8-1-2020

Chlorophyll absorption and phytoplankton size information inferred from hyperspectral particulate beam attenuation

Henry F. Houskeeper
University of California, Santa Cruz

David Draper
University of California, Santa Cruz

Raphael M. Kudela
University of California, Santa Cruz

Emmanuel Boss
University of Maine, emmanuel.boss@maine.edu

Title

Chlorophyll absorption and phytoplankton size information inferred from hyperspectral particulate beam attenuation.

Journal

Applied optics, 59(22)

ISSN

1559-128X

Authors

Houskeeper, Henry F
Draper, David
Kudela, Raphael M
et al.

Publication Date

2020-08-01

DOI

10.1364/ao.396832

Peer reviewed

applied optics

Chlorophyll absorption and phytoplankton size information inferred from hyperspectral particulate beam attenuation

HENRY F. HOUSKEEPER,[1,*] DAVID DRAPER,[2] RAPHAEL M. KUDELA,[1] AND EMMANUEL BOSS[3]

[1]Ocean Sciences Department, University of California, Santa Cruz, 1156 High Street, Santa Cruz, California 95064, USA
[2]Statistics Department, University of California, Santa Cruz, 1156 High Street, Santa Cruz, California 95064, USA
[3]School of Marine Sciences, University of Maine, 360 Aubert Hall, Orono, Maine 04469, USA
*Corresponding author: hhouskee@ucsc.edu

Received 4 May 2020; revised 29 June 2020; accepted 30 June 2020; posted 30 June 2020 (Doc. ID 396832); published 0 MONTH 0000

Electromagnetic theory predicts spectral dependencies in extinction efficiency near a narrow absorption band for a particle with an index of refraction close to that of the medium in which it is immersed. These absorption band effects are anticipated in oceanographic beam-attenuation (beam-c) spectra, primarily due to the narrow red peak in absorption produced by the phytoplankton photopigment, chlorophyll *a* (Chl *a*). Here we present a method to obtain Chl *a* absorption and size information by analyzing an eigendecomposition of hyperspectral beam-c residuals measured in marine surface waters by an automatic underway system. We find that three principal modes capture more than 99% of the variance in beam-c residuals at wavelengths near the Chl *a* red absorption peak. The spectral shapes of the eigenvectors resemble extinction efficiency residuals attributed to the absorption band effects. Projection of the eigenvectors onto the beam-c residuals produces a time series of amplitude functions with absolute values that are strongly correlated to concurrent Chl *a* absorption line height (a_{LH}) measurements (*r* values of 0.59 to 0.83) and hence provide a method to estimate Chl *a* absorption. Multiple linear regression of a_{LH} on the amplitude functions enables an independent estimate of a_{LH}, with RMSE of $3.19 \cdot 10^{-3}$ m^{-1} (3.3%) or $\log_{10}$-RMSE of 18.6%, and a raw-scale R^2 value of 0.894 based on the Tara Oceans Expedition data. Relationships between the amplitude functions and the beam-c exponential slopes are in agreement with theory relating beam-c to the particle size distribution. Compared to multispectral analysis of beam-c slope, hyperspectral analysis of absorption band effects is anticipated to be relatively insensitive to the addition of nonpigmented particles and to monodispersion.

https://doi.org/10.1364/AO.396832

1. INTRODUCTION

Bio-optical characterization of the marine environment supports global ocean monitoring by enabling the use of sensing infrastructure on platforms that range from autonomous floats to satellite imagers. In-water sensors lack the coverage and resample rates of satellite platforms but enable the direct measurement of light absorption and scattering processes. Multispectral absorption measurements have useful applications to measure marine ecosystems, for example, by enabling the estimation of chlorophyll *a* (Chl *a*) concentration [1,2], whereas Chl *a* fluorometers are easily deployed but suffer from uncertainty in phytoplankton assemblage and physiology [3,4]. Multispectral measurements of scattering and attenuation are strongly correlated [5], and their spectral slopes are related to the underlying particle size distribution (PSD). However, they cannot analytically resolve phytoplankton Chl *a* content [6]. Recent advances in hyperspectral instrumentation have enabled the decomposition of *in situ* particulate absorption spectra to discern accessory pigmentation relevant to describing a phytoplankton community composition [7].

For suspended particles with an index of refraction near that of the medium in which they are immersed (which in general includes phytoplankton), electromagnetic theory predicts wavelength (λ) dependencies in scattering or attenuation for spectral regions adjacent to narrow absorption bands, described as *anomalous dispersion* or *absorption band effects* [8–14]. The λ dependencies related to absorption band effects provide information about phytoplankton size, pigmentation, and refractive index at fine spectral scales measured by current hyperspectral sensors. In this paper, we present observations relevant to absorption band effects using a global and methodologically consistent dataset of particulate hyperspectral beam-c residuals, and we consider potential applications to characterize marine ecosystems. In particular, we show that these residuals enable

estimation of Chl *a* absorption and provide size information on phytoplankton.

A. Overview of Absorption Band Effects on Beam-Attenuation Spectra

The modification of the internal marine light field occurs through scattering (elastic and inelastic) and absorption processes, which are defined as the inherent optical properties (IOPs) of the aquatic medium [15]. The beam-attenuation (beam-c) coefficient $c(\lambda)$ describes the decay or directional change in a beam of collimated light and is the sum of the total absorption $a(\lambda)$ and scattering $b(\lambda)$ coefficients. In practice, $c(\lambda)$ and $a(\lambda)$ are more readily measured, and $b(\lambda)$ is obtained through subtraction. Although often treated as such, $a(\lambda)$ and $b(\lambda)$ are not independent properties. The coefficients may be further specified to represent the particulate (algal and non-algal) contributions by subtracting the properties of the dissolved (filtered) materials from those of the whole water [16], expressed as $a_p(\lambda)$, $b_p(\lambda)$, and $c_p(\lambda)$ for the particulate absorption, scattering, and beam-c coefficients, respectively.

For individual particles within a medium, the ratios of the optical cross-sections to the geometric cross-sections define the absorption $Q_a(\lambda)$, scattering $Q_b(\lambda)$, and extinction $Q_{ext}(\lambda)$ efficiency factors, which contribute to the bulk IOPs of a water-mass [i.e., to $a_p(\lambda)$, $b_p(\lambda)$], and $c_p(\lambda)$, respectively. Anomalous diffraction theory approximates $Q_{ext}(\lambda)$ for large (i.e., circumference much greater than the wavelength of light in the medium), non-absorbing, and homogeneous spheres through a phase-lag term, ρ (the change in a ray's phase if it were to travel the full diameter of a spherical particle), defined as [8]

$$\rho = 2x\,(n-1)\,, \quad \text{[unitless]} \tag{1}$$

in which n is the real index of refraction and x is the ratio of the particle's circumference to the wavelength of light in the medium, ranging from 0 to ∞. In this paper, the real index of refraction n is defined relative to seawater, and $(n-1)$ is assumed to be positive. The spectral dependency in x provides a theoretical basis to estimate the PSD from the spectral slope of $c_p(\lambda)$ that is sensitive to pigmented and nonpigmented constituents [17].

