A four-band semi-analytical model for estimating phycocyanin in inland waters from simulated MERIS and OLCI data

Ge Liu, Stefan G.H. Simis, Lin Li, Qiao Wang, Yunmei Li*, Kaishan Song, Heng Lyu, Zhubin

Zheng, and Kun Shi*

Abstract—Existing remote sensing algorithms to estimate the phycocyanin (PC) concentration in turbid inland waters have high associated uncertainties, especially at low PC concentrations in diverse phytoplankton communities. This study provides the theoretical framework for a four-band semi-analytical algorithm (FBA_PC) which isolates PC absorption from second-order variability caused by yellow matter and other phytoplankton pigment absorption. The algorithm suits the band configuration of both the Medium Resolution Imaging Spectrometer (MERIS) and Sentinel-3 Ocean and Land Color Instrument (OLCI). Calibration of the algorithm was based on absorption data from twelve inland water bodies in the USA, The Netherlands, and China, combined with measurements from laboratory-grown cultures, demonstrated that the assumptions underlying FBA-PC are an improvement over existing three-band approaches. Validation of FBA_PC in seven inland water bodies in the USA, The Netherlands, and China showed good agreement of FBA_PC adjusted to the MERIS/OLCI band configuration with measured PC, with root-mean-square error (RMSE) = 27.691 mg m^{-3}, mean absolute percentage error (MAPE) = 172.863 %, and coefficient of determination (R^2) = 0.730). FBA_PC outperformed previously proposed PC algorithms that can be applied to MERIS or OLCI data, and is expected to be more robust when applied to a wider range of water bodies.

Index Terms—remote sensing, turbid inland waters, phytoplankton, cyanobacteria, phycocyanin, MERIS, OLCI

Ge Liu is with the Department of Geography, Nanjing Normal University, Nanjing, China; with Northeast Institute of Geography and Agricultural Ecology, Chinese Academy of Sciences, Changchun, China, and with the State Key laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China (e-mail: liugers@163.com)

Stefan Simis is with the Plymouth Marine Laboratory, Plymouth, United Kingdom (e-mail: stsi@pml.ac.uk).

Lin Li, is with the Department of Earth Science, Indiana University-Purdue University Indianapolis, Indianapolis, USA (e-mail: ll3@iupui.edu).

Yunmei Li, Heng Lyu, and Zhubin Zheng are with the Department of Geography, Nanjing Normal University, Nanjing, China (e-mail: liyunmei@njnu.edu.cn, henglu@njnu.edu.cn, and zhengspace@126.com).

Qiao Wang is with the Satellite Environment Application Center, Ministry of Environmental Protection, Beijing 100029, China (wangqiao@sepa.gov.cn).

Kaishan Song, is with the Northeast Institute of Geography and Agricultural Ecology, Chinese Academy of Sciences, Changchun, China (e-mail: songks@iga.ac.cn).

Kun Shi is with the State Key laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China ((e-mail: kshi@niglas.ac.cn).

* Author to whom correspondence should be addressed

I. INTRODUCTION

The combined effects of increased anthropogenic nutrient loading of freshwater systems and global warming have led to a global increase of cyanobacterial blooms in inland waters [1-4]. A number of bloom-forming species produce toxins (neurotoxins, hepatotoxins, cytotoxins, and irritant and gastrointestinal toxins) which pose a hazard to animal and human health [5, 6]. Cyanobacterial blooms can affect drinking water supply [7], disrupt food webs [8], lead to hypoxia [9], and reduce biodiversity [10, 11]. Identifying and quantifying cyanobacteria biomass is therefore key for monitoring, forecasting, early warning, and risk management of inland water bodies.

Traditional laboratory methods can provide detailed quantitative biological and chemical characterization of cyanobacteria blooms. These methods include cell counts, biovolume, chlorophyll-a (Chl*a*), phycocyanin (PC), microcystin (MC), dissolved oxygen concentration (DO), and chemical oxygen demand (COD). Remote sensing can provide a synoptic complement to the traditional methods by targeting optically active substance concentrations. Cyanobacterial biomass can be characterized from PC, a prominent accessory photosynthetic pigment of freshwater cyanobacteria [12]. PC has a distinct absorption peak at 615 nm which can be targeted with remote sensors for monitoring and assessment of freshwater cyanobacteria. Remote sensing reflectance ($R_{rs}(\lambda)$), with λ denoting the waveband of light, is a function of the inherent absorption ($a(\lambda)$) and backscattering coefficient ($b_b(\lambda)$) [13] :

$$R_{rs}(\lambda) = \frac{f(\lambda)}{Q(\lambda)} \times \frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)} \qquad (1)$$

Where f(λ) describes the sensitivity of the reflectance to variations in the solar angle [14] and Q(λ) describes the bidirectional properties of the reflectance [15]. The ratio $f(\lambda)/Q(\lambda)$ can be approximated by a constant for limited

variation in sun angle, and is only weakly dependent on wavelength [15, 16]. The absorption coefficient $a(\lambda)$ can be divided into absorption by pure water ($a_w(\lambda)$), PC ($a_{PC}(\lambda)$), non-phycocyanin phytoplankton pigment ($a_{ph-PC}(\lambda)$), and yellow matter ($a_{ym}(\lambda)$), the sum of dissolved organic matter and non-pigmented particulates:

$$a(\lambda) = a_{PC}(\lambda) + a_{ph-PC}(\lambda) + a_{ym}(\lambda) + a_w(\lambda) \quad (2)$$

Algorithms previously proposed for estimating PC from $R_{rs}(\lambda)$ can be separated into empirical and semi-analytical categories.

Dekker *et al.* (1993) presented an empirical baseline algorithm which uses the relative height of $R_{rs}(624)$ against two neighboring bands ($R_{rs}(600)$ and $R_{rs}(648)$) as an index of PC concentration in 10 lakes in The Netherlands [17]. Qi *et al.* (2014) developed a band-subtraction model (henceforth denoted Qi14) which can be used with Medium Resolution Imaging Spectrometer (MERIS) data to retrieve PC concentration in Lake Taihu, China (Eq. 3) [18].

$$\text{Qi14} = R_{rs}(560) - R_{rs}(620)$$
$$+ \frac{620-560}{665-560} \times \left(R_{rs}(665) - R_{rs}(560) \right) \quad (3)$$

Schalles *et al.* (2000) used the ratio $R_{rs}(650)/R_{rs}(625)$ (denoted Schalles00) to estimate PC in Lake Carter, USA [19]. Mishra *et al.* (2009) presented a band ratio algorithm based on the ratio $R_{rs}(700)/R_{rs}(600)$, where $R_{rs}(600)$ is used to avoid Chl*a* absorption which overlaps with the PC absorption peak [20]. Woźniak *et al.* (2016) showed a model which was based on linear combination of two variables that is logarithmic transformation of two band ratios for the estimation of PC in the optically complex waters of the Baltic Sea, and formulates solutions for hyperspectral $R_{rs}(\lambda)$ as well as $R_{rs}(\lambda)$ restricted to the wavebands of MERIS and Sentinel-3 Ocean and Land Colour Imager (OLCI) radiometer, respectively [21].

Semi-analytical algorithms include that of Simis *et al.* (2005) who proposed a nested semi-analytical band ratio algorithm (denoted Simis05) [22]. This algorithm combines two ratios of bands compatible with MERIS [23, 24]. Simis05 uses $R_{rs}(778)$ to retrieve b_b, $R_{rs}(709)/R_{rs}(665)$ to estimate Chl*a* absorption at 665 nm ($a_{Chla}(665)$), and $R_{rs}(709)/R_{rs}(620)$ to estimate the total absorption of Chl*a* and PC at 620 nm ($a_{Chla+PC}(620)$). Then, a semi-empirical correction is applied to ($a_{Chla+PC}(620)$) to remove the contribution by Chl*a*, isolating PC absorption. Hunter *et al.* (2008, 2010) used the three band model which was firstly proposed for Chl*a* estimation by Dall'Olmo *et al.* (2003) [25] to retrieve PC in laboratory culture experiments and in Loch Leven and Esthwaite Water, UK, two shallow, eutrophic lakes (denoted Hunter10) [26, 27]. Duan *et al.* (2012) adapted Hunter10 for MERIS for the estimation of PC in three inland lakes (Lake Taihu, Lake Dongjiu, and Lake Gehu) in China (denoted as HD12) [28]. Li *et al.* (2012) proposed a method that combines a three band algorithm with a baseline algorithm (denoted TBBA) which is less sensitive to the interference of absorption due to yellow matter, other pigments and backscattering of the water column, to estimate PC in three reservoirs in central Indiana, USA [29]. Mishra *et al.* (2014) presented a three band algorithm (denoted Mishra14) which accountes for $a_{Chla}(620)$ by using an empirical relationship to

the band ratio $R_{rs}(560)/R_{rs}(665)$, using *in situ* data collected from 15 aquaculture ponds with high PC concentrations [30].

