Remote Sensing of Environment 112 (2008) 3996–4008

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

An evaluation of algorithms for the remote sensing of cyanobacterial biomass

Antonio Ruiz-Verdú [a,b,*], Stefan G.H. Simis [c], Caridad de Hoyos [a], Herman J. Gons [d,1], Ramón Peña-Martínez [a]

[a] Centre for Hydrographic Studies of CEDEX, Paseo Bajo de la Virgen del Puerto, 3 E-28005 Madrid, Spain
[b] National Institute for Aerospace Technology (INTA), Carretera de Ajalvir km. 4.500, E-28850 Torrejón de Ardoz, Spain
[c] Finnish Institute of Marine Research (FIMR), Eric Palménin Aukio 1, 00560 Helsinki, Finland
[d] Netherlands Institute of Ecology (NIOO-KNAW), Centre for Limnology, Rijksstraatweg 6, 3631 AC Nieuwersluis, The Netherlands

ARTICLE INFO

Article history:
Received 2 April 2007
Received in revised form 19 October 2007
Accepted 22 November 2007

Keywords:
Cyanobacteria
Phytoplankton
Phycocyanin
Pigments
Absorption
Remote sensing
Reflectance

ABSTRACT

Most remote sensing algorithms for phytoplankton in inland waters aim at the retrieval of the pigment chlorophyll *a* (Chl *a*), as this pigment is a useful proxy for phytoplankton biomass. More recently, algorithms have been developed to quantify the pigment phycocyanin (PC), which is characteristic of cyanobacteria, a phytoplankton group of relative importance to inland water management due to their negative impact on water quality in response to eutrophication.

We evaluated the accuracy of three published algorithms for the remote sensing of PC in inland waters, using an extensive database of field radiometric and pigment data obtained in the Netherlands and Spain in the period 2001–2005. The three algorithms (a baseline, single band ratio, and a nested band ratio approach) all target the PC absorption effect observed in reflectance spectra in the 620 nm region. We evaluated the sensitivity of the algorithms to errors in reflectance measurements and investigated their performance in cyanobacteria-dominated water bodies as well as in the presence of other phytoplankton pigments.

All algorithms performed best in moderate to high PC concentrations (50–200 mg m^{-3}) and showed the most linear response to increasing PC in cyanobacteria-dominated waters. The highest errors showed at PC<50 mg m^{-3}. In eutrophic waters, the presence of other pigments explained a tendency to overestimate the PC concentration. In oligotrophic waters, negative PC predictions were observed. At very high concentrations (PC>200 mg m^{-3}), PC underestimations by the baseline and single band ratio algorithms were attributed to a non-linear relationship between PC and absorption in the 620 nm region. The nested band ratio gave the overall best fit between predicted and measured PC. For the Spanish dataset, a stable ratio of PC over cyanobacterial Chl *a* was observed, suggesting that PC is indeed a good proxy for cyanobacterial biomass. The single reflectance ratio was the only algorithm insensitive to changes in the amplitude of reflectance spectra, which were observed as a result of different measurement methodologies.

1. Introduction

Cyanobacteria are a group of widely distributed photosynthetic microorganisms common to inland water phytoplankton communities (Gibson & Smith, 1982). They are usually found in eutrophic waters, where they can come to dominate the phytoplankton due to several key adaptations found across cyanobacterial taxa: buoyancy regulation, elementary nitrogen fixing capability, and efficient use of yellow-orange light for photosynthesis (Reynolds & Walsby, 1975). Under favourable growth conditions they may form massive blooms that severely affect water quality for human use ranging from recreation and fishing to drinking water production (Codd et al., 1999a). These negative effects are often emphasized as several cyanobacteria species produce toxins that can affect human health (Codd et al., 1999b). Because many inland waters are under stress from eutrophication, regular monitoring of cyanobacterial biomass and the early detection of potential harmful blooms is of importance for the management of inland waters.

All cyanobacteria contain at least a low concentration of the accessory photosynthetic phycobilin pigment phycocyanin (PC), which only presents itself in high concentrations in this taxonomic group. Although its intracellular content can vary between species and in response to environmental conditions (Tandeau de Marsac & Houmard, 1988), it is the main diagnostic pigment for the presence of cyanobacteria and possibly a useful indicator of cyanobacterial biomass. Unfortunately no extraction or detection method for phycobilin pigments has risen to a standard, so they are often ignored in large scale monitoring programs.

The light absorption properties of PC make the pigment a good candidate for optical detection methods, including remote sensing

* Corresponding author. Present address: National Institute for Aerospace Technology (INTA), Carretera de Ajalvir km. 4.500, E-28850 Torrejón de Ardoz, Spain. Tel.: +34 91 520 1513; fax: +34 91 520 1633.
E-mail address: ruizva@inta.es (A. Ruiz-Verdú).
[1] Deceased.

0034-4257/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2007.11.019

when suitable spectral resolution is available. The maximum absorption *in vivo* (around 615 nm, Bryant, 1981) is located in a region of limited algal absorption, removed from the bulk of absorption of carotenoid pigments and chlorophylls, although the latter exhibit secondary absorption features in the same spectral region as PC (Ficek et al., 2004; Jeffrey et al., 1997; Sathyendranath et al., 1987). When cyanobacteria dominate, the PC absorption feature is easily identified as an enhanced trough in reflectance spectra between 600 and 650 nm (Fig. 1A; Dekker et al., 1991; Gons et al., 1992; Jupp et al., 1994; Kutser et al., 2006) and as a peak in phytoplankton absorption spectra (Fig. 1B). In the presence of other phytoplankton groups, accessory chlorophyll pigments overlap with PC absorption: Chlorophyll *b* (Chl *b*), present in green algae, has a broad maximum around 650 nm and a secondary around 600 nm, the diatom Chlorophylls *c*1 and *c*2 (Chl *c*) have secondary absorption maxima around 590 nm and 640 nm (Ficek et al., 2004; Jeffrey et al., 1997; Sathyendranath et al., 1987). Chlorophyll *a* (Chl *a*), present in all phytoplankton as the main photosynthetic pigment, has an absorption shoulder centred around 623 nm (Bidigare et al., 1990; Ficek et al., 2004; Sathyendranath et al., 1987). The effect of these pigments in the PC absorption region cannot be neglected when interpreting absorption or reflectance spectra with the aim to identify or quantify a cyanobacterial component. Similarly, the absorption by other water constituents (suspended non-algal particles, detritus, and coloured dissolved organic matter) and water itself cannot be ignored as they can reach the same order of magnitude as PC absorption (Dekker, 1993).

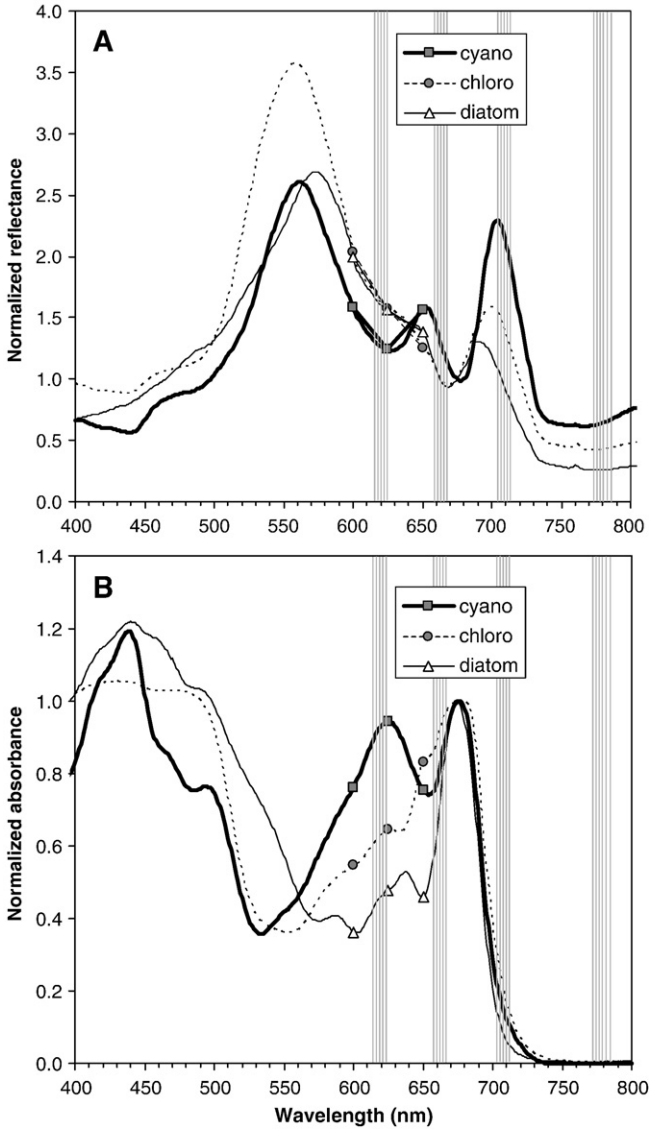


Fig. 1. (A). Reflectance spectra normalized at 675 nm, measured in Spanish water bodies where a single taxonomic group accounted for more than 95% of phytoplankton biomass. Legend: cyano=Cyanobacteria; chloro=Chlorophyceae (green algae); diatom=Bacillariophyceae. The Chl *a* concentrations determined by HPLC were 58.3; 68.9 and 20.7 mg m^{-3}, respectively. Vertical bars indicate the position of the MERIS bands used in the algorithm of Simis et al. (2005), centred at 620, 665, 708.75 and 778.75 nm. Symbols indicate the reflectance values at wavelengths used in the semi-empirical baseline algorithm (600, 624 and 648 nm) and the single reflectance ratio algorithm (625 and 650 nm). (B) Normalized absorbance spectra, obtained by dividing the absorbance at each wavelength by the absorbance at 675 nm. The spectra were measured at NIOO (the Netherlands) from phytoplankton cultures. Legend: cyano=*Limnothrix* sp.; chloro=*Scenedesmus obliquus*; diatom=*Asterionella formosa.* Vertical bars and symbols as in panel A.

Up to date, three algorithm approaches have been proposed for the quantification of PC based on the quantification of the reflectance trough at the 620 nm region in remotely sensed data: a single reflectance band ratio algorithm (Schalles & Yacobi, 2000), a semi-empirical baseline algorithm (Dekker, 1993), and a nested semi-empirical band ratio algorithm (Simis et al., 2005). The algorithms differ in the wavelengths used as reference and in the inclusion or not of a correction for Chl *a* absorption. The single reflectance ratio uses reflectance at 650 nm as reference, and then targets the PC absorption at 625 nm. The baseline approach uses two wavelengths (600 and 648 nm) to draw a reference baseline, and relates PC concentration to the distance from the midpoint of the baseline to the reflectance at 624 nm. The nested band ratio was developed for the spaceborn Medium Resolution Imaging Spectrometer (MERIS) band settings and relates PC to the absorption coefficient at the band centred at 620 nm, which is retrieved from the ratio of the 620 nm band and a near infra-red (NIR) band, centred at 709 nm, as reference. A second, nested ratio (709/665 nm bands) is used to infer a correction for absorption by Chl *a* in the 620 nm band. A fourth band (centred at 779 nm) is used to calculate the backscattering coefficient that is introduced in the two ratios, to retrieve the absorption coefficients through inversion of a reflectance model. Different optical instruments, PC extraction methods and sets of apparent and inherent optical properties were used to develop and parameterize the three algorithms, although all were developed for turbid, shallow, eutrophic temperate lakes, in the Netherlands (Dekker, 1993; Simis et al., 2005) and central USA (Schalles & Yacobi, 2000).

