MDPI

Article

Optimal Cyanobacterial Pigment Retrieval from Ocean Colour Sensors in a Highly Turbid, Optically Complex Lake

Caitlin A.L. Riddick [1,*], **Peter D. Hunter** [1], **José Antonio Domínguez Gómez** [2],
Victor Martinez-Vicente [3], **Mátyás Présing** [4], **Hajnalka Horváth** [4], **Attila W. Kovács** [4],
Lajos Vörös [4], **Eszter Zsigmond** [5] and **Andrew N. Tyler** [1]

1   Biological and Environmental Science, School of Natural Sciences, University of Stirling,
    Stirling FK9 4LA, UK
2   Crop Research Institute, 161 06 Prague, Czech Republic
3   Remote Sensing Group, Plymouth Marine Laboratory, Plymouth PL1 3DH, UK
4   Balaton Limnological Institute, MTA Centre for Ecological Research, 8237 Tihany, Hungary
5   Institute of Genetics, Biological Research Centre Hungarian Academy of Sciences, 6726 Szeged, Hungary
*   Correspondence: caitlin.riddick1@stir.ac.uk

check for updates

Abstract: To date, several algorithms for the retrieval of cyanobacterial phycocyanin (PC) from ocean colour sensors have been presented for inland waters, all of which claim to be robust models. To address this, we conducted a comprehensive comparison to identify the optimal algorithm for retrieval of PC concentrations in the highly optically complex waters of Lake Balaton (Hungary). MEdium Resolution Imaging Spectrometer (MERIS) top-of-atmosphere radiances were first atmospherically corrected using the Self-Contained Atmospheric Parameters Estimation for MERIS data v.B2 (SCAPE-M_B2). Overall, the Simis05 semi-analytical algorithm outperformed more complex inversion algorithms, providing accurate estimates of PC up to ±7 days from the time of satellite overpass during summer cyanobacteria blooms ($RMSE_{log} < 0.33$). Same-day retrieval of PC also showed good agreement with cyanobacteria biomass ($R^2 > 0.66$, $p < 0.001$). In-depth analysis of the Simis05 algorithm using in situ measurements of inherent optical properties (IOPs) revealed that the Simis05 model overestimated the phytoplankton absorption coefficient [$a_{ph}(\lambda)$] by a factor of ~2. However, these errors were compensated for by underestimation of the mass-specific chlorophyll absorption coefficient [$a^*_{chla}(\lambda)$]. This study reinforces the need for further validation of algorithms over a range of optical water types in the context of the recently launched Ocean Land Colour Instrument (OLCI) onboard Sentinel-3.

Keywords: cyanobacteria; phycocyanin; MERIS; Sentinel-3; remote sensing; Lake Balaton

1. Introduction

It has recently been estimated that there are as many as 117 million lakes on Earth covering approximately 3.7% of the planet's non-glaciated land surface [1]. While they comprise only a small fraction of the Earth's land surface, inland water bodies play a fundamental role in many global and regional biogeochemical processes [2,3]. Lakes are also highly sensitive to environmental perturbation and change impacting their airsheds and watersheds. For example, inputs of nutrients derived from anthropogenic sources to lakes have increased in recent years, with eutrophication recognised as one of the most universally widespread ecological, economic and social issues affecting the quality of freshwaters globally [4]. Eutrophication of lakes is acknowledged as a major global ecological concern, particularly in developing countries where annual phosphorus (P) loads are estimated to be ten-fold larger than in developed countries and the risk of P-stimulated eutrophication is higher [5].

Nutrient-enriched lakes have phytoplankton communities that are often dominated by blue-green algae or cyanobacteria, which are notorious bloom-forming prokaryotes. Blooms of cyanobacteria can have profound and often highly adverse impacts on lake ecosystems [6]. Some species of cyanobacteria can fix dissolved dinitrogen gas into organic nitrogen, thus allowing them to outcompete other phytoplankton species and thrive in conditions with a low nitrogen to phosphorus ratio [7–12]. Other adaptations such as the ability to store excess nutrients like phosphorus [13], low light requirements [14], increased growth rates at higher temperatures [15] and buoyancy regulation mechanisms [16] further allow cyanobacteria to prosper in warmer, nutrient enriched waters.

Cyanobacteria can also pose significant risks to animal and human health, as many species produce cyanotoxins with neurotoxic, hepatotoxic, cytotoxic, genotoxic, endotoxin and tumor-promoting properties [17]. Given the health risks they pose, it is vital that methods are developed for the accurate and rapid assessment and monitoring of cyanobacteria blooms in lakes. Traditional water sampling methods are often shore-based and lack sufficient spatial coverage to capture the heterogeneity of cyanobacteria populations. Furthermore, laboratory methods for enumeration and identification of cyanobacteria taxa are time-consuming and costly. In contrast, earth-observing satellites can provide data at a spatial and temporal resolution to permit rapid detection and monitoring of cyanobacteria populations in lakes on an operational basis. Numerous algorithms for retrieval of cyanobacterial pigments from inland waters have been presented to date, however there are no comprehensive validation studies to identify the most robust model(s). To this end, this study aims to assess the capability of semi-empirical, semi-analytical and inversion algorithms for retrieval of cyanobacterial pigments from inland waters using data from the ocean colour sensor MERIS (MEdium Resolution Imaging Spectrometer that was flown on the Envisat satellite) captured over the highly turbid, optically complex waters of Lake Balaton, Hungary. The ultimate aim of this research is to rigorously validate algorithms for future use in near real-time operational detection and monitoring of cyanobacterial blooms in lakes.

Remote sensing is used operationally for monitoring phytoplankton in the global ocean, but remote sensing of inland waters has not progressed as rapidly due to the greater complexity in the atmospheric and in-water optical properties of lakes [18]. Chlorophyll-*a* (Chl-*a*) can be retrieved with algorithms from remote sensing measurements as an indicator of total phytoplankton biomass [19–21]. However, Chl-*a* does not provide information about the phytoplankton community composition or reliably indicate the presence of potentially toxic cyanobacteria blooms. Cyanobacteria-specific information can be acquired by also estimating the concentration of phycocyanin (PC), an indicator pigment for cyanobacteria [22–29]. The unique optical properties of phycocyanin-containing cyanobacteria allow them to be distinguished from other phytoplankton using the shape and magnitude of the remote-sensing reflectance ($R_{rs}(0+,\lambda)$) signal observed during blooms [30].

PC can be most simply estimated from $R_{rs}(0+,\lambda)$ using semi-empirical algorithms incorporating ($R_{rs}(0+,\lambda)$) band-ratios or band-differences which target the main absorption feature at ~620 nm. The earliest approaches for PC estimation from MERIS include the semi-empirical baseline algorithm by Dekker [31] and the band ratio approach introduced by Schalles and Yacobi [32], hereafter referred to as Dekker93 and Schalles00, respectively. The semi-analytical algorithm for PC estimation was subsequently developed by Simis et al. [28] (hereafter referred to as Simis05) adapted from the related algorithm for Chl-*a* retrieval by Gons [33] and Gons et al. [19,34] (hereafter referred to as Gons05). Both Simis05 and Gons05 were specifically tailored to the bands for MERIS and have been validated in a range of inland waters [22,23,35–38]. More recently, other algorithms for PC retrieval compatible with MERIS and OLCI have been developed and published, including the adapted quasi-analytical algorithm (QAA) developed by Mishra et al. [39], the PC Index (PCI) by Qi et al. [40], the IOP Inversion Model of Inland Waters (IIMIW) by Li et al. [41], and the 4-band semi-analytical model by Liu et al. [42] that builds upon the 3-band model developed by Hunter et al. [23], hereafter referred to as Mishra13, Qi14, Li15, Liu18, and Hunter10, respectively. These algorithms were identified as the most relevant PC retrieval methods presently available for MERIS and OLCI data and were thus selected for comparison

in this study. Literature reviews of phycocyanin retrieval algorithms and approaches can be found in Matthews et al. [43], Li and Song et al. [44] and Yan et al. [45].

Despite the growing portfolio of PC retrieval algorithms, the practical use of some of the algorithms developed is limited by the spectral coverage and resolution available with current satellite instruments [27,45]. Although the MERIS sensor is no longer operational, the data archive remains immensely useful for algorithm development and validation studies, particularly because the OLCI instrument on the European Space Agency's recently launched Sentinel-3 satellite has a strong MERIS heritage. Indeed, OLCI is being implemented for the monitoring of water quality parameters in large lakes (e.g., Lake Balaton [46]) and is considered the only operational satellite sensor in orbit applicable for measurement of PC [45]. Importantly, MERIS and its successor OLCI have a waveband centered at 620 nm near the PC absorption maximum, therefore there is great value in the MERIS archive with regard to robust validation of existing PC algorithms over inland waters.

Algorithms tend to be developed for use in specific water bodies and there is a real need to assess whether they are transferable to other lakes with differing optical and biogeochemical properties, and further understanding is required of the uncertainties over the full range of optical water types (OWTs) [18,47,48]. Many validation studies have used in situ measurements of subsurface reflectance, however to date there have simply been few attempts to validate PC algorithms over inland waters using satellite data such as MERIS or OLCI (Sentinel-3). Furthermore, the published PC algorithms for inland waters all claim to provide optimal performance, however no comprehensive validation of all PC algorithms has been undertaken to date. For instance, previous validation studies have focused on performance of a limited set of algorithms (e.g., reflectance-based empirical approaches only [49] or three empirical, semi-empirical and semi-analytical approaches [27]). These studies have tested PC models using in situ [22,27,49] or airborne reflectances [50] rather than satellite-derived reflectances. To address this deficiency in PC algorithm validation, this study aims to comprehensively test existing PC algorithms for ocean colour sensors using a MERIS and in situ dataset on Lake Balaton (Hungary), a site with recurrent summer cyanobacteria blooms, a gradient of phytoplankton biomass and CDOM, and optical properties that are highly influenced by inorganic particulates [51,52]. More specifically, this study aims to firstly test a series of algorithms for PC retrieval from MERIS data over Lake Balaton, comparing retrieval of PC from MERIS data within 1 day of in situ measurements. The best-performing algorithm is investigated in greater detail over a range of matchup windows, and retrieved pigments are validated with a time series of pigment and cell count data from 2007–2011 (Chl-*a*) and 2010–2011 (PC). Finally, the sources of error are examined by validating retrieval of absorption and backscattering coefficients with a dataset of in situ and laboratory inherent optical property (IOP) measurements from August 2010.

2. Materials and Methods

2.1. Study Site

Lake Balaton (46.8°N, 17.7°E; Figure 1) is the largest shallow lake in Europe by surface area (592 km^2), with a mean depth of just 3.2 m. It has a large catchment, dominated by the Zala River, and a history of eutrophication and summer cyanobacteria blooms. The lake itself typically has a gradient in phytoplankton biomass and Chl-*a*, with the highest Chl-*a* concentrations occurring in the western Keszthely basin (~3–45 mg m^{-3}), and lower phytoplankton biomass and Chl-*a* in the eastern Siófok basin (~3–20 mg m^{-3}). Cyanobacteria biomass tends to peak in late summer (anywhere from June-October), with PC values up to ~60–100 mg m^{-3} in the westernmost basins [9]. In recent years, summer cyanobacteria populations in Lake Balaton are dominated by nitrogen-fixing species, including *Raphidiopsis raciborskii* (Wołosz.) (former genus name *Cylindrospermopsis*), *Aphanizomenon flos-aquae* ((L.) Ralfs), *Cuspidothrix issatschenkoi* (Usachev) (former genus *Aphanizomenon*), *Sphaerospermopsis aphanizomenoides* (Forti) (former genus name *Aphanizomenon*) and *Dolichospermum spiroides* (Kleb.)

(former genus name *Anabaena*). During a bloom, cyanobacteria often contribute to >70% of the total phytoplankton biomass [9].

Despite the frequently high concentrations of phytoplankton that occur in Lake Balaton during the summer, light attenuation is largely determined by the frequent wind-driven resuspension of mineral particles from the lake bottom. Total suspended matter in the lake is mostly of inorganic origin, with concentrations typically in the order of 18–28 mg L^{-1} (2010 annual mean), but can exceed 100 mg L^{-1} during windy periods. These high loads of mineral particles contribute significantly to light absorption and scattering; absorption by non-algal particles ($a_{NAP}(440)$) is typically 0.2 m^{-1} and particulate backscattering ($b_{bp}(532)$) varies between 0.04 and 0.2 m^{-1} [51]. These fine particles have a high backscattering efficiency ($b_{bp}:b_p$ up to 0.03) and thus contribute strongly to the water-leaving radiative signal and impart the lake with its characteristic turquoise colour.

Lake Balaton also demonstrates highly localised concentrations of coloured dissolved organic matter (CDOM). CDOM absorption coefficients ($a_{CDOM}(440)$) typically ranging from 0.09 to 1.4 m^{-1}, with the highest CDOM absorption observed at the mouth of the Zala River ($a_{CDOM}(440)$ up to 9.5 m^{-1}) where water rich in dissolved organic carbon produced in the reservoir system Kis-Balaton discharges into the western portion of the lake [51]. However, CDOM is rapidly diluted and bleached through photodegradation as water passes through the system [52].
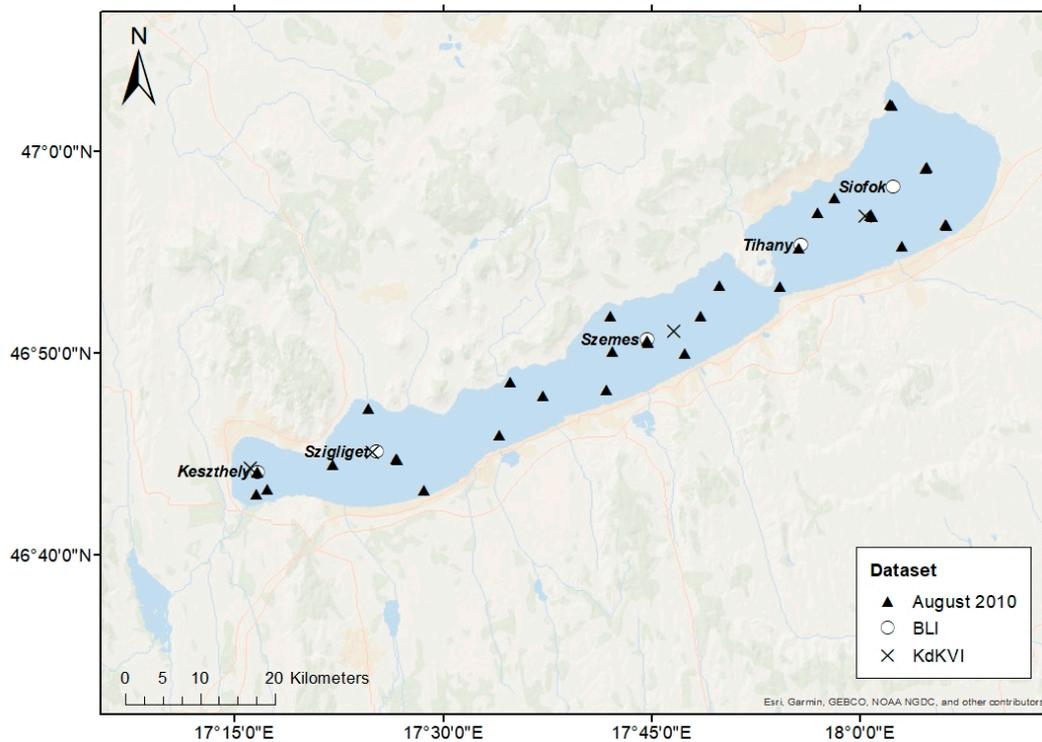

Figure 1. Map of Lake Balaton, indicating the routine monitoring stations (Balaton Limnological Institute, BLI; Central Transdanubian (Regional) Inspectorate for Environmental Protection, KdKVI), and the 35 stations from the August 2010 field campaign (August 2010).

2.2. Validation Datasets

Lake Balaton is regularly monitored by Balaton Limnological Institute (BLI, MTA CER) at a bi-weekly to monthly frequency at up to 5 stations across the lake (Figure 1), thus it has an extensive archive of data for satellite validation studies. In this study, we compiled the routine Chl-*a* concentration and phytoplankton count data from the years 2007–2011 and more recent measurements of the PC concentrations from the years 2010–2011. This dataset is referred to as "BLI" throughout.

Further Chl-*a* data from 2007–2011 were obtained from the Central Transdanubian (Regional) Inspectorate for Environmental Protection, Nature Conservation and Water Management (Közép-dunántúli Környezetvédelmi, Természetvédelmi és Vízügyi Felügyelőség (KDT KTVF)). However, for consistency with previous publications, the abbreviation "KdKVI" is used throughout for this Chl-*a* dataset. These data were collected at 4 stations at the center of the respective basins (Figure 1).

In addition to the routine monitoring programmes, a separate sampling campaign for MERIS validation was conducted from 18 to 26 August 2010 to coincide with an Envisat MERIS overpass on 22 August 2010. IOP measurements and water samples were collected at 35 stations during this campaign (Figure 1). In situ radiometry data were collected at 30 stations only, as the raw data files from stations 31–35 were corrupted. Large volume water samples (5 L) were taken from the surface (~0.5 m) using an acid-rinsed wide-necked polyethylene carboy for subsequent analysis of pigments (Chl-*a* and PC), phytoplankton absorption and phytoplankton counts. Samples were stored on ice in the dark prior to analysis for pigments. This dataset is referred to as "August 2010" throughout.

2.2.1. Chlorophyll-*a*

BLI routine water samples were collected as depth integrated over the first 2–3.5 m of the water column, depending on the maximum sample depth at each station. A 5-L sample was stored in the dark on ice before Chl-*a* extraction, within 24 h. Samples were filtered under low vacuum pressure through GF/C (Whatman) filter papers and subsequently extracted in 90% hot methanol for 1 min. Depending on the sample turbidity, between 500 and 1500 mL was filtered. Sample absorbance was measured spectrophotometrically (Shimadzu UV-1601), as in Iwamura et al. [53]. The concentration (mg m^{-3}) of Chl-*a* was then determined by the following equation, where A_x is the measured absorbance at wavelength x (nm):

$$\text{Chl}-a\left(\text{mg m}^{-3}\right) \ = \ 17.12(A_{666} - A_{750}) - 8.68(A_{653} - A_{750}). \tag{1}$$

KdKVI routine monitoring consists of water samples collected at the surface, followed by filtration under low vacuum pressure through GF/C (Whatman) filter papers. Chl-*a* was extracted in ethanol and measured spectrophotometrically, with concentration calculated as in Equation (1). To investigate any systematic difference between the KdKVI and BLI archive datasets, coincident Chl-*a* data have been compared between BLI depth integrated samples and KdKVI surface samples. Results indicated no significant difference between the medians of the two datasets (non-parametric Mann-Whitney U-test, n = 7, $p = 0.710$) [54].

For the August 2010 sampling campaign, a subsample of a 5 L surface water (0.5 m) sample was filtered on the boat immediately after sample collection under low vacuum pressure through GF/F (Whatman) filter papers. Depending on the water clarity at the location (using Secchi depth), between 20 and 70 mL of sample water was filtered. Filter papers were then flash frozen in liquid nitrogen for <12 h and placed in a −80 °C freezer until analysis (no more than 6 months). Frozen GF/F filter papers were thawed in the dark from −80 °C and chlorophyll-*a* was extracted in 90% hot methanol and measured spectrophotometrically (Shimadzu UV-1601), as for the BLI routine monitoring programme. For all dataset methods, there was no correction included for absorption by phaeopigments, a degradation product of Chl-*a*.

2.2.2. Phycocyanin

The samples collected for BLI routine monitoring were depth integrated over the first 2–3.5 m of the water column, depending on the maximum sample depth at each station. A 5-L sample was stored in the dark on ice before PC extraction, within 24 h. Samples were filtered under low vacuum pressure through GF/C (Whatman) filter papers and subsequently extracted in a solution of 15 mL 0.05M phosphate buffer (pH = 6.8). Depending on the sample turbidity, between 70 and 450 mL of water was

filtered. Filter papers in the buffer solution underwent one freeze-thaw cycle as in Sarada et al. [55], and phycocyanin was extracted by sonication over ice for 15 seconds (Ultrasonic Homogeniizer 4710 Series, Cole-Palmer Instrument Co., USA), as detailed in Horváth et al. [9] (Method E). Finally, extracts were filtered (GF/C Whatman) and the absorption measured on a spectrophotometer (Shimadzu UV-1601, Shimadzu Co., Japan). Phycocyanin concentrations were calculated using the following equation [56], where A_x is the measured absorbance at wavelength x (nm):

$$\text{PC} \left(\text{mg m}^{-3}\right) = \left(A_{615} - 0.474 \times A_{652}\right)/5.34. \tag{2}$$

For the August 2010 dataset, subsamples of a 5 L surface water (~0.5 m) sample were filtered on the boat immediately after collection under low vacuum pressure through GF/F (Whatman) filter papers. Depending on the water clarity at the location (using Secchi depth), between 20 and 70 mL of sample water was filtered. Filter papers were then flash frozen in liquid nitrogen for <12 h and placed in a −80 °C freezer until analysis (no more than 6 months). Frozen GF/F filter papers (2 replicates) were thawed from −80 °C to 9 °C (±1 °C) and the pigment concentration was obtained according to the method described above for the BLI dataset, using Equation (2). As only 2 replicates were analysed for PC concentration, standard error is not presented on these data, however there was generally good agreement between the replicates with good agreement between PC and cyanobacteria biomass for the August 2010 dataset (R^2 = 0.94, CV = 1.7%) [9].

2.2.3. Phytoplankton Biomass

Phytoplankton biomass data for BLI routine monitoring were depth integrated water samples, while data from the August 2010 campaign were collected as subsamples of a 5L surface water sample. All water samples for analysis of phytoplankton biomass were collected in 50 mL polyethylene containers and preserved in Lugol's solution immediately after collection for analysis within 6 months. At least 25 cells (or filaments) of each species were measured to determine biomass and at least 400 were counted using an inverted plankton microscope [57]. The wet weight of each species was then calculated from cell volumes [58].

2.2.4. Measurement of Absorption and Backscattering Coefficients

Measurements of particulate absorption and backscattering coefficients were made during the August 2010 campaign only. For particulate absorption, a subsample of a 5 L water sample was filtered on the boat immediately after sample collection under low vacuum pressure. Depending on the location, 20–70 mL of water was passed through GF/F (Whatman) filter papers and the filters were immediately flash frozen in liquid nitrogen for <12 h prior to storage at −80 °C. All samples were analysed within 6 months of collection. Frozen GF/F filters were defrosted in the lab from −80 °C in the dark. The absorbance of the material on the filter was measured from 350 to 750 nm using a dual beam spectrophotometer retro-fitted with Spectralon coated integrating spheres, according to the 'transmittance-reflectance' method of Tassan and Ferrari [59]. Absorption was measured before and after bleaching with a 1% solution of NaClO to obtain particulate absorption ($a_p(\lambda)$) and absorption by non-algal particles ($a_{NAP}(\lambda)$), respectively. The pathlength amplification correction of Tassan and Ferrari [59] was applied and absorption by phytoplankton ($a_{ph}(\lambda)$) was calculated as the difference between $a_p(\lambda)$ and $a_{NAP}(\lambda)$. Chlorophyll-specific absorption ($a^*_{Chla}(\lambda)$) coefficients were calculated by dividing $a_{ph}(\lambda)$ by the respective Chl-*a* concentration.

Particle backscattering coefficients ($b_{bp}(\lambda)$) were derived from measurements collected with an ECO-BB3 backscatter meter (WET Labs, Inc.). The calibration was performed by the manufacturer prior to sampling, and it was assumed that there was no drift in the calibration coefficients and/or dark counts. The ECO-BB3 measures the total volume scattering function ($\beta_t(\lambda,124°)$) from a centroid angle of scattering (124°) at 3 wavelengths (λ = 470, 532 and 650 nm). The transformation of the raw counts

into $b_b(\lambda)$ was done following the manufacturer's user guide [60] and the methods as in Slade and Boss [61].