MERIS (2002-2012) provided observations at spectral (15 bands from 412 to 900 nm, including 620 nm), radiometric (16-bit), spatial (300 nm at full resolution) and temporal (three-day revisit at the equator) resolutions [31] that are considered adequate for inland water monitoring [32]. The Sentinel-3 mission carrying the Ocean and Land Colour Instrument (OLCI) continues the capability for remote sensing of PC in inland waters, with Sentinel-3A launched in 2016. Of the algorithms listed above only Schalles00, Simis05, HD12, Mishra14, and Qi14 can be used with MERIS/OLCI band configurations. The empirical algorithms Schalles00 and Qi14 are not a monotonic function of PC, so these algorithms may be sensitive to variability in optically active water constituents outside the range considered when the algorithms are developed. Simis05, HD12, and Mishra14 are semi-analytical algorithms, aiming to isolate $a_{pc}(620)$ from $a(620)$. Simis05 uses two empirical coefficients to tune the retrieval of $a_{ph}(620)$ and $a_{ph}(665)$, which may lead to errors when $a_{ym}(620)$ varies outside the range over which Simis05 was calibrated. HD12 uses $a_{ph}(709)$ and $a_{ym}(709)$ to account for non-PC $a_{ph}(620)$ and $a_{ym}(620)$ simultaneously. In turbid inland waters, if the difference of $a_{ym}(\lambda)$ between 620 and 709 nm is non-negligible, subtracting the effect of yellow matter in HD12 may result in erroneous estimates of PC values. Mishra14 attribute the absorption signal at 620 nm to Chl*a*, PC, and water, but ignoring influences of yellow matter absorption. Whereas each algorithm accounts for secondary (non-PC related) variability in specific ways, each presents limitations which are expected to affect their transferability between different optical water types [33-35].

In the present study, the secondary variability around PC retrieval is addressed in a semi-analytical four band algorithm (FBA_PC) with applicability to MERIS/OLCI bands. The primary approach is a modification of the Hunter10 algorithm with the aim to effectively isolate $a_{pc}(620)$ from $a(620)$, based on new assumptions on the variability of $a_{ph}(\lambda)$ and $a_{ym}(\lambda)$. The performance of FBA_PC is evaluated using an extensive database of field radiometric and pigment data obtained in seven inland lakes.

II. DATA AND METHODS

A. Study Areas

(1) Indiana Reservoirs, USA

The sites include three Indiana reservoirs: Eagle Creek Reservoir (39.8°N, 86.3° W), Geist Reservoir (39.9° N, 85.9° W), and Morse Reservoir (40.1° N, 86.0° W). The three reservoirs supply drinking water to the Indianapolis metropolitan area and have similar depth (3.2- 4.7 m), surface area (5.0- 7.5 km^2), volume (21-28 $\times 10^6$ m^3) and residence times (55 – 70 days) [29, 36]. Ten field campaigns were conducted in these reservoirs in 2010, collecting hyperspectral reflectance, PC and Chl*a* concentrations, and absorption measurements. A brief description of measurement protocols is

provided further below. The range of PC and Chl*a* concentrations observed in these campaigns is listed in Table 1.

(2) Lake IJsselmeer, The Netherlands

Lake IJsselmeer (52.8° N, 5.3° E) is the largest freshwater body of Western Europe (1190 km^2). The lake is shallow (mean depth 4.4 m) and eutrophic as a result of anthropogenic activities [37]. The lake has a seasonal phytoplankton succession of microalgae and cyanobacteria, the latter predominant in late summer. Cyanobacteria blooms consist of *Aphanizomenon* sp. and *Microcystis* sp. and occasionally accumulate at the water surface during calm summer days [38, 39]. Six sampling campaigns took place in spring, summer, and late summer during 2004-2005, collecting hyperspectral reflectance and water samples for pigment concentrations and absorption. Measurements and protocols are briefly described further below. The range of PC and Chl*a* concentrations observed in this lake is summarized in Table 1.

(3) Chinese Lakes And Reservoirs

Absorption measurements were carried out on water samples collected from eight inland water bodies in China: Lake Taihu, Lake Chaohu, Lake Hengshui, Lake Dianchi, Lake Dongting, Lake Erhai, Jiajiang Reservoir, and Three Gorges Reservoir. Lake Taihu (31.2° N, 120.1° E) located in the Yangtze River Delta is the third largest freshwater body in China. It is a highly eutrophic lake in which cyanobacterial blooms (*Microcystis* sp.) have occurred frequently over the past twenty years [40]. Lake Chaohu (31.5° N, 117.6° E) is the largest lake in Anhui Province and the fifth largest freshwater body in China. The extensive exploitation of Lake Chaohu in recent years has led to eutrophication and silting [41]. Lake Hengshui (37.6° N, 115.6° E) located in the Hebei Province is one of the important bird habitats in the north of China. Lake Dianchi (24.5° N, 102.3° E) is the largest plateau lake in Yunnan Province, with a coving area of 370 km^2. It is a seriously eutrophic lake with frequent cyanobacterial blooms and extremely high concentrations of total nitrogen (TN), total phosphorus (TP), and Chl*a* [42-44]. Lake Dongting (29.2° N, 113.0° E) is the second largest freshwater body in China. It plays a highly important role in this region for flood protection, drinking water supply, and as water supply for agriculture [45]. The water of Lake Dongting has experienced eutrophication in recent years [41]. Lake Erhai (25.4° N, 99.5° E) located in the Yunnan Province plays a vital role in freshwater supply, agricultural irrigation, fishery, tourism, and navigation [46]. Jiajiang Reservoir (32.0° N, 118.7° E) is located in the lower reaches of the Yangtze River near the city of Nanjing and its main source of drinking water. Three Gorges Reservoir (30.8° E, 110.9° E) is located in the upstream part of the Yangtze River at the boundary of Hubei Province and Chongqing City and is also experiencing eutrophication [47]. Basic statistics on the *in situ* $a_{ph}(620)$, $a_{ym}(620)$, $a_{ph}(620)$:$a_{ym}(620)$, $a_{ym}(620)$:PC, and $a_{ph}(620)$:PC values in Table 2 shows the large difference of the bio-optical properties between the different inland lakes.

B. In Situ Remote Sensing Reflectance Data

An Ocean Optics USA400 unit (Ocean Optics, Inc., Dunedin, FL, USA) radiometer with fibre optics was used to measure $R_{rs}(\lambda)$ in the three Indiana reservoirs in 2010. The measurements of $R_{rs}(\lambda)$ followed the procedure that is described in detail in Gitelson *et al.* (2008) [48]. For the dataset of Dutch lakes, a PR-650 (Photo Research) was used to measure $R_{rs}(\lambda)$ in consecutive measurements of water leaving radiance, sky radiance and radiance reflected by a calibrated Spectralon diffuser plate. Details on the measurement procedure are provided in Simis *et al.* (2005) [22]. The FieldSpec spectroradiometer (Analytical Spectral Devices, Inc., Boulder, CO, USA) was used to measure $R_{rs}(\lambda)$ with wavelength ranges from 350-1050 nm in Lake Taihu, Lake Chaohu, and Lake Hengshui in China in 2016. The measurements of $R_{rs}(\lambda)$ followed the procedure that is described in detail in Mueller *et al.* (2000) [49].

The measured $R_{rs}(\lambda)$ were convolved to represent the spectra of MERIS. This was achieved by applying MERIS's spectral response function (SRF) with *in situ* $R_{rs}(\lambda)$ data [50]. The SRF were downloaded from the U.S NASA Goddard Flight Space Center (GSFC, http://oceancolor.gsfc.nasa.gov).

C. Water Sample Analyses

Surface water samples were collected, filtered onboard and then transferred to cold, dark coolers in the laboratory for analysis within 24h. The absorption coefficient of total particular matter, subsequently separated into phytoplankton and detrital components through bleaching with 80% hot ethanol, were measured following the QFT approach, using a 150-mm integrating sphere with corrections [22]. The absorption coefficient of dissolved organic matter was determined spectrophotometrically against a reference of ultrapure water, following filtration of the water sample over a 0.2 μm membrane filter. The determination of *in situ* PC concentration in the three Indiana reservoirs, Lake Taihu, Lake Chaohu, and Lake Hengshui were same and different from Lake IJsselmeer. For the samples collected in the three Indianan reservoirs, Lake Taihu, Lake Chaohu, and Lake Hengshui, PC concentration was determined based on the method detailed in Randolph *et al.* (2008) [35] which was a modified method of the homogenization method by Sarada *et al.* (1999) [51]. The *in situ* samples were filtered through 0.7 μm pore size glass fibers, a stainless steel spatula was used to transfer to 50 ml polycarbonate centrifuge tubes and broken up in 50 mmol l^{-1} sodium phosphate buffer (pH 7.0 ±0.2) [29]. The broken filters then went through two rounds of grinding and centrifugation. Finally, the PC concentration of the supernatant was fluorometrically determined using a TD700-fluorometer (Turner Designs, Inc.). For the samples collected in Lake IJsselmeer, the freeze-thawing method of Sarada *et al.* (1999) [51] was used but extended to nine cycles of freezing at -20°C with subsequent high speed centrifugation and spectrophotometric quantification, as described in Simis *et al.* (2005) [22].

