



RESEARCH ARTICLE

10.1002/2017JC012859

Estimation of Phytoplankton Accessory Pigments from Hyperspectral Reflectance Spectra: Toward a Global Algorithm

A. P. Chase¹ , E. Boss¹ , I. Cetinić^{2,3}, and W. Slade⁴
¹School of Marine Sciences, University of Maine, Orono, ME, USA, ²GESTAR/Universities Space Research Association, Columbia, MD, USA, ³NASA Goddard Space Flight Center, Greenbelt, MD, USA, ⁴Sequoia Scientific, Inc., Bellevue, WA, USA

Key Points:

- Several phytoplankton accessory pigments are estimated via inversion of hyperspectral reflectance data
- Pigment covariation relationships are also used to estimate accessory pigments from reflectance-derived chlorophyll *a* values
- Both methods show a similar capability for estimating accessory pigments from hyperspectral reflectance on a global scale

Supporting Information:

- Supporting Information S1
- Table S1

Correspondence to:

A. P. Chase,
alison.p.chase@maine.edu

Citation:

Chase, A. P., Boss, E., Cetinić, I., & Slade, W. (2017). Estimation of phytoplankton accessory pigments from hyperspectral reflectance spectra: Toward a global algorithm. *Journal of Geophysical Research: Oceans*, 122. <https://doi.org/10.1002/2017JC012859>

Received 2 MAR 2017

Accepted 16 NOV 2017

Accepted article online 21 NOV 2017

Abstract Phytoplankton community composition in the ocean is complex and highly variable over a wide range of space and time scales. Able to cover these scales, remote-sensing reflectance spectra can be measured both by satellite and by in situ radiometers. The spectral shape of reflectance in the open ocean is influenced by the particles in the water, mainly phytoplankton and covarying nonalgal particles. We investigate the utility of in situ hyperspectral remote-sensing reflectance measurements to detect phytoplankton pigments by using an inversion algorithm that defines phytoplankton pigment absorption as a sum of Gaussian functions. The inverted amplitudes of the Gaussian functions representing pigment absorption are compared to coincident High Performance Liquid Chromatography measurements of pigment concentration. We determined strong predictive capability for chlorophylls *a*, *b*, $c_1 + c_2$, and the photoprotective carotenoids. We also tested the estimation of pigment concentrations from reflectance-derived chlorophyll *a* using global relationships of covariation between chlorophyll *a* and the accessory pigments. We found similar errors in pigment estimation based on the relationships of covariation versus the inversion algorithm. An investigation of spectral residuals in reflectance data after removal of chlorophyll-based average absorption spectra showed no strong relationship between spectral residuals and pigments. Ultimately, we are able to estimate concentrations of three chlorophylls and the photoprotective carotenoid pigments, noting that further work is necessary to address the challenge of extracting information from hyperspectral reflectance beyond the information that can be determined from chlorophyll *a* and its covariation with other pigments.

1. Introduction

The diversity of phytoplankton community composition in the ocean supports a wide range of ocean ecosystems. The ability to observe and monitor this diversity is necessary for understanding complex ocean processes, such as carbon export (e.g., Guidi et al., 2015; Mouw et al., 2016). Numerous methods are used to observe phytoplankton in the oceans, and all have benefits and limitations. Optical data, including spectral absorption and reflectance measurements, have the capability to be sampled at higher spatial and temporal resolution compared to analysis of discrete water samples. However, optical data also present the challenge of indirect observation of phytoplankton metrics such as biomass, composition, or pigments. Optical measurements made either in situ or remotely must be compared with coincidentally measured data such as pigment concentrations or phytoplankton imagery to establish the utility of these data and the associated uncertainties. Remote-sensing reflectance ($R_{rs}(\lambda)$; λ represents wavelength) can be estimated both from in situ data collected using radiometers and remotely by ocean color (OC) sensors on board aircraft and satellites. Estimation of chlorophyll *a* (Chl *a*) concentration, the major phytoplankton pigment used during photosynthesis, is well established from $R_{rs}(\lambda)$ measurements using methods such as the band-ratio model (Gordon et al., 1983) and a three-band reflectance differencing method (Hu et al., 2012).

In addition to Chl *a*, phytoplankton groups contain different assemblages of accessory pigments for both photosynthetic and photoprotective purposes; these include chlorophylls *b* and *c*, carotenoids, and biliproteins. Determining the presence of different accessory pigments can be used to help characterize phytoplankton community composition (e.g., Mackey et al., 1996). Although different algal groups may contain the same pigments, some distinctions between them can be made based on certain pigments. For example, chlorophyll *b* (Chl *b*) is found in green algae (*Chlorophyceae* and *Prasinophyceae*), whereas chlorophyll *c* (Chl *c*) is found in the golden-brown algae, which is a broad group including diatoms, dinoflagellates,

© 2017. The Authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

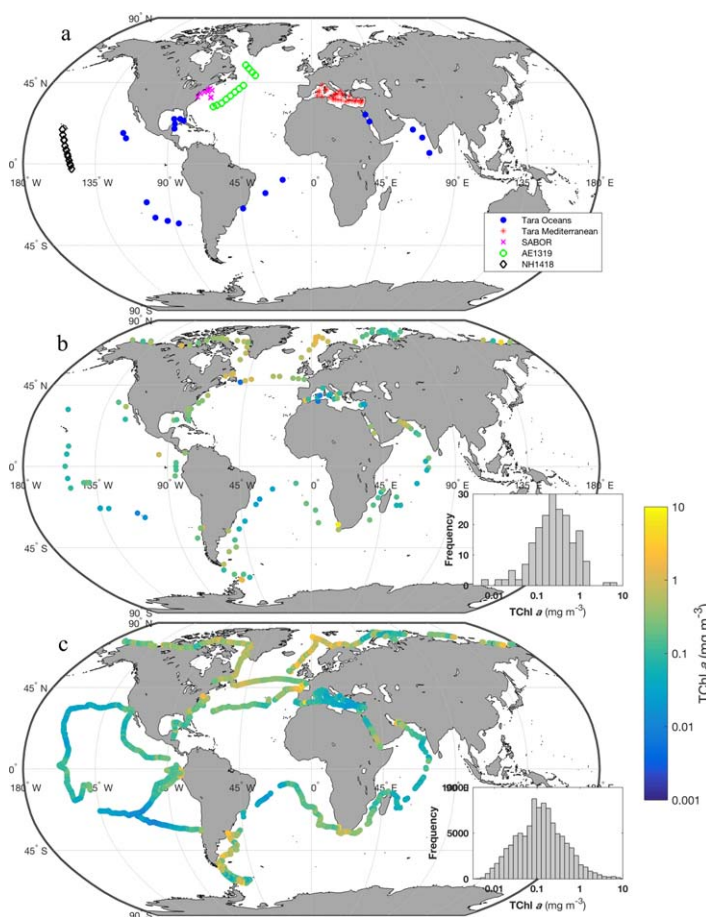


Figure 1. (a) Locations of coincident $R_{rs}(\lambda)$ and HPLC pigment data measured in situ during five different expeditions used in the inversion analysis ($n = 97$; see section 2.3). (b) Locations of the HPLC pigment data from Tara Expeditions used to calculate covariation relationships between Tchl a and each of Tchl b , Tchl $c_1 + c_2$, and PPC ($n = 196$). All data are from the surface. Inset shows the distribution of all Tchl a data. (c) Locations of $a_p(\lambda)$ spectra measured in situ with an ac-s meter deployed on a flow-through system during the Tara Oceans and Mediterranean Expeditions (2009–2014); $n = 96,929$. Inset shows the distribution of Tchl a concentration estimated following the line height method described in Boss et al. (2013).

silicoflagellates, and prymnesiophytes (Jeffrey & Vesik, 1997). Several xanthophyll pigments, a category of the carotenoids, are exclusive to distinct phytoplankton groups: alloxanthin is found in Cryptophyta; 19'-hexanoyloxyfucoxanthin is found in *Prymnesiophyceae* (including coccolithophorids); peridinin is found in dinoflagellates; and fucoxanthin is found in diatoms and golden-brown flagellates (which includes coccolithophorids and silicoflagellates) (Jeffrey & Vesik, 1997). The carotenoid pigments can also be grouped together as photoprotective (PPC) and photosynthetic (PSC) carotenoids; these two groups are made up of pigments that both have related roles in the cell and have similar light absorption spectra (e.g., Figure 1 in Bricaud et al., 2004). Previous efforts to move beyond the detection of Chl a and retrieve information about accessory pigments or different phytoplankton groups directly are described below, following a brief description of the theoretical relationship between pigment absorption and $R_{rs}(\lambda)$.

The total absorption ($a(\lambda)$) and backscattering ($b_b(\lambda)$) by all components in the water, including phytoplankton pigments, are determinants of $R_{rs}(\lambda)$ and their relationship is approximated (while ignoring inelastic scattering) by

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

The parameters f and Q relate measurements made at one angle to the nadir direction and are a function of sun angle, atmospheric conditions, and the optical properties of the water (Morel & Gentili, 1996). Total phytoplankton absorption ($a_p(\lambda)$) includes the absorption spectra of Chl a as well as all accessory pigments present. With multiple, varying phytoplankton pigments affecting $a(\lambda)$ and $a(\lambda)$ influencing $R_{rs}(\lambda)$, it has been the goal of several studies to estimate phytoplankton accessory pigments from $a(\lambda)$ or $R_{rs}(\lambda)$. Hyperspectral particulate absorption data ($a_p(\lambda)$), with information provided every few nanometers, have been used to estimate phytoplankton pigments from spectra measured in the laboratory (Hoepffner & Sathyendranath, 1991, 1993; Lohrenz et al., 2003; Moisan et al., 2011) and in situ (Chase et al., 2013). Extending a similar analysis to hyperspectral $R_{rs}(\lambda)$ data is more challenging and is complicated by the need to account for additional parameters that do not come into play with direct analysis of particulate or phytoplankton absorption spectra. These include inelastic scattering, particulate backscattering ($b_{bp}(\lambda)$) and absorption by dissolved matter and nonalgal particles ($a_{CDOM}(\lambda)$ and $a_{NAP}(\lambda)$, respectively). In addition, there are uncertainties in the relationship between $R_{rs}(\lambda)$ and absorption coefficients (e.g. due to variability in f and Q in equation (1)).

Pigment-based clusters representing different phytoplankton assemblages were previously compared with clusters of absorption spectra, the second derivative of absorption spectra, and $R_{rs}(\lambda)$ data by Torrecilla et al. (2011). The authors found that both the absorption data and the second derivative of $R_{rs}(\lambda)$ compared well with pigment data, suggesting the utility of the optical measurements for describing phytoplankton pigments in the ocean. However, to use their method at another location or time, in situ data are needed to link clusters of pigments with clusters of optical data. Bracher et al. (2015) used an empirical orthogonal function (EOF) analysis to derive linear regression models between pigments and $R_{rs}(\lambda)$ spectra and found strong predictive capabilities for total Chl a (Tchl a) and several accessory pigments in the Atlantic Ocean. A study by Pan et al. (2010) calculated ratios between $R_{rs}(\lambda)$ wavelengths from in situ data to develop algorithms for pigment prediction that were then applied to multispectral satellite data for the U.S. northeast coastal region. They used High Performance Liquid Chromatography (HPLC) pigment data for model development and validation and were able to estimate phytoplankton pigment concentrations from SeaWiFS

and MODIS satellite $R_{rs}(\lambda)$ with mean absolute percent error for most pigments falling between 30% and 50%. Wang et al. (2016) estimated several pigment absorption coefficients from inversion of in situ hyperspectral $R_{rs}(\lambda)$ data for inland lake waters containing high chlorophyll values and cyanobacterial bloom conditions ($\text{Chl } a > 10 \mu\text{g L}^{-1}$).

Another approach for the identification of phytoplankton groups in the ocean is to explore spectral anomalies and residuals that remain after removal of an average chlorophyll-based spectrum. This approach addresses the high covariation among phytoplankton pigments that has been observed on a global scale (Trees et al., 2000) and is used to extract information beyond the average relationship between $\text{Chl } a$ and its covarying parameters, including accessory pigments. Brown et al. (2008) determined that $a_{\text{NAP}}(\lambda)$ and the magnitude of $b_b(\lambda)$ are responsible for the spread around the mean $R_{rs}(\lambda) - \text{Chl } a$ relationship, while Alvain et al. (2012) found that the $\text{Chl } a$ -specific phytoplankton absorption ($a^*(\lambda)$), $a_{\text{CDOM}}(\lambda)$, and $b_b(\lambda)$ all influence reflectance anomalies. Ben Mustapha et al. (2013) built off previous work to improve the PHYSAT method (Alvain et al., 2005, 2008) to use multispectral satellite radiance anomalies to detect the dominance of several phytoplankton groups in the open ocean.

Many previous studies for detecting phytoplankton groups and pigments were conducted using multispectral reflectance data and/or were developed for a limited geographical region (e.g., Alvain et al., 2005, 2008; Ben Mustapha et al., 2013; Farikou et al., 2015; Raitos et al., 2008; Werdell et al., 2014; also see Mouw et al., 2017 for a recent review). However, there is a need to understand both the added value and limitations of hyperspectral data and global algorithms, particularly given the increased spectral resolution of next-generation ocean color satellites (e.g., NASA's Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission and Germany's EnMAP mission). Algorithms that estimate accessory pigments over a broad global range of water types and without prior knowledge of the phytoplankton community composition may be preferred when global hyperspectral satellite $R_{rs}(\lambda)$ data become available within the next decade.

