to the Baltic Sea, the study exploits in situ AERONET-OC data to:

i. investigate $R_{RS}(\lambda)$ spectral features that characterize the temporal development of cyanobacteria blooms;
ii. assess the accuracy of $R_{RS}(\lambda)$ from current satellite ocean colour sensors, focusing on their performance during cyanobacteria blooms;
iii. assess the capability by satellite-derived $R_{RS}(\lambda)$ of reproducing spectral features characteristic of cyanobacteria blooms;
iv. identify, among satellite spectral bands or band combinations, those most robust to atmospheric correction and consequently relevant to detect cyanobacteria.

The satellite data considered in the analyses are restricted to operational products, freely accessible in near-real time by the user community. These include the products from OLCI data from Sentinel-3A (OLCI-A) and Sentinel-3B (OLCI-B), which exhibit a comprehensive spectral band-setting for ocean colour applications (Donlon et al., 2012). MODIS data products from the Aqua (MODIS-A) and Terra (MODIS-T) platforms are also evaluated. These latter data are preferred to those from the Visible Infrared Imaging Radiometer Suite (VIIRS) because of their more suitable band-setting.

The study is organised as follows. The spectral features characterizing cyanobacteria blooms are first identified in AERONET-OC time-series from the Baltic Sea. Satellite radiometric products are successively assessed with respect to these AERONET-OC time-series and then applied to extend the investigation to other regions of the Baltic Sea and to those periods for which AERONET-OC data are not available (e.g., early spring). Finally, the validity of a variety of algorithms for cyanobacteria detection is evaluated and discussed.

2. Materials and methods

2.1. Study area and AERONET-OC data

The study focuses on the Baltic Proper and Gulf of Finland, which exhibit optically complex waters and very low $R_{RS}(\lambda)$ in the blue spectral region because of a high concentration of CDOM (Berthon and Zibordi, 2010; Kratzer and Tett, 2009). This is mainly due to several large watershed discharges and a limited exchange with the North Sea and the Skagerrak through the Danish Straits.

Recurring cyanobacteria blooms in the Baltic Sea are mostly constituted by species belonging to two functional groups: picocyanobacteria, mainly *Synechococcus* sp., and colony-forming filamentous nitrogen-fixing species (Stal et al., 2003). Dominant species in this second group are usually the diazotrophic 'scum-forming' *Nodularia spumigena*, as well as *Aphanizomenon* sp. and *Dolichospermum* (previously called *Anabaena*). *Nodularia spumigena* and *Dolichospermum* are often observed to accumulate in the top 10 m below the surface (Hajdu et al., 2007). Conversely *Aphanizomenon* sp. is usually buoyant at lower depths.

The study relies on in situ AERONET-OC spring-summer time-series of $R_{RS}(\lambda)$ at the centre-wavelengths 400, 412, 443, 490, 510, 560, 620 and 665 nm (Zibordi et al., 2021). It is recalled that AERONET-OC has been established to support satellite ocean colour activities through globally distributed standardized instruments operated on offshore fixed structures: it ensures measurements of radiance from the sea, sky radiance and direct sun irradiance (Zibordi et al., 2009b).

Two AERONET-OC sites were considered in the following analysis (see Fig. 1): Gustaf Dalen Lighthouse (GDLT) located in the northern Baltic Proper (Lat. 58.59° N, Lon. 17.47° E) and Helsinki Lighthouse (HLT) in the Gulf of Finland (Lat. 59.95° N, Lon. 24.93° E).

CE-318 9-band radiometers were operated at the GDLT and HLT sites starting from 2005. Successively, CE-318T 12-band radiometers, replicating the main visible bands of OLCI (Zibordi et al., 2021), have been deployed since 2018 at GDLT and in 2019 at HLT. The new CE-318T 12-band marine instruments provide a larger number of daily AERONET-OC measurements compared to its predecessor because of the potential to increase the number of hourly measurement sequences (i.e., from 2 up to 6). This augmented number of acquisitions allows for describing the evolution of phenomena at the time scale of a few minutes. Taking advantage of CE-318T features, AERONET-OC data from GDLT and HLT were exploited to investigate the cyanobacteria blooms occurring in 2019 and 2020.

Normalized water-leaving radiance $L_{WN}(\lambda)$ data corrected for bidirectional (BRDF) effects following Morel et al. (2002) (which is a standard AERONET-OC radiometric product) were used to determine $R_{RS}(\lambda)$ from:

$$R_{RS}(\lambda) = \frac{L_{WN}(\lambda)}{F_0(\lambda)} \quad (1)$$

where $F_0(\lambda)$ is the mean spectral extra-atmospheric irradiance from Thuillier et al. (2003). It is emphasized that the BRDF correction scheme, consistently applied to both AERONET-OC and satellite data in the visible spectral region, was not conceived for optically complex waters including those affected by blooms of filamentous cyanobacteria. Consequently, the uncertainties affecting these corrections may be rather large.

AERONET-OC sea- and sky-radiance measurements are acquired asynchronously at the different spectral bands during each measurement sequence lasting approximately a few minutes (3–4 min depending on the instrument version). Consequently, any spatial or temporal variability affecting the sea- or sky-radiance may diversely affect measurements at each spectral band. Because of this, the quality control scheme set for AERONET-OC data specifically designed to support satellite ocean colour validation during typical observation conditions characterized by spatial homogeneity and clear sky, discards sea and sky radiance measurement sequences exhibiting high variability likely affected by surface perturbations or cloudiness. In particular, the quality control process classifies data according to increasing quality levels (i.e., Level 1.0, Level 1.5 and fully quality-controlled Level 2.0, as described in Zibordi et al. (2021)) discarding measurement sequences whenever they exhibit variability beyond a given threshold at least in a single spectral band. In the presence of cyanobacteria blooms, which can be the source of high spatial inhomogeneity, a number of AERONET-OC measurements could thus be discarded by the quality control scheme. For this reason, the current study re-assessed AERONET-OC Level 1.0 data during bloom events, not included in previous validation works (e.g., Mélin, 2022; Zibordi et al., 2022). This solution allowed to recover measurements from sequences affected by high variability of the sea radiance and, additionally, those exhibiting $L_{WN}(412) > L_{WN}(443)$ commonly excluded by the AERONET-OC quality control scheme in coastal waters. This increased the number of matchups by 5% for OLCI-A and by 13% for MODIS-A, with respect to those produced solely using the Level 2.0 data. Finally, in agreement with best practice, accurate spectral matching of satellite and in situ radiometric data was ensured through band-shifting of the in situ data (Zibordi et al., 2009a). This

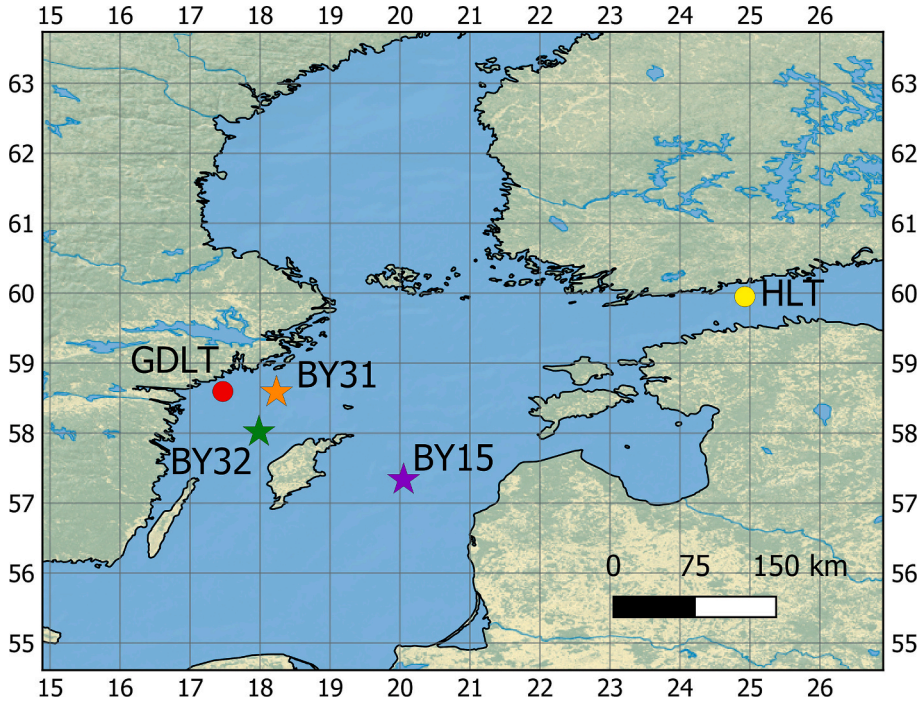

Fig. 1. Study area. The filled circles indicate AERONET-OC sites: Gustav Dalen Lighthouse Tower (GDLT) and Helsinki Lighthouse Tower (HLT). The stars indicate the field measurement sites managed by the Swedish Agency for Marine and Water Management (SwAM) and the Swedish Meteorological and Hydrological Institute (SMHI). Basemap and coastline were obtained from http://www.naturalearthdata.com.

correction, however, is expected to be minor for the OLCI matchups due to the closeness of the OLCI and CE-318T spectral bands.

2.2. In situ measurements of phytoplankton

In situ measurements of *Chl-a*, phytoplankton biovolume and community composition were obtained from the Swedish Agency for Marine and Water Management (SwAM) and the Swedish Meteorological and Hydrological Institute (SMHI), via the SHARKweb database. Data were obtained for the period 2006–2020 for two different measurement sites: BY31 (Landsort Deep, Lat. 58.58° N, Lon. 18.23° E) and BY15 (Gotland Deep, Lat. 57.33° N, Lon. 20.05° E) in the Baltic proper (see Fig. 1). The former is located at approximately 45 km from GDLT, whereas the latter is located east of Gotland Island. No equivalent data were available for the Gulf of Finland. Information on the vertical distribution of the blooms, when available, was instead gathered from the SMHI algal monthly reports at the measuring stations BY15 and BY32 (Norrkoping, Lat. 58.02° N, Lon. 17.98° E, see Fig. 1). In situ data of phytoplankton biovolume and composition were used to evaluate the seasonality of cyanobacteria and algal bloom occurrence, and additionally to confirm the presence of cyanobacteria blooms when observed through satellite images. Nevertheless, biovolume measurements are only available as integrated samples collected by a hose pipe over the first 10 m of the water column, which naturally limits their use in support of remote sensing applications. Finally, biovolume from picoplankton cyanobacteria, which do not have gas vacuoles, was not considered in this study.

2.3. Satellite data

OLCI and MODIS operational ocean colour products are expected to guarantee global coverage with near-real-time access by a multitude of users. MODIS-A and MODIS-T Level-1A data for the area of interest were processed with the *l2gen* processor embedded in the version 7.5 of SeaDAS (Mobley et al., 2016 and references therein). The resulting Level-2 data are consistent with the Ocean Biology Processing Group (OBPG) Reprocessing R2018 of NASA and associated calibration tables (NASA Goddard Space Flight Center, 2018a, 2018b). OLCI-A and OLCI-B Level-2 Reduced-Resolution (1.2-km) products from the Baseline Atmospheric Correction, Collection 3 (OL_L2M.003.01, EUMETSAT, 2021), were obtained from EUMETSAT. OLCI Full-Resolution data (300 m) from Collection 3 were not available at the time this analysis was carried out. It is recalled that, consistently with the in situ data, satellite derived $R_{RS}(\lambda)$ were also corrected for BRDF effects applying the same scheme.

Sentinel-2 MultiSpectral Instrument (MSI, Drusch et al. (2012)) images, which offer a spatial resolution of 10 m in the visible spectral bands, are effective in detecting small superficial accumulations and were consequently used to qualitatively inspect the distribution and patchiness of blooms. However, since consolidated ocean colour products from this sensor do not exist, MSI data were not applied for quantitative analyses.

OLCI-A, OLCI-B, MODIS-A and MODIS-T radiometric products were assessed against AERONET-OC $R_{RS}(\lambda)$. In addition to the analysis of the individual $R_{RS}(\lambda)$ spectral products, the assessment was performed for those band combinations (both ratios and differences) supporting cyanobacteria dedicated algorithms for which corresponding AERONET-OC spectral bands are available (see next section). For each satellite data product, a macro-pixel made of the 3 × 3 pixels centred at each AERONET-OC site was selected and the median of the $R_{RS}(\lambda)$ values was calculated excluding data affected by the standard processing flags recommended by space agencies (NASA, https://oceancolor.gsfc.nasa.gov/atbd/ocl2flags/ for MODIS, and EUMETSAT (2021) for OLCI).

Matchups of AERONET-OC and satellite $R_{RS}(\lambda)$ were constructed applying the following criteria (Mélin et al., 2011): *i.* none of the 9 pixels in the macro-pixel was flagged by the exclusion flags; *ii.* the coefficient of variation (*CV*, i.e., the ratio of macro-pixel standard deviation and average) at 560 nm (or equivalent centre-wavelength) was lower than 20%; *iii.* the time difference between the acquisition time of AERONET-OC data and the satellite overpass was lower than 1 h to minimize the impact of the high spatial and temporal variability characterizing cyanobacteria blooms. The threshold imposed on *CV* is meant to increase comparability of satellite and in situ data products. Nevertheless, in this analysis the criterion resulted in the exclusion of only 1 MODIS-T data point out of 136, which suggests quite homogeneous conditions at the scale of satellite sensors resolution for the potential matchups.

For each sensor, spectral band, and spectral band combination (ratio or difference) relevant to address bloom developments, the mean relative difference Ψ and the mean absolute relative difference $|\Psi|$, were determined as:

$$\Psi = 100\,\frac{1}{N}\sum_{i=1}^{N}\frac{\left(x_i^S - x_i^A\right)}{x_i^A} \tag{2}$$

$$|\Psi| = 100\,\frac{1}{N}\sum_{i=1}^{N}\frac{\left|x_i^S - x_i^A\right|}{x_i^A} \tag{3}$$

where N is the number of matchups, x_i^S indicates the satellite data product (e.g., $R_{RS}(\lambda)$) and x_i^A the equivalent AERONET-OC data product. Linear least-squares regressions were used to estimate the determination coefficient r^2.

Finally, to assess the relevance of the various radiometric indexes for cyanobacteria remote sensing, matchups were also created for the monitoring sites BY15 and BY31 using in situ biovolume and satellite-derived $R_{RS}(\lambda)$ for selected band combinations by applying the same criteria described above. In that case, matchups were created pairing in situ and satellite data by date because the acquisition time was not available for all the in situ measurements. The Pearson correlation (r_p) was calculated for these matchups.

