California State University, Monterey Bay

Digital Commons @ CSUMB

2024

Testing a Hyperspectral, Bio-Optical Approach to Identification of Phytoplankton Community Composition in the Chesapeake Bay Estuary

S. M. McKibben

S. Schollaert Uz

Sherry L. Palacios

Author Contributions:

Conceptualization: S. M. McKibben, S. Schollaert Uz, Sherry L. Palacios
Data curation: Sherry L. Palacios
Formal analysis: S. M. McKibben
Funding acquisition: S. Schollaert Uz
Investigation: S. M. McKibben, S. Schollaert Uz
Methodology: S. M. McKibben, Sherry L. Palacios
Project administration: S. Schollaert Uz

Testing a Hyperspectral, Bio-Optical Approach to Identification of Phytoplankton Community Composition in the Chesapeake Bay Estuary

S. M. McKibben[1,2] [ID], **S. Schollaert Uz[1]** [ID], **and Sherry L. Palacios[3]**

[1]NASA Postdoctoral Program/NASA Goddard Space Flight Center, Greenbelt, MD, USA, [2]University of Maryland, College of Computer, Mathematical, and Natural Sciences, College Park, MD, USA, [3]Department of Marine Science, California State University Monterey Bay, Monterey, CA, USA

Abstract The multi-to hyperspectral evolution of satellite ocean color sensors is anticipated to enable satellite-based identification of phytoplankton biodiversity, a key factor in aquatic ecosystem functioning and upper ocean biogeochemistry. In this work the bio-optical Phytoplankton Detection with Optics (PHYDOTax) approach for deriving taxonomic (class-level) phytoplankton community composition (PCC, e.g. diatoms, dinoflagellates) from hyperspectral information is evaluated in the Chesapeake Bay estuary on the U.S. East Coast. PHYDOTax is among relatively few optical-based PCC differentiation approaches available for optically complex waters, but it has not yet been evaluated beyond the California coastal regime where it was developed. Study goals include: (a) testing the approach in a turbid estuary including novel incorporation of colored dissolved organic matter (CDOM) and non-algal particles (NAP), and (b) performance assessment with both synthetic mixture and field data sets. Algorithm skill was robust on synthetic mixtures. Using field data, cryptophyte and/or cyanophyte phytoplankton groups were predicted, but diatom and dinoflagellate detection was not conclusive. For one field data set, small but significant improvements were observed in predicted PCC groups when tested with incorporation of CDOM and NAP into the algorithm, but not for the second field data set. Sensitivity to three hyperspectral-relevant spectral resolutions (1, 5, 10 nm) was low for all field and synthetic data. PHYDOTax can identify some phytoplankton groups in the estuary using hyperspectral, field-collected measurements, but validation-quality data with broad temporospatial coverage are needed to determine whether the approach is robust enough for science applications.

1. Introduction

As the synoptic-scale of multispectral-based ocean color observations from space have revolutionized our understanding of global ocean phytoplankton distribution and related upper ocean processes over the past few decades, a new generation of current and future hyperspectral satellite sensors show potential for another leap forward in remote sensing capabilities by enabling global-scale measurements of phytoplankton diversity (Dierssen et al., 2023; Werdell et al., 2019). Phytoplankton are single-celled, photosynthesizing microalgae ubiquitous in the sunlit upper layer of aquatic environments worldwide, hence directly influence both ocean food webs and global climate. Where chlorophyll-a (chl-a) concentration derived from multispectral satellite information estimates total phytoplankton community abundance and distribution, hyperspectral remote sensing observations may improve the ability to identify which phytoplankton types comprise that total (Cetinić et al., 2024). Partitioning of phytoplankton community composition (PCC) may be based on cell size, ecological function, or taxonomy (Sathyendranath, 2014). Shifts in PCC fractions reflect changes in ecological and environmental factors such as water column nutrient status (Behrenfeld et al., 2021; Glibert, 2016), energy available to higher trophic levels (e.g., how much, and which, fish could be in area), and upper ocean biogeochemistry (e.g., ocean carbon uptake and sequestration). PCC from hyperspectral satellite sensors has the potential to advance our understanding of, and ability to quantify, such processes at a global scale.

Unlike the relatively optically clear open ocean, far from coasts with low phytoplankton concentrations, coastal waters are optically complex. At the interface of the land and sea, multiple types of optically active constituents such as pigments and sediments from terrestrial runoff and mixing processes, in addition to phytoplankton, may be present in coastal waters at high concentrations. Several approaches to optical identification of phytoplankton diversity using hyperspectral information have been demonstrated (Bracher et al., 2017; Chase et al., 2017; Kramer et al., 2022; Mouw et al., 2017; Sathyendranath, 2014), but most are not applicable to coastal regimes.

PHYtoplankton Detection with Optics (PHYDOTax) is a semi-analytical, bio-optical algorithm unique among PCC-discrimination approaches as it was developed for optically complex waters (Palacios, 2012). Initial development in California's Monterey Bay showed that PHYDOTax successfully partitioned PCC proportions, specifically taxonomic, class-level groups, and showed promise for evaluation in other optically complex regimes (Palacios, 2012).

In this work the theoretical framework of PHYDOTax is evaluated in the Chesapeake Bay, informing efforts to develop bio-optical-based PCC monitoring tools for turbid estuarine waters. As a region which hosts diverse and economically important ecosystem services to the Mid-Atlantic United States, including high-value fisheries and recreation which are impacted by phytoplankton-related issues such as eutrophication and harmful algal blooms (HABs; Wolny et al., 2020), this region exemplifies coastal zones worldwide where stakeholders have a need for synoptic PCC data products (Bracher et al., 2017). As a shallow, semi-enclosed estuary draining a 166,534 km^2 watershed, this tidally dynamic bay is optically complex (Aurin et al., 2013; Tzortziou et al., 2006, 2007): its waters include optically active constituents termed colored dissolved organic matter (CDOM) and non-algal particles (NAP; e.g. sediment) that do not covary with phytoplankton abundance and can be present at high concentrations. CDOM and NAP can contribute non-trivially to optical measurements (Mobley et al., 2004), complicating accurate optical-based PCC identification (Mouw et al., 2015). As the first effort to evaluate PHYDOTax in an optically complex estuary and to incorporate CDOM and NAP, this work tests the adaptive capabilities of PHYDOTax to a new water body.

1.1. Algorithm Background

Two terms are central to the PHYDOTax approach. First, spectral remote sensing reflectance ($R_{rs}(\lambda)$) as described in this study is a measurement of "ocean color," or light emitted from the ocean in the visible region of the electromagnetic spectrum, and is applied as a descriptor of the optically active constituents within a water body. $R_{rs}(\lambda)$, referred to hereafter as Rrs, is derived from water leaving radiance measured just above the sea surface, $L_w(\lambda)$. Depending on measurement approach, $L_w(\lambda)$ may be measured directly or extrapolated from upwelling radiance, $L_u(\lambda)$, which is instead measured just below the surface. $L_w(\lambda)$ is then normalized to downwelling irradiance, $E_d(\lambda)$, measured just above the sea surface yielding the general Equation 1.

$$Rrs(\lambda) = \frac{Lw(\lambda)}{Ed(\lambda)} \tag{1}$$

Second, the Moore-Penrose inverse is a commonly used approach to solve a system of linear equations that lacks a solution. It is distinctive among other generalized inverses (also called pseudo-inverses) of matrices in that it yields a least squares best fit solution for any finite (square or rectangular) matrix of real or complex numbers (Ben-Israel & Greville, 2003). Required inputs are an "unknown" matrix to be solved and "known" matrix of components that are anticipated to comprise the unknown. As an over determined solution the known matrix must have more wavelengths (rows) than spectral components (columns).

2. Data and Methods

2.1. PHYDOTax: Taxonomic PCC From Remote Sensing Reflectance

The PHYDOTax framework consists of the following as summarized from Palacios, 2012 and visually illustrated in Figure S1 in Supporting Information S1. The "unknown" that PCC is derived from is an Rrs spectrum that can be radiometrically measured at a discrete location via handheld, shipboard, or moored platforms, or retrieved by airborne or satellite sensors. The "known" is a library of individual, modeled Rrs spectra that collectively represent the phytoplankton classes and non-algal constituents, that is, the bulk optical components and not specific genera or species, that are expected to be in the water parcel represented by the Rrs measurement. The known and unknown are normalized and subset (described in next paragraph) to yield the signature library matrix, ***S***, and unknown matrix, ***u***, respectively. Matrix algebra then decomposes ***u*** into constituent parts as represented by ***S*** according to Equations 2–4.

The algorithm can be described as the solution for a matrix of coefficients, ***m*** (Equation 2):

$$\boldsymbol{u} = \boldsymbol{S} \cdot \boldsymbol{m} \tag{2}$$

Table 1
Constituents of (a) PHYDOTax Rrs Library and (b) CHEMTAX Input Matrices Grouped by Classes and Non Algal Types (far Left Column)

Component Groups	(a) PHYDOTax remote sensing reflectance library components — Genus, species of phytoplankton or type of non-algal component	(b) CHEMTAX components and identifying pigments — Pigment grouping type[e]	Diagnostic pigment[d,e]	Characteristic pigment suites[d, e]
Diatom	*Thalissiosira* sp.[a]	DIAT-1	–	Chl-a, Chl-c1, Chl-c2[f], Fucoxanthin (Fuco)
	Skeletonema sp.[a]			
Dinoflagellates	*Alexandrium catenella*[a]	DINO-1	Peridinin (Peri)	Chl-a, Chl-c2[f]
	Ceratium sp.[a]			
Cryptophyte	*Rhodomonas* sp.[a, b]	CRYPTO	Alloxanthin (Allo)	Chl-a, Chl-c2[f]
	C. fragilaropsis[a]			
Cyanophyte, (common name cyanobacteria, blue-green algae)	*Synechococcus* sp.[a]	CYANO-2	–	Chl-a, Zeaxanthin (Zea)
	Generic phycocyanin-rich pico-phytoplaknton (average of Synechocystis & Anacystis marina)[a,b]			
Haptophyte, (common name golden brown algae)	*Isochrysis galbana*[a]	HAPTO-8	–	Chl-a, Chl-c3
	Pavlova lutheri[a,b]			
Chlorophyte, (common name green algae)	*Dunaliella* sp.[a]	CHLORO-1	–	Total Chl-b, Zeaxanthin, Lutein (Lut)
Prasinophyte, (green algae)	–	PRASINO-3	Prasinoxanthin (Pras)	Total Chl-b, Lutein, Chl-a
Non algal components	*Colored dissolved organic matter (CDOM)*[c]	–	–	–
	Non algal particles (NAP), brown earth[c]	–	–	–

Note. Column 1 of (b) is CHEMTAX pigment grouping types from Jeffrey et al., 2012: diatoms (DIAT-1), dinoflagellates (DINO-1), cryptophytes (CRYPTO), cyanobacteria (CYANO-2), chlorophytes (CHLORO-1), prasinophytes (PRAS-3), haptophytes (HAPTO-8). Columns 2 and 3 in (b) show the Diagnostic Pigments and Characteristic Pigments Suites applied as identifiers (Jeffrey et al., 2012, Wright & Jeffrey, 2006a, 2006b) for those groups. Chl-a, -b, -c1, and -c2 are abbreviations for chlorophylls a, b, c1, and c2. See also Text S1 in Supporting Information S1. [a](Palacios, 2012). [b](Stramski et al., 2001) cultures and IOPs. [c](Dierssen et al., 2006). [d](Jeffrey, 2006). [e](Jeffrey, 2012). [f]Available in field pigment data as combined Chl-c1 + Chl-c2.