An important distinction among the proposed algorithms is whether they require multi- or hyper-spectral optical data. Algorithms based on the detailed shape of the reflectance spectra could perform better than simplified band ratio algorithms, but their practical use is limited as they cannot be used on a large scale with current spaceborn sensors. The nested band ratio algorithm has been applied to MERIS imagery of inland waters (Simis et al., 2006), whereas the lack of required spectral bands has prevented the application of the other algorithms with spaceborn sensors with global coverage. The distribution of cyanobacteria in inland and open sea waters could be derived with single reflectance ratio and baseline algorithms using the spaceborn sensors Chris/Proba (Ruiz-Verdú et al., 2005) and Hyperion (Kutser, 2004), respectively. While the availability and spatial coverage of these latter sensors is limited, they may be exemplary for future spaceborn sensors with global coverage.

In this study we evaluate the accuracy of the three published algorithms for the remote sensing of phycocyanin in inland waters, by using an extensive database of field radiometric and pigment data obtained in Spain and the Netherlands. In a previous paper we already reported on the influence of various phytoplankton pigments on the performance of the nested band ratio algorithm (Simis et al., 2007). Here, we make an effort to highlight differences between the three different approaches towards quantification of cyanobacterial PC.

Table 1
Summary of sampling campaigns and basic statistics for Chl *a* and PC pigment concentrations

		The Netherlands	Spain
Number of studied lakes and reservoirs		2	62
Lake statistics			
Surface area (km^2)	Min	35[a]	0.5
	Average		19.3
	Max	1190[b]	103
	Standard deviation		24.2
Mean depth (m)	Min	3[a]	0.9
	Average		28.8
	Max	4.4[b]	104
	Standard deviation		19.1
Time frame of field campaigns (years)		2004–2005	2001–2005
Number of visits[c]		6	89
Total samples		290	193
Spectroradiometry samples	PR-650	289	23
	ASD-FR	0	170
Pigment samples	HPLC	203	169
	PC_{FT}	201	16
	PC_{MG}	75	122
	PC_{FL}	0	168
Phytoplankton biomass samples (microscopy)		0	140
Pigment statistics			
Chl *a* (mg m^{-3})	5 Percentile	4.0	1.9
	Median	36.3	18.6
	95 Percentile	88.4	140.1
PC_{REF} (mg m^{-3})	5 Percentile	1.5	0.0
	Median	21.2	17.8
	95 Percentile	118.0	640.6

Abbreviations for PC quantification methods: PC_{FT} — freeze/thaw cycles; PC_{MG} — mechanical grinding; PC_{FL} — fluorescence.
[a] Data from Lake Ketelmeer.
[b] Data from Lake IJsselmeer.
[c] Visits spanning 2 days were counted as 1.

Besides an evaluation of algorithm sensitivity to other phytoplankton pigments, we evaluate their sensitivity to errors in reflectance measurements. Additionally, taxonomic data of the Spanish dataset allows us to evaluate the response of each algorithm in cyanobacteria-dominated water bodies.

2. Methods

2.1. Study sites

The sampling campaigns in Spain were carried out in the period 2001–2005 in 62 lakes and reservoirs, distributed throughout the country and covering a wide gradient of environmental conditions and trophic states. In the Netherlands, sampling took place at Lake Ketelmeer and at Lake IJsselmeer, the largest freshwater body in Western Europe. These were visited 6 times for 2-day cruises in 2004 and 2005. A summary of the study sites in both countries, the instrumentation used, and ranges of pigment concentrations are shown in Tables 1 and 7. Their geographical location is shown in Fig. 2. Further details can be found in Simis et al. (2007).

2.2. Radiometric measurements

Remote sensing reflectance $R_{rs}(\lambda)$ was calculated from above water measurements and defined as the MERIS level-2 standard product 'normalized water-leaving reflectance' (Montagner, 2001):

$$R_{rs}(\lambda) = [\rho_w]_N(0^+,\lambda) = \pi L_w(0^+,\lambda)/E_d(0^+,\lambda) \quad (1)$$

where $L_w(0^+,\lambda)$ is water-leaving radiance corrected for reflected diffuse skylight as detailed in Simis et al. (2007), $E_d(0^+,\lambda)$ is downward irradiance, depth 0^+ points to the situation just above the water surface and wavelength dependence is denoted by λ. Viewing and acceptance angles are within the ranges recommended in NASA (2003), as described below. Subsurface irradiance reflectance $R(0^-,\lambda)$ was also derived from the individual measurements of water-leaving radiance, sky radiance, and downwelling irradiance, following Gons (1999), which applies a correction for diffuse skylight reflection at the air–water interface and conversion from above-surface radiance reflectance to subsurface irradiance reflectance based on field calibrations and theoretical considerations.

Two field spectroradiometers were used: in Spain, an ASD-FR (Analytical Spectral Devices, Inc.), and in the Netherlands a PR-650 (Photo Research, Inc). For an overview of their technical specifications see Simis et al. (2007). In short, the ASD-FR measured with an 8° acceptance angle, with foreoptics a lens mounted on glass fibre,

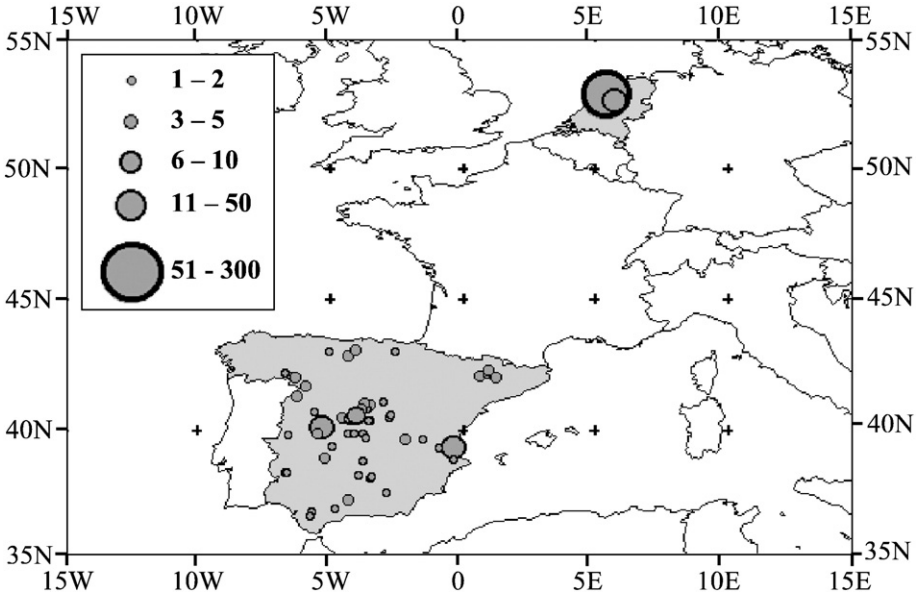


Fig. 2. Location of sampled lakes and reservoirs. Symbol size reflects the number of samples collected at each site.

spectral interval 1.4 nm, full-width half maximum (FWHM) of 3 nm, in the wavelength range 350–1000 nm. Viewing angles for water and sky radiance measurements were 40° from the horizontal plane, at 135° sun azimuth angle. The PR-650 measures at a 1° angle, 4-nm interval, 8-nm FWHM, in the range 380–780 nm. Viewing angles were 42° for water and sky radiance, and 90° sun azimuth angle. In both cases, E_d was obtained by measuring a calibrated reflectance panel. At 13 sampling stations in Spain, concurrent measurements were taken with the two spectroradiometers, yielding an acceptable agreement (R^2=0.94) in the $R_{rs}(\lambda)$ obtained from the two instruments.

2.3. Phytoplankton biomass

Phytoplankton samples were fixed with Lugol's solution and later, in the laboratory, sedimented and counted under an inverted microscope using Utermöhl's method. Various fields were counted until a significant volumetric cell count was achieved (Sournia, 1978). The volume of each species was calculated by using measurements of cell dimensions and suitable geometric formulas (Rott, 1981). Assuming cell density equal to water, calculated volumes were transformed to biomass in mg m^{-3}. The biomass was obtained at least at genus level and the absolute and relative biomasses were computed for each of the following taxonomic Classes: Cyanobacteria, Chlorophyceae, Cryptophyceae, Chrysophyceae, Bacillariophyceae, Dinophyceae, Euglenophyceae, Raphidophyceae and Xanthophyceae.

2.4. Pigment analysis

Chlorophylls and carotenoids were extracted with acetone and quantified using gradient HPLC based on the protocols in Jeffrey et al. (1997). The concentrations of the following pigments were acquired for all samples collected from both countries: Chl *a*, Chl *b*, phaeophytin, peridinin, neoxanthin, violaxanthin, alloxanthin, lutein, zeaxanthin, and fucoxanthin. Chl *c* was not measured in Spanish samples; however, the Dutch samples showed a strong correlation between fucoxanthin and Chl *c* (R^2=0.91), suggesting that fucoxanthin could be used as a proxy for Chl *c*.

There is no standard method of extraction for the water-soluble phycobilin pigments. Two different methods were used here to extract PC. In addition, a fluoroprobe was used *in situ* (discussed below). The first extraction method uses repeated freezing and thawing cycles of samples concentrated by centrifugation. PC concentrations from this method will be referred to as PC_{FT}. The used procedure (Simis et al., 2005, 2007) was based on Sarada et al. (1999). The second method was based on mechanical grinding (PC_{MG}) of samples concentrated on glass fibre filters, and treatment with glycerol and distilled water to induce osmotic stress (Quesada and Vincent, 1993; Wyman and Fay, 1986a,b). After extraction with either the PC_{FT} or PC_{MG} method, the PC concentrations were calculated spectrophotometrically according to Bennett and Bogorad (1973). A third method for PC quantifying was performed only for the Spanish dataset, measuring *in situ* fluorescence (PC_{FL}) using a precalibrated Mini^{tracka} II PA Fluorometer Model HB202 (Chelsea Instruments Ltd., Surrey, UK).

Concurrent measurements for PC_{FT} and PC_{MG} were available for the Netherlands (n=69) and Spain (n=16), and for PC_{MG} and PC_{FL} only in Spain (n=106). It was previously reported that none of the methods consistently yields a higher PC content and all are likely to contain errors (Simis et al., 2007). Therefore, when multiple measurements were available, the average value (PC_{AVG}) was used as the reference value for PC concentration (PC_{REF}). However, where PC_{FL} measurements were available (i.e., the Spanish dataset), a considerable number of samples showed important differences between PC concentrations obtained from the *in situ* fluorescence and extraction methods. In these cases, very low extraction yields could have caused an underestimation of PC_{FT} or PC_{MG}, whereas CDOM fluorescence could have caused an overestimation of PC_{FL}, or absorption by other pigments could have depressed PC_{FL} at high algal biomass. In those cases, the use of the maximum or minimum value instead of the average could yield better estimates of PC_{REF}, but additional criteria are needed to choose between methods. In the Spanish dataset, information on the cyanobacterial biomass was available from microscopy, and this was used to improve the calculation of PC_{REF} as follows. For each sample, Chl *a* contained in the cyanobacterial fraction ($Chla_{CYANO}$) was calculated by multiplying the Chl *a* concentration with the cyanobacterial share of total biovolume. Values of PC_{MG}, PC_{FT}, or PC_{FL} that did not conform to the linear fit (using all samples) of PC_{AVG} as a function of $Chla_{CYANO}$ were then removed before averaging the remaining PC estimates to obtain PC_{REF}.

