Characterizing CDOM Spectral Variability Across Diverse Regions and Spectral Ranges: Characterizing CDOM S....

Article *in* Global Biogeochemical Cycles · December 2017

DOI: 10.1002/2017GB005756

CITATIONS
0

READS
96

3 authors:

Brice Grunert
Michigan Technological University
15 PUBLICATIONS **6** CITATIONS
SEE PROFILE

Colleen B. Mouw
University of Rhode Island
59 PUBLICATIONS **414** CITATIONS
SEE PROFILE

Audrey Ciochetto
University of Rhode Island
14 PUBLICATIONS **19** CITATIONS
SEE PROFILE

AGU PUBLICATIONS

Global Biogeochemical Cycles

RESEARCH ARTICLE
10.1002/2017GB005756

Key Points:
- CDOM spectral slope spatial variability is dependent on spectral range used
- CDOM spectral slope variability is not necessarily reflected in CDOM absorption values in ocean regions with little terrestrial influence
- Gaussian peaks fitting lignin absorption show a distinct blue shift from inland to ocean waters

Supporting Information:
- Supporting Information S1

Correspondence to:
B. K. Grunert,
bgrunert@mtu.edu

Citation:
Grunert, B. K., Mouw, C. B., & Ciochetto, A. B. (2018). Characterizing CDOM spectral variability across diverse regions and spectral ranges. *Global Biogeochemical Cycles*, *32*. https://doi.org/10.1002/2017GB005756

Received 30 JUN 2017
Accepted 19 DEC 2017
Accepted article online 23 DEC 2017

Characterizing CDOM Spectral Variability Across Diverse Regions and Spectral Ranges

Brice K. Grunert[1], Colleen B. Mouw[2], and Audrey B. Ciochetto[2]

[1]Department of Geological and Mining Engineering and Sciences, Michigan Technological University, Houghton, MI, USA, [2]Graduate School of Oceanography, University of Rhode Island, Narragansett, RI, USA

Abstract Satellite remote sensing of colored dissolved organic matter (CDOM) has focused on CDOM absorption (a_{CDOM}) at a reference wavelength, as its magnitude provides insight into the underwater light field and large-scale biogeochemical processes. CDOM spectral slope, S_{CDOM}, has been treated as a constant or semiconstant parameter in satellite retrievals of a_{CDOM} despite significant regional and temporal variabilities. S_{CDOM} and other optical metrics provide insights into CDOM composition, processing, food web dynamics, and carbon cycling. To date, much of this work relies on fluorescence techniques or a_{CDOM} in spectral ranges unavailable to current and planned satellite sensors (e.g., <300 nm). In preparation for anticipated future hyperspectral satellite missions, we take the first step here of exploring global variability in S_{CDOM} and fit deviations in the a_{CDOM} spectra using the recently proposed Gaussian decomposition method. From this, we investigate if global variability in retrieved S_{CDOM} and Gaussian components is significant and regionally distinct. We iteratively decreased the spectral range considered and analyzed the number, location, and magnitude of fitted Gaussian components to understand if a reduced spectral range impacts information obtained within a common spectral window. We compared the fitted slope from the Gaussian decomposition method to absorption-based indices that indicate CDOM composition to determine the ability of satellite-derived slope to inform the analysis and modeling of large-scale biogeochemical processes. Finally, we present implications of the observed variability for remote sensing of CDOM characteristics via S_{CDOM}.

Plain Language Summary A large portion of the global carbon cycle is attributed to dissolved material both used by and produced by aquatic microbial organisms. Tracking changes to this pool is notoriously difficult but necessary to better constrain biogeochemical processes relevant to the global carbon cycle. By considering global variability in a colored component of the dissolved material, we show that there are unique trends in different ocean regions related to the biology, chemistry, and movement of water masses in those systems. Viewing these changes with a next-generation satellite sensor with expanded viewing channels may provide tremendous insight into global patterns relevant to the production and degradation of this globally relevant pool of carbon.