For absorbing particles, including phytoplankton, the imaginary component n' of the complex index of refraction m (defined as $m = n + in'$) corresponds to a particle's pigmentation [18] and is included in the anomalous diffraction approximation of the optical efficiency factors [8]. In the case of phytoplankton cells, Chl *a* and various accessory pigments elevate $Q_a(\lambda)$ and suppress $Q_b(\lambda)$ across a relatively broad range of blue wavelengths [19]. At red wavelengths, a special situation arises from the specific absorption spectrum of Chl *a*, which produces a narrow red absorption band. Changes to n and n' in the vicinity of the absorption band (nominally centered at 676 nm) modify $Q_{ext}(\lambda)$ based on the size and refractivity of the particle (i.e., ρ). Figure 1(a) illustrates the characteristic $Q_{ext}(\lambda)$ spectra in the vicinity of an absorption band using various ρ values and is based on the anomalous diffraction approximation of [8].

The predicted changes in $Q_{ext}(\lambda)$ are described as a function of ρ: For lower ρ values, $Q_{ext}(\lambda)$ is elevated at the absorption band [e.g., resembling an increase in $Q_a(\lambda)$]; for higher ρ values, $Q_{ext}(\lambda)$ is reduced at the absorption band; and for moderate ρ values, an anomalous dispersion curve emerges with $Q_{ext}(\lambda)$ reduced at shorter wavelengths and elevated at longer wavelengths (relative to the center of the absorption band). We use the more general term *absorption band effects*, following [13], to describe the spectral features in $Q_{ext}(\lambda)$ or $c_p(\lambda)$ that are observed near the absorption bands.

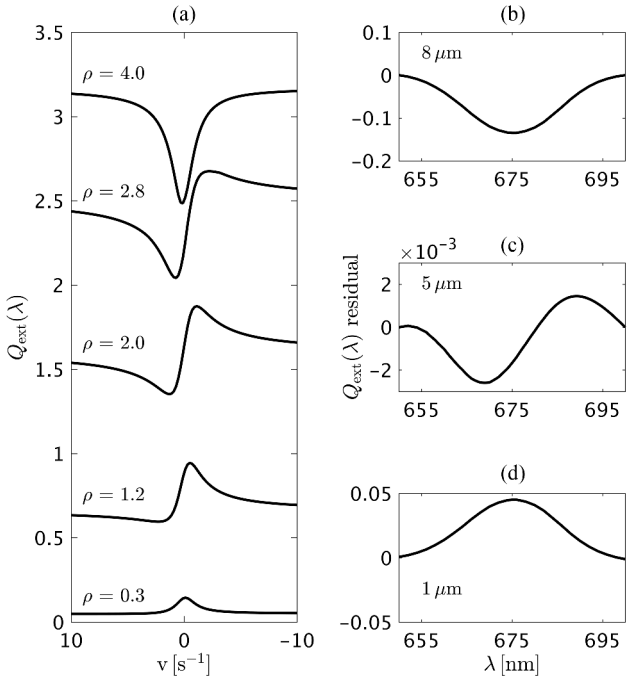

Fig. 1. (a) Anomalous diffraction approximation for $Q_{ext}(\lambda)$ at a narrow absorption band as a function of light frequency for various ρ values, recreated from van de Hulst [8] using a lookup table; and (b)–(d) illustrative examples of $Q_{ext}(\lambda)$ residuals for various small-sized phytoplankton (diameters 8, 5, and 1 μm, respectively), with a fixed n of 1.0344 and spectral n' with a maximum value of 0.0024 at the Chl *a* red absorption peak. The sizes presented in (b)–(d) are sensitive to the selection of real and imaginary refractive index; for example, as [14] illustrates, an anomalous dispersion curve for $Q_b(\lambda)$ using a 1 μm absorbing sphere and spectral dependencies in both n and n'.

B. Relevance of Absorption Band Effects to Phytoplankton Composition

The approximation that the real index of refraction is near that of seawater is valid for many types of phytoplankton, although natural variability exists due to differences in cellular composition [14]. For example, calcification generally corresponds to a higher refractivity of coccolithophores [20]. Cell size is relevant to the phase lag parameterization in such a way that, for constant cellular composition, smaller cells are associated with lower ρ values and larger cells with higher ρ values. The combined effects of cell size and index of refraction, therefore, yield the result that $Q_{ext}(\lambda)$ residuals in the spectral vicinity of a narrow absorption band can be positive for small phytoplankton with n near seawater, or negative for larger or more refractive phytoplankton. Within a narrow, intermediate range in size and refractivity, predicted $Q_{ext}(\lambda)$ residuals resemble an anomalous

dispersion curve. Thresholds for ρ have been approximated (e.g., $\rho < 3$ can correspond to an anomalous dispersion curve), but these limits are not particularly useful due to intracellular variability in refractivity, pigmentation, and cellular shape, as well as uncertainties in the actual size distribution when representing ρ for a theoretical mean equivalent particle ($\tilde{\rho}$) [10]. Illustrative $Q_{ext}(\lambda)$ residuals for various sizes of phytoplankton are shown in panels of Figs. 1(b–d) based on the anomalous diffraction approximation following [8,11].

Considering IOPs of polydisperse systems rather than single particle efficiencies, the $c_p(\lambda)$ and $b_p(\lambda)$ spectra are anticipated to be smoothed by the diversity of refractivity and cell sizes (i.e., polydispersion), in natural phytoplankton communities [10]. However, communities dominated by small cell sizes, characteristic of many oligotrophic marine environments, are anticipated to produce peaks and anomalous dispersion curves in $c_p(\lambda)$ measurements. Communities where absorption is dominated by large phytoplankton (i.e., microplankton) are anticipated to produce local minima in $c_p(\lambda)$ measurements. The objective of this paper is to advance understanding of how absorption band effects influence $c_p(\lambda)$ spectra at red wavelengths; in particular, by partitioning the contributions from polydisperse systems of phytoplankton to enable inference of biomass or community information. We do not focus here on $b_p(\lambda)$ or on its backward component, $b_{bp}(\lambda)$, as hyperspectral $c_p(\lambda)$ is more routinely measured *in situ*. However, we note that $b_{bp}(\lambda)$ is more strongly affected by absorption band effects compared to $c_p(\lambda)$ and $b_p(\lambda)$ [13], and is likewise most relevant to a remote sensing perspective [21].

2. MATERIALS AND METHODS

A. Description of Bio-optical Dataset

Bio-optical oceanographic data was collected during the Tara Oceans Expedition, in which an aluminum-hulled schooner sailed through the Pacific, Atlantic, and Indian Ocean basins, as well as the Caribbean, Mediterranean, and Red Seas, while continually sampling surface waters using a flow-through system [22]. Briefly, seawater was routed to a WET Labs ac-s meter, which measures $c(\lambda)$ and $a(\lambda)$ by passing water through separate columns illuminated by collimated and diffuse light sources, respectively. The ac-s instrument is hyperspectral and measures approximately 80 wavelengths spanning 400–730 nm.