D. Phytoplankton Culture Experiments

Cultures of the chlorophyte *Chlorella vulgaris*, the cryptophyta *Cryptomonas ovata*, and the diatom *Cyclotella meneghiniana*, were grown on BG11, AF-6, and CSI media,

respectively. All three algae strains were maintained in a culture cabinet at 25 ±1 °C under a white light intensity of 25 μmol $m^{-2}s^{-1}$ photons on a 12 h : 12 h light/dark cycle. Because none of these algae species contained PC, $a_{ph}(\lambda)$ can be considered equal to $a_{ph-PC}(\lambda)$.

E. Accuracy Assessment

The relative error (RE), mean absolute percentage error (MAPE) and root-mean-square error (RMSE) were used to evaluate the accuracy of the presented algorithms:

$$RE = \left| \frac{value_{observed,i} - value_{predicted,i}}{value_{observed,i}} \right| \times 100\% \qquad (4)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N} \left| \frac{value_{observed,i} - value_{predicted,i}}{value_{observed,i}} \right| \times 100\% \qquad (5)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(value_{observed,i} - value_{predicted,i})^2}{N}} \qquad (6)$$

Where $value_{observed,i}$ is the variable that is considered the reference observed from *in situ* measurements, $value_{predicted,i}$ represents the same variable of interest derived from proposed models or algorithms, and N is the number of observations.

III. THEORY OF THE FBA_PC MODEL

The four bands of FBA_PC were selected based on the following assumptions, which extends the theory of a three-band retrieval algorithm as described by Dall'Olmo [52] and applied by Hunter *et al.* (2010) for PC retrieval [26]:

(1) The first band λ_1 should be maximally sensitive to the targeted optically active substance, in this case PC absorption ($a_{PC}(\lambda_1) \gg a_{PC}(\lambda_2)$, $a_{PC}(\lambda_1) \gg a_{PC}(\lambda_3)$). Waveband λ_1 should thus be narrow and located near the absorption maximum of PC at 615 nm.

(2) The yellow matter absorption contained at λ_1 ($a_{ym}(\lambda_1)$), as well as the non-PC phytoplankton pigment absorption at λ_1 ($a_{ph-PC}(\lambda_1)$) should be captured using a band combination that captures the variability in these absorption components. The contribution to absorption in bands λ_2 and λ_3 is expressed with scaling parameter η that we let vary between -0.1 and 1.1. The function of this parameter is to capture the influence of both $a_{ym}(\lambda_1)$ and $a_{ph-PC}(\lambda_1)$, based on two reflectance bands, in a single algorithm parameter η. Band λ_2 and λ_3 are located at shorter and longer wavelengths than λ_1, respectively:

$$\begin{cases} a_{ym}(\lambda_1) = \eta a_{ym}(\lambda_2) + (1-\eta)a_{ym}(\lambda_3) & (7a) \\ a_{ph-PC}(\lambda_1) = \eta a_{ph-PC}(\lambda_2) + (1-\eta)a_{ph-PC}(\lambda_3) & (7b) \end{cases}$$

Calibration of η is covered in the next section.

(3) The fourth band λ_4 is minimally sensitive to absorption by phytoplankton and yellow matter, whereas the absorption of pure water ($a_w(\lambda_4)$) dominates [22, 52, 53].

(4) The backscattering coefficient $b_b(\lambda)$ is not significantly different over the wavelength range between bands λ_1 to λ_4. Because this assumption will deviate from known spectral dependency of $b_b(\lambda)$ found in many natural waters [22, 29], the best combination of bands will be those found in relatively close proximity, while still isolating the dominant absorption signals.

Mathematically, the FBA_PC indices is defined as:

$$FBA\_PC = \left[\frac{1}{R_{rs}(\lambda_1)} - \frac{\eta}{R_{rs}(\lambda_2)} - \frac{1-\eta}{R_{rs}(\lambda_3)} \right] R_{rs}(\lambda_4) \qquad (8)$$

Applying Eqs. 1-2 to Eq. 8 yields the long form of the algorithm with explicit absorption and backscattering terms for each band:

$$FBA\_PC \propto \left[\frac{a_w(\lambda_1) + a_{ym}(\lambda_1) + a_{PC}(\lambda_1) + a_{ph-PC}(\lambda_1) + b_b(\lambda_1)}{b_b(\lambda_1)} \right.$$
$$- \frac{\eta a_w(\lambda_2) + \eta a_{ym}(\lambda_2) + \eta a_{PC}(\lambda_2) + \eta a_{ph-PC}(\lambda_2) + \eta b_b(\lambda_2)}{b_b(\lambda_2)}$$
$$\left. - \frac{(1-\eta)a_w(\lambda_3)+(1-\eta)a_{ym}(\lambda_3)+(1-\eta)a_{PC}(\lambda_3)+(1-\eta)a_{ph-PC}(\lambda_3)+(1-\eta)b_b(\lambda_3)}{b_b(\lambda_3)} \right]$$
$$\times \frac{b_b(\lambda_4)}{a_w(\lambda_4)+a_{ym}(\lambda_4)+a_{PC}(\lambda_4)+a_{ph-PC}(\lambda_4)+b_b(\lambda_4)} \qquad (9)$$

With assumptions $1-4$, the FBA_PC index is expressed as:
$$FBA\_PC \propto \left[a_{pc}(\lambda_1) + a_w(\lambda_1) - \eta a_w(\lambda_2) - (1-\eta)a_w(\lambda_3) \right]/ a_w(\lambda_4) \qquad (10)$$

The PC concentration can subsequently be derived by dividing $a_{pc}(\lambda_1)$ by the PC-specific absorption coefficient at λ_1 ($a_{PC}^*(\lambda_1)$). From Eq. 10, we can see that the relationship between FBA_PC and the PC concentration is monotonic under the assumptions of FBA_PC.

IV. RESULTS AND DISCUSSION

From Lake IJsselmeer, Lake Taihu, Lake Chaohu, Lake Hengshui, Morse Reservoir, Geist Reservoir, and Eagle Creek Reservoir, there are a total of 431 samples for which reflectance and PC pigment data are available. The PC concentration of 430 samples was < 350 mg m^{-3} while one sample measured 710.29 mg m^{-3}. This sample constitutes an outlier that was omitted from the analysis of the performance of FBA and previous MERIS algorithms. The remaining 430 samples were divided into random subsets of 215 samples for algorithm calibration and 215 samples was used for algorithm validation. In addition, a geographical rather than a random split between the datasets is described in section D, below.

A. Optimizing η In FBA_PC

FBA_PC for MERIS is obtained by adapting MERIS bands (λ_1=620, λ_2=560, λ_3=709, and λ_4=754 nm). The successful, robust application of FBA_PC requires a value of η to hold for a variety of optical conditions found in inland water bodies. For discrete values of η between -0.1 and 1.1 (evaluated at intervals of 0.1), we optimized η with respect to the smallest MAPE while evaluating the term $\eta a_{ym}(560) + (1-\eta)a_{ym}(709)$ to estimate $a_{ym}(620)$ as in Eq. 7a. All absorption data collected from the study sites are used in this analysis. It is then found that for $\eta = 0.4$, the term $0.4a_{ym}(560) + 0.6a_{ym}(709)$ is in good agreement with $a_{ym}(620)$, following a highly significant linear relationship (slope of 1.073, R^2 = 0.998) (Figure 1). The term used to estimate $a_{ym}(620)$ from two bands outperformed previous attempts using a single band, as shown in Figure 1 with the present model yielding 10.368% error compared to 61.942% in HD12. From Table 3 we can see that, in all the 12 study areas, the term $0.4a_{ym}(560) + 0.6a_{ym}(709)$ in FBA_PC has a better performance than only using $a_{ym}(709)$ in HD12 in the estimation of $a_{ym}(620)$.

Next, the optimized $\eta = 0.4$ is tested in the estimation of $a_{ph-PC}(620)$ using absorption data from laboratory-grown

algal cultures. Figure 2 shows how the term $0.4a_{ph-PC}(560) + 0.6a_{ph-PC}(709)$ corresponds to $a_{ph-PC}(620)$ as expressed in Eq. 7b. The MAPEs for *Chlorella vulgaris*, *Cryptomonas ovata*, and *Cyclotella meneghiniana* are 74.268%, 26.631%, and 67.718%, respectively. In comparison, when only $a_{ph-PC}(709)$ is used to estimate $a_{ph-PC}(620)$, the MAPEs for *Chlorella vulgaris*, *Cryptomonas ovata*, and *Cyclotella meneghiniana* are 86.159%, 91.823%, and 85.165%, respectively. The term used to estimate $a_{ph-PC}(620)$ in FBA_PC still has large error for *Chlorella vulgaris* and *Cyclotella meneghiniana*, but improves upon using only $a_{ph-PC}(709)$ in HD12.