Firstly, the instrument raw counts were converted into an uncorrected value of the volume scattering function, $(\beta_u(\lambda,124°))$:

$$\beta_u(\lambda, 124°) \ = \ s(\lambda)[raw(\lambda) - d(\lambda)], \tag{3}$$

where s and d are scaling and dark counts factors, respectively, with values supplied by the manufacturer. The scaling factors change with time at a typical rate of 10%, 4% and 3% per year for blue, green and red wavelengths, respectively [62]. As this sampling campaign was conducted over 10 days, the expected maximum change over time is about 0.3% of the signal, and laboratory tracking from 2013 to 2015 of the scaling factor on this ECO-BB3 corroborated this assumption (Martinez-Vicente, pers. comm.). Secondly, $\beta_u(\lambda,124°)$ was corrected to account for the light attenuation within the instrument path length using an attenuation factor, $K(\lambda)$:

$$K(\lambda) \ = \ exp\big(La_{pg}(\lambda)\big), \tag{4}$$

where L is the effective pathlength (0.015 m) and $a_{pg}(\lambda)$ is the absorption due to particulate and dissolved matter. $a_{pg}(\lambda)$ was simultaneously measured in situ by a AC-9 (WET Labs, Inc.), and these data were processed using the temperature *proportional*$_c^e$ scattering correction from Röttgers et al. [63] and pure water calibrations measured before and after the campaign. In each case, the nearest wavelengths to the ECO-BB3 wavelengths were applied (i.e., 488, 532 and 660). $\beta_t(\lambda,124°)$ is then calculated as:

$$\beta_t(\lambda, 124°) \ = \ \beta_u(\lambda, 124°) \times K(\lambda). \tag{5}$$

Thirdly, the value of the volume scattering function of water was subtracted from $\beta_t(\lambda,124°)$ using the model by Zhang et al. [64] with fixed temperature and salinity values of 24.0 °C and 0.4 psu, respectively. Finally, the volume scattering function of the particles at 124° is extrapolated to the backwards direction to derive the particulate backscattering coefficient, $b_{bp}(\lambda)$, using a conversion factor of $\chi_p(124°) = 1.08$ as in Sullivan et al. [62]:

$$b_{bp}(\lambda) \ = \ 2\pi\chi_p(124°) \times [\beta_t(\lambda, 124°) - \beta_w(\lambda, 124°)]. \tag{6}$$

Mean $b_{bp}(\lambda)$ values were calculated over 1–5 min measurement periods. As CDOM is assumed to be non-scattering, $b_{bp}(\lambda)$ is referred to here as $b_b(\lambda)$.

2.2.5. In Situ Radiometry

In situ radiometry data were collected during the August 2010 campaign only at Stations 1–30 (n = 30), using a HyperSAS (Hyperspectral Surface Acquisition System; Satlantic) for validation of the atmospheric correction of MERIS data. Three radiometers were positioned at a height of 3.5 m from a pole at the bow of the boat. The boat was positioned on station to point radiance sensors at a relative azimuth angle of 135° from the sun, and the zenith angles of the sea- and sky-viewing radiance sensors were fixed at 40°. The three radiometers measured downwelling irradiance ($E_s(\lambda)$), surface radiance of the water ($L_t(\lambda)$) and sky radiance ($L_{sky}(\lambda)$). Measurements were made at each station for 30 min with an integration time of 21 seconds. The sensors measured over a wavelength range of 348.9–801.6 nm with a sampling interval of 3.3 nm (137 bands) and a spectral width of 10 nm.

$L_t(\lambda)$ and $L_{sky}(\lambda)$ were used to calculate the water-leaving radiance ($L_w(\lambda)$) after correction for air-sea interface reflection as follows [65,66]:

$$L_w(\lambda) \ = \ L_t(\lambda) - \rho_{sky}L_{sky}(\lambda), \tag{7}$$

where ρ_{sky}, the air–water interface reflection coefficient, is estimated for sunny conditions as a function of wind speed (W; m s^{-1}) (Ruddick et al., 2006):

$$\rho_{sky} \ = \ 0.0256 \ + \ 0.00039 \times W + 0.000034 \times W^2. \tag{8}$$

The HyperSAS raw data were processed to level 3a using Satlantic ProSoft software (v.7.7.10). After processing, all spectra underwent quality control for any outliers due to variable cloudiness or sun glint due to lake surface state. Remote sensing reflectance, $R_{rs}(\lambda)$ (sr^{-1}), was then calculated as follows:

$$R_{rs}(\lambda) \ = \ L_w(\lambda)/E_s(\lambda). \tag{9}$$

2.3. MERIS Data Processing

MERIS full resolution full swath (FRS) 300 m Level-1b data for Lake Balaton were obtained for a period of 5 years (2007–2011) from the European Space Agency's Merci system (https://earth.esa.int/web/guest/data-access/online-archives). The geolocation of the data was improved using the AMORGOS (Accurate MERIS Ortho-Rectified Geo-location Operational Software) v.4.0 processor. Images with high cloud cover were discarded, leaving 34 images with matching Chl-*a* data, 21 images with matching cell biomass data and 5 images with matching PC data (within 7 days of overpass date). The MERIS data were atmospherically corrected using the Self-Contained Atmospheric Parameters Estimation for MERIS data (SCAPE-M) automatic atmospheric correction processor developed by Guanter et al. [67], following the SCAPE-M_B2 implementation as in Domínguez Gómez et al. [68]. SCAPE-M_B2 is an improved version of SCAPE-M, which corrects MERIS band 2 with an interpolation between the values of band 1 and band 3 [68]. It is noted that this is the first implementation of the SCAPE-M_B2 correction to MERIS FRS data, as opposed to FR (full resolution) data, and that SCAPE-M_B2 will be adapted to OLCI data in the future. SCAPE-M and SCAPE-M_B2 have been shown to compute accurate water-leaving reflectances for lakes, particularly for highly turbid waters [69–73].

Following atmospheric correction with SCAPE-M_B2, water-leaving reflectance ($\rho_w(\lambda)$) was extracted from the MERIS images using BEAM VISAT v.4.11 (Brockmann Consult, v.4.11). $\rho_w(\lambda)$ was converted (denormalised) to remote sensing reflectance ($R_{rs}(\lambda)$) through division by pi (π). Each algorithm was then applied to the extracted $R_{rs}(\lambda)$, and the best performing algorithm was implemented using the Graph Processing Framework in BEAM to produce mapped time-series products.

2.3.1. Validation of Atmospheric Correction

MERIS $R_{rs}(\lambda)$ data for a 3×3 pixel kernel (i.e., 9 pixels) were extracted for the corresponding in situ sampling locations, and the mean and standard deviation for each kernel was calculated. The SCAPE-M_B2 atmospheric correction was validated using same-day (±1 day) matchups with in situ $R_{rs}(\lambda)$ only from the August 2010 dataset (n = 7), as this was considered a suitable approach for a validation over a dynamic lake system. However, all available matchups of MERIS SCAPE-M_B2 and in situ $R_{rs}(\lambda)$ are presented here to investigate the ability to use a broader temporal window for validation of the atmospheric correction (n = 30; ±4 days). All matchups were cloud-free for the August 2010 dataset, and no spatial homogeneity method was employed.

$R_{rs}(\lambda)$ was compared at all 15 MERIS bands with in situ $R_{rs}(\lambda)$ collected by HyperSAS radiometry. For this, HyperSAS $R_{rs}(\lambda)$ data were simulated to the MERIS spectral response using the band centres and full width half maximum (FWHM), assuming a Gaussian distribution for the spectral response function (SRF). At each station, $R_{rs}(\lambda)$ for each MERIS band was also plotted over the range of in situ $R_{rs}(\lambda)$ measured by the HyperSAS at the respective station (n = 30). This allowed for comparison of the spectral shape and magnitude of $R_{rs}(\lambda)$ between the MERIS SCAPE-M_B2 and HyperSAS data for each station.

2.3.2. Validation Matchup Data

MERIS $R_{rs}(\lambda)$ data for a 3×3 pixel kernel (i.e., 9 pixels) were extracted for the corresponding in situ sampling locations, and the mean, standard deviation and coefficient of variation (CV%) for each kernel was calculated, where CV is the ratio of the standard deviation to mean value expressed as a percent. To reduce the effects of spatial variability on the validation, any MERIS $R_{rs}(\lambda)$ or retrieved parameter with CV>20% was removed, similar to the methods for spatial homogeneity employed in Goyens et al. [74] and Jamet et al. [75]. This resulted in MERIS data which were generally discarded due to cloud cover or where there was interference from land pixels from near-shore sampling locations (adjacency effect). MERIS-retrieved parameters were then validated with in situ data at time windows of ± 1, 3 and 7 days from a MERIS overpass, with duplicate matchups discarded (i.e., where more than one sampling date corresponded to the same MERIS image for a particular station). All in situ $R_{rs}(\lambda)$ and IOP measurements for validation were part of the August 2010 dataset and thus collected ± 4 days of the MERIS overpass (22 August 2010). The in situ $R_{rs}(\lambda)$ data measured by HyperSAS radiometry are shown in Figure 2, and a summary of the resulting available pigment matchup data from each in situ dataset is shown in Table 1.
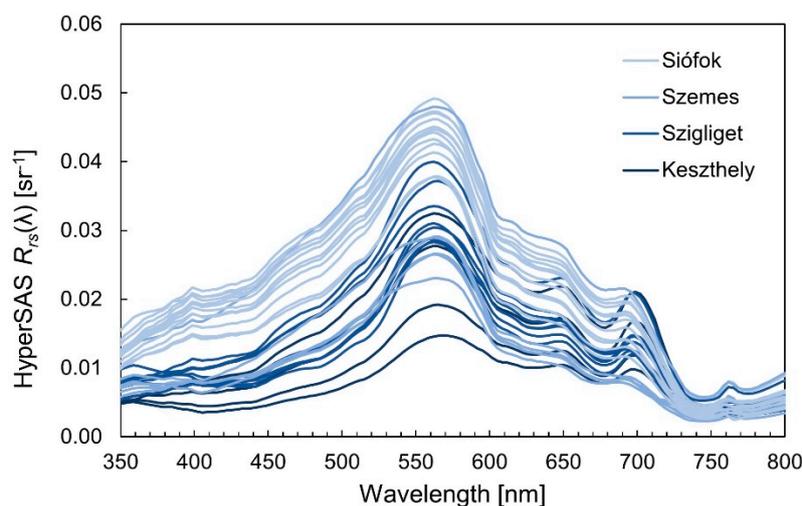

Figure 2. Mean in situ HyperSAS $R_{rs}(\lambda)$ spectra measured in August 2010 in Lake Balaton (n = 30), showing the variability by basin from west (Keszthely) to east (Siófok).

Table 1. Descriptive statistics of Chl-*a*, PC, phytoplankton biomass, cyanobacteria biomass, $a_{ph}(665)$, $a_{ph}(620)$ and $b_b(650)$ from each in situ dataset used for validation of MERIS-retrieved parameters. The sample size (n) is the resulting number of matchups after removal of duplicate matchups, cloud cover or interference from land pixels (i.e., standard deviation of a 3×3 pixel window value >2).

Matchup Window	Parameter	Dataset	n	Min	Max	Mean	St Dev	Units
±1 day	Chl-*a*	August 2010	13	8.31	34.4	19.1	9.60	mg m^{-3}
		BLI	18	2.43	33.8	13.4	9.39	mg m^{-3}
		KdKVI	105	1.50	57.0	12.1	10.5	mg m^{-3}
	PC	August 2010	14	2.34	31.8	11.8	8.26	mg m^{-3}
		BLI	8	3.20	83.1	29.2	31.7	mg m^{-3}
	Phytoplankton biomass	August 2010	13	2047	8368	4240	1840	mg m^{-3}
		BLI	15	482	8078	3334	2158	mg m^{-3}
	Cyanobacteria biomass	August 2010	14	510	7590	2996	1762	mg m^{-3}
		BLI	19	158	7050	1848	1975	mg m^{-3}

Table 1. *Cont.*

Matchup Window	Parameter	Dataset	n	Min	Max	Mean	St Dev	Units
±1 day	Chl-*a*	August 2010	13	8.31	34.4	19.1	9.60	mg m^{-3}
		BLI	18	2.43	33.8	13.4	9.39	mg m^{-3}
		KdKVI	105	1.50	57.0	12.1	10.5	mg m^{-3}
	PC	August 2010	14	2.34	31.8	11.8	8.26	mg m^{-3}
		BLI	8	3.20	83.1	29.2	31.7	mg m^{-3}
	Phytoplankton biomass	August 2010	13	2047	8368	4240	1840	mg m^{-3}
		BLI	15	482	8078	3334	2158	mg m^{-3}
	Cyanobacteria biomass	August 2010	14	510	7590	2996	1762	mg m^{-3}
		BLI	19	158	7050	1848	1975	mg m^{-3}
±3 days	Chl-*a*	August 2010	23	5.45	39.1	16.7	10.1	mg m^{-3}
		BLI	20	2.43	33.8	13.2	9.06	mg m^{-3}
		KdKVI	113	1.50	57.0	12.0	10.3	mg m^{-3}
	PC	August 2010	22	2.34	31.8	11.8	7.57	mg m^{-3}
		BLI	8	3.20	83.1	29.2	31.7	mg m^{-3}
	Phytoplankton biomass	August 2010	23	859	8794	3670	2098	mg m^{-3}
		BLI	17	482	8078	3184	2117	mg m^{-3}
	Cyanobacteria biomass	August 2010	22	210	7590	2831	1845	mg m^{-3}
		BLI	19	158	7050	1848	1975	mg m^{-3}
±4 days (IOPs only)	$a_{ph}(665)$	August 2010	29	0.035	0.339	0.123	0.084	m^{-1}
	$a_{ph}(620)$	August 2010	29	0.020	0.238	0.078	0.055	m^{-1}
	$b_b(650)$	August 2010	29	0.027	0.108	0.080	0.025	m^{-1}
±7 days	Chl-*a*	August 2010	29	5.45	39.1	15.7	9.22	mg m^{-3}
		BLI	52	2.43	45.1	15.8	11.5	mg m^{-3}
		KdKVI	113	1.50	57.0	12.0	10.3	mg m^{-3}
	PC	August 2010	28	2.34	31.8	11.6	6.80	mg m^{-3}
		BLI	12	3.20	99.6	37.0	39.2	mg m^{-3}
	Phytoplankton biomass	August 2010	29	859	8794	3549	1914	mg m^{-3}
		BLI	32	482	11097	3371	2638	mg m^{-3}
	Cyanobacteria biomass	August 2010	28	210	7590	2654	1708	mg m^{-3}
		BLI	36	0	9449	1856	2340	mg m^{-3}

2.3.3. Algorithm Implementation and Performance Assessment

A summary of the algorithms implemented for PC retrieval is provided in Table 2, with further details of their implementation to MERIS data for this study included in Appendix A. These algorithms were compiled based on a literature review and selected based upon the criteria that they: (1) are applicable or easily adapted to MERIS/OLCI data, and (2) were specifically developed or calibrated for inland waters.

Table 2. Summary of algorithms for retrieval of phycocyanin (PC) applicable to MERIS and OLCI data (further details provided in Appendix A). Models are a function of remote sensing reflectance (R_{rs}), absorption by phycocyanin (a_{pc}), absorption by water (a_w), backscattering (b_b), absorption by chlorophyll-*a* (a_{chla}), mass-specific absorption by phycocyanin ($a*_{pc}$), absorption by phytoplankton (a_{ph}), total minus water absorption (a_{t-w}), absorption by coloured detrital matter (a_{CDM}) and absorption of phytoplankton minus phycocyanin (a_{ph-pc}) at the respective wavelengths (λ). Coefficient values, where relevant, are provided within the table.

Model	Formula(e)	Reference(s)
Dekker93	$PC\left(mg\ m^{-3}\right) \propto 0.5 \times \left[\left(R_{rs}(560) + R_{rs}(665)\right) - R_{rs}(620)\right]$	[31]
Dekker93_modified	$PC\left(mg\ m^{-3}\right) \propto 0.5 \times \left[\left(R_{rs}(560) + R_{rs}(665)\right) - R_{rs}(620)\right] - R_{rs}(754)$	[23,31,76]
Schalles00	$PC\left(mg\ m^{-3}\right) \propto \frac{R_{rs}(665)}{R_{rs}(620)}$	[32]

Table 2. *Cont.*

Model	Formula(e)	Reference(s)
Simis05	$a_{pc}(620) = \left[\frac{R_{rs}(709)}{R_{rs}(620)} \times (a_w(709) + b_b) - b_b - a_w(620)\right] \times$ $\delta^{-1} - (\varepsilon \times a_{Chla}(665)),$ where $\delta = 0.84$ and $\varepsilon = 0.24$, and $PC\left(mg\ m^{-3}\right) = \frac{a_{pc}(620)}{a^*_{pc}(620)},$ where $a^*_{pc}(620) = 0.007\ m^2\ mg^{-1}$.	[28,37]
Hunter10_Duan	$PC\left(mg\ m^{-3}\right) \propto \left[R_{rs}^{-1}(620) - R_{rs}^{-1}(709)\right] \times R_{rs}(753)$	[22,23]
Mishra13	$a_{pc}(620) = \frac{\psi_1 a_{ph}(620) - a_{ph}(665)}{\psi_1 - \psi_2},$ where $\psi_1 = a_{chla}(665)/a_{chla}(620)$ and $\psi_2 = a_{pc}(665)/a_{pc}(620)$, and $PC\left(mg\ m^{-3}\right) = \frac{a_{pc}(620)}{a^*_{pc}(620)},$ where $a^*_{pc} = 0.0048\ m^2\ mg^{-1}$.	[39]
Mishra13_Simis	As in Mishra13 except $a^*_{pc} = 0.007\ m^2\ mg^{-1}$.	[37,39]
Qi14	$PC\ Index\ (PCI) = \quad (R_{rs}(560) + \frac{620-560}{665-560}$ $\times [R_{rs}(665) - R_{rs}(560)]) - R_{rs}(620)$ $PC\left(mg\ m^{-3}\right) = ae^{b*PCI},$ where $a = 3.87$ and $b = 1154$.	[40]
Qi14_Balaton	As in Qi14 except calibrated to Lake Balaton, where $a = 21.26$ and $b = -139.3$.	[40]
Li15	$a_{ph}(\lambda) = a_{t-w}(\lambda) - a_{CDM}(\lambda)$ $a_{ph-pc}(\lambda) = 1.1872\ C1(\lambda)\ a_{t-w}(665) + C2(\lambda)$ Where C1 and C2 are wavelength dependent regression coefficients outlined in Table A1 in Li et al. [41]. $a_{pc}(620) = a_{ph}(620) - a_{ph-pc}(620)$ $PC\left(mg\ m^{-3}\right) = \frac{a_{pc}(620)}{a^*_{pc}(620)}$ where $a^*_{pc}(620) = 0.0046\ m^2 mg^{-1}$.	[41]
Li15_Simis	As in Li et al. (2015) except $a^*_{pc}(620) = 0.0007\ m^2 mg^{-1}$.	[37,41]
Liu18	Four band semi-analytical algorithm for PC $(FBA_{PC}) = \left[\frac{1}{R_{rs}(620)} - \frac{0.4}{R_{rs}(560)} - \frac{0.6}{R_{rs}(709)}\right] \times R_{rs}(754)$ $PC\left(mg\ m^{-3}\right) = m \times FBA_{PC} + B$ where $m = 462.5$ and $B = 22.598$.	[42]
Liu18_Balaton	As in Liu18 except calibrated to Lake Balaton, where $m = 76.7$ and $B = 23.09$.	[42]

The performance of each algorithm for PC retrieval from MERIS data was assessed using matchups within 1 day of in situ PC measurements from 2010 and 2011 (n = 22; Table A1). The strength of the relationship between the algorithm estimated values and measured values was evaluated using linear least squares regression analysis. The agreement (goodness of fit) is reported by the coefficient of determination (R^2) and *p* value. Algorithm accuracy was quantified with measures of error, including the Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Bias. RMSE and Bias were also calculated in log space ($RMSE_{log}$ and $Bias_{log}$, respectively), following recent recommendations on error metrics for the performance assessment of satellite ocean colour data products [77]. The Median Absolute Percentage Error (MdAPE) and Symmetric Mean Absolute Percentage Error (SMAPE) were also quantified, as these metrics are less sensitive to outliers. Performance metrics were additionally calculated for MERIS-retrieved PC values validated with in situ PC< 50 mg m^{-3} only. The formulae for the error metrics are provided in Table 3.

Table 3. Formulae for error metrics, where *M, O* and *n* represent the modelled value, observed value (from in situ measurements) and number of samples, respectively.

Error Metric	Formula
RMSE	$\sqrt{\dfrac{\sum_{i=1}^{n}(O_i - M_i)^2}{n}}$
Bias	$\dfrac{\sum_{i=1}^{n}O_i - M_i}{n}$
MAPE	$100 \times \dfrac{1}{n} \times \sum_{i=1}^{n}\left|\dfrac{O_i - M_i}{O_i}\right|$
MdAPE	$\text{median}\left[100 \times \left|\dfrac{O_i - M_i}{O_i}\right|\right]$
SMAPE	$100 \times \dfrac{1}{n} \times \sum_{i=1}^{n}\left[\dfrac{|O_i - M_i|}{(|O_i| + |M_i|)/2}\right]$

3. Results

3.1. Pigments and Cell Counts

Inter-annual variations in cell counts are shown from 2007 to 2011 for total phytoplankton and cyanobacteria biomass, alongside Chl-*a* and PC pigment data (Figure 3). Lake Balaton generally has a spring (January–March) diatom bloom, which is indicated by the smaller peaks in phytoplankton biomass and Chl-*a*. This is followed by a late summer (August–October) bloom in cyanobacteria, as indicated by the larger peaks in phytoplankton biomass and Chl-*a*. Keszthely is the westernmost basin, which has higher Chl-*a* concentrations than Tihany in the eastern basin. Both Keszthely and Tihany show the same timing of peaks in phytoplankton abundance, although the late summer cyanobacteria bloom is more prolific in Keszthely, with cyanobacteria biomass reaching nearly an order of magnitude higher than that in Tihany.

3.2. Validation of Atmospheric Correction

MERIS remote sensing reflectance spectra atmospherically corrected with SCAPE-M_B2 were compared to in situ measurements made with a system of HyperSAS radiometers, which were resampled to the MERIS bands, assuming a Gaussian distribution for the SRFs. Scatterplots of data from the same day as the MERIS overpass (22 August 2010) showed acceptable agreement at all 12 bands compared ($\text{RMSE}_{\text{log}} < 0.205$, $\text{Bias}_{\text{log}} < 0.197$, MAPE < 58.7%, n = 7; Figure 4). Agreement was poorest over the NIR wavelengths (Bands 10–12) and blue portion of the spectrum (Band 1), while agreement was better for Bands 2–9 ($\text{RMSE}_{\text{log}} < 0.113$, $\text{Bias}_{\text{log}} < 0.107$, MAPE < 28.5%). Agreement declined when the entire dataset was considered (±4 day matchups; $\text{RMSE}_{\text{log}} < 0.316$, $\text{Bias}_{\text{log}} < 0.152$, MAPE < 98.6%, n = 30; Figure 4), however errors were still lowest for Bands 2–9 ($\text{RMSE}_{\text{log}} < 0.176$, $\text{Bias}_{\text{log}} < -0.000348$, MAPE < 37.7%). Despite the small same-day matchup dataset (n = 7), the SCAPE-M_B2 atmospheric correction shows acceptable performance, particularly from 443 to 709 nm (Bands 2–9).

MERIS SCAPE-M_B2 $R_{\text{rs}}(\lambda)$ spectra are presented alongside in situ $R_{\text{rs}}(\lambda)$ for all stations in Figures 5 and 6. For same-day matchups (stations 16–22), MERIS bands generally fell reasonably within the range of in situ $R_{\text{rs}}(\lambda)$. However, for matchups out with the ±1 day matchup window (±4 days), $R_{\text{rs}}(\lambda)$ was frequently under (stations 1–15) or overestimated (e.g., stations 29–30).