2.5. PC retrieval from radiometry

Three different algorithms were used to retrieve PC concentrations from reflectance data (denoted PC_{RAD}); they are outlined below.

2.6. Single reflectance ratio algorithm

This algorithm (Schalles & Yacobi, 2000) is based on an empirical relationship between the ratio of reflectances at λ=650 and λ=625 nm (Fig. 1) and PC concentration (mg m^{-3}), obtained from a four-year study in Carter Lake, a hypereutrophic floodplain lake in Nebraska, USA. The reflectance quantities used in the algorithm are the ratio of the radiance just beneath the water surface, $L_u(0^-,\lambda)$, and the radiance of a Lambertian Spectralon panel which equals $E_d(0^+,\lambda)/\pi$. Both quantities were measured with an acceptance angle of 23°. This combination of measurements is not available for our whole dataset, although $L_u(0^-,\lambda)$ can be calculated from $L_w(0^+,\lambda)$ using a conversion factor proposed by Austin (1980). However, since the ratio of reflectances at the two wavelengths is not influenced by such a transformation, we used $R_{rs}(\lambda)$ values directly. The algorithm was thus applied as:

$$PC_{RAD} = \{[R_{rs}(650)/R_{rs}(625)]-0.97\} \times 1096.5 \quad (2)$$

The spectral resolutions of data from ASD-FR (1 nm) and PR-650 (4 nm) were averaged and interpolated, respectively, to 2 nm, the spectral resolution at which the algorithm was originally developed.

2.7. Semi-empirical baseline algorithm

The basis of the baseline algorithm (Dekker, 1993) is to obtain a value of the decrease in $R(0^-,624)$ relative to the surrounding spectrum, observed with increasing PC concentration. The algorithm draws a baseline between two bands located at 600 and 648 nm, i.e. around the absorption feature in reflectance spectra around 624 nm, in this case attributed to PC (Fig. 1). Dekker (1993) based the algorithm coefficients on averaged inherent optical properties from 10 shallow eutrophic lakes in the Netherlands. The proposed algorithm (for 2-nm band width) was:

$$PC_{RAD} = -24.6 + 13686 \times \{0.5 \times [R(0^-,600) + R(0^-,648)] - R(0-,624)\} \quad (3)$$

To apply this algorithm, our dataset was converted to $R(0^-,\lambda)$ at 2 nm intervals.

2.8. Nested semi-empirical band ratio algorithm

This algorithm (Simis et al., 2005, 2007) was developed for the band configuration of the Medium Resolution Imaging Spectrometer (MERIS), using MERIS channels 6, 7, 9, and 12 centred at 620, 665, 709, and 779 nm, respectively. Absorption in the 620 and 665 nm bands is assumed to be dominated by water and phytoplankton pigments (PC and Chl *a* at 620 nm, and Chl *a* alone at 665), while the absorption in

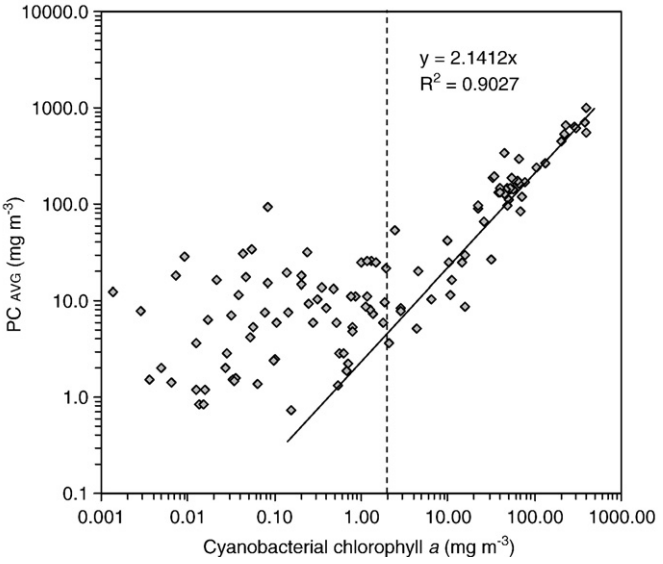


Fig. 3. PC concentration (sample average of measurement techniques) plotted as a function of Chl *a* biomass present in cyanobacteria. The regression intercept was forced to 0 ($p<0.001$ at 99% confidence level). Only data from the Spanish dataset are shown ($n=131$).

the 709 and 779 nm bands is assumed to be dominated by water alone. A spectrally neutral backscattering coefficient is derived from the 779 nm band through inversion of a commonly used relationship between inherent optical properties and reflectance (Gordon et al., 1975), as described in detail in Gons et al. (2005):

$$b_b(779) = 1.61R_{rs}(779)/\{0.082-0.6 \times R_{rs}(779)\} \quad (4)$$

Absorption coefficients in the 620 and 665 nm bands are then retrieved from the same reflectance model, through a ratio of the 709-nm band over each of the red bands. The absorption by PC at 620 is finally obtained after correcting for absorption by Chl *a* in the same band, proportional to the Chl *a* absorption coefficient derived at 665 nm. The

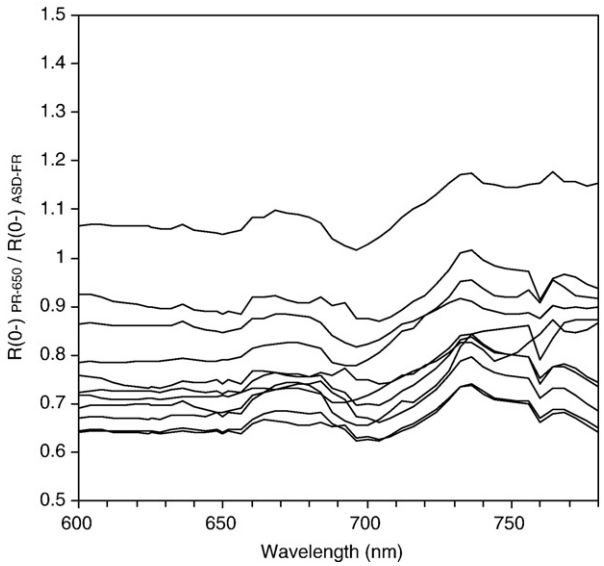


Fig. 4. Ratio of reflectances measured concurrently with the field spectroradiometers PR-650 and ASD-FR, at 13 sampling stations in Spain. The graph covers the spectral region used by the three tested algorithms.

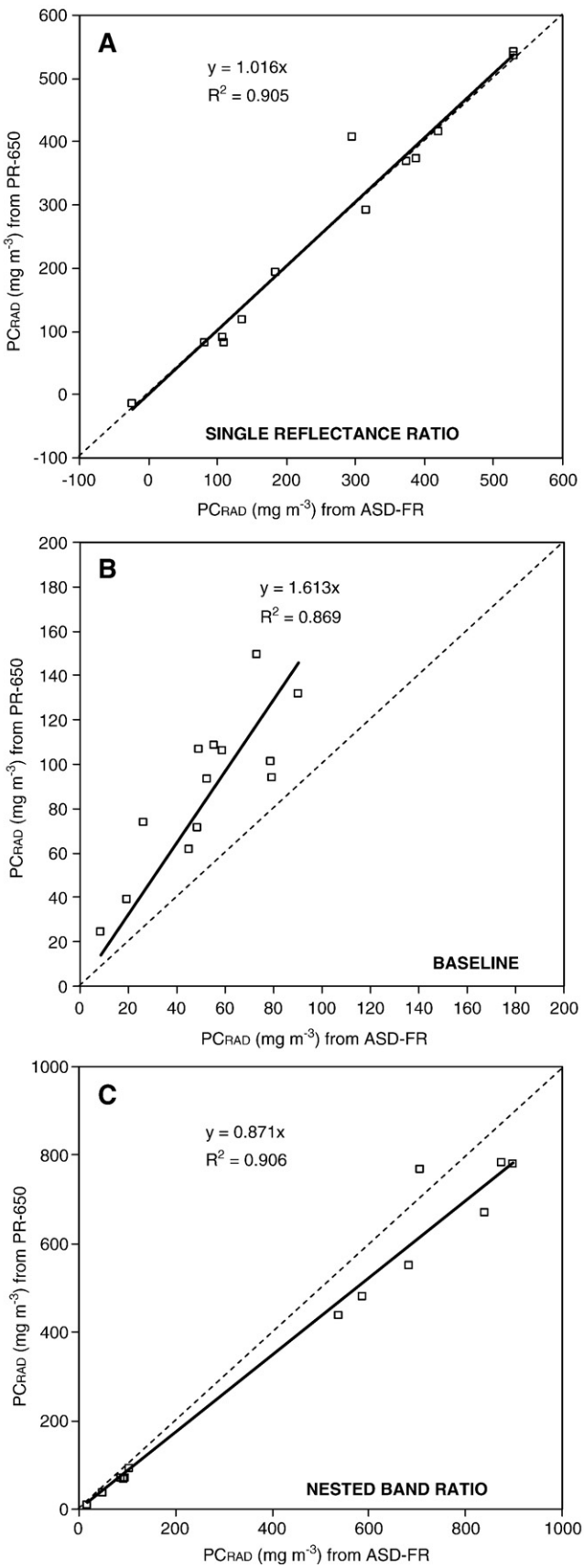


Fig. 5. Comparison of retrieved PC concentrations (PC_{RAD}) in Spanish water bodies in which concurrent measurements were taken with the PR-650 and ASD-FR field spectroradiometers. PC_{RAD} were calculated with (A) the single reflectance ratio algorithm, (B) the baseline algorithm, and (C) the nested band ratio algorithm. For all regressions $n=13$ and $p<0.001$ at 99% confidence level.