The relative calibration of the ac-s meter in the underway configuration was achieved during the Tara Oceans Expedition by periodically filtering (0.2 μm) the flow-through samples (every 30 or 60 min) and subtracting the measurements of the dissolved samples from the total, as described in [23]. The difference corresponds to the particulate contributions, $c_p(\lambda)$ and $a_p(\lambda)$, with scattering corrections performed following [24]. The ac-s instrument has a nonnegligible acceptance angle of 0.93 deg for beam-c measurements, which decreases the sensitivity to scattering by large particles. It can introduce a bias by decreasing the contribution of large particles in measured $c_p(\lambda)$ relative to theoretical $c_p(\lambda)$ [25].

We accessed Tara Oceans Expedition $c_p(\lambda)$ and $a_p(\lambda)$ measurements at one-minute temporal resolution through the NASA SeaWiFS Bio-Optical Archive and Storage System (SeaBASS; seabass.gsfc.nasa.gov) data repository [26–28]. The global sampling locations of the concurrent $a_p(\lambda)$ and $c_p(\lambda)$ measurements used in this study (245,277 in total) are shown in Fig. 2.

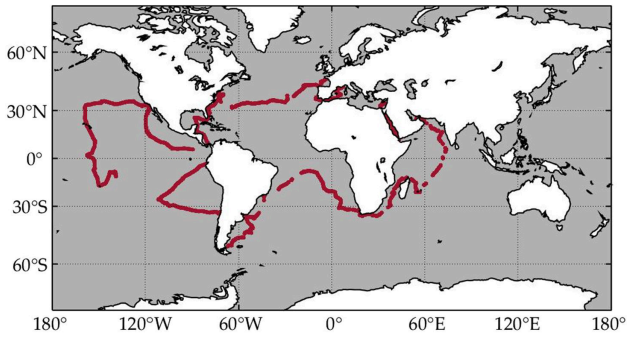

Fig. 2. Sampling locations for bio-optical measurements obtained from the Tara Oceans Expedition archive in SeaBASS.

The exact spectral locations of the ac-s wavebands differ, both between instruments and within the same instrument following factory recalibration; in response, we linearly interpolated all $c_p(\lambda)$ spectra onto a consistent waveband set. Following [7], we removed filter artifacts that resulted from default smoothing of the ac-s spectra across $c_p(\lambda)$ channels. We modeled broad spectral dependencies in $c_p(\lambda)$ that result from the relationship between particle size and wavelength as a power-law function [6] using the wavelengths within 500–650 and 700–720 nm, and subtracted the power-law model from the $c_p(\lambda)$ spectra to obtain residuals, denoted $c_p'(\lambda)$, as

$$c_p'(\lambda) = c_p(\lambda) - A\left[\frac{532\ \text{nm}}{\lambda}\right]^{\gamma}, \quad [\text{m}^{-1}], \tag{2}$$

where A and γ are free parameters obtained using a least-squares optimization. The PSD of natural oceanic particles also is represented by a power-law function [16], with exponential slope, ξ, that may be predicted from the $c_p(\lambda)$ exponential slope, γ [6]. Increases in γ have been shown to correspond to increases in ξ; e.g., steeper (more negative) $c_p(\lambda)$ slopes indicate greater relative proportions of smaller particles.

We obtained an independent proxy for Chl *a* absorption from coincident $a_p(\lambda)$ measurements based on the height of the red peak that is attributed to Chl *a* absorption. Briefly, we linearly interpolated an absorption baseline between shoulder wavebands and subtracted this from the region of maximum Chl *a* red absorption, as

$$a_{LH} = a_p(\lambda_2) - a_p(\lambda_1) - \left(\frac{\lambda_2 - \lambda_1}{\lambda_3 - \lambda_1}\right) \times \left[a_p(\lambda_3) - a_p(\lambda_1)\right], \quad [\text{m}^{-1}], \tag{3}$$

in which λ_2 is the red peak (nominally 676 nm), and λ_1 and λ_3 are adjacent legacy wavelengths (e.g., 650 and 715 nm). Estimation of Chl *a* from a_{LH} has been shown to be relatively insensitive to pigment packaging effects [29,30] compared to blue wavelength algorithms, as well as to changes in phytoplankton physiology compared with Chl *a* fluorescence-based techniques [31] (see [32] for field data).

B. Reduction of $c_p'(\lambda)$ Spectra Using Empirical Orthogonal Functions

We reduced the dimensionality of the beam-c residuals dataset using an eigenanalysis of the $c_p'(\lambda)$ dataset within a narrow spectral subset (13 wavebands spanning approximately 50 nm) centered on the Chl *a* red absorption peak (nominally 676 nm). Briefly, we performed an eigendecomposition of the $c_p'(\lambda)$ covariance matrix of the form,

$$\boldsymbol{C}\boldsymbol{\psi} = \boldsymbol{\Lambda}\boldsymbol{\psi}, \quad [\mathrm{m}^{-2}], \tag{4}$$

in which $\boldsymbol{C}$ is the covariance matrix of the $c_p'(\lambda)$ dataset and $\boldsymbol{\psi}$ is the eigenfunction matrix, with 13 columns $\boldsymbol{\psi}_i(\lambda)$ describing modes of variability across the wavelength domain of the $c_p'(\lambda)$ data. In Section 3.B below we compare the spectral shape of the $\boldsymbol{\psi}_i(\lambda)$ eigenvectors with the $Q_{ext}(\lambda)$ residuals predicted for absorption band effects. The diagonal matrix $\boldsymbol{\Lambda}$ contains eigenvalues relating scalar information for each eigenvector, with the sum of the eigenvalues equal to the sum of the wavelength-specific variances in the $c_p'(\lambda)$ dataset (the diagonal elements of the covariance matrix $\boldsymbol{C}$), expressed as

$$\sum_{i=1}^{k=13} \Lambda_{i,i} = \sum_{i=1}^{k=13} \sigma^2_{(\lambda_i,\lambda_i)}. \quad [\mathrm{m}^{-2}]. \tag{5}$$

Comparison of the eigenvalues $\Lambda_{i,i}$ enables consideration of the variance captured by each eigenfunction. We reduced the spectra of the $c_p'(\lambda)$ dataset to scalar amplitudes by projecting the $c_p'(\lambda)$ data onto the eigenfunction matrix, as

$$\boldsymbol{S} = c_p'\boldsymbol{\psi}, \quad [\mathrm{m}^{-1}], \tag{6}$$

in which the 13 columns of the $\boldsymbol{S}$ matrix are a time series of amplitude functions that quantify the stretching and compressing necessary to represent the $c_p'(\lambda)$ dataset in the new coordinates defined by the eigenvector basis functions.