In this study, we explore the utility of hyperspectral $R_{rs}(\lambda)$ for estimating phytoplankton pigments using a global database of in situ $R_{rs}(\lambda)$ and HPLC measurements. We attempt to exploit the spectral differences in absorption by different phytoplankton accessory pigments, which in turn should influence the spectral shape of $R_{rs}(\lambda)$. As a comparison, we also examine the covariation between $\text{Chl } a$ and accessory pigments and the capabilities of estimating accessory pigments using covariation relationships. Finally, we conduct a brief analysis of spectral residuals defined by deviations from the average global relationship between $\text{Chl } a$ and $a_p(\lambda)$. These approaches help us understand the potential and the limitations for extracting information on phytoplankton pigments from hyperspectral $R_{rs}(\lambda)$ data on a broad global scale.

2. Data and Methods

2.1. Data Sets

Data sets from five different expeditions are used in our study (Table 1). For all data sets described below, HPLC pigment data were collected using discrete surface water samples (depth ≤ 5 m) that were filtered and preserved on board. When duplicate or triplicate HPLC samples were available the mean value is used. We studied four HPLC pigment groups (Table 2). $a_p(\lambda)$ data from an ac-s spectrophotometer (WET Labs, Inc., Philomath, OR, USA) deployed in a flow-through setup (Slade et al., 2010) are available for three of the expeditions (Tara Oceans, Tara Mediterranean, and SABOR). The ac-s data measured coincidentally with $R_{rs}(\lambda)$ are used for calculating an attenuation correction when processing $R_{rs}(\lambda)$ spectra; see section 2.3 for ac-s and $R_{rs}(\lambda)$ processing details.

2.1.1. Tara Expeditions

Data from two extended expeditions on the R/V Tara are used: Tara Oceans (including Tara Oceans Polar Circle; global coverage from 2009–2013; Boss et al., 2014; Picheral et al., 2014) and Tara Mediterranean (Mediterranean Sea; June–September 2014). $R_{rs}(\lambda)$ spectra and HPLC pigment data collected within four hours of each other were used in the inversion analysis. There are 58 data points of coincident $R_{rs}(\lambda)$ and HPLC data from varied water types around the globe from the Tara Expeditions (Figure 1a). The complete HPLC data sets from the Tara Expeditions are also used in our development of global pigment covariation relationships ($n = 196$; Figure 1b); these data are independent from the data used in the inversion analysis. Laboratory analysis of HPLC data was carried out at Laboratoire d'Océanographie de Villefranche-sur-Mer (LOV) according to the method described by Ras et al. (2008), which is adapted from Van Heukelem and

Table 1
Summary of Data Sets Used in This Study

Expedition	Dates	Location(s)	Data type and number of points			
			$R_{rs}(\lambda)$ and HPLC for inversion analysis	HPLC global covariation	$a_p(\lambda)$ for residual analysis	$a_p(\lambda)$ for $R_{rs}(\lambda)$ processing
Tara Oceans ^a	2009–2013	Global	29	196	89,421	25
Tara Mediterranean ^b	June–September 2014	Mediterranean Sea	29		7,508	29
SABOR ^b	July–August 2014	Gulf of Maine/North Atlantic/Mid-Atlantic coast	12			12
AE1319 ^b	August–September 2013	North Atlantic and Labrador Sea	13			
NH1418 ^b	September–October 2014	Equatorial Pacific	14			
Totals			97	196	96,929	66^c

^aIncludes Tara Oceans Polar Circle Expedition; data available at <https://doi.pangaea.de/10.1594/PANGAEA.836318> and <https://doi.pangaea.de/10.1594/PANGAEA.836319>.

^bData available at <https://seabass.gsfc.nasa.gov/> (Werdell et al., 2003).

^cSee section 2.2 for an explanation of $R_{rs}(\lambda)$ processing for the 31 data points without matching $a_p(\lambda)$ data.

Thomas (2001). Finally, the global data set of ac-s spectral absorption measurements from the Tara Expeditions was used to calculate the shape and magnitude of average global particulate absorption spectra used in the residual analysis (Figure 1c; $n = 96,929$; see section 2.5).

2.1.2. SABOR, AE1319, and NH1418 Expeditions

Coincident HPLC and $R_{rs}(\lambda)$ data from three additional expeditions were used in the inversion analysis: SABOR (Gulf of Maine/North Atlantic/Mid-Atlantic coast; July–August 2014), AE1319 (North Atlantic and Labrador Sea; August–September 2013), and NH1418 (Equatorial Pacific; September–October 2014); see Figure 1a for data locations. $R_{rs}(\lambda)$ spectra and HPLC pigment data collected within 4 h of each other were used in the inversion analysis. The Ocean Ecology Laboratory at NASA Goddard Space Flight Center performed HPLC analysis for the SABOR, AE1319, and NH1418 expeditions following methods in Van Heukelem and Thomas (2001) and further described in Hooker et al. (2009). The three expeditions combined contribute 39 matching HPLC and $R_{rs}(\lambda)$ data points (Table 1).

2.2. $R_{rs}(\lambda)$ Data and Processing

$R_{rs}(\lambda)$ spectra were calculated from upwelling radiance ($L_u(\lambda)$) and downwelling irradiance ($E_d(\lambda)$) spectra measured with a Profiler II radiometer suite using HyperOCR sensors (Satlantic, Halifax, NS, Canada) deployed with detachable float collar in HyperTSRB “buoy mode” (Figure 2). In this configuration, the in-situ radiometers are deployed while the instrument is tethered to the vessel away from the ship shadow and floating at the ocean surface. The sensor that collects downwelling light is above the ocean surface in the air, and the sensor that collects the upwelling light is approximately 0.2 m below the ocean surface. We only keep the $E_d(\lambda)$ and $L_u(\lambda)$ data that fall between the 25th and 75th percentiles to eliminate any outliers

Table 2
Pigment Groups From HPLC Analysis

Pigment group	Abbreviation	Sum
Total Chlorophyll <i>a</i>	TChl <i>a</i>	Monovinyl Chl <i>a</i> + divinyl Chl <i>a</i> + chlorophyllide <i>a</i> + Chl <i>a</i> allomers + Chl <i>a</i> epimers
Total Chlorophyll <i>b</i>	TChl <i>b</i>	Monovinyl Chl <i>b</i> + divinyl Chl <i>b</i> + Chl <i>b</i> epimers
Chlorophylls <i>c</i> ₁ and <i>c</i> ₂	Chl <i>c</i> ₁ + <i>c</i> ₂	Chlorophyll <i>c</i> ₁ + chlorophyll <i>c</i> ₂
Photoprotective carotenoids	PPC	$\alpha\beta$ -carotene + zeaxanthin + alloxanthin + diadinoxanthin

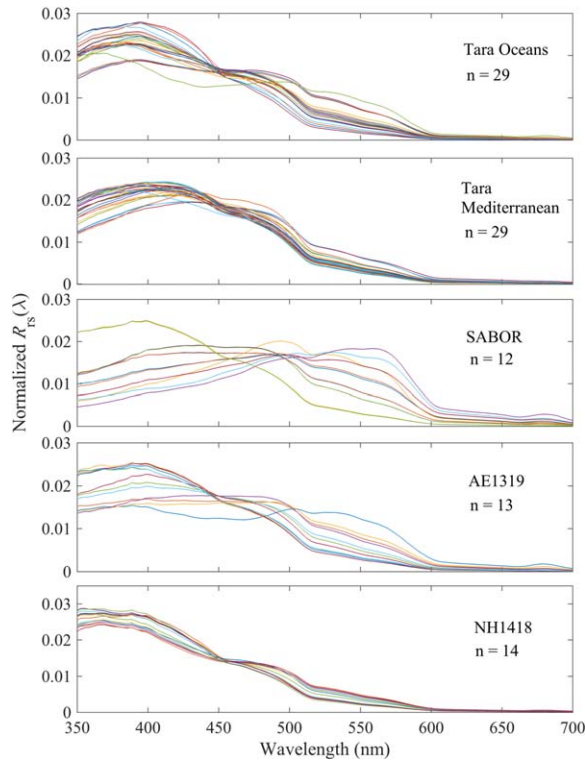


Figure 2. $R_{rs}(\lambda)$ spectra measured in situ with a Satlantic HyperTSRB. Each spectrum is normalized to the area under the curve to emphasize differences in spectral shape. See Figure 1a for the locations of each expedition.

before computing $R_{rs}(\lambda)$, which is the desired parameter for ultimately modeling the absorption of phytoplankton pigments via inversion. $R_{rs}(\lambda)$ is defined as

$$R_{rs}(\lambda) = \frac{L_w(\lambda, 0^+)}{E_d(\lambda)}, \quad (2)$$

where $L_w(\lambda, 0^+)$ is the water-leaving radiance just above the sea surface. $L_w(\lambda, 0^+)$ is derived from $L_u(\lambda)$ by first extrapolating $L_u(\lambda)$ from the sensor depth ($z = 0.2$ m) to just below the sea surface ($L_u(\lambda, z_{0-})$) using

$$L_u(\lambda, z_{0-}) = L_u(\lambda) e^{-K_{Lu}(\lambda) \cdot z}, \quad (3)$$

where $K_{Lu}(\lambda)$ is the upwelling attenuation coefficient approximated as

$$K_{Lu}(\lambda) \cong \frac{a_w(\lambda) + a_p(\lambda)}{\mu_u}, \quad (4)$$

and μ_u is the average cosine for the upwelling light and is approximated as 0.5.

Absorption by seawater ($a_w(\lambda)$) in the UV and visible is known (Mason et al., 2016; Pope & Fry, 1997) and corrected for temperature and salinity (Sullivan et al., 2006), which are measured coincidentally. The $a_p(\lambda)$ spectra are measured with the ac-s meter deployed on a flow-through system that records both total and dissolved absorption, and then $a_p(\lambda)$ spectra are calculated by difference (Slade et al., 2010). This method provides $a_p(\lambda)$ measurements that are independent of calibration and instrument drift problems, allowing extended continuous at-sea deployment. However, the method also results in un-calibrated total and dissolved absorption measurements; as a result, equation (4)

neglects the contribution of CDOM. Additionally, equation (4) does not account for inelastic scattering, which will decrease the effective attenuation. There were 31 cases where no coincident $a_p(\lambda)$ data were available for the correction in equation (4). For these data we developed an iterative method to estimate the appropriate $a_p(\lambda)$ spectra, as follows: first, the Chl *a* value for a given $R_{rs}(\lambda)$ signal was estimated using the generalized inherent optical properties (GIOP) model (Werdell et al., 2013). Second, the $a_p(\lambda)$ spectrum for the estimated Chl *a* value was defined using the A_{chl} and B_{chl} coefficients determined in this study (see section 2.5.1 for details), and used to calculate $K_{Lu}(\lambda)$ (equation (4)). Finally, the $K_{Lu}(\lambda)$ value was used in the calculation of $R_{rs}(\lambda)$ (equations (2–4)). The process was iterated upon until the Chl *a* value determined using GIOP before the attenuation correction was within 1% of the Chl *a* value following the attenuation correction. A self-shading correction was also calculated and applied to the $L_u(\lambda)$ spectra by following the methods in Leathers et al. (2001).

The water-leaving radiance exiting the sea surface, $L_w(\lambda, 0^+)$, is then calculated as

$$L_w(\lambda, 0^+) = \frac{t L_u(\lambda, z_{0-})}{n^2}, \quad (5)$$

where n is the refractive index of seawater and t is the radiance transmittance of the surface (assumed to be 1.34 and 0.98, respectively; Mobley, 1994).

2.2.1. Raman Scattering Correction

Raman scattering by water molecules contributes to the water-leaving radiance ($L_w(\lambda, 0^+)$), and can therefore influence $R_{rs}(\lambda)$, particularly in the blue wavelengths in clear waters (McKinna et al., 2016 and references therein). To account for Raman scattering, the Raman-specific $L_w(\lambda, 0^+)$ spectra are computed and subtracted from the original $L_w(\lambda, 0^+)$ spectra. Briefly, the steps to complete this are as follows: (1) use a radiation model to estimate the ultraviolet (UV) and visible $E_d(\lambda)$ for the day, time and location on the globe where the in situ $R_{rs}(\lambda)$ data were collected (Gregg & Carder, 1990; as coded by Richard Davis, 1998, personal communication); (2) force the measured and modeled $E_d(\lambda)$ spectra to match at the visible wavelengths, and use the corresponding modeled UV values (as the radiometer does not collect data at the necessary UV

wavelengths); (3) use the GIOP model (Werdell et al., 2013) to generate inherent optical properties ($a(\lambda)$ and $b_b(\lambda)$; extrapolation used for UV values) for use in calculating $K_{Lu}(\lambda)$ and the downwelling attenuation coefficient ($K_d(\lambda)$); (4) finally, using the calculated $K_{Lu}(\lambda)$ and $K_d(\lambda)$ values and following methods in Westberry et al. (2013), calculate the Raman-specific $L_w(\lambda, 0^+)$ spectra and subtract them from the corresponding original $L_w(\lambda, 0^+)$ spectra. The resulting Raman-corrected $L_w(\lambda, 0^+)$ spectra are used to calculate $R_{rs}(\lambda)$ (equation (2)). The spectral resolution of the final $R_{rs}(\lambda)$ data is approximately 3.35 nm (± 0.05 nm).