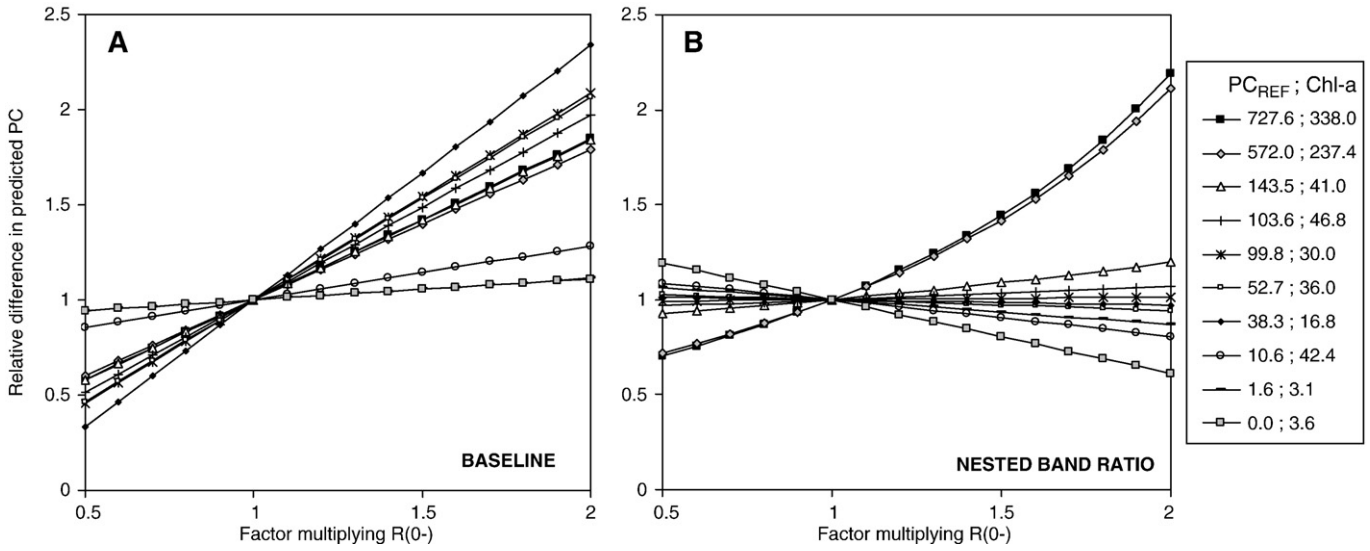


Fig. 6. Effect of changes in $R(0^-)$ in the retrieval of PC for (A) the baseline algorithm and (B) the nested band ratio algorithm. The algorithms were applied to reflectance values at several Spanish water bodies after multiplying by a factor ranging from 0.5 to 2. Values in the legend represent field PC and Chl *a* concentrations (mg m^{-3}).

detailed procedure is given in Simis et al. (2005, 2007). The equation applied here was (Simis et al., 2007, Eqs. (5a)–(5b)):

$$a_{ph}(665) = 147 \times \left[\frac{R(709) \times (0.727 + b_b)}{R(665)} - b_b - 0.401\right] \quad (5a)$$

$$PC_{RAD} = 170 \times \left\{\left[\frac{R(709) \times (0.727 + b_b)}{R(620)} - b_b - 0.281\right] - [0.24 \times a_{ph}(665)]\right\} \quad (5b)$$

3. Results

3.1. Correction of reference PC values

Poor extraction yields for PC_{MG} or PC_{FT}, or interference of background fluorescence with PC_{FL} may give rise to erroneous PC_{REF} values. In order to identify anomalous PC measurements, the relationship between PC_{AVG} (the average of available reference PC measurements) and $Chla_{CYANO}$ (Chl *a* contained in the cyanobacterial fraction of the sample) was analysed for the Spanish dataset, where such taxonomical data was available. Fig. 3 shows that the PC_{AVG}: $Chla_{CYANO}$ ratio was less variable than may be expected when taking into account the role of PC as an accessory pigment. A significant linear regression ($p<0.001$) was obtained for the whole concentration range, indicating a stable relationship between the two pigments in the studied cyanobacterial populations. For very low cyanobacterial biomass (below 2 mg m^{-3} $Chla_{CYANO}$, marked with a dashed line in Fig. 3), a cluster of points that clearly deviate from the regression line was observed. In these samples PC_{AVG} was two orders of magnitude higher than $Chla_{CYANO}$, whereas the ratio was 2.14 for the dataset as a whole ($R^2=0.90$). The high PC: Chl *a* ratios found for low $Chla_{CYANO}$ concentrations cannot occur in cyanobacterial cells and must be considered overestimations in one of the available reference PC measurements. Indeed, in most cases, the relatively high PC_{AVG} values were correlated with high PC_{FL} measurements, whereas the corresponding PC_{MG} values were very low or zero. Thus, to calculate PC_{REF}, PC_{FL} for samples with $Chla_{CYANO}<2$ mg m^{-3} were omitted; for $Chla_{CYANO}\geq 2$ mg m^{-3}, we selected the maximum value of PC_{FL} and PC_{MG} instead of the average. Microscope counts and PC_{FL} measurements were not available for the Dutch dataset, so PC_{AVG} was adopted for PC_{REF}.

3.2. Effects of differences in reflectance measurement protocols

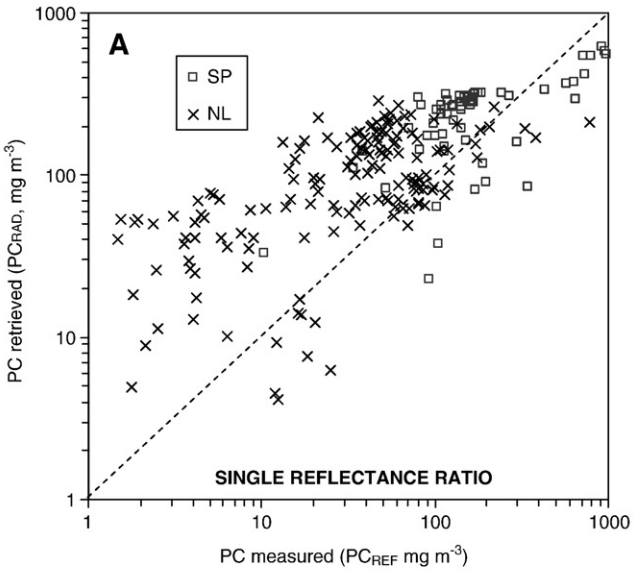
Two different spectroradiometers and measurement setups were used to collect the present dataset, and different methods were used for the calculation of remote sensing reflectances and subsurface irradiance reflectance, as described in Methods. For a limited number ($n=13$) of sampling stations in Spain, concurrent measurements with the two approaches were obtained in May 2005. In this subset, the ratio of reflectances for the two instruments was nearly flat for the spectral bands used in the three algorithms (Fig. 4). This indicates that the differences in instrumentation and measurement angles affected the amplitude but not the spectral shape of the reflectance spectra. The similarity in spectral shape suggests that the single reflectance ratio algorithm should yield similar results with both measurement approaches, whereas the baseline algorithm is likely to be sensitive to differences in the amplitude of concurrent spectra. For the nested band algorithm the interpretation is more complicated, as there are four bands involved and the effects will be contained in any of the terms that are subtracted from the band ratios in Eqs. (5a)–(5b).

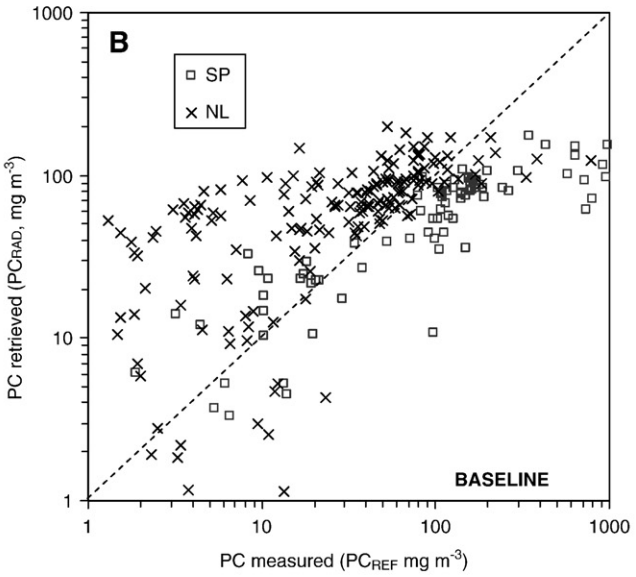
PC_{RAD} values obtained from the three algorithms for the concurrent reflectance measurements are shown in Fig. 5. In all cases the

Table 2
Linear regression of reflectances measured concurrently with the PR-650 spectroradiometer and the ASD-FR spectroradiometer

Band	β_0	β	R^2
$R_{rs}(625)$	0.0001	0.7374	0.883
$R_{rs}(650)$	0.0001	0.7368	0.859
$R(0^-,600)$	0.0006	0.7406	0.858
$R(0^-,624)$	0.0001	0.7389	0.883
$R(0^-,648)$	0.0006	0.7334	0.858
$R_{rs}(M6)$	0.0014	0.7812	0.876
$R_{rs}(M7)$	0.004	0.6243	0.841
$R_{rs}(M9)$	0.0028	0.7909	0.942
$R_{rs}(M12)$	0.0012	0.7902	0.958

The resulting equations take the form: $R\_(PR\text{-}650)=\beta_0+\beta\times[R\_(ASD\text{-}FR)]$. In all cases, $n=13$ and $p<0.001$ at 99% confidence level. Symbols and abbreviations: $R_{rs}(\lambda)$ — remote sensing reflectance for the wavelengths used in the single reflectance ratio algorithm (Schalles & Yacobi, 2000); $R(0^-,\lambda)$ — subsurface irradiance reflectance for wavelengths used in the semi-empirical baseline algorithm (Dekker, 1993); $R_{rs}(M\_)$ — remote sensing reflectance for MERIS bands used in the nested semi-empirical band ratio algorithm (Simis et al., 2005).





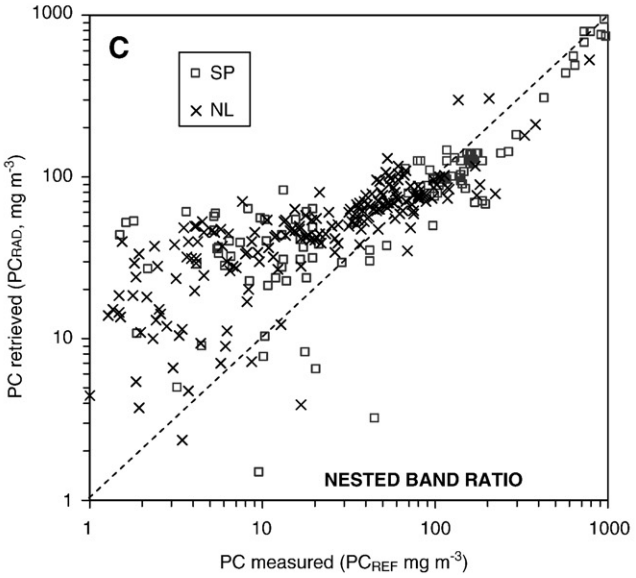
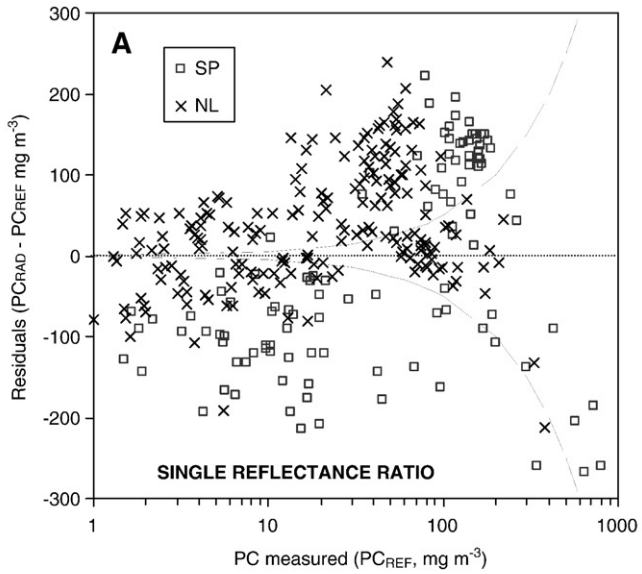
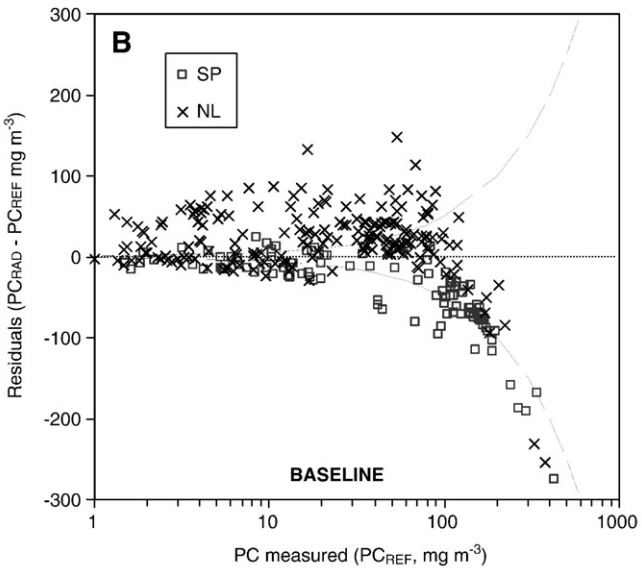


Fig. 7. Scatter plots of retrieved PC concentrations against measured PC, separated by country, for (A) the single reflectance ratio algorithm, (B) the baseline algorithm, and (C) the nested band ratio algorithm. Only values higher than 1 mg m^{-3} were plotted. Negative predictions represented 36.4%, 20.9% and 0.3% of data points for panels A–C, respectively.