Solving for m re-arranges the equation such that m is the dot product of the inverse matrix of S and u (Equation 3):

$$m = S^{-1} \cdot u \tag{3}$$

S is a non-square matrix, so a pseudo-inverse (Moore-Penrose inverse) of S is used (Equation 4):

$$m = S^{+} \cdot u \tag{4}$$

Equation 4 outputs a matrix of coefficients for each of the spectral signatures in S; the sum of the coefficients equals 1. Output coefficients for the individual phytoplankton spectra were first summed for each class. For example, the coefficients of two diatom genera in m were summed to represent a total diatom coefficient. This was repeated for the dinoflagellate, cyanobacteria, cryptophyte, "other" phytoplankton (chlorophyte + haptophyte; a bin of more rarely observed taxa), and binned non algal (CDOM + NAP) components (Table 1a and Figures 1a–1g). Each class-level coefficient was then normalized to the sum of all phytoplankton coefficients, yielding PCC proportions for each phytoplankton group (diatoms, dinoflagellates, cryptophytes, cyanophytes, and "other" phytoplankton).

PHYDOTax algorithm settings include the number and type of signatures in S as well as how S and u are normalized and subset prior to application of the algorithm. Normalization is an approach to standardizing the variability in spectral magnitude attributed to different concentrations of optically active constituents for inter-comparison. Subsetting, or choosing a minimum, maximum, and spectral resolution value for the wavelengths of, S and u defines the wavelengths at which the Rrs values are input into the algorithm (see also Section 2.2.3).
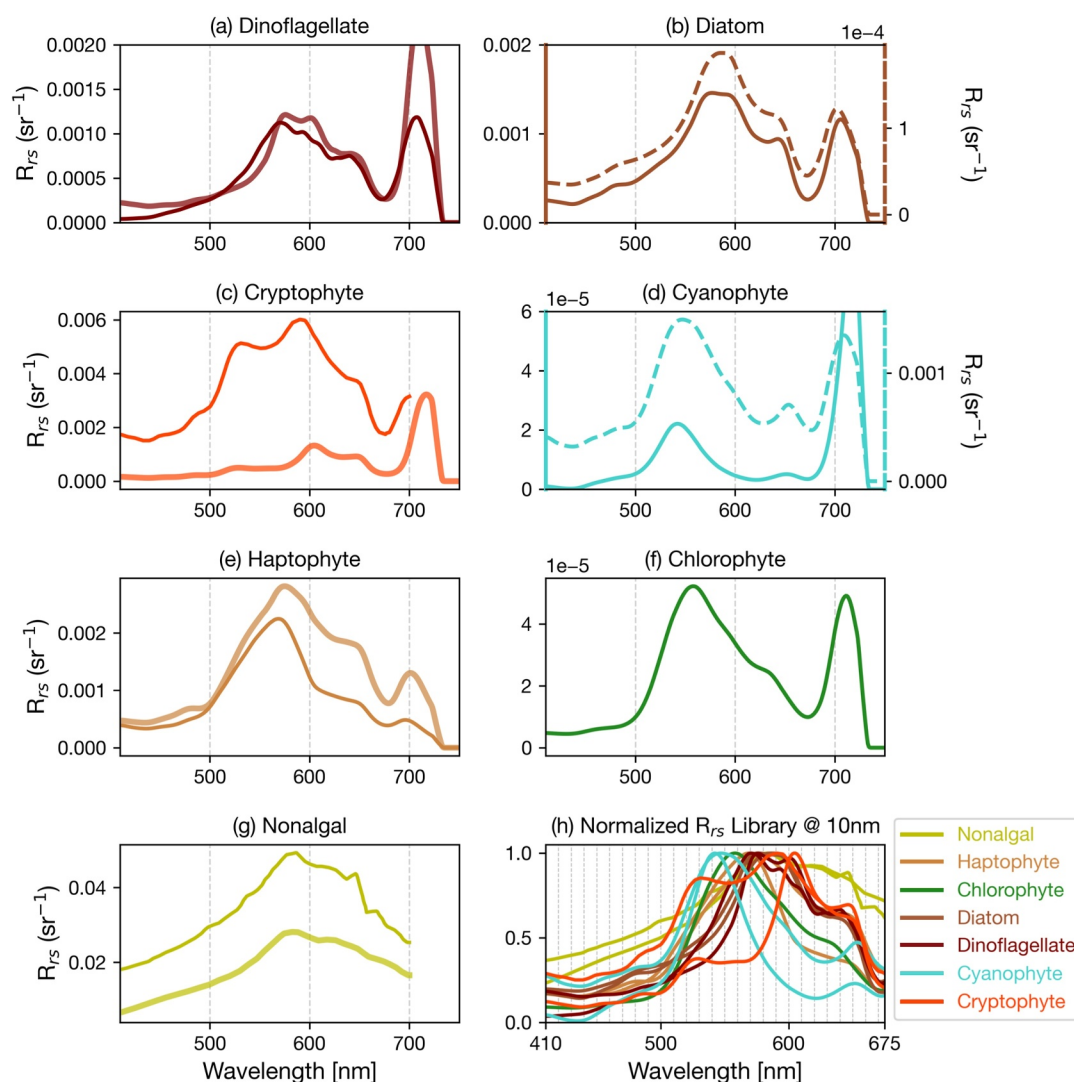
Figure 1. Spectral remote sensing reflectance (Rrs) signature library. Includes phytoplankton groups (a) dinoflagellates, (b) diatoms, (c) cryptophytes, (d) cyanophytes, (e) haptophytes, and (f) chlorophytes, and (g) the non-algal components colored dissolved organic matter (CDOM, bold line) and non-algal particles (NAP, thin line). Line colors are groupings by class as shown in legend. Line styles indicate different types within each group (Table 1a). Y-axes are units of Rrs for plots (a–g). Spectra in panels (b, d) with dashed lines have values on a secondary, dashed y-axis. Plot (h) is an example of the signature library components (a–g) after being normalized and subset for PHYDOTax runs at 10 nm spectral resolution from 410 to 675 nm. Plot (h) y-axis is Rrs normalized to the spectral maxima below 675 nm. Vertical lines illustrate the 10 nm increments input into algorithm; this was varied at 1 and 5 nm for spectral sensitivity testing (Table 2; Figure 5, Figures S2 and S3 in Supporting Information S1).

Rrs spectral library: The list of spectra within the Rrs signature library (Table 1a, Figure 1) were determined according to: (a) Rrs signatures available from previous work; and (b) the bulk, optically active groups (e.g., phytoplankton classes, non-algal constituents) that were anticipated to be present in the water parcel being measured with Rrs (Table 1a, first column). The Rrs signatures were modeled for each phytoplankton type from lab-measured inherent optical properties (IOPs) of monocultures, or modeled IOPs when lab-based IOPs were not available, using Hydrolight™ software as described in Palacios, 2012; Stramski et al., 2001. CDOM and NAP Rrs signatures were also modeled with Hydrolight™ from their IOPs (Dierssen et al., 2006).
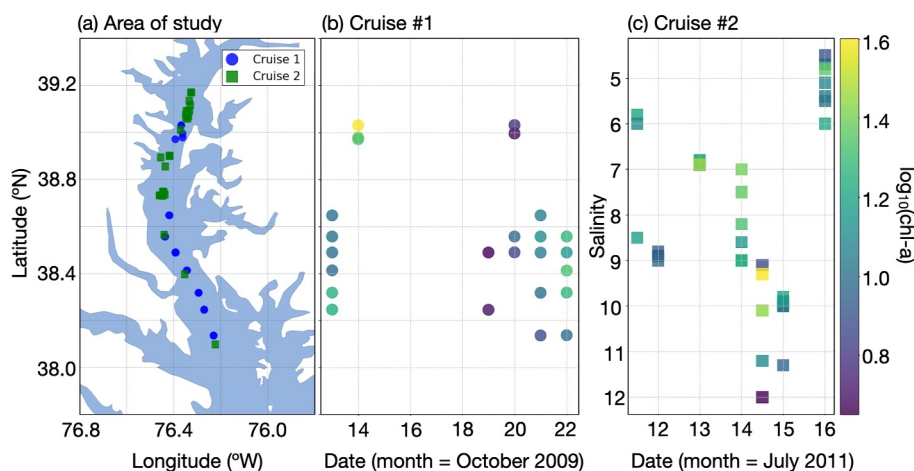
Figure 2. Area of study. (a) Locations of ship sampling in lower oligohaline and upper mesohaline regions of the mainstem Chesapeake Bay. Blue circles and green square symbols indicate Cruises 1 and 2, respectively. Plot (b) shows the latitude of each Cruise 1 sample collected (y axis) during 6 dates in October 2009 (x axis). Salinity values were not available for these points, but latitude from north to south, from the primary river input toward the Atlantic Ocean, is a rough approximation of lower to higher salinity. Plot (c) shows Cruise 2 salinity values for each sample collected (y axis) during 7 dates (x axis) in July of 2011. Colors in panels (b, c) indicate log10(chla).

2.2. Data Sets

2.2.1. Synthetic Remote Sensing Reflectance Mixtures for Algorithm Testing

Expanding on previous methods for testing PHYDOTax on synthetic mixtures (Palacios, 2012), synthetic mixtures in this work were composed to investigate algorithm skill given differing: (a) spectral resolution and (b) number of constituents in synthetic Rrs mixtures. The number, c, of constituents, that is, Rrs signatures, used to create the mixtures differentiated the data sets where c ranged from 2 to 18 in increments of two for a total of nine data sets. Individual Rrs mixtures were created by randomly choosing c Rrs signatures, normalizing each to its maximum value, then summing them. As such the proportional composition of each mixture was known and recorded for statistical testing. This was repeated 500 times creating 500 mixtures in each of the nine data sets. Rrs signatures from Table 1 plus 7 additional, modeled and approximated, phytoplankton and non algal signatures (Table S1 in Supporting Information S1) were used to create the mixtures.

2.2.2. Field Data and Area of Study

Chesapeake Bay field data were downloaded from NASA's SeaWiFS Bio-optical Archive and Storage System (SeaBASS, https://seabass.gsfc.nasa.gov), a standardized system for archiving and distributing in situ data collected during satellite ocean color field validation efforts (Werdell et al., 2003). All instances of time- and location-coincident, surface-collected, shipboard Rrs and phytoplankton pigment concentration data from the Chesapeake Bay were downloaded. Only samples from the mainstem bay region, not the tidal tributaries, were included. The mainstem spans ~320 km north to south from its primary freshwater input, the Susquehanna River, to its marine endmember, the Atlantic Ocean. The tidal tributaries are the tidally-influenced portions of the rivers entering the Bay. Data from 2 cruises detailed below fulfilled the requirements above.

Shipboard Rrs and pigment data: Cruise 1 and 2 sampling covered the mid to northern mainstem of the Chesapeake Bay (Figure 2a) during October 13–22, 2009 and July 11–16, 2011, respectively (Figures 2b and 2c), with multiple sampling efforts per day. Cruise 1 was conducted by the Naval Research Lab (SeaBASS, 2009). Cruise 2 (SeaBASS, 2011) was the shipboard component of broad field efforts in support of NASA's GEOstationary Coastal and Air Pollution Events (GEO-CAPE) mission (Fishman et al., 2012). Phytoplankton pigments were collected by filtering a known volume of water onto glass fiber filters, which were immediately stored in aluminum foil pouches and placed in liquid nitrogen aboard the ship, then stored at −80°C in the lab until analysis. Pigment concentrations were determined with high performance liquid chromatography following Van Heukelem & Thomas, 2001; Hooker et al., 2005. Shipboard Rrs collection and processing methods for Cruise 1 data were not

available (SeaBASS, 2009). For Cruise 2, radiometric measurements were collected with a custom-built, hand-held, above water SPECTRIX SPX300 spectrometer equipped with an extension tube permitting a 10° field-of-view. These data were collected, processed, and glint-corrected to yield Rrs following methods in (Lee et al., 1997, 1999, Zhang et al., 2018), and the data repository documentation (SeaBASS, 2011).