2.4. Cyanobacteria-specific algorithms

The spectral features of cyanobacteria were investigated using AERONET-OC data acquired during bloom events. Techniques for the detection of cyanobacteria are generally based on the quantification of either chlorophyll-*a* or other accessory pigments typical of cyanobacteria (e.g., zeaxanthin and phycocyanin). Schalles and Yacobi (2000) and references therein detailed the natural water reflectance spectral features caused by the various pigments. In the blue spectral region, the minimum near 430 nm is due to the chlorophyll-*a* blue Soret absorption, whereas the inflection near 480 nm in the left shoulder of the green reflectance peak is produced by carotenoids (i.e., zeaxanthin) absorption. The reflectance peak occurring in the green spectral region is originated by a reduced pigment absorption between 550 and 580 nm. Phycocyanin absorption and secondary chlorophyll-*a* absorption peaks cause a trough around 620 nm and a more pronounced one near 670 nm, respectively (e.g., Mishra and Mishra, 2014; Qi et al., 2014). These absorption features are only effective in the presence of high biomass concentrations with blooms characterized by *Chl-a* generally higher than 10 mg m^{-3} (Reinart and Kutser, 2006). Schalles and Yacobi (2000) still proposed a semi-empirical equation based on a band-ratio using the local peak near 650 nm and the trough feature near 625 nm as a predictor of phycocyanin. Additionally, to reduce the influence of the spectral absorption and scattering caused by pigments other than phycocyanin, Dekker (1993) proposed a baseline subtraction algorithm based on the decrease in subsurface irradiance reflectance at 624 nm weighted through the subtraction of a baseline value calculated from two close bands not sensitive to phycocyanin absorption (i.e., 600 and 648 nm). Simis et al. (2005) proposed a nested band-ratio relying on $R_{RS}(709)/R_{RS}(665)$ to determine the absorption of phycocyanin in turbid inland waters concurrently correcting for the contribution by chlorophyll-*a* absorption at 620 nm. The phycocyanin index (PCI) by Qi et al. (2014) also uses $R_{RS}(620)$ normalized values against a baseline obtained from the linear regression of $R_{RS}(560)$ and $R_{RS}(665)$ to estimate phycocyanin concentration. PCI can also be calculated from the

Rayleigh corrected reflectance $R_{RC}(\lambda)$, circumventing the effects of inaccurate atmospheric corrections. Finally, the feature at 620 nm is also used by the three-band algorithm (PC3), proposed by Mishra and Mishra (2014) for inland waters to derive phycocyanin concentration from $R_{RS}(\lambda)$, but introducing a correction factor for the chlorophyll-a absorption at 620 nm.

Wynne et al. (2008) applied an approximation of the second derivative of $R_{RS}(\lambda)$ to MERIS data to exploit the negative spectral slope at 681 nm ($SS(681)$) with respect to the values at the 665 nm and 709 nm centre-wavelengths in the presence of cyanobacteria. The same method was proposed for MODIS data, but based on $R_{RC}(\lambda)$ at the 678, 667 and 748 nm centre-wavelengths (Wynne et al., 2013). The prominent gradient in phycocyanin absorption in the green was used by Dash et al. (2011) to estimate phycocyanin concentration with an empirical algorithm relying on the $R_{RS}(556)$-$R_{RS}(510)$ band-difference applied to the Oceansat-1 satellite Ocean Colour Monitor (OCM). Finally, the position and height of the Maximum Peak Height (MPH) was used to identify algal or cyanobacteria-dominated blooms (Matthews et al., 2012; Matthews and Odermatt, 2015). Other algorithms dedicated to the Baltic Sea (i.e., those from Kahru et al., 2007; Kahru and Elmgren, 2014; Karabashev, 2021; Woźniak et al., 2016) were already put forward in the introduction.

This study focuses on those approaches that use bands or band combinations among those available in the band set of AERONET-OC instruments in the visible spectral region where the in situ radiometric data have quantified uncertainties. It includes the analysis of the response to cyanobacteria presence of the bands and band combinations applied by Dash et al. (2011), Kahru et al. (2007), Kahru and Elmgren (2014), Karabashev (2021), Qi et al. (2014) and Woźniak et al. (2016). The band-ratios applied in the standard algorithms OC4ME and OC3M to determine *Chl-a* (Morel et al., 2007; O'Reilly et al., 1998), together with a number of additional band-ratios and band-differences detailed in Table 1, are also examined. Finally, a few algorithms based on red and near-infrared (NIR) spectral bands for which equivalent AERONET-OC data are not available, are also considered for qualitative evaluations during cyanobacteria blooms.

Table 1
Band combinations analysed in this study using both AERONET-OC and OLCI or MODIS $R_{RS}(\lambda)$ data.

Band combination	Reference
$R_{RS}(442)$-$R_{RS}(412)$	(Karabashev, 2021)
$R_{RS}(560)$-$R_{RS}(510)$	(Dash et al., 2011)
$R_{RS}(620)/R_{RS}(665)$	(Schalles and Yacobi, 2000; Woźniak et al., 2016)
$R_{RS}(560)+\frac{620-560}{665-560}*(R_{RS}(665)\text{-}R_{RS}(560))\text{-}R_{RS}(620)$	(Qi et al., 2014)
$R_{RS}(667)$ (turbid water flag)	(Kahru et al., 2007; Kahru and Elmgren, 2014)
$R_{RS}(443)/R_{RS}(560)$, $R_{RS}(490)/R_{RS}(560)$, $R_{RS}(510)/R_{RS}(560)$	OC4ME *Chl-a* algorithm for OLCI bands (Morel et al., 2007)
$R_{RS}(443)/R_{RS}(547)$, $R_{RS}(488)/R_{RS}(547)$	OC3M *Chl-a* algorithm for MODIS bands (O'Reilly et al., 1998)
$R_{RS}(560)/R_{RS}(490)$, $R_{RS}(620)/R_{RS}(490)$, $R_{RS}(665)/R_{RS}(510)$, $R_{RS}(560)$-$R_{RS}(490)$, $R_{RS}(560)$-$R_{RS}(620)$, $R_{RS}(560)$-$R_{RS}(665)$	From this study for OLCI bands
$R_{RS}(547)/R_{RS}(488)$, $R_{RS}(667)/R_{RS}(531)$, $R_{RS}(547)$-$R_{RS}(488)$, $R_{RS}(547)$-$R_{RS}(667)$	From this study for MODIS bands

2.5. Identification of bloom events in the Baltic Sea

As already mentioned in Section 2.2, in situ measurements of phytoplankton biovolume were used to confirm the occurrence of cyanobacteria or algal blooms and the related dominant species. The occurrence of surface accumulation was also occasionally confirmed by the SMHI monthly reports describing algae state as part of the Swedish national monitoring program. However, considering the relatively low frequency of in situ measurements, additional techniques were also applied to identify cyanobacteria blooms and their spatial distribution. Following previous works on the Baltic Sea (Kahru et al., 2007; Kahru and Elmgren, 2014), the threshold of 0.0012 sr^{-1} was applied to AERONET-OC $R_{RS}(665)$ to identify as a first guess any bloom event. Additionally, true colour MODIS, VIIRS, OLCI and MSI imagery of the study regions were also examined to verify the presence of blooms. Furthermore, Surface Algal Bloom products by the Finnish Environment Institute (SYKE, 2021) were used as a reference. Despite its name, this SYKE product was conceived as a cyanobacterial surface accumulation indicator and classifies waters in four states: *i. no* surface accumulations, *ii. potential* surface accumulations, *iii. likely* surface accumulations and *iv. evident* surface accumulations (Anttila et al., 2018).

Finally, in situ measurements and algal reports were used to identify generic algal blooms occurring in spring. This allowed to investigate the spectral features of these blooms with respect to those of cyanobacteria and to verify the robustness of the conclusions reached in this work.

3. Results

3.1. Cyanobacteria blooms

Fig. 2 shows biovolume concentrations measured at BY31 from 2005 and 2020. As clearly visible from the boxplot in panel b, cyanobacteria are usually detected in large amounts from June to August and occasionally in September. The maximum is usually registered in July with a few exceptions (e.g., in June 2018). Still, the accuracy of statistics is affected by the measurement date that does not necessarily coincides with the occurrence of the peak. Noteworthy, biovolume concentrations of phytoplankton other than cyanobacteria displayed in panel c show the largest values in April.

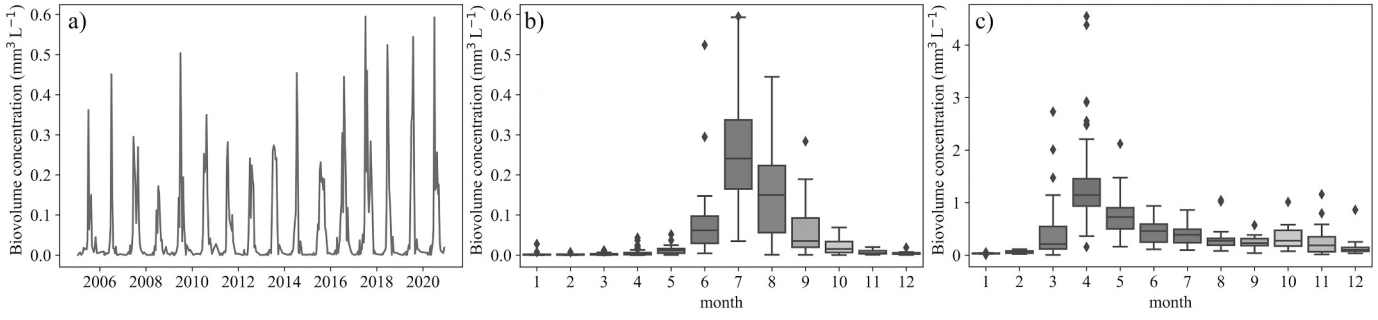

Fig. 2. (a) In situ cyanobacteria biovolume concentration at the BY31 measurement site from 2005 to 2020. (b) Corresponding monthly boxplot of cyanobacteria concentration. (c) Monthly boxplot for phytoplankton biovolume concentration other than cyanobacteria. Each box extends from the first to the third quartile with a line indicating the median value. The whiskers extend from the box by 1.5 times the inter-quartile range. Diamonds show outliers.

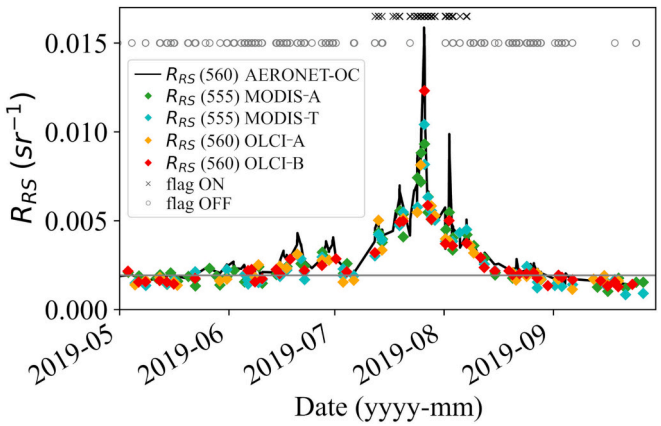

Fig. 3. AERONET-OC (black solid line) and satellite-derived $R_{RS}(\lambda)$ at 560 nm, or corresponding centre-wavelength, at GDLT in 2019. The grey line indicates the median AERONET-OC $R_{RS}(560)$ calculated for all available data outside the summer bloom periods. Crosses and circles indicate that the turbid water flag is ON or OFF, respectively.

3.2. Evolution of in situ $R_{RS}(\lambda)$ during cyanobacteria blooms

AERONET-OC data show that $R_{RS}(\lambda)$ increases at all visible bands during summer blooms. $R_{RS}(560)$, which always exhibits the highest value across the $R_{RS}(\lambda)$ spectra regardless of the presence of cyanobacteria, generally exceeds 0.004 sr^{-1} with the turbid water flag activated (Figs. 3, 6 and 8). The highest $R_{RS}(560)$ values were determined at GDLT in 2005 (not shown) and at HLT in 2019. Large and abrupt variations of $R_{RS}(\lambda)$, which are common during bloom events, are certainly explained by the heterogeneity of cyanobacteria distributions typical of colony-forming and buoyant species. Blooms do not equally affect the two regions investigated: as expected, the magnitude and duration of the events, as well as their starting dates, are quite different. However, in some cases the unavailability of AERONET-OC data limits a comprehensive analysis of these differences. For example, at the beginning of July 2019 (see Fig. 3), AERONET-OC data are not available at GDLT during the period preceding the systematic activation of the turbid water flag, which may potentially hinder the accurate identification of the starting date.

3.2.1. Bloom event at GDLT during summer 2019

AERONET-OC and satellite $R_{RS}(560)$ data at the GDLT site are displayed in Fig. 3 for the year 2019 where the cyanobacteria bloom period is indicated by black crosses according to the turbid water flag determined from AERONET-OC $R_{RS}(665)$. These data show high values already in the second half of June whereas the related SYKE products available since June 27th, classify the area immediately south of GDLT as *potential* surface accumulation (see definition in Section 2.5). At the beginning of July, $R_{RS}(560)$ drops and the SYKE products do not indicate the presence of cyanobacteria until July 10th. The highest values of 0.016 sr^{-1} for $R_{RS}(560)$ occur on July 26th. A second peak appears on August 2nd after a signal drop observed between July 27th and August 1st. With a few exceptions, the turbid water flag is activated from July 12th (no data are available on the 10th) till August 7th in correspondence of a $R_{RS}(560)$ maximum. In general, both in situ and satellite $R_{RS}(560)$ exhibit comparable temporal variations.