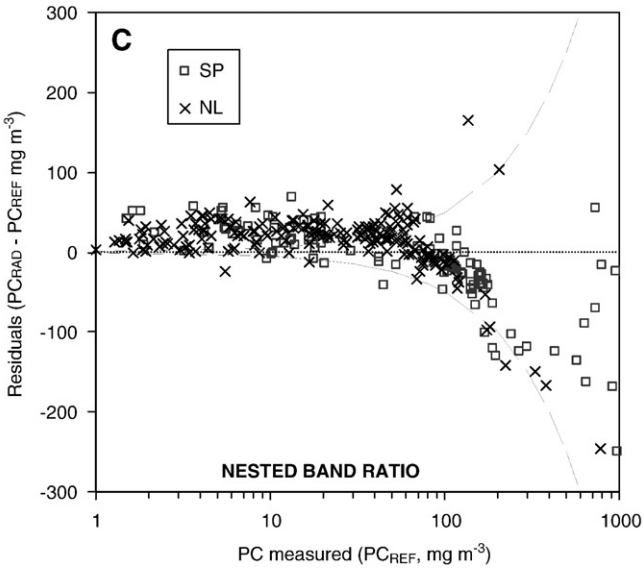


Fig. 8. Scatter plot of the residuals (calculated as $PC_{RAD} - PC_{REF}$) against PC_{REF} for the 1–1000 mg m^{-3} range, for (A) the single reflectance ratio algorithm (B) the baseline algorithm, and (C) the nested band ratio algorithm. The area between the dashed lines indicates relative errors in PC retrieval below 50%.

correlation between the reflectances obtained by the two measurement approaches was high, but slopes varied. For the simple band ratio (Fig. 5A) the differences were minimal and the slope was close to unity, as expected. The most affected algorithm was the baseline algorithm (Fig. 5B), which depends most on the amplitude of the reflectance spectra, where PC_{RAD} values obtained with the PR-650 were approximately 60% higher than those obtained with the ASD-FR. For the nested band algorithm (Fig. 5C), which combines the reflectances of four bands in a non-linear manner, PC_{RAD} values were higher for the ASD-FR measurements.

A sensitivity analysis on the effect of changes in the amplitude of the reflectance spectrum on PC retrieval was carried out. By definition, the single reflectance ratio algorithm is not affected by this treatment. The response of the baseline and nested band ratio algorithms are plotted for 10 sampling stations along a wide gradient of PC_{REF} and Chl *a* concentrations, chosen from the Spanish dataset (Fig. 6). To obtain these results, ASD-FR values were multiplied by spectrally constant factors ranging from 0.5 to 2, and subsequent algorithm responses were compared to the original PC_{RAD} (equal to a multiplication factor of 1). The relative errors derived from the baseline algorithm (Fig. 6A) increased linearly with the factor multiplying $R(0^-)$. The slope was lower at low PC concentrations and increased with PC concentration, although a clear trend was not evident. In the nested band ratio algorithm (Fig. 6B) the differences were not linear. For intermediate values of PC and Chl *a* the errors were small, with differences around 10–15% in PC_{RAD} observed over the range of multiplication factors. The differences tended to be higher for very low PC and Chl *a* values, and especially for very high values, where the predicted values can be doubled if the reflectances are doubled. It is noted that such extreme amplitudinal differences between concurrent reflectance measurements were not observed during this study.

3.3. Intercalibration of radiometric measurements

In order to minimize the errors due to instrumental differences in subsequent analyses, the reflectances measured with the ASD-FR spectroradiometer were calibrated, taking the small set of concurrent PR-650 spectroradiometer reflectances ($n=13$) as reference. The relation between the different measurement approaches was obtained from a linear regression fit, the resulting equations applied to the whole Spanish dataset. The regression coefficients (Table 2) were very similar for all the spectral bands used in the algorithms, as expected from the spectrally flat reflectance ratios between the two radiometers (Fig. 4). PR-650 reflectances were chosen as reference as their measurement protocol was most similar to those used in the development of the baseline and nested band ratio algorithms. Although the coefficient of determination was high in all cases, it should be taken into account that the number of samples used for this calibration is relatively low and many of them are from the hypereutrophic lake L'Albufera. However, the calibration is justified given the effect of instrumental differences observed in the performance of two of the considered algorithms.

3.4. Performance of the PC retrieval algorithms

The results of each reflectance algorithm (PC_{RAD}) were compared to reference PC measurements (PC_{REF}) and plotted in Fig. 7. Residual errors ($PC_{RAD}-PC_{REF}$) were plotted against PC_{REF} for the 1–1000 mg m^{-3} range (Fig. 8). Fig. 7 shows that all algorithms exhibited a substantial amount of scatter, increasing towards lower PC concentrations. The regressions of PC_{RAD} against PC_{REF} for the three algorithms are summarized in Table 3. The relationship between PC_{RAD} and PC_{REF} could not be fitted well with a linear regression model for the single reflectance ratio (Schalles & Yacobi, 2000) and baseline (Dekker, 1993) algorithms (Fig. 7A and B, Table 3), therefore the coefficients of determination were low and regression coefficients were far from 1. The nested band ratio algorithm (Simis et al., 2005) gave the best fit ($R^2=0.92$) and the lowest standard error of the three algorithms, and linear regression coefficient β close to unity (Fig. 7C, Table 3). The coefficients of determination of all algorithms were highest for the Spanish subset, where regression coefficients were also closest to 1.

In the three algorithms, the highest relative errors (under- and overestimations) occurred at low PC_{REF} concentrations ($PC_{REF}<50$ mg m^{-3}). For $PC>200$ mg m^{-3} relative errors were lower and PC tended to be underestimated. The best predictions were found in the intermediate concentration range (50–200 mg m^{-3}) in all algorithms (Fig. 8). For the three PC_{REF} concentration ranges, the highest errors occurred in the single reflectance ratio algorithm and the lowest in the nested band ratio algorithm.

For $PC_{REF}<50$ mg m^{-3}, 54% of the PC_{RAD} values obtained for the whole dataset were negative for the single reflectance ratio algorithm; 31% for the baseline algorithm and 0.4% for the nested band ratio algorithm. When considering only the Spanish dataset, the percentages were higher, except for the nested band ratio algorithm (92%, 55% and 0% respectively).

The scatter plot of residuals against PC_{REF} concentrations (Fig. 8) showed significant differences between the Dutch and Spanish datasets for the single reflectance ratio and baseline algorithms, but not for the nested band ratio algorithm. In the first two algorithms, retrieved PC concentrations tended to be overestimated in the Dutch samples, and clearly underestimated in the Spanish samples. Most residuals for the baseline algorithm were negative in the Spanish samples and linearly correlated with PC_{REF}, as $[PC_{RAD}-PC_{REF}]=-0.82\times PC_{REF}$ ($R^2=0.96$; $p<0.001$; $n=152$).

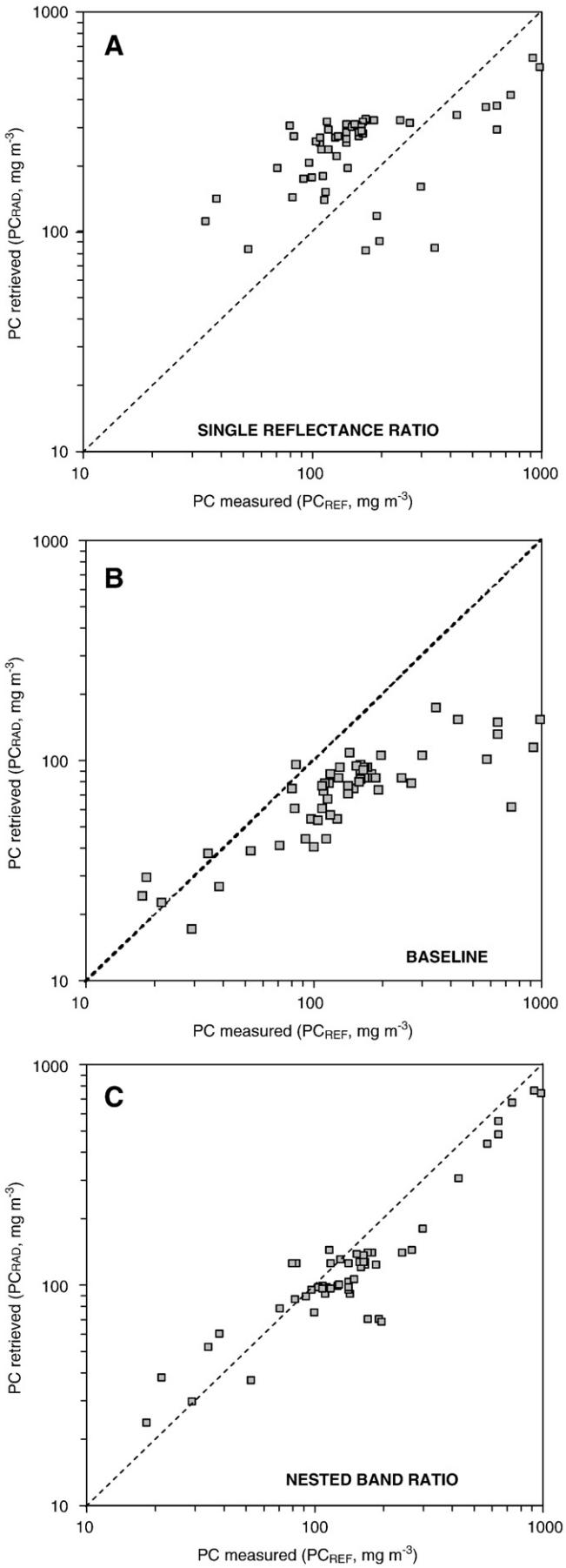
3.5. Algorithm performance in cyanobacteria-dominated waters

The performance of the algorithms in cyanobacteria-dominated waters was examined for a subset of the Spanish dataset for which the cyanobacterial biomass was higher than 80% of the total biomass, as determined with microscopy (Fig. 9). For these samples, PC estimates were positive with few exceptions, as expected for waters where the PC absorption feature is clearly present in the reflectance spectra. The trends of residual errors resulting from the single reflectance ratio and

Table 3
Regression of PC_{RAD} (PC retrieved from algorithms) against PC_{REF} (measured PC) for the three algorithms

	Single reflectance ratio			Baseline			Nested band ratio		
	NL+SP	NL	SP	NL+SP	NL	SP	NL+SP	NL	SP
n	352	202	150	352	202	150	352	202	150
R^2	0.461	0.21	0.594	0.211	0.249	0.484	0.923	0.753	0.948
Adjusted R^2	0.459	0.206	0.591	0.208	0.246	0.480	0.922	0.752	0.948
RMSE	119.66	89.70	150.20	135.47	66.76	191.31	47.48	40.48	55.94
Model-p	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
β PC_{REF}	**0.679**	**0.458**	**0.764**	**0.346**	**0.499**	**0.696**	**0.961**	**0.868**	**0.974**

Significant (>99% confidence) p and β values are printed in bold face. Symbols and abbreviations: NL — Samples from The Netherlands; S — Samples from Spain; n — number of valid cases; RMSE — Square root of the mean squared error.



baseline algorithms became more pronounced when using this high cyanobacterial biomass subset, compared to the full dataset. The single reflectance ratio algorithm (Fig. 9A) yielded overestimations for moderate concentrations and increasing underestimation with higher PC_{REF}. The baseline algorithm (Fig. 9B) showed constant relative errors (underestimation increasing with PC_{REF}) suggesting that sensitivity of this algorithm is limited to a PC_{REF} range up to 200 mg m^{-3}. The only significant regression between PC_{RAD} and PC_{REF} for the cyanobacterially dominated waters was obtained with the nested band ratio algorithm ($PC_{RAD}=4.17+0.76\times PC_{REF}$; $R^2=0.97$; $p<0.001$; n=n =59). This algorithm also showed a tendency to overestimate PC at low PC_{REF} and underestimate PC at high PC_{REF}.