It is noted that there was a high standard deviation for MERIS $R_{\text{rs}}(\lambda)$ in the NIR range for 3 stations (10, 14 and 24) because these 3×3 pixel windows were impacted by the adjacency effect (i.e., were influenced by land pixels). When these 3 stations were removed from the validation dataset as outliers, the validation errors using the entire dataset were reduced by 51–84% in the 3 NIR bands (Bands 10–12). Therefore, data from these stations were disregarded in this study. Thus, with this improvement, the SCAPE-M_B2 model shows good performance across all MERIS bands with the exception of Band 1. Importantly, Band 1 (413 nm) is not used in any of the PC models tested, and therefore is not of concern for the purpose of this study.
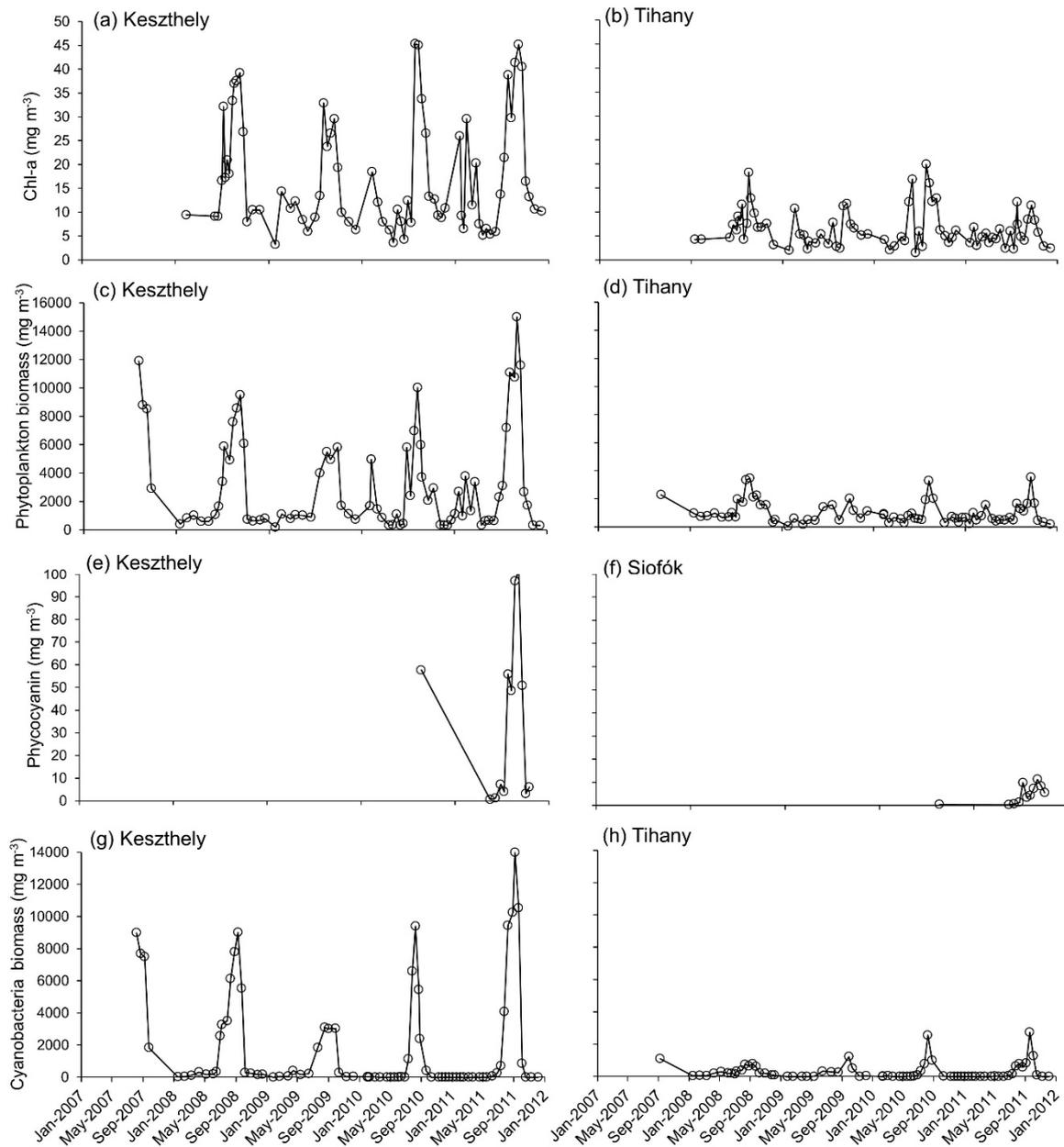
Figure 3. Inter-annual variations in (**a**,**b**) Chl-*a*, (**c**,**d**) phytoplankton biomass, (**e**,**f**) PC and (**g**,**h**) cyanobacteria biomass at Keszthely (westernmost basin) and Tihany or Siofók (easternmost basin; data from BLI routine monitoring).
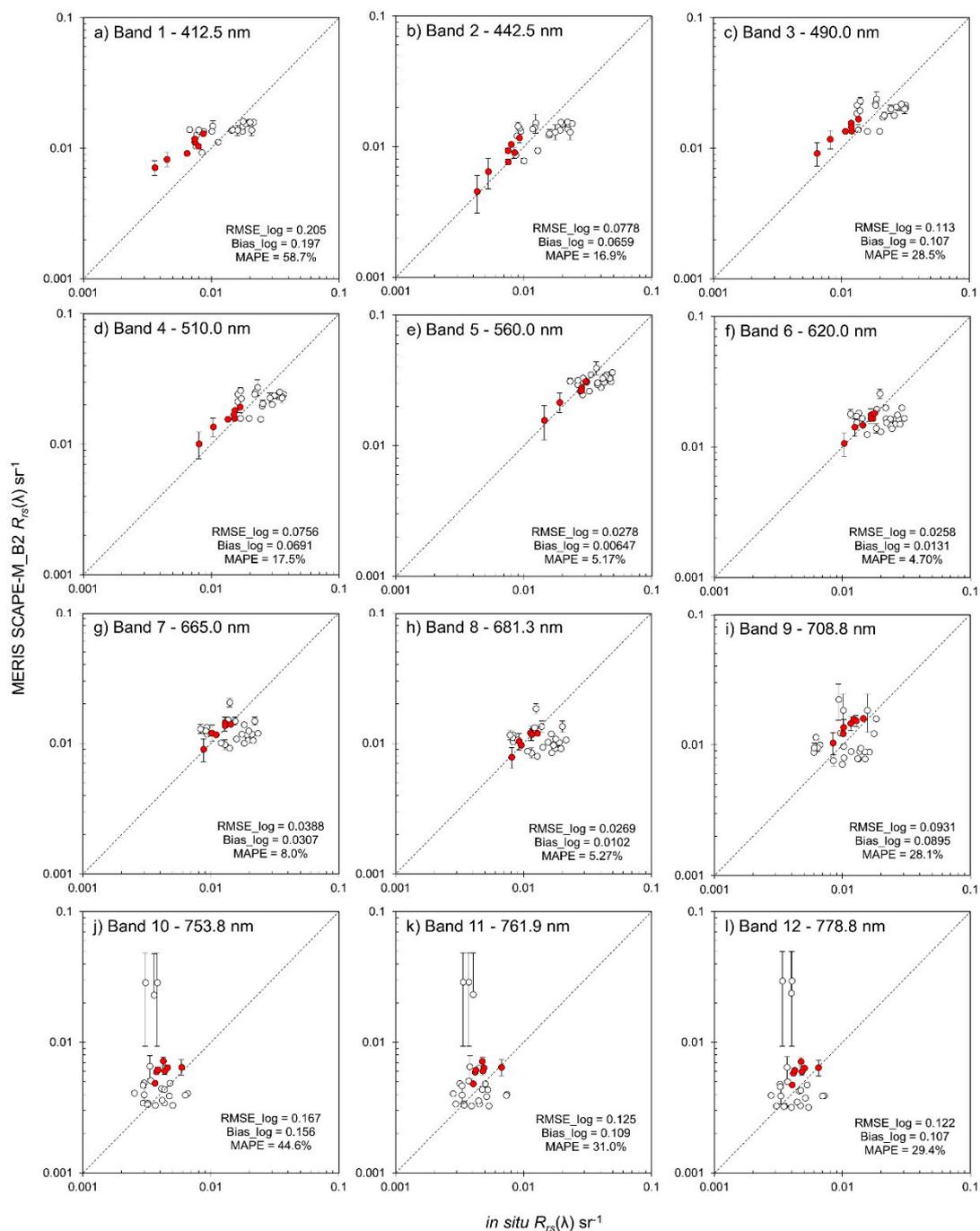
Figure 4. Validation scatter plots of MERIS $R_{rs}(\lambda)$ atmospherically corrected with SCAPE-M_B2 as a function of in situ $R_{rs}(\lambda)$ measured by the HyperSAS for each of the 12 MERIS bands (**a**)–(**i**). *In situ* HyperSAS $R_{rs}(\lambda)$ was simulated to each MERIS band using the spectral response functions, assuming a Gaussian distribution. Scatterplots include all stations with *in situ* data (n = 30), with same day matchups shown in red (n = 7). Statistics correspond to the same day matchups only.
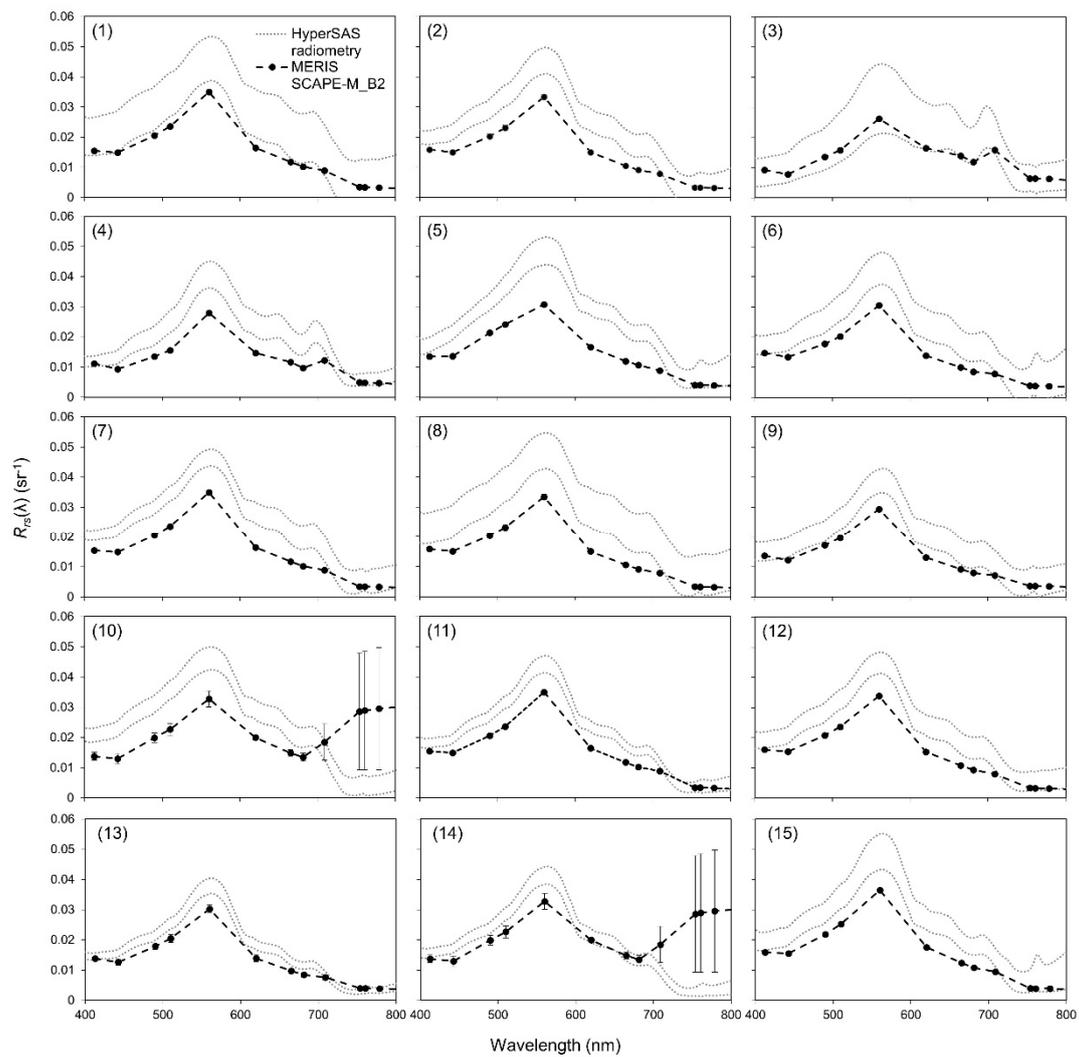
Figure 5. MERIS $R_{rs}(\lambda)$ spectra atmospherically corrected with SCAPE-M_B2 and in situ HyperSAS radiometry $R_{rs}(\lambda)$ data for stations 1–15. MERIS $R_{rs}(\lambda)$ are the mean of a 3×3 pixel window with error bars indicating standard deviation. HyperSAS $R_{rs}(\lambda)$ are shown as a range, with lines indicating the minimum to maximum $R_{rs}(\lambda)$ recorded at each station.
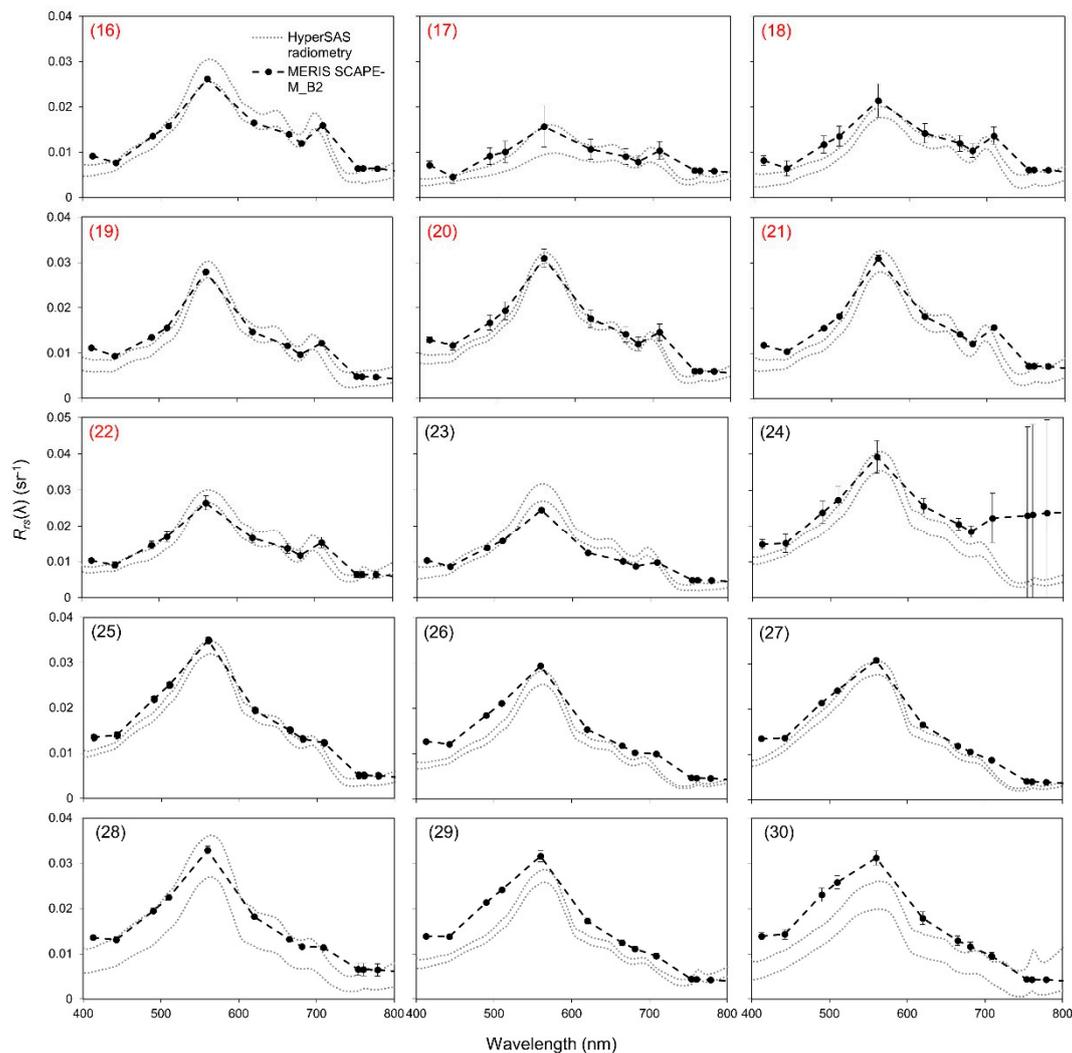
Figure 6. MERIS $R_{rs}(\lambda)$ spectra atmospherically corrected with SCAPE-M_B2 and in situ HyperSAS radiometry $R_{rs}(\lambda)$ data for stations 16–30. MERIS $R_{rs}(\lambda)$ are the mean of a 3×3 pixel window with error bars indicating standard deviation. HyperSAS $R_{rs}(\lambda)$ are shown as a range, with lines indicating the minimum to maximum $R_{rs}(\lambda)$ recorded at each station. Same-day matchups include stations 16–22 only, shown in red (n = 7). Note the different ranges for y-axis values.

3.3. Phycocyanin Algorithm Performance

The performance results for each algorithm tested are provided in Figure 7, with associated error metrics shown in Table A1. Dekker93 retrieved PC concentrations with low accuracy ($RMSE_{log}$ = 0.444, $Bias_{log}$ = 0.148, MAPE = 224%, MdAPE = 84.4%, SMAPE = 88.3%), however errors reduced markedly when the fourth term was added (Dekker_93_modified; $RMSE_{log}$ = 0.354, $bias_{log}$ = 0.0542, MAPE = 142%, MdAPE = 69.3%, SMAPE = 72.3%). The Schalles00 band ratio algorithm retrieved PC concentrations relatively well ($RMSE_{log}$ = 0.370), however bias ($bias_{log}$ = 0.238) and percentage error metrics (MAPE = 124%, MdAPE = 111%, SMAPE = 112%) were high due to poor performance for retrieval over low PC values. The Simis05 algorithm retrieved PC concentrations with the lowest RMSE (11.8 mg m^{-3}), MdAPE (50.8%) and SMAPE (48.2%), with low values for the other error metrics ($RMSE_{log}$ = 0.272, $bias_{log}$ = 0.147, MAPE = 77%) and a high coefficient of determination (R^2 = 0.710, p < 0.0001). The Hunter10_Duan12 algorithm-retrieved PC concentrations were strongly correlated with in situ values, however errors were high as low concentrations were overestimated and high concentrations were underestimated ($RMSE_{log}$ = 0.356, $bias_{log}$ = 0.107, MAPE = 153%, MdAPE = 114%,

SMAPE = 77%). The Mishra13 algorithm had the poorest relationship between retrieved and measured PC, with an R^2 value of 0.00836 using $a^*_{pc}(620)$ values from both Simis et al. [37] and Mishra et al. [39], and this was mainly driven by poor prediction at higher PC concentrations. However, Mishra13_Simis had the lowest $RMSE_{log}$ (0.246) and MAPE (61.2%) of all algorithms tested. Qi14 demonstrated the poorest performance overall, with a poor relationship to in situ PC and markedly higher error metrics ($RMSE_{log}$ = 1.55, $bias_{log}$ = 1.38, MAPE = 6197%, MdAPE = 3512%, SMAPE = 158%), although errors were reduced considerably when the coefficients for Qi14 were calibrated for Lake Balaton ($RMSE_{log}$ = 0.403, $bias_{log}$ = 0.0563, MAPE = 101%, MdAPE = 52.8%, SMAPE = 64.4%). While PC retrieved with Li15 had a high coefficient of determination (R^2 = 0.716, $p < 0.0001$) and the lowest bias ($bias_{log}$ = 0.0349), agreement was poor ($RMSE_{log}$ = 0.503, MAPE = 205%, MdAPE = 170%, SMAPE = 141%) with negative PC retrieved for in situ concentrations <10 mg m^{-3}. Furthermore, using the value of $a^*_{pc}(620)$ from Simis et al. (2007) made little difference to algorithm performance (Li15_Simis; $RMSE_{log}$ = 0.523, $bias_{log}$ = −0.147, MAPE = 150%, MdAPE = 131%, SMAPE = 136%). Similarly, PC retrieved with the Liu18 semi-analytical algorithm had a strong correlation with in situ PC (R^2 = 0.634, $p < 0.0001$), however errors were typically high ($RMSE_{log}$ = 0.841, $bias_{log}$ = −0.305, MAPE = 184%, MdAPE = 122%, SMAPE = 134%) with low PC values poorly predicted. Calibration of Liu18 coefficients to Lake Balaton did improve algorithm performance ($RMSE_{log}$ = 0.354, $bias_{log}$ = 0.088, MAPE = 155%, MdAPE = 113%, SMAPE = 78.2%), however the model overpredicted for low PC concentrations and underpredicted high concentrations.

All algorithms were additionally validated against same day matchups for PC concentrations less than 50 mg m^{-3} (Table A2). These results indicated lower RMSE for all algorithms with the exception of Qi14, however the rest of the error metrics typically increased for algorithms validated against low PC concentrations. The main exception was Mishra13_Simis, which demonstrated improved performance with lower RMSE (5.14 mg m^{-3}), $RMSE_{log}$ (0.224), MAPE (59.5%), MdAPE (36.3%) and SMAPE (48.2%) when validated with low PC concentrations. This highlights the variable performance of PC retrieval algorithms over different PC concentration ranges, and should be considered when choosing an algorithm for use in a specific optical water type [78]. While the Mishra_Simis model performed best for low values of PC, the Simis05 algorithm performed best with consistently low errors for the full range of PC values in Lake Balaton. Thus, the Simis05 semi-analytical algorithm was investigated in further detail for the remainder of this study.
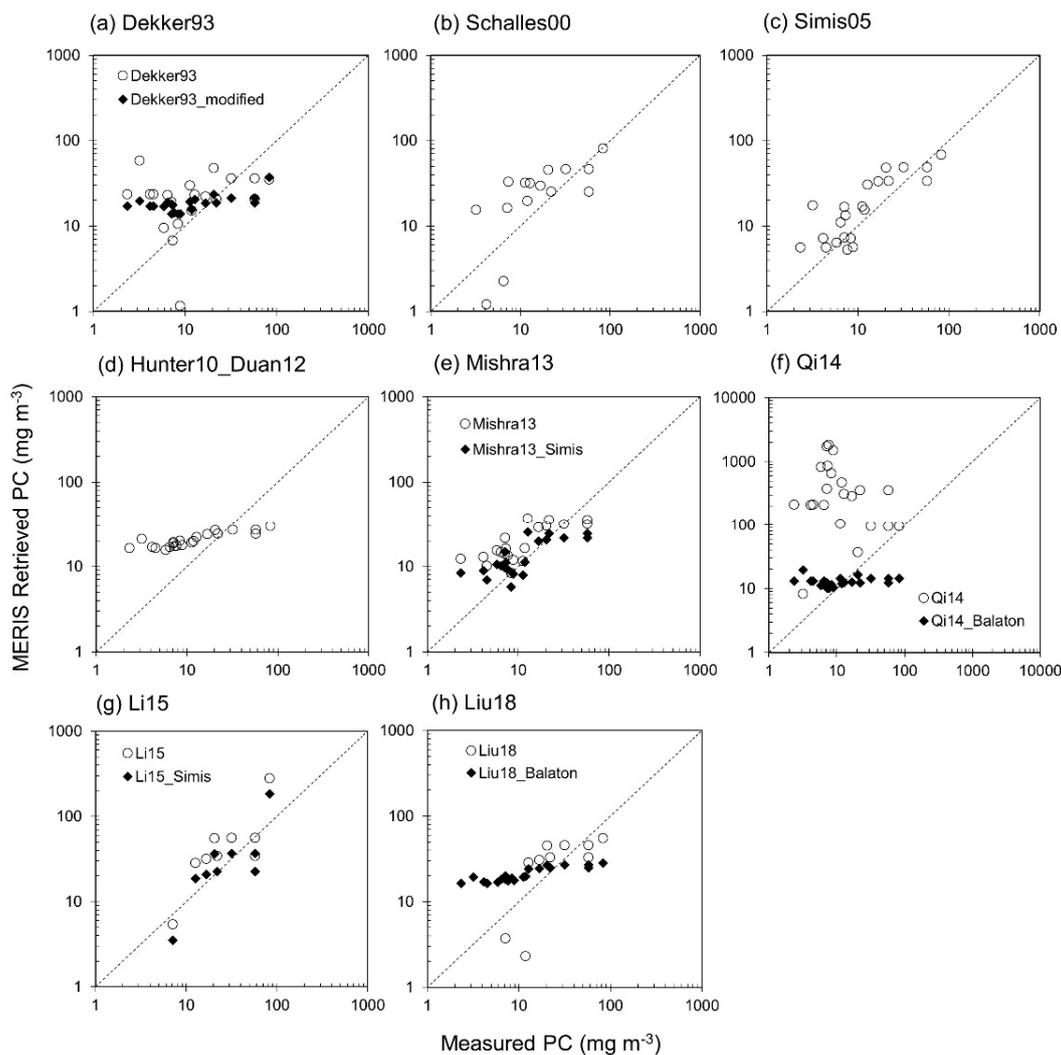
Figure 7. Algorithm performance for phycocyanin (PC) retrieval from MERIS data over Lake Balaton from 2010 to 2011 using (**a**) Dekker93, (**b**) Schalles00, (**c**) Simis05, (**d**) Hunter10_Duan12, (**e**) Mishra13, (**f**) Qi14, (**g**) Li15 and (**h**) Liu18. Matchups are ±1 day of in situ measurements (n = 22). Note the different axis scale for Qi14 due to high retrieved PC values >1000 mg m^{-3}. Negative retrieved PC values are not shown in log scale for Schalles00, Mishra13 and Li15. For algorithm details see Table 1 and Appendix A.