CHEMTAX, Taxonomic PCC from pigment concentrations: CHEMical TAXonomy, or CHEMTAX, was downloaded from the Australian Antarctic Data Center (Wright, 2017). This software derives the percentage of each phytoplankton class present using identifying ratios of marker carotenoid pigment concentrations to total chl-a concentration, termed here marker:chla ratios (Mackey et al., 1996). Briefly, CHEMTAX methods include first building an input matrix of marker:chl-a ratios for the phytoplankton classes anticipated to be present in the sample (classes and identifying pigments in Table 1b), optimization of this matrix, then factorization of the pigment data. Resulting class-level proportions are reported as percent of total chl-a and used to evaluate PHYDOTax-based PCC proportions. Detailed CHEMTAX methods and input matrices reported in Text S1 and Table S2 in Supporting Information S1.

2.2.3. Algorithm Performance and Sensitivity Analysis

Algorithm runs: PHYDOTax was run by applying the settings in Section 2.1 to synthetic and field data sets, then output was tested against either the (a) PCC proportions of the synthetic Rrs mixtures, or (b) PCC proportions determined by CHEMTAX from coincident, in-water pigment concentrations, respectively. The spectral range and resolution for each Rrs spectra in both the library, S, and the unknown measurement, u, were subset from 410 to 675 nm and normalized to their individual maximum value within that range. Sensitivity to spectral resolution was evaluated by running PHYDOTax at 3 settings relevant to current and planned hyperspectral NASA missions: 1 nm, the highest resolution available from the data; 5 nm, the bandwidth for NASA's Plankton Aerosol Cloud ocean Ecosystem (PACE) Ocean Color Imager (Werdell et al., 2019); and 10 nm, the anticipated resolution of both the optical spectrometer on NASA's upcoming Surface Biology and Geology mission (SBG; Stavros et al., 2023) and the selected Earth Venture instrument, the Geosynchronous Littoral Imaging and Monitoring Radiometer (GLIMR, https://eos.unh.edu/glimr/about). For field data, runs were conducted with and without CDOM and NAP signature input to test sensitivity to their inclusion.

Statistical tests for synthetic mixture data sets: PHYDOTax-predicted PCC proportions (y) versus known proportions used to create each synthetic mixture (x) were tested. As output proportion values within each data set (500 in each) were normally distributed, Pearson's test was used to describe PHYDOTax performance. A critical value of $p < 0.05$ was set for rejecting the null hypothesis that x and y are not related and accepting the alternative hypothesis they are significantly correlated.

Statistical tests for field data sets: Proportions for the four phytoplankton classes most frequently observed in the mainstem Chesapeake Bay and typically responsible for the majority of bloom biomass (Adolf et al., 2006; Harding et al., 2015) were the focus of analysis: diatoms, dinoflagellates, cryptophytes, and cyanobacteria. More rarely observed chlorophytes, prasinophytes, and haptophytes (Table 1a) were pooled and termed "other phytoplankton." The output non algal coefficients could not be evaluated as sufficient field measurements were not available. PHYDOTax-predicted PCC proportions (y) versus in water PCC (x) were tested for the four groups with Spearman Rank Correlation (SRC), a non-parametric analog of Pearson's correlation that is appropriate for testing these non-normally distributed, continuous data. Here the SRC coefficient, ρ_s, is applied to describe whether the populations x and y are monotonically increasing together. This occurs when larger x values are on average associated with larger y values, or positively related ($0 < \rho_s \leq 1$). Otherwise, if $-1 \leq \rho_s < 0$ values were monotonically decreasing together, or negatively related, and not a successful prediction. A *p value* of <0.05 was set for rejecting the null hypothesis that x and y are not monotonically increasing (H_0: when $\rho_s \leq 0$) and accepting the alternative hypothesis that they are (H_1: $\rho_s > 0$). Due to small sample size, critical values for ρ_s were also set as a function of n samples for each data set. The following ρ_s values yielded 80% power, where power is the percent likelihood that the SRC test is detecting an effect when one actually exists, for a SRC two tailed t-test with a significance level of $p < 0.05$ (May & Looney, 2020): $\rho_s > 0.55$ for Cruise 1 ($n = 25$); and: $\rho_s > 0.65$ for Cruise 2 ($n = 32$). For example, if a SRC test of PHYDOTax-predicted diatom proportions versus the in-water proportions yielded both $\rho_s > 0.65$ *and* $p < 0.05$, the relationship would be significantly, positively related and a successful PHYDOTax result, but a result with $\rho_s < 0.65$ and $p < 0.05$ would not.

Field PCC results above were further interpreted using expected seasonal averages and regression plots. For both data sets, the mean, standard deviation and standard error were calculated and then compared to literature-based phenological PCC for the region. This served as a first order test of whether the CHEMTAX and PHYDOTax PCC fractions were within anticipated ranges for each cruise. Regression plots of the data that were tested with SRC visually informed performance differences with and without CDOM and NAP.

2.3. Evaluative PCC Maps Derived From Hyperspectral Remote Sensing

Airborne data: Hyperspectral data were collected via an Airborne Compact Atmospheric Mapper (ACAM) on board a NASA UC-12 aircraft on 20 July 2011. Airborne Compact Atmospheric Mapper data collection and processing are fully described in (Zhang et al., 2018) and references therein. Briefly, the ACAM setup included the Ocean Color (OC, 460–900 nm range) and Air Quality (AQ, 300–520 nm) spectrometers. Collection was at a flight altitude of ~8.5 km with ~0.75 km cross track and ~0.5 km along-track spatial resolution. Atmospheric correction is described in (Zhang et al., 2018) and was conducted using the traditional "black pixel" approach (Gordon & Wang, 1994) with two adjustments. First, due to the low altitude of the plane, radiative transfer-based look-up-tables (LUTs) customized for the Chesapeake Bay were created in place of the standard LUTs used for satellite or higher altitude airborne imagery. Second, non-zero Rrs observed in the near-infrared (NIR) due to turbid Bay waters was addressed with an iterative, 2-band NIR (750 and 865 nm) approach (Stumpf et al., 2003). PHYDOTax was run on this swath at 10 nm spectral resolution.

Two satellite-based physical conditions were also obtained for 20 July 2011 in the northern Chesapeake: sea surface temperature (SST) and K_d (490), the diffuse attenuation coefficient for downwelling irradiance at 490 nm in m^{-1}, a metric of water clarity. Multi-scale Ultra-high Resolution (MUR) SST data (v04.1, Global, 0.01° resolution) were downloaded from NOAA Coastwatch ERDDAP (https://coastwatch.pfeg.noaa.gov/erddap). This product is created by fusing and interpolating SST data from multiple platforms each day (e.g., nighttime infrared & microwave-based satellite SST, in situ SST; Chin et al., 2017). K_d (490) data were downloaded from NOAA CoastWatch East Coast Node (https://www.star.nesdis.noaa.gov/pub/socd1/ecn/data/modis/k490noaa/daily/cd/) and was derived from MODIS Aqua 1 km resolution data according to methods of (Wang et al., 2009).

3. Results and Discussion

3.1. Algorithm Performance on Synthetic Mixtures

Testing algorithm skill given a variety of controlled data sets showed robust performance. All optical groups of interest were predicted in all runs (Figure 3, $p << 0.05$). Cryptophyte and cyanophyte groups were detected with the highest accuracy, irrespective of the number of components, c, in the Rrs mixtures suggesting their spectral shapes are distinguished better than the others, discussed further in Section 3.3.2. Generally, the more complex the mixtures ($c > 8$), the lower the correlation coefficients for that data set. Skill at differentiating diatom and haptophyte groups was the lowest of all runs when $c > 8$ (Figure 3, darker blue regions), perhaps due to their similar spectral shape (Figures 1b and 1e). Little sensitivity to spectral resolution is evident (Figure 3, gray bar on x axis), suggesting PHYDOTax functions well at all hyperspectral resolutions on these data.

Caution is necessary when interpreting these results relative to the algorithm test results on field data below. While synthetic and field PHYDOTax runs use the same input library, algorithm, and settings, their data sets are not related. Synthetic Rrs mixtures are not intended to approximate a full radiative transfer model of the various optical constituents in the Chesapeake Bay or approximate algorithm performance in the field. Rather these tests inform algorithm skill given differing spectral resolution and mixture complexities, aiding interpretation of field-based results.

3.2. Customization of PHYDOTax for Application to Field Data

A comparison of the methods in the Monterey Bay study (MBS; Palacios, 2012) and this Chesapeake Bay study (CBS) illustrates customizing PHYDOTax settings based on region, study question, and characteristics of the data sets being applied. For the MBS, S included 13 phytoplankton spectral signatures. Most phytoplankton types in the spectral library are cosmopolitan genera or species. Two signatures represented bloom-forming phytoplankton types associated with HABs in the region (*Pseudo-nitzschia spp., Akashiwo sanguinea;* Palacios, 2012). The other phytoplankton signature types were determined by data availability, specifically which species were
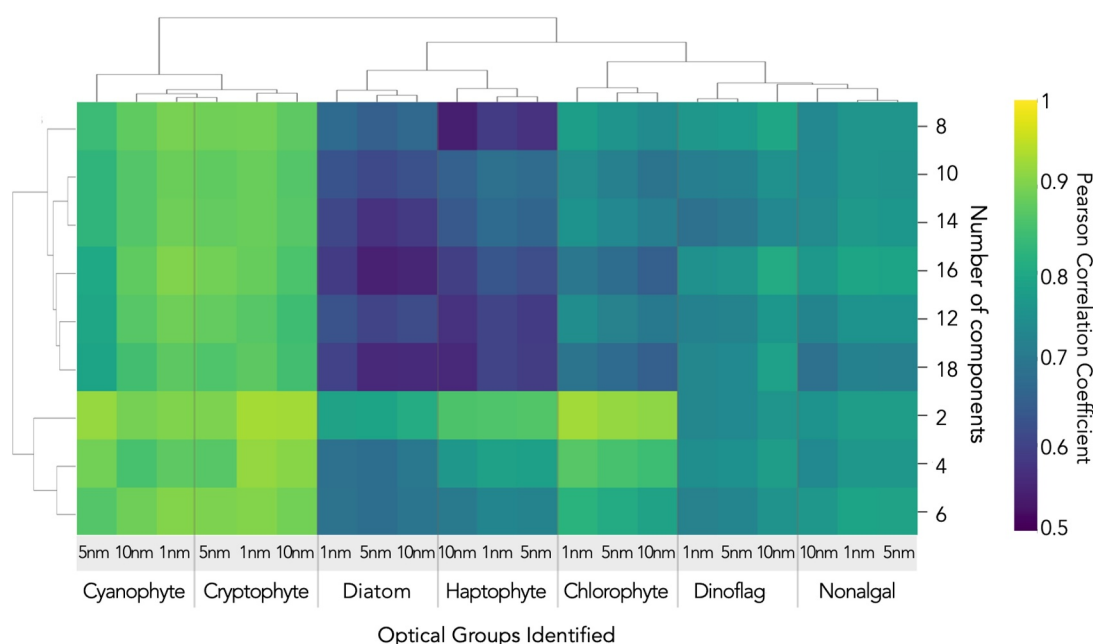
Figure 3. Results of PHYDOTax runs on synthetic Rrs mixtures. Clustered heatmap of Pearson correlation coefficients (colors) demonstrate significant positive relationships for all classes (grouped columns) in all runs with varying numbers of components (rows) and all resolution settings (individual columns); $p << 0.05$ and $n = 500$ for each run (i.e., each square): all runs were successful. Dendrograms at top and left side show hierarchical clustering of results according to similarity. Please note the clustering defined the order of all rows and columns. Axes display how run settings differed: **X axis**, spectral resolution in nanometers (1, 5, and 10 nm, gray bar) and name of optical group identified (diatom, cyanophyte, etc); **Y axis**, number of optically-active constituents present in synthetic mixtures.

available as lab cultures as well as data available from previous work (Stramski et al., 2001). As this region is a deep, open bay in a productive coastal eastern boundary upwelling regime, CDOM and NAP from terrestrial sources was presumed to be small and not included in S. Both S and u were normalized to the spectral peak of chl-a fluorescence above 673 nm then subset to 455–675 nm at 10 nm spectral resolution. The 410–455 nm spectral region was excluded due to instrument noise in imaging sensor data applied in the MBS (Palacios, 2012; Palacios et al., 2015).