We assessed the relationships between the components of the $\boldsymbol{S}$ matrix time series and the a_{LH} and γ data products through univariate and multivariate regression. The positive and negative phases of each $\boldsymbol{S}$ component i were treated as separate $\boldsymbol{S}$ matrix predictors because positive and negative signs in $\boldsymbol{S}_i$ indicate corresponding phase shifts in $\boldsymbol{\psi}_i(\lambda)$ (i.e., phase shifts mirror the eigenvector spectral shape in the y-dimension). The positive and negative phases in $\boldsymbol{\psi}_i(\lambda)$ are anticipated to resemble the different shapes of the $Q_{ext}(\lambda)$ residuals (i.e., the maxima or minima shown in Fig. 1) due to the absence of other significant $c_p(\lambda)$ spectral dependencies within this wavelength domain. The predictors corresponding to the positive and negative values of each $\boldsymbol{S}_i$ function, $\boldsymbol{P}_i^{(+)}$ and $\boldsymbol{P}_i^{(-)}$, are defined as

$$\begin{aligned} &\boldsymbol{P}_i^{(+)} = \boldsymbol{S}_i; \quad \boldsymbol{P}_i^{(-)} = 0, \quad (\text{if } \boldsymbol{S}_i > 0); \quad [\mathrm{m}^{-1}], \\ &\boldsymbol{P}_i^{(+)} = 0; \quad \boldsymbol{P}_i^{(-)} = |\boldsymbol{S}_i|, \quad (\text{if } \boldsymbol{S}_i \le 0). \quad [\mathrm{m}^{-1}]. \end{aligned} \tag{7}$$

We evaluated regressions using the root mean squared error of prediction (RMSE), which–when represented as a percentage for raw-scale values—was normalized by the range in the a_{LH} dataset. We analyzed the $\boldsymbol{S}$ matrix using a thinned dataset to reduce autocorrelation related to the relatively long spatial decorrelation scales of marine waters compared to the average speed of the Tara vessel. We subsampled the dataset across approximate length scales of 11.1 km and 33.3 km for coastal (within 200 km from the shore) and oceanic (over 200 km from the shore) water masses, respectively. Using the subsampled dataset, we evaluated the accuracy of the combined $\boldsymbol{S}$ matrix predictors to estimate a_{LH} using multiple linear regression over 10,000 cross-validation replications. In each replication, we randomly partitioned the dataset into modeling and validation subsets using an 80%/20% split, which corresponded to 2262 and 565 data points, respectively.

3. RESULTS

A. Effects of Autocorrelation

Autocorrelation is a persistently challenging topic in oceanography because spatial and temporal decorrelation scales are variable among regions and seasons, and because large differences between marine provinces (e.g., coastal zones, oligotrophic gyres, upwelling regions) often overshadow smaller-scale variability within each region. Our spatial-scale thinning of the Tara $c_p'(\lambda)$ dataset decreased the total number of observations by more than 98%. The coastal zones were measured less frequently than oceanic waters in the raw dataset because the Tara vessel generally maintained a trans-oceanic course. Our spatial thinning, which used different length scales for coastal and oceanic measurements, increased the proportion of coastal waters (i.e., within 200 km of shore) within our dataset to 50%, compared to 27% within the original dataset. Despite the changes in size and regional representation due to thinning, the spectral shapes and ordering of the eigenfunctions were not significantly altered by subsampling, and the eigenvalues corresponding to the first three modes of the subsampled dataset were each within 2% of those derived from the full (not subsampled) dataset.

B. Interpretation of the Eigenfunctions

The bulk optical properties of a watermass integrate contributions from various constituents, as well as from the medium. An idealized equation relating $c_p(\lambda)$ to the size-dependent $Q_{ext}(\lambda, s)$ contributions is of the form [8],

$$c_p(\lambda) = \int_0^\infty Q_{ext}(\lambda, s)N(s)\pi s^2 \mathrm{d}s, \quad [\mathrm{m}^{-1}], \tag{8}$$

in which $N(s)$ is the number of particles with radius s. The $c_p(\lambda)$ spectra is a bulk property that arises from the addition of the underlying particle properties, making the spectral dependencies in $Q_{ext}(\lambda, s)$ directly related to the spectral shape of $c_p(\lambda)$. Our eigenanalysis of the $c_p'(\lambda)$ covariance matrix quantified the primary modes of variability in the $c_p'(\lambda)$ spectra, which are shown in Fig. 3. Although both positive and negative phases of the eigenfunctions may resemble the spectral residuals associated with absorption band effects, only one representative phase is shown for each eigenvector.

The first three modes of the eigenanalysis captured more than 99% of the variance, with the first, second, and third

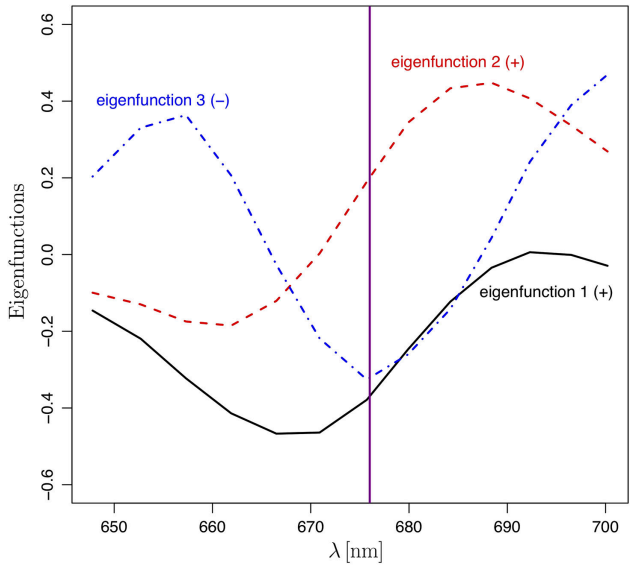

Fig. 3. Eigenfunctions of the $c'_p(\lambda)$ covariance matrix, with the nominal location of the Chl *a* red absorption maximum indicated by the vertical line.

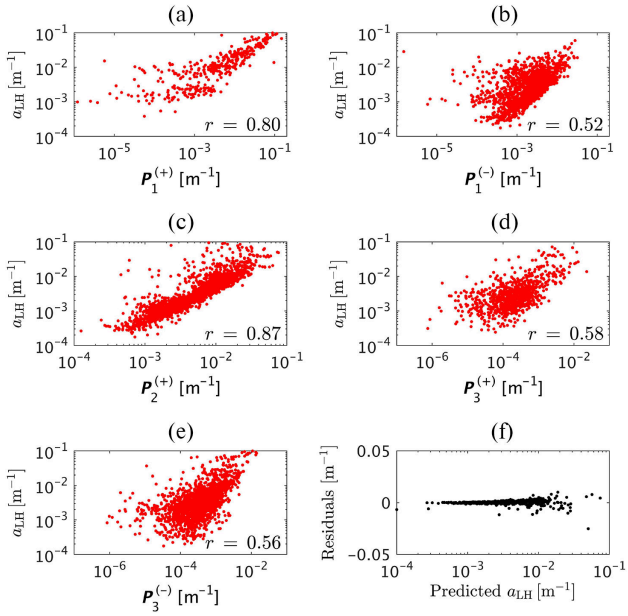

Fig. 4. Relationships between a_{LH} and the $\boldsymbol{S}$ matrix predictors: (a) $\boldsymbol{P}_1^{(+)}$; (b) $\boldsymbol{P}_1^{(-)}$; (c) $\boldsymbol{P}_2^{(+)}$; (d) $\boldsymbol{P}_3^{(+)}$; and (e) $\boldsymbol{P}_3^{(-)}$. (f) Residuals against predicted a_{LH} values from a multivariate linear regression of the $\boldsymbol{S}$ matrix predictors. The horizontal and vertical scales in (a)–(e) and the horizontal scales in (f) are $\log_{10}$.