3.4. Chlorophyll-a Retrieval

The Gons05 algorithm for Chl-*a* retrieval is the predecessor and foundation for the Simis05 algorithm for PC retrieval, and therefore is considered here. More specifically, the Gons05 algorithm estimates $a_{ph}(665)$, which is assumed to comprise absorption from Chl-*a* only ($a_{ph}(665)= a_{Chla}(665)$). $a_{Chla}(665)$ is extrapolated to 620 nm and subtracted from the combined pigment absorption to obtain $a_{PC}(620)$ (see Appendix A for the Gons05 and Simis05 algorithm descriptions). Thus, errors in the retrieval of $a_{Chla}(665)$ using Gons05 will propagate through to the estimate of PC. It is beyond the scope of this study to comprehensively test Chl-*a* algorithms, and moreover, this has recently been done in turbid, optically complex waters elsewhere (e.g. [54,79–81]). Thus, only the Gons05 algorithm is considered here purely as a basis for the best performing PC algorithm, Simis05.

Chl-*a* was retrieved from MERIS after estimating the absorption coefficient of phytoplankton at 665 nm. At ±1, 3- and 7-day matchup intervals, the Gons05 algorithm retrieved Chl-*a* with good performance using all datasets for validation, however MERIS-retrieved values typically overestimated measured values (RMSE$_{log}$ = 0.369–0.394, Bias$_{log}$ = 0.296–0.329, MAPE = 132–151%,

MdAPE = 74.2–86.4%, SMAPE = 63.0–68.5%; Figure 8; Table A3). There was also increased scatter in the relationship between retrieved and measured values at Chl-*a* concentrations <10 mg m^{-3} (±7 days: RMSE$_{log}$ = 0.476, MAPE = 208%) compared to higher Chl-*a* concentrations, >10 mg m^{-3} (±7 days: RMSE$_{log}$ = 0.384, MAPE = 142%). Furthermore, retrieved Chl-*a* values validated with surface samples from August 2010 only had higher R^2 values (R^2 = 0.906–0.943, *p* < 0.0001) and better agreement (RMSE$_{log}$ = 0.185–0.203, Bias$_{log}$ = 0.167–0.199, MAPE = 50.0–58.9%, MdAPE = 50.6–52.7%, SMAPE = 38.1–45.0%; Table A4) than those validated with integrated samples routinely taken by BLI (R^2 = 0.723–0.810, *p* < 0.0001, RMSE$_{log}$ = 0.331–0.337, Bias$_{log}$ = 0.249–0.271, MAPE = 107–112%, MdAPE = 61.5–76.5%, SMAPE = 55.7–59.9%; Table A5) or surface samples from the KdKVI dataset (R^2 = 0.784–0.786, *p* < 0.0001, RMSE$_{log}$ = 0.416–0.421, Bias$_{log}$ = 0.352–0.355, MAPE = 165–169%, MdAPE = 101%, SMAPE = 72.7–73.3%; Table A6). When retrieved Chl-*a* values were restricted to those stations also used for PC validation, there was better agreement between MERIS-retrieved and in situ Chl-*a* concentrations (RMSE$_{log}$ = 0.211–0.241, Bias$_{log}$ = 0.174–0.216, MAPE = 55.9–70.0%, MdAPE = 50.1–51.9%, SMAPE = 39.2–48.0%; Figure 9 and Table A7).

Gons05 retrieved Chl-*a* values were also partially validated with phytoplankton counts (Figure 10). There was good agreement between the Chl-*a* values retrieved from MERIS and total phytoplankton biomass, with little difference in the R^2 values for ±1 (R^2 = 0.667, *p* < 0.001), ±3 (R^2 = 0.698, *p* < 0.001) and ±7 (R^2 = 0.636, *p* < 0.001) day matchups (Table A3). There was no marked difference in agreement for validation using surface (August 2010) versus integrated (BLI) samples (Tables A4 and A5, respectively).
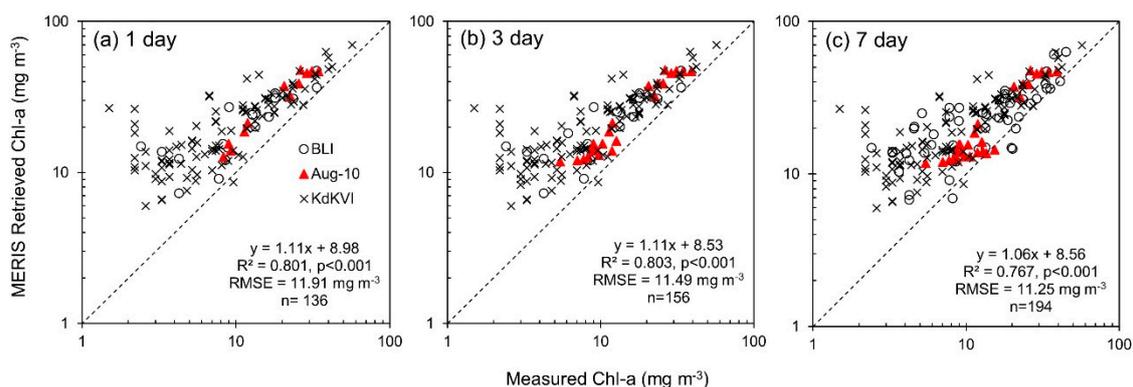
Figure 8. Gons05 retrieval of Chl-*a* from MERIS at matchups with measured Chl-*a* within (**a**) ±1 day, (**b**) ±3 days and (**c**) ±7 days. Retrieved values shown are mean pixel values within two standard deviations. Linear regression results shown are for all data. Dashed line is 1:1.
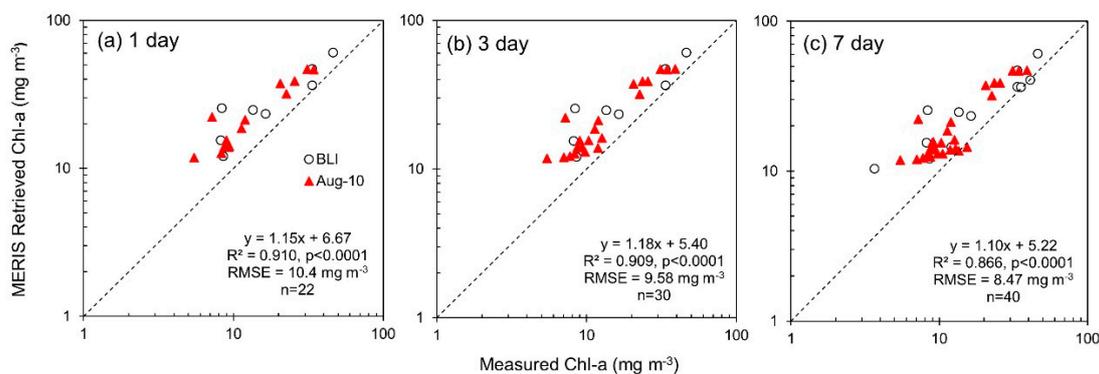

Figure 9. Gons05 retrieval of Chl-*a* from MERIS at matchups with measured Chl-*a* within (**a**) ±1 day, (**b**) ±3 days and (**c**) ±7 days, including only those matchups used for retrieval of PC using Simis05. Retrieved values shown are mean pixel values within two standard deviations. Linear regression results shown are for all data. Dashed line is 1:1.
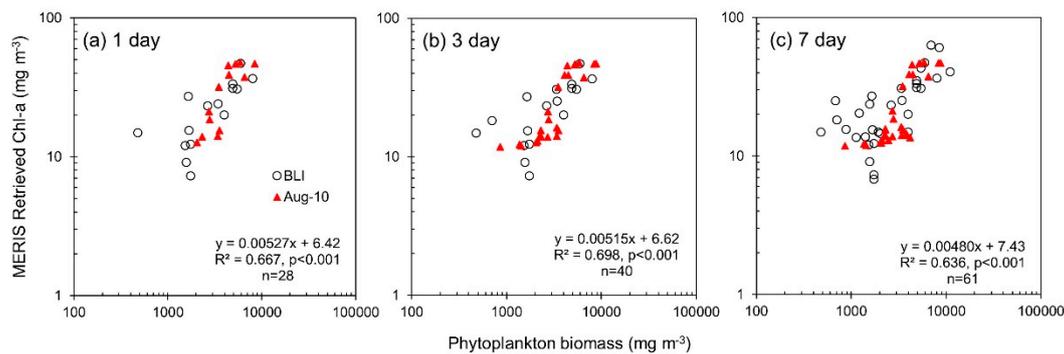
Figure 10. Gons05 retrieval of Chl-*a* from MERIS at matchups with phytoplankton biomass from (**a**) ±1 day, (**b**) ±3 days and (**c**) ±7 days. Retrieved values shown are mean pixel values within two standard deviations. Linear regression results are shown for all data.

3.5. Phycocyanin Retrieval

Phycocyanin concentrations were retrieved from MERIS data after estimation of $a_{ph}(620)$, and were validated with measured concentrations from the August 2010 campaign and BLI routine monitoring data. The PC concentrations retrieved using Simis05 showed good agreement with the in situ data for matchups within 1, 3 or 7 days of the overpass ($RMSE_{log}$ = 0.261–0.273, $Bias_{log}$ = 0.0218–0.147, MAPE = 63.0–77.0%, MdAPE = 49.5–50.8%, SMAPE = 48.2–49.8%; Figure 11 and Table A3). At all three time windows, retrieved PC values were a better fit with August 2010 surface samples for matchups ±7 days (R^2 = 0.718, p < 0.0001; Table A4) than those validated with integrated samples routinely taken by BLI (R^2 = 0.580, p < 0.01; Table A5). Determination coefficients were high for matchups with BLI data at ±1 and 3 days (R^2 = 0.910, p < 0.001), however this is likely driven by the small sample size at these windows (n = 8). Overall, PC retrieved at all three time windows showed better agreement with the August 2010 surface samples ($RMSE_{log}$ = 0.246–0.254, $Bias_{log}$ = 0.042–0.189, MAPE = 59.3–71.6%, MdAPE = 52–61.2%, SMAPE = 49.4–49.9%) than the BLI-integrated samples ($RMSE_{log}$ = 0.301–0.324, $Bias_{log}$ = −0.026–0.100, MAPE = 71.7–86.4%, MdAPE = 33.1%, SMAPE = 45.4–50.7%). As with Gons05, the Simis05 algorithm did not accurately retrieve PC values below approximately 5 mg m^{-3}, and this was most noticeable with the greatest matchup window of ±7 days (PC < 5 mg m^{-3}: $RMSE_{log}$ = 0.366, $Bias_{log}$ = 0.284, MAPE = 126%, MdAPE = 69.7%, SMAPE = 58.7%; PC > 5 mg m^{-3}: $RMSE_{log}$ = 0.253, $Bias_{log}$ = −0.0244, MAPE = 51.9%, MdAPE = 45.6%, SMAPE = 48.2%).
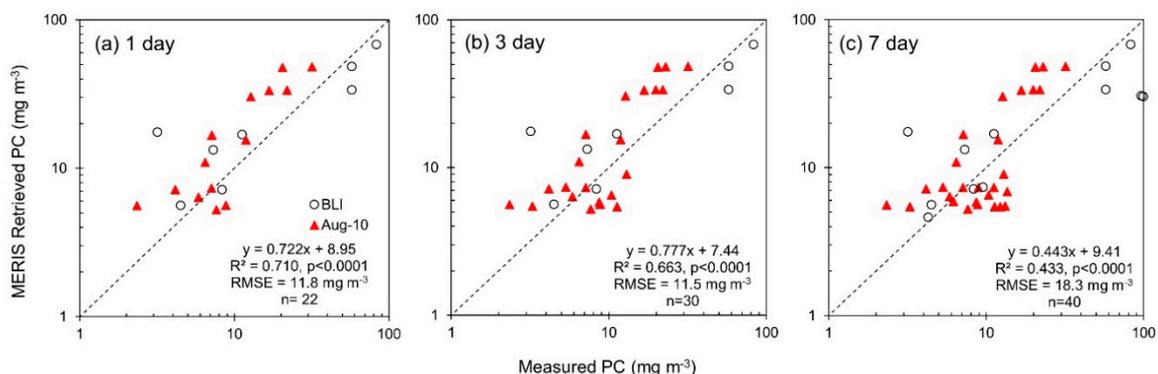
Figure 11. Simis05 retrieval of PC from MERIS at matchups with measured PC from (**a**) ±1 day, (**b**) ±3 days and (**c**) ±7 days. Retrieved values shown are mean pixel values within two standard deviations. Linear regression results shown are for all data. Dashed line is 1:1.

PC concentrations retrieved from the Simis05 algorithm were also partially validated with measured cyanobacteria biomass (Figure 12). There was significant scatter around the linear regression, however PC estimates were significantly related to cyanobacteria counts with R^2 values from 0.462 to

0.525 ($p < 0.001$) depending on the temporal window used for matchup. There was no evidence of degradation in the relationship between retrieved PC and cyanobacteria biomass, with correlation coefficients actually improving from ±1 to 7 days. Although, this is perhaps simply due to the greater number of matchups at ±3 and ±7 days and the increased data range as compared to ±1 day. It should also be noted that determination coefficients improved very slightly when a linear regression was performed using retrieved PC as a function of combined cyanobacteria and cryptophyte biomass (R^2 = 0.489-0.554, $p < 0.0001$). The relationship between retrieved Chl-*a* and phytoplankton biomass was stronger than the relationship between retrieved PC and cyanobacterial biomass at all matchup intervals (see Table A3).
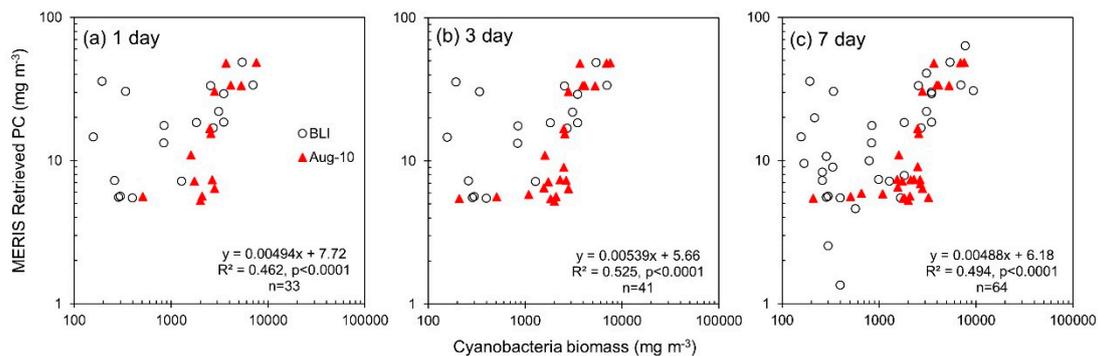
Figure 12. Simis05 retrieval of PC from MERIS at matchups with cyanobacteria biomass from (**a**) ±1 day, (**b**) ±3 days and (**c**) ±7 days. Retrieved values shown are mean pixel values within two standard deviations. Linear regression results are shown for all data.

An example time-series of the mapped PC products produced using the Simis05 algorithm is shown in Figure 13 for 2008. Retrieved PC concentrations show a significant peak in late summer during the cyanobacteria bloom, from mid-August into September. A gradient in PC is indicated across the lake, with the highest concentration generally in the western basins. It is also noted that a smaller spike in estimated PC is shown in western basins in the image for 26 May 2008 (Figure 13), however this is not mirrored by the cyanobacteria cell counts (Figure 3g,h).
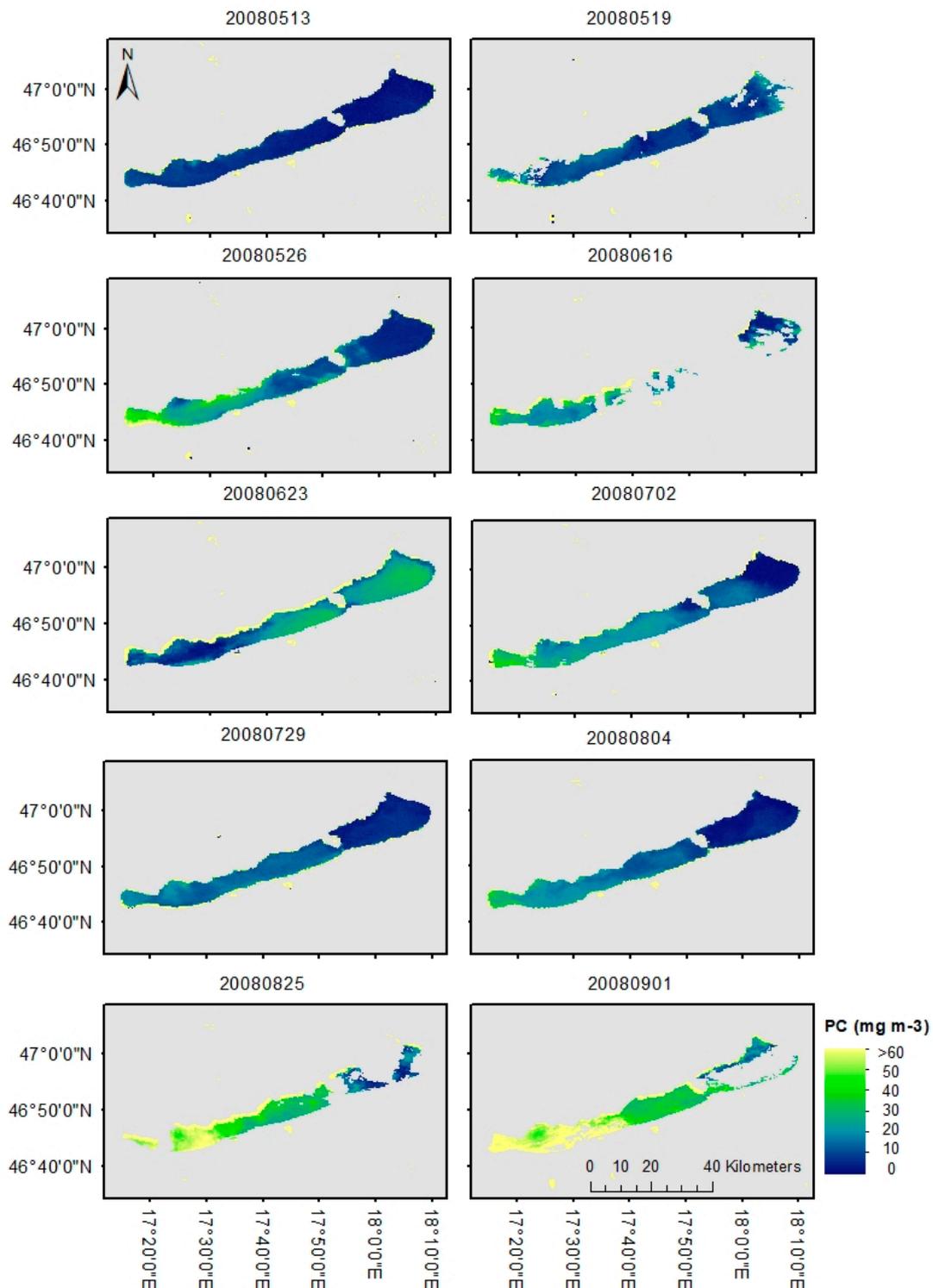
Figure 13. Time series of estimated PC from MERIS (Simis05 algorithm) for May–September 2008.

3.6. IOP Retrieval

The Gons05 semi-analytical algorithm was deconstructed and the component parts were validated using the IOP measurements from the August 2010 sampling campaign on Lake Balaton. The Gons05 algorithm retrieves $a_{ph}(665)$, which is assumed to comprise Chl-*a* absorption only ($a_{ph}(665) = a_{Chla}(665)$), using a correction factor, γ [28,34]. $a_{ph}(620)$ is then extrapolated from $a_{ph}(665)$, using a correction factor,

ε, which is then used to estimate PC concentration using the Simis05 algorithm (see Appendix A for Gons05 and Simis05 algorithm descriptions).

This study compared the retrieved $a_{ph}(665)$ against those determined using the filter-pad absorption measurements (Figure 14a). Although the coefficient of determination for least squares linear regression was high ($R^2 = 0.836$, $p < 0.0001$), the retrieved values for $a_{ph}(665)$ were found to overestimate the measured $a_{ph}(665)$ by a factor of 2 with large errors ($RMSE_{log} = 0.444$, $Bias_{log} = 0.430$, MAPE = 178%, MdAPE = 175%, SMAPE = 90.6%; Table A4). Summed absorption by Chl-*a* and PC at 620 nm was also retrieved with Simis05 and compared to measured $a_{ph}(620)$ from August 2010 (Figure 14b). Again, the coefficient of determination was high ($R^2 = 0.834$, $p < 0.0001$), however the values retrieved with Simis05 were found to overestimate measured pigment absorption by a factor of 3 ($RMSE_{log} = 0.645$, $Bias_{log} = 0.635$, MAPE = 346%, MdAPE = 332%, SMAPE = 123%; Table A4). Similarly, the bio-optical models of Mishra13 and Li15 also overestimated $a_{ph}(\lambda)$ (Figure A1), and $b_b(\lambda)$ was poorly represented by both models for same-day matchups (Figure A2). However, the Mishra13 and Li15 estimates of $a_{ph}(\lambda)$, especially Li15 retrieved $a_{ph}(665)$, showed better agreement with in situ data than Simis05. It is important to note, however, that filter-pad absorption measurements can be subject to significant errors due to unresolved issues with path-length amplification [82], therefore some of this error may be a result of the method itself.
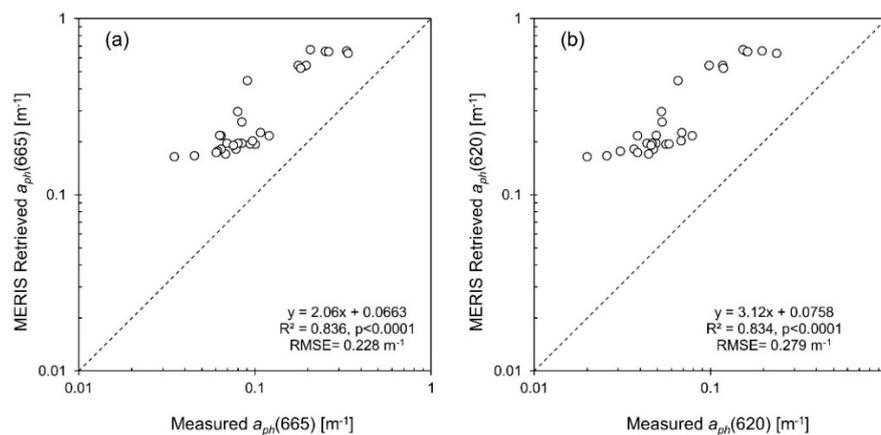
Figure 14. Retrieval of (**a**) $a_{ph}(665)$ and (**b**) $a_{ph}(620)$ using Simis05 for 22 August 2010 with matchups ± 4 days (n = 29).

The Gons05 and Simis05 algorithms estimate $b_b(\lambda)$ from reflectance at 778.75 nm based on the assumption that it is spectrally invariant over the wavelengths used in the algorithm [33]. This assumption was tested using $b_b(\lambda)$ coefficients measured at three wavelengths with an ECO-BB3 at 35 stations during the August 2010 sampling campaign (Figure 15a). Backscattering was highest at 470 nm, with lower coefficients at 532 and 650 nm. Measured $b_b(\lambda)$ coefficients at 650 nm were then compared with the retrieved $b_b(\lambda)$ coefficient from MERIS band 12 (778.75 nm) using the image from 22 August 2010 (Figure 15b). $b_b(650)$ was chosen because it lies between the 620 nm and 665 nm bands used in the Gons05 and Simis 05 algorithms. If backscattering was spectrally neutral and the algorithms were accurately retrieving the $b_b(\lambda)$ coefficient, a correlation between MERIS-retrieved $b_b(778.75)$ and measured $b_b(650)$ may be expected. However, there was no linear relationship between the measured and retrieved $b_b(\lambda)$ coefficients ($R^2 < 0.001$, $p > 0.5$; Table A4). Similarly, $b_b(\lambda)$ was poorly retrieved by the bio-optical models, Mishra13 and Li15 (Figure A2). The assumption of spectral neutrality may be reasonable based on the flat trend between 532 and 650 nm (Figure 15a), although it is unknown how $b_b(\lambda)$ varies between 650 and 780 nm. In general, the retrieved $b_b(\lambda)$ coefficients also showed significantly less variation across the stations sampled in Lake Balaton than was actually measured.
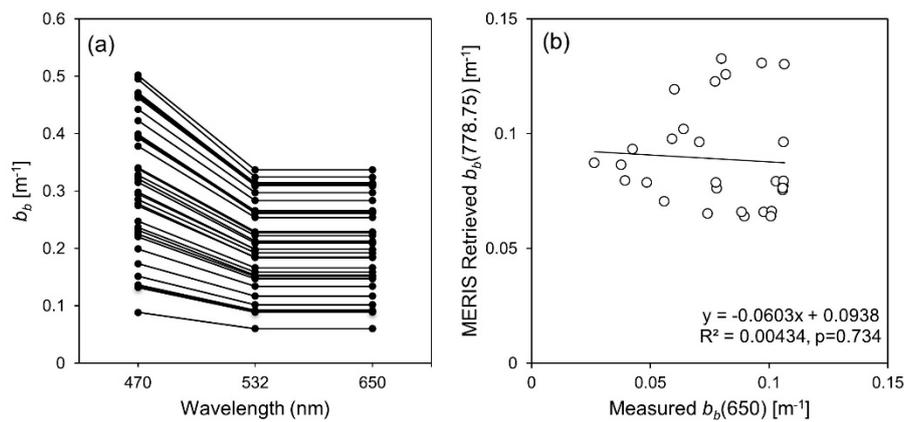
Figure 15. (**a**) Spectra of backscattering at three wavelengths as measured by the BB3 during the August 2010 field campaign (n = 35) and (**b**) MERIS-retrieved b_b(778.75) using Simis05 with matchups of in situ measured b_b(650) ± 4 days (n = 29).