In contrast, for the CBS, S included 11 spectral signatures selected from the MBS library plus two signatures to approximate CDOM and NAP (Figures 1a–1g; Table 1a). Compared to the MBS library, the HAB-related *Pseudo-nitzschia spp.* and *Akashiwo sanguinea* signatures were excluded as they are not known to bloom in high concentrations, or attributable to HAB events, in the Chesapeake. Also excluded was the coccolithophorid haptophyte *Emilliania huxleyi* due to its unique, highly-scattering Rrs signature and no evidence it blooms commonly in the Bay. The remaining 11 Rrs signatures were applied to collectively represent the optical features of the anticipated classes as well as CDOM and NAP in the region, these assumptions discussed in Section 3.4. Spectra in S and u were subset to a larger spectral range than in the MBS, 410–675 nm, and then normalized to their individual maxima as neither CDOM nor NAP spectra exhibit the chl-a peak that was applied to normalizing in the MBS. These similarities and differences of CBS and MBS methods demonstrate adjustment of the algorithm in differing studies and data sets.

3.3. Algorithm Performance on Field Data

PHYDOTax PCC predictions were tested against coincident, in water PCC with field data sets in the northern Chesapeake (Figure 2). Both cruises had similar, log-normally distributed ranges of chl-a values (Figures 2b and 2c: Cruise 1, 3.53–40.0 mg m^{-3}; Cruise 2, 5.38–57.8 mg m^{-3}). PHYDOTax successfully predicted 2 of the 4 phytoplankton classes of interest: cyanophytes and cryptophytes; Table 2. For Cruise 1 runs with CDOM and NAP included in the spectral library, cryptophytes were detected by PHYDOTax at all 3 spectral resolutions (Table 2a, $\rho_s = 0.808$–0.865, $p << 0.05$). With significant p values but correlation coefficients (ρ_s) just below threshold, runs without CDOM and NAP did not quite meet detection criteria (Table 2a, $\rho_s = 0.490$–0.628, $p = 0.002$–1.18×10^{-4}). For Cruise 2, irrespective of CDOM and NAP inclusion, cryptophytes were detected at

Table 2

Spearman Rank Correlation Values for PHYDOTax-Predicted Versus in Water Phytoplankton Community Composition Values at 3 Spectral Resolution Settings (1, 5, and 10 nm) for Cruise 1 (a) and Cruise 2 (b)

| | | 1 nm resolution | | | | 5 nm resolution | | | | 10 nm resolution | | | |
| | | no CDOM, NAP | | with CDOM, NAP | | no CDOM, NAP | | with CDOM, NAP | | no CDOM, NAP | | with CDOM, NAP | |
	n	ρ_s	**p**	ρ_s	**P**	ρ_s	**p**	ρ_s	**p**	ρ_s	**p**	ρ_s	**p**
(a) Cruise #1–Spearman rank correlation values													
Dinoflagellate	25	−0.339	0.951	−0.085	0.657	−0.401	0.977	−0.072	0.634	−7.70E−03	0.515	0.467	9.26E−03
Diatom	25	0.275	0.0920	−0.308	0.933	0.301	0.0720	−0.155	0.770	0.148	0.239	−0.149	0.762
Cryptophyte	25	0.628	3.91E−04	**0.828**	**1.58E−07**	0.490	6.45E−03	**0.865**	**1.18E−08**	0.557	0.002	**0.808**	**5.06E−07**
Cyanophyte	25	−0.418	0.981	−0.161	0.779	−0.422	0.982	−0.00692	0.513	−0.363	0.963	−0.431	0.984
(b) Cruise #2–Spearman rank correlation values													
Dinoflagellate	32	−0.104	0.715	−0.319	0.962	−0.0913	0.690	−0.260	0.924	−0.423	0.992	−0.346	0.974
Diatom	32	0.354	0.0233	−0.110	0.726	0.340	0.0285	0.0433	0.407	0.391	0.0134	−0.107	0.720
Cryptophyte	32	**0.582**	**0.000**	**0.737**	**7.60E−07**	**0.495**	**1.98E−03**	**0.743**	**5.46E−07**	**0.628**	**5.97E−05**	**0.647**	**3.14E−05**
Cyanophyte	32	0.540	7.06E−04	0.540	7.06E−04	0.537	7.65E−04	0.521	1.13E−03	**0.583**	**2.32E−04**	**0.659**	**2.07E−05**

Note. The first column, **n**, is the number of samples. Whether colored dissolved organic matter (CDOM) and non algal particles (NAP) were included in the spectral library is indicated in column headers. The SRC critical value of $p < 0.05$ was set to accept the alternative hypothesis that PHYDOTax-based PCC was monotonically increasing with predicted PCC. To account for the small sample sizes, critical values for the correlation coefficient, ρ_s, were also set as a condition for significance: $\rho_s > 0.65$ (Cruise 1) or $\rho_s > 0.55$ (Cruise 2). Both thresholds are a function of **n** following (May & Looney, 2020). PHYDOTax correctly predicted the phytoplankton group when both p and ρ_s thresholds for significance were met, indicated by bold text. Regression plots for results above are found in Figure 5 (10 nm spectral resolution results), and Figures S2 and S3 in Supporting Information S1 (1 and 5 nm, respectively).

all 3 spectral resolutions (Table 2b, ρ_s = 0.495–0.743, $p \ll 0.05$) as well as cyanophytes at 10 nm resolution (Table 2b, ρ_s = 0.583–0.659, $p \ll 0.05$). Only Cruise 2 shows slight, but significant, spectral sensitivity for cyanophytes with critical values met at 10 nm resolution but ρ_s values were just below the threshold for significance at 1 and 5 nm (Table 2b). PHYDOTax exhibits low sensitivity for certain phytoplankton groups when tested at different spectral resolutions. Sensitivity to non algal library components varied and is discussed in Section 3.4.

For greater statistical confidence given sample sizes, PCC prediction was considered successful if critical values set for *both* the p value and correlation coefficient, ρ_s, were satisfied. Across all spectral resolutions, some cryptophytes in both cruises and all Cruise 2 cyanophyte runs had significant p values alongside positive, but too low to be significant, ρ_s values. Detection of these groups in all runs may have been statistically significant if sample size had been larger, but more data are needed to verify this.

3.3.1. Are PCC Estimates in Cruises 1 & 2 Reasonable Given Their Region/Season?

As both CHEMTAX and PHYDOTax apply an a priori set of input information to derive results, it is prudent to perform an assessment of the output data using independent PCC data from the area to determine whether results fall within a reasonable range. For CHEMTAX it is recommended to compare results against coincident cell count based PCC proportions (Higgins et al., 2012; Lewitus et al., 2005; Wright, 2017). Instances of time-space coincident cell count *and* pigment concentration data from the field, however, are quite rare, including the data sets in this study. Instead, seasonal mean PCC fractions based on long term pigment- and microscopy-based observations in the northern, mainstem Chesapeake Bay (Harding et al., 2015) were used as a measure of expected class-level PCC proportions against which field PCC values were compared. Cruises 1 and 2 were in fall and summer, respectively, which have contrasting seasonal PCC useful for analysis.

Both CHEMTAX- and PHYDOTax-based phytoplankton fractions generally agreed with seasonal expectations (Figure 4, Table S4 in Supporting Information S1). Cruise 1 mean PCC (Figure 4a) was overall similar to expected fall mean PCC (Figure 4b) with diatoms as the largest proportion, cryptophytes second largest, and comparatively low values for dinoflagellates and cyanophytes. For Cruise 2 no single group comprised >40% of the total (Figure 4c), grossly matching the mixed summer assemblages seen in summer (Figure 4d) with an
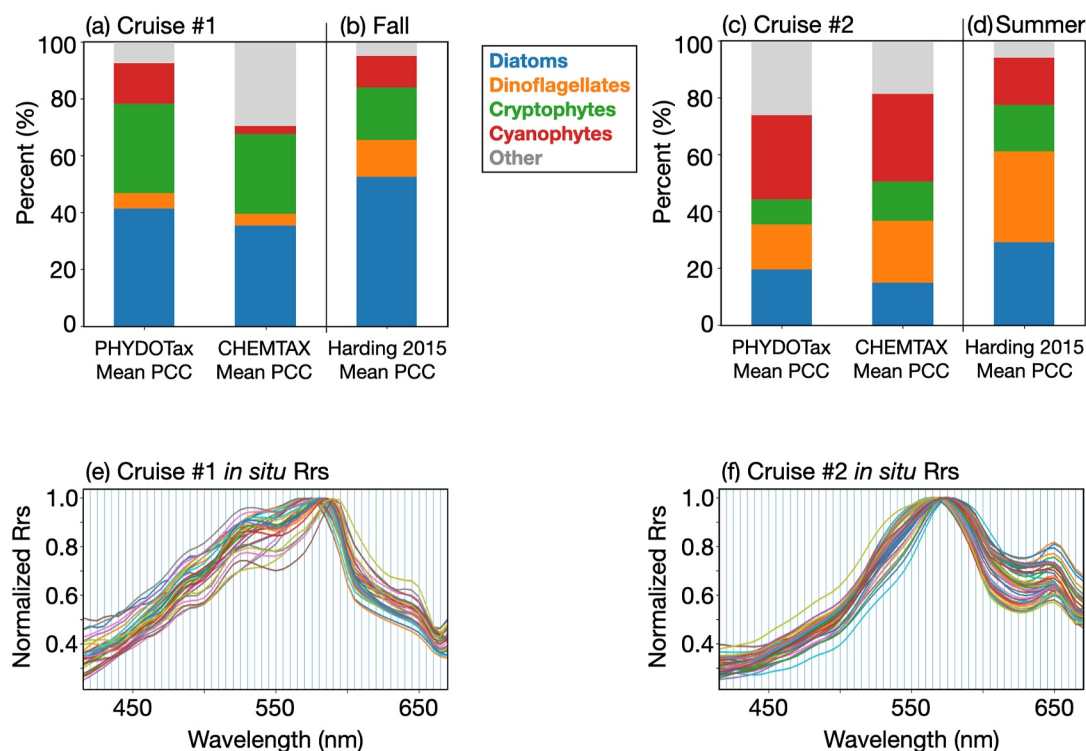
Figure 4. Mean phytoplankton community composition (PCC) for Cruise 1 (a–b) and Cruise 2 (c–d). Subplots (a, c) show the observed PCC fractions measured in the field and are compared to expected, long-term phenological mean PCC for fall (b) & summer (d) from Harding et al., 2015. Number values in panels (a–d) reported in Table S3 in Supporting Information S1. Plots (e, f) show all normalized shipboard Rrs spectra for Cruise 1 and Cruise 2, respectively, at 5 nm resolution.

important difference. In summer dinoflagellate and diatom fractions are, on average, roughly equal and slightly more prevalent than cyanophytes and cryptophytes (Figure 4d), but PCC from Cruise 2 showed higher than average cyanophyte proportions (Figure 4c). This difference is explained by regional and interannual deviation from seasonal means due to hydrology (Adolf et al., 2006; Harding et al., 2015, 2019) and longer-scale variability in climate and nutrient enrichment (Harding et al., 2016) that can cause periods of high cyanobacterial dominance (Adolf et al., 2006; Harding et al., 2015, 2019; Marshall et al., 2005; Paerl et al., 2006). This variability is not captured by the long-term seasonal means (Figure 4d). As there is agreement between the both Cruise 2 field-based PCC fractions (Figure 4c) and numerous cyanophytes were observed during this cruise (Zhang et al., 2018), it is reasonable to conclude that both CHEMTAX and PHYDOTax were producing PCC values representative of the in water conditions during Cruise 2 as well.