2.2.2. Correction for Angular Effects

Following the correction for Raman scattering the $R_{rs}(\lambda)$ spectra are normalized to eliminate the angular effect of the sun position in the sky relative to nadir. Following the methods described in Lee et al. (2011), we first use the quasianalytical algorithm (Lee et al., 2002) to estimate $a(\lambda)$ and $b_b(\lambda)$ for a given $R_{rs}(\lambda)$ spectrum. The normalized $R_{rs}(\lambda)$ spectra are then calculated using equations (14–20) in Lee et al. (2011), and resulting $R_{rs}(\lambda)$ spectra are used in the inversion algorithm described in section 2.3. The absolute percent differences between the original and normalized $R_{rs}(\lambda)$ spectra are 0–5%, 0–6%, and 0.5–9.5% at 440, 490, and

550 nm, respectively. The effect of normalization is small and does not significantly change the outcome during estimation of pigment concentrations; however, we include it as a known effect on the $R_{rs}(\lambda)$ spectra that when accounted for results in spectra that more closely represent the data calculated from satellite information.

Table 3

Parameters Allowed to Vary During the Inversion of $R_{rs}(\lambda)$ Spectra^a

Parameter	Units	First guess	Lower bound	Upper bound	Mean \pm SD after inversion
C_{NAP}	m^{-1}	0.005	0	0.05	0.004 ± 0.005
S_{NAP}	nm^{-1}	0.011^b	0.005	0.016	0.013 ± 0.003
C_{CDOM}	m^{-1}	0.1	0.01	0.8	0.047 ± 0.044
S_{CDOM}	nm^{-1}	0.0185	0.005	0.02	0.018 ± 0.002
b_{bp}/b_p		0.01^c	0.005	0.015	0.007 ± 0.002
C_{cp}	m^{-1}	0.1	0.01	1	0.103 ± 0.070
γ		1	0	1.3	1.156 ± 0.221
$a_{gaus}(\lambda)$ center λ	nm	384^d	383	385	383.81 ± 0.46
	nm	413^d	412	414	413.44 ± 0.73
	nm	435^d	434	436	435.50 ± 0.55
	nm	461^d	460	462	460.15 ± 0.42
	nm	464^d	463	465	464.13 ± 0.80
	nm	490^d	489	491	489.23 ± 0.49
	nm	532^d	531	533	531.77 ± 0.63
	nm	583^d	582	584	582.55 ± 0.68
$a_{gaus}(384)$	m^{-1}	0.01	0	0.5	0.014 ± 0.036
$a_{gaus}(413)$	m^{-1}	0.01	0	0.5	0.005 ± 0.006
$a_{gaus}(435)$	m^{-1}	0.01	0	0.5	0.014 ± 0.015
$a_{gaus}(461)$	m^{-1}	0.01	0	0.5	0.004 ± 0.004
$a_{gaus}(464)$	m^{-1}	0.01	0	0.5	0.007 ± 0.006
$a_{gaus}(490)$	m^{-1}	0.01	0	0.5	0.010 ± 0.007
$a_{gaus}(532)$	m^{-1}	0.01	0	0.5	0.014 ± 0.006
$a_{gaus}(583)$	m^{-1}	0.01	0	0.5	0.022 ± 0.012
$\sigma(384)$	nm	23^d	22	24	22.81 ± 0.51
$\sigma(413)$	nm	9^d	8	10	9.86 ± 0.37
$\sigma(435)$	nm	14^d	13	15	14.77 ± 0.45
$\sigma(461)$	nm	11^d	10	12	10.22 ± 0.44
$\sigma(464)$	nm	19^d	18	20	19.85 ± 0.40
$\sigma(490)$	nm	19^d	18	20	18.24 ± 0.50
$\sigma(532)$	nm	20^d	19	21	19.63 ± 0.70
$\sigma(583)$	nm	20^d	19	21	20.80 ± 0.47

^aSee section 2.3 for parameter definitions. $n = 97$ inversion runs used to calculate mean and standard deviation (SD). b_{bp}/b_p and γ are dimensionless parameters. MATLAB code used for the inversion is available at <http://misclab.umeoce.maine.edu/software.php>.

^bRoesler et al. (1989).

^cTwardowski et al. (2001).

^dHoepffner and Sathyendranath (1991); Bricaud et al. (2004).

2.3. Inversion of $R_{rs}(\lambda)$ Spectra

To directly model the backscattering and absorption components, we used a model developed by Gordon et al. (1988), where the term $u(\lambda)$ is defined as

$$u(\lambda) \equiv \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}, \quad (6)$$

and then used in the quadratic equation

$$r_{rs}(\lambda) = g_1 u(\lambda) + g_2 u(\lambda)^2, \quad (7)$$

where $g_1 = 0.0949$ and $g_2 = 0.0794$ are constants computed by Gordon and Brown (1988) and $r_{rs}(\lambda)$ is the remote-sensing reflectance just below the sea surface and can be calculated from $R_{rs}(\lambda)$ using the method from Lee et al. (2002):

$$r_{rs}(\lambda) = \frac{R_{rs}(\lambda)}{0.52 + 1.7R_{rs}(\lambda)}. \quad (8)$$

The quadratic formula is used to solve equation (7) for $u(\lambda)$ and the one positive solution, denoted $u_{meas}(\lambda)$, is used in a weighted nonlinear least squares inversion to determine the combination of functions representing absorption and backscattering by different constituents in the water that most closely matches the $u_{meas}(\lambda)$ spectrum. The inversion algorithm iterates to find the best fit while allowing for variation of 31 different parameters (Table 3), by minimizing the function

$$\chi^2 = \sum_{i=1}^{60} \left(\frac{u_{meas}(\lambda_i) - u_{mod}(\lambda_i)}{u_{std}(\lambda_i)} \right)^2, \quad (9)$$

where the data are summed over the angles 400–600 nm at approximately 3.35 nm resolution (60 total angles), and $u_{mod}(\lambda)$ is the reconstructed spectrum using the spectral components described below. The total backscattering and absorption are broken down into their component spectra, and the expanded version of $u_{mod}(\lambda)$ is defined as

$$u_{mod}(\lambda) = \frac{b_{bp}(\lambda) + b_{bw}(\lambda)}{a_{\varphi}(\lambda) + a_{CDOM}(\lambda) + a_{NAP}(\lambda) + a_w(\lambda) + b_{bp}(\lambda) + b_{bw}(\lambda)}, \quad (10)$$

where $b_{bw}(\lambda)$ is the known backscattering by seawater (Zhang et al., 2009), $a_{\varphi}(\lambda)$ is the absorption by phytoplankton, and $a_w(\lambda)$ is the known absorption by seawater (see references in section 2.2). The $b_{bp}(\lambda)$ spectrum is defined in terms of the particulate backscattering ratio (b_{bp}/b_p ; assumed to be spectrally constant), $a_p(\lambda)$, and particulate attenuation ($c_p(\lambda)$) after Roesler and Boss (2003):

$$b_{bp}(\lambda) = \frac{b_{bp}}{b_p} (c_p(\lambda) - a_p(\lambda)), \quad (11)$$

where $b_p(\lambda) = c_p(\lambda) - a_p(\lambda)$; $a_p(\lambda) = a_{\varphi}(\lambda) + a_{NAP}(\lambda)$; and $c_p(\lambda)$ is defined as a decreasing power law function:

$$c_p(\lambda) = C_{cp} \left(\frac{\lambda}{\lambda_0} \right)^{-\gamma}. \quad (12)$$

The spectra for $a_{NAP}(\lambda)$ and $a_{CDOM}(\lambda)$ are defined as decreasing exponential functions:

$$a_{NAP}(\lambda) = C_{NAP} e^{-S_{NAP}(\lambda - \lambda_0)}, \quad (13)$$

$$a_{CDOM}(\lambda) = C_{CDOM} e^{-S_{CDOM}(\lambda - \lambda_0)}, \quad (14)$$

where λ_0 is set to 400 nm. In equations (11–14) the following parameters are allowed to vary: b_{bp}/b_p ; all of the components in $a_p(\lambda)$ (see equation (15) for details on how we define the $a_{\varphi}(\lambda)$ components); the magnitude and slope (C_{cp} and γ , respectively) of $c_p(\lambda)$; and the magnitudes (C_{NAP} and C_{CDOM}) and exponential slopes (S_{NAP} and S_{CDOM}) of $a_{NAP}(\lambda)$ and $a_{CDOM}(\lambda)$.

The spectrum for $a_{\varphi}(\lambda)$ is defined as a linear combination of eight Gaussian functions, rather than a single eigenvector as is usually the case during inversion of $R_{rs}(\lambda)$ spectra into absorbing and scattering components. The individual Gaussian functions ($a_{gaus}(\lambda)$) represent absorption by different phytoplankton pigments or pigment groups, and are defined as:

$$a_{\varphi}(\lambda) = \sum_{i=1}^8 a_{gaus}(peak_i, \lambda) e^{\left(-0.5 \left(\frac{\lambda - peak_i}{\sigma_i} \right)^2 \right)}, \quad (15)$$

with $peak_i$ and σ_i representing the center wavelengths and widths of each Gaussian, respectively, where σ is related to the full width half maximum (FWHM) by $FWHM = 2.355\sigma$. Although 12 $a_{gaus}(\lambda)$ functions were initially defined, the signal in the red part of the $R_{rs}(\lambda)$ spectra (>600 nm) is relatively low due to the strong absorption by water. We found improved inversion results in terms of reduced median errors when the $R_{rs}(\lambda)$ data were restricted to wavelengths of 400–600 nm; this prevents the inversion from fitting Gaussian functions in the red wavelengths (between 600 and 730 nm) at the cost of poor fitting at the blue end of the spectrum where the $R_{rs}(\lambda)$ signal has lower relative uncertainty. As a result, we use eight Gaussian functions (Table 3). A similar finding by Isada et al. (2015) showed that the $R_{rs}(\lambda)$ information >547 nm was dominated by water absorption and therefore not useful for phytoplankton group detection in a Northern Japan bay. We tested several methods for choosing the optimal $peak_i$ and σ_i values, as evaluated by final correlation and error statistics. These included a derivative analysis and an iterative method in which the $peak_i$ and σ_i values were allowed to change by ± 5 nm. However, we ultimately defined the $peak_i$ and σ_i values based on known pigment absorption shapes (see Figure 1, Bricaud et al. 2004) and limited the allowable shift of the $peak_i$ position during inversion to ± 1 nm (Table 3). The first guess and upper and lower bound values for each parameter in the inversion are based on known values from previous studies of various water types, or from our testing of the inversion algorithm using a range of values (Table 3).

The inversion results were improved by normalizing $u_{meas}(\lambda) - u_{mod}(\lambda)$ by the uncertainty of $u_{meas}(\lambda)$ at each wavelength (equation (9)). We calculated the uncertainty by first calculating $u_{meas}(\lambda)$ for every $E_d(\lambda)$ and $L_u(\lambda)$ spectral pair from a given deployment; for example, a typical deployment from Tara Oceans resulted in roughly 200 spectra from a 3 min time series. We then used the standard deviation of all $u_{meas}(\lambda)$ spectra from one deployment as the uncertainty value ($u_{std}(\lambda)$). These calculations did not include the upper and lower 25th percentiles of $E_d(\lambda)$ and $L_u(\lambda)$ data; their removal is described previously in section 2.2.

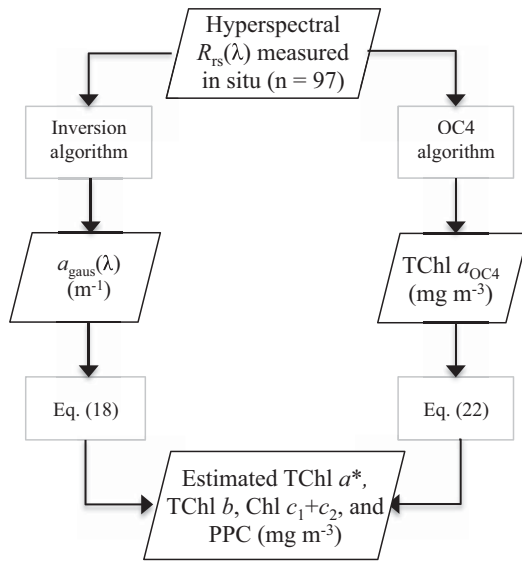


Figure 3. Schematic of the data and processing steps used to estimate accessory pigments from hyperspectral $R_{rs}(\lambda)$ data. *Note that TChl a in the bottom box is estimated only from $a_{\text{gaus}}(\lambda)$; TChl a estimated from $R_{rs}(\lambda)$ is computed using the OC4 algorithm (TChl a_{OC4} in middle right box).

2.4. Estimation of Pigment Concentrations

The concentrations of TChl a , TChl b , Chl $c_1 + c_2$, and PPC were initially estimated using the relationships between $a_{\text{gaus}}(\lambda)$ magnitudes derived from $R_{rs}(\lambda)$ spectra and HPLC pigments (section 2.4.1). As a comparison, the concentrations of TChl b , Chl $c_1 + c_2$, and PPC were estimated from $R_{rs}(\lambda)$ -derived TChl a (denoted TChl a_{OC4}) using the covariation relationships among pigments as calculated from a global HPLC data set (section 2.4.2). The TChl a_{OC4} values were calculated using NASA's OC4 algorithm (O'Reilly et al. 2000; see Appendix A for details). A schematic of the two approaches used to estimate pigment concentrations is provided (Figure 3).