eigenvectors comprising 61.1%, 37.0%, and 1.2% of the total, respectively. We did not examine the additional components, which individually corresponded to 0.25% or less of the variance. We propose the following interpretations of the spectral shapes of the relevant positive (+) and negative (−) eigenfunctions, with the local maximum for Chl *a* absorption at red wavelengths defined as the *absorption band* for brevity:

- Eigenfunction 1 (+): a negative anomaly shifted ∼10 nm shorter than the absorption band.
- Eigenfunction 1 (−): a positive anomaly shifted ∼10 nm shorter than the absorption band.
- Eigenfunction 2 (+): a positive anomaly shifted (∼12 nm) longer than the absorption band, a negative anomaly shifted (∼15 nm) shorter than the absorption band, and an inflection near the absorption band.
- Eigenfunction 2 (−): a negative anomaly shifted (∼12 nm) longer than the absorption band, a positive anomaly shifted (∼15 nm) shorter than the absorption band, and an inflection near the absorption band.
- Eigenfunction 3 (+): a positive anomaly centered at the absorption band.
- Eigenfunction 3 (−): a negative anomaly centered at the absorption band.

Considering the theoretical parameterizations governing absorption band effects on $Q_{ext}(\lambda)$, the eigenvector basis functions also may be interpreted through a transition in ρ values (i.e., by comparing Figs. 1 and 3), as follows: The lowest ρ (an increase at the absorption band) corresponds to eigenfunctions 1 (−) and 3 (+); intermediate ρ (the anomalous dispersion curve) corresponds to eigenfunction 2 (+); and the highest ρ (a decrease at the absorption band) corresponds to eigenfunctions 1 (+) and 3 (−). The spectral shape of eigenfunction 2 (−) was not in agreement with the $Q_{ext}(\lambda)$ residuals predicted for absorption band effects, as shown in Fig. 1, and we therefore regard it as potentially nonphysical within the scope of this work. This conclusion is supported by the low expression of eigenfunction 2 (−) in the $\boldsymbol{S}$ matrix time series, with less than 1% of the $\boldsymbol{S}_2$ amplitudes negative. The same is true, but to a lesser extent, for eigenfunction 1 (−), for which the maximum was less spectrally separate from the absorption band compared to eigenfunction 2 (−). Similarly, less than 11% of the $\boldsymbol{S}_1$ amplitudes were negative.

C. Interpretation of the S Matrix Predictors

The relationships between the $\boldsymbol{S}$ matrix predictors and the a_{LH} dataset are shown individually in Fig. 4, with the negative amplitudes from the time series for component 2 (26 out of 2827 observations) omitted. Each $\boldsymbol{S}$ matrix predictor had highly significant positive correlation with a_{LH} (posterior probability <0.001 of nonpositive correlation in the population to which our results generalize), except for the negative amplitudes in the time series for component 2, which indicated a positive but insignificant relationship.

We found that, for each eigenanalysis component, slope coefficients derived by linear univariate regression of a_{LH} and the $\boldsymbol{S}$ matrix predictors were greater for the phases that corresponded to higher ρ values based on the interpretations described in Section 3.B. For example, the positive phase of component 1, which corresponds to a negative anomaly near the absorption band (higher ρ), produced a significantly steeper slope than the negative phase of component 1, which corresponds to a positive anomaly near the absorption band (lower ρ). Similarly, the positive phase of component 2 (anomalous dispersion curve; intermediate ρ) produced a significant positive slope, while the slope of the negative phase (nonphysical) was flattened and

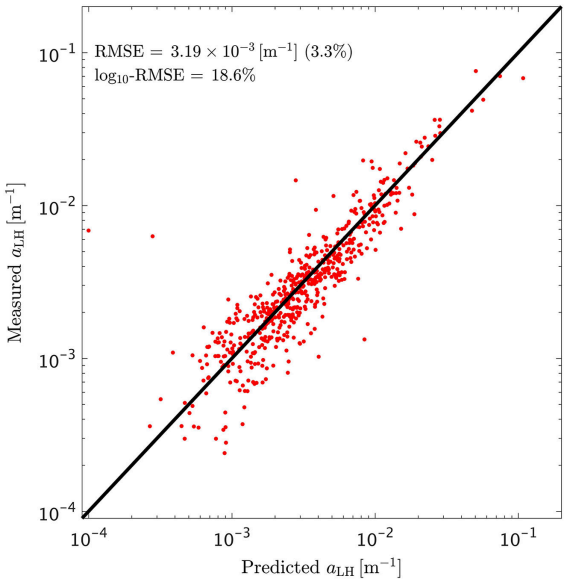

Fig. 5. Median validation scatterplot relating measured (vertical) and predicted (horizontal) a_{LH} values from the $\boldsymbol{S}$ matrix predictors, with both axes on the $\log_{10}$ scale.

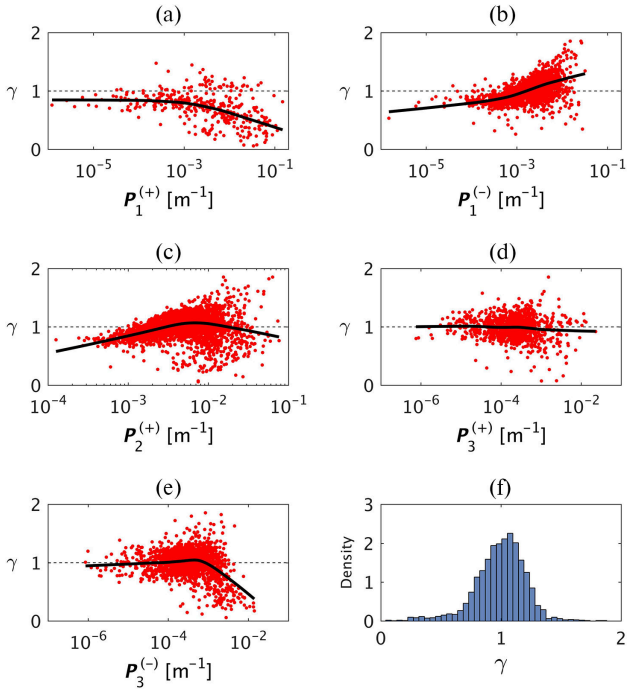

Fig. 6. Relationships between γ and the $\boldsymbol{S}$ matrix predictors: (a) $\boldsymbol{P}_1{}^{(+)}$; (b) $\boldsymbol{P}_1{}^{(-)}$; (c) $\boldsymbol{P}_2{}^{(+)}$; (d) $\boldsymbol{P}_3{}^{(+)}$; and (e) $\boldsymbol{P}_3{}^{(-)}$, with locally weighted scatterplot smoothing (lowess) functions overlaid in solid black. The horizontal scales are $\log_{10}$. (f) Histogram estimate of the probability density function of the γ values derived from the $c_p(\lambda)$ dataset.

insignificant. The negative phase of component 3 (absorption band minima; higher ρ) produced a greater slope than the positive phase (absorption band maxima; lower ρ).