3.3.2. Spectral Shape, Phytoplankton Abundance, and Class Discrimination

Rrs data from Cruises 1 and 2 (Figures 4e and 4f) exhibit spectral shapes with absorption features similar to predicted cryptophyte and cyanophyte groups (Figures 2c and 2d). Cryptophytes were detected best across the two cruises (Table 2). Their spectra exhibit 3 unique troughs between ~440 and 550 nm and a maxima shifted toward 600 nm (Figure 1c), features plainly visible in the Cruise 1 Rrs spectra (Figure 4e). Cryptophytes were among the smallest PCC fractions in the mixed Cruise 2 assemblage yet were still successfully predicted (Table 2). Cruise 2 spectra Rrs also exhibit the cryptophyte features below ~550 nm (Figure 4f), albeit at a much smaller magnitude than Cruise 1 (Figure 4e). This spectral shape may permit their detection even when they are present at a lower abundance. For Cruise 2 a unique trough at ~625 nm in cyanophyte spectra (Figure 1d) due to phycocyanin absorption, a pigment characteristic of this group (Lewitus et al., 2005), is also visible in the ship-based Rrs spectra (Figure 4f). Previous findings in the MBS differed, with cryptophyte and cyanophyte concentrations typically very low to zero and not predicted by PHYDOTax (Palacios, 2012). The relatively higher cryptophyte PCC fraction in Cruise 1 and elevated Cruise 2 cyanophytes, plus their unique spectral shapes, likely
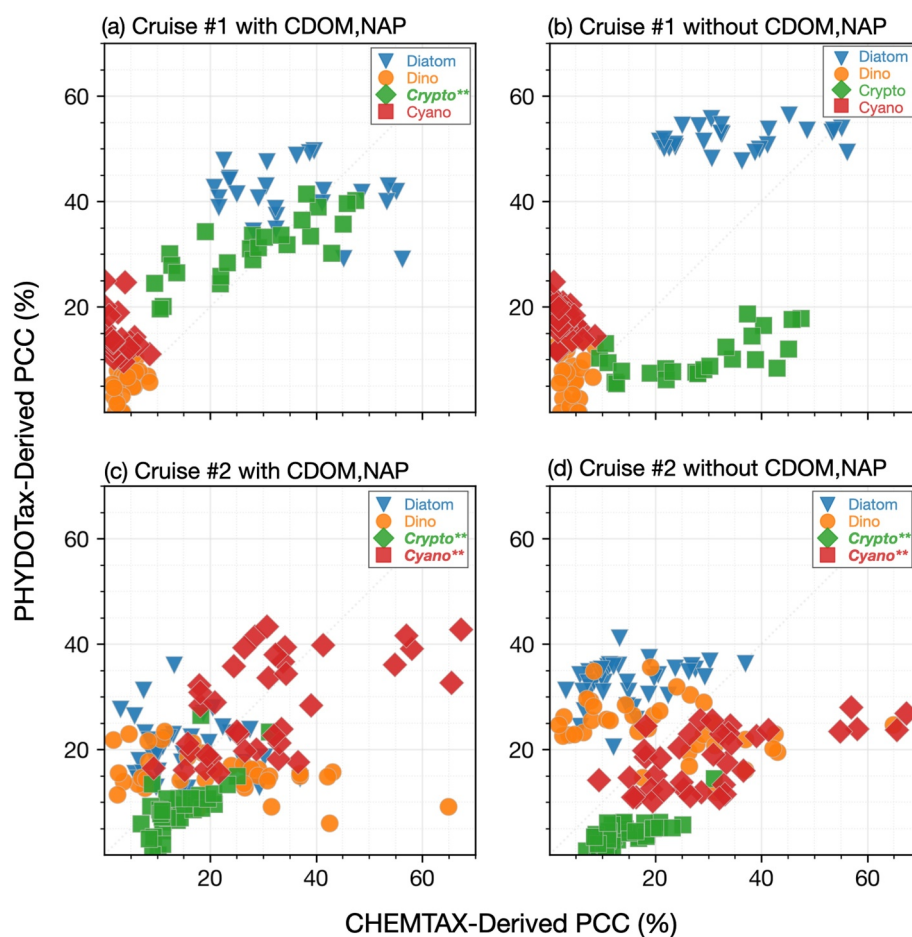
Figure 5. Regressions of PHYDOTax-predicted proportions for each phytoplankton type (*y* axis) versus expected proportions (*x* axis) of total phytoplankton community composition for Cruise 1 (plots a, b) and Cruise 2 (plots c, d). Runs are with (a, c) and without (b, d) colored dissolved organic matter and non-algal particles incorporated in the algorithm. Colors & shapes represent phytoplankton classes according to the legend. Significant results highlighted with bold text and asterisks, see Table 2 for values. Above are results from PHYDOTax runs at 10 nm spectral resolution, see Figures S2 and S3 in Supporting Information S1 for 1 and 5 nm.

permitted successful detection. This is supported by the synthetic tests where these two groups were also predicted the best (Figure 3).

Neither diatom or dinoflagellate proportions were predicted based on SRC testing (Table 2). This is also in contrast to MBS findings demonstrating successful discrimination between them, which was a notable result given their similar brown-green spectral shapes (Figures 1a and 1b). Chl-a concentration also differed with MBS having typically higher values. Total CBS chl-a magnitude ranged from 10^0 to 10^1 mg m^{-3}, while the MBS chl-a varied considerably by year and location of station on the order of 10^0–10^2 mg m^{-3}. The highest chl-a values (10^2 mg m^{-3}) were in extremely dense dinoflagellate blooms. High concentrations of dinoflagellates, as well as diatoms, often dominated MBS PCC. The high abundance, the assumed low levels of CDOM and NAP, and perhaps the specificity of the spectral library applied to the MBS (the Rrs library included at least 1 of the dinoflagellate species that were at high bloom abundances), may explain successful predictions of diatoms and dinoflagellates in MBS, but not CBS, efforts.

3.3.3. Impacts of CDOM & NAP Addition to Spectral Library

Sensitivity to the addition of CDOM and NAP spectral signatures to the Rrs library was higher in Cruise 1 results than Cruise 2 (Figure 5). Predicted diatom and cryptophyte proportions in Cruise 1 increased in spread on the *y* axis with CDOM and NAP (Figure 5a) relative to their omission from runs (Figure 5b). This change significantly

improved cryptophyte detection, but did not for diatoms (Figures 5a and 5b; Table 2a). Regardless of library composition, dinoflagellate and cyanophyte predicted fractions remained low and largely between about 10% and 25% and <10% for the in water, pigment-based PCC fractions. This could be due to their low abundance hence presumably lower impact on the Rrs signal. Statistically, PHYDOTax only predicted cryptophyte fractions for Cruise 1, but diatoms, cyanophytes, and dinoflagellates cluster together in similar ranges (Figure 5a). Given the large differences in the PCC methodologies, and their respective assumptions, uncertainties, and potential for error (Section 3.4), and the similar ranges of the phenological PCC fractions (Figure 5a) suggest that PHYDOTax is predicting reasonable PCC fractions for all 4 groups in Cruise 1 when CDOM and NAP are accounted for (Figure 5a).

In contrast, Cruise 2 results showed small, non-significant differences with differing input library components. Cryptophytes and cyanophytes were identified by PHYDOTax in both Cruise 2 runs, and ρ_s was slightly improved with CDOM and NAP (Figures 5c and 5d, Table 2b). Significance was similar but PCC fractions differed in these runs: PHYDOTax predicted more diatoms and dinoflagellates and less cyanophytes and cryptophytes when CDOM and NAP were omitted (Figures 5c and 5d). Relative to other seasons, summer PCC assemblages are typically mixed, that is, have similar proportions of these 4 groups (Figure 4d) potentially convoluting their spectral fingerprints relative to seasons with 1–2 dominant classes (e.g., Figure 4b). Between Cruises 1 and 2, only Cruise 1 was sensitive to incorporation of CDOM and NAP. This difference may be due to the phenologically-differing PCC proportions (mixed vs. 2 dominant groups) or differences in CDOM and NAP concentration during the two cruises.

3.3.4. PCC From Cruise #2 Airborne Data on 20 July 2011

Airborne hyperspectral imagery shows qualitative spatial patterns of PCC (Figures 6a–6d). This region's spatial heterogeneity is evident (Figures 6a–6f). Three patches with differing PCC can be discerned, with low cyanobacteria in the middle (~38.8°N–39°N) along with higher dinoflagellate fractions and, to a lesser extent, elevated fractions for the other groups. To the north and south of this area is an inverse pattern of higher cyanophyte proportions and lower proportions for other groups (Figures 6a–6d), similar to Cruise 2 mean PCC (Figure 4c). This potential for 3 optically-different water masses is supported by satellite-based sea surface temperature (SST; Figure 6e) and the diffuse attenuation coefficient (K_d(490), Figure 6f). In the north the largest freshwater input into the estuary, the Susquehanna River, enters the bay. Here K_d(490) and SST are highest above a front located at ~39.1°N that is likely due river input (note front location differs as SST is a 1-day, modeled average, while K_d(490) is a swath at a specific time in the tidal cycle). From there K_d(490) and SST are both lower in the middle and lowest in the south.

One pixel of airborne data (Figure 6e, star) had two location-coincident, shipboard measures of Rrs and pigments termed Time A (Figures 7a and 7c) and Time B (Figures 7b and 7d) to compare. In situ and airborne Rrs show similar magnitudes and maxima, but airborne spectra were considerably noisier (Figure 7). Small troughs at ~580, 630, 650, and 690 nm (Figures 7c and 7d) are likely artifacts from sensor calibration and atmospheric correction issues. A discontinuity at ~505–515 nm is due to overlap of differing measurements from the ACAM's two spectrometers (Zhang et al., 2018). Also ACAM Rrs from 400 to 500 nm, a spectral region with larger uncertainty due to atmospheric correction, was overall higher than in the in situ Rrs (Figures 7c and 7d). Features from ~525 to 690 nm (Figures 1a–1g) useful for phytoplankton discrimination and present in ship Rrs are not observed in airborne Rrs (Figure 7). Sensor calibration and atmospheric correction were responsible for the uncertainty in ACAM Rrs (Zhang et al., 2018) and likely the major factor in the differences between airborne and ship based spectral shape. By extension, no match was found between ship- and ACAM-measured PCC proportions. PCC from both Time A and B ACAM spectra (Figures 7c and 7d) predicted diatom fractions roughly twice as large or more than the relative fractions of the other 3 groups (Time A: 33% diatoms, 18% dinoflagellates, 16% cryptophytes, 7% cyanobacteria; Time B: 49% diatoms, 21% cryptophytes, 10% dinoflagellates, and 8% cyanobacteria). Shipboard predicted PCC at Time A (Figure 7a) was evenly distributed, ranging ~13–21%, and Time B (Figure 7b) PCC was also evenly mixed between 16% and 24%, except the cyanophyte fraction at 7%.