3.7. Time Series Analysis

The pigment concentrations retrieved from MERIS (± 7 days) using Gons05 and Simis05 were compared alongside the available time series of pigment and cell biomass data from the BLI dataset (Figure 16). Generally, MERIS retrieved concentrations of Chl-*a* and PC followed the interannual patterns of in situ Chl-*a*, phytoplankton biomass and cyanobacteria biomass. However, retrieved pigments tended to match the variability better in Keszthely (Figure 16a,c,e) than in Tihany (Figure 16b,d,f).
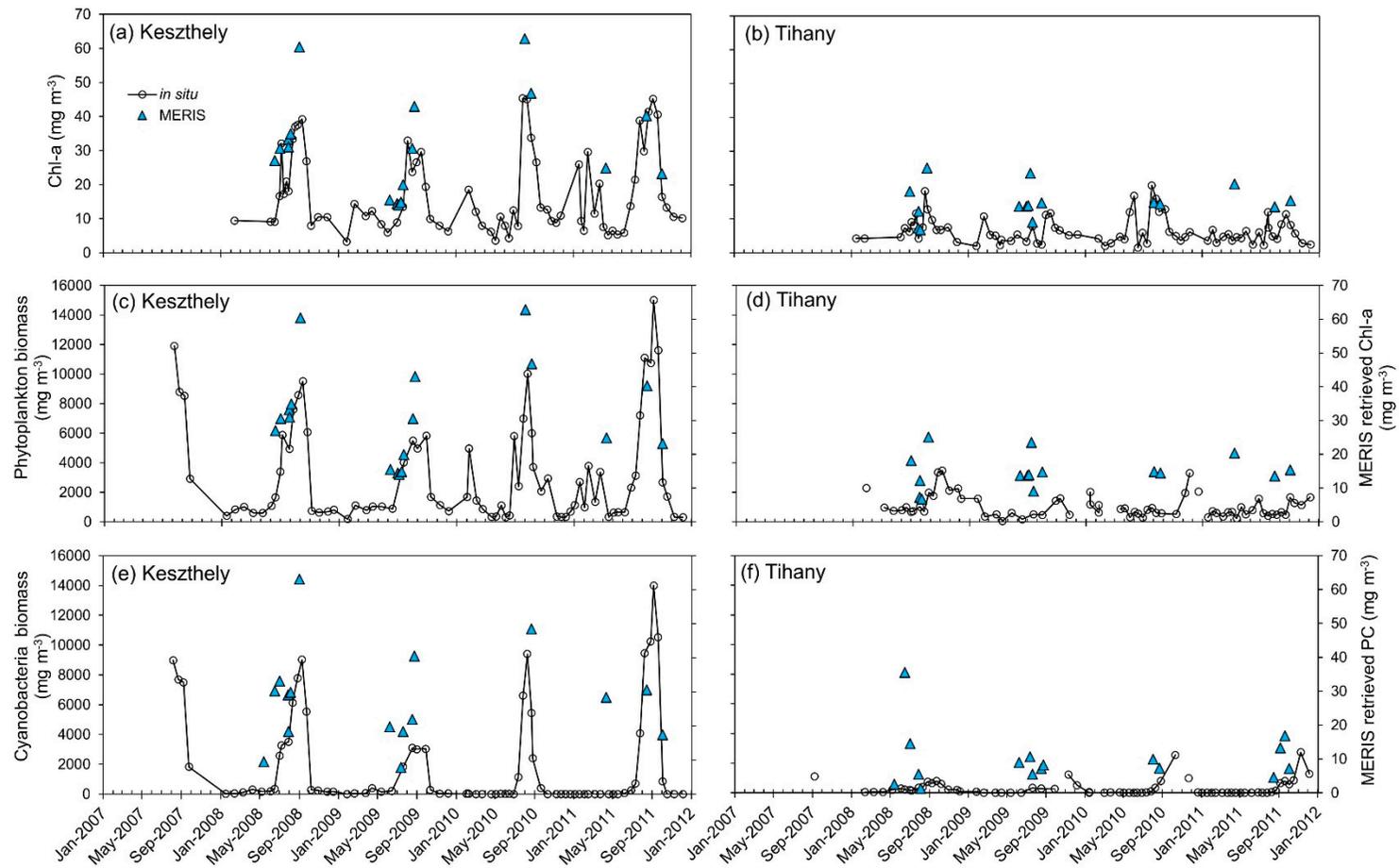
Figure 16. MERIS-retrieved concentrations of Chl-*a* (**a**–**d**) and PC (**e**–**f**) over the time series of in situ measurements of Chl-*a* (**a**–**b**), phytoplankton biomass (**c**–**d**) and cyanobacteria biomass (**e**–**f**) at Keszthely (westernmost basin) and Tihany (easternmost basin).

4. Discussion

4.1. Algorithm Performance

The Gons05 and Simis05 algorithms retrieved pigment concentrations (Chl-*a* and PC, respectively) in Lake Balaton with acceptable accuracy (see Table A3 to Table A7 in Appendix B for a summary of retrieval parameters and errors). Using ±1 day matchups, Chl-*a* ($RMSE_{log}$ = 0.394, $Bias_{log}$ = 0.329, MAPE = 151%, MdAPE = 86.4%, SMAPE = 68.5%) and PC concentrations ($RMSE_{log}$ = 0.272, $Bias_{log}$ = 0.147, MAPE = 77%, MdAPE = 50.8%, SMAPE = 48.2%) could be retrieved with from MERIS, thus this study can attest to the good performance of Gons05 and Simis05 algorithms for pigment retrieval in Lake Balaton.

Several models for PC retrieval from MERIS data were compared for Lake Balaton, and of these Simis05 provided the most accurate PC estimates over the full dataset. The other empirical, semi-empirical and semi-analytical models (Dekker93, Schalles00, Hunter10_Duan and Liu18) had generally larger errors than Simis05 or the bio-optical models, with the Qi14 model performing the poorest overall (Table A1). However, errors could be markedly reduced by using site-specific adaptations of each model, for example by the addition of an additional term to account for absorption by water (Dekker92_modified) or calibration of model coefficients with the 2010–2011 Balaton dataset (Qi14_Balaton or Liu18_Balaton).

Lower error metrics were observed for Simis05 and the bio-optical models (Mishra13 and Li15). However, although a stronger linear relationship was found between MERIS-retrieved and in situ PC using Li15, this model significantly underestimated PC values <10 mg m^{-3} and overestimated for concentrations ~10–100 mg m^{-3}. The Li15 model was developed using three reservoirs in Indiana (Eagle Creek, Geist and Morse Reservoirs), where PC concentrations ranged from 0.73 to 370 mg m^{-3}, while the PC validation dataset for Lake Balaton varied over a narrower range (2.34 to 113 mg m^{-3}). However, the discrepancy in performance using Li15 is more likely related to the high concentration of inorganic particles in Lake Balaton (in August 2010, 3–30 mg L^{-1} and comprising up to 91% of the total suspended matter or TSM), therefore the model coefficients developed for the reservoirs may not be appropriate in Lake Balaton. Although TSM concentrations were relatively similar between the Indiana reservoirs and Lake Balaton, we are unaware as to what fraction of the TSM in the Indiana reservoirs was comprised of organic or inorganic particles and cannot therefore compare. In contrast, retrieved PC concentrations from the Mishra13 algorithm had no significant relationship with in situ PC, although agreement with in situ PC improved greatly for lower PC concentrations (<50 mg m^{-3}). In fact, Mishra13_Simis demonstrated the lowest errors across the board when validated with the low PC dataset only (Table A2). This is unexpected as Mishra13 was developed using a training set from hypereutrophic aquaculture ponds (Chl-*a* = 59.4–1376.6 mg m^{-3}; PC = 68.1–3032.5 mg m^{-3}). The Mishra13 model is based on the well-validated QAA for Chl-*a* retrieval in ocean waters [83], and uses physics-based relationships to invert reflectances into absorption and backscattering properties of the optically active substances. In this way, the Mishra13 model should be adaptable to variable OWTs, and indeed, it performed well in Lake Balaton where PC concentrations were <50 mg m^{-3}. However, further work is needed to test the applicability of Mishra13 at lower cyanobacterial biomass and particularly in conditions representative of the early stages of bloom development. It is also noted that both the Mishra13 and Li15 inversion algorithms were also developed using in situ reflectance measurements and recommended different methods for conversion of above surface to below surface $R_{rs}(\lambda)$.

Ultimately, the Simis05 algorithm was found to outperform both simple empirical algorithms and more complex inversion algorithms over full the range of PC concentrations found in Lake Balaton (Table A1). Recent studies on Chl-*a* retrieval have recommended a return to simple empirical models [84], however our results suggest semi-analytical approaches are more accurate for PC retrieval. Although there were also differences in the bio-geo-optical properties of the lakes used to calibrate Simis05, Lake Loosdrecht and Lake IJsselmeer are noted to have generally high absorption by non-algal particles (up to 50% of water constituents), which is comparable to that in Lake Balaton (up to 40%) [51].

Therefore, it may be that semi-analytical and analytical models for PC retrieval are more appropriate for use in lakes with high concentrations of suspended inorganic particles.

4.2. Biomass Retrieval

The estimated PC and Chl-*a* concentrations from MERIS were also partially validated using microscopy-based estimates of cyanobacteria and total phytoplankton biomass, respectively. PC and Chl-*a* retrieval with Gons05 and Simis05 was a good prediction of cell biomass measured from surface and/or integrated samples for ±1, 3- and 7-day temporal windows ($R^2 > 0.4$, $p < 0.01$; Tables A3–A5). The coefficients of determination were lower for the linear regressions between retrieved pigments and cell biomass ($R^2 = 0.462$–0.698, $p < 0.0001$) in contrast to measured pigments ($R^2 = 0.433$–0.803, $p < 0.0001$), when using the complete dataset for validation (Table A3). This was expected due to the variability in PC cell quotas among freshwater cyanobacteria taxa [9]. The exception to this was for the ±7-day PC matchups, where MERIS-retrieved PC was explained marginally better by cyanobacteria biomass ($R^2 = 0.494$, $p < 0.0001$) than by measured PC ($R^2 = 0.433$, $p < 0.0001$).

These results demonstrate that satellite retrieved PC concentrations can be used as an indicator of cyanobacterial biomass, and allows the use of cyanobacterial cell counts for partial-validation of the PC algorithms. Lake Balaton, like many other lakes, has longer and more complete data on phytoplankton cell counts than on PC measurements. However, the relationship between PC and cyanobacteria biomass in Balaton may be seasonally variable and largely dependent on nutrient availability. It is also important to note that this strong relationship may not be found in other lakes, where intracellular PC content in cyanobacteria can vary in response to environmental stressors, such as nitrogen limitation or increased irradiance (photoacclimation). Recently, cell counts were successfully retrieved from numerous lakes in the eastern US, with a significant positive correlation between MERIS-retrieved and in situ cyanobacteria cell counts for matchups within 7 days of MERIS overpass [85]. Although both Lunetta et al. [85] and the present study found significant scatter around this relationship, both studies confirm the capability of remote sensing for monitoring cyanobacteria cell counts in lakes.

Additionally, there was little variation in accuracy at the three matchup ranges when validating with cell counts, with no apparent degradation in PC or Chl-*a* retrieval capability within ±7 days of image acquisition. It was noted, however, that the relationship between retrieved Chl-*a* and phytoplankton biomass was better than that for retrieved PC and cyanobacterial biomass. This is perhaps because PC may be more spatially variable within a pixel. Although in its time MERIS provided the best available resolution for satellite application to lakes, the pixel size (300 m) is still quite large given that cyanobacteria blooms can vary markedly over a few metres or hours [86]. Therefore, it is likely the time difference between in situ samples and the MERIS overpass causes this discrepancy. Furthermore, the validation approach for PC in inland waters requires further investigation, as application of methods from ocean and coastal waters (e.g., use of a 3×3 pixel window) may be unsuitable for use in heterogeneous waters.

4.3. Temporal Window for Validation

When the complete in situ pigment dataset is considered, Chl-*a* retrieval deteriorated only slightly over the ±1 to 7-day windows, with only a small reduction in errors (Figure 8; Table A3). $RMSE_{log}$ ranged from 0.369 to 0.394, while MAPE ranged from 132 to 151% for the three temporal windows, confirming that the Gons05 algorithm can retrieve Chl-*a* concentrations in Lake Balaton within ±7 days of MERIS overpass with acceptable accuracy. When considering the KdKVI surface Chl-*a* dataset only, there was no notable decrease in retrieval accuracy (Table A7). This is simply due to the consistently large sample size for the KdKVI validation dataset, which varied little with increasing temporal windows (n = 105, 113 and 113 for ±1, 3 and 7 days, respectively). As with Chl-*a*, PC concentrations were retrieved with comparable accuracy at all three time windows when the full dataset is considered (Figure 11; Table A3). Surprisingly, the lowest errors were typically for the largest temporal window of ±7 days, however validation with same day in situ data yielded the highest R^2 of 0.710 ($p < 0.0001$)

which deteriorated to R^2 = 0.433 (p < 0.0001) at ±7 days (Table A3). This limited change in pigment retrieval over time may be simply driven by the increase in sample numbers with the increase in temporal window (Table A3). Specifically for PC, all matchups with MERIS were from August and October, typically during a cyanobacteria bloom (Figure 16). For this dataset, the cyanobacteria population was typically formed of N_2-fixing species (e.g., *Cylindrospermopsis raciborskii*) which are known in Balaton to form persistent planktonic populations that can last up to 2 months [87]. In these conditions, cyanobacteria biomass is unlikely to change significantly over a 7-day period, which could explain the minimal difference in PC retrieval capability over the range of time windows. This indicates that there is little impact of the matchup time window for MERIS retrieval of PC during late summer cyanobacteria blooms (August–October), however in non-bloom conditions with lower and more variable concentrations of pigments, a shorter time window is optimal. Therefore, the Simis05 algorithm is likely limited to retrieve PC at larger time intervals (~7 days) only during cyanobacteria blooms, whereas reliable retrieval of PC during non-bloom conditions should be no more than ±1 day from the satellite overpass.

4.4. Impact of Dataset and Sampling Methods on Validation

MERIS-retrieved Chl-*a* from the Gons05 algorithm validated with the August 2010 surface samples only had the highest accuracy, with the lower errors for ±1 day matchups than when validated with either the BLI or KdKVI datasets (Tables A4–A6). The August 2010 Chl-*a* samples were intentionally collected at the surface during the dedicated field campaign for the purpose of satellite validation, and it is therefore unsurprising that these data align well with MERIS-retrieved Chl-*a*. PC was also retrieved with the highest accuracy within ±1 day when validated with August 2010 surface samples only as compared to BLI-integrated samples only (Tables A4 and A5). As sunlight penetration in turbid lakes can be very limited, it is expected that surface samples would be more comparable with satellite retrieved values. We found this to be the case in Lake Balaton, where low Secchi depths indicated turbid conditions (e.g., August 2010 mean Secchi depth = 0.74 m).

4.5. Sources of Error Explained with IOP Measurements

In the case of both Chl-*a* and PC retrieval, concentrations were overestimated by the models, and this is most notable for retrieval of Chl-*a*. This can be explained, in part, due to the application of constant $a^*_{ph}(665)$ and $a^*_{pc}(620)$ coefficients for calculating Chl-*a* and PC concentrations, respectively. These coefficients are taken from previous studies [34,37] and may not be representative of the phytoplankton community in Lake Balaton during the period of analysis, which was often dominated by the subtropical cyanobacterium *C. raciborskii.* The specific absorption coefficient (or absorption efficiency) may also vary across the lake or between phytoplankton communities due to differences in pigmentation, physiology, cell size and pigment packaging effects [88]. However, pigment packaging has a stronger impact on a^*_{ph} in the blue wavelengths with more consistent a^*_{ph} values in cyanobacteria over the red portion of the spectrum [89], therefore pigment packaging is unlikely to be the cause for variability in $a^*_{ph}(665)$. The mean measured $a^*_{ph}(665)$ during August 2010 in Lake Balaton was 0.0075 m^2 mg^{-1} Chl-*a* (±0.0015), a much lower value than employed by Gons et al. [34] (0.0139 m^2 mg^{-1} Chl-*a*). Subsequent studies have also investigated measurement of $a^*_{pc}(620)$ in laboratory algal cultures using a timed bleaching procedure, although the authors deemed this method unsuitable for mixed phytoplankton communities and thus it was not employed in this study [89]. More recently, $a^*_{pc}(625)$ was estimated by subtracting the effect of chlorophylls at 625 nm, and $a^*_{pc}(625)$ was found to vary widely for low PC concentrations (<10 mg m^{-3}) [90]. The phycocyanin specific absorption coefficient published in Simis and Kauko [89] was $a^*_{pc}(622)$ = 0.0071 m^2 mg^{-1} PC and the range recommended by Yacobi et al. [90] was $a^*_{pc}(625)$ = 0.007–0.008 m^2 mg^{-1} PC, which both agree with the value of 0.007 m^2 mg^{-1} PC recommended in Simis et al. [28,37] and was adopted in this study. However, given the lower value of $a^*_{ph}(665)$ measured in Lake Balaton, it is plausible that the published data for $a^*_{pc}(620)$ may also misrepresent the phytoplankton community in Lake Balaton.

The error in the Gons05 and Simis05 algorithms is further explained by the retrieval of the phytoplankton absorption coefficients. $a_{ph}(665)$ was overestimated compared to the measured values from the August 2010 sampling campaign (Figure 14). However, because the standard algorithm uses a higher $a^*_{ph}(665)$ coefficient (0.0139 m^2 mg^{-1} Chl-*a*) than that measured in Lake Balaton (mean = 0.0075 m^2 mg^{-1} Chl-*a*), the errors effectively cancel out, resulting in relatively good agreement between the measured and MERIS-retrieved Chl-*a* concentrations. Substitution of the $a^*_{ph}(665)$ coefficient measured in Lake Balaton (August 2010) into Gons05 actually results in greater errors, with Chl-*a* overestimated by a factor of 1.9 for matchups ±7 days (RMSE$_{log}$ = 0.621, Bias$_{log}$ = 0.579, MAPE = 347%; compare with results in Table A3). In fact, for the August 2010 dataset only, substitution of the Balaton $a^*_{ph}(665)$ results in an even greater overestimate by a factor of 2.5. Thus, using the site-specific $a^*_{ph}(665)$ measured in Balaton does not compensate for the fundamental issue with the Gons05 and Simis05 algorithms, i.e., overestimation of $a_{ph}(665)$, and results in a poorer Chl-*a* estimate. Mishra13- and Li15-retrieved $a_{ph}(\lambda)$ values were also overestimated, however these models still performed better for $a_{ph}(\lambda)$ retrieval than Simis05 (Figure A1). Given the poorer performance of Mishra13 and Li15 for PC retrieval, it suggests that these bio-optical models have errors in other model components which propagate to the PC estimates. It is also important to note that Lake Balaton is subject to changes in the relative contributions of IOPs as a result of wind-driven resuspension events [51]. However, MERIS-retrieved $a_{ph}(665)$ and $a_{ph}(620)$ on the sampling date following a wind event (26 August 2010, Stations 31–35) showed no difference in performance from the rest of the dataset. Thus, this suggests that the errors within the application of the Gons05 and Simis05 algorithms to Lake Balaton lie within the estimates of the IOPs ($a_{ph}(\lambda)$ and $b_b(\lambda)$) rather than the SIOP ($a^*_{ph}(\lambda)$).

The $b_b(\lambda)$ coefficient was estimated by the Simis05 algorithm at MERIS band 12, or 778.75 nm, as in Gons [33]. During the August 2010 campaign, backscattering was measured by an ECO-BB3, which does not measure explicitly at this wavelength (778.75 nm), however it is useful to examine the assumptions that are also employed in the Simis05 semi-analytical algorithm. From the in situ measurements, it was found that there was little difference in $b_b(\lambda)$ across the green and red wavelengths (Figure 15a), therefore it is a valid assumption that backscattering is spectrally neutral [33]. However, it was also shown that there is no significant linear relationship (R^2 < 0.001, p > 0.5) between $b_b(\lambda)$ measured in situ and that estimated with Simis05 (Figure 15b), and a similar result was found for Mishra13 and Li15 retrieved $b_b(\lambda)$ (Figure A2). Although the in situ backscattering is measured at a different wavelength, it is reasonable to expect to see a relationship between retrieved and measured $b_b(\lambda)$. As this is not the case, the $b_b(\lambda)$ estimation by the Simis05 algorithm is a potential source of error to PC retrieval from MERIS.

The measured particulate $b_b(532)$ in Lake Balaton during August 2010 varied over a wide range from 0.06 to 0.34 m^{-1}, with a high mean ratio of backscattering to total suspended matter ($b_b(532)$:TSM = 0.016). This was mainly due to the high percentage (~40–90%) of inorganic mineral sediments that were frequently resuspended in Lake Balaton. The distinctive biogeochemical composition and optical properties of Lake Balaton may then explain why the backscattering coefficient is not estimated well by the Gons05 and Simis05 algorithms. However, it is also worth noting that there were greater errors in the atmospheric correction of MERIS data for band 12, $R_{rs}(778.8)$ (see Figure 4l), the band at which Gons05 and Simis05 estimate b_b. Therefore, the atmospheric correction may also contribute to the poor estimation of b_b by the Gons05 and Simis05 algorithms.

When measured $b_b(650)$ is substituted into the Gons05 algorithm for each station in the August 2010 dataset, Chl-*a* is overestimated by a factor of 2.8 for a ±7-day matchup window (RMSE$_{log}$ = 0.376, Bias$_{log}$ = −0.341, MAPE=133%). Thus, as with substitution of $a^*_{ph}(665)$, applying the measured $b_b(650)$ actually results in higher errors than Chl-*a* retrieval using estimated b_b at MERIS band 12 and does not improve the algorithm (see Table A4). In the case that $b_b(\lambda)$ was measured at the same wavelength as MERIS band 12 (778.75 nm), the retrieved Chl-*a* concentrations would perhaps be more precisely compared. However, adjustment of the $b_b(\lambda)$ coefficient does not seem to appreciably alter the retrieved Chl-*a* concentrations from Gons05, or at least not as significantly as with the modification of $a^*_{ph}(665)$.

4.6. Applicability of Pigment Algorithms

Maps of MERIS-retrieved PC concentrations showed a peak in the western basins (26 May 2008; Figure 13), however this was not mirrored by cyanobacteria cell counts for this date (Figure 3g–h). It is known that algae containing chlorophyll *b* or *c* (1+2) also have a reflectance minimum (absorption peak) around the same wavelength as PC absorption (~640 nm). In particular, laboratory measurements of diatom cultures have reported a reflectance minimum (absorption peak) around the same wavelength as PC absorption (~640nm), possibly related to the high concentrations of chlorophyll *c* (1+2) found in diatoms [30]. Thus, it could be that in large enough numbers (i.e., a diatom bloom), the pigments of other algal groups are incorrectly interpreted by the Simis05 algorithm as PC. Indeed, BLI routine monitoring data indicate an increase in total diatom biomass in Keszthely (western basin) between 7 May and 16 June 2008 (100–759 mg m^{-3}) while nitrogen-fixing cyanobacteria biomass was low (0–60 mg m^{-3}), therefore this diatom peak may be responsible for the false positive for PC retrieval. Similarly, Simis et al. [37] found that PC was significantly overestimated in the presence of chlorophyll *b* and fucoxanthin, particularly for low concentrations of PC. This suggests that the Simis05 algorithm performs better for retrieval of PC during low diatom abundance.