1. Introduction

Dissolved organic matter (DOM) is the largest pool of reduced carbon in the ocean (Hansell et al., 2009). Slight alterations in DOM remineralization can result in significant increases or decreases in the release of CO_2 (Mannino et al., 2014) with remineralization rates highly dependent on DOM composition. DOM constituents vary from labile, readily metabolized compounds (e.g., phytoplankton exudates) important to short-term climate dynamics and food web processes to refractive compounds (e.g., microbial by-products) decoupled from modern climate dynamics and largely precluded from food web dynamics (Hansell, 2013). DOM contains optically active components in colored and fluorescent dissolved organic matter (CDOM and FDOM, respectively), with CDOM a small, variable portion of the total DOM pool and FDOM a small, variable portion of the CDOM pool (Stedmon & Nelson, 2014). CDOM impacts the spectral quality and intensity of the underwater light field, surface ocean heating and plays direct and indirect roles in biogeochemical cycling (Andrew et al., 2013; Hickman et al., 2010; Kim et al., 2016).

Advances in optical characterization through absorption and fluorescence techniques have illuminated mechanistic relationships between CDOM molecular composition and optical properties such as the shape of the CDOM absorption spectra indicated by the spectral slope of CDOM (S_{CDOM}), where S_{CDOM} is derived from fitting the CDOM absorption spectra to an exponential curve (see equation (1)) (Helms et al., 2008;

Stedmon & Markager, 2005; Walker et al., 2013). Changes to CDOM optical characteristics have been related to general classifications of the composition of CDOM including molecular weight and origin (Spencer et al., 2008). The strongest absorption-based indicators of these properties rely on S_{CDOM} from 275 to 295 nm and 350 to 400 nm ($S_{275:295}$ and $S_{350:400}$, respectively) since the largest divergence in spectral properties across compositionally unique CDOM samples are found in these spectral regions (Helms et al., 2008). Low values of $S_{275:295}$ (~0.01 nm^{-1}) generally represent high molecular weight material (e.g., lignin and lignin derivatives) that decreases in molecular size primarily from photodegradation and secondarily from microbial degradation (Del Vecchio & Blough, 2002). $S_{275:295}$ increases with photodegradation, with terrestrial material often showing markedly different $S_{275:295}$ values due to the presence of lignin. Conversely, $S_{350:400}$ generally decreases with photodegradation, potentially due to the presence of photorefractory chromophores that absorb in this portion of the spectrum (Helms et al., 2013).

Linking a_{CDOM} optical proxies with future hyperspectral satellite data presents a potentially powerful tool to assess spatially and temporally distinct biogeochemical processes associated with a_{CDOM} variability. To date, this connection is complicated by a lack of consistency in the spectral range used to calculate S_{CDOM} and a general focus on a spectral range that best represents the entire a_{CDOM} spectra over the utility of that spectral range (Schwarz et al., 2002; Twardowski et al., 2004). Broad range S_{CDOM} (e.g., $S_{240:700}$) that best characterize the entire a_{CDOM} spectra have not been considered extensively for mechanistic relationships with CDOM composition. Most studies track S_{CDOM} across salinity gradients (Kowalczuk, Stedmon, & Markager, 2006; Stedmon et al., 2011; Stedmon & Markager, 2003) or distinct in-water processes (Nelson, Siegel, & Michaels, 1998; Organelli et al., 2014), producing correlations between S_{CDOM} and CDOM source or diagenetic state across geographically or temporally unique environments (Organelli et al., 2014; Stedmon & Markager, 2003). Some work has suggested that $S_{300:600}$ is a useful indicator of CDOM molecular weight (Stedmon & Nelson, 2014), while the general understanding is that broad range S_{CDOM} increases with photodegradation as chromophores are bleached, particularly in the UV region (Del Vecchio & Blough, 2002; Reche et al., 2000). Optical properties of source material have been considered (Hansen et al., 2016; Visser, 1983), and two major components of CDOM, humic acid and fulvic acid, display slopes of approximately 0.011 nm^{-1} and 0.019 nm^{-1} (Carder et al., 1989). However, observed slope values across broad spectral ranges can vary above and below these benchmarks and the optical properties of source material are rapidly altered through poorly understood processes (Sharpless et al., 2014), suggesting that the complexity of material and its spectral signal is still poorly constrained.