B. Calibration Of FBA_PC And Other Algorithms

The calibration dataset with *in situ* PC concentrations ranging from 0.327 to 317.743 mg m^{-3} was used to establish the expression of FBA_PC (Eq. 11). In addition to FBA_PC, we also apply Schalles00, HD12, and Qi14 algorithms. Algorithm calibration of a single algorithm against the present data set would likely lead to an unfair advantage over the other algorithms used in comparison. Therefore, each algorithm was calibrated against the same calibration dataset for an unbiased comparison of PC retrieval skill. From the calibration data, the FBA_PC algorithm was established as Eq. 11 and the expressions of the other three algorithms were shown in Figure 3.

$$\text{FBA\_PC} = \left[\frac{1}{R_{rs}(620)} - \frac{0.4}{R_{rs}(560)} - \frac{0.6}{R_{rs}(709)}\right]R_{rs}(754)$$

$$\text{PC} = 462.5 \times \text{FBA\_PC} + 22.598 \qquad (11)$$

Figure 3 shows strong and similar relationships between each reflectance or absorption indices of FBA_PC, Schalless00, HD12, and the *in situ* PC concentration (R^2 = 0.686, 0.640, 0.700, respectively). The relationship between the reflectance indices of Qi14 and *in situ* PC concentration are weaker (R^2 = 0.095) in comparison.

C. Validation Of FBA_PC And Previous MERIS Algorithms

Statistics on RMSE and MAPE were used to evaluate the performance of each algorithm in their final re-calibrated form. Results illustrated in Figure 4 shows that in terms of RMSE, FBA_PC gave the best performance (RMSE = 27.691 mg m^{-3}), whereas HD12 and Schalles00 algorithms follow closely (RMSE = 30.050 and 31.836 mg m^{-3}, respectively). The Qi14 shows relatively worse performance (RMSE = 48.222 mg m^{-3}). The MAPE shows the clearest differences in algorithm performance with Qi14 > 400%, Schalles00 and HD12 in the 200-300% range (214.856% and 284.199%, respectively) and FBA_PC only giving 172.863%.

Many attempts at PC estimation have shown poor performance in the low PC concentration range in turbid inland waters in previous studies [29, 38, 54]. Therefore, we evaluated the accuracy of each algorithm not only for the whole PC range but also for the low PC range ≤ 50.000 mg m^{-3} (Table 4). The retrieval performance of each algorithm decreased markedly when estimating low PC concentration. The estimated PC values of Qi14 had next to no correlation with the *in situ* PC values (R^2 = 0.028) which illustrates that this algorithm is not

suitable for the low PC concentration range. Schalles00 and HD12 perform relatively better (R^2 = 0.320 and 0.295, respectively, RMSE = 23.282 and 23.656 mg m^{-3}), but the MAPEs of these algorithms exceed 300% (Schalles00 = 323.046%, and HD12 = 431.900%). For FBA_PC, we observed somewhat higher R^2 (0.364), lower RMSE (18.439 mg m^{-3}), and MAPE below 300 % (255.798%).

D. Geographic Transferability Of FBA_PC And Previous MERIS Algorithms

The geographic transferability of FBA_PC and other MERIS algorithms was tested by calibrating the algorithm explicitly using a single dataset (*in situ* data collected from the Netherlands) and validating against other datasets (*in situ* data collected from China and the United States, respectively). The re-calibration results of these algorithms are shown in Figure 5. FBA_PC, Schalles00, and HD12 all had strong and similar relationships between their indices and the *in situ* PC concentration (R^2=0.657, 0.563, and 0.662, respectively). The relationship between the reflectance indices of Qi14 and *in situ* PC concentration are weaker (R^2 = 0.176) in comparison.

The validation results of these algorithms in China are shown in Figure 6. The RMSE of FBA_PC, Schalles00, HD12, and Qi14 were 32.494, 37.980, 34.102, and 63.602 mg m^{-3}, respectively. The MAPE of these four algorithms were 114.846%, 158.157%, 136.417%, and 164.351%, respectively. FBA_PC outperformed the other algorithms both in terms of RMSE or MAPE for this independent dataset. The validation results of these algorithms in the United States are shown in Figure 7. In terms of RMSE, Schalles00 gave the best performance (RMSE = 25.216 mg m^{-3}), followed closely by HD12 and FBA_PC with RMSE = 29.285 and 29.788 mg m^{-3}, respectively. Qi14 showed relatively worse performance with RMSE = 37.609 mg m^{-3}. Although Schalles00 and HD12 both showed relatively lower RMSE, the MAPE of these two algorithms were >100% (Schalles00: 104.783%, and 108.681%). The MAPE of FBA_PC was < 50% (49.183%) illustrating that FBA_PC can get more reliable inversion results than the existing algorithms, for this *in situ* dataset. Figure 6 and Figure 7 both show that the FBA_PC model calibrated using the Dutch *in situ* dataset consistently underestimates PC concentration in both *in situ* dataset collected in China and the United States, especially when the PC concentration is higher than 50 mg m^{-3}.

E. Factors Influencing The Performance Of PC Retrieval From MERIS/OLCI Bands

Interference from other optically active constituents (OACs) such as yellow matter and non-PC pigments will cause error in estimated PC [22, 35, 38, 55]. The poor performance of Schalles00 indicates that directly omitting the signal of other optically constituents (OACs) may result in large instabilities in estimated PC concentration. The low correlation between retrieved PC concentrations from Qi14 and *in situ* PC concentrations also shows that the "baseline" method may not have the PC estimation ability at broad geographic scales. With the exception of the Schalles00 and Qi14 algorithms, FBA_PC

and HD12 algorithms described in this paper both have specific ways to remove the signal of other OACs and isolate the PC absorption signal (a_{PC}).

To investigate the influence of yellow matter and non-PC pigments absorption to the accuracy of PC estimation for FBA_PC and HD12 algorithms, the relationship between the REs of these algorithms and two ratios of $a_{ym}(620)$: PC and $a_{ph}(620)$: PC are shown in Figure 8. HD12 showed significant correlations with the ratios $a_{ym}(620)$: PC ($R^2 = 0.627$, p<0.05). When $a_{ym}(620)$: PC > 0.1, it is clear that the REs of HD12 increased with $a_{ym}(620)$: PC, so that $a_{ym}(620)$ influences the error in retrieved PC. For the FBA_PC algorithm, the REs showed a positive correlation with the ratio $a_{ym}(620)$: PC, but the relationship is no longer significant ($R^2 = 0.221$, p<0.05). Figure 5b shows the relationship between REs and the ratios $a_{ph}(620)$: PC of FBA_PC and HD12 algorithms, respectively. HD12 gave a weaker correlation ($R^2 = 0.282$, p<0.05), while FBA_PC outperformed HD12 ($R^2 = 0.374$, p<0.05). The analysis of laboratory-grown cultures described in the previous section have slightly improved the removal of $a_{ph-PC}(620)$ in FBA_PC compared to HD12, based on three species of algae (Figure 2). However, the *in situ* data showed opposite results and the influence of the absorption signal of phycoerythrin (PE) of some species algae of Cyanophyta at ~565 nm may be the reason. Using the current set of wavebands, FBA_PC may have large uncertainties in the estimation of PC when PE-bearing phytoplankton is present.

We only measured $a_{ph}(\lambda)$ of algae for the laboratory-grown cultures and *in situ* samples. Therefore, we cannot evaluate the performance of the term $0.4a_{ph-PC}(560) + 0.6a_{ph-PC}(709)$ for the estimation of $a_{ph-PC}(620)$ for Cyanophyta which contain PC and PE. Ideally this should be included in further work, potentially targeting additional satellite wavebands on future sensors which may better capture natural variability in both PC and PE.

V. Conclusion

FBA_PC, an extension of HD12 algorithm was developed to estimate PC concentration in turbid inland waters. *In situ* absorption data collected from various inland water bodies demonstrate the feasibility of estimating absorption terms of non-PC components at 620 nm from a single band combination, with a robust parameterization between the study sites. FBA_PC shows improved performance over three previously described algorithms for MERIS/OLCI.

Acknowledgment

References

[1] J. Huisman, H. C. Matthijs, and P. M. Visser, *Harmful cyanobacteria*: Springer Science & Business Media, 2006.

[2] A. M. Michalak, E. J. Anderson, D. Beletsky *et al.*, "Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions," *Proceedings of the National Academy of Sciences,* vol. 110, no. 16, pp. 6448-6452, 2013.

[3] H. W. Paerl, and J. Huisman, "Climate change: a catalyst for global expansion of harmful cyanobacterial blooms," *Environmental Microbiology Reports,* vol. 1, no. 1, pp. 27-37, 2009.

[4] H. W. Paerl, and V. J. Paul, "Climate change: links to global expansion of harmful cyanobacteria," *Water research,* vol. 46, no. 5, pp. 1349-1363, 2012.

[5] W. W. Carmichael, "Health effects of toxin-producing cyanobacteria:"The CyanoHABs","" *Human and ecological risk assessment: An International Journal,* vol. 7, no. 5, pp. 1393-1407, 2001.

[6] G. A. Codd, S. G. BELL, K. Kaya *et al.*, "Cyanobacterial toxins, exposure routes and human health," *European Journal of Phycology,* vol. 34, no. 04, pp. 405-415, 1999.

[7] L. Guo, "Doing battle with the green monster of Taihu Lake," *Science,* vol. 317, no. 5842, pp. 1166-1166, 2007.