2.4.1. Pigments Estimated from $a_{\text{gaus}}(\lambda)$ –HPLC Relationships

Data sets with coincident $a_{\text{gaus}}(\lambda)$ derived from inverted $R_{rs}(\lambda)$ and HPLC pigments ($n = 97$; Table 1) were compared and used to estimate phytoplankton pigment concentrations: 435 nm and TChl a ; 464 nm and TChl b ; 461 nm and Chl $c_1 + c_2$; and 490 nm and PPC (Figure 4). We used a nonlinear fit of each $a_{\text{gaus}}(\lambda)$ and corresponding HPLC pigment concentration ($[\text{pigment}_{\text{HPLC}}]$) to determine the two coefficients (A_{inv} and B_{inv} ; Table 4) in the equation

$$a_{\text{gaus}}(\lambda) = A_{\text{inv}} [\text{pigment}_{\text{HPLC}}]^{B_{\text{inv}}}, \quad (16)$$

by minimizing the cost function

$$\chi^2 = \sum_{i=1}^{97} \left(a_{\text{gaus}}(\lambda)_i - A_{\text{inv}} * [\text{pigment}_{\text{HPLC}}]_i^{B_{\text{inv}}} \right)^2, \quad (17)$$

where there are 97 coincident $a_{\text{gaus}}(\lambda)$ and HPLC data points and the function is minimized over the entire data set. The A_{inv} and B_{inv} coefficients obtained from equations (16) and (17) are then rearranged to solve for the estimated pigment concentration ($[\text{pigment}_{\text{inv}}]$; mg m^{-3}):

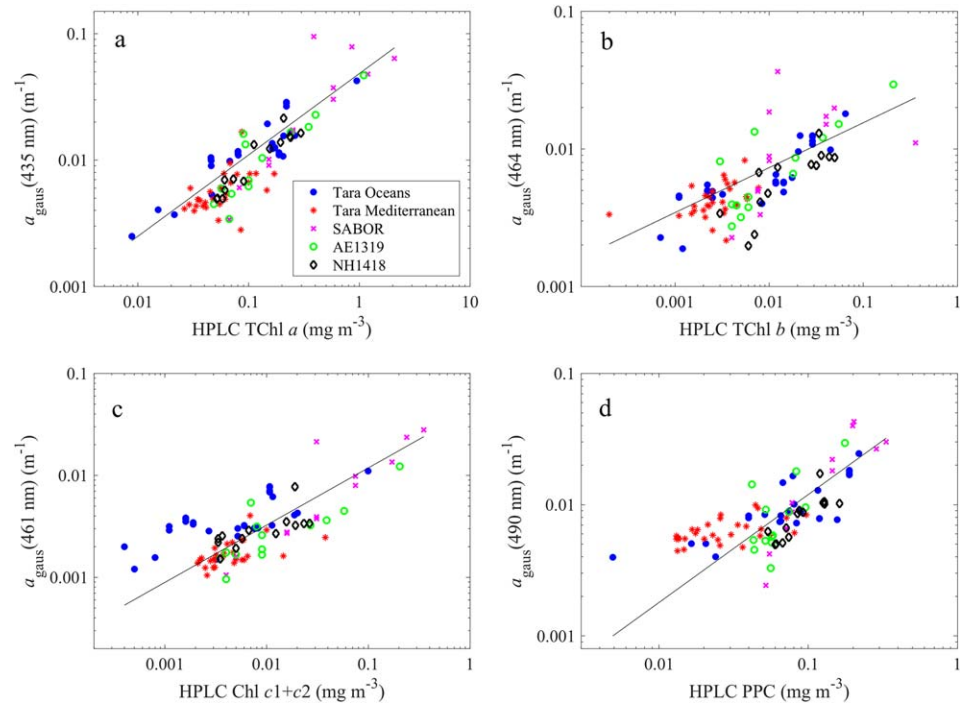


Figure 4. Comparisons between HPLC pigment concentrations (x-axes) and $a_{\text{gaus}}(\lambda)$ magnitudes (y-axes) for five different field campaigns ($n = 97$). The solid black lines show the type 1 nonlinear fits to the data; see Table 4 for coefficients of the best fit lines. Locations of each field campaign are shown in Figure 1a.

Table 4
Coefficients of $a_{\text{gaus}}(\lambda)$ /HPLC Relationships, and HPLC Covariation Relationships^a

Wavelength (nm)	HPLC pigment(s)	From $R_{\text{rs}}(\lambda)$ inversion to $a_{\text{gaus}}(\lambda)$		From global pigment covariation	
		$A_{\text{inv}} (\text{m}^{-1})$	B_{inv}	$A_{\text{cov}} (\text{mg m}^{-3})$	B_{cov}
435	TChl <i>a</i>	0.048 ± 0.008	0.643 ± 0.068		
461	Chl $c_1 + c_2$	0.043 ± 0.009	0.561 ± 0.059	6.27 ± 1.08	0.81 ± 0.02
464	TChl <i>b</i>	0.033 ± 0.013	0.327 ± 0.074	5.44 ± 1.14	0.86 ± 0.04
490	PPC ^b	0.079 ± 0.024	0.823 ± 0.105	11.10 ± 1.16	1.44 ± 0.06

^a A_{inv} and B_{inv} calculated using equation (17); A_{cov} and B_{cov} calculated using equation (20). Uncertainties were derived using bootstrapping and iterative fitting methods; see section 2.4.3. The relationships of pigment covariation compare TChl *a* to HPLC accessory pigments; thus there are no A_{cov} and B_{cov} values for TChl *a*. Note that the wavelengths are only relevant to the inversion analysis, and that B_{inv} and B_{cov} are exponents and therefore unitless.

$$[\text{pigment}]_{\text{inv}} = \left(\frac{a_{\text{gaus}}}{A_{\text{inv}}} \right)^{\frac{1}{B_{\text{inv}}}}. \quad (18)$$

The estimated pigment concentrations are used to calculate the median percent error (ME):

$$ME = \left(\text{median} \left(\frac{\text{abs}([\text{pigment}]_{\text{inv}} - [\text{pigment}]_{\text{HPLC}})}{[\text{pigment}]_{\text{HPLC}}} \right) \right) * 100. \quad (19)$$

Note that our use of a type 1 nonlinear fit for the data assumes constant uncertainties for all values of $a_{\text{gaus}}(\lambda)$ and that HPLC pigment concentrations are error free. Additionally, the relationships determined here are specific to the units of the data used in this study; the A_{inv} values cannot be transferred to estimate pigments concentrations from absorption measurements provided in units other than m^{-1} without proper unit conversion. We also note that throughout this paper, we implicitly assume that any nonlinear coefficients within our reported equations and relationships are normalized, as it is not mathematically possible to exponentiate or compute the logarithm of a dimensionful value. For example, $[\text{pigment}]_{\text{HPLC}}$ in equation (16) must be normalized by 1 mg m^{-3} (the units of HPLC pigment data).

2.4.2. Pigments Estimated from TChl a_{OC4} Using HPLC Covariation Relationships

As a comparison to the inversion algorithm, pigment concentrations were also estimated using TChl a_{OC4} and covariation relationships between TChl *a* and accessory pigments from HPLC data. The covariation relationships were determined using an HPLC data set spanning various parts of the world ocean (Figures 1b and 5). To do this, we applied a type 2 linear least squares fit of the log-normalized HPLC data that is weighted by the uncertainties in the data (MATLAB script for the type 2 fit by E. T. Peltzer, MBARI, 2016, <http://www.mbari.org/index-of-downloadable-files/>). HPLC uncertainties were calculated using the average percent error for each pigment or pigment group (TChl *a*, TChl *b*, Chl $c_1 + c_2$, and PPC), which is determined by dividing the standard deviation from replicate HPLC samples by the associated measurement. The A_{cov} and B_{cov} coefficients for the covariation pigment relationships (Table 4) were determined from the fit between TChl *a* and each of the three accessory pigments ($[\text{pigment}]_{\text{HPLC}}$):

$$\log(\text{TChl } a) = \log A_{\text{cov}} + B_{\text{cov}} * \log([\text{pigment}]_{\text{HPLC}}), \quad (20)$$

which is equivalent to

$$\text{TChl } a = A_{\text{cov}} [\text{pigment}]_{\text{HPLC}}^{B_{\text{cov}}}. \quad (21)$$

We used the calculated A_{cov} and B_{cov} coefficients from the global HPLC pigment covariation to estimate TChl *b*, Chl $c_1 + c_2$, and PPC from TChl a_{OC4} :

$$[\text{pigment}]_{\text{cov}} = \left(\frac{\text{TChl } a_{\text{OC4}}}{A_{\text{cov}}} \right)^{\frac{1}{B_{\text{cov}}}}. \quad (22)$$

The $[\text{pigment}]_{\text{cov}}$ concentrations were then used to calculate median error (ME) values using equation (19), where $[\text{pigment}]_{\text{inv}}$ is replaced by $[\text{pigment}]_{\text{cov}}$.

2.4.3. Calculation of Uncertainties in A and B Coefficients and in Estimated Pigments

To estimate uncertainty values in the A_{inv} and B_{inv} coefficients, we used a bootstrapping method to iteratively subsample all 97 points (with replacement; $n = 10,000$ iterations) during the fitting between $a_{\text{gaus}}(\lambda)$ magnitudes and HPLC pigment concentrations ([pigment]_{HPLC}; equation (17)). For the A_{cov} and B_{cov} coefficients, the type 2 linear least squares fit described above in section 2.4.2 also uses an iterative method to find the best fit between TChl a from HPLC and the three accessory pigments ([pigment]_{HPLC}). The number of iterations in this case is determined by the fitting routine reaching a user-set convergence limit. Following these iterative methods for both sets of A and B coefficients, we calculated the standard deviation of all iterations for each coefficient (A_{inv} , B_{inv} , A_{cov} , and B_{cov}); results are provided as uncertainty values in Table 4.

We propagated the A and B uncertainties using a Monte Carlo approach during the calculation of [pigment]_{inv} and [pigment]_{cov} to produce uncertainties in estimated pigment values. We multiplied the uncertainty values for each A and B coefficient (Table 4) by a random number (10,000 simulations) pulled from a standard normal distribution. These values are then used to calculate [pigment]_{inv} and [pigment]_{cov}, and the middle 68th percentiles of all values are used to provide uncertainty estimates on the median [pigment]_{inv} and [pigment]_{cov} estimates (Figures 8 and 9).

2.5. Analysis of Spectral Residuals

2.5.1. Calculation of $a_{\text{p-global}}(\lambda)$

To investigate spectral residuals remaining after removal of a Chl a -based $a_{\text{p}}(\lambda)$ spectrum, we calculate an “average global” particulate absorption spectrum ($a_{\text{p-global}}(\lambda)$) as a function of Chl a concentration. We used a large database of $a_{\text{p}}(\lambda)$ spectra ($n = 96,929$; Table 1; Figure 1c) and their corresponding Chl a values to determine the relationship between Chl a and $a_{\text{p}}(\lambda)$ at each wavelength. The Chl a value for each $a_{\text{p}}(\lambda)$ spectrum was calculated using the ac-s line height method and coefficients described in Boss et al. (2013). This data set covers a range of oceanographic conditions and is well representative of the global distribution of chlorophyll values observed by satellite (e.g., Figure 3 in Morel et al. [2007]), with chlorophyll values centered on approximately 0.1 mg m^{-3} (Figure 1c inset).

For each wavelength of $a_{\text{p}}(\lambda)$, we regressed the magnitude of $a_{\text{p}}(\lambda)$ against the Chl a concentration from the line height calculation; see Figure 6 for an example using $a_{\text{p}}(440 \text{ nm})$. We tested both a linear fit of the log-transformed data as well as a nonlinear fit that includes normalization of each $a_{\text{p}}(\lambda)$ value by the associated uncertainty. A linear regression of the log of the variables implicitly assumes that relative uncertainties are constant, which gives more weight to smaller values. However, this is not consistent with what we know about spectrophotometry, where at low values there is an absolute uncertainty resulting from factors such as instrument resolution and calibration error. To account for uncertainties, we used a nonlinear fit that is normalized by the uncertainties in particulate absorption, denoted $a_{\text{pUNC}}(\lambda)$. The $a_{\text{pUNC}}(\lambda)$ values were determined by calculating the standard deviation associated with the binning of $a_{\text{p}}(\lambda)$ data to 1 km^2 resolution. We regressed the $a_{\text{p}}(\lambda)$ spectra against Chl a by minimizing the following cost function for each wavelength (400–700 nm; 2 nm resolution):

$$\chi^2 = \sum_{\lambda=400}^{700} \left(\frac{a_{\text{p}}(\lambda) - A_{\text{chl}}(\lambda) \text{Chl}^{B_{\text{chl}}(\lambda)}}{a_{\text{pUNC}}(\lambda)} \right)^2. \quad (23)$$

We determined the values of the A_{chl} and B_{chl} coefficients, as well as associated median error values, for each wavelength (Figure 7; supporting information Table S1) similarly to the analysis performed by Bricaud et al. (1995, 1998). We then calculated an $a_{\text{p-global}}(\lambda)$ spectrum that corresponds to each measured $R_{\text{rs}}(\lambda)$ by using the A_{chl} and B_{chl} coefficients and TChl a_{OC4} in the equation:

$$a_{\text{p-global}}(\lambda) = A_{\text{chl}}(\lambda) \text{TChl} a^{B_{\text{chl}}(\lambda)} \quad (24)$$

Note that as with the relationships derived in section 2.4, it is implicitly assumed that the exponentiated value TChl a is normalized by 1 mg m^{-3} ; with B_{chl} unitless and the units of A_{chl} equal to m^{-1} , the resulting $a_{\text{p-global}}(\lambda)$ will also have units of m^{-1} . We also tested the data for the influence of coastal processes and found no significant change in the fit after removing all data within 10 km of shore (5% of data). Values of median error (ME) for each wavelength were calculated similarly to equation (19) by comparing the measured $a_{\text{p}}(\lambda)$ and the $a_{\text{p-global}}(\lambda)$ spectra.