Based on this, optically estimating CDOM composition from hyperspectral ocean color data requires validation of significant temporal and regional variabilities in S_{CDOM} and a potential link between observable spectral features and spectral indices capable of estimating CDOM composition. Maximizing CDOM compositional information that can be retrieved by optical methods will enhance our ability to track changes in the CDOM pool through autonomous and remote sensing platforms, providing a way to observe large-scale changes in ocean biogeochemical processes and circulation (Nelson et al., 2010). However, this first requires determination of what information regarding CDOM composition is garnered from a specific spectral range and fully utilizing the information contained within a_{CDOM} spectra to optically estimate CDOM composition. Here we utilize a_{CDOM} spectra available from public data repositories to address knowledge gaps that have evolved from an inconsistency in spectral range used. We utilize the Gaussian decomposition approach of Massicotte and Markager (2016) to better estimate S_{CDOM} and further characterize the CDOM pool and report differences between this approach and the standard exponential model traditionally used for fitting a_{CDOM} spectra. We describe global variability in S_{CDOM} within the context of biogeochemical provinces and relate several broad range S_{CDOM} to $S_{275:295}$, $S_{350:400}$, and a_{CDOM} at 350 nm ($a_{CDOM}(350)$) to define the efficacy of each of these parameters to optically estimate CDOM composition based on observations in the literature. Finally, we offer insight into what can be determined about CDOM composition from $S_{350:400}$ and $S_{350:550}$ to step toward defining a methodology that can be applied to hyperspectral, satellite remotely sensed retrievals of a_{CDOM}.

2. Methods

2.1. CDOM Absorption Spectra

CDOM absorption spectra were obtained from NASA SeaWiFS Bio-optical Archive and Storage System (SeaBASS, https://seabass.gsfc.nasa.gov/) on 13 April 2016 (Werdell et al., 2003). Only data measured on a

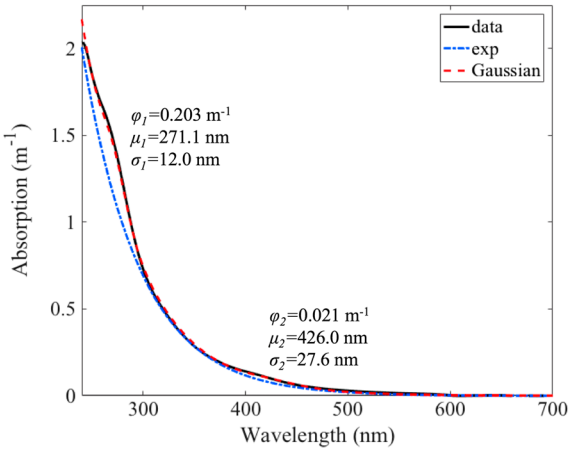

Figure 1. Example of exponential (equation (1)) and Gaussian decomposition fitting (equation (2)) of a_{CDOM} spectra from the Chukchi Sea. The black line is the observed data, the blue dash-dot line is the exponential fit for the spectra from 240 to 700 nm ($S_{E240:700}$ = 0.0178 nm^{-1}), and the red dashed line is the Gaussian decomposition fit from 240 to 700 nm ($S_{G240:700}$ = 0.0180 nm^{-1}). Two Gaussian components were fitted to this spectra, with φ, μ, and σ for each component indicated within the figure.

bench top spectrophotometer were utilized. The CDOM absorption spectrum is typically modeled using an exponential decay function:

$$a_{CDOM}(\lambda) = a_{CDOM}(\lambda_0)^{e-S_E(\lambda-\lambda_0)} + K \quad (1)$$

where λ_0 represents a reference wavelength for initializing the spectrum (nm), S_E is the spectral slope coefficient for an exponential CDOM model (nm^{-1}), and K is a constant addressing background noise and potential instrument bias (m^{-1}), calculated as the average a_{CDOM} from 690 to 700 nm (Figure 1). Spectra were quality controlled by selecting spectra representing a broad quasi-exponential function: if the fit with equation (1) displayed an $r^2 > 0.9$ (suggesting no contamination or instrumentation issues), spectra were kept for further analysis (Massicotte & Frenette, 2011).