[8] F. Micheli, "Eutrophication, fisheries, and consumer-resource dynamics in marine pelagic ecosystems," *Science,* vol. 285, no. 5432, pp. 1396-1398, 1999.

[9] R. J. Diaz, and R. Rosenberg, "Spreading dead zones and consequences for marine ecosystems," *science,* vol. 321, no. 5891, pp. 926-929, 2008.

[10] B. Qin, W. Li, G. Zhu *et al.*, "Cyanobacterial bloom management through integrated monitoring and forecasting in large shallow eutrophic Lake Taihu (China)," *Journal of hazardous materials,* vol. 287, pp. 356-363, 2015.

[11] P. Vonlanthen, D. Bittner, A. Hudson *et al.*, "Eutrophication causes speciation reversal in whitefish adaptive radiations," *Nature,* vol. 482, no. 7385, pp. 357-362, 2012.

[12] D. Bryant, "The cyanobacterial photosynthetic apparatus: comparison to those of higher plants and photosynthetic bacteria," *Can. Bull. Fish. Aquat. Sci,* vol. 214, pp. 423-500, 1986.

[13] H. GORDON, J. BROWN, O. BROWN *et al.*, "A semianalytic radiance model of ocean color," *Journal of Geophysical Research,* vol. 93, no. D9, pp. 10909-10924, 1988.

[14] A. Morel, and B. Gentili, "Diffuse reflectance of oceanic waters: its dependence on Sun angle as influenced by the molecular scattering contribution," *Applied optics,* vol. 30, no. 30, pp. 4427-4438, 1991.

[15] A. Morel, and B. Gentili, "Diffuse reflectance of oceanic waters. II. Bidirectional aspects," *Applied Optics,* vol. 32, no. 33, pp. 6864-6879, 1993.

[16] A. Morel, and B. Gentili, "Diffuse reflectance of oceanic waters. III. Implication of bidirectionality for the remote-sensing problem," *Applied Optics,* vol. 35, no. 24, pp. 4850-4862, 1996.

[17] A. G. Dekker, "Detection of optical water quality parameters for eutrophic waters by high resolution remote sensing," 1993.

[18] L. Qi, C. Hu, H. Duan *et al.*, "A novel MERIS algorithm to derive cyanobacterial phycocyanin pigment concentrations in a eutrophic lake: Theoretical basis and practical considerations," *Remote Sensing of Environment,* vol. 154, no. 154, pp. 298-317, 2014.

[19] J. F. Schalles, and Y. Z. Yacobi, "Remote detection and seasonal patterns of phycocyanin, carotenoid and chlorophyll pigments in eutrophic waters," *Ergebnisse Der Limnologie,* vol. 55, pp. 153-168, 2000.

[20] S. Mishra, D. R. Mishra, and W. M. Schluchter, "A Novel Algorithm for Predicting Phycocyanin Concentrations in Cyanobacteria: A Proximal Hyperspectral Remote Sensing Approach," *Remote Sensing,* vol. 1, no. 4, pp. 758-775, 2009.

[21] M. Woźniak, K. M. Bradtke, M. Darecki *et al.*, "Empirical Model for Phycocyanin Concentration Estimation as an Indicator of Cyanobacterial Bloom in the Optically Complex Coastal Waters of the Baltic Sea," *Remote Sensing,* vol. 8, no. 3, pp. 212, 2016.

[22] S. G. Simis, S. W. Peters, and H. J. Gons, "Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water," *Limnology and Oceanography,* vol. 50, no. 1, pp. 237-245, 2005.

[23] H. R. Gordon, O. B. Brown, and M. M. Jacobs, "Computed relationships between the inherent and apparent optical properties of a flat homogeneous ocean," *Applied optics,* vol. 14, no. 2, pp. 417-427, 1975.

[24] H. J. Gons, M. Rijkeboer, and K. G. Ruddick, "Effect of a waveband shift on chlorophyll retrieval from MERIS imagery of inland and coastal waters," *Journal of Plankton Research,* vol. 27, no. 1, pp. 125-127, 2005.

[25] G. Dall'Olmo, A. A. Gitelson, and D. C. Rundquist, "Towards a unified

approach for remote estimation of chlorophyll‐a in both terrestrial vegetation and turbid productive waters,” *Geophysical Research Letters,* vol. 30, no. 18, 2003.

[26] P. D. Hunter, A. N. Tyler, L. Carvalho *et al.*, "Hyperspectral remote sensing of cyanobacterial pigments as indicators for cell populations and toxins in eutrophic lakes," *Remote Sensing of Environment,* vol. 114, no. 11, pp. 2705-2718, 2010.

[27] P. D. Hunter, A. N. Tyler, M. Présing *et al.*, "Spectral discrimination of phytoplankton colour groups: The effect of suspended particulate matter and sensor spectral resolution," *Remote Sensing of Environment,* vol. 112, no. 4, pp. 1527-1544, 2008.

[28] H. Duan, R. Ma, and C. Hu, "Evaluation of remote sensing algorithms for cyanobacterial pigment retrievals during spring bloom formation in several lakes of East China," *Remote Sensing of Environment,* vol. 126, pp. 126-135, 2012.

[29] L. Li, L. Li, K. Shi *et al.*, "A semi-analytical algorithm for remote estimation of phycocyanin in inland waters," *Science of the Total Environment,* vol. 435, pp. 141-150, 2012.

[30] S. Mishra, and D. R. Mishra, "A novel remote sensing algorithm to quantify phycocyanin in cyanobacterial algal blooms," *Environmental Research Letters,* vol. 9, no. 11, 2014.

[31] M. Rast, J. Bezy, and S. Bruzzi, "The ESA Medium Resolution Imaging Spectrometer MERIS a review of the instrument and its mission," *International Journal of Remote Sensing,* vol. 20, no. 9, pp. 1681-1702, 1999.

[32] S. C. Palmer, P. D. Hunter, T. Lankester *et al.*, "Validation of Envisat MERIS algorithms for chlorophyll retrieval in a large, turbid and optically-complex shallow lake," *Remote Sensing of Environment,* vol. 157, pp. 158-169, 2015.

[33] I. Ogashawara, D. R. Mishra, S. Mishra *et al.*, "A performance review of reflectance based algorithms for predicting phycocyanin concentrations in inland waters," *Remote Sensing,* vol. 5, no. 10, pp. 4774-4798, 2013.

[34] A. Ruiz-Verdú, S. G. H. Simis, C. d. Hoyos *et al.*, "An evaluation of algorithms for the remote sensing of cyanobacterial biomass," *Remote Sensing of Environment,* vol. 112, no. 11, pp. 3996-4008, 2008.

[35] K. Randolph, J. Wilson, L. Tedesco *et al.*, "Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments, chlorophyll a and phycocyanin," *Remote Sensing of Environment,* vol. 112, no. 11, pp. 4009-4019, 2008.

[36] L. Li, L. Li, K. Song *et al.*, "An inversion model for deriving inherent optical properties of inland waters: Establishment, validation and application," *Remote Sensing of Environment,* vol. 135, pp. 150-166, 2013.

[37] M. Chawira, "Monitoring blue-green algae in the IJsselmeer using remote sensing and in-situ measurements," 2012.

[38] S. G. Simis, A. Ruiz-Verdú, J. A. Domínguez-Gómez *et al.*, "Influence of phytoplankton pigment composition on remote sensing of cyanobacterial biomass," *Remote Sensing of Environment,* vol. 106, no. 4, pp. 414-427, 2007.

[39] M. Scheffer, S. Rinaldi, A. Gragnani *et al.*, "On the dominance of filamentous cyanobacteria in shallow, turbid lakes," *Ecology,* vol. 78, no. 1, pp. 272-282, 1997.

[40] C. Hu, Z. Lee, R. Ma *et al.*, "Moderate resolution imaging spectroradiometer (MODIS) observations of cyanobacteria blooms in Taihu Lake, China," *Journal of Geophysical Research: Oceans (1978–2012),* vol. 115, no. C4, 2010.

[41] X. Yang, N. Anderson, X. Dong *et al.*, "Surface sediment diatom assemblages and epilimnetic total phosphorus in large, shallow lakes of the Yangtze floodplain: their relationships and implications for assessing long‐term eutrophication," *Freshwater Biology,* vol. 53, no. 7, pp. 1273-1290, 2008.

[42] J. Luo, J. Yin, L. Cai *et al.*, "Freshwater fungi in Lake Dianchi, a heavily polluted lake in Yunnan, China," *Fungal Diversity*, 2004.

[43] H. Li, G. Hou, F. Dakui *et al.*, "Prediction and elucidation of the population dynamics of Microcystis spp. in Lake Dianchi (China) by means of artificial neural networks," *Ecological Informatics,* vol. 2, no. 2, pp. 184-192, 2007.

[44] Y. Shen, Y. Liu, G. Wu *et al.*, "Mechanical removal of heavy cyanobacterial bloom in the hyper-eutrophic lake Dianchi," *Acta hydrobiologica Sinica/Chung-kuo ko hsueh yuan shui sheng sheng wu yen chiu so,* vol. 28, no. 2, pp. 131-136, 2003.