PHYDOTax is sensitive to spectral shape; its accuracy depends on atmospheric correction and sensor calibration that preserve phytoplankton absorption and reflectance features (Palacios et al., 2015). Time A and B comparisons indicate the selected pixels of ACAM Rrs data are too noisy to predict PCC. Previous work also postulated that mismatches between PHYDOTax-based airborne PCC and coincident in situ PCC were in part attributable to
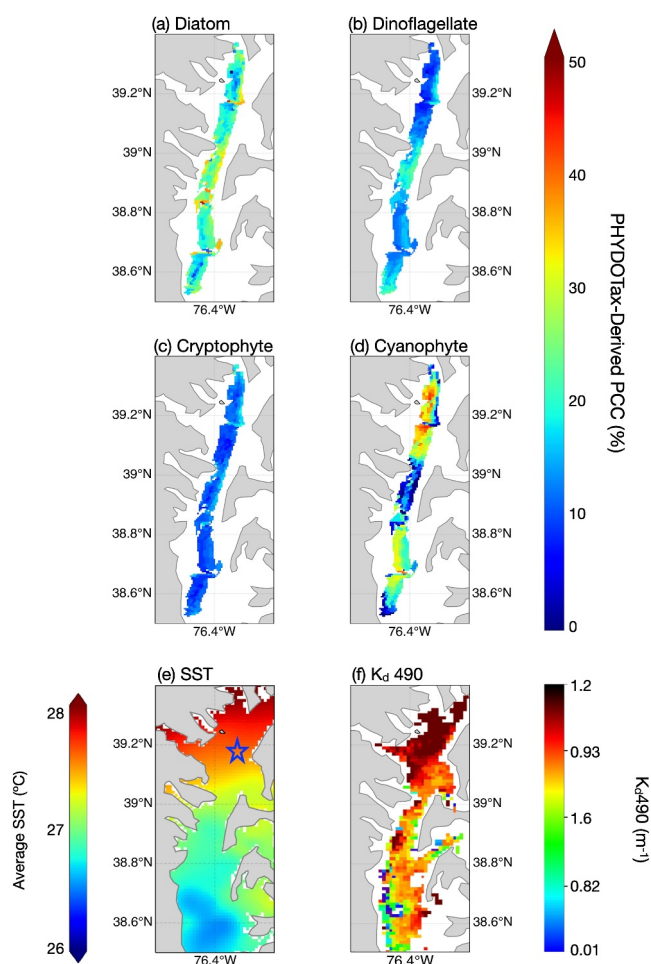
Figure 6. Evaluative maps of phytoplankton community composition from airborne data collected on 20 July 2011 for the northern Chesapeake Bay and satellite-based physical conditions (plots e, f) on 20 July 2011 for qualitative analysis. Maps (a–d) show relative proportions of (a) diatoms, (b) dinoflagellates, (c) cryptophytes, and (d) cyanophytes. For the same date, plot (e) is satellite-based daily average sea surface temperature and plot (f) is the diffuse attenuation coefficient at 490 nanometers ($K_d(490)$), a metric of water clarity. Higher Kd(490) values indicate greater attenuation of light, hence more turbid water. Station location of Times A and B in Figure 7 indicated by a star in panel (e).

atmospheric correction and sensor issues (Palacios et al., 2015). Physical metrics do suggest that PCC patterns were associated with differing spectral shapes of the 3 water masses (Figure 6), but more ship-to-airborne matchups are required to assess to what extent these patterns represent actual in water distribution.

3.4. Assumptions and Uncertainties

Uncertainty is tied to the error in the field and lab data from which PCC fractions are derived. All data matchups between platforms (e.g., ship, airplane, boat) used to test PCC predictions are subject to error due to the temporospatial variability associated with sampling in a shallow, tidally dynamic estuary. Shipboard Rrs- and pigment data collected at the approximate same time and location could be subject to meter-scale patchiness in blooms from tidal and wind dynamics. Patchiness of pigment-based PCC and phytoplankton abundance with depth in this region of the Chesapeake Bay (Keller et al., 2014; Latasa et al., 2022) is also not captured in the shipboard surface data. For airborne data each pixel represents average Rrs over a 3×3 box (~1.5–2 km^2) while ship-based matchups are a small point in within ±1 hr of airborne flyover. These temporospatial differences in coverage contribute to uncertainty in all field PCC estimates.

There is growing consensus that 4–6 degrees of freedom is the limit of optically active components identifiable with hyperspectral Rrs (Cael et al., 2020; Kramer & Siegel, 2019). The higher values are possible at regional, but not global, scales (Kramer & Siegel, 2019), and may be more likely to achieve when semi-analytical algorithms such as PHYDOTax are applied in optically complex waters (Cael et al., 2023). Here 6 total groups were applied: diatoms, dinoflagellates, cyanophytes, cryptophytes, "other" phytoplankton, and non algal components (Table 1), with algorithm performance tested using the 4 most commonly blooming phytoplankton classes. Verifying whether 4–6 degrees of freedom are indeed available requires confirmation of whether sufficient difference exists between the spectral shapes of the individual optical groups (here this means phytoplankton classes, not individual species) to differentiate them from each other given their calculated error (Cael et al., 2020). The uncertainty information needed to answer this question, however, was not available from the data sets applied in this effort.

PCC from spectral shape with PHYDOTax: A key question addressed by algorithm testing is whether lab-measured IOPs of monocultures (single species), and by extension the spectral library of Rrs spectra modeled from them, can represent a taxonomic class and then be applied to predict class-level proportional abundance from Rrs measurements, as PHYDOTax does. Several assumptions to simplify biological and optical complexity are part of this approach. The spectral library is assumed to broadly represent the numerous phytoplankton, NAP (sediment), and CDOM types potentially in the water. To achieve this, output coefficients for each of the several phytoplankton *genera* and *species* are binned into their respective *classes*, a broader taxonomic level than species or genera which are the highest and second highest levels of taxon specificity, respectively. Species level identification is not a goal of PHYDOTax. Rather it assumes that genera/species share class-level physiological similarities which influence the light field, that is, cell wall structure, cell size, suite of accessory pigments, etc. producing broad, common spectral features useful for identification. This can be mostly true for classes such as cyanobacteria, which generally have phycoerythrin, creating their identifying spectral trough at ~625 nm. This assumption, however, may not hold for other classes. Dinoflagellates, for example, have a highly complex lineage and hence considerable variety in potential accessory pigment configurations, cell shapes, and metabolic modes which all can differentially influence spectral shape (Van Heukelem & Hooker, 2012). Also, the matrix factorization approach assumes linearity between the proportion of taxon-specific spectral shapes (a proportion of the
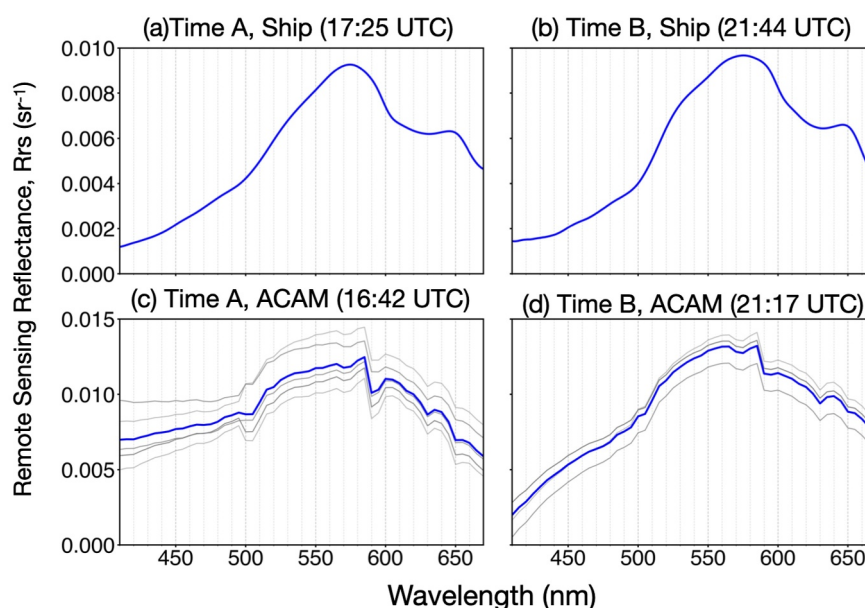
Figure 7. Shipboard remote sensing reflectance (Rrs) collected at Time A, 17:25 UTC (a), and Time B, 21:44 UTC (b), compared to airborne-based Rrs at 16:42 UTC (c) and 21:17 UTC (d), respectively on 20 July 2011. Gray lines in panels (c, d) represent pixel values within ta 3×3 pixel box centered over the ship station. Shipboard sampling and airborne flyovers both occur within ± 1 hr of each other at the location indicated by a star on Figure 6e.

total Rrs signal) and their actual in-water abundance. This is a large simplification compared to the complex, non-linear interaction of light with multiple optically active constituents and the water itself (Mobley, 1994). Even so, results support these assumptions for cryptophytes and cyanophytes in this work as well as dinoflagellates, diatoms, and others in the Monterey Bay study (Palacios, 2012).

PCC from pigments with CHEMTAX: Among methods that utilize quantitative phytoplankton pigment concentrations to determine PCC (e.g., Chase et al., 2020; Higgins et al., 2012; Kramer & Siegel, 2019; reviewed in Higgins et al., 2012), the commonly-utilized CHEMTAX software was chosen for this analysis. Uncertainties in this method include variability in pigment concentrations at intra- and interspecific levels *not* associated with changes in abundance (Armbrecht et al., 2015; Kozlowski et al., 2011; Latasa et al., 2022), which can add error to the pigment-based PCC proportions the algorithm is tested against. Changes in marker:chla ratios may reflect cellular-level changes in photophysiology over the course of the day (e.g., photoacclimation, photoprotection) or nutrient condition (Lewitus et al., 2005) rather than abundance. CHEMTAX-based PCC proportions can be inconsistent compared to cell count-based PCC, particularly for sub-class level taxonomic identification (Armbrecht et al., 2015; Kozlowski et al., 2011; Latasa et al., 2022). At the class level, however, CHEMTAX-based PCC fractions show significant correlation with cell count-based PCC in the Chesapeake Bay (Harding et al., 2015), the nearby NW Atlantic (Pan et al., 2011), and studies across a variety of aquatic regimes (e.g., Kozlowski et al., 2011). Even with these uncertainties, past work and results of this study support application of CHEMTAX-based PCC proportions as representative of in-water PCC.