Deviations from the decaying expression in equation (1) can result from absorption by specific chromophores (e.g., lignin) and can vary in degree and location (Massicotte & Markager, 2016). To detect regions where such deviations exist, spectra were fit with the standard exponential in equation (1) and then trimmed by removing data with residuals greater than the mean absolute residual multiplied by the weighting factor, C. C is dependent on the quality of the data (signal-to-noise ratio), the environment, and spectral range considered and governs the number of residuals considered for Gaussian fitting by defining the threshold for excluding these points from the baseline exponential fit; we utilized a C of 1 to maintain consistency across spectra. With the remaining points, spectra were again fit with equation (1) with the result defined as the baseline exponential. Massicotte and Markager (2016) utilized a spectral range from 240 to 700 nm with Gaussian fitting performed on residuals from 250 to 500 nm. S_{CDOM} is sensitive to the spectral range used for calculation when using the exponential model (e.g., Twardowski et al., 2004). To determine the influence of spectral range on S_{CDOM} using the Gaussian decomposition model, we considered spectral ranges of 240–700 nm, 300–700 nm, and 350–550 nm, with Gaussian fitting performed on residuals below 500 nm to avoid concerns with a low signal-to-noise (SNR) ratio at longer wavelengths (Massicotte & Markager, 2016). For spectral ranges below 300 nm, we found including wavelengths down to 240 nm, rather than a 250 nm cutoff, better accounted for absorption deviations likely due to lignin or lignin derivatives. This and altering the spectral range considered were the only deviations from the methodology as described by Massicotte and Markager (2016). The optimal number of Gaussian components modeled for each spectrum was chosen based on minimizing the Bayesian information criterion (BIC) score, with the final equation represented as

$$a_{CDOM}(\lambda) = a_{CDOM}(\lambda_0)^{e-S_G(\lambda-\lambda_0)} + K + \sum_{i=0}^{n} \varphi_i e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}} + \varepsilon \quad (2)$$

where S_G (nm^{-1}) is the spectral slope coefficient for a Gaussian decomposition CDOM model, φ (m^{-1}) is the height of the Gaussian peak, μ (nm) is the position of the center of the peak, σ (nm) is the width of the peak, and ε are the residuals after fitting of the full model. An example spectrum fitted with and without the Gaussian decomposition approach is shown (Figure 1), illustrating the change in slope. For this approach, fitted K as described for equation (1) is also used as a threshold for the minimum height of fitted Gaussian components to add certainty that the Gaussian components are fitting chromophores and not instrument noise. We also acknowledged that many different instruments with varying accuracy were used to obtain this large data set. Considering this, we calculated the mean K value across all measured spectra (0.0004) and applied this as a threshold value for fitting Gaussian components: if the spectra-specific K value was smaller than the mean K, the mean K was used for fitting Gaussian components; the spectra-specific K value was used for all other spectra.

The inclusion of K in equations (1) and (2) is often debated, with many researchers considering it an unknown variable that has no physical basis. Others have pointed out its ability to account for instrument noise in regions of the spectrum that should have no, or very little, signal from CDOM absorption. We have included it in this analysis for several reasons: (1) many of the spectra that we considered had some noise at longer

wavelengths (>600 nm) within an otherwise good absorption spectra, suggesting some instrument noise in measuring a very low signal; (2) our analysis focuses primarily on the shape of the spectra, which is not altered by including a K term; (3) our a_{CDOM} analyses are focused on shorter wavelengths (e.g., 350 nm) where the percent contribution of the K term to the overall signal analyzed is generally quite small (mean of 0.6%); and (4) K was used as the threshold for fitting Gaussian components, leading to stricter fitting.