2.5.2. Calculation and Inversion of Spectral Residuals

Following the calculation of $a_{p\text{-global}}(\lambda)$, we repeated the inversion of $R_{rs}(\lambda)$ with the combined $a_{\text{NAP}}(\lambda)$ and $a_{\text{gaus}}(\lambda)$ functions in equation (10) replaced by $a_{p\text{-global}}(\lambda)$ to represent the shape of particulate absorption. During the inversion of each spectrum, we set the initial guess for the magnitude of $a_p(\lambda)$ to the magnitude of $a_{p\text{-global}}(\lambda)$ at 440 nm. The resulting $u_{\text{mod}}(\lambda)$ (defined in equation (10)) is differenced with $u_{\text{meas}}(\lambda)$ to calculate a residual spectrum, $u_{\text{resid}}(\lambda)$:

$$u_{\text{resid}}(\lambda) = u_{\text{meas}}(\lambda) - u_{\text{mod}}(\lambda). \quad (25)$$

The inversion described in section 2.3 was repeated for each $u_{\text{resid}}(\lambda)$ spectrum ($n = 97$) and the results were used to compare the $a_{\text{gaus}}(\lambda)$ magnitudes with HPLC pigment data, similarly to the analysis described in section 2.4. Although the use of $a_{p\text{-global}}(\lambda)$ to calculate $u_{\text{resid}}(\lambda)$ removes an average Chl-based absorption spectrum, the residual absorption information that should be contained in $u_{\text{resid}}(\lambda)$ can include the influence of Chl a as well as any of the accessory pigments, if their relationship to Chl a deviates from the global average.

3. Results

3.1. Pigments Estimated from $a_{\text{gaus}}(\lambda)$ and Covariation Relationships

Of the eight $a_{\text{gaus}}(\lambda)$ functions used in the inversion, the results of four are presented: $a_{\text{gaus}}(435)$, $a_{\text{gaus}}(461)$, $a_{\text{gaus}}(464)$, and $a_{\text{gaus}}(490)$. The $a_{\text{gaus}}(384)$ function influences the inversion via the right-hand tail of the Gaussian function that has values ≥ 400 nm, but because we evaluate the inversion from 400 to 600 nm only, we do not compare $a_{\text{gaus}}(384)$ with HPLC pigment data. The remaining three $a_{\text{gaus}}(\lambda)$ functions ($a_{\text{gaus}}(413)$, $a_{\text{gaus}}(532)$, and $a_{\text{gaus}}(583)$) did not show significant correlations with any of the pigments or pigment groups we investigated. The $a_{\text{gaus}}(413)$, $a_{\text{gaus}}(532)$, and $a_{\text{gaus}}(583)$ functions represent absorption by Tchl a , PSC, and Chl $c_1 + c_2$, respectively. Of these, we can estimate Tchl a and Chl $c_1 + c_2$ using $a_{\text{gaus}}(435)$ and $a_{\text{gaus}}(461)$, as described above. Using our inversion method, estimates of PSC concentrations are unable to be retrieved due to lack of significant correlation between PSC pigments from HPLC and $a_{\text{gaus}}(532)$ magnitudes.

We found positive and significant correlations between HPLC pigment concentrations and [pigment]_{inv} that range from 0.65 to 0.87 (Spearman's rank correlation coefficient, ρ) and 0.25 to 0.86 (Pearson's linear correlation coefficient, r^2). Median relative errors in prediction of Tchl a , Tchl b , Chl $c_1 + c_2$, and PPC using [pigment]_{inv} range from 36% to 65% (Table 5; Figures 8 and 9). PPC is estimated with the least error, followed by Tchl a , Chl $c_1 + c_2$, and Tchl b . Similar positive and significant correlations were found between pigments estimated from Tchl a_{OC4} using covariation relationships ([pigment]_{cov}) and HPLC pigment concentrations. The correlation coefficients range from 0.67 to 0.89 (ρ) and 0.51 to 0.96 (r^2). We found increased median error when predicting PPC (52%), the same error for Tchl a (37%), and lower median errors for Chl $c_1 + c_2$ and Tchl b (40% and 56%, respectively) for [pigment]_{cov} versus [pigment]_{inv} (Table 5; Figures 8 and 9).

The increased spread in low values of [pigment]_{inv} for all three accessory pigments (Figure 8) is also reflected in the spread in the data used to calculate A_{inv} and B_{inv} (Figure 4). Chl $c_1 + c_2$ is predicted with the lowest error when using [pigment]_{cov}, which is also reflected in the low amount of spread around the best-fit line in the data used to calculate A_{cov} and B_{cov} (Figure 5b). Several outlying points from the Tara Oceans Expedition (Figure 4c) are likely the cause of increased errors in prediction of Chl $c_1 + c_2$ using [pigment]_{inv} relative to [pigment]_{cov}. Uncertainties in the estimated [pigment]_{inv} and [pigment]_{cov} values as shown by error bars in Figures 8 and 9 are relatively small in all cases except for Tchl b estimated by [pigment]_{inv}. This is also reflected in the high ME values for Tchl b from [pigment]_{inv} (Table 5).

3.2. Pigment Ratios

As a test of consistency with previous studies, we compared the distributions of pigment ratios determined using HPLC pigments, the [pigment]_{inv} concentrations from $a_{\text{gaus}}(\lambda)$ magnitudes, and the

Table 5
Statistics of Pigment Estimation^a

Wavelength (nm)	HPLC pigment(s)	ρ	r^2	ME (%)
		from $a_{\text{gaus}}(\lambda)$ magnitudes ([pigment] _{inv})		
435	Tchl a	0.87	0.69	37
461	Chl $c_1 + c_2$	0.65	0.86	48
464	Tchl b	0.75	0.25	65
490	PPC ^b	0.75	0.77	36
		from Tchl a_{OC4} ([pigment] _{cov})		
	Tchl a	0.89	0.83	37
	Chl $c_1 + c_2$	0.84	0.96	40
	Tchl b	0.67	0.51	56
	PPC ^b	0.67	0.70	52

^aCorrelations are Spearman's rank correlation coefficient (ρ) and Pearson's linear correlation coefficient (r^2). ME = median error (equation (19)). Wavelengths are only relevant to inversion analysis where $a_{\text{gaus}}(\lambda)$ magnitudes are used to estimate pigment concentrations. Note that the statistics comparing Tchl a with Tchl a_{OC4} are a direct comparison; no covariation relationships are used.

^bPPC = $\alpha\beta$ -carotene + zeaxanthin + alloxanthin + diadinoxanthin.

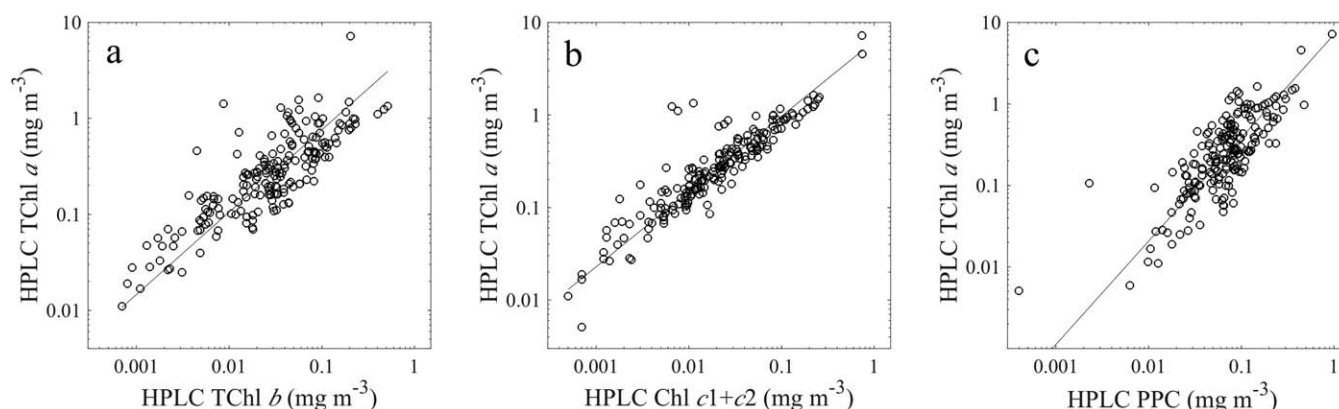


Figure 5. Relationships between (a) TChl *a* and TChl *b*, (b) TChl *a* and Chl *c*₁+*c*₂, and (c) TChl *a* and PPC from HPLC pigment concentrations for the global HPLC data set (*n* = 196). The locations of all data points are shown in Figure 1b. The coefficients of the best-fit lines are shown in Table 4.

[pigment_{cov}] concentrations from TChl *a*_{OC4} using global covariation relationships (Figure 10). The distribution of TChl *b*:TChl *a* from TChl *a*_{OC4} ([pigment_{cov}]) are shifted slightly higher than TChl *b*:TChl *a* from the other two data sets (Figure 10a). The distributions of Chl *c*₁+*c*₂:TChl *a* from HPLC and both [pigment_{inv}] and [pigment_{cov}] follow a similar pattern (Figure 10b). The mode values of the PPC:TChl *a* distribution from both [pigment_{inv}] and [pigment_{cov}] are slightly higher than that from the HPLC pigment data (Figure 10c). The distributions of all values are within the ranges of previously published values (see Table 3 in Chase et al., 2013).

3.3. Spectral Residuals

The values of spectral *A*_{chl} and *B*_{chl} are shown in Figures 7a and 7b and provided in supporting information Table S1 (Table S1 can be found in supporting information). Deviations from the values reported in Bricaud et al. (1998) are likely the result of both an increased number of data points at the lower end of the Chl *a* value range in our data set, and the nonlinear fit used in our calculation of *A*_{chl} and *B*_{chl}. Note that the difference in scattering correction of the *a*_p(*λ*) data (our nonlinear correction versus a spectrally constant scattering correction by Bricaud et al. 1998) cannot explain the difference as ours is expected to remove more scattering out of absorption in the blue wavelengths, which would preferentially decrease the *a*_p(*λ*) values in that region of the spectrum. The relative error values are highest in the region of the spectrum where absorption is generally the lowest (Figure 7c).

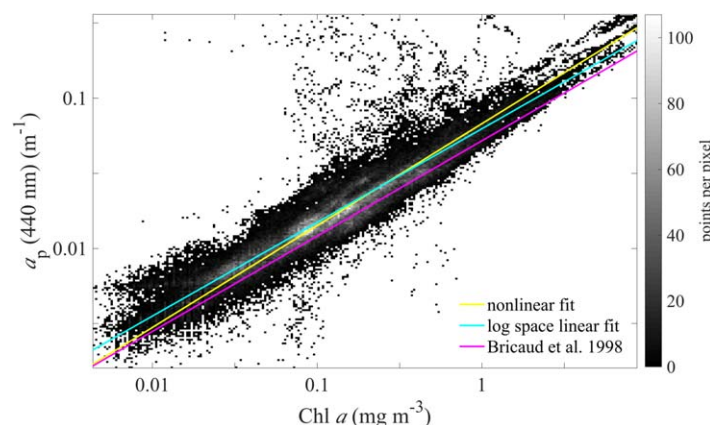


Figure 6. Chl *a* estimated using *a*_p(676 nm) line height regressed against in situ-measured *a*_p(440 nm). Both the linear fit in log space and the nonlinear fit that also accounts for uncertainties are shown, as well as the fit from Bricaud et al. (1998). The 99.9th percentile of all data are shown for ease of viewing. The gray scale bar represents the number of data points within each 1-km² after spatial binning. The outlying points above the best-fit lines are likely caused by a large influence of *a*_{NAP}(440) to *a*_p(440).

The inversion of *u*_{resid}(*λ*) spectra (defined in section 2.5.2) was conducted to allow for the extraction of pigment information contained in any spectral features in *a*_p(*λ*) that deviate from *a*_{p-global}(*λ*). We found no significant correlations between the *a*_{gaus}(*λ*) magnitudes following the inversion of *u*_{resid}(*λ*) spectra and HPLC pigment concentrations (not shown). In the inversion analysis, *a*_{CDOM}(*λ*) and *b*_b(*λ*) are allowed to vary during the inversion as they were not measured in situ and so are not known *a priori*. Because of this, during the inversion analysis of *u*_{resid}(*λ*) the inverted *a*_{CDOM}(*λ*) and *b*_b(*λ*) functions compensated for the differences between the measured *a*_p(*λ*) and *a*_{p-global}(*λ*) spectra, which prevented any extraction of pigment information beyond the average global relationships between TChl *a* and accessory pigments.