Lessons learned from CHEMTAX: Previous research on CHEMTAX, the conceptual ancestor of PHYDOTax (Palacios, 2012), may inform future PHYDOTax development. CHEMTAX (and potentially PHYDOTax) can be sensitive to the composition of the input matrix (Mackey et al., 1996). For PHYDOTax this matrix is the Rrs spectral signature library (Figure 1), and for CHEMTAX it's a matrix of identifying pigment ratios (Table S2c and S2d in Supporting Information S1). The process of deciding which phytoplankton types comprise these input matrices is subjective to some degree and can be limited in scope due to the low availability of the highly-specialized data types that they require. Optimization of the input matrix to the pigment data set, prior to its factorization, has been found to reduce this source of CHETMAX error attributable to composition and increase statistical power of the results (Latasa, 2007; Mackey et al., 1996). The software includes one optimization approach (Mackey et al., 1996; Wright, 2017), and others have since been developed that also proved successful (e.g., Kozlowski et al., 2011; Latasa, 2007). PHYDOTax functionality may similarly benefit from developing and

testing optimization methods for the Rrs spectral library. This includes understanding how differing the composition of phytoplankton taxa in the library may impact performance of PHYDOTax. For example, the library in this work includes mostly cosmopolitan, generalized coastal phytoplankton types commonly grown in oceanography labs. Could PHYDOTax performance in the Chesapeake improve if wild isolates of genera and species from the region were cultured to develop a new set of spectral signatures to represent the phytoplankton classes of interest? Or are spectral library signatures from broadly coastal, but not highly region-specific, phytoplankton taxa sufficient to yield reliable results from PHYDOTax? Investigation of these and related questions will improve understanding of how best to build and apply the input Rrs spectral library for PHYDOTax.

3.5. Conclusions

Development of approaches to leverage upcoming hyperspectral satellite-based information in coastal and estuarine zones represents an opportunity to better understand these critical environments and to develop advanced satellite products that will inform decisions which address societal challenges such as such as health, safety, and food security (Schollaert Uz et al., 2019). This study contributed to these goals by testing a bio-optical approach to PCC identification with field data from two cruise campaigns in the Chesapeake Bay. PHYDOTax discriminated cryptophyte and cyanophyte classes using ship-based Rrs measurements and synthetic data. Spectral sensitivity of the algorithm to input at differing hyperspectral resolutions of 1, 5, or 10 nm was low. Sensitivity to addition of non algal components varied according to field data set and possibly seasonal PCC composition. Due to the considerable temporal and spatial variability of the Chesapeake Bay relative to our data set, our results do not conclude that PHYDOTax is an operational algorithm ready for Bay applications. These results do provide the first support to the theory that PHYDOTax may be applied beyond the coastal upwelling zone that it was developed in. Further development and testing is encouraged where sufficient data are available in this and other optically complex regimes.

Data Availability Statement

Phytoplankton pigment and remote sensing reflectance (Rrs) data sets for algorithm performance evaluation were downloaded from the NASA SeaWIFS Bio-optical Archive and Storage System (SeaBASS; Werdell et al., 2003) repository for in situ oceanographic data, maintained by the NASA Ocean Biology Processing Group (OBPG). Original cruise 1 and 2 data are available at SeaBASS (2009) [https://doi.org/10.5067/SeaBASS/2009OCT_CHESAPEAKE/DATA001] and SeaBASS (2011) [https://doi.org/10.5067/SeaBASS/GEO-CAPE/DATA001], respectively. Phytoplankton community composition was derived from pigment data with CHEMical TAXonomy (CHEMTAX) software, available at https://data.aad.gov.au/metadata/CHEMTAX (Wright, 2017) and licensed under the CCBY Attribution License (http://creativecommons.org/licenses/by/4.0/). Hyperspectral airborne data were provided by Chuanmin Hu (Univ. of South Florida) and Minwei Zhang (NASA/SAIC), link to repository with data file below.

A comprehensive Code and Data Repository is located at Palacios and McKibben (2023) [https://doi.org/10.5281/zenodo.10360063]. For this work the original MATLAB PHYDOTax code from Palacios, 2012 was converted to Python. This Python code plus all data related to this manuscript are within this repository. Specific data files include the Rrs spectral libraries, the paired Rrs and pigment concentrations from ship-based field data, the 20 July 2011 swath of airborne Rrs data, and the synthetic mixture data sets.

All manuscript data were processed with Python (v. 3.9.7) using open source packages NumPy (v. 1.20.3; https://numpy.org), Pandas (v. 1.3.4; https://pandas.pydata.org), and/or Xarray (v. 0.16.1; https://xarray.dev). Figures were plotted with open source Matplotlib (v. 3.2.1; https://matplotlib.org/) and Cartopy (v. 0.20.0; https://pypi.org/project/Cartopy/) packages. See links for download and license information.

References

Adolf, J. E., Yeager, C. L., Miller, W. D., Mallonee, M. E., & Harding, L. W. (2006). Environmental forcing of phytoplankton floral composition, biomass, and primary productivity in Chesapeake Bay, USA. *Estuarine, Coastal and Shelf Science*, *67*(1–2), 108–122. https://doi.org/10.1016/j.ecss.2005.11.030

Armbrecht, L. H., Wright, S. W., Petocz, P., & Armand, L. K. (2015). A new approach to testing the agreement of two phytoplankton quantification techniques: Microscopy and CHEMTAX. *Limnology and Oceanography: Methods*, *13*(8), 425–437. https://doi.org/10.1002/lom3.10037

Aurin, D., Mannino, A., & Franz, B. (2013). Spatially resolving ocean color and sediment dispersion in river plumes, coastal systems, and continental shelf waters. *Remote Sensing of Environment*, *137*, 212–225. https://doi.org/10.1016/j.rse.2013.06.018

Behrenfeld, M. J., Boss, E. S., & Halsey, K. H. (2021). Phytoplankton community structuring and succession in a competition-neutral resource landscape. *ISME Communications*, *1*, 1–8. https://doi.org/10.1038/s43705-021-00011-5

Ben-Israel, A., & Greville, T. (2003). *Generalized inverses: Theory and applications* (2nd ed.). Springer Science & Business Media, Springer. https://doi.org/10.1007/b97366

Bracher, A., Bouman, H. A., Brewin, R. J. W., Bricaud, A., Brotas, V., Ciotti, A. M., et al. (2017). Obtaining phytoplankton diversity from ocean color: A scientific roadmap for future development. *Frontiers in Marine Science*, *4*, 1–15. https://doi.org/10.3389/fmars.2017.00055

Cael, B., Chase, A., & Boss, E. (2020). Information content of absorption spectra and implications for ocean color inversion. *Applied Optics*, *59*(13), 3971–3984. https://doi.org/10.1364/ao.389189

Cael, B. B., Bisson, K., Boss, E., & Erickson, Z. K. (2023). How many independent quantities can be extracted from ocean color? *Limnology and Oceanography Letters*, *8*(4), 603–610. https://doi.org/10.1002/lol2.10319

Cetinić, I., Rousseaux, C. S., Carroll, I. T., Chase, A. P., Kramer, S. J., Werdell, P. J., et al. (2024). Phytoplankton composition from sPACE: Requirements, opportunities, and challenges. *Remote Sensing of Environment*, *302*, 113964. https://doi.org/10.1016/j.rse.2023.113964

Chase, A. P., Boss, E., Cetinic, I., & Slade, W. (2017). Estimation of phytoplankton accessory pigments from hyperspectral reflectance spectra: Toward a global algorithm. *Journal of Geophysical Research: Oceans*, *122*(12), 9725–9743. https://doi.org/10.1002/2017JC012859

Chase, A. P., Kramer, S. J., Haëntjens, N., Boss, E. S., Karp-boss, L., Edmondson, M., & Graff, J. R. (2020). Evaluation of diagnostic pigments to estimate phytoplankton size classes. *Limnology and Oceanography: Methods*, *18*(10), 570–584. https://doi.org/10.1002/lom3.10385

Chin, T. M., Vazquez-Cuervo, J., & Armstrong, E. M. (2017). A multi-scale high-resolution analysis of global sea surface temperature. *Remote Sensing of Environment*, *200*, 154–169. https://doi.org/10.1016/j.rse.2017.07.029

Dierssen, H. M., Gierach, M., Guild, L. S., Mannino, A., Salisbury, J., Schollaert Uz, S., et al. (2023). Synergies between NASA's hyperspectral aquatic missions PACE, GLIMR, and SBG: Opportunities for new science and applications. *Journal of Geophysical Research: Biogeosciences*, *128*(10). https://doi.org/10.1029/2023JG007574

Dierssen, H. M., Kudela, R. M., Ryan, J. P., & Zimmerman, R. C. (2006). Red and black tides: Quantitative analysis of water-leaving radiance and perceived color for phytoplankton, colored dissolved organic matter, and suspended sediments. *Limnology & Oceanography*, *51*(6), 2646–2659. https://doi.org/10.4319/lo.2006.51.6.2646

Fishman, J., Iraci, L. T., Al-Saadi, J., Chance, K., Chavez, F., Chin, M., et al. (2012). The United States' next generation of atmospheric composition and coastal ecosystem measurements: NASA's geostationary coastal and air pollution events (GEO-CAPE) mission. *Bulletin of the American Meteorological Society*, *93*(10), 1547–1566. https://doi.org/10.1175/BAMS-D-11-00201.1

Glibert, P. M. (2016). Margalef revisited: A new phytoplankton mandala incorporating twelve dimensions, including nutritional physiology. *Harmful Algae*, *55*, 25–30. https://doi.org/10.1016/j.hal.2016.01.008

Gordon, H. R., & Wang, M. (1994). Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: A preliminary algorithm. *Applied Optics*, *33*(3), 443–452. https://doi.org/10.1364/ao.33.000443

Harding, L. W., Adolf, J. E., Mallonee, M. E., Miller, W. D., Gallegos, C. L., Perry, E. S., et al. (2015). Climate effects on phytoplankton floral composition in Chesapeake Bay. *Estuarine, Coastal and Shelf Science*, *162*, 53–68. https://doi.org/10.1016/j.ecss.2014.12.030

Harding, L. W., Mallonee, M. E., Perry, E. S., Miller, W. D., Adolf, J. E., Gallegos, C. L., & Paerl, H. W. (2016). Variable climatic conditions dominate recent phytoplankton dynamics in Chesapeake Bay. *Scientific Reports*, *6*, 1–16. https://doi.org/10.1038/srep23773

Harding, L. W., Mallonee, M. E., Perry, E. S., Miller, W. D., Adolf, J. E., Gallegos, C. L., & Paerl, H. W. (2019). Long-term trends, current status, and transitions of water quality in Chesapeake Bay. *Scientific Reports*, *9*, 1–19. https://doi.org/10.1038/s41598-019-43036-6

Higgins, H. W., Wright, S. W., & Schlüter, L. (2012). Quantitative interpretation of chemotaxonomic pigment data. In S. Roy, C. A. Llewellyn, E. S. Egeland, & G. Johnsen (Eds.), *Phytoplankton pigments: Characterization, chemotaxonomy and applications in oceanography. Cambridge environmental chemistry series* (pp. 257–313). Cambridge University Press. https://doi.org/10.1017/CBO9780511732263.010

Hooker, S. B., Van Heukelem, L. V., Thomas, C. S., Claustre, H., Ras, J., Barlow, R., et al. (2005). *Second SeaWiFS HPLC analysis round-robin experiment (SeaHARRE-2)* (Vol. 212785). National Aeronautics and Space Administration, Goddard Space Flight Center. Retrieved from https://oceancolor.gsfc.nasa.gov/fsg/hplc/SH2_TM2005_212785.pdf