Fig. 4 displays the daily median $R_{RS}(\lambda)$ for selected dates and the related values normalized at 560 nm, $nR_{RS}(\lambda)$. The daily standard deviations associated with each spectrum provide an index of the variability characterizing the AERONET-OC data included in the analysis. In

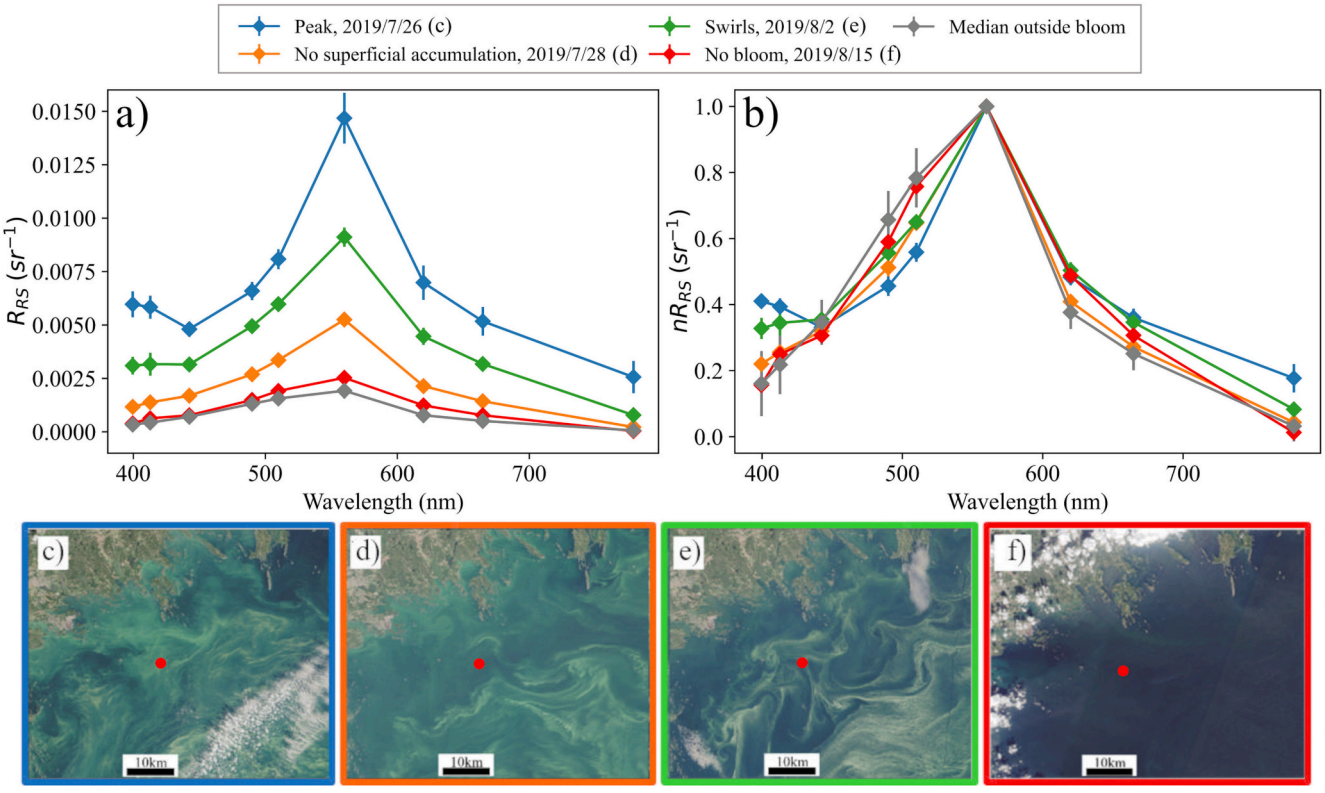

Fig. 4. Daily median AERONET-OC $R_{RS}(\lambda)$ (a) and $nR_{RS}(\lambda)$ (b) during the 2019 bloom event at GDLT. The error-bar on the y-axis indicates the standard deviation of the data applied (in a few cases, their values are smaller than the size of the symbols). The grey line indicates the median spectrum outside the summer bloom period. The Sentinel-2A or -2B MSI true colour image generated through SYKE TARKKA web application is displayed for each date (i.e., July 26th and 28th, August 2nd and 15th, extending from 58.35° N, 16.97° E to 58.85° N, 18.07° E) in panels (c-f). The colour of the square framing each satellite image is that associated with the daily AERONET-OC spectra in panels (a)-(b). The red point indicates the location of the GDLT AERONET-OC site. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

each plot, a reference spectrum constructed from AERONET-OC data outside typical bloom periods (i.e., excluding data from June to August) for the years 2018–2020 is also shown in grey. Values at 779 nm are reported for completeness while recalling that large uncertainties may affect these data and that BRDF corrections are not applied beyond 709 nm (Zibordi et al., 2021).

The spectra displayed in panels a and b of Fig. 4 show significant variations in the blue at 400 nm and 413 nm, in the red at 620 nm and 665 nm, and in the NIR at 779 nm. The highest $R_{RS}(\lambda)$ values are observed on July 26th concurrent to evident accumulations at the surface and likely very high cyanobacteria concentrations. These specific conditions are qualitatively documented through the MSI true colour image displayed in Fig. 4c. On August 2nd (see Fig. 4e) swirls moving around the site cause changes at a fine temporal scale in $R_{RS}(\lambda)$ explained by variations in surface accumulation. For this specific day, the coefficient of variation of the radiance $L_T(\lambda)$ from the sea within AERONET-OC measurement sequences shows values up to 7% at 560 nm (see Zibordi et al., 2009b for details on the AERONET-OC measurement protocol). Moreover, $R_{RS}(\lambda)$ values from successive measurement sequences performed within approximately 4 min from each other, exhibit coefficient of variations up to 24% at 560 nm. This temporal variability affecting AERONET-OC measurements could be explained by the small-scale spatially inhomogeneous conditions characterizing the surface accumulation confirmed by the MSI image acquired on that same day. This high spatial variability suggests the difficulty for either AERONET-OC point measurements or satellite products with coarser spatial resolution, to capture the whole range of changes during such events (see Fig. S1 and the dedicated paragraph in the supplementary material for further details).

In situ biovolume data at BY31 confirm the occurrence of cyanobacteria in the area. In particular, Fig. 5 shows an increase in cyanobacteria concentration since mid-June (0.15 mm^3 L^{-1} on June 17th), with a peak on July 31st (0.54 mm^3 L^{-1}) and again values close to zero at the end of August. Cyanobacteria biovolume concentrations are initially dominated by *Aphanizomenon* with *Chl-a* at BY32 indicating a peak at a depth lower than 10 m on June 16th. Conversely, in correspondence with the peak observed at the end of July, the bloom is dominated by 'scum-forming' *Nodularia spumigena* (measured on July 31st). In the same period, MODIS $R_{RS}(\lambda)$ and $R_{RS}(547)$-$R_{RS}(667)$ (see Fig. 5) increase with the maximum value occurring on July 24th (no satellite image is available for July 31st due to cloudiness). MODIS $R_{RS}(547)/R_{RS}(531)$ instead shows enhanced values in correspondence of any phytoplankton bloom, including the algal spring one. Conversely, in correspondence with the algal spring blooms, neither $R_{RS}(\lambda)$ nor $R_{RS}(547)$-$R_{RS}(667)$ increase and the water appears dark in satellite true-colour images. This peculiarity can be explained by the fact that algal spring blooms usually do not accumulate at the surface as cyanobacteria do. Additionally, cyanobacteria species furnished with gas vacuoles exhibit enhanced back-scattering properties with respect to most algae species.

3.2.2. Bloom event at HLT during summer 2019

The $R_{RS}(\lambda)$ evolution at HLT during the 2019 bloom event is displayed in Fig. 6. Although the SYKE product indicates the *potential* presence of a bloom in the area surrounding HLT during the second week of July 2019, the presence of surface accumulations and swirls is already visible in satellite imagery at the end of June. Still, the bloom does not appear to strictly influence the water at the AERONET-OC site. On July 7th and 8th, when the turbid water flag applied to AERONET-OC data is activated, the satellite imagery is largely affected by clouds. Nevertheless, the cloud-free pixels show the presence of accumulations and swirls in the Gulf of Finland explaining the increased AERONET-OC $R_{RS}(\lambda)$ at HLT. Successively, $R_{RS}(560)$ increases and decreases until the peak occurring between July 27th and 28th. High values of $R_{RS}(\lambda)$ are observed for a few hours on July 27th: the related median spectrum in

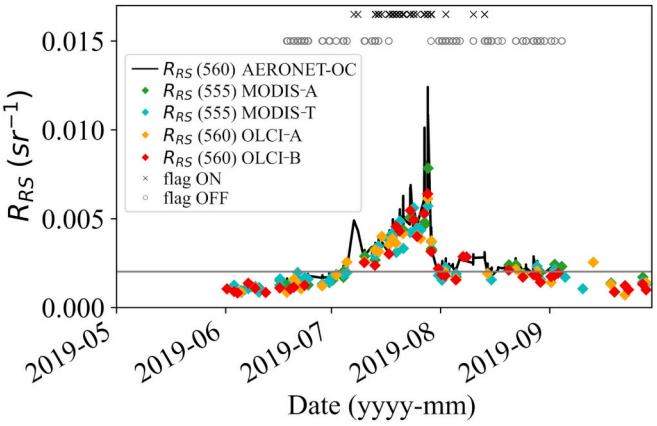

Fig. 6. AERONET-OC (black solid line) and satellite-derived $R_{RS}(\lambda)$ at 560 nm or corresponding centre-wavelengths at HLT in 2019. The grey line indicates the median $R_{RS}(560)$ calculated for all available data outside the summer bloom periods. Crosses and circles indicate that the turbid water flag is ON or OFF, respectively.

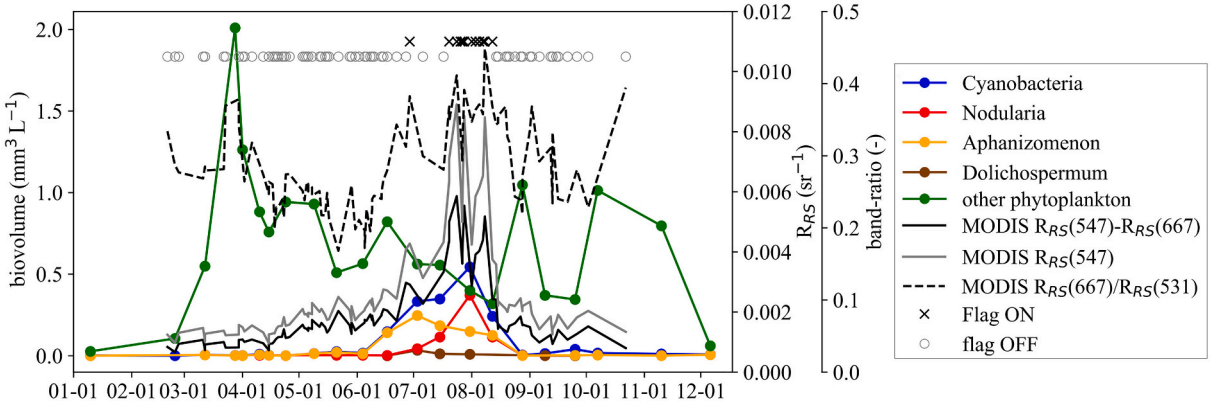

Fig. 5. In situ measurements of biovolume concentration of cyanobacteria (in blue), *Nodularia spumigena* (in red), *Aphanizomenon* (in orange), *Dolichospermum* (in brown) and other phytoplankton species (in green) at BY31 in 2019. MODIS-A $R_{RS}(547)$, $R_{RS}(547)$-$R_{RS}(667)$ and $R_{RS}(667)/R_{RS}(531)$ are provided on the right axes. Crosses and circles indicate that the MODIS-A turbid water flag is ON (i.e., $R_{RS}(667) \geq 0.0012$ sr^{-1}) or OFF, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

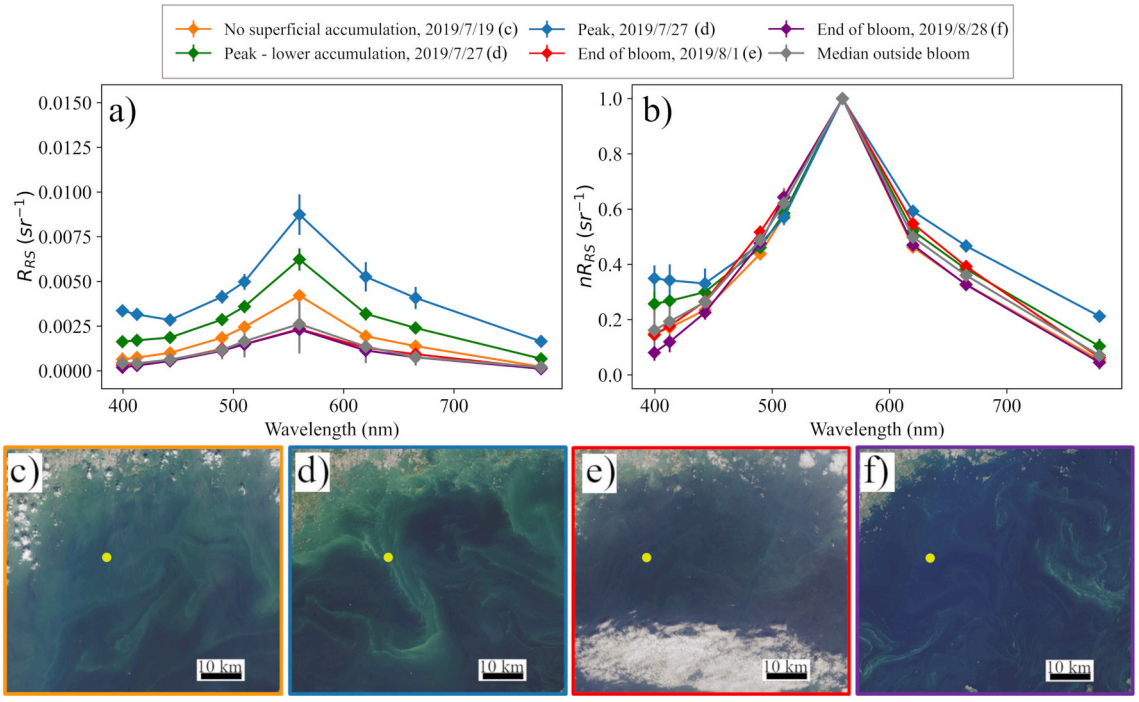

Fig. 7. Median AERONET-OC $R_{RS}(\lambda)$ (a) and $nR_{RS}(\lambda)$ (b), calculated on a daily basis or shorter time frame expected to represent homogeneous surface conditions during the 2019 bloom event at HLT. The error-bar on the y-axis indicates the standard deviation of the data applied (in a few cases, their values are smaller than the size of the symbols). The grey line indicates the median spectrum outside the summer bloom period. The Sentinel-2A or -2B MSI true colour image from 59.67° N,24.49° E to 60.19° N,25.68° E and generated through SYKE TARKKA web application, is displayed for each date (i.e., July 19th and 27th, August 1st and 28th) in panels (c-f). The colour of the square framing each satellite image is that associated with the daily AERONET-OC spectra in panels (a)-(b). Panel d relates to the same day of both blue and green spectra. The yellow point indicates the location of the HLT AERONET-OC site. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 7 exhibits a shape very similar to that displayed in correspondence of the peak occurring at GDLT. The $R_{RS}(\lambda)$ values rapidly decrease at the end of July. Indeed, from July 27th to 29th the bloom is visible from satellite imagery as a big filamentous structure just in front of the Helsinki Gulf (see Fig. 7d) and disappears on August 1st (see Fig. 7e). The wind speed, included in the AERONET-OC data records, indicates daily mean values increasing from 4.3 m s^{-1} on July 27th to 7.3 m s^{-1} on 30th, likely hindering the accumulation of cyanobacteria at the surface. Successively, until the end of September, high-resolution imagery shows the presence of cyanobacteria accumulation at the surface appearing as swirls in the Gulf of Finland. However, as revealed by the turbid water flag occasionally activated until August 13th, these cyanobacteria are only intermittently detected by AERONET-OC data.