There was also a distinct lower limit for MERIS-retrieved pigment concentrations (Figure 8; Figure 11). This is perhaps most noticeable on the PC plots, and especially in the figures for ±7-day matchups. This may suggest differences in the water sampling methods, where the routine monitoring data for Chl-*a* is derived from to depth-integrated samples but the August 2010 Chl-*a* data were determined from surface samples. The effect is a much higher accuracy for the August 2010 samples as opposed to the routine monitoring (Figure 8). Additionally, the euphotic depth (Z_{eu}) in Lake Balaton in August 2010 was typically less than 1.7 m. Thus, it is not surprising that MERIS-retrieved pigment estimates show stronger agreement with surface rather than depth-integrated pigment concentrations.

Quite possibly, this minimum in retrieved pigments is due to limitations of MERIS itself or the Gons05 and Simis05 algorithms. The retrieved pigment concentrations from Gons05 and Simis05 correspond well with the interannual patterns of in situ Chl-*a*, phytoplankton biomass and cyanobacteria biomass (Figure 16). However, it was noted that the most accurate retrieved concentrations were in the western basin, where Chl-*a* is generally higher (<60 mg m^{-3}) than in the east of Lake Balaton (Tihany), where Chl-*a* levels are predominantly lower (<20 mg m^{-3}). The Gons05 algorithm may therefore not be appropriate for Chl-*a* retrieval in waters with low phytoplankton biomass (e.g., where Chl-*a* <10 mg m^{-3}). This is further supported by the poor accuracy and high scatter for retrieval when Chl-*a* is low (Figure 8). Gons et al. [91] found that MERIS-retrieved Chl-*a* values were poorly predicted by Gons05 at Chl-*a* <5 mg m^{-3} in oligotrophic lakes, which was attributed to weak emerging light flux, masking of the red Chl-*a* absorption peak by absorption of water, and an increased influence of Chl-*a* fluorescence that is not taken into account by the algorithm. Domínguez Gómez et al. [68] applied a different version of the Normalised Difference Chlorophyll Index (NDCI) for values of Chl-*a* greater or less than a threshold of 17 mg m^{-3}, as water-leaving reflectance is very low for Chl-*a*<17 mg m^{-3}. This is further corroborated by a more recent study where the fluorescence line height (FLH) algorithm retrieved Chl-*a* from MERIS most accurately at higher Chl-*a* concentrations [54]. It is therefore possible that Gons05 and Simis05 algorithms have limited applicability. These algorithms were designed for and perform optimally in eutrophic lakes as opposed to mesotrophic or oligotrophic waters, and are insensitive to changes in Chl-*a* below 10 mg m^{-3}. In addition, PC retrieval with Simis05 may also be limited by interference of chlorophyll-*b*, particularly when cyanobacteria biomass is low and the community is dominated by other eukaryotic phytoplankton groups, and this could be a factor in the overestimation of PC when Chl-*a* is low [37]. Indeed, in this study the Mishra13_Simis algorithm was found to outperform Simis05 at low PC concentrations (Table A2). It may be that the Simis05 and Gons05 algorithms are better considered as part of an ensemble approach, as a single algorithm is unlikely to work in all conditions, even within a single lake system.

5. Conclusions

With the recent launch of ESA's Sentinel-3 Ocean and Land Colour Instrument (OLCI), it is an opportune time to explore the historic archive of data from satellite instruments, such as MERIS, to investigate and improve retrieval of inland water quality parameters. In recent years, there has been vast improvement to both the availability of ESA instrument data (e.g., MERCI and EUMETSAT systems) and the tools for data analysis (e.g., Beam, SNAP), and this has allowed for corresponding advances in the remote sensing of inland waters. There have also been developments in the atmospheric correction of MERIS data, including this study as the first instance of implementing SCAPE-M_B2 for the correction of MERIS FRS data. Further measurements of the optical properties of inland waters and the corresponding development of bio-optical models have also been advanced with the accessibility of MERIS and OLCI data and associated analytical tools.

Specific to this study, the Simis05 and Gons05 semi-analytical algorithms performed well to retrieve PC and Chl-*a* pigment concentrations in Lake Balaton, respectively. Phytoplankton and cyanobacteria biomass were also retrieved, albeit to a lesser degree of accuracy, demonstrating a critical step towards the monitoring medium-term changes in cyanobacteria abundance in lakes. Accuracy for both Chl-*a* and PC retrieval was fairly consistent within ±7 days of the MERIS overpass, however it remains imperative that samples for pigment analysis are collected as temporally close as possible to the date and time of the satellite overpass. Additionally, pigment retrieval from Gons05 and Simi05 algorithms matched the interannual patterns in Chl-*a* and biomass, although in situ data were better predicted in the more eutrophic western basin than the meso- to oligotrophic eastern basin of Lake Balaton. Although Gons05 and Simis05 performed well in Lake Balaton, in situ IOP measurements crucially highlighted the sources of error in estimates of $a_{ph}(\lambda)$ and $b_b(\lambda)$, precursors to pigment retrieval in these algorithms. Estimated $b_b(\lambda)$ was found to be unrelated to measured $b_b(\lambda)$, while a_{ph} at 665 and 620 nm was overestimated by the algorithms, then subsequently overcorrected using a higher $a^*_{ph}(665)$ than that measured in Lake Balaton. While substituting Balaton specific IOPs into the algorithms did not improve retrieval of Chl-*a*, knowledge of the IOPs did elucidate the sources of error within the Gons05 and Simis05 algorithms.

This study provides additional evidence for the success of Chl-*a* and PC retrieval algorithms using MERIS data in an emblematic European lake, with a view to facing new challenges for remote sensing of inland waters with the recently launched OLCI at match-up timescales of up to 7 days. Furthermore, this study proves the effectiveness of collecting IOP measurements alongside pigment and satellite data, in order to analyse the performance of pigment retrieval algorithms and the ability to tune a semi-analytical algorithm to the waterbody of interest. Other recent studies have promoted the effectiveness of simple empirical models for retrieval of phytoplankton pigments, however we emphasize the utility of semi-analytical models which account for variations in phytoplankton IOPs on the water-leaving signal. Through continued characterisation of bio-geo-optical properties, future work can generate more informed decisions on the parameterisation of pigment retrieval algorithms and achieve model transferability among inland waters models, e.g., via OWTs [78,81]. There remains abundant opportunity for future work on the success of water quality algorithms for MERIS and OLCI, and the ability to accurately retrieve parameters in a range of inland waters within the constraints of contemporary satellite instruments.

Author Contributions: Conceptualization, C.A.L.R. and P.D.H.; Data curation, C.A.L.R., J.A.D.G., V.M.-V., H.H. and L.V.; Formal analysis, C.A.L.R., J.A.D.G. and V.M.-V.; Funding acquisition, P.D.H. and A.N.T.; Investigation, C.A.L.R., M.P., A.W.K., E.Z. and A.N.T.; Methodology, H.H.; Project administration, A.N.T.; Resources, M.P., A.W.K. and A.N.T.; Supervision, P.D.H. and A.N.T.; Validation, C.A.L.R.; Visualization, C.A.L.R.; Writing—original draft, C.A.L.R.; Writing—review and editing, C.A.L.R., P.D.H., J.A.D.G., V.M.-V and A.N.T.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A. MERIS Phycocyanin Algorithms

Appendix A.1. Dekker93

Dekker93 is the earliest known method for estimating PC concentrations in lakes, calibrated with data from eutrophic Vecht lakes in the Netherlands [31]. This algorithm is a semi-empirical baseline approach which uses R_{rs} at 600, 648 and 624 nm, however it can be adapted to MERIS and OLCI bands centred at 560 nm, 620 nm and 665 nm as follows:

$$\text{PC}\left(\text{mg m}^{-3}\right) \propto 0.5 \times \left[(R_{rs}(560) + R_{rs}(665)) - R_{rs}(620)\right]. \tag{A1}$$

In this study, Dekker93 was also modified to include an additional term at a fourth wavelength at which absorption is dominated by water and is minimally sensitive to absorption by pigments, minerals, detritus and CDOM [23,76]. This is herein referred to as Dekker93_modified and is denoted as follows:

$$\text{PC}\left(\text{mg m}^{-3}\right) \propto 0.5 \times \left[(R_{rs}(560) + R_{rs}(665)) - R_{rs}(620)\right] - R_{rs}(754). \tag{A2}$$

Appendix A.2. Schalles00

Schalles00 is a simple band-ratio algorithm that employs a ratio of the reflectance peak at 650 nm to the trough at 625 nm [32]. Schalles00 was calibrated over Carter Lake, USA, a eutrophic groundwater-fed lake. Although not specifically developed for MERIS, this band-ratio can be adapted to MERIS and OLCI bands centred at 665 nm and 620 nm as follows:

$$\text{PC}\left(\text{mg m}^{-3}\right) \propto \frac{R_{rs}(665)}{R_{rs}(620)}. \tag{A3}$$

Appendix A.3. Gons05 and Simis05

The Gons05 algorithm is naturally considered here for retrieval of Chl-*a*, as it is the foundation upon which the Simis05 PC algorithm evolved [28]. Specifically, the Gons05 algorithm calculates $a_{Chla}(665)$, and from this the summed absorption of Chl-*a* and PC at 620 nm is obtained. $a_{Chla}(620)$ is subtracted from summed Chl-*a* and PC pigment absorption at 620 nm in order to obtain $a_{PC}(620)$.

Gons05 was calibrated in the shallow eutrophic freshwaters of Ijssel Lagoon, Netherlands (Chl-*a* = 4-185 mg m^{-3}), with initial validation in well-mixed and optically deep lakes in The Netherlands, the Scheldt Estuary (The Netherlands and Belgium), Lake Taihu (China), the Hudson/Raritan Estuary (USA) and the North Sea (Belgian coast) [19,33,34]. As a progression of Gons05, Simis05 was initially calibrated in the well-mixed and eutrophic Lakes Loosdrecht (Chl-*a* = 48–98 mg m^{-3}, PC = 22–80 mg m^{-3}) and IJsselmeer (Chl-*a* = 23–92 mg m^{-3}, PC = 0.8–65 mg m^{-3}) in The Netherlands [28], and has subsequently obtained accurate retrieval of PC in mostly eutrophic inland waters with moderate to high cyanobacterial biomass, including Spanish lakes and reservoirs [37], Indiana reservoirs [26,29,36], Spanish and Dutch lakes and reservoirs [27], shallow eutrophic UK lakes [23] and eutrophic lakes in East China [22]. The Simis05 algorithm in particular has undergone limited testing with independent datasets of inland waters, and there is a need for validation in waters with differing optical properties.

The Simis05 algorithm can be applied to MERIS or OLCI data using a commonly acknowledged relationship between inherent optical properties and reflectance [92], where backscattering (b_b) is

assumed to be spectrally neutral and is derived from R_{rs} at a single wavelength in the near infra-red (NIR) [33], as detailed in Gons et al. [34]:

$$b_b(779) = \frac{1.61 \times R_{rs}(779)}{0.082 - 0.6 \times R_{rs}(779)}. \tag{A4}$$

Total absorption at a particular wavelength can then be calculated from a reflectance ratio [$R_{rs}(\lambda_1)$/ $R_{rs}(\lambda_2)$], b_b and absorption at λ_2 [$a(\lambda_2)$]:

$$a(\lambda_1) = \frac{R_{rs}(\lambda_2)}{R_{rs}(\lambda_1)} \times [a(\lambda_2) + b_b] - b_b, \tag{A5}$$

where the reflectance ratio of $\lambda_1 = 665$ nm and $\lambda_2 = 709$ nm is effective for retrieval of Chl-a for Gons05 [19,31,93] and $\lambda_1 = 620$ nm and $\lambda_2 = 709$ nm is applied for the retrieval of PC for Simis05 [28]. The Gons05 algorithm for Chl-a absorption can then be described as a function of the R_{rs} ratio, the absorption coefficient of water (a_w) and b_b. A correction factor, γ, was also introduced by Simis et al. [28] to relate the R_{rs} ratio to measured pigment absorption, thus obtaining the following algorithm for Chl-a absorption, modified from Gons et al. [19,34]:

$$a_{Chla}(665) = \left[\left(\frac{R_{rs}(709)}{R_{rs}(665)} \times (a_w(709) + b_b) - b_b - a_w(665) \right) \right] \times \gamma^{-1}, \tag{A6}$$

where $a_w(709) = 0.727$ m^{-1}, $a_w(665) = 0.401$ m^{-1} and $\gamma = 0.68$.

The Chl-a concentration can then be calculated by dividing the Chl-a absorption coefficient by the specific absorption coefficient, $a^*_{Chla}(665)$:

$$\text{Chl} - a \left(\text{mg m}^{-3} \right) = \frac{a_{Chla}(665)}{a^*_{Chla}(665)}, \tag{A7}$$

where $a^*_{Chla}(665)$ is 0.0139 m^{2}mg^{-1} for uncorrected Chl-a [34].

It is assumed that PC and Chl-a comprise absorption at 620 nm, as apparent from reflectance spectra of cyanobacteria-dominated waters. Thus, in order to estimate PC absorption alone at this wavelength, summative pigment absorption at 620 nm is calculated first. A factor δ is introduced for the correction of $a(620)$, and the summed absorption of PC and Chl-a at 620 nm can be estimated as:

$$a_{Chla}(620) + a_{pc}(620) = \left[\frac{R_{rs}(709)}{R_{rs}(620)} \times (a_w(709) + b_b) - b_b - a_w(620) \right] \times \delta^{-1}, \tag{A8}$$

where $a_w(620) = 0.281$ m^{-1} and $\delta = 0.84$.

Absorption by PC can then be derived by subtracting the absorption by Chl-a at 620 nm, using a conversion factor, ε, which relates in vivo absorption by Chl-a at 665 nm to its absorption at 620 nm:

$$a_{pc}(620) = \left[\frac{R_{rs}(709)}{R_{rs}(620)} \times (a_w(709) + b_b) - b_b - a_w(620) \right] \times \delta^{-1} - (\varepsilon \times a_{Chla}(665)), \tag{A9}$$

where $\varepsilon = 0.24$.

Finally, the concentration of PC can then be calculated by dividing the solution PC absorption coefficient by the specific absorption coefficient, $a^*_{pc}(620)$:

$$\text{PC} \left(\text{mg m}^{-3} \right) = \frac{a_{pc}(620)}{a^*_{pc}(620)}, \tag{A10}$$

where $a^*_{pc}(620)$ is 0.007 m^{2}mg^{-1} [37].

Appendix A.4. Hunter10_Duan

Hunter et al. [23] developed a semi-analytical 3-band model, based on the theory from Dall'Olmo et al. [76]. Hunter10 was calibrated using two shallow eutrophic lakes in the UK, Loch Leven and Esthwaite Water, and has since then been adapted to MERIS bands by Duan et al. [22] and validated in Lake Taihu, China. The Hunter10 algorithm can be implemented for MERIS and OLCI data as follows:

$$Chl - a \left(mg \ m^{-3} \right) \propto \left[R_{rs}^{-1}(\lambda_1) - R_{rs}^{-1}(\lambda_2) \right] \times R_{rs}(\lambda_3), \tag{A11}$$

where λ_1 = 620 nm, λ_2 = 709 nm and λ_3 = 753 nm, as in Duan et al. [22].

Appendix A.5. Mishra13

As an inversion model, the Mishra13 algorithm builds upon the QAA algorithm for Chl-*a* retrieval [83], which was developed for oceans and has been widely validated in ocean, coastal and more recently inland waters. The Mishra13 model was calibrated for PC retrieval in turbid and highly productive aquaculture ponds [39]. The wavelengths used for Mishra13 were adapted to coincide with MERIS band centres (and can be similarly implemented for OLCI), and the subsequent retrieval of PC concentrations is summarised below.

Above surface remote sensing reflectance (R_{rs}) is first converted to subsurface remote sensing reflectance (r_{rs}) using the following equation:

$$r_{rs} = R_{rs} / (0.52 + 1.7 R_{rs}). \tag{A12}$$

r_{rs} is a function of u, which is defined as the ratio of the backscattering coefficient (b_b) to the sum of total absorption (a) and backscattering coefficients:

$$u(\lambda) = \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}. \tag{A13}$$

u is then empirically derived from r_{rs} as in Gordon et al. [92]:

$$u(\lambda) = \frac{-g_0 + \sqrt{(g_0)^2 + 4 g_1 r_{rs}(\lambda)}}{2 g_1}, \tag{A14}$$

where g_0 = 0.089 and g_1 = 0.125.

Total absorption coefficients can then be estimated at a reference wavelength (λ_0) as a function of the absorption coefficient of water [$a_w(709)$]:

$$a(\lambda_0) = a_w(709) + 10^{-0.8125 - 2.3404\chi + 1.24\chi^2}, \tag{A15}$$

where λ_0 = 708 nm, as parameterised for turbid and productive waters in Mishra et al. (2014), $a_w(709)$ = 0.7204 and

$$\chi = log_{10}\left(\frac{0.01 * r_{rs}(442.5) + r_{rs}(620)}{r_{rs}(708.75) + 0.005 * \frac{r_{rs}(620)}{r_{rs}(442.5)} * r_{rs}(620)} \right). \tag{A16}$$

Next, the particulate backscattering coefficient at the reference wavelength [$b_{bp}(\lambda_0)$] is retrieved as follows:

$$b_{bp}(\lambda_0) = \frac{u(\lambda_0) a(\lambda_0)}{1 - u(\lambda_0)} - b_{bw}(\lambda_0). \tag{A17}$$

The particulate backscattering coefficients at other wavelengths [$b_{bp}(\lambda)$] are then estimated from $b_{bp}(\lambda_0)$:

$$b_{bp}(\lambda) = b_{bp}(\lambda_0)\left(\frac{\lambda_0}{\lambda}\right)^{\eta}, \tag{A18}$$

where the spectral power, η, is empirically estimated as:

$$\eta = 2.0\left\{1 - 1.2exp\left[-0.9\frac{r_{rs}(442.5)}{r_{rs}(560)}\right]\right\}. \tag{A19}$$

The total absorption coefficient [$a(\lambda)$] can thus be calculated as:

$$a(\lambda) = \frac{(1 - u(\lambda))\big(b_{bw}(\lambda) + b_{bp}(\lambda)\big)}{u(\lambda)}. \tag{A20}$$

The total absorption coefficient is then further decomposed into the combined absorption by coloured dissolved organic matter (CDOM) and coloured detrital matter [$a_{CDM}(\lambda)$] and phytoplankton absorption [$a_{ph}(\lambda)$]:

$$a_{CDM}(443) = \frac{[a(412.5) - \zeta a(442.5)] - [a_w(411) - \zeta a_w(411)]}{\xi - \zeta}, \tag{A21}$$

where

$$\zeta = \frac{a_{ph}(411)}{a_{ph}(443)} = 0.74 + \frac{0.2}{0.8 + r_{rs}(442.5)/r_{rs}(560)}, \tag{A22}$$

$$\xi = \frac{a_{CDM}(411)}{a_{CDM}(443)} = e^{S(443-411)}, \tag{A23}$$

and $a_w(411) = 0.0068$.

$a_{CDM}(\lambda)$ can then be calculated using the exponential function:

$$a_{CDM}(\lambda) = a_{CDM}(443)e^{-S(\lambda-443)}, \tag{A24}$$

where the slope S is 0.02 nm^{-1}. $a_{ph}(\lambda)$ is subsequently calculated by subtraction of $a_w(\lambda)$ and $a_{CDM}(\lambda)$ from total absorption:

$$a_{ph}(\lambda) = a(\lambda) - a_w(\lambda) - a_{CDM}(\lambda). \tag{A25}$$

It is assumed that $a_{ph}(\lambda)$ is approximately equal to absorption by chlorophyll-*a* and phycocyanin [$a_{ph}(\lambda) \approx a_{chla}(\lambda) + a_{pc}(\lambda)$] at $\lambda = 665$ and 620 nm, and $a_{pc}(620)$ is retrieved as:

$$a_{pc}(620) = \frac{\psi_1 a_{ph}(620) - a_{ph}(665)}{\psi_1 - \psi_2}, \tag{A26}$$

where $\psi_1 = a_{chla}(665)/a_{chla}(620)$ and $\psi_2 = a_{pc}(665)/a_{pc}(620)$.

Finally, PC concentrations are calculated by dividing $a_{pc}(620)$ by the specific absorption coefficient of PC, $a^*_{pc}(620)$:

$$\text{PC}\left(\text{mg m}^{-3}\right) = \frac{a_{pc}(620)}{a*_{pc}(620)}. \tag{A27}$$

For the application of Mishra13 in this study, PC concentration was calculated using two different values of $a^*_{pc}(620)$ from Mishra et al. [39] and Simis et al. [37] (0.0048 m^2 mg^{-1} and 0.007 m^2 mg^{-1}, respectively, herein referred to as Mishra13 and Mishra13_Simis).

Appendix A.6. Qi14

Qi14, or the PCI, is a band subtraction model which was calibrated in Lake Taihu, China [40]. It has subsequently been tested in three Indiana reservoirs, USA, Lake IJsselmeer, the Netherlands, and eight Chinese lakes and reservoirs, however no correlation was found with in situ PC values (R^2 = 0.028) [42]. This algorithm was applied in this study for MERIS data (and can be similarly implemented for OLCI) as follows:

$$\text{PCI} = \left(R_{rs}(560) + \frac{620 - 560}{665 - 560} \times [R_{rs}(665) - R_{rs}(560)]\right) - R_{rs}(620), \tag{A28}$$

where PC concentration can be calculated from the PCI according to an exponential relationship of the form:

$$\text{PC}\left(\text{mg m}^{-3}\right) = ae^{b*PCI}. \tag{A29}$$

In Qi et al. [40], coefficients were defined as a = 3.87 and b=1154 for Lake Taihu (herein referred to as Qi14), however these coefficients were also calibrated to Lake Balaton 2010–2011 PC dataset as a = 21.26 and b = −139.3 (herein referred to as Qi14_Balaton).

Appendix A.7. Li15

The Li15 algorithm is an extension of the IIMIW model first presented in Li et al. [26] for the retrieval of Chl-*a*. Li15 was calibrated with three central Indiana reservoirs Li et al. [41], however to our knowledge at the time of writing the Li15 algorithm has not been validated with independent datasets. The steps for the application of the Li15 are as in Table 3 of Li et al. [41], with wavelengths adapted to correspond with MERIS band centres (and can be similarly implemented for OLCI), as summarised below.

Above surface remote sensing reflectance (R_{rs}) is first converted to subsurface remote sensing reflectance (r_{rs}) using the following equation, as in Li et al. [41]:

$$r_{rs}(\lambda) = R_{rs}(\lambda)/0.54. \tag{A30}$$

The backscattering coefficient at MERIS band 12 (or for OLCI, band 16) [$b_b(778.75)$] is then estimated as a function of $r_{rs}(778.75)$ and absorption by pure water [$a_w(778)$]:

$$b_b(778.75) = \frac{r_{rs}(778.75)a_w(778)}{0.082 - r_{rs}(778.75)}. \tag{A31}$$

The spectral power, η, is empirically estimated as:

$$\eta = 2.0\left\{1 - 1.2exp\left[-0.9\frac{r_{rs}(442.5)}{r_{rs}(560)}\right]\right\}. \tag{A32}$$

The particulate backscattering coefficient at 560 nm [$b_{bp}(560)$] can then be retrieved as follows:

$$b_{bp}(560) = [b_{bp}(778.75) - b_{bw}(778.75)]/(0.7198)^\eta, \tag{A33}$$

and the backscattering coefficients at other wavelengths [$b_b(\lambda)$] are then estimated from $b_{bp}(560)$:

$$b_b(\lambda) = b_{bp}(560)\left(\frac{560}{\lambda}\right)^\eta + b_{bw}(\lambda). \tag{A34}$$

Subsequently, total minus water absorption $[a_{t-w}(\lambda)]$, or absorption of particulate (p) and dissolved (g) matter $[a_{pg}(\lambda)]$, can be calculated as:

$$a_{t-w}(\lambda) = \frac{r_{rs}(708.75)b_b(\lambda)[a_w(708.75) + b_b(708.75)]}{r_{rs}(\lambda)b_b(708.75)} - b_b(\lambda) - a_w(\lambda), \qquad (A35)$$

At this point in the algorithm steps, Chl-*a* concentration can be estimated using three different approaches, as outlined in Li et al. [26]. Total absorption is further partitioned into $a_{CDM}(\lambda)$ and $a_{ph}(\lambda)$ by first calculating the in vivo phytoplankton absorption without the contribution from PC $[a_{ph\text{-}pc}(\lambda)]$:

$$a_{ph-pc}(\lambda) = 1.1872 * C1(\lambda)a_{t-w}(665) + C2(\lambda), \qquad (A36)$$

where C1 and C2 are the wavelength dependent regression coefficients outlined in Table A1 of Appendix B in Li et al. (2015).