Jeffrey, S. W., Wright, S. W., & Zapata, M. (2012). Microalgal classes and their signature pigments. In S. Roy, C. A. Llewellyn, E. S. Egeland, & G. Johnsen (Eds.), *Phytoplankton pigments: Characterization, chemotaxonomy and applications in oceanography. Cambridge environmental chemistry series* (Vol. 2011, pp. 257–313). Cambridge University Press. https://doi.org/10.1017/cbo9780511732263.004

Keller, D. P., Lee, D. Y., & Hood, R. R. (2014). Turbidity maximum entrapment of phytoplankton in the Chesapeake Bay. *Estuaries and Coasts*, *37*(2), 279–298. https://doi.org/10.1007/s12237-013-9692-2

Kozlowski, W. A., Deutschman, D., Garibotti, I., Trees, C., & Vernet, M. (2011). An evaluation of the application of CHEMTAX to Antarctic coastal pigment data. *Deep Sea Res 1 Oceanogr Res Pap*, *58*(4), 350–364. https://doi.org/10.1016/j.dsr.2011.01.008

Kramer, S. J., & Siegel, D. A. (2019). How can phytoplankton pigments be best used to characterize surface ocean phytoplankton groups for ocean color remote sensing algorithms? *Journal of Geophysical Research: Oceans*, *124*(11), 7557–7574. https://doi.org/10.1029/2019JC015604

Kramer, S. J., Siegel, D. A., Maritorena, S., & Catlett, D. (2022). Modeling surface ocean phytoplankton pigments from hyperspectral remote sensing reflectance on global scales. *Remote Sensing of Environment*, *270*, 112879. https://doi.org/10.1016/j.rse.2021.112879

Latasa, M. (2007). Improving estimations of phytoplankton class abundances using CHEMTAX. *Marine Ecology Progress Series*, *329*, 13–21. https://doi.org/10.3354/meps329013

Latasa, M., Scharek, R., Morán, X. A. G., Gutiérrez-Rodríguez, A., Emelianov, M., Salat, J., et al. (2022). Dynamics of phytoplankton groups in three contrasting situations of the open NW Mediterranean Sea revealed by pigment, microscopy, and flow cytometry analyses. *Progress in Oceanography*, *201*, 102737. https://doi.org/10.1016/j.pocean.2021.102737

Lee, Z., Carder, K. L., Mobley, C. D., Steward, R. G., & Patch, J. S. (1999). Hyperspectral remote sensing for shallow waters: 2. Deriving bottom depths and water properties by optimization. *Applied Optics*, *38*(18), 3831–3843. https://doi.org/10.1364/AO.38.003831

Lee, Z., Carder, K. L., Steward, R. G., Peacock, T. G., Davis, C. O., & Mueller, J. L. (1997). Remote sensing reflectance and inherent optical properties of oceanic waters derived from above-water measurements. In *Ocean Optics XIII* (Vol. 2963, pp. 160–166). SPIE. https://doi.org/10.1117/12.266436

Lewitus, A. J., White, D. L., Tymowski, R. G., Geesey, M. E., Hymel, S. N., & Noble, P. A. (2005). Adapting the CHEMTAX method for assessing phytoplankton taxonomic composition in southeastern U.S. estuaries. *Estuaries*, *28*(1), 160–172. https://doi.org/10.1007/BF02732761

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. https://doi.org/10.3354/meps144265

Marshall, H. G., Burchardt, L., & Lacouture, R. (2005). A review of phytoplankton composition within Chesapeake Bay and its tidal estuaries. *Journal of Plankton Research*, *27*(11), 1083–1102. https://doi.org/10.1093/plankt/fbi079

May, J. O., & Looney, S. W. (2020). Sample size charts for Spearman and Kendall coefficients methods for sample size determination. *Biometrics & Biostatistics*, *11*, 5–11. https://doi.org/10.37421/jbmbs.2020.11.440

Mobley, C. D. (1994). *Light and water: Radiative transfer in natural waters*. Academic press.

Mobley, C. D., Stramski, D., Bissett, W., & Boss, E. (2004). Optical modeling of ocean water. *Oceanography*, *17*(2), 60–67. https://doi.org/10.5670/oceanog.2004.48

Mouw, C. B., Greb, S., Aurin, D., DiGiacomo, P. M., Lee, Z., Twardowski, M., et al. (2015). Aquatic color radiometry remote sensing of coastal and inland waters: Challenges and recommendations for future satellite missions. *Remote Sensing of Environment*, *160*, 15–30. https://doi.org/10.1016/j.rse.2015.02.001

Mouw, C. B., Hardman-Mountford, N. J., Alvain, S., Bracher, A., Brewin, R. J. W., Bricaud, A., et al. (2017). A consumer's guide to satellite remote sensing of multiple phytoplankton groups in the global ocean. *Frontiers in Marine Science*, *4*. https://doi.org/10.3389/fmars.2017.00041

Paerl, H. W., Valdes, L. M., Peierls, B. L., Adolf, J. E., & Harding, L. W. (2006). Anthropogenic and climatic influences on the eutrophication of large estuarine ecosystems. *Limnology & Oceanography*, *51*(1part2), 448–462. https://doi.org/10.4319/lo.2006.51.1_part_2.0448

Palacios, S. L. (2012). *Identifying and tracking evolving water masses in optically complex aquatic environments*. UC Santa Cruz. ProQuest ID: Palacios_ucsc_0036E_10002. Merritt ID: ark:/13030/m55h7jzx. https://escholarship.org/uc/item/19c9r3w1

Palacios, S. L., Kudela, R. M., Guild, L. S., Negrey, K. H., Torres-Perez, J., & Broughton, J. (2015). Remote sensing of phytoplankton functional types in the coastal ocean from the HyspIRI Preparatory Flight Campaign. *Remote Sensing of Environment*, *167*, 269–280. https://doi.org/10.1016/j.rse.2015.05.014

Palacios, S. L., & McKibben, S. M. (2023). [Python code and data for manuscript] PHYDOTax: PHYtoplankton detection with optics (PHYDOTax_v1.0.0). https://doi.org/10.5281/zenodo.10360063

Pan, X., Mannino, A., Marshall, H. G., Filippino, K. C., & Mulholland, M. R. (2011). Remote sensing of phytoplankton community composition along the northeast coast of the United States. *Remote Sensing of Environment*, *115*(12), 3731–3747. https://doi.org/10.1016/j.rse.2011.09.011

Sathyendranath, S. (2014). Phytoplankton functional types from space. *Reports of the International Ocean-Colour Coordinating Group. Dartmouth, Canada*. https://doi.org/10.25607/OBP-106

Schollaert Uz, S., Kim, G. E., Mannino, A., Werdell, P. J., & Tzortziou, M. (2019). Developing a community of practice for applied uses of future PACE data to address marine food security challenges. *Frontiers of Earth Science*, *7*, 1–13. https://doi.org/10.3389/feart.2019.00283

SeaBASS. (2009). 2009OCT_CHESAPEAKE [Dataset]. *NASA Ocean Biology Processing Group*. https://doi.org/10.5067/SeaBASS/2009OCT_CHESAPEAKE/DATA001

SeaBASS. (2011). GEO-CAPE [Dataset]. *NASA Ocean Biology Processing Group*. https://doi.org/10.5067/SeaBASS/GEO-CAPE/DATA001

Stavros, E. N., Chrone, J., Cawse-Nicholson, K., Freeman, A., Glenn, N. F., Guild, L., et al. (2023). Designing an observing system to study the surface Biology and Geology (SBG) of the Earth in the 2020s. *Journal of Geophysical Research: Biogeosciences*, *128*, 1–18. https://doi.org/10.1029/2021jg006471

Stramski, D., Bricaud, A., & Morel, A. (2001). Modeling the inherent optical properties of the ocean based on the detailed composition of the planktonic community. *Applied Optics*, *40*(18), 2929. https://doi.org/10.1364/ao.40.002929

Stumpf, R. P., Holderied, K., & Sinclair, M. (2003). Determination of water depth with high-resolution satellite imagery over variable bottom types. *Limnology & Oceanography*, *48*(1part2), 547–556. https://doi.org/10.4319/lo.2003.48.1_part_2.0547

Tzortziou, M., Herman, J. R., Gallegos, C. L., Neale, P. J., Subramaniam, A., Harding, L. W., & Ahmad, Z. (2006). Bio-optics of the Chesapeake Bay from measurements and radiative transfer closure. *Estuarine, Coastal and Shelf Science*, *68*(1–2), 348–362. https://doi.org/10.1016/j.ecss.2006.02.016

Tzortziou, M., Subramaniam, A., Herman, J. R., Gallegos, C. L., Neale, P. J., & Harding, L. W. (2007). Remote sensing reflectance and inherent optical properties in the mid Chesapeake Bay. *Estuarine, Coastal and Shelf Science*, *72*(1–2), 16–32. https://doi.org/10.1016/j.ecss.2006.09.018

Van Heukelem, L., & Hooker, S. B. (2012). The importance of a quality assurance plan for method validation and minimizing uncertainties in the HPLC analysis of phytoplankton pigments. *Phytoplankton Pigments*, 195–256. https://doi.org/10.1017/cbo9780511732263.009

Van Heukelem, L., & Thomas, C. S. (2001). Computer-assisted HPLC method development with applications to the isolation and analysis of marine phytoplankton pigments. *Journal of Chromatography, A*, *910*(1), 31–49. https://doi.org/10.1016/S0378-4347(00)00603-4

Wang, M., Son, S. H., & Harding, L. W. (2009). Retrieval of diffuse attenuation coefficient in the Chesapeake Bay and turbid ocean regions for satellite ocean color applications. *Journal of Geophysical Research*, *114*(C10). https://doi.org/10.1029/2009JC005286

Werdell, P. J., Bailey, S., Fargion, G., Pietras, C., Knobelspiesse, K., Feldman, G., & Mcclain, C. (2003). Unique data repository facilitates ocean color satellite validation. *Eos*, *84*(38), 377–387. https://doi.org/10.1029/2003EO380001

Werdell, P. J., Behrenfeld, M. J., Bontempi, P. S., Boss, E., Cairns, B., Davis, G. T., et al. (2019). The Plankton, Aerosol, Cloud, ocean Ecosystem mission status, science, advances. *Bulletin of the American Meteorological Society*, *100*(9), 1775–1794. https://doi.org/10.1175/BAMS-D-18-0056.1

Wolny, J. L., Tomlinson, M. C., Schollaert Uz, S., Egerton, T. A., McKay, J. R., Meredith, A., et al. (2020). Current and future remote sensing of harmful algal blooms in the Chesapeake Bay to support the shellfish Industry. *Frontiers in Marine Science*, *7*, 1–16. https://doi.org/10.3389/fmars.2020.00337

Wright, S. (2017). *CHEMTAX for calculating the taxonomic composition of phytoplankton populations, Ver. 1.95*. Australian Antarctic Data Centre. https://doi.org/10.4225/15/59fff1c5ea8fc

Wright, S. W., & Jeffrey, S. W. (2006a). Marine organic matter: Biomarkers, isotopes and DNA. *The Handbook of Marine Chemistry*, 71–104.

Wright, S. W., & Jeffrey, S. W. (2006b). Pigment markers for phytoplankton production. In J. K. Volkman (Ed.), *Marine organic matter: Biomarkers, isotopes and DNA* (pp. 71–104). Springer-Verlag. https://doi.org/10.1007/698_2_003

Zhang, M., Hu, C., Cannizzaro, J., Kowalewski, M. G., & Janz, S. J. (2018). Diurnal changes of remote sensing reflectance over Chesapeake bay: Observations from the airborne Compact atmospheric mapper. *Estuarine, Coastal and Shelf Science*, *200*, 181–193. https://doi.org/10.1016/j.ecss.2017.10.021