Spectral slope was retrieved for the following wavelength ranges: 275–295 nm, 350–400 nm, 240–700 nm, 300–700 nm, and 350–550 nm. We also considered spectral slope from 300 to 600 nm, 350 to 600 nm, and 350 to 700 nm; however, slope values and Gaussian statistics for these spectral ranges were not significantly different from 300 to 700 nm and thus were not considered further. Throughout the manuscript, slope coefficient subscripts indicate fitting procedure (E for exponential decay with equation (1) and G for Gaussian with equation (2)) along with the spectral range of data utilized for the fit. For example, $S_{G300:700}$ would indicate the results from a fit with equation (2) from 300 to 700 nm.

2.2. Data Distribution

We focused on aggregating a_{CDOM} spectra into biogeochemical provinces following Longhurst (2006) biogeochemical province designations (provided by VLIZ (2009)). We considered all 54 Longhurst biogeochemical provinces and aggregated all inland water samples into an "INLAND" biogeochemical province, resulting in 55 potential regions. Of these, 35 were represented within the SeaBASS data set: 34 Longhurst provinces and the INLAND province (Figure 2). Grouping samples for the INLAND province was justified by two considerations: (1) sampling locations were all in the eastern half of the United States or the Laurentian Great Lakes, and (2) retrieved spectral slope was tightly grouped, ranging from 0.01 to 0.025 m^{-1} (see the discussion below regarding Figure 5). To be considered for analysis, we required each province to contain 5% of the maximum number of samples found within the most sampled province. While sample size is typically treated through consideration of group variance, we found that undersampled provinces displayed little variability typically due to all samples coming from a single oceanographic cruise. In effect, our threshold considered provinces with at least 100 spectra for spectral ranges <300 nm (less frequently sampled spectral range) and provinces with at least 200 spectra for spectral ranges ≥300 nm (more frequently sampled spectral range).

Absorption spectra were also considered within three depth classes defined as the first and second optical depths, calculated as 2.3 and 4.6 divided by $K_d(490)$, respectively, and below the photic zone (BPZ), comprising depths greater than the second optical depth but less than 1,500 m (Kirk, 1994). The majority of spectra below 1,500 m were sampled at the Bermuda Atlantic Times Series site; hence, we imposed the 1,500 m threshold to avoid a potential bias from deepwater spectra in this region. The depth classes for each province were identified by the average light extinction coefficient at 490 nm, $K_d(490)$, determined from MODIS-Aqua seasonal climatologies for each province. Seasonal $K_d(490)$ was used to ensure that samples from different seasons fell within a similar underwater light field to maintain relatively constant photodegradation rates across samples. Thus, a boreal winter sample near the lower limit of the first optical depth in a province could be a few meters deeper than a boreal summer sample near the upper limit of the second optical depth in the same province. Biogeochemical provinces for each depth class were included in the analysis using the same 5% of maximum sampling criterion described above for the complete data set, resulting in similar sample sizes. For most analyses, only the first optical depth is presented in the figures due to space constraints. Additional depth data for these analyses can be found in Tables S1–S3 in the supporting information. Seasonal distribution of data within the provinces did not allow representation across all four seasons; thus, results focus only on spatial trends in CDOM metrics.

2.3. Analysis of CDOM Metrics

Spectral slopes calculated for the Gaussian decomposition and standard exponential methodologies were compared using linear regression. $S_{G240:700}$, $S_{G300:700}$, and $S_{G350:550}$ were compared to $S_{E275:295}$ and $S_{E350:400}$ to determine their potential for estimating CDOM molecular weight, source, and degradation state using linear regression. Nonlinear least squares fits were used to assess the ability of log normalized $a_{CDOM}(\lambda)$ at wavelengths of 350, 412, and 443 nm to estimate $S_{G240:700}$, $S_{G300:700}$, $S_{G350:550}$, $S_{E350:400}$, and $S_{E275:295}$.