4. Discussion

Changes in phytoplankton community composition are known to be correlated with changes in Chl *a*, as smaller cells are linked to regenerated production and dominating in oligotrophic waters (Chl *a* < 0.1 mg m⁻³), and larger cells (namely diatoms) are associated with the “new” production of higher biomass regions (Chisholm, 1992;

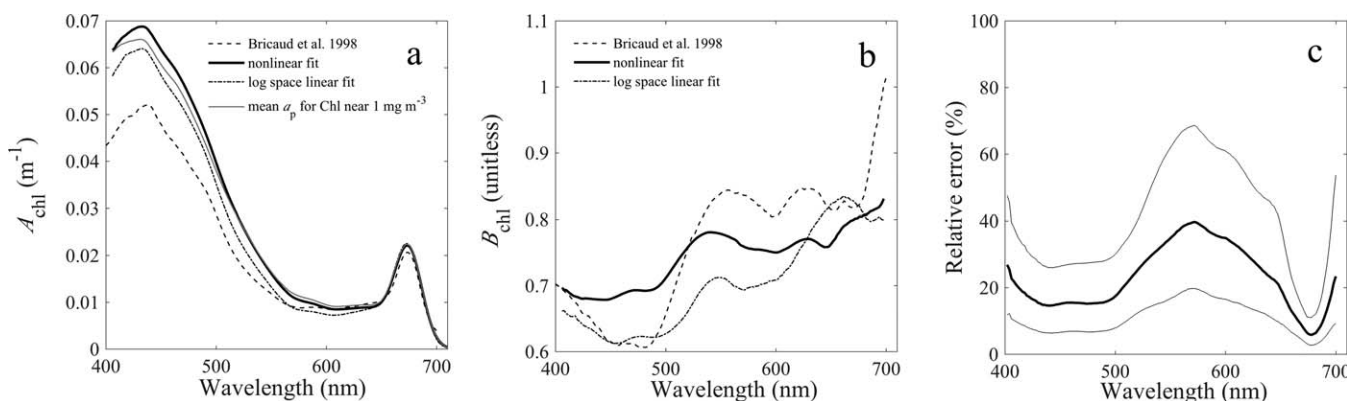


Figure 7. Spectral (a) A_{chl} and (b) B_{chl} derived from the relationship: $a_p(\lambda) = A_{chl} * Chl * a^{B_{chl}}$. Values of A_{chl} and B_{chl} are provided in supporting information Table S1. Comparisons between the nonlinear and linear fits are shown, as well as values from Bricaud et al. (1998). Figure 7a shows the $a_p(\lambda)$ spectrum calculated with all values of Chl < 1.05 and $> 0.95 \text{ mg m}^{-3}$ for comparison with the nonlinear fit, which represents the spectrum for Chl = 1.0 mg m^{-3} . (c) Spectral values of the relative error associated with the use of the A_{chl} and B_{chl} coefficients for prediction of $a_p(\lambda)$ from chlorophyll. Thick black line shows the median value (reported as ME in supporting information Table S1) and the bottom and top lines are the 25th and 75th percentiles of the error values, respectively.

Eppey & Peterson, 1979). Surface Chl a values have also been related to phytoplankton size classes determined by HPLC diagnostic pigments (Ras et al., 2008; Uitz et al., 2006). Understanding the relationships between hyperspectral reflectance measurements and underlying properties is critical for developing algorithms that can be used to study phytoplankton community composition on a global scale. In the work presented here, we explored the limits of using global, hyperspectral $R_{rs}(\lambda)$ data to estimate phytoplankton pigments, which in turn can be used to help define phytoplankton composition. We also used global covariation relationships between TChl a and HPLC pigments to estimate accessory pigment concentrations from $R_{rs}(\lambda)$ -derived Chl a ; these estimates provide a bench-mark against which any other method for accessory pigment estimation should be tested.

4.1. Estimation of Pigment Concentrations

In the comparison of our results to previous studies, we consider both the median relative errors resulting from the spectral inversion algorithm and from the pigment covariation method (Table 5). The median relative errors of estimated pigments from our study are higher than those reported in Wang et al. (2016), but several key differences are noted. First, their study compared only $a_{gaus}(\lambda)$ magnitudes from the decomposition of $a_p(\lambda)$ and $R_{rs}(\lambda)$ spectra (i.e., no HPLC data were available), and therefore the method was not

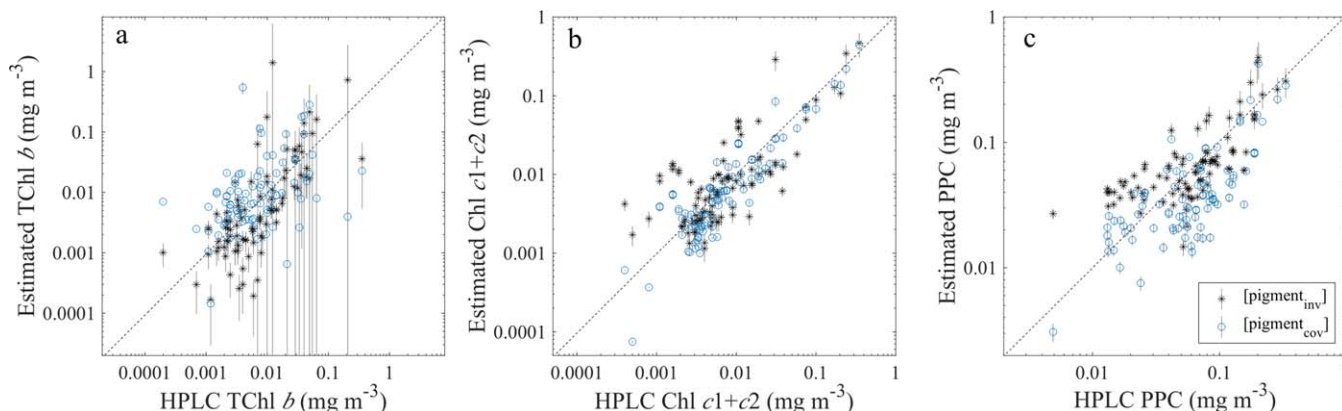


Figure 8. Comparisons between HPLC pigment concentrations (x-axes) and pigments estimated both from $a_{gaus}(\lambda)$ ([pigment_{inv}], black “*”) and TChl a_{OC4} ([pigment_{cov}], blue open circles) (y-axes). Dotted lines show one-to-one relationships; $n = 97$. Vertical lines show the first standard deviation of estimated pigments; see section 2.4.3 for more details.

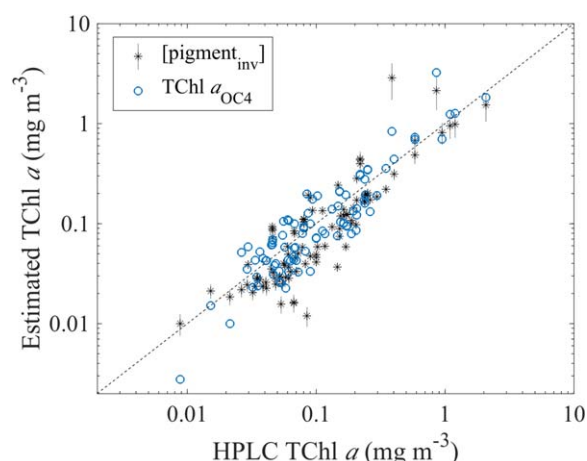


Figure 9. Comparison of TChl a computed from 1) $a_{\text{gaus}}(435)$ ([$\text{pigment}_{\text{inv}}$], black “*”), and 2) $R_{\text{rs}}(\lambda)$ using the OC4 algorithm (TChl a_{OC4} , blue open circles) with HPLC TChl a . Dotted line shows the one-to-one relationship; $n = 97$. Vertical lines on the [$\text{pigment}_{\text{inv}}$] values show the first standard deviation of estimated pigments; see section 2.4.3 for more details. Uncertainties are not provided for the TChl a_{OC4} values as we do not have information on the uncertainty of parameters used in the OC4 algorithm.

evaluated using measured pigment concentrations. Second, their data set consisted of measurements made in fresh inland waters with Chl a values ranging from 10 to 1,377 $\mu\text{g L}^{-1}$. Additionally, the high reflectance values in the red region (Figure 4 in Wang et al., 2016) allow their use of the entire visible spectrum for the inversion analysis; this is not possible with open ocean $R_{\text{rs}}(\lambda)$ spectra that have a much lower signal above 600 nm (e.g., Figure 2, this study). When compared with our study, median errors in pigments estimated by Bracher et al. (2015) are lower for estimation of TChl a , Chl $c_1 + c_2$, and PPC. The study by Bracher et al. (2015) did not include TChl b , as a result of a high percentage of samples with no detected TChl b (24% and 34% for satellite and field samples, respectively; see supplement of Bracher et al. [2015]). The mean absolute percent error values found by Pan et al. (2010) are slightly lower for TChl a , TChl b , and TChl c (which includes Chl c_3), and higher for PPC (in their study each PPC pigment is estimated individually). Both the Pan et al. (2010) and Bracher et al. (2015) studies were regionally adapted for the United States northeast coast and the Atlantic Ocean, respectively. Although the median error values we calculated were slightly higher than the values from these two previous studies, our method is theoretically applicable at a global scale.

4.2. Pigment Covariation

The study by Pan et al. (2010) showed high correlation coefficients between TChl a and TChl c , and TChl a and each of the PPC pigments used in our study except for zeaxanthin. The global covariation between Chl a and phytoplankton accessory pigments was also observed previously by Trees et al. (2000), with a similar slope value (0.93) of the log-linear relationship between Chl a and total summed accessory pigments (both in mg m^{-3}) to what we observed in the global HPLC data (column of B_{cov} values in Table 4; mean slope value of 1.04). A previous study by Uitz et al. (2015) found strong correlations between the first and second EOF mode (describing a combined 95% of the variance) of $R_{\text{rs}}(\lambda)$ spectra and several pigments (Chl a , zeaxanthin, 19'-hexanoyloxyfucoxanthin, and fucoxanthin). Bracher et al. (2015) showed that the second EOF is the most important for the prediction of several pigments and pigment groups: TChl a , monovinyl Chl a (MVChl a), PSC, and PPC. In the case of both of these previous studies, the prediction of Chl a and other pigments from the same EOF modes further supports the covariation of Chl a and accessory pigments in ocean waters.

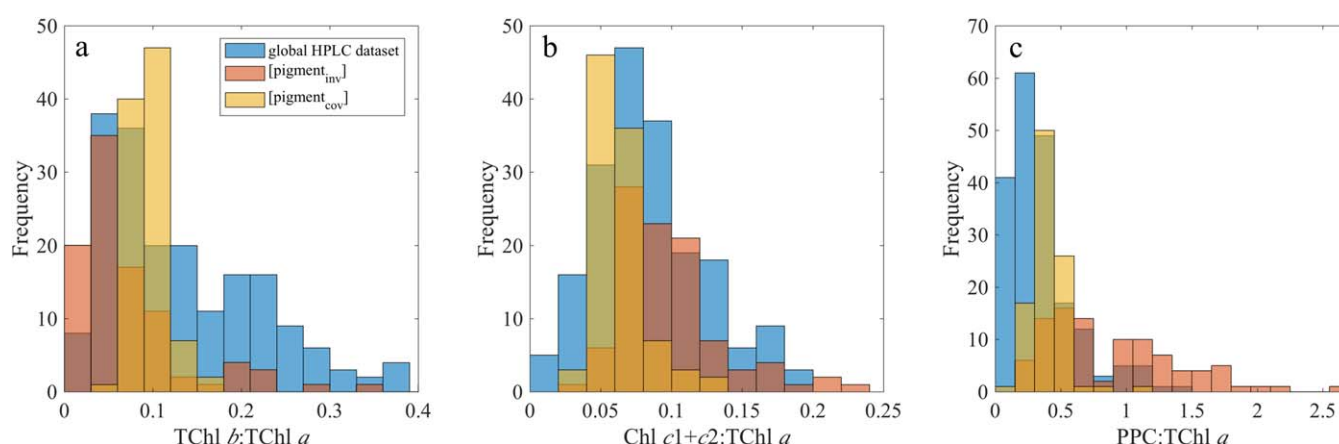


Figure 10. Frequency distributions of (a) TChl b :TChl a , (b) Chl $c_1 + c_2$:TChl a , and (c) PPC:TChl a determined from the global HPLC data set ($n = 196$, blue), the pigment concentrations estimated from $a_{\text{gaus}}(\lambda)$ ([$\text{pigment}_{\text{inv}}$]; $n = 97$, orange), and the pigment concentrations estimated from TChl a_{OC4} ([$\text{pigment}_{\text{cov}}$]; $n = 97$, yellow). The details of all data sets are presented in Table 1. Three high [$\text{pigment}_{\text{inv}}$] data points are excluded for ease of viewing, with ratio values of 0.48 and 0.75 in Figure 10a and 4.2 in Figure 10c.

Our previous work to estimate pigments from $a_p(\lambda)$ spectra (Chase et al., 2013) did not examine the use of pigment covariation for estimating accessory pigments; however, we can compare the results of that study with the statistics generated here for estimated pigments. There are similar or slightly lower errors using $a_{\text{gaus}}(\lambda)$ from decomposition of ac-s spectra as reported in Chase et al. (2013) compared to the median errors in the present work, with the exception of PPC, which is estimated with lower error using the inversion algorithm in the current study. The lower errors in estimating chlorophylls a , b , and c directly from $a_p(\lambda)$ is likely made possible by the distinct absorption peak positions of the chlorophylls compared to other pigments in the red part of the visible spectrum, where there is sufficient information in $a_p(\lambda)$ but not $R_{rs}(\lambda)$ spectra. One source of uncertainty in pigment covariation relationships is the spread around the best-fit line that results from natural variations in phytoplankton pigment composition. In addition, pigment covariation relationships are limited by the accuracy of the TChl a that is subsequently used to estimate the other accessory pigments. Uncertainties in the estimation of TChl a are variable depending on the method used (e.g., HPLC analysis, spectral band ratio, ac-s line height) and should be taken into consideration when using TChl a and pigment covariation parameters to estimate accessory pigments.

4.3. Spectral Resolution Considerations

Our use of spectral inversion and Gaussian functions to estimate phytoplankton pigments from hyperspectral $R_{rs}(\lambda)$ is an analytical approach that attempts to exploit any influence of pigment absorption on the spectral shape of $R_{rs}(\lambda)$. One appeal of such an approach is that it does not require training data sets or any previous knowledge of the water of interest, thereby making it more robust in an ocean experiencing large scale climate-driven changes. In addition, it eliminates the need to define the shape of $a_{\varphi}(\lambda)$. However, there are still assumptions made when defining the shapes of $a_{\text{NAP}}(\lambda)$ and $a_{\text{CDOM}}(\lambda)$ spectra used in the inversion algorithm. This could present a problem if, for example, $a_{\text{NAP}}(\lambda)$ spectra are not exactly exponential, which has been seen in previous studies (Babin & Stramski, 2004; Estapa et al., 2012; Iturriaga & Siegel, 1989). Uncertainties in the prescribed spectral shapes, as well as in the Raman correction (section 2.2.1), are potential sources of error in the inversion and subsequent magnitudes of $a_{\text{gaus}}(\lambda)$ spectra.