Fig. 7, similar to Fig. 4, shows the daily median $R_{RS}(\lambda)$ and $nR_{RS}(\lambda)$ spectra, except for the peak date of July 27th exhibiting high values of the standard deviation suggesting fast-varying conditions. For this date, the daily values from 8:00 to 12:00 and from 12:00 to 14:00 GMT are separately presented through the green and blue spectra. Panels a and b show that the $R_{RS}(\lambda)$ spectral shape does not exhibit significant differences during the first part of the bloom event (see the orange spectrum). Indeed, in the absence of large surface accumulation (as shown by the MSI image in Fig. 7c), the $R_{RS}(\lambda)$ values indicate a large increase with the bloom across the whole spectrum in the visible while exhibiting minor spectral variations in the $nR_{RS}(\lambda)$ values. Conversely, at the peak, $nR_{RS}(\lambda)$ exhibits increased values in the blue, red and NIR, with a local minimum at 443 nm.

3.2.3. Bloom event at GDLT during summer 2020

The bloom affecting GDLT in 2020 also exhibits a complex temporal evolution. Thin accumulations are visible in MSI images of early June (unfortunately AERONET-OC data are not available for that specific period). Successively, the AERONET-OC $R_{RS}(560)$ time-series displayed in Fig. 8 shows two peaks corresponding to sequential episodes of the bloom separated by a period exhibiting less pronounced effects. $R_{RS}(560)$ gradually increases till reaching a peak on June 28th contemporary to the appearance of greenish waters in satellite imagery (e.g., see Fig. 9c-d). During August, $R_{RS}(560)$ shows repeated maxima with satellite imagery displaying very heterogeneous accumulation at the surface near GDLT (see Fig. 9f). Looking at the spectra displayed in Fig. 9a, the $R_{RS}(\lambda)$ median spectrum determined during the peak of June looks similar to those observed in 2019 with a relative minimum at 443 nm and the highest value at 560 nm. Looking at $nR_{RS}(\lambda)$ in Fig. 9b, the

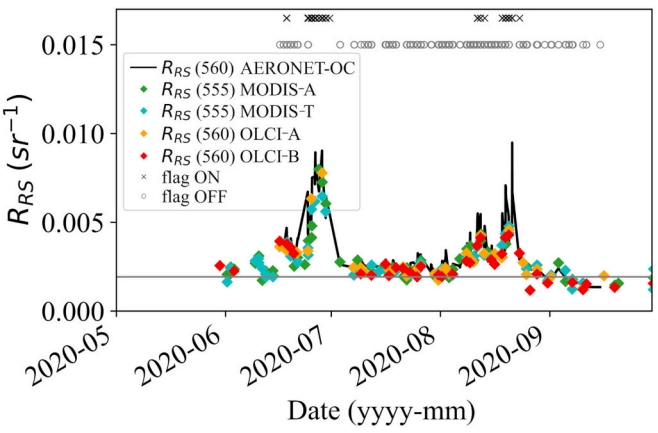

Fig. 8. AERONET-OC (black solid line) and satellite-derived $R_{RS}(\lambda)$ at 560 nm or corresponding centre-wavelengths, at GDLT in 2020. The grey line indicates the median $R_{RS}(560)$ calculated for all available data outside the summer bloom periods. Crosses and circles indicate the turbid water flag is ON or OFF, respectively.

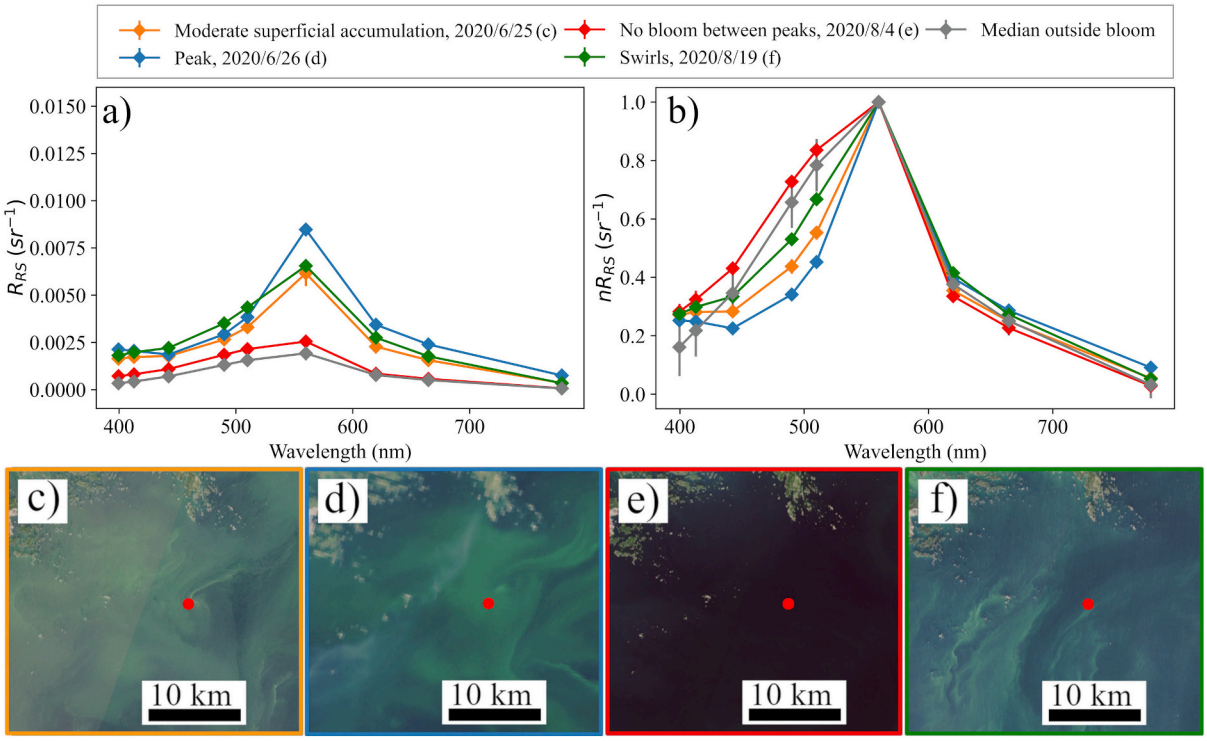

Fig. 9. Daily median AERONET-OC $R_{RS}(\lambda)$ (a) and $nR_{RS}(\lambda)$ (b) during the 2020 bloom event at GDLT. The error-bar on the y-axis indicates the standard deviation of the data applied (in a few cases, their values are smaller than the size of the symbols). The grey line indicates the median spectrum outside the summer bloom period. The Sentinel-2A or -2B MSI or OLCI-A true colour image extending from 58.47° N,17.13° E to 58.73° N, 17.67° E and generated through the SYKE TARKKA web application is displayed for each date (i.e., June 25th and 26th, August 4th and 19th) in panels (c-f). The colour of the square framing each satellite image is that associated with the daily AERONET-OC spectra in panels (a)-(b). The red point indicates the location of the GDLT AERONET-OC site. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

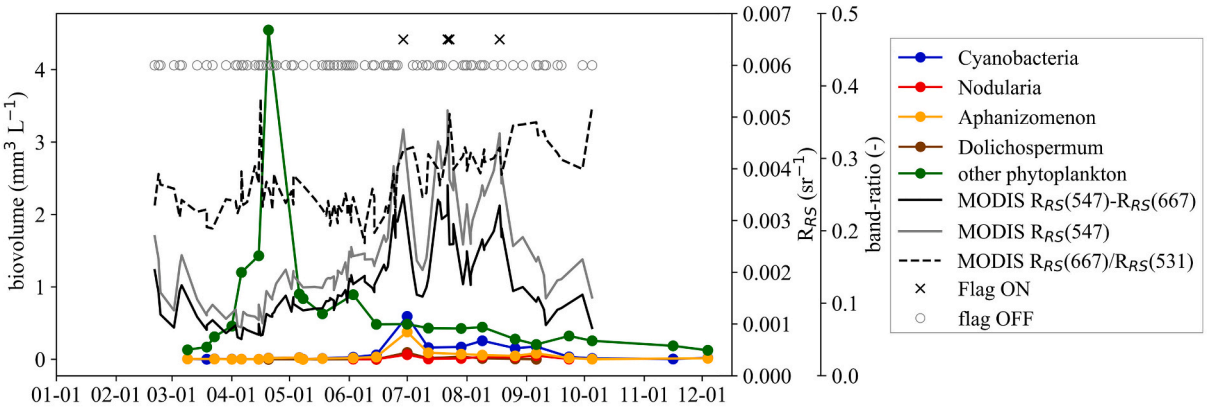

Fig. 10. In situ measurements of biovolume concentration of cyanobacteria (in blue), *Nodularia spumigena* (in red), *Aphanizomenon* (in orange), *Dolichospermum* (in brown) and other phytoplankton species (in green) at BY31 in 2020. MODIS-A $R_{RS}(547)$, $R_{RS}(547)$-$R_{RS}(667)$ and $R_{RS}(667)/R_{RS}(531)$ are provided on the right axes. Crosses and circles indicate that the MODIS turbid water flag is ON (i.e. $R_{RS}(667) \geq 0.0012$ sr^{-1}) or OFF, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

change in slope in the blue spectral region is notable, well characterizing the "bloom" with respect to the "non-bloom" conditions.

Biovolume data at BY31 indicate that *Nodularia spumigena* appears in very low concentration in mid-June. However, Fig. 10 shows cyanobacteria peak at the beginning of July, with some presence until the beginning of September in conjunction with other phytoplankton species still dominating the bloom. The evolution of MODIS-A $R_{RS}(547)$ exhibits a sequence of peaks and troughs suggesting temporal and spatial changes of surface accumulation. Unfortunately, there were no in situ biovolume measurements available matching these peaks.

3.2.4. *In situ $R_{RS}(\lambda)$ spectral features during blooms*

The time-series of various $R_{RS}(\lambda)$ spectral values, band-ratios and band-differences introduced in Section 2.4, were evaluated. As expected, $R_{RS}(\lambda)$ shows increasing values in correspondence of blooms. $R_{RS}(665)$, used to raise the turbid water flag, increases but the threshold of 0.0012 sr^{-1} seems to be ineffective (i.e., too high) to identify blooms in their early stage as shown in Fig. 11 by the intermittent activation of the turbid water flag at the beginning of blooms. The band-differences shown in panels b, d, f of Fig. 11 ($R_{RS}(560)$-$R_{RS}(510)$, $R_{RS}(560)$-$R_{RS}(490)$, $R_{RS}(560)$-$R_{RS}(620)$, and $R_{RS}(560)$-$R_{RS}(665)$), as well as the PCI band combination (not displayed), exhibit very similar temporal

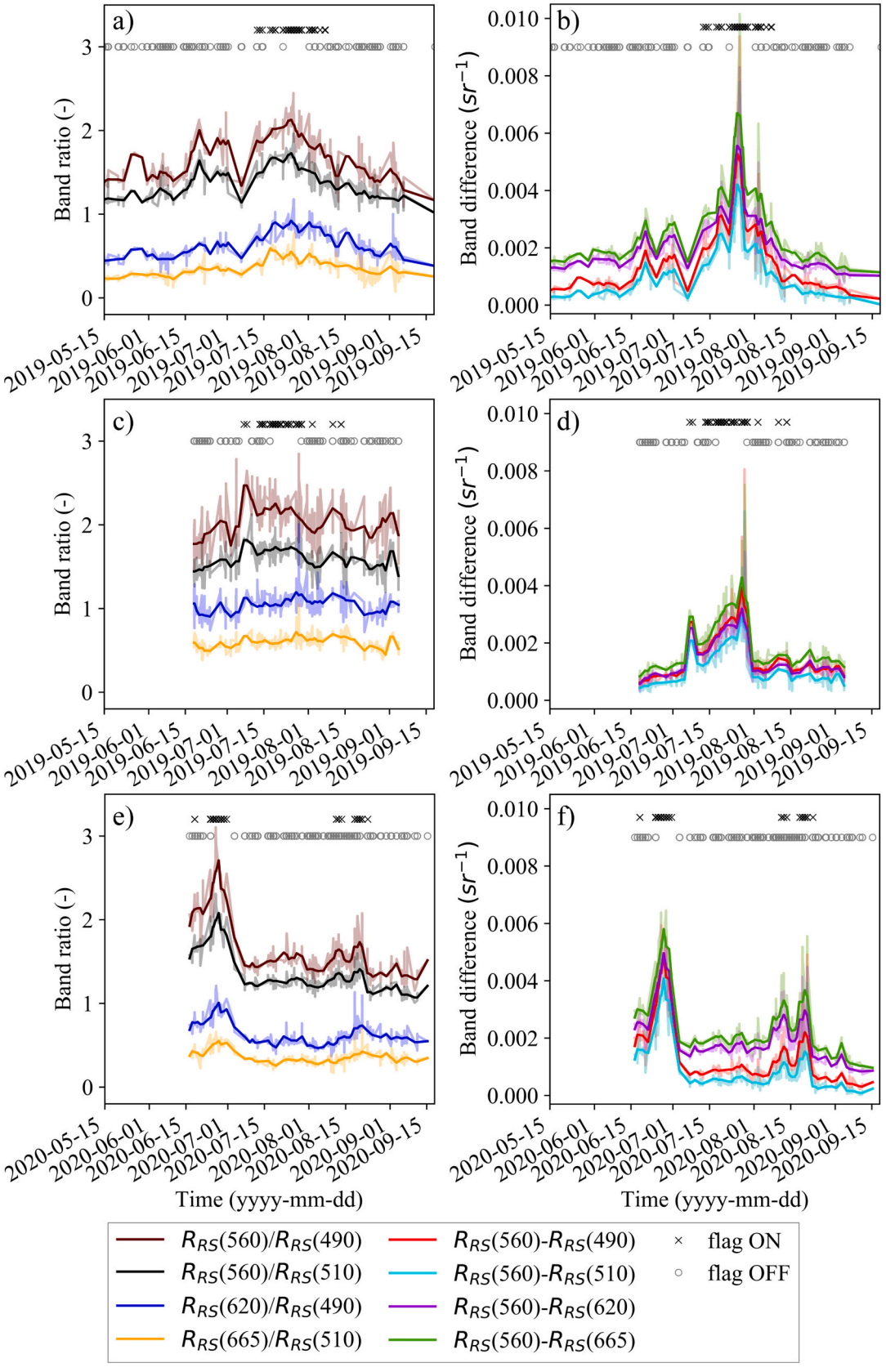

Fig. 11. AERONET-OC (a,c,e) band-ratios $R_{RS}(560)/R_{RS}(490)$, $R_{RS}(620)/R_{RS}(490)$, $R_{RS}(665)/R_{RS}(510)$ and (b,d,f) band-differences $R_{RS}(560)$-$R_{RS}(490)$, $R_{RS}(560)$-$R_{RS}(620)$, $R_{RS}(560)$-$R_{RS}(665)$ at GDLT (a,b,e,f) and HLT (c,d) in 2019 (a-d) and at GDLT in 2020 (e,f). The solid lines indicate moving average values determined with a time window of 3 days.