We sought to determine if observed differences in S_{CDOM} between biogeochemical provinces were significant. We first applied a one-way analysis of variance (ANOVA) to S_{CDOM} categorized by biogeochemical province to determine if the mean slope values between biogeochemical provinces were significantly different

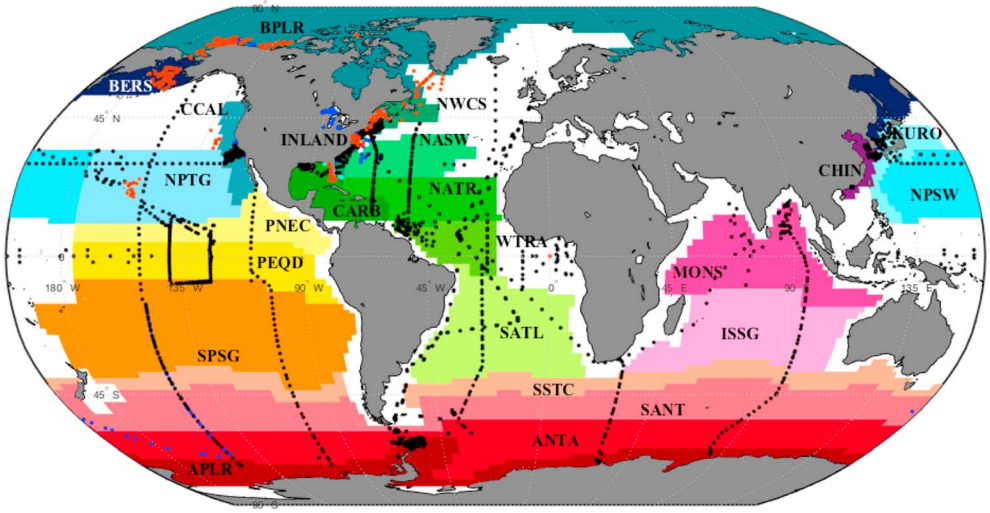

Figure 2. Map of a_{CDOM} observation locations for all spectral ranges considered: 350–550 nm (cyan, blue, black, and red), 300–700 nm (blue, black, and red), 275–295 nm (black and red), and 240–700 nm (red). The majority of INLAND samples are taken from rivers and lakes along the east and southeastern coasts of the United States or from the Laurentian Great Lakes. The colored regions indicate Longhurst provinces used for analysis of spatial trends in the data set (Longhurst, 2006). SATL, KURO, CHIN, NPSW, SSTC, ISSG, and ANTA were only used for $S_{350:400}$ and $S_{275:295}$ analyses due to data restrictions described in section 2.2. Following those restrictions, the data points indicated here in the uncolored regions are only used for aggregate slope comparisons depicted in Figures 3a–3c and are not used for any other analysis.

($p < 0.01$). The least significant difference multiple comparison of mean test was then applied using the results of the ANOVA to determine if the distribution of spectral slope mean values between provinces was significantly different ($p < 0.05$) (Rafter et al., 2002). It is recognized that the least significant difference (LSD) methodology offers looser statistical thresholds for significance than the Tukey-Kramer methodology (Hayter, 1986). However, as an exploratory exercise aimed at defining potential differences, the more liberal LSD methodology was deemed more suitable here.

Throughout section 3, multiple comparisons of mean tests are presented on bioplots as a way to visualize the degree of similarity in S_{CDOM} between different biogeochemical provinces. A line between boxes on a bioplot indicates mean S_{CDOM} values between those provinces are statistically similar, while no connection indicates that mean S_{CDOM} values between considered provinces are statistically unique. These results are further clarified within maps of the regions displaying which provinces are statistically similar and how similar (the number of related provinces) within groups displaying connectivity (statistical similarity). The bioplots illustrate which provinces displayed mean S_{CDOM} values that are explicitly similar, while the maps give a global representation of provinces that are statistically affiliated (all provinces are related to at least one other province within that group).