The peak locations of the functions representing TChl b and Chl $c_1 + c_2$ absorption in our analysis (464 nm and 461 nm, respectively) are located only three nanometers apart; however, the defined widths of the two Gaussian functions are different (Table 3). Both the peak locations and widths are based on laboratory-measured pigment absorption (Bricaud et al., 2004). The method of spectral decomposition into multiple Gaussian functions uses both the peak location and width, which can allow for the separation of absorption features that are close spectrally. Additionally, hyperspectral data provide the capability for inversion methods using Gaussian decomposition, while multispectral data do. Previous studies have used spectral derivative methods (Lee et al., 2007; Wolanin et al., 2016) to determine the optimal wavelength locations needed to maximize the information extracted from $R_{rs}(\lambda)$ spectra. Wolanin et al. (2016) concluded that hyperspectral data is the most effective for discriminating between three phytoplankton groups (diatoms, coccolithophores, and cyanobacteria). The derivative analysis by Lee et al. (2007) showed that a multispectral approach could be used to identify the location of local extremes or inflections. However, an inversion method using Gaussian decomposition can be used to estimate pigments that may be influencing the reflectance signal without creating local features that can be enhanced by derivative analysis.

4.4. Spectral Residuals Analysis

We suspect that factors other than phytoplankton accessory pigments are driving the deviation of $a_p(\lambda)$ spectra from the $a_{p\text{-global}}(\lambda)$ expected for a given Chl a value. For example, as mentioned above, $a_{\text{NAP}}(\lambda)$ spectra may not be exactly exponential. Reflectance anomalies calculated using multispectral satellite data by Huot and Antoine (2016) showed that several $R_{rs}(\lambda)$ anomalies were well correlated with either $a_{\text{CDOM}}(\lambda)$ anomalies or $b_{\text{bp}}(\lambda)$ estimates. However, some of these anomalies (namely those determined using the ratio of $R_{rs}(488 \text{ nm})$ to other bands) were not well correlated with any other satellite ocean color products, but rather are likely due to some combined variation of $a_{\text{CDOM}}(\lambda)$, $a_{\varphi}(\lambda)$, and/or $b_{\text{bp}}(\lambda)$. Of these three, $a_{\varphi}(\lambda)$ contains the most spectral variability, due to the presence of different accessory pigments. This supports the possibility of using hyperspectral reflectance residuals to

detect phytoplankton accessory pigment signatures; however, more work is needed to understand how to best use the residuals of hyperspectral $R_{rs}(\lambda)$ data for obtaining information on phytoplankton community composition.

5. Conclusions

Direct estimation of phytoplankton pigments from $R_{rs}(\lambda)$ spectra has the appeal of application to a wide range of ocean waters; this is especially desirable for use with satellite remote-sensing data that covers much of the globe. Here, we have estimated phytoplankton pigments from in situ hyperspectral $R_{rs}(\lambda)$ data that spans various oceanic water types. We determined that several phytoplankton accessory pigments can be predicted via an inversion algorithm with median errors ranging from 36% to 65%. Additionally, we calculated similar predictive capabilities for pigments when a global relationship of pigment covariation is applied to TChl a concentrations derived from $R_{rs}(\lambda)$ spectra. The median error values provided in this study are a crucial component in efforts to estimate phytoplankton groups from optical and ultimately satellite data, and a lack of uncertainty information has been an area identified as a gap in the current research (Bracher et al., 2017). However, we note the similar values in the predictive capabilities of the two methods, and that there will be variability in the predicted pigment uncertainties due to factors such as the algorithm used to estimate Chl a from $R_{rs}(\lambda)$. The high covariation of accessory pigments with TChl a , on a global scale, should be carefully considered when using algorithms that are designed to estimate phytoplankton accessory pigments from $R_{rs}(\lambda)$ spectra; we propose that covariation relationships should be used as a bench-mark against which to evaluate the utility of novel methods. Although changes in phytoplankton biomass may be correlated with phytoplankton community composition, extracting knowledge on community composition beyond information based on TChl a and its covarying parameters is not trivial. While the analysis of spectral residuals reported here did not show the utility of residuals from hyperspectral $R_{rs}(\lambda)$ to estimate pigment concentrations, this approach should be explored further as one way to move beyond the information provided from the covariation of accessory pigments with TChl a . Finally, automated phytoplankton imagery and molecular data are rapidly becoming more available and will be important for comparison with both HPLC data and optical spectra. Future research should include analysis and validation work to compare HPLC pigment information with phytoplankton community composition from other data such as molecular information or imagery, thus connecting the estimation of pigments with phytoplankton community in a framework with known uncertainties.

Table A1

Statistics Calculated from the Comparison of the OC Algorithms with HPLC TChl a^a

Chl a algorithm	ρ	r^2	ME (%)	MeanE (%)	MBE (%)	RMSE
OC4	0.86	0.8	35	39	4	0.25
OCI	0.89	0.8	39	53	43	0.25

Note. Parameters shown are: ρ = Spearman's rank correlation coefficient, r^2 = Pearson's linear correlation coefficient, ME = median percent error (see equation (19)), MeanE = same as ME but mean instead of median, MBE = mean bias error (same as ME but without taking the absolute value), RMSE = root mean square error.

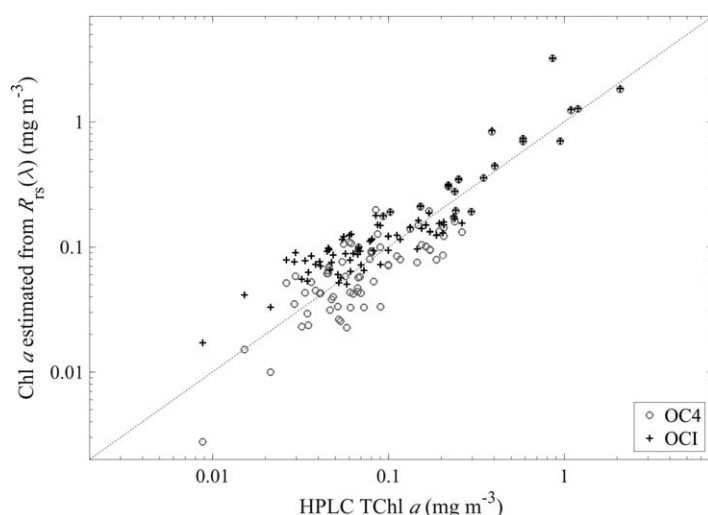


Figure A1. Chl a estimated from $R_{rs}(\lambda)$ spectra using OC4 (grey open circles) and OCI (black plus symbols) compared to TChl a from HPLC analysis. $R_{rs}(\lambda)$ spectra are from the same five data sets used in the inversion analysis ($n = 97$; Table 1; Figure 2).

Appendix A: Comparison of Chlorophyll Algorithms

We tested both NASA's current OCI algorithm (https://oceancolor.gsfc.nasa.gov/atbd/chlor_a/) and the standard OC4 algorithm (O'Reilly et al., 2000) when calculating the $R_{rs}(\lambda)$ -derived estimates of TChl a . The OCI algorithm is a combination of the standard OC4 band ratio algorithm combined with the Color Index (CI) from Hu et al. (2012). At Chl a values above 0.2 mg m^{-3} OC4 is used, at values below 0.15 mg m^{-3} CI is used, and a weighted combination of the two algorithms is used for values in between. We found that the comparison between $R_{rs}(\lambda)$ -derived TChl a and HPLC TChl a was slightly improved when using OC4 (Figure A1 and Table A1); as a result, we opted to use the OC4 algorithm through the study. Note that we also tested the use of Chl a versus TChl a and the statistics and results are similar.

Acknowledgments

We thank the Tara Foundation and the crew and scientists of the R/V Tara for data collection and processing (for data in this paper labeled as Tara Oceans and Tara Mediterranean), as well as the crews and scientists of the SABOR (R/V Endeavor), NH1418 (R/V New Horizon), and AE1319 (R/V Atlantic Explorer) cruises. Thank you to Annick Bricaud for providing the data for comparison of the $\text{Chl} - a_p(\lambda)$ relationship. Thanks to Collin Roesler for helpful discussions regarding interpretation of our results. Data used throughout the manuscript can be found at the following locations: <https://doi.pangaea.de/10.1594/PANGAEA.836318> and <https://doi.pangaea.de/10.1594/PANGAEA.836319> (Tara Oceans); [https://seabass.gsfc.nasa.gov/\(Tara Mediterranean, SABOR, AE1319, and NH1418\)](https://seabass.gsfc.nasa.gov/(Tara%20Mediterranean,%20SABOR,%20AE1319,%20and%20NH1418)). A NASA Earth and Space Science Fellowship (NESSF) funded this research as well as NASA Ocean Biology and Biogeochemistry program grants NNX15AC08G and NNX13AC42G. Finally, the authors wish to thank Astrid Bracher and an anonymous reviewer for their comments and feedback, which helped us to greatly improve the manuscript.

References

- Alvain, S., Loisel, H., & Dessailly, D. (2012). Theoretical analysis of ocean color radiances anomalies and implications for phytoplankton groups detection in Case 1 waters. *Optics Express*, 20(2), 1070–1083.
- Alvain, S., Moulin, C., Dandonneau, Y., & Bréon, F. M. (2005). Remote sensing of phytoplankton groups in Case 1 waters from Global SeaWiFS Imagery. *Deep-Sea Research Part I: Oceanographic Research Papers*, 52(2005), 1984–2004.
- Alvain, S., Moulin, C., Dandonneau, Y., & Loisel, H. (2008). Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: A satellite view. *Global Biogeochemical Cycles*, 22, GB3001.
- Babin, M., & Stramski, D. (2004). Variations in the mass-specific absorption coefficient of mineral particles suspended in water. *Limnology and Oceanography*, 49, 756–767.
- Ben Mustapha, Z., Alvain, S., Jamet, C., Loisel, H., & Dessailly, D. (2013). Automatic classification of water-leaving radiance anomalies from global SeaWiFS imagery: Application to the detection of phytoplankton groups in open ocean waters. *Remote Sensing of Environment*, 146, 97–112.
- Boss, E., Picheral, M., Leeuw, T., Chase, A., Karsenti, E., Gorsky, G., . . . Claustre, H. (2013). The characteristics of particulate absorption, scattering and attenuation coefficients in the surface ocean; contribution of the Tara oceans expedition. *Methods in Oceanography*, 7, 52–62.
- Boss, E., Picheral, M., Slade, W., Taylor, L., Leeuw, T., & Chase, A., & Tara oceans consortium, coordinators; Tara Oceans Expedition, Participants (2014). *Properties of seawater and particulate matter from a WETLabs AC-S spectrophotometer and a WETLabs chlorophyll fluorometer mounted on the continuous surface water sampling system during the Tara Oceans expedition 2009–2013*. PANGAEA. <https://doi.org/10.1594/PANGAEA.836318>
- Bracher, A., Bouman, H. A., Brewin, R. J., Bricaud, A., Ciotti, A. M., Clementson, L., . . . Wolanin, A. (2017). Obtaining phytoplankton diversity from ocean color: A scientific roadmap for future development. *Frontiers in Marine Science*, 4(March), 1–15. <https://doi.org/10.3389/fmars.2017.00055>
- Bracher, A., Taylor, M. H., Taylor, B., Dinter, T., Röttgers, R., & Steinmetz, F. (2015). Using empirical orthogonal functions derived from remote-sensing reflectance for the prediction of phytoplankton pigment concentrations. *Ocean Science*, 11(1), 139–158. <https://doi.org/10.5194/os-11-139-2015>
- Bricaud, A., Babin, M., Morel, A., & Claustre, H. (1995). Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: Analysis and parameterization. *Journal of Geophysical Research*, 100(C7), 13321–13332.
- Bricaud, A., Claustre, H., Ras, J., & Oubelkheir, K. (2004). Natural variability of phytoplanktonic absorption in oceanic waters: Influence of the size structure of algal populations. *Journal of Geophysical Research*, 109, C11010. <https://doi.org/10.1029/2004JC002419>
- Bricaud, A., Morel, A., Babin, M., Allali, K., & Claustre, H. (1998). Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (Case 1) waters: Analysis and implications for bio-optical models. *Journal of Geophysical Research*, 103(C13), 31033–31044.
- Brown, C. A., Huot, Y., Werdell, P. J., Gentili, B., & Claustre, H. (2008). The origin and global distribution of second order variability in satellite ocean color and its potential applications to algorithm development. *Remote Sensing of Environment*, 112(12), 4186–4203. <https://doi.org/10.1016/j.rse.2008.06.008>
- Chase, A., Boss, E., Zaneveld, R., Bricaud, A., Claustre, H., Ras, J., . . . Westberry, T. K. (2013). Decomposition of in situ particulate absorption spectra. *Methods in Oceanography*, 7, 110–124.
- Chisholm, S. W. (1992). Phytoplankton size. In P. G. Falkowski, A. D. Woodhead, & K. Virvito (Eds.), *Primary productivity and biogeochemical cycles in the sea* (Vol. 43). Boston, MA: Environmental Science Research, Springer. https://doi.org/10.1007/978-1-4899-0762-2_12
- Eppley, R. W., & Peterson, B. J. (1979). Particulate organic matter flux and planktonic new production in the deep ocean. *Nature*, 282(5740), 677–680. <https://doi.org/10.1038/282677a0>
- Estapa, M. L., Boss, E., Mayer, L. M., & Roesler, C. S. (2012). Role of iron and organic carbon in mass-specific light absorption by particulate matter from Louisiana Coastal Waters. *Limnology and Oceanography*, 57(1), 97–112. <https://doi.org/10.4319/lo.2012.57.1.0097>
- Farikou, O., Sawadogo, S., Niang, A., & Diouf, D. (2015). Inferring the seasonal evolution of phytoplankton groups in the Senegalo-Mauritanian upwelling region from satellite ocean-color spectral measurements. *Journal of Geophysical Research Oceans*, 120, 6581–6601. <https://doi.org/10.1002/2015JC010738>
- Gordon, H. R., Brown, O. B., Evans, R. H., Brown, J. W., Smith, R. C., Baker, K. S., & Clark, D. K. (1988). A semianalytic radiance model of ocean color. *Journal of Geophysical Research*, 93(D9), 10909–10924.
- Gordon, H. R., Clark, D. K., Brown, J. W., Brown, O. B., Evans, R. H., & Broenkow, W. W. (1983). Phytoplankton pigment concentrations in the Middle Atlantic Bight: Comparison of ship determinations and CZCS estimates. *Applied Optics*, 22, 3929–3931.
- Gregg, W. W., & Carder, K. L. (1990). A simple spectral solar irradiance model for cloudless maritime atmospheres. *Limnology and Oceanography*, 35(8), 1657–1675.
- Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., . . . Gorsky, G. (2015). Plankton networks driving carbon export in the oligotrophic ocean. *Nature*, 532, 465–470.
- Hoepffner, N., & Sathyendranath, S. (1991). Effect of pigment composition on absorption properties of phytoplankton. *Marine Ecology Progress Series*, 73, 11–23.
- Hoepffner, N., & Sathyendranath, S. (1993). Determination of the major groups of phytoplankton pigments from the absorption spectra of total particulate matter. *Journal of Geophysical Research*, 98(C12), 22789–22803.
- Hooker, S. B., Heukelem, L. V., Thomas, C. S., Claustre, H., Ras, J., Schlüter, L., . . . Perl, J. (2009). *The Third SeaWiFS HPLC analysis round-robin experiment (SeaHARRE-3)*. Goddard, MD: National Aeronautics and Space Administrations.
- Hu, C., Lee, Z., & Franz, B. (2012). Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. *Journal of Geophysical Research*, 117, C01011. <https://doi.org/10.1029/2011JC007395>
- Huot, Y., & Antoine, D. (2016). Remote sensing reflectance anomalies in the ocean. *Remote Sensing of Environment*, 184, 101–111. <https://doi.org/10.1016/j.rse.2016.06.002>
- Isada, T., Hirawake, T., Kobayashi, T., Nosaka, Y., & Natsuike, M. (2015). Remote sensing of environment hyperspectral optical discrimination of phytoplankton community structure in Funka Bay and its implications for ocean color remote sensing of diatoms. *Remote Sensing of Environment*, 159, 134–151.
- Iturriaga, R., & Siegel, D. A. (1989). Microphotometric characterization of phytoplankton and detrital absorption properties in the Sargasso Sea. *Limnology and Oceanography*, 34, 1706–1726.
- Jeffrey, S. W., & Vesk, M. (1997). Introduction to marine phytoplankton and their pigment signatures. In S. W. Jeffrey, R. F. C. Mantoura, & S. W. Wright (Eds.), *Phytoplankton pigment in oceanography: Guidelines to modern methods* (pp. 33–84). Paris, France: UNESCO.
- Leathers, R., Downes, T. V., & Mobley, C. (2001). Self-shading correction for upwelling sea-surface radiance measurements made with buoyed instruments. *Optics Express*, 8(10), 561–570. <https://doi.org/10.1364/OE.8.000561>
- Lee, Z., Carder, K. L., & Arnone, R. A. (2002). Deriving inherent optical properties from water color: A multiband quasi-analytical algorithm for optically deep waters. *Applied Optics*, 41(27), 5755. <https://doi.org/10.1364/AO.41.005755>