The repeated cross-validation within the thinned Tara Oceans Expedition dataset produced a median RMSE to estimate a_{LH} of $3.19 \cdot 10^{-3} \pm 0.55 \cdot 10^{-3}$ m^{-1}, corresponding to 3.3% of the range in a_{LH}. RMSE derived from $\log_{10}$-transformed variables ($\log_{10}$-RMSE) indicated uncertainty of 18.6% ± 1.7%. The median R^2 value in the validation datasets was 0.894. A link to code based on the median-performing model is provided in the supplemental materials. The log-log scatter plot of measured and predicted a_{LH} values in the median validation subset is shown in Fig. 5.

The relationships for the $\boldsymbol{S}$ matrix predictors and the exponential slopes of the $c_p(\lambda)$ dataset, γ, were evaluated using the nonparametric scatterplot smoother lowess [33], shown in Fig. 6 with $\log_{10}$ horizontal scales. As in Fig. 4, we omitted the predictor corresponding to the negative phase of $\boldsymbol{S}_2$, due to the low number of observations and lack of physical interpretability. The nonparametric smoothers indicate that large positive expressions of mode 1 and large negative expressions of mode 3 in the $\boldsymbol{S}$ matrix correspond with decreasing γ, and large negative expressions of mode 1 correspond with increasing γ. Based on our Section 3.B interpretation of the eigenvector spectra, the relationships in Figs. 6 (a–c) also could be expressed in terms of ρ, indicating a negative association between γ and ρ. Ignoring the variability in m (i.e., $\rho \propto x$), the results are in agreement with theory relating γ to the PSD. For example, an increase in the concentration of larger cells corresponds to a lower γ and a greater expression of the high-ρ amplitude functions $\boldsymbol{P}_1^{(+)}$ and $\boldsymbol{P}_3^{(-)}$. However, the theoretical relationship between PSD and γ corresponds to the full particle population, while the relationship to the absorption band effects corresponds to the pigmented particle fraction.

4. DISCUSSION

A. Relevance of Absorption Band Effects to Phytoplankton Dynamics

The development of hyperspectral IOP sensors and their deployment in a continuous, underway configuration provided us with a large surface ocean IOP dataset with 13 wavebands within about 25 nm of the Chl *a* red absorption peak. Our eigendecomposition found that three principal modes captured more than 99% of the variance in the $c_p'(\lambda)$ dataset. The shapes of these principal modes resembled the $Q_{ext}(\lambda)$ spectral residuals predicted by the electromagnetic theory to arise in the vicinity of narrow absorption bands.

The first eigenfunction, which captured 61.1% of the variance in the $c_p'(\lambda)$ dataset, was primarily expressed in its positive phase in the $\boldsymbol{S}$ matrix time series, which corresponded to a minimum at wavelengths shorter (~10 nm) than the Chl *a* red absorption maximum. The second eigenfunction, which captured 37.0% of the $c_p'(\lambda)$ variance, most closely resembled the anomalous dispersion curve illustrated in Fig. 1(c). Absorption band effects are not anticipated to produce spectral shapes resembling the reflection of this curve (recall Section 3.B and Fig. 1), and likewise the negative phase of the second eigenfunction was rarely present in the $\boldsymbol{S}$ matrix time series (less than 1% of the $\boldsymbol{S}_2$ values were negative). The variance captured by the second eigenfunction is interesting because previous analyses have indicated that the anomalous dispersion result is only relevant to very small phytoplankton(e.g., [14]). Although caution is warranted in relating ρ to phytoplankton size because

of uncertainty in cellular shape, refractivity, and pigmentation [10], the results of our eigenanalysis interpretation are in general agreement with the ecological perspectives that picoplankton (diameters generally less than 1 or 2 μm) are ubiquitous across oceanic gyre ecosystems [34] and that increases in biomass often result from the addition of larger cells over a more stable background of smaller cells [35]. The third eigenfunction captured only 1.2% of the variance in $c_p'(\lambda)$, but was retained because its spectral shape resembled the $Q_{ext}(\lambda)$ anomalies corresponding to the ρ end members (i.e., the maximum and minimum were centered on the absorption band).

The eigenvector projections onto the $c_p'(\lambda)$ dataset indicate that the separate phases of the $\boldsymbol{S}$ matrix predictors [i.e., $\boldsymbol{P}_i^{(+)}$ and $\boldsymbol{P}_i^{(-)}$ in Eq. (7)], are each significantly and positively correlated with a_{LH}, with the exception of $\boldsymbol{P}_2^{(-)}$, which corresponded to an eigenvector spectral shape that was not predicted from the absorption band effects. Predictors associated with higher ρ domains (based on interpretations in Section 3.B) indicate steeper relationships with a_{LH} compared to those associated with lower ρ domains. The results suggest that additions of larger or more refractive phytoplankton [see Eq. (1)] correspond to greater increases in a_{LH} than additions of smaller or less refractive phytoplankton. In general, this perspective is in agreement with the phytoplankton ecological paradigm that larger phytoplankton predominantly occupy more productive water masses, while smaller phytoplankton are ubiquitous across less productive, more oligotrophic regimes [36,37]. Considering the relationships between γ and the $\boldsymbol{S}$ matrix predictors, our findings are consistent with electromagnetic theory, although we could not separate the effects of refractivity and size with the approach taken here.

Our theoretical description of absorption band effects is based on the treatment of phytoplankton as simple, homogeneous spheres. Phytoplankton are often nonspherical, but the $Q_{ext}(\lambda)$ approximations discussed here can be generalized to describe the average efficiency factors for nonspherical particles, provided that the particles are randomly oriented within the medium [38]. Phytoplankton also contain various internal cellular structures that produce a large variability in the refractive index, particularly for structures bound by lipid membranes, those containing gases, or housing pigmented molecules, as well as layering by plates and frustules that encase coccolithophore and diatom cells, respectively. As a conceptualization, our theoretical description provides a basis to interpret the eigenanalysis and is consistent with previous approaches that have advanced fundamental concepts in marine optics using simple representations of phytoplankton cells as homogeneous [39] or layered [40] spheres.

Despite the terminology *anomalous dispersion*, the theory described in this paper is in agreement with basic principles (e.g., ρ decreases with increasing wavelength), with all else being equal and consistent with normal dispersion. However, key differences between the absorption band effects and γ analysis are relevant to measuring marine systems. First, while $c_p(\lambda)$ slope methods are sensitive to the sizes of pigmented and nonpigmented particles [6], absorption band effects are only anticipated for pigmented particles, as spectral dependencies in particle n' values arise primarily through absorption by photopigments. Second, the relationship between the $c_p(\lambda)$ and PSD slope is robust in polydisperse environments within an appropriate PSD size range, but is challenged in monodisperse systems [41]. Absorption band effects are confounded by polydispersion [10], but we show that hyperspectral resolution enables the decomposition of overlapping signals to extract key biological information from polydisperse environments.