3.6. Algorithm performance in relation to pigment presence

The influence of other phytoplankton groups on the residual errors of the three algorithms was analysed with multiple linear regressions in which the residual error (calculated as $PC_{RAD}-PC_{REF}$) was the dependent variable and the concentration of several diagnostic pigments were the independent variables. The amount by which the residual errors increase when the concentration of a diagnostic pigment is increased is given by its partial regression coefficient (β). Chlorophyll *a* was included in the analysis to represent the effects of total phytoplankton biomass, as well as Phaeophytin *a*, its main degradation product and with similar optical properties. Chlorophyll *b* (Chl *b*) was used as a marker for green algae (Chlorophyceae); Fucoxanthin (a proxy for chlorophyll *c*) for diatoms; Peridinin for dinoflagellates and Alloxanthin for Cryptophyceae. The regression was carried out for three concentration ranges ($PC_{REF}<50$, 50–200, and >200 mg m^{-3}) and for the whole dataset as well as subsets divided by country. The confidence interval was fixed at 99% unless otherwise indicated. A summary of the results is presented in Tables 4–6 for the single reflectance ratio algorithm, the baseline algorithm and nested band ratio algorithm, respectively. A previous analysis was made with this dataset on the nested band ratio algorithm (Simis et al., 2007), but then no correction for instrumental differences was applied.

When the whole dataset is examined, all algorithms yielded significant regressions at $PC_{REF}<50$ mg m^{-3}, indicating that part of the observed variability in the residual errors can be explained by the presence of other pigments. The significance of the regressions was lower (or no significant at 99% confidence level) for the intermediate concentration range (50–200 mg m^{-3}). For $PC_{REF}>200$ mg m^{-3}, the regressions were again significant, except for the nested band ratio algorithm. The effect of pigments on the prediction error was stronger in the Dutch dataset, while in the Spanish dataset most regressions were not significant.

Chl *a* (and to a lesser extent, Phaeophytin) had the strongest overestimating effect in all cases except for $PC_{REF}>200$ mg m^{-3} where Chl *a* was negatively correlated with prediction errors. Chl *b* was negatively correlated with residuals in the single reflectance ratio algorithm, but positively in the Dutch dataset for baseline and nested band ratio algorithms, whereas for the Spanish dataset the correlation was not significant. The other accessory pigments had weaker or non-significant correlations and did not show clear trends in their effect in prediction errors. The exception was the single reflectance ratio algorithm, where all pigments, except Chl *a* and Phaeophytin, had positive and significant regression coefficients.

4. Discussion

The performance of three algorithms for the retrieval of the cyanobacterial pigment phycocyanin was tested with pigment and

Fig. 9. Scatter plot of retrieved PC concentrations against measured PC, at Spanish stations where Cyanobacteria represented more than 80% of total phytoplankton biomass, for (A) the single reflectance ratio algorithm, (B) the baseline algorithm, and (C) the nested band ratio algorithm.

Table 4
Regression of residuals against Chl *a* and marker pigments, for the single reflectance ratio algorithm (Schalles & Yacobi, 2000)

	$PC_{REF}<50$ mg m^{-3}			$50 \le PC_{REF} \le 200$ mg m^{-3}			$PC_{REF}>200$ mg m^{-3}
	NL	SP	NL+SP	NL	SP	NL+SP	NL+SP
n	141	79	220	55	27	82	12
Multiple R^2	0.646	0.591	0.669	0.530	0.820	0.528	0.927
Adjusted R^2	0.633	0.557	0.660	0.482	0.777	0.497	0.866
SE of estimate	42.15	39.97	54.50	50.15	36.86	53.34	76.60
Whole-model p	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.002**
β Chlorophyll *a*	**1.138**	**3.157**	**1.540**	−0.215	**0.451**	0.271	**−2.412**
β Phaeophytin	**0.315**	**0.597**	**0.220**	0.234	−0.117	0.120	**2.371**
β Chlorophyll *b*	**−0.447**	**−1.952**	**−0.695**	0.205	−0.642	−0.232	−1.438
β Fucoxanthin	**−0.429**	**−0.747**	**−0.468**	−0.312	−0.183	**−0.588**	0.612
β Peridinin	–	**−3.053**	**−0.941**	–	–	–	–
β Alloxanthin	**−0.583**	**−1.068**	**−0.480**	−0.286	−0.256	**−0.370**	0.301

Significant (>99% confidence) *p* and β values are printed in bold face. Symbols and abbreviations: NL — Samples from The Netherlands; S — Samples from Spain; n — number of valid cases; SE — standard error.

Table 5
Regression of residuals against Chl *a* and marker pigments, for the semi-empirical baseline algorithm (Dekker, 1993)

	$PC_{REF}<50$ mg m^{-3}			$50 \le PC_{REF} \le 200$ mg m^{-3}			$PC_{REF}>200$ mg m^{-3}
	NL	SP	NL+SP	NL	SP	NL+SP	NL+SP
n	141	79	220	55	27	82	12
Multiple R^2	0.615	0.257	0.454	0.313	0.304	0.426	0.990
Adjusted R^2	0.601	0.196	0.439	0.243	0.138	0.388	0.982
SE of estimate	18.37	10.80	21.21	35.36	23.77	40.73	41.86
Whole-model p	**0.000**	**0.001**	**0.000**	**0.002**	0.150	**0.000**	**0.000**
β Chlorophyll *a*	**0.337**	1.007	**0.948**	−0.394	−0.396	**−0.384**	**−1.378**
β Phaeophytin	**0.299**	−0.217	**0.259**	−0.259	−1.795	−0.148	**0.830**
β Chlorophyll *b*	**0.310**	−0.626	**−0.245**	**0.489**	1.542	**0.817**	0.010
β Fucoxanthin	−0.012	−0.143	**−0.314**	0.215	1.183	0.245	−0.231
β Peridinin	–	−0.676	**−0.531**	–	–	–	–
β Alloxanthin	**−0.607**	−0.634	**−0.474**	**−0.443**	−0.548	**−0.423**	−0.292

Significant (>99% confidence) *p* and β values are printed in bold face. Symbols and abbreviations: NL — Samples from The Netherlands; S — Samples from Spain; n — number of valid cases; SE — standard error.

field spectroradiometric data collected in the period 2001–2005 in the Netherlands and Spain, and the effects of several potential sources of errors were analyzed. These error sources include instrumental error and measurement methodology, and the presence of pigments with overlapping light absorption features. The different distribution of errors observed in the Dutch and the Spanish samples could be attributed to trophic conditions and phytoplankton assemblages, as indicated by pigment statistics (Table 7). The samples from the Netherlands came from two shallow and eutrophic big lakes, while the Spanish samples were collected in 62 lakes and reservoirs along a wide environmental gradient (Fig. 2, Table 1). For low PC concentrations ($PC_{REF}<50$ mg m^{-3}), the Dutch samples had much higher Chl *a*, Chl *b* and Fucoxanthin concentrations than the Spanish samples, suggesting eutrophic conditions and a mixture of phytoplankton groups at high density. In contrast, the Spanish samples within this concentration range came mostly from oligotrophic reservoirs, with few eutrophic exceptions with low cyanobacterial presence. For the range 50–200 mg m^{-3} of PC_{REF}, most samples came from cyanobacteria-dominated waters, but the pigment composition indicated a significant presence of other phytoplankton groups in the Dutch samples while the Spanish samples came from waters with almost exclusively cyanobacteria. The dominance of cyanobacteria is evident in the two subsets for $PC_{REF}>200$ mg m^{-3}. It should be taken into account that most of the data for very high PC_{REF} concentration came from a single hypereutrophic lake in Spain (L'Albufera).

Table 6
Regression of algorithm residuals against Chl *a* and marker pigments, for the nested semi-empirical band ratio algorithm (Simis et al., 2005)

	$PC_{REF}<50$ mg m^{-3}			$50 \le PC_{REF} \le 200$ mg m^{-3}			$PC_{REF}>200$ mg m^{-3}
	NL	SP	NL+SP	NL	SP	NL+SP	NL+SP
n	141	79	220	55	27	82	12
Multiple R^2	0.597	0.137	0.153	0.079	0.300	0.097	0.576
Adjusted R^2	0.582	0.066	0.129	−0.016	0.133	0.037	0.222
SE of estimate	9.09	20.29	15.86	39.42	23.13	35.45	83.02
Whole-model p	**0.000**	0.090	**0.000**	0.531	0.157	0.162	0.284
β Chlorophyll *a*	**0.277**	−1.105	0.090	0.166	−0.325	−0.141	−1.989
β Phaeophytin	0.132	0.132	0.047	−0.043	−1.227	−0.003	3.064
β Chlorophyll *b*	**0.529**	0.548	0.257	−0.022	2.087	0.432	−2.005
β Fucoxanthin	0.069	0.341	0.044	−0.299	0.337	−0.174	0.639
β Peridinin	–	1.044	0.171	–	–	–	–
β Alloxanthin	**−0.188**	0.293	0.041	−0.146	−0.863	−0.151	0.315

Significant (>99% confidence) *p* and β values are printed in bold face. Symbols and abbreviations: NL — Samples from The Netherlands; S — Samples from Spain; *n* — number of valid cases; SE — standard error.