- Lee, Z., Carder, K., Arnone, R., & He, M. (2007). Determination of primary spectral bands for remote sensing of aquatic environments. *Sensors*, 7(12), 3428–3441. <https://doi.org/10.3390/s7123428>
- Lee, Z., Du, K., Voss, K. J., Zibordi, G., Lubac, B., Arnone, R., & Weidemann, A. (2011). An inherent-optical-property-centered approach to correct the angular effects in water-leaving radiance. *Applied Optics*, 50(19), 3155–3167.
- Lohrenz, S. E., Weidemann, A., & Tuel, M. (2003). Phytoplankton spectral absorption as influenced by community size structure and pigment composition. *Journal of Plankton Research*, 25(1), 35–61.
- Mackey, M. D., Mackey, D. J., Higgins, H. W., & Wright, S. W. (1996). CHEMTAX — A program for estimating class abundances from chemical markers: Application to HPLC measurements of phytoplankton. *Marine Ecology Progress Series*, 144, 265–283.
- Mason, J. D., Cone, M. T., & Fry, E. S. (2016). Ultraviolet (250–550 Nm) absorption spectrum of pure water. *Applied Optics*, 55(25), 7163. <https://doi.org/10.1364/AO.55.007163>
- McKinna, L. I. W., Werdell, P. J., & Proctor, C. W. (2016). Implementation of an analytical Raman scattering correction for satellite ocean-color processing. *Optics Express*, 24(14), 4241–4255. <https://doi.org/10.1364/OE.24.0A1123>
- Mobley, C. D. (1994). *Light and water: Radiative transfer in natural waters*. San Diego, CA: Academic.
- Moisan, J. R., Moisan, T. A. H., & Linkswiler, M. A. (2011). An inverse modeling approach to estimating phytoplankton pigment concentrations from phytoplankton absorption spectra. *Journal of Geophysical Research*, 116, C09018. <https://doi.org/10.1029/2010JC006786>
- Morel, A., & Gentili, B. (1996). Diffuse reflectance of oceanic waters. III. Implication of bidirectionality for the remote-sensing problem. *Applied Optics*, 35 (24), 4850–4862. <https://doi.org/10.1364/AO.35.004850>.
- Morel, A., Huot, Y., Gentili, B., Werdell, P. J., Hooker, S. B., & Franz, B. A. (2007). Examining the consistency of products derived from various ocean color sensors in open ocean (Case 1) waters in the perspective of a multi-sensor approach. *Remote Sensing of Environment*, 111(1), 69–88. <https://doi.org/10.1016/j.rse.2007.03.012>
- Mouw, C. B., Barnett, A., Mckinley, G. A., Gloege, L., & Pilcher, D. (2016). Phytoplankton size impact on export flux in the global ocean. *Global Biogeochemical Cycles*, 30, 1542–1562. <https://doi.org/10.1002/2015GB005355>
- Mouw, C. B., Hardman-Mountford, N., Alvain, S., Bracher, A., Brewin, R., Bricaud, A., . . . Uitz, J. (2017). A consumer's guide to satellite remote sensing of phytoplankton groups in the global ocean. *Frontiers in Marine Science*, 4, 41. <https://doi.org/10.1002/2015GB005355>.
- O'Reilly, J. E., & 24 Coauthors (2000). SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3. In S. B. Hooker & E. R. Firestone (Eds.), *SeaWiFS technical report series* (NASA Tech. Memo. 2000–206892, Vol. 11, 49 pp.). Greenbelt, MD: NASA Goddard Space Flight Center.
- Pan, X., Mannino, A., Russ, M. E., Hooker, S. B., & Harding, L. W. (2010). Remote sensing of phytoplankton pigment distribution in the United States Northeast Coast. *Remote Sensing Environment*, 114(11), 2403–2416. <https://doi.org/10.1016/j.rse.2010.05.015>.
- Picheral, M., Searson, S., Taillandier, V., Bricaud, A., Boss, E., Ras, J., Gabriel, H., . . . Tara Oceans Consortium, Coordinators; Tara Oceans Expedition, Participants. (2014). *Vertical profiles of environmental parameters measured on discrete water samples collected with Niskin bottles during the Tara Oceans expedition 2009–2013*. PANGAEA. <https://doi.org/10.1594/PANGAEA.836319>
- Pope, R. M., & Fry, E. S. (1997). Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements. *Applied Optics*, 36(33), 8710–8723.
- Raitos, D. E., Lavender, S. J., Maravelias, C. D., Haralabous, J., Richardson, A. J., & Reid, P. C. (2008). Identifying four phytoplankton functional types from space: An ecological approach. *Limnology and Oceanography*, 53(2), 605–613.
- Ras, J., Claustre, H., & Uitz, J. (2008). Spatial variability of phytoplankton pigment distributions in the subtropical South Pacific Ocean: Comparison between in situ and predicted data. *Biogeosciences*, 5, 353–369.
- Roesler, C. S., & Boss, E. (2003). Spectral beam attenuation coefficient retrieved from ocean color inversion. *Geophysical Research Letters*, 30(9), 1468. <https://doi.org/10.1029/2002GL016185>
- Roesler, C. S., Perry, M. J., & Carder, K. L. (1989). Modeling in situ phytoplankton absorption from total absorption spectra in productive inland marine waters. *Limnology and Oceanography*, 34(8), 1510–1523. <https://doi.org/10.4319/lo.1989.34.8.1510>
- Slade, W. H., Boss, E., Dall'olmo, G., Langner, M. R., Loftin, J., Behrenfeld, M. J., . . . Westberry, T. K. (2010). Underway and moored methods for improving accuracy in measurement of spectral particulate absorption and attenuation. *Journal of Atmospheric and Oceanic Technology*, 27(10), 1733–1746. <https://doi.org/10.1175/2010JTECH0755.1>
- Sullivan, J. M., Twardowski, M. S., Zaneveld, J. R. V., Moore, C. M., Barnard, A. H., Donaghay, P. L., & Rhoades, B. (2006). Hyperspectral temperature and salt dependencies of absorption by water and heavy water in the 400–750 Nm spectral range. *Applied Optics*, 45(21), 5294. <https://doi.org/10.1364/AO.45.005294>
- Torreca, E., Stramski, D., Reynolds, R. A., Millán-Núñez, E., & Piera, J. (2011). Cluster analysis of hyperspectral optical data for discriminating phytoplankton pigment assemblages in the open ocean. *Remote Sensing Environment*, 115(10), 2578–2593. <https://doi.org/10.1016/j.rse.2011.05.014>
- Trees, C. C., Clark, D. K., Bidigare, R. R., Ondrusek, M. E., & Mueller, J. L. (2000). Accessory pigments versus chlorophyll a concentrations within the euphotic zone: A ubiquitous relationship. *Limnology and Oceanography*, 45(5), 1130–1143. <https://doi.org/10.4319/lo.2000.45.5.1130>
- Twardowski, M. S., Boss, E., Macdonald, J. B., Pegau, W. S., Barnard, A. H., & Zaneveld, J. R. V. (2001). A model for estimating bulk refractive index from the optical backscattering ratio and the implications for understanding particle composition in Case I and Case II waters. *Journal of Geophysical Research*, 106(C7), 14129–14142. <https://doi.org/10.1029/2000JC000404>
- Uitz, J., Claustre, H., Morel, A., & Hooker, S. B. (2006). Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll. *Journal of Geophysical Research*, 111, C08005. <https://doi.org/10.1029/2005JC003207>
- Uitz, J., Stramski, D., Reynolds, R. A., & Dubranna, J. (2015). Assessing phytoplankton community composition from hyperspectral measurements of phytoplankton absorption coefficient and remote-sensing reflectance in open-ocean environments. *Remote Sensing Environment*, 171, 58–74.
- Van Heukelem, L., & Thomas, C. S. (2001). Computer-assisted high-performance liquid chromatography method development with applications to the isolation and analysis of phytoplankton pigments. *Journal of Chromatography A*, 910, 31–49.
- Wang, G., Lee, Z., Mishra, D. R., & Ma, R. (2016). Retrieving absorption coefficients of multiple phytoplankton pigments from hyperspectral remote sensing reflectance measured over cyanobacteria bloom waters. *Limnology and Oceanography Methods*, 14(7), 425–489. <https://doi.org/10.1002/lom3.10102>
- Werdell, P. J., Bailey, S. W., Fargion, G. S., Pietras, C., Knobelspiesse, K. D., Feldman, G. C., & McClain, C. R. (2003). Unique data repository facilitates ocean color satellite validation. *EOS Transactions American Geophysical Union*, 84(38), 377.
- Werdell, P. J., Franz, B. A., Bailey, S. W., Feldman, G. C., Boss, E., Brando, V. E., . . . Mangin, A. (2013). Generalized ocean color inversion model for retrieving marine inherent optical properties. *Applied Optics*, 52(10), 2019–2037.
- Werdell, P. J., Roesler, C. S., & Goes, J. I. (2014). Discrimination of phytoplankton functional groups using an ocean reflectance inversion model. *Applied Optics*, 53(22), 4833–4849.

- Westberry, T. K., Boss, E., & Lee, Z. (2013). Influence of Raman scattering on ocean color inversion models. *Applied Optics*, 52(22), 5552–5561.
- Wolanin, A., Soppa, M., & Bracher, A. (2016). Investigation of spectral band requirements for improving retrievals of phytoplankton functional types. *Remote Sensing*, 8(10), 871. <https://doi.org/10.3390/rs8100871>
- Zhang, X., Hu, L., & He, M. (2009). Scattering by pure seawater: Effect of salinity. *Optics Express*, 17(7), 5698–5710.