evolutions with values increasing in correspondence of blooms. Conversely, the band-difference $R_{RS}(412)$-$R_{RS}(442)$ exhibiting very noisy values, was not included in Fig. 11. The band-ratio $R_{RS}(620)/R_{RS}(665)$ (not shown) also exhibits noisy values at GDLT, with maxima occurring away from peak dates. Among the band-ratios used in OC4ME, the inverse of $R_{RS}(490)/R_{RS}(560)$ and $R_{RS}(510)/R_{RS}(560)$ displayed in Fig. 11a, c and e, show increasing values in correspondence of cyanobacteria blooms. Finally, other band-ratios leading to apparently relevant results are $R_{RS}(620)/R_{RS}(490)$ and $R_{RS}(665)/R_{RS}(510)$ (panels a, c, e). Results illustrated in panels a, b, e and f of Fig. 11 for the GDLT site, show an increase of values with the bloom occurrence between the end of July and the beginning of August 2019, and additionally at the end of June and in August 2020. For the 2020 event at GDLT, $R_{RS}(560)/R_{RS}(490)$ and band-differences show a more pronounced increase during the early phase of the event in June with respect to the final one occurring in August. This could be explained by the higher cyanobacteria biovolume characterizing that early stage of the event. A general increase of band-ratios and band-differences with bloom appearance, in addition to those already shown in Fig. 11, is also observed for $R_{RS}(412)/R_{RS}(442)$, $R_{RS}(412)/R_{RS}(490)$, $R_{RS}(412)/R_{RS}(510)$, $R_{RS}(560)/R_{RS}(510)$, $R_{RS}(620)/R_{RS}(510)$, $R_{RS}(665)/R_{RS}(490)$, $R_{RS}(620)$-$R_{RS}(400)$ and $R_{RS}(620)$-$R_{RS}(665)$.

When looking at HLT (see Fig. 11d), band-differences and PCI (not shown) exhibit a peak at the end of July with features comparable to those observed at GDLT. Conversely, band-ratios do not provide any clear indication of bloom occurrence (see Fig. 11c). At both sites, the additional ratio $R_{RS}(400)/R_{RS}(443)$ indicates increasing values until the bloom peak (not shown). However, due to the very low signal at 400 nm leading to a decrease of the signal-to-noise ratio, the trend determined by $R_{RS}(400)/R_{RS}(443)$ should be considered with caution. Overall, these results clearly indicate that band-differences, and in particular PCI and $R_{RS}(560)$-$R_{RS}(620)$, but also $R_{RS}(560)$-$R_{RS}(665)$ even though not benefitting of the 620-nm centre-wavelength, appear relevant indexes for the detection and quantification of cyanobacteria in the Baltic Sea.

The optical properties of *Aphanizomenon* and *Nodularia spumigena* well justify the amplified $R_{RS}(\lambda)$ spectra and the enhanced band-difference values. The specific absorption coefficients of *Aphanizomenon* and *Nodularia spumigena* (as shown in Fig. 2 by Wojtasiewicz and Stoń-Egiert (2016) and in Fig. 1 by Kutser et al. (2007)) in fact exhibit a minimum around 550–560 nm and an increase towards longer wavelengths until approximately 670 nm, which enhance the differences $R_{RS}(560)$-$R_{RS}(665)$, $R_{RS}(560)$-$R_{RS}(620)$ and $R_{RS}(620)$-$R_{RS}(665)$ with cyanobacteria accumulation. Additionally, the backscattering coefficient of *Aphanizomenon* and *Nodularia spumigena*, mainly due to gas vacuoles and exhibiting fairly flat spectral values (Wojtasiewicz and Stoń-Egiert, 2016), is responsible for the increase of $R_{RS}(\lambda)$ at all wavelengths, which potentially allows to distinguish cyanobacteria from algal blooms in the Baltic Sea.

3.3. Assessment of ocean colour radiometric data products

The cumulative number of matchups produced for GDLT and HLT is 80, 93, 137, 135 for OLCI-A, OLCI-B, MODIS-A and MODIS-T respectively. Among them, the matchups affected by cyanobacteria presence are 12, 18, 29 and 26, respectively, as identified relying on the turbid water flag applied to AERONET-OC data as described in Section 3.2. When considering the whole dataset, results from the analysis of matchup data confirm the poor performance of the atmospheric correction process reported in previous studies in the Baltic Sea (e.g., Mélin, 2022; Zibordi et al., 2022), particularly in the blue bands, with respect to other sites. The low performance of the atmospheric correction could be explained by a combination of high solar zenith angles, low $R_{RS}(\lambda)$ values and issues with the NIR corrections, as reported by Mélin (2022), Zibordi et al. (2009a) and Kutser et al. (2018). Specifically a systematic underestimation of OLCI-A and OLCI-B derived $R_{RS}(\lambda)$ across most of the spectrum is noted, whereas MODIS-A and MODIS-T $R_{RS}(\lambda)$ show a more complex spectral dependence with overestimations at the blue centre-wavelengths and underestimations at the red ones. The matchup data analysis shows that results are not usually degraded in the presence of cyanobacteria. The scatterplots and statistics for single bands and sites are provided in the supplementary material (Fig. S2). It is worth, however, mentioning that the results for the red bands usually show a large underestimation of the satellite value (with Ψ typically of about −41% for OLCI-A and OLCI-B and −8% for MODIS-A and MODIS-T), likely leading to the omission of some cyanobacteria cases if the turbid water flag is applied to detect cyanobacteria using satellite data.

Conversely, when considering the band combinations proposed in Section 3.2.4, some of them show appreciable agreement with the in situ data. Among band-ratios, only $R_{RS}(620)/R_{RS}(510)$ and $R_{RS}(665)/R_{RS}(510)$ show reasonably good results for either OLCI-A and OLCI-B, with values well distributed near the 1:1 line, even though largely underestimated (see the results of band-ratio analyses in the supplementary material's Figs. S3 and S4). A remarkable agreement between satellite and in situ data is instead shown by the band-differences $R_{RS}(560)$-$R_{RS}(620)$ and $R_{RS}(560)$-$R_{RS}(665)$ (see Fig. 12) with $|\Psi|$ of

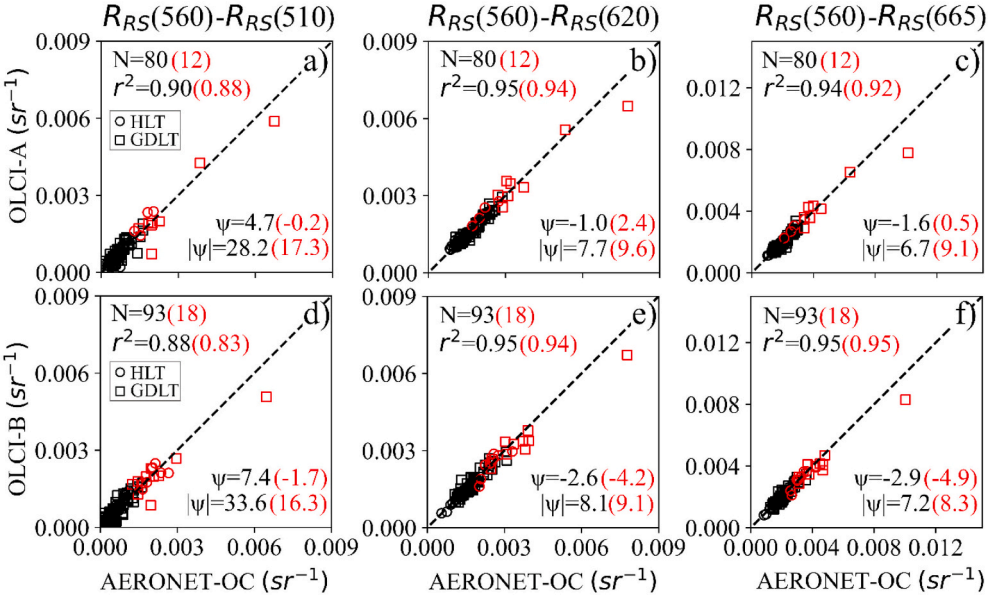

Fig. 12. Scatterplots illustrating the matchup comparison of OLCI-A or OLCI-B with AERONET-OC for $R_{RS}(\lambda)$ band-differences from GDLT (indicated by squares) and HLT (indicated by circles). Ψ is the mean of relative differences and $|\Psi|$ the mean of absolute relative differences, both in %. N is the number of matchups. The red symbols and numbers indicate matchups and corresponding statistics related to cyanobacteria blooms identified applying the turbid water flag to AERONET-OC data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

7–8%. Matchups affected by the presence of cyanobacteria show slightly worse statistics (see the red numbers in brackets in Fig. 12). Reasonable results are also obtained for $R_{RS}(560)$-$R_{RS}(510)$, with r^2 approaching 0.90, and $|\Psi|$ equal to 28.2% and 33.6% for OLCI-A and OLCI-B, respectively (down to 17.3% and 16.3% in the presence of cyanobacteria). Slightly lower performance is also obtained for the PCI band combination with r^2 between 0.88 and 0.92, and $|\Psi|$ equal to 16.4% and 15.5% for OLCI-A and OLCI-B, respectively (down to 11.3% and 11.9% in the presence of cyanobacteria). Finally, $R_{RS}(620)$-$R_{RS}(665)$ (not shown) exhibits a slightly worse performance with r^2 equal to 0.72 and 0.84, and $|\Psi|$ equal to 16.7% and 18.4% for OLCI-A and OLCI-B, respectively, while performing better in the presence of cyanobacteria (r^2 equal to 0.78 and 0.89 and $|\Psi|$ equal to 13.3% and 10.9% for OLCI-A and OLCI-B, respectively).

Unsurprisingly, band-differences involving any centre-wavelength in the 400–443 nm spectral region (not shown) exhibit a very limited performance.

Similar to OLCI, MODIS band-ratios also perform rather poorly (see Fig. S5 in supplementary material). Good agreement is instead shown for $R_{RS}(547)$-$R_{RS}(667)$ in Fig. 13 (panels b and d), slightly worsening in the presence of cyanobacteria (still with $|\Psi|$ below 13% for both MODIS-A and MODIS-T). $R_{RS}(547)$-$R_{RS}(488)$ (see Fig. 13a and c) shows a slight underestimation with Ψ of −5.9% and −6.2% for MODIS-A and MODIS-T, respectively, more pronounced for the larger values characterizing the bloom events (−9.8% and −7.2%, respectively).

The in situ and satellite data displayed in Fig. 14 illustrate $R_{RS}(\lambda)$ matching spectra in the presence of cyanobacteria blooms. These qualitative comparisons suggest consistent spectral shapes between in situ and satellite $R_{RS}(\lambda)$, despite some large biases. The characteristic peak at 709 nm (Matthews et al., 2012) is occasionally shown by OLCI data products.

The above results suggest the potential use of OLCI and MODIS data products to investigate cyanobacteria blooms in the Baltic Sea. This is best achievable by exploiting the green-red part of the spectrum through the band-differences $R_{RS}(560)$-$R_{RS}(620)$ and $R_{RS}(560)$-$R_{RS}(665)$ for OLCI-A and OLCI-B, or $R_{RS}(547)$-$R_{RS}(667)$ for MODIS-A and MODIS-T. On the contrary, the noise affecting satellite derived $R_{RS}(\lambda)$ at the blue centre-wavelengths does not support a reliable identification and spectral characterization of blooms, as suggested by Fig. S2 in supplementary material and the spectra displayed in Fig. 14. This figure also shows that OLCI data do not always exhibit the expected minimum at the 443 nm centre-wavelength observed at bloom peaks (see also Sections 3.2.1–3.2.4, and Karabashev (2021)). Conversely, this minimum may characterize satellite data products outside bloom periods (see for example the spectra of August 28th 2019 for OLCI-A and MODIS-A and -T in Fig. S6), once more highlighting potential atmospheric correction issues in this spectral region.

3.4. Satellite-derived $R_{RS}(\lambda)$ temporal evolution and spectral features during blooms

AERONET-OC data in the Baltic Sea allowed to investigate $R_{RS}(\lambda)$ at specific locations in the northern Baltic Proper and in the Gulf of Finland from late spring to early autumn. These data, however, were generally not available during the month of April when the spring bloom consisting of diatoms and dinoflagellates usually occurs. The application of satellite-derived $R_{RS}(\lambda)$ allowed to extend the analysis in time and space. Figs. 5 and 10 show MODIS-A $R_{RS}(547)$, the band-difference $R_{RS}(547)$-$R_{RS}(667)$ and the band-ratio $R_{RS}(667)/R_{RS}(531)$ for 2019 and 2020. Results clearly indicate that $R_{RS}(547)$ and the band-difference increase in the presence of cyanobacteria. Conversely, they are not sensitive to algal blooms, while the band-ratio shows relative maxima also during algal spring blooms.

To further evaluate the effectiveness of the indexes presented in Section 2.4 and 3.2, in situ measurements of cyanobacteria biovolume concentration were matched with satellite-derived $R_{RS}(\lambda)$. This analysis was restricted to bands or band combinations showing good performance in the assessment of radiometric data products: $R_{RS}(665)$ used for the turbid water flag; the PCI band combination and the band-differences $R_{RS}(560)$-$R_{RS}(510)$, $R_{RS}(560)$-$R_{RS}(620)$ and $R_{RS}(560)$-$R_{RS}(665)$ for OLCI or the corresponding $R_{RS}(547)$-$R_{RS}(667)$ for MODIS; and finally, the band-ratio $R_{RS}(665)/R_{RS}(510)$ for OLCI or the corresponding $R_{RS}(667)/R_{RS}(531)$ for MODIS.