Table 7
Descriptive statistics of major photosynthetic pigments in the dataset, separated by country and by concentration ranges for PC_{REF}

PC_{REF} range		PC_{REF}		Chlorophyll *a*		Phaeophytin		Chlorophyll *b*		Fucoxanthin		Alloxanthin	
		NL	SP	NL	SP	NL	SP	NL	SP	NL	SP	NL	SP
<50	*n*	138	92	137	92	137	92	137	92	137	92	137	92
	Average	16.3	7.7	29.0	14.6	1.7	0.3	2.5	1.2	2.1	1.1	0.8	1.2
	Median	10.1	3.4	31.0	7.2	0.4	0.2	2.4	0.3	1.2	0.4	0.5	0.3
	SD	15.1	10.6	15.8	22.2	2.8	0.4	1.5	2.8	3.3	2.4	0.8	2.7
	Minimum	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Maximum	49.6	45.1	74.5	171.9	13.3	2.2	6.7	19.1	21.8	14.2	4.2	15.5
50–200	*n*	55	50	55	37	55	37	55	37	55	37	55	37
	Average	80.9	131.6	65.2	53.1	2.1	1.1	3.2	1.4	3.7	1.9	0.9	1.1
	Median	74.1	128.9	62.6	52.0	1.6	0.8	3.5	0.4	2.2	0.6	0.8	0.5
	SD	30.8	36.9	22.4	22.3	1.5	1.2	1.4	2.7	3.1	5.6	0.6	1.9
	Minimum	51.6	52.7	32.7	15.5	0.1	0.3	0.6	0.0	0.4	0.0	0.0	0.0
	Maximum	183.3	198.0	136.1	144.8	6.2	5.4	6.1	11.0	13.4	34.2	2.9	11.0
>200	*n*	5	16	5	16	5	16	5	16	5	16	5	16
	Average	383.3	617.2	154.8	294.1	3.1	6.1	1.7	4.9	2.7	5.1	0.3	1.4
	Median	329.4	642.0	128.1	296.0	3.2	5.3	1.9	3.9	2.9	2.6	0.3	1.0
	SD	233.0	282.0	104.7	173.3	0.7	4.1	0.5	3.4	0.9	5.3	0.4	1.6
	Minimum	206.4	242.1	72.6	48.2	2.1	1.6	1.1	0.6	1.6	0.0	0.0	0.0
	Maximum	779.1	1040	334.2	674.7	3.8	18.6	2.1	12.0	3.9	18.8	0.9	5.8

All concentrations are in mg m^{-3}. Symbols and abbreviations: NL — Samples from The Netherlands; S — Samples from Spain; *n* — number of valid cases; SD — standard deviation.

The baseline and single reflectance ratio algorithms assume a linear relationship between the depth of the PC absorption trough at 620 nm and PC concentrations, and have been developed from optical data obtained in eutrophic, cyanobacteria-dominated lakes. The observation of the residual errors from our dataset suggests that both at low PC concentrations and at very high PC concentration this relationship is no longer linear. It follows from Eqs. (2) and (3) that negative PC predictions will be obtained when the reflectance ratio $R_{rs}(650)/R_{rs}(625)$ drops below 0.97 for the single reflectance ratio algorithm, or when the difference between the midpoint of the 600–648 nm baseline and $R(0^-)$ at 624 nm is smaller than 0.0018 (for the baseline algorithm). Such situations are common in turbid waters with low cyanobacterial biomass (see Fig. 1) and in clearer (coastal) waters where pigment absorption in the red spectral region can be weak (Kutser et al., 2006). Nearly all negative residuals observed with the single reflectance ratio algorithm for sites in Spain indeed had $Chla_{CYANO}<2$ mg m^{-3}, whereas the majority of Spanish samples resulted in negative residuals with the baseline algorithm. The increasing negative residual error with higher PC concentrations (Fig. 8) resembles an apparent reduction of pigment-specific absorption coefficients known as the package effect. The nested band ratio algorithm appears unaffected, which can be explained by its use of a NIR band around 709 nm instead of the red band around 650 nm as reference wavelength. The reflectance in the NIR band tends to increase with phytoplankton biomass, and compensates the decrease in the specific absorption coefficient at 620 nm.

The systematic relative errors of the baseline algorithm in the Spanish samples suggest that regional parameterization of the algorithm could improve its performance. For the subset of cyanobacteria-dominated samples with $PC_{REF}<200$ mg m^{-3}, the empirical parameterization of the algorithm was recalculated by linear regression of the distance of $R(0^-,624)$ above the baseline as a function of PC_{REF} ($R^2=0.63$; $p<0.001$; $n=48$). The resulting relationship ($PC_{REF}=-20.0+16{,}224\times\{0.5\times[R(0^-,600)+R(0^-,648)]-R(0^-,624)\}$) was then substituted for Eq. (3) and applied to all Spanish samples. The fit between predicted and measured PC for $PC<200$ mg m^{-3} was clearly improved ($PC_{RAD}=-3.9+0.56\times PC_{REF}$; $R^2=0.81$ for the original parameterization and $PC_{RAD}=4.3+0.88\times PC_{REF}$; $R^2=0.82$ for the optimized algorithm; $p<0.001$, $n=136$ in both cases).

Effects of the presence of other phytoplankton pigments on the retrieval of PC can be mostly explained by their absorption overlapping the target waveband for PC around 620–624 nm. The effects are expected to be relatively more important at low PC concentrations. Chl *a* concentration showed high positive correlations with residuals in all algorithms for $PC_{REF}<200$ mg m^{-3}. The nested band ratio algorithm was least affected, as may be expected due to the incorporated correction for Chl *a* absorption. Chl *b* was negatively correlated with residuals for the Dutch dataset in the single reflectance ratio algorithm, but positively correlated for the other algorithms, while the correlation was not significant for the Spanish dataset. The presence of Chl *b* increases absorption around 650 nm, leading to a reduced reflectance ratio of 650 over 625 nm, and therefore lower predictions by the single reflectance ratio. At the same time, increasing Chl *b* concentrations would lower the baseline between 600 and 648 nm due to lowered reflectance around 650 nm, which should increase the PC prediction. The influence of Chl *b* on the nested band ratio is more obscure and it has been argued that a different correction for Chl *a* absorption at 620 nm is needed when green algae are present (Simis et al., 2007). Correction for Chl *a* absorption around 625 nm is not included in the other algorithms, but if the expression of *in vivo* Chl *a* absorption indeed differs between green algae and cyanobacteria, a similar effect would apply to these algorithms. The effects of fucoxanthin were weaker and generally negative, suggesting that a small number of samples with significant diatom presence revealed an effect of the presence of Chl *c*, but it is not possible to draw clear conclusions from the relatively low concentrations of the diatom pigments contained in this dataset. The same restriction holds for the remaining pigments, peridinin and alloxanthin. Finally, underestimations of PC_{REF} due to low extraction yields may also account for some apparent overestimates of PC_{RAD}, especially in the Dutch samples where the PC_{REF} could not be corrected with cyanobacterial biomass data.

The ratio of PC over cyanobacterial Chl *a* was nearly constant over a wide PC concentration range in the Spanish dataset. If these results can be extrapolated to water bodies elsewhere, a good estimate of PC would provide a predictor of cyanobacterial biomass in terms of their share in total Chl *a*. Laboratory studies have shown the dependence of PC production to light quality (Tandeau de Marsac and Houmard, 1988) and nutrient availability (Lemasson et al., 1973), so the stable relationship between PC and cyanobacterial Chl *a* found in the field in the current dataset is not necessarily evidence for an invariable intracellular PC: Chl *a* ratio. However, in nature and under light limited growth conditions, as commonly observed in cyanobacteria-dominated eutrophic waters, the internal pigment ratio may indeed prove fairly stable. In situations of cyanobacterial bloom, PC: Chl *a*

ratios in the water body rarely exceeded 2.5 in the Netherlands, while ratios above 4 where regularly observed in Spain. Therefore, the same stability of the PC: Chl *a* ratio in cyanobacteria may yet rely on differences in, for example, average solar irradiance or temperature, which would correlate with latitude. This is a testable hypothesis and investigating this issue would be useful for the further development of PC as a proxy of cyanobacterial biomass in remote sensing studies.

When deciding on the applicability of the algorithms for operational use, we should consider several factors. First, the quality of the input data should be carefully examined. If uncertainties in measurement methodology can be expected to affect the quality of the reflectance data, it may be best not to opt for the band ratio approaches but choose a baseline algorithm instead. This consideration is of less importance when the expected amplitude of the reflectance signal is high (in turbid waters) because spectrally neutral offset errors will have a limited influence on band ratio values. Second, all algorithms performed relatively poorly when cyanobacteria were a minor component of the phytoplankton community, as may be expected from overlapping absorption signatures. It is therefore important at all times to monitor both the results of the PC retrieval algorithm as well as Chl *a* measurements or predictions obtained from reflectance (e.g. Gons et al., 2005; Ruddick et al., 2001). In all the cases observed here, typical PC : Chl *a* pigment weight ratios limited to the cyanobacterial component of phytoplankton biomass (as far as discernable) were in the range 2–4. When a significantly higher ratio is observed from algorithm response, the retrieved PC value should probably be discarded. Third, in the most eutrophic conditions, the nested band ratio algorithm outperformed the other approaches. We suspect this is due to the use of a reference band at 709 nm. Reflectance at this wavelength was positively correlated with phytoplankton biomass, which cancelled out part of the errors caused by pigment packaging. This unintentional behaviour of the algorithm is useful, but it would be better if the effect were properly described and accounted for. This may be also possible using a single band ratio or baseline approach. Fourth, it should be noted that both the algorithm development and this validation used data from temperate lakes and reservoirs, so the extrapolation of these results to other climatic regions should be done with care. Eutrophic lakes, especially where cyanobacteria tend to dominate the phytoplankton populations, should not pose unexpected problems. The current dataset covers a large number of lakes that exhibit a long thermal stratification period and where nitrogen or light availability often limits algal growth. Cases of cyanobacterial presence in oligotrophic or mesotrophic lakes, as well as the early stages of cyanobacterial bloom and bloom termination are probably underrepresented. Finally, it should also be considered that the retrieved PC concentrations correspond to the visible surface layer of the water column. Approximately 90% of the remote sensed signal from a water body comes from the first optical depth (Gordon & McCunley, 1975). In the current dataset the average for visible light (400–700 nm) ranges from 4 m in the most oligotrophic reservoir to only 8 cm in the most hypereutrophic lake. Typical values for eutrophic waters are in the range of 0.2 to 1 m. The retrieved pigment concentrations will only ever be representative of the first optical depth, and only when the optical depth does not exceed the mixing depth. When floating scum is present, the observed reflectance is no longer influenced by the absorption by water and algorithms that rely on characterizing the absorption by water in the NIR, such as the nested band ratio algorithm, will fail (Kutser, 2004). The current dataset does include cases with shallow mixing depth and accumulation of biomass near the water surface, but does not include floating layers, which are however common in eutrophic water bodies, especially under calm weather conditions.

In all cases where operational optical monitoring is considered, the added value of remote sensing lies mostly in integrating information in space and time, and the added value is highest when complemented by point-based *in situ* data. The final choice of algorithm will depend on the sensor and platform from which reflectance data is derived. The current limitation for the use of these algorithms is a lack of spaceborn sensors with sufficient spectral and radiometric resolution. When the sensors do conform to the algorithm input requirements, such as in the case of MERIS and the nested band ratio algorithm, accurate atmospheric correction methods are required to avoid unexpected errors in the retrieval of PC.

5. Conclusions

The accuracy of three different algorithms for the remote sensing of the cyanobacterial pigment phycocyanin (PC) was evaluated with an extensive dataset of field radiometric and pigment data obtained in the period 2001–2005 in the Netherlands and Spain. The three algorithms yielded their best predictions in a range of PC concentrations corresponding to eutrophic conditions, higher errors were found for waters with either low or very high PC. When cyanobacteria formed the dominant phytoplankton group the accuracy of the algorithms was clearly higher, while the presence of other taxonomic groups led to increased errors, mainly due to partly overlapping absorption signatures. The relationship between PC and cyanobacterial Chl *a* biomass for a subset of Spanish samples was stable, suggesting that PC is a good proxy for cyanobacterial biomass. Pigment packaging effects at high concentrations were not accounted for by any of the algorithms, and led to marked underestimations at high PC concentrations, except for the nested band ratio algorithm where the use of a NIR band that correlated with biomass relaxed this effect somewhat. A correction for Chl *a* absorption in the latter algorithm also reduced errors in the low concentration range compared to the other algorithms. Chl *b* and Chl *c* have absorption features that overlap with PC and pigments characteristic of green algae and diatoms, respectively, indeed deteriorated PC retrieval accuracy in the low PC concentration range. This calls for a novel, perhaps holistic approach to separate the influence of the various red light absorbing pigments. Development of such algorithms will rely on hyper-spectral data and is not expected to be developed for spaceborn sensors.