The combined absorption of CDOM, NAP and PC $[a_{CDM+pc}(\lambda)]$ is then calculated by subtraction:

$$a_{CDM+pc}(\lambda) = a_{t-w}(\lambda) - a_{ph-pc}(\lambda). \qquad (A37)$$

Absorption by coloured dissolved matter (CDM) $[a_{CDM}(\lambda)]$ can then be calculated using the spectral slope of CDM (S_{CDM}):

$$a_{CDM}(\lambda) = a_{CDM}(412.5)\exp[-S_{CDM} \times (\lambda - 412.5)], \qquad (A38)$$

where it is assumed $a_{CDM}(412.5){=}a_{CDM+pc}(412.5)$ and S_{CDM} was set to a constant value of 0.020 nm^{-1}, the mean measured value of S_{CDOM} for Lake Balaton in August 2010. $a_{CDM}(708.75)$ is set to 0, with $a_{CDM}(\lambda)$ values adjusted accordingly.

The phytoplankton absorption coefficients $[a_{ph}(\lambda)]$ can then be calculated by subtraction:

$$a_{ph}(\lambda) = a_{t-w}(\lambda) - a_{CDM}(\lambda). \qquad (A39)$$

The phycocyanin absorption coefficient $[a_{pc}(620)]$ is then calculated as the difference between phytoplankton absorption and non-PC phytoplankton absorption at 620 nm:

$$a_{pc}(620) = a_{ph}(620) - a_{ph-pc}(620). \qquad (A40)$$

Finally, PC concentration is calculated by division of $a_{pc}(620)$ by the PC specific absorption coefficient $[a*_{pc}(620)]$:

$$\text{PC}\left(\text{mg m}^{-3}\right) = a_{pc}(620)/a*_{pc}(620), \qquad (A41)$$

where $a*_{pc}(620)$ is set to 0.0046 m^{2}mg^{-1} as in Li et al. [41] (herein referred to as Li15) or 0.007 m^{2}mg^{-1} as in Simis et al. [37] (herein referred to as Li15_Simis).

Appendix A.8. Liu18

Liu18 is a 4 band semi-analytical model, which builds upon the approach by Hunter et al. [23]. This was calibrated using three Indiana reservoirs, USA, Lake IJsselmeer, the Netherlands, and eight Chinese lakes and reservoirs [42]. To our knowledge at the time of writing, Liu18 has not been validated with independent datasets. Liu18 was implemented in this study by calculation of the FBA_PC index for MERIS data (and can be similarly implemented for OLCI) as follows:

$$\text{FBA}_{\text{PC}} = \left[\frac{1}{R_{rs}(620)} - \frac{0.4}{R_{rs}(560)} - \frac{0.6}{R_{rs}(709)}\right] \times R_{rs}(754), \qquad (A42)$$

where PC concentration is then calculated according to linear least squares regression:

$$\text{PC}\left(\text{mg m}^{-3}\right) \;=\; m \times \text{FBA}_{\text{PC}} + B. \tag{A43}$$

The coefficients for a regression to relate PC concentration to the FBA_PC were as in Liu et al. [42] (m = 462.5 and B = 22.598, herein referred to as Liu18) as well as those calibrated to the 2010–2011 dataset for Lake Balaton (m = 76.7 and B = 23.09, herein referred to as Liu18_Balaton).

Appendix B. Validation Assessment Results

Table A1. Accuracy assessment results for PC retrieval algorithms using MERIS data from 2010–2011 (n = 22), including intercept (*b*), slope (*m*), coefficient of determination (R^2) and p-value (*p*) for a linear least squares regression (y = *m*x+*b*), and error metrics Root Mean Square Error (RMSE), log-space RMSE (RMSE$_{log}$), Bias, log-space Bias (Bias$_{log}$), Mean Absolute Percentage Error (MAPE), Median Absolute Percentage Error (MdAPE) and Symmetric Mean Absolute Percentage Error (SMAPE). Matchups are ± 1 day of *in situ* PC measurements. Bold values indicate lowest error for all algorithms.

Model	*b* mg m^{-3}	*m*	R^2	*p*	RMSE mg m^{-3}	RMSE $_{log}$	Bias mg m^{-3}	Bias $_{log}$	MAPE %	MdAPE %	SMAPE %
Dekker93	18.1	0.218	0.0992	0.153	21.6	0.444	3.94	0.148	224	84.4	88.3
Dekker93_modified	15.6	0.172	0.576	<0.0001	17.5	0.354	0.614	0.0542	142	69.3	72.3
Schalles00	4.58	0.822	0.595	<0.0001	14.6	0.370	1.35	0.238	124	111	112
Simis05	8.95	0.722	0.710	<0.0001	**11.8**	0.272	3.92	0.147	77.0	**50.8**	**48.2**
Hunter10_Duan12	17.9	0.160	0.662	<0.0001	17.8	0.356	2.66	0.107	153	114	77.0
Mishra13	17.3	0.0522	0.00836	0.686	22.9	0.330	**0.0987**	0.230	104	87.0	66.1
Mishra13_Simis	11.8	0.0358	0.00836	0.686	22.3	**0.246**	−5.62	0.0664	**61.2**	53.2	58.2
Qi14	632	−7.68	0.0910	0.172	715	1.55	475	1.38	6197	3512	158
Qi14_Balaton	12.6	0.0211	0.0433	0.352	21.1	0.403	−5.18	0.0563	101	52.8	64.4
Li15	−26.2	2.55	**0.716**	<0.0001	46.3	0.503	1.84	**0.0349**	205	170	141
Li15_Simis	−17.2	1.67	**0.716**	<0.0001	26.4	0.523	−5.00	−0.147	150	131	136
Liu18	**−6.40**	**0.906**	0.634	<0.0001	16.5	0.841	−8.10	−0.305	184	122	134
Liu18_Balaton	18.3	0.150	0.634	<0.0001	18.0	0.354	2.88	0.088	155	113	78.2

Table A2. Accuracy assessment results for PC retrieval algorithms using MERIS data from 2010–2011 for low phycocyanin concentrations (PC<50 mg m^{-3}, n = 19), including intercept (*b*), slope (*m*), coefficient of determination (R^2), p-value (*p*), Root Mean Square Error (RMSE), log-space RMSE (RMSE$_{log}$), Bias, log-space Bias (Bias$_{log}$), Mean Absolute Percentage Error (MAPE), Median Absolute Percentage Error (MdAPE) and Symmetric Mean Absolute Percentage Error (SMAPE). Matchups are ± 1 day of *in situ* PC measurements. Bold values indicate lowest error for all algorithms.

Model	*b* mg m^{-3}	*m*	R^2	*p*	RMSE mg m^{-3}	RMSE $_{log}$	Bias mg m^{-3}	Bias $_{log}$	MAPE %	MdAPE %	SMAPE %
Dekker93	15.0	0.543	0.0719	0.267	17.84	0.467	10.2	0.295	251	92.0	90.5
Dekker93_modified	15.7	0.187	0.281	0.0195	9.52	0.328	7.11	0.198	155	86.6	69.5
Schalles00	−4.16	1.78	0.554	0.117	13.57	0.405	4.01	0.357	139	124	124
Simis05	**−0.146**	1.69	0.793	<0.0001	10.7	0.286	7.09	0.191	85.1	53.0	51.1
Hunter10_Duan12	15.5	0.407	0.770	<0.0001	10.4	0.346	9.28	0.255	168	134	76.0
Mishra13	9.66	**0.793**	0.440	<0.001	9.76	0.355	7.69	0.296	119	110	62.8
Mishra13_Simis	4.90	0.715	0.605	<0.001	**5.14**	**0.224**	**1.90**	**0.120**	**59.5**	**36.3**	**48.2**
Qi14	734	−18.3	0.0588	0.317	766	1.66	532	1.54	7144	4595	173
Qi14_Balaton	12.3	0.0452	0.0225	0.539	7.66	0.339	2.28	0.172	105	43.8	54.0
Li15	−26.0	2.82	0.697	<0.0001	20.2	0.561	−6.90	0.00889	222	171	155
Li15_Simis	−17.1	1.85	0.697	<0.0001	13.6	0.587	−8.14	−0.173	162	146	147
Liu18	−21.9	2.51	**0.814**	<0.0001	15.3	0.965	−5.96	−0.349	208	138	149
Liu18_Balaton	15.7	0.417	**0.814**	<0.0001	10.6	0.340	9.56	0.247	170	129	77.4

Table A3. Least squares linear regression results (y = mx+b) for Chl-*a*, PC, a_{ph} and b_b retrieval using Gons05 or Simis05 algorithms for all data, including integrated and surface samples.

	Match-up Interval	n	x	y	b mg m^{-3}	m	R^2	p	RMSE mg m^{-3}	RMSE log	Bias mg m^{-3}	Bias log	MAPE %	MdAPE %	SMAPE %
PC	1 day	22	Measured PC	MERIS Retrieved PC	8.95	0.722	0.710	<0.0001	11.8	0.272	3.92	0.147	77.0	50.8	48.2
	3 day	30	Measured PC	MERIS Retrieved PC	7.44	0.777	0.663	<0.0001	11.5	0.261	3.77	0.110	71.0	50.6	48.5
	7 day	40	Measured PC	MERIS Retrieved PC	9.41	0.443	0.433	<0.0001	18.3	0.273	−1.30	0.0218	63.0	49.5	49.8
	1 day	33	Cyanobacteria biomass	MERIS Retrieved PC	7.72	0.00494	0.462	<0.0001	-	-	-	-	-	-	-
	3 day	41	Cyanobacteria biomass	MERIS Retrieved PC	5.66	0.00539	0.525	<0.0001	-	-	-	-	-	-	-
	7 day	64	Cyanobacteria biomass	MERIS Retrieved PC	6.18	0.00488	0.494	<0.0001	-	-	-	-	-	-	-
Chl-*a*	1 day	136	Measured Chl-*a*	MERIS Retrieved Chl-*a*	8.98	1.11	0.801	<0.0001	11.9	0.394	10.4	0.329	151	86.4	68.5
	3 day	156	Measured Chl-*a*	MERIS Retrieved Chl-*a*	8.53	1.11	0.803	<0.0001	11.5	0.382	9.97	0.319	143	82.2	66.6
	7 day	194	Measured Chl-*a*	MERIS Retrieved Chl-*a*	8.56	1.06	0.767	<0.0001	11.2	0.369	9.38	0.296	132	74.2	63.0
	1 day	28	Phytoplankton biomass	MERIS Retrieved Chl-*a*	6.42	0.00527	0.667	<0.0001	-	-	-	-	-	-	-
	3 day	40	Phytoplankton biomass	MERIS Retrieved Chl-*a*	6.62	0.00515	0.698	<0.0001	-	-	-	-	-	-	-
	7 day	61	Phytoplankton biomass	MERIS Retrieved Chl-*a*	7.43	0.00480	0.636	<0.0001	-	-	-	-	-	-	-

Table A4. Least squares linear regression results (y = mx+b) for Chl-a, PC, a_{ph}, $a_{Chla+PC}$, and b_b retrieval using Gons05 or Simis05 algorithms, validated with surface samples only (August 2010 only).

	Match-up Interval	n	x	y	b mg m^{-3} or m^{-1}	m	R^2	p	RMSE mg m^{-3} or m^{-1}	RMSE log	Bias mg m^{-3} or m^{-1}	Bias log	MAPE %	MdAPE %	SMAPE %
PC	1 day	14	Measured PC	MERIS Retrieved PC	−1.28	1.77	0.836	<0.0001	11.7	0.254	7.79	0.174	71.6	61.2	49.9
	3 day	22	Measured PC	MERIS Retrieved PC	−4.15	1.87	0.799	<0.0001	11.3	0.246	6.18	0.189	65.4	52.6	49.7
	7 day	28	Measured PC	MERIS Retrieved PC	−5.95	1.85	0.718	<0.0001	10.3	0.248	3.89	0.042	59.3	52.0	49.4
	1 day	14	Cyanobacteria biomass	MERIS Retrieved PC	−2.48	0.00737	0.660	<0.001	-	-	-	-	-	-	-
	3 day	22	Cyanobacteria biomass	MERIS Retrieved PC	−3.04	0.00743	0.745	<0.0001	-	-	-	-	-	-	-
	7 day	28	Cyanobacteria biomass	MERIS Retrieved PC	−3.88	0.00731	0.709	<0.0001	-	-	-	-	-	-	-
Chl-a	1 day	13	Measured Chl-a	MERIS Retrieved Chl-a	2.22	1.45	0.943	<0.0001	12.2	0.203	10.9	0.199	58.9	52.7	45.0
	3 day	23	Measured Chl-a	MERIS Retrieved Chl-a	2.38	1.38	0.930	<0.0001	10.2	0.198	8.71	0.189	55.9	52.7	42.6
	7 day	29	Measured Chl-a	MERIS Retrieved Chl-a	0.995	1.41	0.906	<0.0001	9.19	0.185	7.36	0.167	50	50.6	38.1
	1 day	13	Phytoplankton biomass	MERIS Retrieved Chl-a	3.24	0.00633	0.656	<0.001	-	-	-	-	-	-	-
	3 day	23	Phytoplankton biomass	MERIS Retrieved Chl-a	3.88	0.00588	0.736	<0.0001	-	-	-	-	-	-	-
	7 day	29	Phytoplankton biomass	MERIS Retrieved Chl-a	2.22	0.00587	0.682	<0.0001	-	-	-	-	-	-	-
a_{ph} (665)	7 day	29	Measured a_{ph}(665)	MERIS Retrieved a_{ph}(665)	0.0663	2.06	0.836	<0.0001	0.228	0.444	0.197	0.430	178	175	90.6
a_{ph} (620)	7 day	29	Measured a_{ph}(620)	MERIS Retrieved $a_{Chla+PC}$ (620)	0.0758	3.12	0.834	<0.0001	0.279	0.645	0.242	0.635	346	332	123
b_b (λ)	7 day	29	Measured b_b(650)	MERIS Retrieved b_b(778.75)	0.0938	−0.0603	0.00434	0.734	-	-	-	-	-	-	-

Table A5. Least squares linear regression results (y = *m*x+*b*) for Chl-*a* and PC retrieval using Gons05 and Simis05 algorithms, respectively, validated with integrated samples only (BLI, 2007–2011).

	Match-up Interval	n	x	y	b mg m^{-3}	m	R^2	p	RMSE mg m^{-3}	RMSE log	Bias mg m^{-3}	Bias log	MAPE %	MdAPE %	SMAPE %
PC	1 day	8	Measured PC	MERIS Retrieved PC	6.94	0.664	0.910	<0.001	12.1	0.301	−2.86	0.100	86.4	33.1	45.4
	3 day	8	Measured PC	MERIS Retrieved PC	6.94	0.664	0.910	<0.001	12.1	0.301	−2.86	0.100	86.4	33.1	45.4
	7 day	12	Measured PC	MERIS Retrieved PC	9.55	0.379	0.580	<0.01	29.5	0.324	−13.4	−0.026	71.7	33.1	50.7
	1 day	19	Cyanobacteria biomass	MERIS Retrieved PC	11.2	0.00421	0.405	<0.01	-	-	-	-	-	-	-
	3 day	19	Cyanobacteria biomass	MERIS Retrieved PC	11.2	0.00421	0.405	<0.01	-	-	-	-	-	-	-
	7 day	36	Cyanobacteria biomass	MERIS Retrieved PC	10.1	0.00430	0.474	<0.0001	-	-	-	-	-	-	-
Chl-*a*	1 day	18	Measured Chl-*a*	MERIS Retrieved Chl-*a*	8.26	1.03	0.810	<0.0001	9.73	0.331	8.60	0.271	109	70.6	57.5
	3 day	20	Measured Chl-*a*	MERIS Retrieved Chl-*a*	8.81	1.02	0.798	<0.0001	10.2	0.337	9.09	0.281	112	76.5	59.9
	7 day	52	Measured Chl-*a*	MERIS Retrieved Chl-*a*	9.12	0.950	0.723	<0.0001	10.7	0.334	8.33	0.249	107	61.5	55.7
	1 day	15	Phytoplankton biomass	MERIS Retrieved Chl-*a*	8.27	0.00438	0.692	<0.001	-	-	-	-	-	-	-
	3 day	17	Phytoplankton biomass	MERIS Retrieved Chl-*a*	9.82	0.00417	0.660	<0.0001	-	-	-	-	-	-	-
	7 day	37	Phytoplankton biomass	MERIS Retrieved Chl-*a*	10.3	0.00432	0.640	<0.0001	-	-	-	-	-	-	-

Table A6. Least squares linear regression results (y=*m*x+*b*) for Chl-*a* retrieval using the Gons05 algorithm, validated with KdKVI surface samples only (2007-2011).

	Match-up Interval	n	x	y	b mg m^{-3}	m	R^2	p	RMSE mg m^{-3}	RMSE log	Bias mg m^{-3}	Bias log	MAPE %	MdAPE %	SMAPE %
Chl-*a*	1 day	105	Measured Chl-*a*	MERIS Retrieved Chl-*a*	9.52	1.09	0.784	<0.0001	12.2	0.421	10.6	0.355	169	101	73.3
	3 day	113	Measured Chl-*a*	MERIS Retrieved Chl-*a*	9.27	1.09	0.786	<0.0001	12.0	0.416	10.4	0.352	165	101	72.7
	7 day	113	Measured Chl-*a*	MERIS Retrieved Chl-*a*	9.27	1.09	0.786	<0.0001	12.0	0.416	10.4	0.352	165	101	72.7

Table A7. Least squares linear regression results (y=*m*x+*b*) for Chl-*a* retrieval using the Gons05 algorithm, validated with stations from all datasets used for PC validation only (2010–2011).

	Match-up Interval	n	x	y	b mg m^{-3}	m	R^2	p	RMSE mg m^{-3}	RMSE log	Bias mg m^{-3}	Bias log	MAPE %	MdAPE %	SMAPE %
Chl-*a*	1 day	22	Measured Chl-*a*	MERIS Retrieved Chl-*a*	6.67	1.15	0.910	<0.0001	10.4	0.241	9.33	0.216	70.0	51.9	48.0
	3 day	30	Measured Chl-*a*	MERIS Retrieved Chl-*a*	5.40	1.18	0.909	<0.0001	9.58	0.222	8.39	0.199	62.7	50.9	44.3
	7 day	40	Measured Chl-*a*	MERIS Retrieved Chl-*a*	5.22	1.10	0.866	<0.0001	8.47	0.211	6.84	0.174	55.9	50.1	39.2
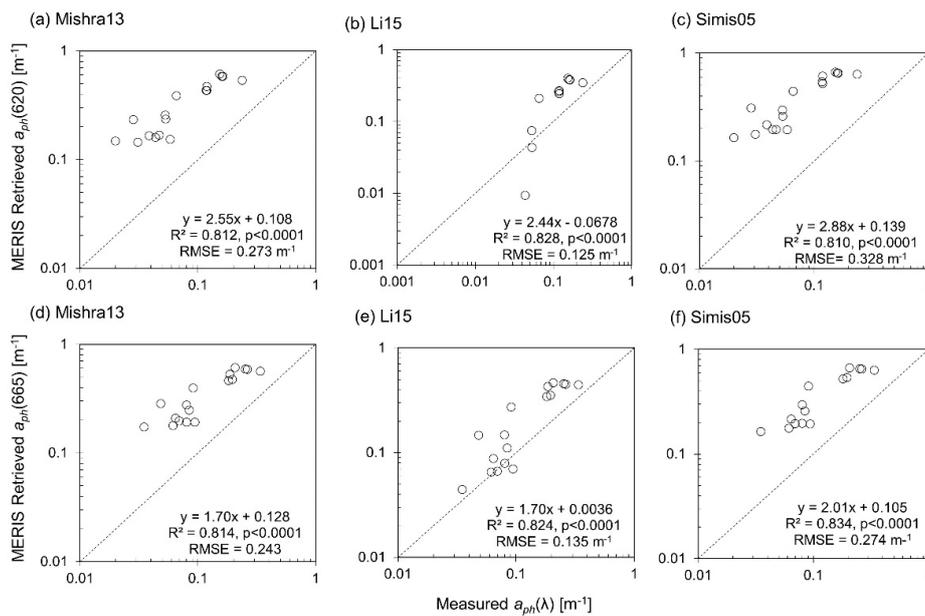
Figure A1. Validation plots of MERIS retrieved $a_{ph}(\lambda)$ from the Mishra13 (**a**,**d**), Li15 (**b**,**e**) and Simis05 (**c**,**f**) models for matchups within 1 day of satellite overpass.
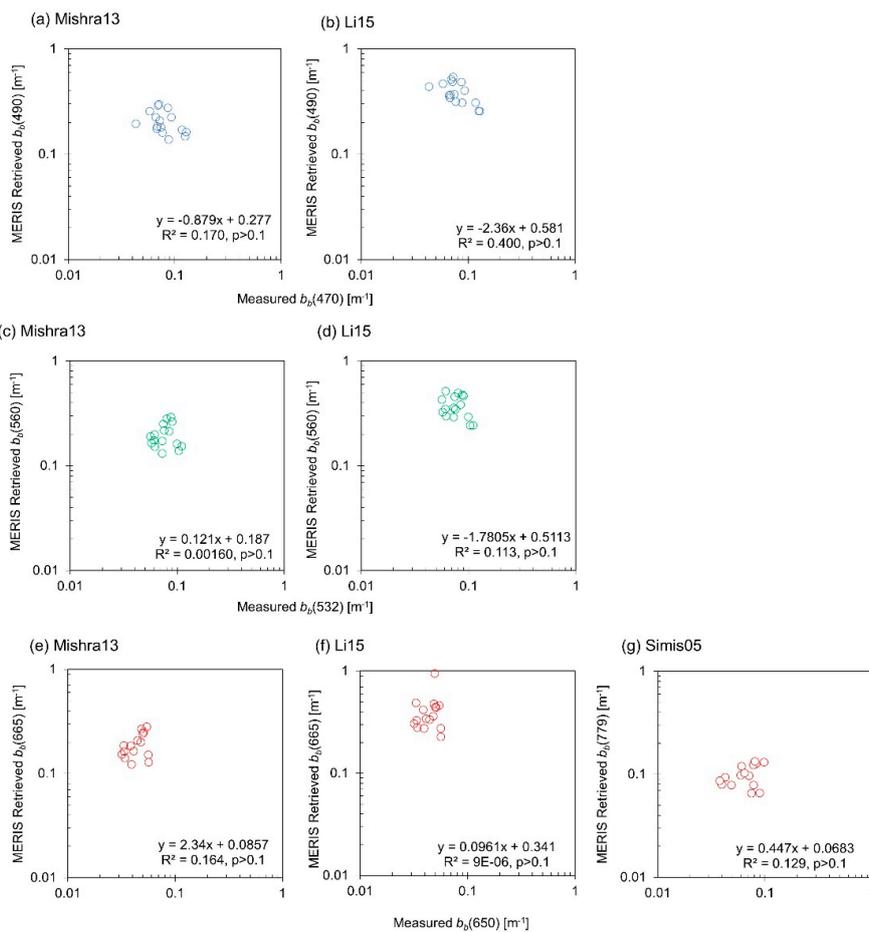
Figure A2. Validation plots of MERIS retrieved $b_b(\lambda)$ from the Mishra13 (**a**,**c**,**e**) and Li15 (**b**,**d**,**f**) bio-optical models, and the Simis05 semi-analytical model (**g**) for matchups within 1 day of satellite overpass (n = 16). Note that negative retrieved values are not shown in log scale.