Fig. 15 displays the scatterplots obtained for OLCI-A and OLCI-B and

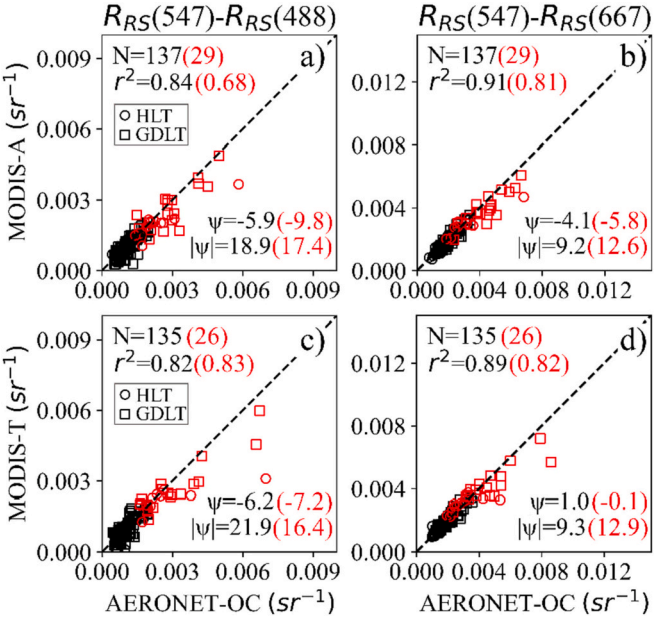

Fig. 13. Scatterplots illustrating the comparison of MODIS-A or MODIS-T with AERONET-OC $R_{RS}(\lambda)$ band-differences for GDLT (indicated by squares) and HLT (indicated by circles). Ψ is the mean of relative differences and $|\Psi|$ the mean of absolute relative differences, both in %. N is the number of matchups. The red symbols and numbers indicate matchups and corresponding statistics related to cyanobacteria blooms identified applying the turbid water flag to AERONET-OC data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

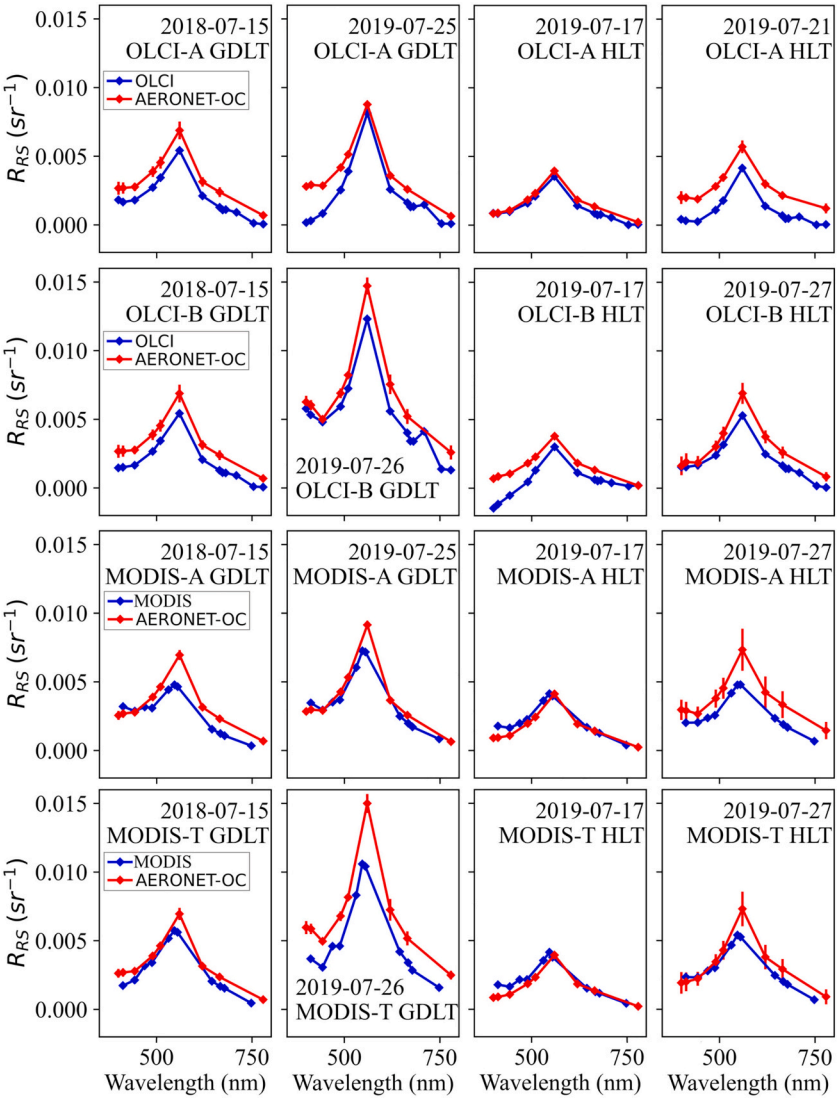

Fig. 14. Sample AERONET-OC and corresponding OLCI-A, OLCI-B, MODIS-A or MODIS-T $R_{RS}(\lambda)$ matching spectra during cyanobacteria blooms at GDLT and HLT. Median and standard deviation are reported for the AERONET-OC spectra available within a ± 1 h interval from the satellite overpass.

for MODIS-A and MODIS-T at the two in situ measurements locations BY31 and BY15. As already mentioned in the previous section, the turbid water flag activated by $R_{RS}(665) > 0.0012$ sr^{-1} in several cases does not detect the presence of cyanobacteria partially due to an underestimation of the $R_{RS}(665)$ values. Still, it shows some proportionality between biovolume concentration and $R_{RS}(\lambda)$ magnitude ($r_p = 0.79$ for OLCI and $r_p = 0.71$ for MODIS, *p-value* < 0.01, see panels a and b). There is also a good relation ($r_p = 0.70$, *p-value* < 0.01) between biovolume concentration and $R_{RS}(560)$-$R_{RS}(510)$ (panel c), and a stronger one between cyanobacteria biovolume and $R_{RS}(560)$-$R_{RS}(620)$ (not shown, $r_p = 0.78$, *p-value* < 0.01) or $R_{RS}(560)$-$R_{RS}(665)$ (panel e, $r_p = 0.77$, *p-value* < 0.01), or the corresponding MODIS bands (panel f, $r_p = 0.74$, *p-value* < 0.01). Similar results are also proposed for PCI (panel d, $r_p = 0.78$, *p-value* $<$ 0.01). As expected, band-ratios exhibit a weak relation with cyanobacteria biovolume concentrations (see panels g and h, $r_p < 0.5$, *p-value* $<$ 0.01), which is partly explained by the low accuracy shown by satellite-derived data for these band combinations and partly by the fact that they are also sensitive to algal blooms. To evaluate the performance of those existing algorithms for which no equivalent could be reproduced through AERONET-OC data, values obtained from OLCI $R_{RS}(\lambda)$ through PC3 (Mishra and Mishra, 2014) and *SS*(681) (Wynne et al., 2013) and from OLCI $R_{RC}(\lambda)$ through MPH (Matthews et al., 2012) and PCI, were also matched with biovolume concentration. No significant relationships could be established for MPH ($r_p = -0.23$, *p-value* > 0.05), *SS*(681) ($r_p < 0.1$, *p-value* > 0.05), and PC3 ($r_p = -0.16$, *p-value* > 0.05). These results are probably explained by the different concentrations characterizing cyanobacteria blooms in the basins for which they were developed, responsible for very different $R_{RS}(\lambda)$ shapes in the red and NIR spectral regions (see for example the $R_{RS}(\lambda)$ spectra in Fig. 1 in Mishra and Mishra (2014)). Conversely, a quite good relation is observed for PCI calculated from $R_{RC}(\lambda)$ ($r_p = 0.61$, *p-value* < 0.01, see Fig. S7), but still weaker than that computed with $R_{RS}(\lambda)$.

It is further remarked that the in situ measurements of biovolume refer to integrated values within the first 10 m below the surface, and that blooms occurring away from the surface are not seen from satellite sensors due to the high attenuation of light in these waters. For this reason, the definition of any relation between concentrations and $R_{RS}(\lambda)$ or even the determination of any threshold for cyanobacteria identification solely based on these data, would be speculative.

A few of these indexes were also applied to a test case, and specifically to an OLCI-B image acquired during a cyanobacteria bloom on August 19th, 2020, when the biovolume phytoplankton concentration available at station BY15 indicated the presence of a cyanobacteria bloom dominated by *Nodularia spumigena*. From this test case, it is possible to confirm that the underestimation of satellite-derived $R_{RS}(665)$ does not allow to fully identify the presence of cyanobacteria. Band-differences and the PCI band

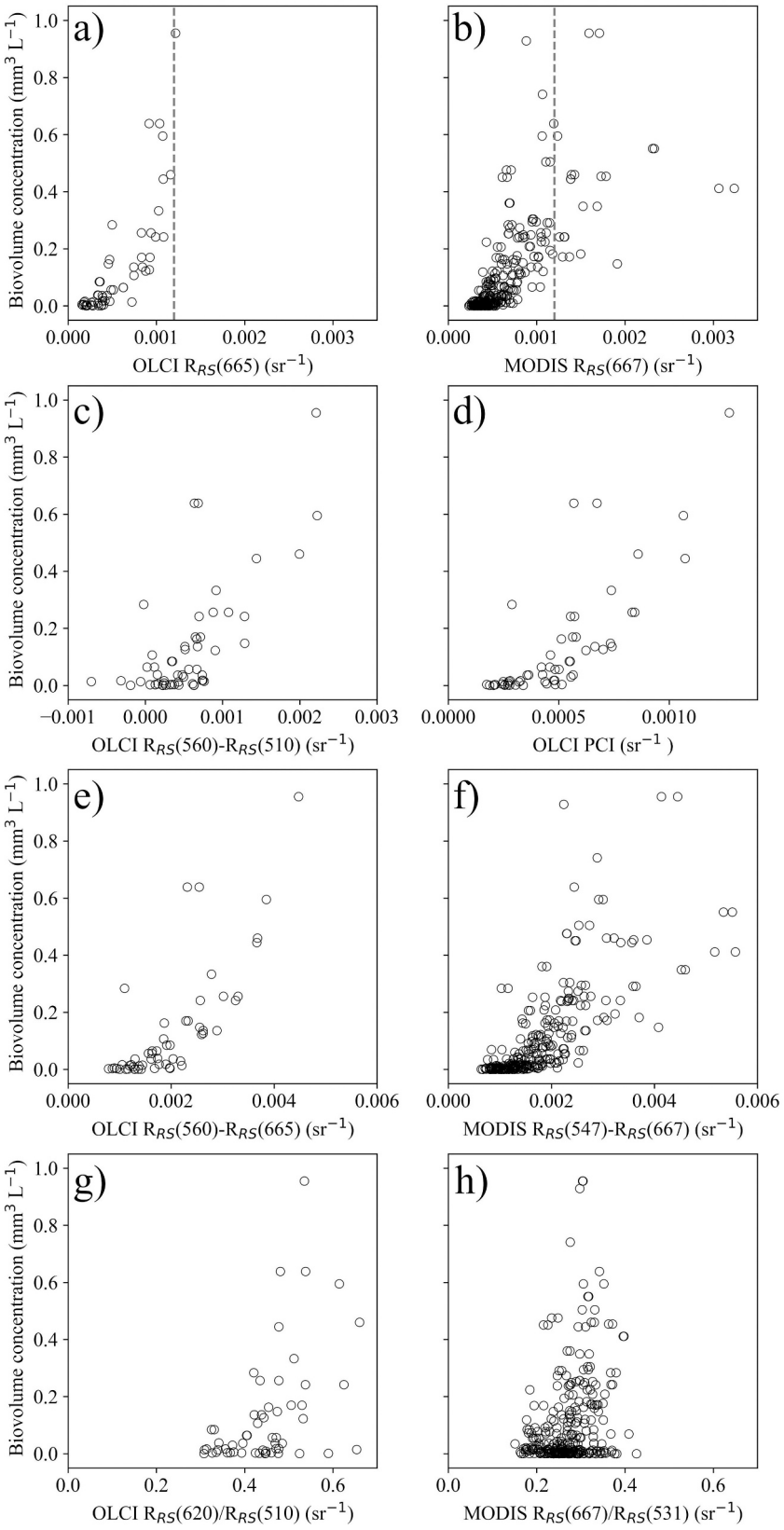

Fig. 15. Scatterplot for cyanobacteria biovolume concentrations measured in situ at BY31 and BY15 versus a) OLCI R_{RS}(665) and b) MODIS R_{RS}(667) where the vertical lines indicate the turbid water flag threshold (i.e. 0.0012 sr^{-1}), c) OLCI R_{RS}(560)-R_{RS}(510), d) OLCI PCI band combination from R_{RS}(λ), e) OLCI R_{RS}(560)-R_{RS}(665), f) MODIS R_{RS}(547)-R_{RS}(667), g) OLCI R_{RS}(620)/R_{RS}(510) and h) MODIS R_{RS}(667)/R_{RS}(531).

combination instead appear capable of capturing cyanobacteria and their spatial distribution with increased sensitivity with respect to the turbid water flag. Equivalent distributions are obtained for PCI calculated with $R_{RS}(\lambda)$ or $R_{RC}(\lambda)$, although the $R_{RS}(\lambda)$ Level-2 product flags allow to exclude pixels affected by clouds.

It is generally not possible to reliably distinguish between the enhanced scattering caused by cyanobacteria from that due to inorganic particles (see Kratzer et al. (2020) and Kratzer and Moore (2018)) unless applying dedicated techniques (e.g., Miller et al., 2006). In fact, high concentrations of Total Suspended Matter (TSM, including mineral constituents) are usually translated into high values of $R_{RS}(\lambda)$ and potentially of band-differences too. Nevertheless, PCI and $R_{RS}(560)$-$R_{RS}(620)$ appear to be less affected by the presence of TSM-rich turbid waters mostly occurring near the coastline. Still, away from these very coastal areas, the patterns displayed by band-differences are quite similar. This suggests that comparable mapping results could be obtained with satellite sensors not specifically benefitting from the 620-nm centre-wavelength (at least for the bloom conditions characterizing this study). The maps produced for this test case are available in the supplementary material (Fig. S8).