B. Potential for Application of Absorption Band Effects

The methods presented here enabled us to assess spectral variability in $c_p(\lambda)$ near a photopigment absorption band without requiring *a priori* decisions about the shapes of the extracted signals. Optimization of the spectral shapes of extracted signals (e.g., using theoretical response functions), may improve estimates of a_{LH} from $c_p(\lambda)$ datasets. We compared our analysis with an eigendecomposition using $c_p'(\lambda)$ spectra that were peak-normalized, and we found that the accuracy of our estimates of Chl *a* absorption decreased, although the relationships between the $\boldsymbol{S}$ matrix predictors and γ were similar. Advancing the capability to estimate Chl *a* absorption from beam-c would provide useful redundancies in instances in which both beam-c and absorption are measured concurrently (e.g., by an ac-s) and would be useful in turbid waterbodies, where absorption meters are more prone to fouling than beam-c meters. In general, beam-c is more easily measured than absorption.

Information on PSD can be derived from the $c_p(\lambda)$ exponential slope using legacy multispectral instrumentation because the approach only requires measurement of two wavelengths [41]. Decomposition methods require greater spectral resolution, because the targeted signals are often more complex and spectrally overlapping. Because the signals associated with absorption band effects are most apparent within a narrow range in λ, even the hyperspectral ac-s instrument only provided up to 13 relevant wavebands in this study. Improvement in spectral resolution generally coincides with a trade-off in radiometric accuracy, which is problematic for decomposition of low amplitude signals. For example, the signals associated with absorption band effects are relatively low compared to the ranges in $c_p(\lambda)$ that result from variability in the refractivity and size distributions of pigmented or nonpigmented particles in natural marine systems.

Low signal of the spectral anomalies related to absorption band effects may be partially mitigated by the measurement of backscattering, rather than beam-attenuation or total scattering, because the backscattering spectra are more sensitive to absorption band effects [13]. Recently, a commercial hyperspectral $b_b(\lambda)$ instrument has been developed, which could potentially be applied to advance this topic [42]. For water bodies dominated by large refractive phytoplankton, $c_p(\lambda)$ and $b_{bp}(\lambda)$ are anticipated to produce local minima at the Chl *a* red absorption band (e.g., consider Fig. 1 for high ρ), and an estimation of phytoplankton biomass could be performed using a line height approach to reduce spectral resolution requirements. Research to develop multispectral backscatter instruments that target absorption band effects to predict Chl *a* is underway [43]. In the work summarized here, we did not investigate the importance of accessory pigmentation on absorption band effects, because

most phytoplankton photopigments are not active in the vicinity of the Chl *a* red peak. Chlorophyll *b*, which can form an absorption plateau with Chl *a* at high concentrations, is one notable exception.

5. CONCLUSIONS

We reduced the dimensionality of a surface ocean hyperspectral beam-c dataset with minimal loss of information by identifying three principal modes of spectral variability, which were similar in shape to the spectra of theoretical particle extinction residuals associated with absorption band effects. The results indicate that at wavelengths adjacent to the Chl *a* red absorption peak, absorption band effects are a primary source of variability in beam-c spectra, due in part to the absence of other strong spectral dependencies within the region. Challenges to our approach include low signal and high spectral requirements, polydispersion of natural marine ecosystems, and variability in the pigmentation, refractivity, and shape of marine phytoplankton.

The positive and negative amplitudes of the major eigenfunctions we found provided useful predictors for a_{LH} in our study, indicating that analysis of absorption band effects in $c_p(\lambda)$ spectra can enable an alternative estimate of Chl *a* absorption. The relationships between the eigenfunctions and the $c_p(\lambda)$ exponential slopes are in agreement with electromagnetic theory and suggest that useful size parameters could be estimated from the decomposition of hyperspectral beam-c measurements. Future improvements in measuring the index of refraction of phytoplankton will be useful to advance these topics, and the upper and lower limits for interpreting absorption band effects should be explored. Culture work, in particular, could help better elucidate the bio-optical relationships described by this analysis.

Finally, in this study we considered absorption band effects using an observational approach that was made possible by advances in hyperspectral IOP instrumentation. Basic, rather than applied, scientific research was first necessary to develop an understanding of the optics of narrow absorption bands, and our work relies on advances achieved through electromagnetic theory (e.g., [9–13]. Our principal conclusion is that, with recent advances in IOP instrumentation, current hyperspectral beam-c datasets enable accurate estimation of Chl *a* absorption based on information captured from absorption band effects.

Funding. National Aeronautics and Space Administration (80NSSC19K0812, NNX17AK89G)

Acknowledgment. We are grateful to Kenneth Voss (University of Miami), Collin Roesler (Bowdoin College), Curtis Mobley (Sequoia Scientific), and Ivona Cetinić (NASA GSFC) for useful discussions that helped clarify topics related to this work. Numerous student participants in the NASA Ocean Optics summer course also provided valuable feedback for this project. The collection of the Tara data was possible with the support of NASA's Ocean Biology and Biogeochemistry program and the Tara Oceans Foundation.

Disclosures. The authors declare no conflicts of interest.