[45] F. Li, J. Huang, G. Zeng *et al.*, "Spatial risk assessment and sources identification of heavy metals in surface sediments from the Dongting Lake, Middle China," *Journal of Geochemical Exploration,* vol. 132, pp. 75-83, 2013.

[46] H. Guo, L. Liu, G. Huang *et al.*, "A system dynamics approach for regional environmental planning and management: a study for the Lake Erhai Basin," *Journal of Environmental Management,* vol. 61, no. 1, pp. 93-111, 2001.

[47] Q. Cai, and Z. Hu, "Studies on eutrophication problem and control strategy in the Three Gorges Reservoir," *Acta hydrobiologica sinica/Shuisheng Shengwu Xuebao,* vol. 30, no. 1, pp. 7-11, 2006.

[48] A. A. Gitelson, G. Dall'Olmo, W. Moses *et al.*, "A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: Validation," *Remote Sensing of Environment,* vol. 112, no. 9, pp. 3582-3593, 2008.

[49] J. L. Mueller, C. Davis, R. Arnone *et al.*, "Above-water radiance and remote sensing reflectance measurements and analysis protocols," *Ocean Optics protocols for satellite ocean color sensor validation Revision,* vol. 2, pp. 98-107, 2000.

[50] M. P. Handbook, "ESA," Paris, 2002.

[51] R. Sarada, M. G. Pillai, and G. Ravishankar, "Phycocyanin from Spirulina sp: influence of processing of biomass on phycocyanin yield, analysis of efficacy of extraction methods and stability studies on phycocyanin," *Process Biochemistry,* vol. 34, no. 8, pp. 795-801, 1999.

[52] G. Dall'Olmo, and A. A. Gitelson, "Effect of bio-optical parameter variability on the remote estimation of chlorophyll-a concentration in turbid productive waters: experimental results," *Applied optics,* vol. 44, no. 3, pp. 412-422, 2005.

[53] D. Gurlin, A. A. Gitelson, and W. J. Moses, "Remote estimation of chl-a concentration in turbid productive waters—Return to a simple two-band NIR-red model?," *Remote Sensing of Environment,* vol. 115, no. 12, pp. 3479-3490, 2011.

[54] A. Ruiz-Verdú, S. G. Simis, C. de Hoyos *et al.*, "An evaluation of algorithms for the remote sensing of cyanobacterial biomass," *Remote Sensing of Environment,* vol. 112, no. 11, pp. 3996-4008, 2008.

[55] L. Li, L. Li, and K. Song, "Remote sensing of freshwater cyanobacteria: An extended IOP Inversion Model of Inland Waters (IIMIW) for partitioning absorption coefficient and estimating phycocyanin," *Remote Sensing of Environment,* vol. 157, pp. 9-23, 2015.

Table 1. Descriptive statistics of water quality parameters in Eagle Creek Reservoir (ECR), Geist Reservoir (GR), Morse Reservoir (MR), Lake IJsselmeer (LIJ), Lake Taihu (LTH), Lake Chaohu (LCH) and Lake Hengshui (LHS).

Location		Chl$a$ (mg m^{-3})	PC (mg m^{-3})
	Min	22.029	8.580
ECR (n=60)	Mean	55.194	32.588
	Max	122.573	72.240
	Min	19.835	6.570
GR (n=37)	Mean	38.697	69.314
	Max	62.121	157.200
	Min	1.850	1.388
MR (n=54)	Mean	60.209	66.439
	Max	129.389	146.100
	Min	1.899	0.327
LIJ (n=187)	Mean	334.175	710.287
	Max	44.616	43.089
	Min	4.338	4.005
LTH (n=23)	Mean	189.586	61.739
	Max	48.617	299.443
	Min	13.578	5.004
LCH(n=65)	Mean	37.607	29.383
	Max	98.184	217.674
	Min	37.2	22.708
LHS (n=9)	Mean	52.533	36.617
	Max	43.133	50.153

Table 2. Statistics of $a_{ph}(620)$, $a_{ym}(620)$ and $a_{ph}(620)$: $a_{ym}(620)$ values from the Eagle Creek Reservoir (ECR), Geist Reservoir (GR), Morse Reservoir (MR), Lake IJsselmeer (LIJ), Lake Taihu (LTH), Lake Hengshui (LHS), Lake Chaohu (LCH), Lake Dianchi(LDC), Lake Dongting (LDT), Lake Erhai (LEH), Jiajiang Reservoir (JR), and Three Gorges Reservoir (TGR).

Location		$a_{ph}(620)$ (m^{-1})	$a_{ym}(620)$ (m^{-1})	$a_{ph}(620)/a_{ym}(620)$
ECR (n=81)	Min	0.0270	0.0133	0.250
	Mean	0.250	0.153	3.252
	Max	0.640	0.579	8.444
GR (n=58)	Min	0.052	0.018	0.540
	Mean	0.336	0.143	2.747
	Max	0.616	0.786	5.368
MR (n=54)	Min	0.088	0.190	0.159
	Mean	0.976	0.408	1.630
	Max	0.588	0.617	3.788
LIJ (n=105)	Min	0.0162	0.029	0.054
	Mean	0.366	0.355	1.402
	Max	0.983	1.397	7.838
LTH (n=224)	Min	0.005	0.028	0.002
	Mean	0.741	0.638	1.709
	Max	6.532	4.861	15.808
LCH (n=62)	Min	0.051	0.124	0.267
	Mean	0.804	0.514	1.584
	Max	4.293	1.294	5.566
LHS (n=9)	Min	0.059	1.935	0.027
	Mean	0.905	3.047	0.289
	Max	1.124	3.928	0.436
LDC (n=33)	Min	0.472	0.081	1.140
	Mean	1.316	0.328	4.850
	Max	2.545	0.592	15.684
LDT (n=87)	Min	0.148	0.076	0.157
	Mean	1.652	1.088	2.271
	Max	8.462	3.361	7.434
LEH (n=42)	Min	0.0375	0.009	0.491
	Mean	0.074	0.049	2.083
	Max	0.160	0.096	6.464
JR (n=29)	Min	0.787	3.729	0.079
	Mean	2.591	7.268	0.400
	Max	5.740	17.648	1.197
TGR (n=25)	Min	0.020	0.0367	0.0487
	Mean	0.331	0.395	1.410
	Max	1.435	0.939	5.280

Table 3. The MRE of FBA_PC and HD12 in the estimation of $a_{ym}(620)$ with their specific assumption in Eagle Creek Reservoir (ECR), Geist Reservoir (GR), Morse Reservoir (MR), Lake IJsselmeer (LIJ), Lake Taihu (LTH) , Lake Hengshui (LHS), Lake Chaohu (LCH), Lake Dianchi(LDC), Lake Dongting (LDT), , Lake Erhai (LEH), Jiajiang Reservoir (JR), and Three Gorges Reservoir (TGR), respectively.

Location	MRE(%) of FBA_PC	MRE(%) of HD12
ECR (n=81)	2.780	75.871
GR (n=58)	5.171	79.131
MR (n=54)	3.297	75.394
LIJ (n=105)	7.347	58.566
LTH (n=224)	14.633	53.169
LCH (n=62)	5.105	37.757
LHS (n=9)	31.841	46.370
LDC(n=33)	28.780	57.973
LDT (n=87)	13.707	64.990
LEH (n=42)	17.186	90.963
JR (n=29)	9.940	30.199
TGR (n=25)	21.009	73.532

Table 4. Error analysis of empirical and semi-analytical algorithms for estimating PC. Numbers in side square brackets are the PC ranges (mg m^{-3}) when the algorithm was presented.