References

1. Verpoorter, C.; Kutser, T.; Seekell, D.A.; Tranvik, L.J. A global inventory of lakes based on high-resolution satellite imagery. *Geophys. Res. Lett.* **2014**, *41*, 6396–6402. [CrossRef]

2. Bastviken, D.; Tranvik, L.J.; Downing, J.A.; Crill, P.M.; Enrich-Prast, A. Freshwater Methane Emissions Offset the Continental Carbon Sink. *Science* **2011**, *331*, 50. [CrossRef] [PubMed]

3. Cole, J.J.; Prairie, Y.T.; Caraco, N.F.; McDowell, W.H.; Tranvik, L.J.; Striegl, R.G.; Duarte, C.M.; Kortelainen, P.; Downing, J.A.; Middelburg, J.J.; et al. Plumbing the Global Carbon Cycle: Integrating Inland Waters into the Terrestrial Carbon Budget. *Ecosystems* **2007**, *10*, 172–185. [CrossRef]

4. Smith, V.H. Eutrophication of freshwater and coastal marine ecosystems a global problem. *Environ. Sci. Pollut. Res.* **2003**, *10*, 126–139. [CrossRef]

5. Fink, G.; Alcamo, J.; Flörke, M.; Reder, K. Phosphorus Loadings to the World's Largest Lakes: Sources and Trends. *Glob. Biogeochem. Cycles* **2018**, *32*, 617–634. [CrossRef]

6. Paerl, H.W.; Hall, N.S.; Calandrino, E.S. Controlling harmful cyanobacterial blooms in a world experiencing anthropogenic and climatic-induced change. *Sci. Total Environ.* **2011**, *409*, 1739–1745. [CrossRef] [PubMed]

7. Downing, J.A.; Watson, S.B.; McCauley, E. Predicting Cyanobacteria dominance in lakes. *Can. J. Fish. Aquat. Sci.* **2001**, *58*, 1905–1908. [CrossRef]

8. Ferber, L.R.; Levine, S.N.; Lini, A.; Livingston, G.P. Do cyanobacteria dominate in eutrophic lakes because they fix atmospheric nitrogen? *Freshw. Biol.* **2004**, *49*, 690–708. [CrossRef]

9. Horváth, H.; Kovács, A.W.; Riddick, C.; Présing, M. Extraction methods for phycocyanin determination in freshwater filamentous cyanobacteria and their application in a shallow lake. *Eur. J. Phycol.* **2013**, *48*, 278–286. [CrossRef]

10. Présing, M.; Preston, T.; Takátsy, A.; Sprőber, P.; Kovács, A.W.; Vörös, L.; Kenesi, G.; Kóbor, I. Phytoplankton nitrogen demand and the significance of internal and external nitrogen sources in a large shallow lake (Lake Balaton, Hungary). *Hydrobiologia* **2008**, *599*, 87–95. [CrossRef]

11. Schindler, D.W. Evolution of Phosphorus Limitation in Lakes. *Science* **1977**, *195*, 260–262. [CrossRef] [PubMed]

12. Smith, V.H. Low Nitrogen to Phosphorus Ratios Favor Dominance by Blue-Green Algae in Lake Phytoplankton. *Science* **1983**, *221*, 669–671. [CrossRef] [PubMed]

13. Antenucci, J.P.; Ghadouani, A.; Burford, M.A.; Romero, J.R. The long-term effect of artificial destratification on phytoplankton species composition in a subtropical reservoir. *Freshw. Biol.* **2005**, *50*, 1081–1093. [CrossRef]

14. Reynolds, C.S.; Huszar, V.; Kruk, C.; Naselli-Flores, L.; Melo, S. Towards a functional classification of the freshwater phytoplankton. *J. Plankton Res.* **2002**, *24*, 417–428. [CrossRef]

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

16. Wagner, C.; Adrian, R. Cyanobacteria dominance: Quantifying the effects of climate change. *Limnol. Oceanogr.* **2009**, *54*, 2460–2468. [CrossRef]

17. Codd, G.A.; Morrison, L.F.; Metcalf, J.S. Cyanobacterial toxins: risk management for health protection. *Toxicol. Appl. Pharmacol.* **2005**, *203*, 264–272. [CrossRef]

18. Palmer, S.C.J.; Kutser, T.; Hunter, P.D. Remote sensing of inland waters: Challenges, progress and future directions. *Remote Sens. Environ.* **2015**, *157*, 1–8. [CrossRef]

19. Gons, H.J.; Rijkeboer, M.; Ruddick, K.G. A chlorophyll-retrieval algorithm for satellite imagery (Medium Resolution Imaging Spectrometer) of inland and coastal waters. *J. Plankton Res.* **2002**, *24*, 947–951. [CrossRef]

20. Kutser, T. Quantitative detection of chlorophyll in cyanobacterial blooms by satellite remote sensing. *Limnol. Oceanogr.* **2004**, *49*, 2179–2189. [CrossRef]

21. Tyler, A.N.; Svab, E.; Preston, T.; Présing, M.; Kovács, W.A. Remote sensing of the water quality of shallow lakes: A mixture modelling approach to quantifying phytoplankton in water characterized by high-suspended sediment. *Int. J. Remote Sens.* **2006**, *27*, 1521–1537. [CrossRef]

22. Duan, H.; Ma, R.; Hu, C. Evaluation of remote sensing algorithms for cyanobacterial pigment retrievals during spring bloom formation in several lakes of East China. *Remote Sens. Environ.* **2012**, *126*, 126–135. [CrossRef]

23. Hunter, P.D.; Tyler, A.N.; Carvalho, L.; Codd, G.A.; Maberly, S.C. Hyperspectral remote sensing of cyanobacterial pigments as indicators for cell populations and toxins in eutrophic lakes. *Remote Sens. Environ.* **2010**, *114*, 2705–2718. [CrossRef]

24. Kutser, T. Passive optical remote sensing of cyanobacteria and other intense phytoplankton blooms in coastal and inland waters. *Int. J. Remote Sens.* **2009**, *30*, 4401–4425. [CrossRef]

25. Kutser, T.; Metsamaa, L.; Strömbeck, N.; Vahtmäe, E. Monitoring cyanobacterial blooms by satellite remote sensing. *Estuar. Coast. Shelf Sci.* **2006**, *67*, 303–312. [CrossRef]

26. Li, L.; Li, L.; Shi, K.; Li, Z.; Song, K. A semi-analytical algorithm for remote estimation of phycocyanin in inland waters. *Sci. Total Environ.* **2012**, *435–436*, 141–150. [CrossRef] [PubMed]

27. Ruiz-Verdú, A.; Simis, S.G.H.; de Hoyos, C.; Gons, H.J.; Peña-Martínez, R. An evaluation of algorithms for the remote sensing of cyanobacterial biomass. *Remote Sens. Environ.* **2008**, *112*, 3996–4008. [CrossRef]

28. Simis, S.G.H.; Peters, S.W.M.; Gons, H.J. Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. *Limnol. Oceanogr.* **2005**, *50*, 237–245. [CrossRef]

29. Song, K.; Li, L.; Tedesco, L.; Clercin, N.; Hall, B.; Li, S.; Shi, K.; Liu, D.; Sun, Y. Remote estimation of phycocyanin (PC) for inland waters coupled with YSI PC fluorescence probe. *Environ. Sci. Pollut. Res.* **2013**, *20*, 5330–5340. [CrossRef]

30. Hunter, P.; Tyler, A.; Presing, M.; Kovacs, A.; Preston, T. Spectral discrimination of phytoplankton colour groups: The effect of suspended particulate matter and sensor spectral resolution. *Remote Sens. Environ.* **2008**, *112*, 1527–1544. [CrossRef]

31. Dekker, A.G. Detection of Optical Water Quality Parameters for Eutrophic Waters by High Resolution Remote Sensing. Ph.D. Thesis, Proefschrift Vrije Universiteit (Free University), Amsterdam, The Netherlands, 1993.

32. Schalles, J.F.; Yacobi, Y.Z. Remote detection and seasonal patterns of phycocyanin, carotenoid and chlorophyll pigments in eutrophic waters. *Ergeb. Limnol.* **2000**, *55*, 153–168.

33. Gons, H.J. Optical Teledetection of Chlorophyll *a* in Turbid Inland Waters. *Environ. Sci. Technol.* **1999**, *33*, 1127–1132. [CrossRef]

34. Gons, H.J.; Rijkeboer, M.; Ruddick, K.G. Effect of a waveband shift on chlorophyll retrieval from MERIS imagery of inland and coastal waters. *J. Plankton Res.* **2005**, *27*, 125–127. [CrossRef]

35. Le, C.; Li, Y.; Zha, Y.; Wang, Q.; Zhang, H.; Yin, B. Remote sensing of phycocyanin pigment in highly turbid inland waters in Lake Taihu, China. *Int. J. Remote Sens.* **2011**, *32*, 8253–8269. [CrossRef]

36. Randolph, K.; Wilson, J.; Tedesco, L.; Li, L.; Pascual, D.L.; Soyeux, E. Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments, chlorophyll a and phycocyanin. *Remote Sens. Environ.* **2008**, *112*, 4009–4019. [CrossRef]

37. Simis, S.G.H.; Ruiz-Verdú, A.; Domínguez-Gómez, J.A.; Peña-Martinez, R.; Peters, S.W.M.; Gons, H.J. Influence of phytoplankton pigment composition on remote sensing of cyanobacterial biomass. *Remote Sens. Environ.* **2007**, *106*, 414–427. [CrossRef]

38. Wheeler, S.M.; Morrissey, L.A.; Levine, S.N.; Livingston, G.P.; Vincent, W.F. Mapping cyanobacterial blooms in Lake Champlain's Missisquoi Bay using QuickBird and MERIS satellite data. *J. Gt. Lakes Res.* **2012**, *38*, 68–75. [CrossRef]

39. Mishra, S.; Mishra, D.R.; Lee, Z.; Tucker, C.S. Quantifying cyanobacterial phycocyanin concentration in turbid productive waters: A quasi-analytical approach. *Remote Sens. Environ.* **2013**, *133*, 141–151. [CrossRef]

40. Qi, L.; Hu, C.; Duan, H.; Cannizzaro, J.; Ma, R. A novel MERIS algorithm to derive cyanobacterial phycocyanin pigment concentrations in a eutrophic lake: Theoretical basis and practical considerations. *Remote Sens. Environ.* **2014**, *154*, 298–317. [CrossRef]

41. Li, L.; Li, L.; Song, K. Remote sensing of freshwater cyanobacteria: An extended IOP Inversion Model of Inland Waters (IIMIW) for partitioning absorption coefficient and estimating phycocyanin. *Remote Sens. Environ.* **2015**, *157*, 9–23. [CrossRef]

42. Liu, G.; Simis, S.G.H.; Li, L.; Wang, Q.; Li, Y.; Song, K.; Lyu, H.; Zheng, Z.; Shi, K. A Four-Band Semi-Analytical Model for Estimating Phycocyanin in Inland Waters From Simulated MERIS and OLCI Data. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 1374–1385. [CrossRef]

43. Matthews, M.W. A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. *Int. J. Remote Sens.* **2011**, *32*, 6855–6899. [CrossRef]

44. Li, L.; Song, K. Bio-optical Modeling of Phycocyanin. In *Bio-optical Modeling and Remote Sensing of Inland Waters*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 233–262, ISBN 978-0-12-804644-9.

45. Yan, Y.; Bao, Z.; Shao, J. Phycocyanin concentration retrieval in inland waters: A comparative review of the remote sensing techniques and algorithms. *J. Gt. Lakes Res.* **2018**, *44*, 748–755. [CrossRef]

46. Blix, K.; Pálffy, K.; Tóth, V.R.; Eltoft, T. Remote Sensing of Water Quality Parameters over Lake Balaton by Using Sentinel-3 OLCI. *Water* **2018**, *10*, 1428. [CrossRef]

47. Mouw, C.B.; Greb, S.; Aurin, D.; DiGiacomo, P.M.; Lee, Z.; Twardowski, M.; Binding, C.; Hu, C.; Ma, R.; Moore, T.; et al. Aquatic color radiometry remote sensing of coastal and inland waters: Challenges and recommendations for future satellite missions. *Remote Sens. Environ.* **2015**, *160*, 15–30. [CrossRef]

48. Tyler, A.N.; Hunter, P.D.; Spyrakos, E.; Groom, S.; Constantinescu, A.M.; Kitchen, J. Developments in Earth observation for the assessment and monitoring of inland, transitional, coastal and shelf-sea waters. *Sci. Total Environ.* **2016**, *572*, 1307–1321. [CrossRef] [PubMed]

49. Ogashawara, I.; Mishra, D.; Mishra, S.; Curtarelli, M.; Stech, J. A Performance Review of Reflectance Based Algorithms for Predicting Phycocyanin Concentrations in Inland Waters. *Remote Sens.* **2013**, *5*, 4774–4798. [CrossRef]

50. Beck, R.; Xu, M.; Zhan, S.; Liu, H.; Johansen, R.; Tong, S.; Yang, B.; Shu, S.; Wu, Q.; Wang, S.; et al. Comparison of Satellite Reflectance Algorithms for Estimating Phycocyanin Values and Cyanobacterial Total Biovolume in a Temperate Reservoir Using Coincident Hyperspectral Aircraft Imagery and Dense Coincident Surface Observations. *Remote Sens.* **2017**, *9*, 538. [CrossRef]

51. Riddick, C.A.L.; Hunter, P.D.; Tyler, A.N.; Martinez-Vicente, V.; Horváth, H.; Kovács, A.W.; Vörös, L.; Preston, T.; Présing, M. Spatial variability of absorption coefficients over a biogeochemical gradient in a large and optically complex shallow lake: LIGHT ABSORPTION IN LAKE BALATON. *J. Geophys. Res. Oceans* **2015**, *120*, 7040–7066. [CrossRef]

52. Aulló-Maestro, M.E.; Hunter, P.; Spyrakos, E.; Mercatoris, P.; Kovács, A.; Horváth, H.; Preston, T.; Présing, M.; Torres Palenzuela, J.; Tyler, A. Spatio-seasonal variability of chromophoric dissolved organic matter absorption and responses to photobleaching in a large shallow temperate lake. *Biogeosciences* **2017**, *14*, 1215–1233. [CrossRef]

53. Iwamura, T.; Nagai, H.; Ichimura, S.-E. Improved Methods for Determining Contents of Chlorophyll, Protein, Ribonucleic Acid, and Deoxyribonucleic Acid in Planktonic Populations. *Int. Revue ges. Hydrobiol. Hydrogr.* **1970**, *55*, 131–147. [CrossRef]

54. Palmer, S.C.J.; Hunter, P.D.; Lankester, T.; Hubbard, S.; Spyrakos, E.; Tyler, A.N.; Présing, M.; Horváth, H.; Lamb, A.; Balzter, H.; et al. Validation of Envisat MERIS algorithms for chlorophyll retrieval in a large, turbid and optically-complex shallow lake. *Remote Sens. Environ.* **2015**, *157*, 158–169. [CrossRef]

55. Sarada, R.; Pillai, M.G.; Ravishankar, G.A. Phycocyanin from Spirulina sp: influence of processing of biomass on phycocyanin yield, analysis of efficacy of extraction methods and stability studies on phycocyanin. *Process Biochem.* **1999**, *34*, 795–801. [CrossRef]

56. Siegelman, H.; Kycia, J.H. Algal biliproteins. In *Handbook of Phycological Methods: Physiological and Biochemical Methods*; Hellebust, J.A., Craigie, J.S., Eds.; Cambridge University Press: New York, NY, USA, 1978; pp. 71–79.

57. Utermöhl, H. *Zur Vervollkommnung der quantitativen Phytoplankton-Methodik*; Schweizerbart: Stuttgart, Germany, 1958.

58. Németh, J.; Vörös, L. Koncepció és módszertan felszíni vizek algológiai monitoringjához [Concepts and methodics for algological monitoring of surface water]. In *Környezetés természetvédelmi kutatások*; Katona, S., Ed.; Országos Környezet és Termeszetvédelmi Hivatal: Budapest, Hungary, 1986; ISBN 978-963-02-4647-7.

59. Tassan, S.; Ferrari, G.M. Measurement of light absorption by aquatic particles retained on filters: determination of the optical pathlength amplification by the 'transmittance-reflectance' method. *J. Plankton Res.* **1998**, *20*, 1699–1709. [CrossRef]

60. Wet Labs, Inc. *ECO 3-Measurement Sensor (Triplet)*; Wet Labs, Inc.: Philomath, Oregon, 2010; pp. 1–23.

61. Slade, W.H.; Boss, E. Spectral attenuation and backscattering as indicators of average particle size. *Appl. Opt.* **2015**, *54*, 7264. [CrossRef]

62. Sullivan, J.M.; Twardowski, M.S.; Ronald, J.; Zaneveld, V.; Moore, C.C. Measuring optical backscattering in water. In *Light Scattering Reviews 7: Radiative Transfer and Optical Properties of Atmosphere and Underlying Surface*; Kokhanovsky, A.A., Ed.; Springer Praxis Books; Springer: Berlin, Heidelberg, Germany, 2013; pp. 189–224, ISBN 978-3-642-21907-8.

63. Röttgers, R.; McKee, D.; Woźniak, S.B. Evaluation of scatter corrections for ac-9 absorption measurements in coastal waters. *Methods Oceanogr.* **2013**, *7*, 21–39. [CrossRef]

64. Zhang, X.; Hu, L.; He, M.-X. Scattering by pure seawater: Effect of salinity. *Opt. Express* **2009**, *17*, 5698. [CrossRef]

65. Mueller, J.L.; Davis, C.; Arnone, R.; Frouin, R.; Carder, K.; Lee, Z.P.; Steward, R.G.; Hooker, S.; Mobley, C.D.; McLean, S. Above-Water Radiance and Remote Sensing Reflectance Measurement and Analysis Protocols. In *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation*; Fargion, G.S., Mueller, J.L., Eds.; NASA: Greenbelt, Maryland, 2000; Chapter 10; p. 184.

66. Mueller, J.L.; Morel, A.; Frouin, R.; Davis, C.; Arnone, R.; Carder, K.; Lee, Z.P.; Steward, R.G.; Hooker, S.; Mobley, C.D.; et al. Radiometric Measurements and Data Analysis Protocols. In *Ocean Optics Protocols for Satellite Color Sensor Validation*; Mueller, J.L., Fargion, G.S., McClain, C.R., Eds.; NASA: Greenbelt, Maryland, 2003; Volume III, p. 78.

67. Guanter, L.; Ruiz-Verdú, A.; Odermatt, D.; Giardino, C.; Simis, S.; Estellés, V.; Heege, T.; Domínguez-Gómez, J.A.; Moreno, J. Atmospheric correction of ENVISAT/MERIS data over inland waters: Validation for European lakes. *Remote Sens. Environ.* **2010**, *114*, 467–480. [CrossRef]

68. Domínguez Gómez, J.A.; Alonso Alonso, C.; Alonso García, A. Remote sensing as a tool for monitoring water quality parameters for Mediterranean Lakes of European Union water framework directive (WFD) and as a system of surveillance of cyanobacterial harmful algae blooms (SCyanoHABs). *Environ. Monit. Assess.* **2011**, *181*, 317–334. [CrossRef]

69. Agha, R.; Cirés, S.; Wörmer, L.; Domínguez, J.A.; Quesada, A. Multi-scale strategies for the monitoring of freshwater cyanobacteria: Reducing the sources of uncertainty. *Water Res.* **2012**, *46*, 3043–3053. [CrossRef]

70. Jaelani, L.M.; Matsushita, B.; Yang, W.; Fukushima, T. Evaluation of four MERIS atmospheric correction algorithms in Lake Kasumigaura, Japan. *Int. J. Remote Sens.* **2013**, *34*, 8967–8985. [CrossRef]

71. Medina-Cobo, M.; Domínguez, J.A.; Quesada, A.; de Hoyos, C. Estimation of cyanobacteria biovolume in water reservoirs by MERIS sensor. *Water Res.* **2014**, *63*, 10–20. [CrossRef] [PubMed]

72. Yang, W.; Matsushita, B.; Chen, J.; Fukushima, T. A Relaxed Matrix Inversion Method for Retrieving Water Constituent Concentrations in Case II Waters: The Case of Lake Kasumigaura, Japan. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 3381–3392. [CrossRef]

73. Yang, W.; Matsushita, B.; Chen, J.; Fukushima, T. Estimating constituent concentrations in case II waters from MERIS satellite data by semi-analytical model optimizing and look-up tables. *Remote Sens. Environ.* **2011**, *115*, 1247–1259. [CrossRef]

74. Goyens, C.; Jamet, C.; Schroeder, T. Evaluation of four atmospheric correction algorithms for MODIS-Aqua images over contrasted coastal waters. *Remote Sens. Environ.* **2013**, *131*, 63–75. [CrossRef]

75. Jamet, C.; Loisel, H.; Kuchinke, C.P.; Ruddick, K.; Zibordi, G.; Feng, H. Comparison of three SeaWiFS atmospheric correction algorithms for turbid waters using AERONET-OC measurements. *Remote Sens. Environ.* **2011**, *115*, 1955–1965. [CrossRef]

76. Dall'Olmo, G.; Gitelson, A.A.; Rundquist, D.C. Towards a unified approach for remote estimation of chlorophyll-a in both terrestrial vegetation and turbid productive waters: UNIFIED APPROACH FOR CHLOROPHYLL ESTIMATION. *Geophys. Res. Lett.* **2003**, *30*. [CrossRef]

77. Seegers, B.N.; Stumpf, R.P.; Schaeffer, B.A.; Loftin, K.A.; Werdell, P.J. Performance metrics for the assessment of satellite data products: an ocean color case study. *Opt. Express* **2018**, *26*, 7404. [CrossRef]

78. Spyrakos, E.; O'Donnell, R.; Hunter, P.D.; Miller, C.; Scott, M.; Simis, S.G.H.; Neil, C.; Barbosa, C.C.F.; Binding, C.E.; Bradt, S.; et al. Optical types of inland and coastal waters: Optical types of inland and coastal waters. *Limnol. Oceanogr.* **2017**, *63*, 846–870. [CrossRef]

79. Binding, C.E.; Greenberg, T.A.; McCullough, G.; Watson, S.B.; Page, E. An analysis of satellite-derived chlorophyll and algal bloom indices on Lake Winnipeg. *J. Gt. Lakes Res.* **2018**, *44*, 436–446. [CrossRef]

80. Salem, S.; Strand, M.; Higa, H.; Kim, H.; Kazuhiro, K.; Oki, K.; Oki, T. Evaluation of MERIS Chlorophyll-a Retrieval Processors in a Complex Turbid Lake Kasumigaura over a 10-Year Mission. *Remote Sens.* **2017**, *9*, 1022. [CrossRef]

81. Neil, C.; Spyrakos, E.; Hunter, P.D.; Tyler, A.N. A global approach for chlorophyll-a retrieval across optically complex inland waters based on optical water types. *Remote Sens. Environ.* **2019**, *229*, 159–178. [CrossRef]

82. McKee, D.; Röttgers, R.; Neukermans, G.; Calzado, V.S.; Trees, C.; Ampolo-Rella, M.; Neil, C.; Cunningham, A. Impact of measurement uncertainties on determination of chlorophyll-specific absorption coefficient for marine phytoplankton. *J. Geophys. Res. Oceans* **2014**, *119*, 9013–9025. [CrossRef]

83. Lee, Z.; Carder, K.L.; Arnone, R.A. Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. *Appl. Opt.* **2002**, *41*, 5755. [CrossRef] [PubMed]

84. Gurlin, D.; Gitelson, A.A.; Moses, W.J. Remote estimation of chl-a concentration in turbid productive waters — Return to a simple two-band NIR-red model? *Remote Sens. Environ.* **2011**, *115*, 3479–3490. [CrossRef]

85. Lunetta, R.S.; Schaeffer, B.A.; Stumpf, R.P.; Keith, D.; Jacobs, S.A.; Murphy, M.S. Evaluation of cyanobacteria cell count detection derived from MERIS imagery across the eastern USA. *Remote Sens. Environ.* **2015**, *157*, 24–34. [CrossRef]

86. Hunter, P.D.; Tyler, A.N.; Willby, N.J.; Gilvear, D.J. The spatial dynamics of vertical migration by Microcystis aeruginosa in a eutrophic shallow lake: A case study using high spatial resolution time-series airborne remote sensing. *Limnol. Oceanogr.* **2008**, *53*, 2391–2406. [CrossRef]

87. Kovács, A.W.; Tóth, V.R.; Vörös, L. Light-dependent germination and subsequent proliferation of N_2-fixing cyanobacteria in a large shallow lake. *Ann. Limnol. Int. J. Limnol.* **2012**, *48*, 177–185. [CrossRef]

88. Bricaud, A.; Babin, M.; Morel, A.; Claustre, H. Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: Analysis and parameterization. *J. Geophys. Res.* **1995**, *100*, 13321. [CrossRef]

89. Simis, S.G.H.; Kauko, H.M. In vivo mass-specific absorption spectra of phycobilipigments through selective bleaching: Selective bleaching of phytoplankton pigments. *Limnol. Oceanogr. Methods* **2012**, *10*, 214–226. [CrossRef]

90. Yacobi, Y.Z.; Köhler, J.; Leunert, F.; Gitelson, A. Phycocyanin-specific absorption coefficient: Eliminating the effect of chlorophylls absorption: Phycocyanin-specific absorption coefficient. *Limnol. Oceanogr. Methods* **2015**, *13*, e10015. [CrossRef]

91. Gons, H.J.; Auer, M.T.; Effler, S.W. MERIS satellite chlorophyll mapping of oligotrophic and eutrophic waters in the Laurentian Great Lakes. *Remote Sens. Environ.* **2008**, *112*, 4098–4106. [CrossRef]

92. Gordon, H.R.; Brown, O.B.; Jacobs, M.M. Computed Relationships Between the Inherent and Apparent Optical Properties of a Flat Homogeneous Ocean. *Appl. Opt.* **1975**, *14*, 417. [CrossRef] [PubMed]

93. Mittenzwey, K.-H.; Ullrich, S.; Gitelson, A.A.; Kondratiev, K.Y. Determination of chlorophyll a of inland waters on the basis of spectral reflectance. *Limnol. Oceanogr.* **1992**, *37*, 147–149. [CrossRef]