REFERENCES

1. R. Davis, C. Moore, J. Zaneveld, and J. Napp, "Reducing the effects of fouling on chlorophyll estimates derived from long-term deployments of optical instruments," J. Geophys. Res. Oceans **102**, 5851–5855 (1997).
2. E. S. Boss, R. Collier, G. Larson, K. Fennel, and W. Pegau, "Measurements of spectral optical properties and their relation to biogeochemical variables and processes in Crater lake, Crater Lake National Park, OR," Hydrobiologia **574**, 149–159 (2007).
3. J. J. Cullen, "The deep chlorophyll maximum: comparing vertical profiles of chlorophyll a," Can. J. Fish. Aquat. Sci. **39**, 791–803 (1982).
4. C. Roesler, J. Uitz, H. Claustre, E. Boss, X. Xing, E. Organelli, N. Briggs, A. Bricaud, C. Schmechtig, A. Poteau, F. D'Ortenzio, J. Ras, S. Drapeau, N. Haëntjens, and M. Barbieux, "Recommendations for obtaining unbiased chlorophyll estimates from in situ chlorophyll fluorometers: a global analysis of wet labs eco sensors," Limnol. Oceanogr. Methods **15**, 572–585 (2017).
5. G. Dall'Olmo, T. Westberry, M. Behrenfeld, E. Boss, and W. Slade, "Significant contribution of large particles to optical backscattering in the open ocean," Biogeosciences **6**, 947–967 (2009).
6. J. Kitchen, J. Zaneveld, and H. Pak, "Effect of particle size distribution and chlorophyll content on beam attenuation spectra," Appl. Opt. **21**, 3913–3918 (1982).
7. A. Chase, E. Boss, R. Zaneveld, A. Bricaud, H. Claustre, J. Ras, G. Dall'Olmo, and T. K. Westberry, "Decomposition of in situ particulate absorption spectra," Methods Oceanogr. **7**, 110–124 (2013).
8. H. Van de Hulst, *Light Scattering by Small Particles* (Wiley, 1957).
9. A. Morel and A. Bricaud, "Theoretical results concerning light absorption in a discrete medium, and application to specific absorption of phytoplankton," Deep Sea Res. A **28**, 1375–1393 (1981).
10. A. Bricaud, A. Morel, and L. Prieur, "Optical efficiency factors of some phytoplankters," Limnol. Oceanogr. **28**, 816–832 (1983).
11. A. Morel and A. Bricaud, "Inherent optical properties of algal cells including picoplankton: theoretical and experimental results," Can. Bull. Fish. Aquat. Sci. **214**, 521–559 (1986).
12. D. Stramski, A. Morel, and A. Bricaud, "Modeling the light attenuation and scattering by spherical phytoplanktonic cells: a retrieval of the bulk refractive index," Appl. Opt. **27**, 3954–3956 (1988).
13. J. R. V. Zaneveld and J. C. Kitchen, "The variation in the inherent optical properties of phytoplankton near an absorption peak as determined by various models of cell structure," J. Geophys. Res. Oceans **100**, 13309–13320 (1995).
14. E. Aas, "Refractive index of phytoplankton derived from its metabolite composition," J. Plankton Res. **18**, 2223–2249 (1996).
15. R. W. Preisendorfer and J. E. Tyler, "The measurement of light in natural waters; radiometric concepts and optical properties," Technical Report (Scripps Institution of Oceanography, 1958) .
16. N. G. Jerlov, *Marine Optics* (Elsevier, 1976).
17. P. Diehl and H. Haardt, "Measurement of the spectral attenuation to support biological-research in a plankton tube experiment," Oceanologica Acta **3**, 89–96 (1980).
18. D. Stramski, "Refractive index of planktonic cells as a measure of cellular carbon and chlorophyll a content," Deep Sea Res. I **46**, 335–351 (1999).
19. J. L. Mueller, "The influence of phytoplankton on ocean color spectra," Ph.D. thesis, (Oregon State University, 1973).
20. A. Bricaud, J. R. V. Zaneveld, and J. C. Kitchen, "Backscattering efficiency of coccolithophorids: use of a three-layered sphere model," Proc. SPIE **1750**, 27–33 (1992).
21. H. R. Gordon, O. B. Brown, R. H. Evans, J. W. Brown, R. C. Smith, K. S. Baker, and D. K. Clark, "A semianalytic radiance model of ocean color," J. Geophys. Res. Atmos. **93**, 10909–10924 (1988).
22. E. Boss, M. Picheral, T. Leeuw, A. Chase, E. Karsenti, G. Gorsky, L. Taylor, W. Slade, J. Ras, and H. Claustre, "The characteristics of particulate absorption, scattering and attenuation coefficients in the surface ocean; contribution of the Tara oceans expedition," Methods Oceanogr. **7**, 52–62 (2013).
23. W. H. Slade, E. Boss, G. Dall'Olmo, M. R. Langner, J. Loftin, M. J. Behrenfeld, C. Roesler, and T. K. Westberry, "Underway and moored

methods for improving accuracy in measurement of spectral particulate absorption and attenuation," J. Atmos. Oceanic Technol. **27**, 1733–1746 (2010).
24. J. R. V. Zaneveld, J. C. Kitchen, and C. C. Moore, "Scattering error correction of reflection-tube absorption meters," Proc. SPIE **2258**, 44–55 (1994).
25. E. Boss, W. H. Slade, M. Behrenfeld, and G. Dall'Olmo, "Acceptance angle effects on the beam attenuation in the ocean," Opt. Express **17**, 1535–1550 (2009).
26. S. B. Hooker, C. R. McClain, J. K. Firestone, T. L. Westphal, E.-N. Yeh, Y. Ge, and E. R. Firestone, "The SeaWiFS bio-optical archive and storage system (SeaBASS), part 1," Technical Report TM-104566 (National Aeronautics and Space Administration, Goddard Space Flight Center, 1994).
27. P. J. Werdell and S. Bailey, "The SeaWiFS bio-optical archive and storage system (Seabass): current architecture and implementation," Technical Report Tech. Memo 2002-211617 (NASA Goddard Space Flight Center, 2002).
28. P. J. Werdell, S. Bailey, G. Fargion, C. Pietras, K. Knobelspiesse, G. Feldman, and C. McClain, "Unique data repository facilitates ocean color satellite validation," EOS Trans. Am. Geophys. Union **84**, 377–387 (2003).
29. L. Duysens, "The flattening of the absorption spectrum of suspensions, as compared to that of solutions," Biochim. Biophys. Acta **19**, 1–12 (1956).
30. J. Kirk, "A theoretical analysis of the contribution of algal cells to the attenuation of light within natural waters I. General treatment of suspensions of pigmented cells," New Phytol. **75**, 11–20 (1975).
31. C. S. Roesler and A. H. Barnard, "Optical proxy for phytoplankton biomass in the absence of photophysiology: rethinking the absorption line height," Methods Oceanogr. **7**, 79–94 (2013).
32. R. J. Brewin, G. Dall'Olmo, S. Pardo, V. van Dongen-Vogels, and E. S. Boss, "Underway spectrophotometry along the Atlantic Meridional transect reveals high performance in satellite chlorophyll retrievals," Remote Sens. Environ. **183**, 82–97 (2016).
33. W. Cleveland, "Robust locally weighted regression and smoothing scatterplots," J. Am. Stat. Assoc. **74**, 829–836 (1979).
34. G. E. Fogg, "Some comments on picoplankton and its importance in the pelagic ecosystem," Aquat. Microb. Ecol. **9**, 33–39 (1995).
35. A. M. Ciotti, M. R. Lewis, and J. J. Cullen, "Assessment of the relationships between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient," Limnol. Oceanogr. **47**, 404–417 (2002).
36. R. Margalef, "Life-forms of phytoplankton as survival alternatives in an unstable environment," Oceanol. Acta **1**, 493–509 (1978).
37. A. Bricaud, H. Claustre, J. Ras, and K. Oubelkheir, "Natural variability of phytoplanktonic absorption in oceanic waters: Influence of the size structure of algal populations," J. Geophys. Res. Oceans **109**, C11010 (2004).
38. W. Clavano, E. Boss, and L. Karp-Boss, "Inherent optical properties of non-spherical marine-like particles–from theory to observation," Oceanogr. Marine Biol. **45**, 1–38 (2007).
39. A. Bricaud and A. Morel, "Light attenuation and scattering by phytoplanktonic cells: a theoretical modeling," Appl. Opt. **25**, 571–580 (1986).
40. J. C. Kitchen and J. R. V. Zaneveld, "A three-layered sphere model of the optical properties of phytoplankton," Limnol. Oceanogr. **37**, 1680–1690 (1992).
41. E. Boss, M. S. Twardowski, and S. Herring, "Shape of the particulate beam attenuation spectrum and its inversion to obtain the shape of the particulate size distribution," Appl. Opt. **40**, 4885–4893 (2001).
42. O. A. Mikkelsen, W. H. Slade, T. Leeuw, D. R. Dana, and C. Pottsmith, "A new submersible instrument for measuring hyperspectral backscattering in natural waters (hyper-bb)," in *Ocean Sciences Meeting* (AGU, 2020).
43. S. Graban, S. Goult, R. Sauzede, and G. Dall'Olmo, "Accurate deep-learning estimation of chlorophyll-a concentration from the spectral particulate beam-attenuation coefficient," Opt. Express (Submitted) .