4. Discussion

Since the deployment of the new CE-318T 12-band marine instruments, the number of daily AERONET-OC measurements has strongly increased. For example, the number of Level 1.0 data at GDLT between June and August are almost 10 times higher. This allows a better description of the development of phenomena varying at the time scale of a few minutes and provides an increased number of matchups with satellite data. It also allows to identify spectral features that characterize different development stages (e.g., the local minima in the blue region observed in correspondence of intense cyanobacteria accumulation at surface). However, some considerations are necessary on the quality assessment of AERONET-OC Level 1.5 and Level 2.0 data for sites characterized by the presence of blooms. The exclusion criteria implemented in the current AERONET-OC processor aim at ensuring the highest quality to in situ data in view of best supporting the validation of satellite data products. In this work, however, the quality control criterion imposing $L_{WN}(412)$ lower than $L_{WN}(443)$ at coastal sites has been ignored for data acquired during blooms. In fact, in the presence of blooms, several measurements show $L_{WN}(412)$ and $L_{WN}(400)$ higher than $L_{WN}(443)$, which may lead to the exclusion of relevant measurement sequences (as observed for July 28th 2019 at HLT, and July 26th 2019 at GDLT). This remark is not meant to question the quality control criteria that must ensure the creation of robust and consistent validation datasets, but rather to point out the risk that the specific generalized criterion can reduce the possibility of rising to higher quality levels (i.e., Level 1.5 and Level 2.0) those data relevant for bloom investigations.

Ultimately, the availability of AERONET-OC data from CE-318T instruments, including spectral bands replicating all the main visible bands of OLCI, benefit of supplementary bands with respect to those available with the instruments deployed in the Baltic Sea up to 2017. These include a spectral band at 400 nm and one at 620 nm, which were expected to have particular relevance for the study of cyanobacteria blooms. The analysis led to the identification of basic indexes best applicable for cyanobacteria detection such as the spectral band-differences $R_{RS}(560)$-$R_{RS}(620)$ and $R_{RS}(560)$-$R_{RS}(665)$ for OLCI or the equivalent $R_{RS}(547)$-$R_{RS}(667)$ for MODIS, which exhibit promising performance on the detection and possibly quantification of cyanobacteria superficial blooms in the Baltic Sea. By relying on the turbid water flag to heuristically determine the presence of cyanobacteria, the index $R_{RS}(560)$-$R_{RS}(620)$ computed from AERONET-OC data exhibits values varying in the range of 0.0011–0.0083 sr^{-1} with a median of 0.0030 sr^{-1} when the flag is raised, and conversely in the range of 0.0003–0.0041 sr^{-1} with a median of 0.0015 sr^{-1} when not raised. The index $R_{RS}(560)$-$R_{RS}(665)$ shows values in the range of 0.0013–0.0102 sr^{-1} with a median of 0.0036 sr^{-1} in the presence of cyanobacteria and conversely in the range of 0.0007–0.0043 sr^{-1} with a median of 0.0018 sr^{-1} in their absence. The ambiguity raised by the partially overlapping $R_{RS}(560)$-$R_{RS}(665)$ values in the presence or absence of cyanobacteria, can only be addressed by matching in situ measurements of phytoplankton and pigments concentrations, also essential to determine a relationship linking concentrations and radiometric indexes. The use of these indexes, when compared to the turbid water flag relying on a threshold applied to a single spectral band value, would reduce the impact of biases affecting $R_{RS}(\lambda)$ such as those due to an inaccurate atmospheric correction. The impact of a poor atmospheric correction could also be minimized in the green-red spectral region by relying on $R_{RC}(\lambda)$. Preliminary results from this study, however, do not show any benefit from the use of $R_{RC}(\lambda)$ for both the PCI and MPH algorithms.

It is also mentioned that the use of near and shortwave infrared bands has shown some potential to detect cyanobacteria scum or other surface features (e.g., Gower and King, 2020; Hu et al., 2019; Hu et al., 2010, 2023), but this has admittedly not been explored here because the study focused on bands common to both AERONET-OC and satellite data. Still, the test performed with red-NIR-based indexes (MPH, *SS* (681), PC3) did not show promising results.

It is finally stressed that the results obtained in this work relates to blooming conditions typical of the Baltic Sea region. They may not be applicable to other regions characterized by different phytoplankton compositions and concentrations during cyanobacteria blooms and consequently different $R_{RS}(\lambda)$ spectral shapes. Examples are in inland waters such as Lake Erie and Lake Taihu or the Curonian Lagoon, where blooms are dominated by *Microcystis* (Paldavičienė et al., 2009; Qi et al., 2014; Wynne et al., 2008) characterized by a quite different absorption spectrum (see Fig. 1 in Wojtasiewicz and Stoń-Egiert (2016)). It is finally noted that band-differences and PCI alone would not allow to distinguish cyanobacteria from highly-scattering phytoplankton species such as coccolithophores, whose concentration is strongly correlated to these indexes as shown for $R_{RS}(560)$-$R_{RS}(665)$ by Mitchell et al. (2017). Nevertheless, coccolithophores are currently not a relevant species in the Baltic Sea.

5. Conclusions

This work presented an unprecedented dataset of $R_{RS}(\lambda)$ spectra acquired during cyanobacteria blooms in the Baltic Sea from the GDLT and HLT AERONET-OC sites between 2018 and 2020. It benefitted of the most recent AERONET-OC instrument series providing data at additional centre-wavelengths and at higher frequency with respect to those operated up to 2017, fundamental to identify rapid changes in $R_{RS}(\lambda)$ features. AERONET-OC data usually discarded by the quality assurance process, but useful to detect cyanobacteria attributes during bloom events, were also considered. The analysis of AERONET-OC data allowed to investigate the radiometric features and the temporal evolution characterizing cyanobacteria blooms in the Baltic Sea and additionally to evaluate the capability of satellite data products to reproduce these features. Specifically, the analysis of in situ AERONET-OC data led to the following conclusions:

a. In situ $R_{RS}(\lambda)$ in the Baltic Sea, as determined from time-series from the GDLT and HLT sites, show an almost regular increase across the spectrum as a function of the accumulation of cyanobacteria during summer blooms. The observed spectral changes are more pronounced at 560 nm, which corresponds to the maximum of the $R_{RS}(\lambda)$ spectra.
b. In situ $R_{RS}(\lambda)$ referring to bloom with respect to non-bloom occurrence, exhibit spectral features that can be augmented through $R_{RS}(\lambda)$ band-differences. When solely considering the AERONET-OC visible spectral bands, cyanobacteria appear to significantly affect the band-differences $R_{RS}(560)$-$R_{RS}(490)$, $R_{RS}(560)$-$R_{RS}(510)$, $R_{RS}(560)$-$R_{RS}(620)$, $R_{RS}(560)$-$R_{RS}(665)$, $R_{RS}(620)$-$R_{RS}(665)$ and the PCI band combination. Nevertheless, it is recognized that in situ

matching measurements of phytoplankton concentrations and $R_{RS}(\lambda)$ are needed to quantitatively define relationships between cyanobacteria concentrations and radiometric indexes.

The assessment of OLCI and MODIS radiometric products in the presence of cyanobacteria instead showed that:

c. $R_{RS}(\lambda)$ from satellite ocean colour sensors exhibit low accuracy more marked in the blue spectral regions, which largely limits the exploitation of blue bands data for quantitative applications. Additionally, some band-ratios specifically proposed for the identification of cyanobacteria exhibit low accuracy.
d. OLCI $R_{RS}(665)$ and the corresponding MODIS $R_{RS}(667)$ show large underestimation (up to −41% for OLCI), potentially impacting the effectiveness of the turbid water flag exclusively relying on their values.
e. Conversely, the outstanding agreement between satellite and in situ data shown by the spectral differences $R_{RS}(560)$-$R_{RS}(620)$ and $R_{RS}(560)$-$R_{RS}(665)$ for OLCI or the equivalent $R_{RS}(547)$-$R_{RS}(667)$ for MODIS, indicate that these band-differences could potentially be used for cyanobacteria bloom detection and for any related quantitative application. Still, caution is suggested in applying these features in sediment-rich coastal waters because the current data set did not allow to evaluate their efficacy in these specific conditions.

These conclusions are focused on bands common to AERONET-OC and ocean colour satellite sensors, but do not exclude the potential offered by other bands, for instance in the near-infrared, for cyanobacteria applications.

Finally, in situ measurements of phytoplankton biovolume allowed to qualitatively evaluate those relations between cyanobacteria concentration and band combinations potentially relevant for cyanobacteria-specific algorithms. A quite strong relation with biovolume concentrations was found for $R_{RS}(665)$, $R_{RS}(560)$-$R_{RS}(510)$, $R_{RS}(560)$-$R_{RS}(620)$, $R_{RS}(560)$-$R_{RS}(665)$, $R_{RS}(620)$-$R_{RS}(665)$ and the PCI band combination applied to OLCI, and for $R_{RS}(667)$ and $R_{RS}(547)$-$R_{RS}(667)$ applied to MODIS. Conversely, these bands and band combinations did not appear affected by the presence of other phytoplankton blooms in the Baltic Sea, which usually develop deeper in the water column and are not characterized by high backscattering properties. Further progress could be achieved by the concurrent collection of optical and biological field measurements in cyanobacteria bloom conditions.

CRediT authorship contribution statement

Ilaria Cazzaniga: Formal analysis, Software, Investigation, Validation, Visualization, Writing – original draft, Writing – review & editing. **Giuseppe Zibordi:** Investigation, Supervision, Writing – review & editing. **Frédéric Mélin:** Software, Supervision, 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.

Data availability

The data used in this work are already available on the websites of the various databases mentioned in the manuscript

Acknowledgments

The authors would like to thank the AERONET Team for processing and distributing the data from the Ocean Color component of the Aerosol Robotic Network. The authors would also like to thank NASA OBPG for granting access to the MODIS and VIIRS data, EUMETSAT for providing the OLCI data, and the Finnish Environment Institute's (SYKE) Geoinformatics Systems Group and the Geoinformatics Research Group for providing true colour images and Surface Algal Bloom products. Also, SwAM and SMHI are acknowledged for making accessible in situ measurements of *Chl-a* and phytoplankton biovolume. This work has received funding from the EMPIR programme (grant 19ENV07 for METEOC-4) co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme. The support provided by DG DEFIS, i.e., the European Commission Directorate-General for Defence, Industry and Space, and the Copernicus programme is also gratefully acknowledged. Finally, four anonymous Reviewers are acknowledged for their comments and suggestions leading to an improvement of the work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2023.113464.

References

Anttila, S., Fleming-Lehtinen, V., Attila, J., Junttila, S., Alasalmi, H., Hällfors, H., Kervinen, M., Koponen, S., 2018. A novel earth observation based ecological indicator for cyanobacterial blooms. Int. J. Appl. Earth Obs. Geoinf. 64, 145–155. https://doi.org/10.1016/j.jag.2017.09.007.

Banks, A.C., Mélin, F., 2015. An assessment of cloud masking schemes for satellite ocean colour data of marine optical extremes. Int. J. Remote Sens. 36, 797–821. https://doi.org/10.1080/01431161.2014.1001085.

Beltrán-Abaunza, J.M., Kratzer, S., Brockmann, C., 2014. Evaluation of MERIS products from Baltic Sea coastal waters rich in CDOM. Ocean Sci. 10, 377–396. https://doi.org/10.5194/os-10-377-2014.

Berthon, J.F., Zibordi, G., 2010. Optically black waters in the northern Baltic Sea. Geophys. Res. Lett. 37, 1–6. https://doi.org/10.1029/2010GL043227.

Carmichael, W.W., 1992. Cyanobacteria secondary metabolites—the cyanotoxins. J. Appl. Bacteriol. 72, 445–459. https://doi.org/10.1111/j.1365-2672.1992.tb01858.x.

Dash, P., Walker, N.D., Mishra, D.R., Hu, C., Pinckney, J.L., D'Sa, E.J., 2011. Estimation of cyanobacterial pigments in a freshwater lake using OCM satellite data. Remote Sens. Environ. 115, 3409–3423. https://doi.org/10.1016/j.rse.2011.08.004.

Dekker, A.G., 1993. Detection of optical water quality parameters for eutrophic waters by high resolution remote sensing. Vrije Universiteit Amsterdam.

Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M.-H., Féménias, P., Frerick, J., Goryl, P., Klein, U., Laur, H., Mavrocordatos, C., Nieke, J., Rebhan, H., Seitz, B., Stroede, J., Sciarra, R., 2012. The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission. Remote Sens. Environ. 120, 37–57. https://doi.org/10.1016/j.rse.2011.07.024.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's optical high-resolution Mission for GMES operational services. Remote Sens. Environ. 120, 25–36. https://doi.org/10.1016/j.rse.2011.11.026.

EUMETSAT, 2021. Sentinel-3 Product Notice – OLCI Level-2 Ocean Colour - S3.PN-OLCI-L2M.003.00.

Gower, J., King, S., 2020. The distribution of pelagic Sargassum observed with OLCI. Int. J. Remote Sens. 41, 5669–5679. https://doi.org/10.1080/01431161.2019.1658240.

Hajdu, S., Höglander, H., Larsson, U., 2007. Phytoplankton vertical distributions and composition in Baltic Sea cyanobacterial blooms. Harmful Algae 6, 189–205. https://doi.org/10.1016/j.hal.2006.07.006.

Huisman, J., Codd, G.A., Paerl, H.W., Ibelings, B.W., Verspagen, J.M.H., Visser, P.M., 2018. Cyanobacterial blooms. Nat. Rev. Microbiol. 16, 471–483. https://doi.org/10.1038/s41579-018-0040-1.

Hu, C., Lee, Z., Ma, R., Yu, K., Li, D., Shang, S., 2010. Moderate resolution imaging spectroradiometer (MODIS) observations of cyanobacteria blooms in taihu Lake, China. J. Geophys. Res. Ocean. 115 https://doi.org/10.1029/2009JC005511.

Hu, L., Zeng, K., Hu, C., He, M.X., 2019. On the remote estimation of Ulva prolifera areal coverage and biomass. Remote Sens. Environ. 223, 194–207. https://doi.org/10.1016/j.rse.2019.01.014.

Hu, C., Qi, L., English, D.C., Wang, M., Mikelsons, K., Barnes, B.B., Pawlik, M.M., Fick, D., 2023. Pollen in the Baltic Sea as viewed from space. Remote Sens. Environ. 284, 113337 https://doi.org/10.1016/j.rse.2022.113337.