The nested band ratio algorithm gave the best fit in most cases. The baseline algorithm agreed well with the Dutch data for moderate concentrations (within the range used for development of the algorithm) and could be reparameterized to better fit the Spanish dataset. The errors produced with the single reflectance ratio algorithm were generally larger, while it was least sensitive to instrumental errors affecting the absolute value of reflectance. In this context it is noted that both the baseline (Dekker, 1993) and nested band ratio algorithms (Simis et al., 2005) were developed for turbid, shallow, and eutrophic lakes in the Netherlands, using comparable PC extraction techniques and optical instruments, while the single reflectance ratio algorithm (Schalles & Yacobi, 2000) was developed and calibrated in a different setting. Overall, when comparing the magnitude of residual error produced by each algorithm, the use of the relatively complex nested band ratio approach is most obvious, while the baseline approach may prove useful when abovementioned variations in methodology and instrumentation can be well accounted for. In addition, the sensitivity of each algorithm to errors that are intrinsic to remotely sensed data rather than from handheld spectroradiometers must be investigated. For example, it was shown that ratios of MERIS band 709 over bands 620 or 665 nm are affected by poor atmospheric correction, leading to unexpected, high predictions of PC and Chl *a* (Simis et al., 2006). The other algorithms can not yet be applied to operational satellite sensors with global coverage as the required wavebands are lacking.

Over the past years, the possibilities for the remote sensing of phytoplankton group specific parameters have increased, both through the advance of sensor technology and the acquisition of

datasets covering a broad range of phytoplankton pigments, of which some will be useful as diagnostic pigments. The nature of aquatic photosynthesis itself, where a community of different algal and cyanobacterial species makes efficient use of a large part of the spectrum of visible solar irradiation, is unlikely to permit complete deconvolution of reflectance spectra into the contribution of all individual pigments. Most of these pigments have too much overlap in their absorption curves to be unambiguously resolved from the absorption or reflectance envelope. Similarly, the collection of water colour data prior to any interpretation in terms of phytoplankton content, is subject to various steps of image correction that always lead to some level of error. In many cases, simple empirical algorithms tuned to geographically limited areas or seasons may yield the most accurate interpretation of this imagery, which however does not reduce the need for globally valid algorithms that give non-experts the possibility to monitor these water quality parameters with at least a reasonable level of accuracy. It is not possible to highlight one approach as ideal for all causes. The three approaches compared in the current study all have definite advantages and drawbacks, ranging from ease of use and local calibration, spectral requirements, to sensitivity to methodological biases. The results of this study may therefore be used to decide on the optimal approach for cyanobacterial monitoring studies where the use of optical remote sensing tools is intended.

Acknowledgments

The authors thank the drivers and field technicians at the Centre for Hydrographic Studies (CEDEX, Spain) and the crew of R/V *Luctor* (NIOO, The Netherlands). José Antonio Domínguez-Gómez (CEDEX, Spain) is thanked for his support and cooperation in the preparation and execution of field campaigns. Antonio Quesada (Universidad Autónoma, Madrid, Spain) is thanked for the advice and the use of lab facilities for PC extractions. Further, thanks are due to Ana Alonso and Covadonga Alonso (CEDEX) and Cobie Kleppe-van Zetten (NIOO) for the HPLC pigment analysis. SGHS was funded through grant EO-053 from the User Support Programme managed by the programme office External Research of the Netherlands Organization for Scientific Research—National Institute for Space Research (NWO-SRON) and a grant from the Schure-Beijerinck Popping fund (KNAW).

References

Austin, R. W. (1980). Gulf of Mexico, ocean-colour surface-truth measurements. *Boundary — layer Meteorology*, *18*, 269–285.

Bidigare, R. R., Ondrusek, M. E., Morrow, J. H., & Kiefer, D. A. (1990). In vivo absorption properties of algal pigments. *SPIE Ocean Optics X*, *1302*, 290–302.

Bennett, A., & Bogorad, L. (1973). Complementary chromatic adaptation in a filamentous blue-green alga. *The Journal of Cell Biology*, *58*, 419–435.

Bryant, D. A. (1981). The photoregulated expression of multiple phycocyanin species. *European Journal of Biochemistry*, *119*, 425–429.

Codd, G. A., Bell, S. G., Kaya, K., Ward, C. J., Beattie, K. A., & Metcalf, J. S. (1999). Cyanobacterial toxins, exposure routes and human health. *European Journal of Phycology*, *34*, 405–415.

Codd, G. A., Chorus, I., & Burch, M. (1999b). Design of monitoring programmes. In: Toxic cyanobacteria in water: a guide to their public health consequences, monitoring and management, I. Chorus & J. Bartram (Eds.). WHO. E&FN Spon. London and New York. 416 pp.

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

Dekker, A. G., Malthus, T. J., & Seyhan, E. (1991). Quantitative modeling of inland water-quality for high-resolution MSS systems. *IEEE Transactions on Geoscience and Remote Sensing*, *29*, 89–95.

Ficek, D., Kaczmarek, S., Ston-Egiert, J., Wozniak, B., Majchrowski, R., & Dera, J. (2004). Spectra of light absorption by phytoplankton pigments in the Baltic; conclusions to be drawn from a Gaussian analysis of empirical data. *Oceanologia*, *46*, 533–555.

Gibson, C. E., & Smith, R. V. (1982). Freshwater plankton. In N. G Carr & B. A. Whitton (Eds.), *The Biology of Cyanobacteria* Oxford: Blackwell Science Publications.

Gordon, H. R., & McCunley, W. R. (1975). Estimation of the depth of sunlight penetration in the sea for remote sensing. *Applied Optics*, *14*(2), 413–416.

Gordon, H. R., Brown, O. B., & Jacobs, M. M. (1975). Computed relationships between the inherent and apparent optical properties of a flat homogeneous ocean. *Applied Optics*, *14*, 417–427.

Gons, H. J. (1999). Optical teledetection of chlorophyll a in turbid inland waters. *Environmental Science & Technology*, *33*, 1127–1132.

Gons, H. J., Kromkamp, J., Rijkeboer, M., & Schofield, O. (1992). Characterization of the light field in laboratory scale enclosures of eutrophic lake water (Lake Loosdrecht, The Netherlands). *Hydrobiologia*, *238*, 99–109.

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

Jeffrey, S. W., Mantoura, R. F. C., & Wright, S. W. (1997). *Phytoplankton pigments in oceanography : guidelines to modern methods.* Paris: Unesco.

Jupp, D. L. B., Kirk, J. T. O., & Harris, G. P. (1994). Detection, identification and mapping of cyanobacteria — Using remote-sensing to measure the optical-quality of turbid inland waters. *Australian Journal of Marine and Freshwater Research*, *45*, 801–828.

Kutser, T. (2004). Quantitative detection of chlorophyll in cyanobacterial blooms by satellite remote sensing. *Limnology and Oceanography*, *49*, 2179–2189.

Kutser, T., Metsamaa, L., Strombeck, N., & Vahtmae, E. (2006). Monitoring cyanobacterial blooms by satellite remote sensing. *Estuarine, Coastal and Shelf Science*, *67*, 303–312.

Lemasson, C., Tandeau de Marsac, N., & Cohen-Bazire, G. (1973). Role of allophycocyanin as a light-harvesting pigment in cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *70*, 3130–3133.

Montagner, F. (2001). *Reference model for MERIS Level 2 processing. European Space Agency, document PO-TN-MEL-GS-0026.*

NASA (2003). *Ocean Optics Protocols For Satellite Ocean Color Sensor Validation. Revision 4. NASA/TM-2003-21621/Rev-Vol III.*

Quesada, A., & Vincent, W. F. (1993). Adaptation to the light regime within Antarctic cyanobacterial mats. *Verhandlungen Internationale Vereinigung für Theoretische und Angewandte Limnologie*, *25*, 960–965.

Reynolds, C. S., & Walsby, A. E. (1975). Water-blooms. *Biological Reviews*, *50*, 437–481.

Rott, E. (1981). Some results from phytoplankton counting intercalibrations. *Schweizerische Zeitschrift für Hydrologie*, *43/1*, 34–59.

Ruddick, K. G., Gons, H. J., Rijkeboer, M., & Tilstone, G. (2001). Optical remote sensing of chlorophyll a in case 2 waters by use of an adaptive two-band algorithm with optimal error properties. *Applied Optics*, *40*, 3575–3585.

Ruiz-Verdú, A., Domínguez, J. A., & Peña-Martínez, R. (2005). Use of CHRIS for monitoring water quality in Rosarito reservoir. *Proceedings of the Third Chris Proba Workshop. ESA-ESRIN, Frascati, Italy, March 2005*.

Sarada, R., Pillai, M. G., & Ravishankar, G. A. (1999). 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*, *34*, 795–801.

Sathyendranath, S., Lazzara, L., & Prieur, L. (1987). Variations in the spectral values of specific absorption of phytoplankton. *Limnology and Oceanography*, *32*, 403–415.

Schalles, J. F., & Yacobi, Y. Z. (2000). Remote detection and seasonal patterns of phycocyanin, carotenoid and chlorophyll pigments in eutrophic waters. *Archiv fur Hydrobiologie Special Issues Advances in Limnology*, *55*, 153–168.

Simis, S. G. H., Peters, S. W. M., & Gons, H. J. (2005). Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. *Limnology and Oceanography*, *50*, 237–245.

Simis, S. G. H., Peters, S. W. M., & Gons, H. J. (2006). MERIS potential for remote sensing of water quality parameters for turbid inland water, p. 97–120. In: Simis, S.G.H., Blue-green catastrophe: remote sensing of mass viral lysis of cyanobacteria. Ph.D. thesis, Vrije Universiteit Amsterdam. doi: hdl.handle.net/1871/10641

Simis, S. G. H., Ruiz-Verdú, A., Domínguez-Gómez, J. A., Peña-Martinez, R., Peters, S. W. M., & Gons, H. J. (2007). Influence of phytoplankton pigment composition on remote sensing of cyanobacterial biomass. *Remote Sensing of Environment*, *106*, 414–427.

Sournia, A. (1978). Phytoplankton manual. *Monographs on Oceanographic Methodology*, *Vol. 6*. (pp.)Paris, France: UNESCO Publishing 337 pp.

Tandeau de Marsac, N., & Houmard, J. (1988). Complementary chromatic adaptation: physiological conditions and action spectra. *Methods in Enzymology*, *167*, 318–328.

Wyman, M., & Fay, P. (1986). Underwater light climate and the growth and pigmentation of planktonic blue-green-algae (cyanobacteria).1. The influence of light quantity. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *227*, 367–380.

Wyman, M., & Fay, P. (1986). Underwater light climate and the growth and pigmentation of planktonic blue-green-algae (cyanobacteria). 2. The influence of light quality. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *227*, 381–393.