Algorithm	Range of PC concentration (mg m^{-3})					
	$0 \leq PC \leq 329.411$			$0 \leq PC \leq 50.0$		
	R^2	MAPE (%)	RMSE (mg m^{-3})	R^2	MAPE (%)	RMSE (mg m^{-3})
FBA_PC [0.33-710.29]	0.730	172.863	27.691	0.364	255.798	18.439
Schalles00 [0-530]	0.630	214.856	31.836	0.320	323.046	23.282
HD12 [0.09-7.71]	0.701	284.199	30.050	0.295	431.900	23.656
Qi14 [1.58-263.71]	0.178	425.730	48.222	0.028	644.655	30.379
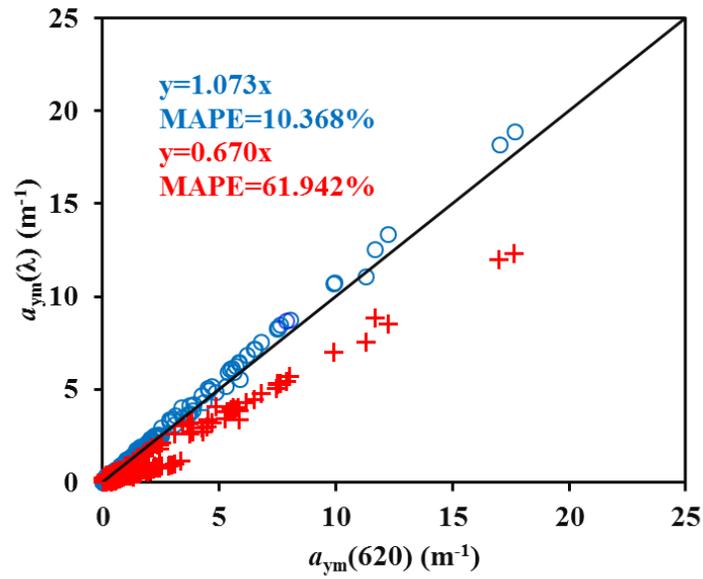
Figure 1. Scatterplot of $a_{ym}(709)$ (blue "circle") and $0.4a_{ym}(560) + 0.6a_{cdm}(709)$ (red "plus") versus $a_{ym}(620)$. The solid line is 1:1 line.
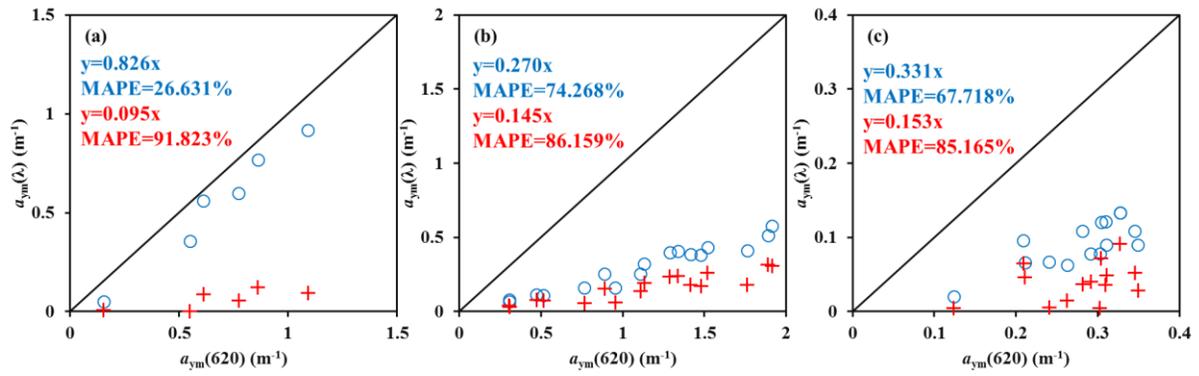
Figure 2. Scatterplot of $a_{ph-PC}(709)$ (blue "circle") and $0.4a_{ph-PC}(560) + 0.6a_{ph-PC}(709)$ (red "plus") versus $a_{ph-PC}(620)$ for (a) *Chlorella vulgaris*, (b) *Cryptomonas ovata*, and (c) *Cyclotella meneghiniana*. The solid lines are the 1:1 lines.
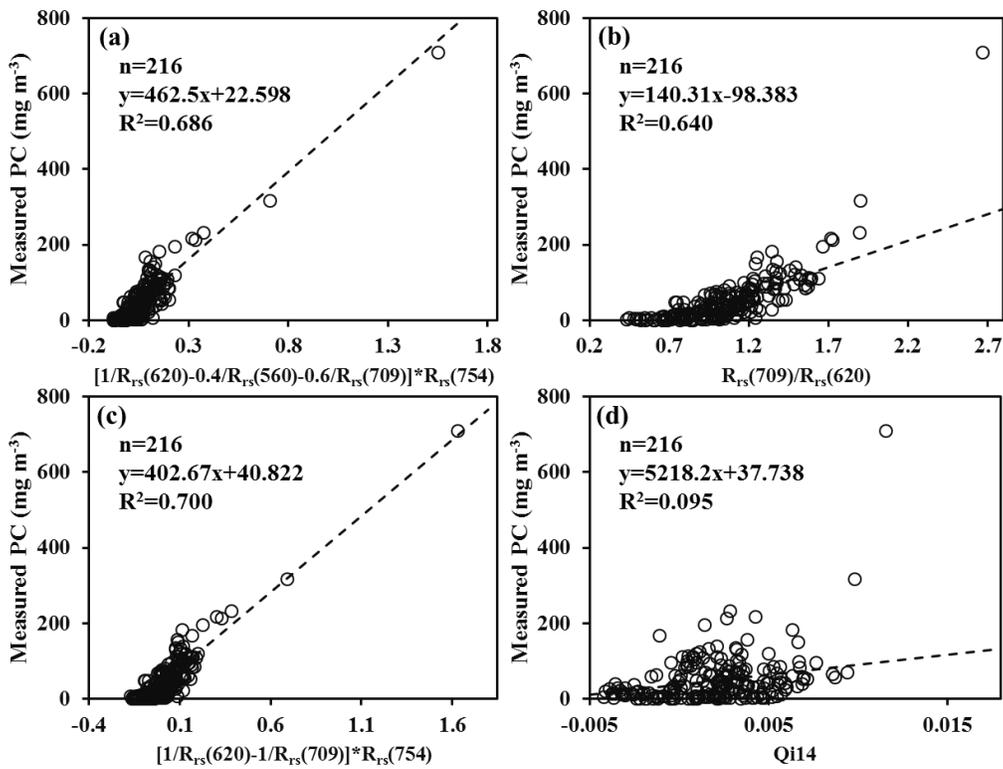
Figure 3. Indices of the models adjusted for the MERIS wavelengths versus PC concentrations: (a) FBA_PC, (b) Schalles00, (c) HD12, (d) Qi14. The dashed lines are the linear regression lines. The extreme value point in the dotted circle is not included in the model calibration.
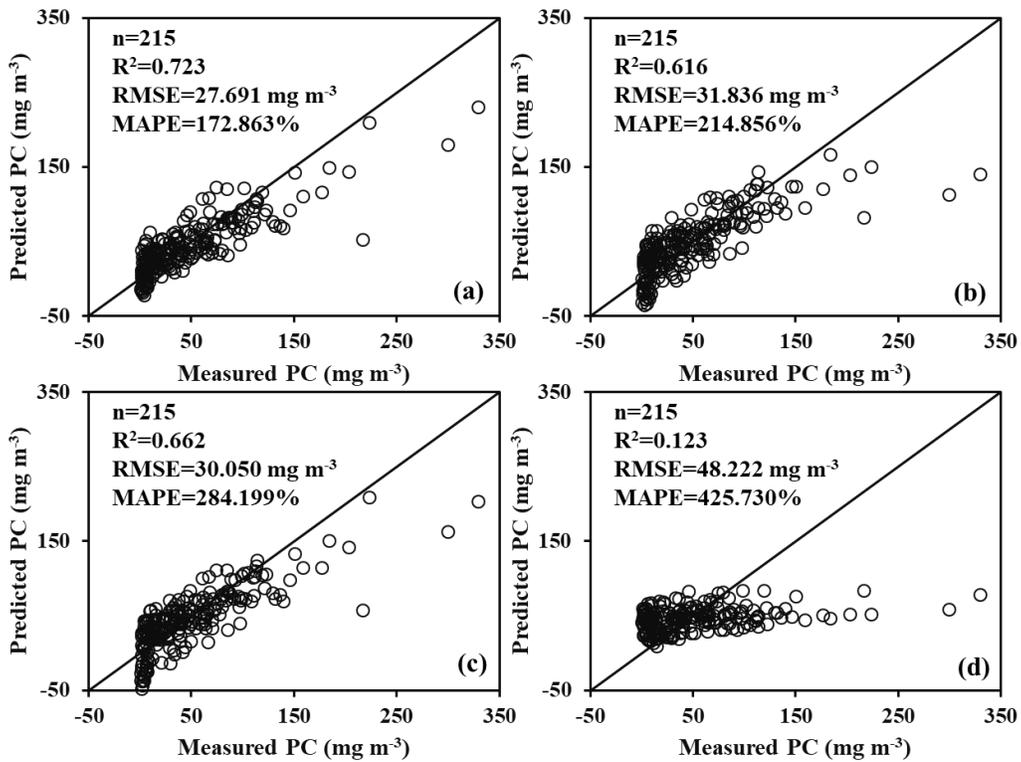