Kahru, M., Elmgren, R., 2014. Multidecadal time series of satellite-detected accumulations of cyanobacteria in the Baltic Sea. Biogeosciences 11, 3619–3633. https://doi.org/10.5194/bg-11-3619-2014.

Kahru, M., Savchuk, O.P., Elmgren, R., 2007. Satellite measurements of cyanobacterial bloom frequency in the Baltic Sea: interannual and spatial variability. Mar. Ecol. Prog. Ser. 343, 15–23. https://doi.org/10.3354/meps06943.

Karabashev, G.S., 2021. Spectral indexation of pixels of MODIS Sea surface images for detecting inconstancy of phytopigment composition in water. Oceanologia 63, 482–496. https://doi.org/10.1016/j.oceano.2021.06.001.
Kratzer, S., Kyryliuk, D., Brockmann, C., 2020. Inorganic suspended matter as an indicator of terrestrial influence in Baltic Sea coastal areas — algorithm development and validation, and ecological relevance. Remote Sens. Environ. 237, 111609 https://doi.org/10.1016/j.rse.2019.111609.
Kratzer, S., Moore, G., 2018. Inherent optical properties of the Baltic Sea in comparison to other seas and oceans. Remote Sens. 10, 418. https://doi.org/10.3390/rs10030418.
Kratzer, S., Tett, P., 2009. Using bio-optics to investigate the extent of coastal waters: a Swedish case study. Hydrobiologia 629, 169–186. https://doi.org/10.1007/s10750-009-9769-x.
Kratzer, S., Vinterhav, C., 2010. Improvement of MERIS level 2 products in Baltic Sea coastal areas by applying the improved contrast between ocean and land processor (ICOL) - data analysis and validation. Oceanologia 52, 211–236. https://doi.org/10.5697/oc.52-2.211.
Kutser, T., 2004. Quantitative detection of chlorophyll in cyanobacterial blooms by satellite remote sensing. Limnol. Oceanogr. 49, 2179–2189. https://doi.org/10.4319/lo.2004.49.6.2179.
Kutser, T., Metsamaa, L., Dekker, A.G., 2008. Influence of the vertical distribution of cyanobacteria in the water column on the remote sensing signal. Estuar. Coast. Shelf Sci. 78, 649–654. https://doi.org/10.1016/j.ecss.2008.02.024.
Kutser, T., Metsamaa, L., Strömbeck, N., Vahtmäe, E., 2006. Monitoring cyanobacterial blooms by satellite remote sensing. Estuar. Coast. Shelf Sci. 67, 303–312. https://doi.org/10.1016/j.ecss.2005.11.024.
Kutser, T., Metsamaa, L., Vahtmae, E., Stroembeck, N., 2007. Suitability of MODIS 250 m resolution band data for quantitative mapping of cyanobacterial blooms. Proc. Est. Acad. Sci. Biol. Ecol. 55, 318–328.
Kutser, T., Soomets, T., Toming, K., Uiboupin, R., Arikas, A., Vahter, K., Paavel, B., 2018. In: Assessing the Baltic Sea Water Quality with Sentinel-3 OLCI Imagery, in: 2018 IEEE/OES Baltic International Symposium, BALTIC 2018. IEEE, pp. 1–6. https://doi.org/10.1109/BALTIC.2018.8634849.
Matthews, M.W., Bernard, S., Robertson, L., 2012. An algorithm for detecting trophic status (chlorophyll-a), cyanobacterial-dominance, surface scums and floating vegetation in inland and coastal waters. Remote Sens. Environ. 124, 637–652. https://doi.org/10.1016/j.rse.2012.05.032.
Matthews, M.W., Odermatt, D., 2015. Improved algorithm for routine monitoring of cyanobacteria and eutrophication in inland and near-coastal waters. Remote Sens. Environ. 156, 374–382. https://doi.org/10.1016/j.rse.2014.10.010.
Mélin, F., 2022. Validation of ocean color remote sensing reflectance data: analysis of results at European coastal sites. Remote Sens. Environ. 280, 113153 https://doi.org/10.1016/j.rse.2022.113153.
Mélin, F., Zibordi, G., Berthon, J.-F., Bailey, S., Franz, B., Voss, K., Flora, S., Grant, M., 2011. Assessment of MERIS reflectance data as processed with SeaDAS over the European seas. Opt. Express 19, 25657. https://doi.org/10.1364/OE.19.025657.
Metsamaa, L., Kutser, T., Strömbeck, N., 2006. Recognising cyanobacterial blooms based on their optical signature: a modeling study. Boreal Environ. Res. 11, 493–506.
Miller, P., Shutler, J., Moore, G., Groom, S., 2006. SeaWiFS discrimination of harmful algal bloom evolution. Int. J. Remote Sens. 27, 2287–2301. https://doi.org/10.1080/01431160500396816.
Mishra, S., Mishra, D.R., 2014. A novel remote sensing algorithm to quantify phycocyanin in cyanobacterial algal blooms. Environ. Res. Lett. 9 https://doi.org/10.1088/1748-9326/9/11/114003.
Mitchell, C., Hu, C., Bowler, B., Drapeau, D., Balch, W.M., 2017. Estimating particulate inorganic carbon concentrations of the Global Ocean from ocean color measurements using a reflectance difference approach. J. Geophys. Res. Ocean. 122, 8707–8720. https://doi.org/10.1002/2017JC013146.
Mobley, C.D., Werdell, J., Franz, B., Ahmad, Z., Bailey, S., 2016. Atmospheric Correction for Satellite Ocean Color Radiometry.
Morel, A., Antoine, D., Gentili, B., 2002. Bidirectional reflectance of oceanic waters: accounting for Raman emission and varying particle scattering phase function. Appl. Opt. 41, 6289–6306. https://doi.org/10.1364/AO.41.006289.
Morel, A., Huot, Y., Gentili, B., Werdell, P.J., Hooker, S.B., Franz, B.A., 2007. Examining the consistency of products derived from various ocean color sensors in open ocean (Case 1) waters in the perspective of a multi-sensor approach. Remote Sens. Environ. 111, 69–88. https://doi.org/10.1016/j.rse.2007.03.012.
Munkes, B., Löptien, U., Dietze, H., 2021. Cyanobacteria blooms in the Baltic Sea: a review of models and facts. Biogeosciences 18, 2347–2378. https://doi.org/10.5194/bg-18-2347-2021.
NASA Goddard Space Flight Center Ocean Biology Processing Group, Ocean Ecology Laboratory, 2018. Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Ocean Color Data; 2018a Reprocessing. NASA OB.DAAC, Greenbelt, MD, USA data/10.5067/AQUA/MODIS/L2/OC/2018. Last accessed on 03/16/2021.
NASA Goddard Space Flight Center Ocean Biology Processing Group, Ocean Ecology Laboratory, 2018. Moderate-resolution Imaging Spectroradiometer (MODIS) Terra Ocean Color Data; 2018b Reprocessing. NASA OB.DAAC, Greenbelt, MD, USA data/10.5067/TERRA/MODIS/L2/OC/2018. Last accessed on 03/16/2021.
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. Ocean. 103, 24937–24953. https://doi.org/10.1029/98JC02160.
Paerl, H.W., Paul, V.J., 2012. Climate change: links to global expansion of harmful cyanobacteria. Water Res. 46, 1349–1363. https://doi.org/10.1016/j.watres.2011.08.002.
Paldavičienė, A., Mazur-Marzec, H., Razinkovas, A., 2009. Toxic cyanobacteria blooms in the Lithuanian part of the Curonian Lagoon. OCEANOLOGIA 51, 203–216. https://doi.org/10.5697/oc.51-2.203.
Qi, L., Hu, C., Duan, H., Cannizzaro, J., Ma, R., 2014. A novel MERIS algorithm to derive cyanobacterial phycocyanin pigment concentrations in a eutrophic lake: theoretical basis and practical considerations. Remote Sens. Environ. 154, 298–317. https://doi.org/10.1016/j.rse.2014.08.026.
Reinart, A., Kutser, T., 2006. Comparison of different satellite sensors in detecting cyanobacterial bloom events in the Baltic Sea. Remote Sens. Environ. 102, 74–85. https://doi.org/10.1016/j.rse.2006.02.013.
Riha, S., Krawczyk, H., 2011. Development of a remote sensing algorithm for cyanobacterial phycocyanin pigment in the Baltic Sea using neural network approach. In: Jr., C.R.B., Mertikas, S.P., Neyt, X., Velez-Reyes, M. (Eds.), Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions 2011. SPIE, pp. 42–48. https://doi.org/10.1117/12.898081.
Schalles, J.F., Yacobi, Y.Z., 2000. Remote detection and seasonal patterns of phycocyanin, carotenoid and chlorophyll pigments in eutrophic waters. Arch. Hydrobiol. Spec. Issues Advanc. Limnol. 55, 153–168.
Simis, S.G.H., Peters, S.W.M., Gons, H.J., 2005. Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. Limnol. Oceanogr. 50, 237–245. https://doi.org/10.4319/lo.2005.50.1.0237.
Stal, L.J., Albertano, P., Bergman, B., Von Bröckel, K., Gallon, J.R., Hayes, P.K., Sivonen, K., Walsby, A.E., 2003. BASIC: Baltic Sea cyanobacteria. An investigation of the structure and dynamics of water blooms of cyanobacteria in the Baltic Sea - responses to a changing environment. Cont. Shelf Res. 23, 1695–1714. https://doi.org/10.1016/j.csr.2003.06.001.
SYKE, 2021. Daily surface algal blooms of the Baltic Sea (Sentinel-2 MSI, Landsat-8 OLI, Sentinel-3 OLCI) - 2017. Contains modified Copernicus data and data from USGS / NASA Landsat Program (2021). Accessed using TARKKA web application (https://syke.fi/tarkka/en) Licensed under CC BY 4.0 International.
Thuillier, G., Hersé, M., Labs, D., Foujols, T., Peetermans, W., Gillotay, D., Simon, P.C., Mandel, H., 2003. The solar spectral irradiance from 200 to 2400 nm as measured by the SOLSPEC spectrometer from the Atlas and Eureca Missions. Sol. Phys. 214, 1–22. https://doi.org/10.1023/A:1024048429145.
Toming, K., Kutser, T., Uiboupin, R., Arikas, A., Vahter, K., Paavel, B., 2017. Mapping water quality parameters with Sentinel-3 ocean and land colour instrument imagery in the Baltic Sea. Remote Sens. 9 https://doi.org/10.3390/rs9101070.
Wojtasiewicz, B., Stoń-Egiert, J., 2016. Bio-optical characterization of selected cyanobacteria strains present in marine and freshwater ecosystems. J. Appl. Phycol. 28, 2299–2314. https://doi.org/10.1007/s10811-015-0774-3.
Woźniak, M., Bradtke, K.M., Darecki, M., Kręzel, A., 2016. Empirical model for phycocyanin concentration estimation as an indicator of cyanobacterial bloom in the optically complex coastalwaters of the Baltic Sea. Remote Sens. 8 https://doi.org/10.3390/rs8030212.
Wynne, T.T., Stumpf, R.P., Briggs, T.O., 2013. Comparing MODIS and MERIS spectral shapes for cyanobacterial bloom detection. Int. J. Remote Sens. 34, 6668–6678. https://doi.org/10.1080/01431161.2013.804228.
Wynne, T.T., Stumpf, R.P., Tomlinson, M.C., Warner, R.A., Tester, P.A., Dyble, J., Fahnenstiel, G.L., 2008. Relating spectral shape to cyanobacterial blooms in the laurentian Great Lakes. Int. J. Remote Sens. 29, 3665–3672. https://doi.org/10.1080/01431160802007640.
Zibordi, G., Berthon, J.-F., Mélin, F., D'Alimonte, D., 2011. Cross-site consistent in situ measurements for satellite ocean color applications: the BiOMaP radiometric dataset. Remote Sens. Environ. 115, 2104–2115. https://doi.org/10.1016/j.rse.2011.04.013.
Zibordi, G., Berthon, J.F., Mélin, F., D'Alimonte, D., Kaitala, S., 2009a. Validation of satellite ocean color primary products at optically complex coastal sites: northern Adriatic Sea, northern Baltic proper and gulf of Finland. Remote Sens. Environ. 113, 2574–2591. https://doi.org/10.1016/j.rse.2009.07.013.
Zibordi, G., Holben, B.N., Talone, M., D'Alimonte, D., Slutsker, I., Giles, D.M., Sorokin, M.G., D'Alimonte, D., Slutsker, I., Giles, D.M., Sorokin, M.G., D'Alimonte, D., Slutsker, I., Giles, D.M., Sorokin, M.G., D'Alimonte, D., Slutsker, I., Giles, D.M., Sorokin, M.G., 2021. Advances in the ocean color component of the aerosol robotic network (AERONET-OC). J. Atmos. Ocean. Technol. 38, 725–746. https://doi.org/10.1175/jtech-d-20-0085.1.
Zibordi, G., Kwiatkowska, E., Mélin, F., Talone, M., Cazzaniga, I., Dessailly, D., Gossn, J. I., 2022. Assessment of OLCI-A and OLCI-B radiometric data products across European seas. Remote Sens. Environ. 272, 112911 https://doi.org/10.1016/j.rse.2022.112911.
Zibordi, G., Mélin, F., Berthon, J.-F.F., Holben, B., Slutsker, I., Giles, D., D'Alimonte, D., Vandemark, D., Feng, H., Schuster, G., Fabbri, B.E., Kaitala, S., Seppälä, J., D'alimonte, D., Mélin, F., Berthon, J.-F.F., Vandemark, D., Feng, H., Schuster, G., Fabbri, B.E., Kaitala, S., Seppälä, J., 2009b. AERONET-OC: a network for the validation of ocean color primary products. J. Atmos. Ocean. Technol. 26, 1634–1651. https://doi.org/10.1175/2009JTECHO654.1.
Zibordi, G., Strömbeck, N., Mélin, F., Berthon, J.F., 2006. Tower-based radiometric observations at a coastal site in the Baltic proper. Estuar. Coast. Shelf Sci. 69, 649–654. https://doi.org/10.1016/j.ecss.2006.05.022.
Zolfaghari, K., Pahlevan, N., Binding, C., Gurlin, D., Simis, S.G.H., Verdu, A.R., Li, L., Crawford, C.J., Vanderwoude, A., Errera, R., Zastepa, A., Duguay, C.R., 2022. Impact of spectral resolution on quantifying cyanobacteria in lakes and reservoirs: a machine-learning assessment. IEEE Trans. Geosci. Remote Sens. 60, 1–20. https://doi.org/10.1109/TGRS.2021.3114635.