Figure 4. Correlations between the measured and estimated PC concentrations by (a) FBA_PC, (b) Schalles00, (c) HD12, (d) Qi14. The solid lines are the 1:1 lines.
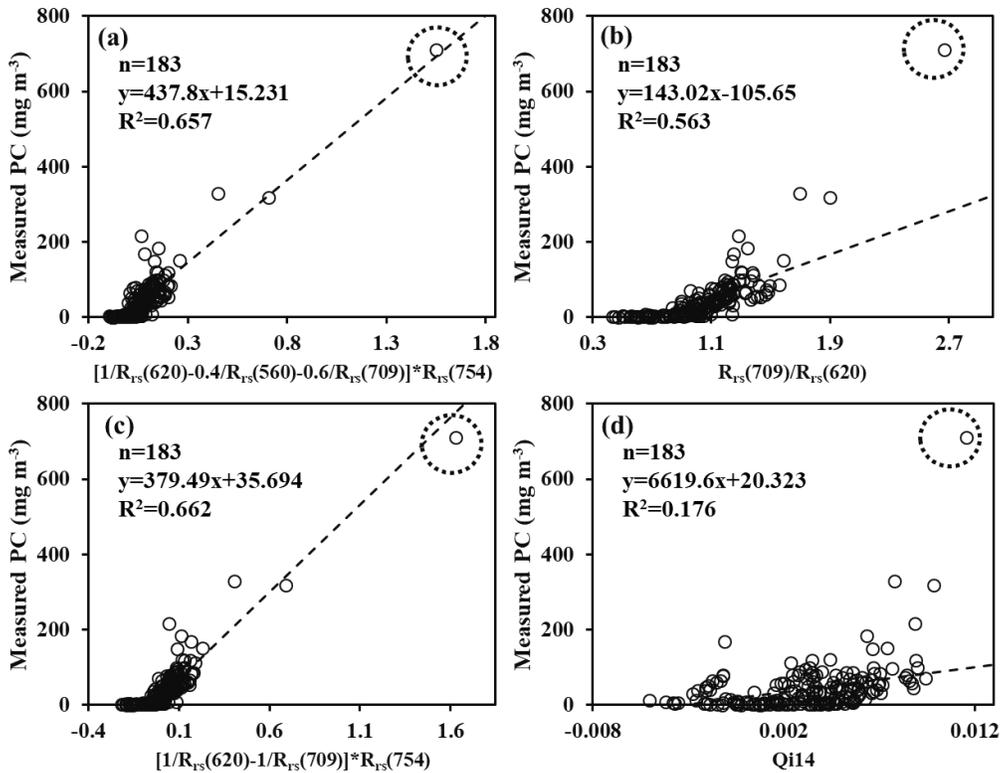
Figure 5. Indices of the models adjusted for the MERIS/OLCI wavelengths versus PC concentrations: (a) FBA_PC, (b) Schalles00, (c) HD12, (d) Qi14 only using *in situ* data collected in The Netherlands. The dashed lines are the linear regression lines. The extreme value point in the dotted circle is not included in the model calibration.
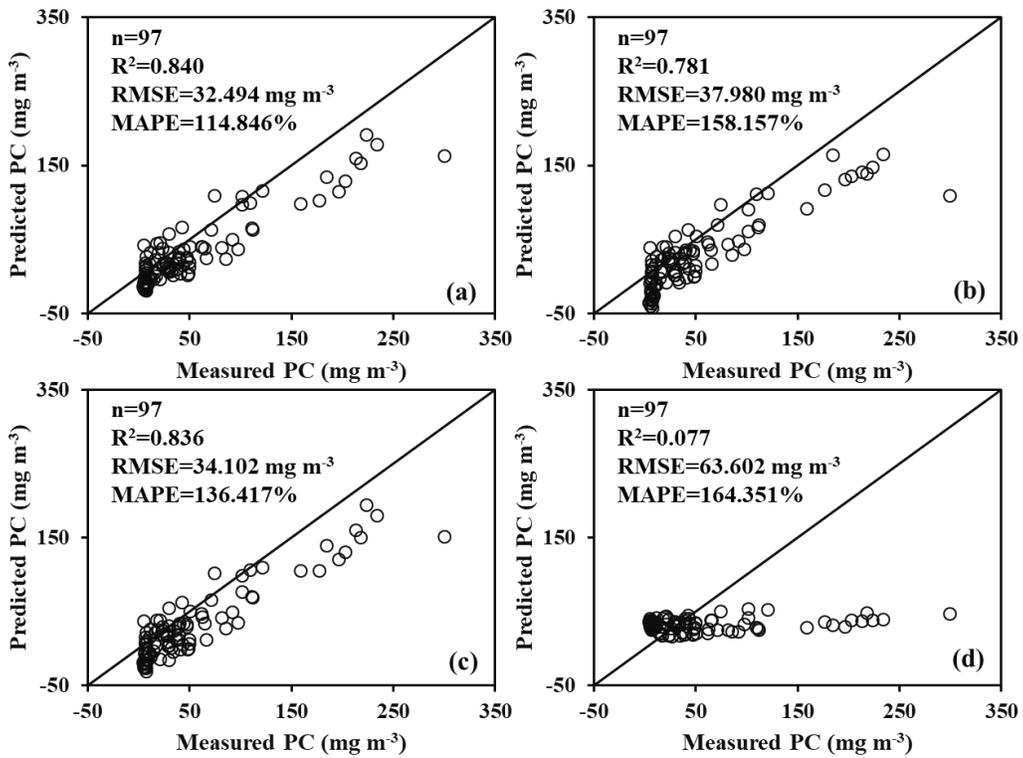
Figure 6. Correlations between the measured and estimated PC concentrations by (a) FBA_PC, (b) Schalles00, (c) HD12, (d) Qi14 for the *in situ* data collected in China. The solid lines are the 1:1 lines.
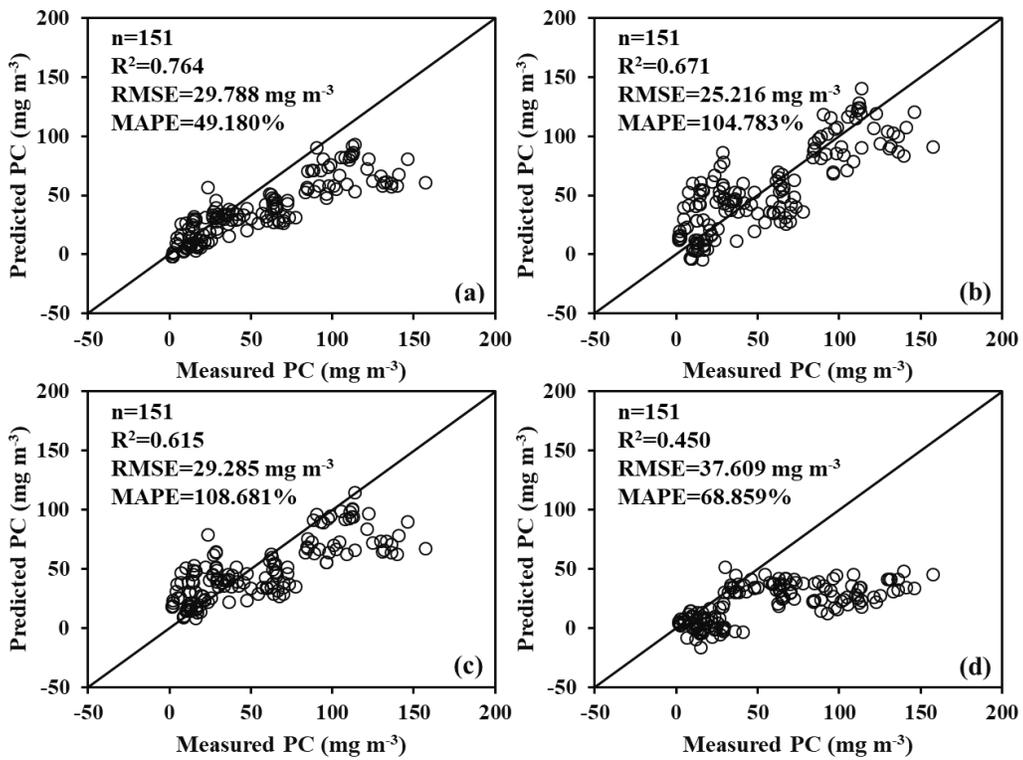
Figure 7. Correlations between the measured and estimated PC concentrations by (a) FBA_PC, (b) Schalles00, (c) HD12, (d) Qi14 for the *in situ* data collected in the United States. The solid lines are the 1:1 lines.
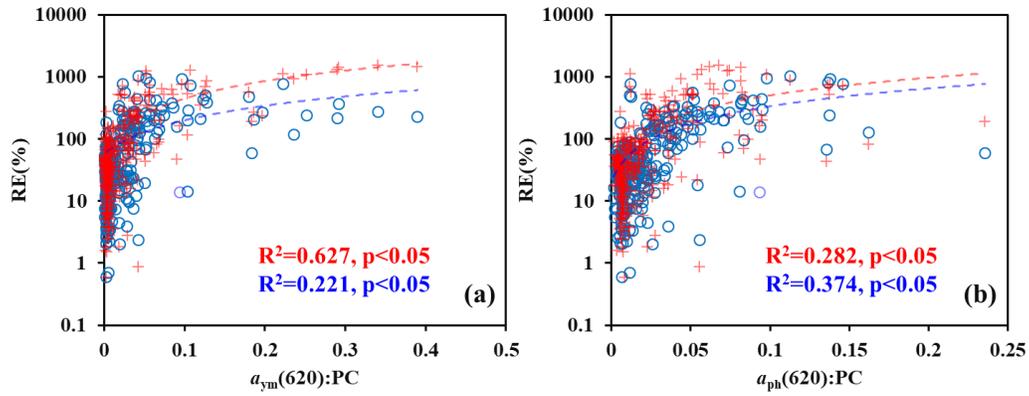
Figure 8. Correlations between the relative error (RE) and the possible interfering factors: (a) $a_{ym}(620)$: PC, (b) $a_{ph}(620)$: PC of different PC estimation models (FBA_PC in blue "circle" and HD12 in red "plus"). The dashed lines are the linear regression lines.