3. Results

3.1. Comparison of CDOM Models

We compared spectral slope retrieved with the standard exponential to the Gaussian decomposition approach (Figure 3). For all spectral ranges, the majority of points fell at or above the 1:1 line in Figure 3, demonstrating that the Gaussian decomposition fit resulted in identical or larger slope values for a given spectral range. When considering the spectral ranges investigated, differences between $S_{G240:700}$ and $S_{E240:700}$ were smallest for inland and coastal provinces and largest for open ocean provinces (Figure 4a), although this comparison relies on only one open ocean province (NPTG). Absolute differences between $S_{G300:700}$ and $S_{E300:700}$ did not show a distinct spatial trend, as both NWCS (coastal) and NASW (open ocean) contained the largest absolute differences between the methodologies (0.025 and 0.0279 nm^{-1}, respectively). However, overall mean differences between $S_{G300:700}$ and $S_{E300:700}$ were smallest for inland and coastal regions and increased in open ocean regions, while environments characterized by a greater

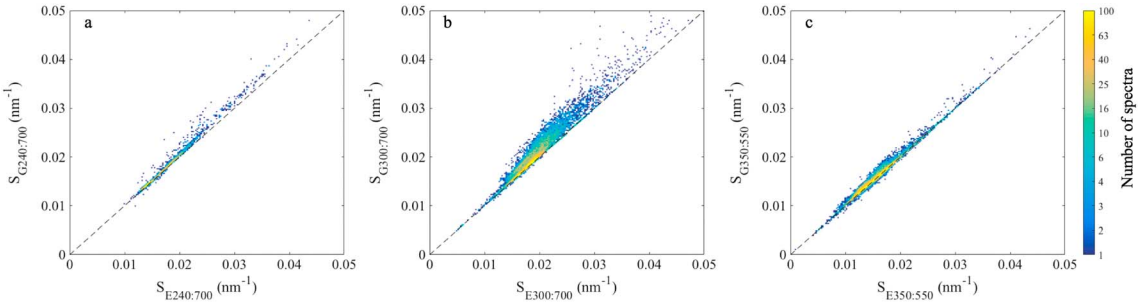

Figure 3. Probability density functions comparing S_E (equation (1)) versus S_G (equation (2)) for (a) 240–700 nm, (b) 300–700 nm, and (c) 350–550 nm spectral ranges. The dashed lines indicate a 1:1 relationship.

magnitude of a_{CDOM} (Arctic, inland, and coastal waters) more frequently displayed no difference in calculated slope between the two methodologies (Figure 4b). This is not surprising, as the Gaussian peaks fitted in the 300–700 and 350–550 nm spectral ranges for these environments were relatively small in magnitude. Following this expectation, differences between $S_{G350:550}$ and $S_{E350:550}$ were quite small overall and did not display any regional trends (Figure 4c).

3.1.1. Slope Values: Spectral and Spatial Trends

$S_{E275:295}$ displayed the largest variability in values across all biogeochemical provinces with coastal waters displaying smaller values, typically 0.02–0.03 nm^{-1}. Oceanic, Atlantic regions displayed larger values relative to all Pacific regions except NPSW (Figures 5a and 5f and Table S1). $S_{G240:700}$ displayed large differences between provinces influenced by terrestrial CDOM (INLAND, coastal, and Arctic waters) and slope values observed in NPTG. $S_{G300:700}$ displayed similar trends, with INLAND, coastal, and Arctic waters displaying slope values typically $<$0.025 nm^{-1}; however, it should be noted that coastal, sub-Arctic, and Arctic provinces displayed more variability in slope values than the INLAND province, particularly with a longer tailed distribution toward higher $S_{G300:700}$ values (Figure 5d and Table S1). Open ocean regions tended to display higher slope values, but variability and distribution of variables were higher and irregular. In contrast to $S_{G300:700}$, $S_{G350:550}$ trended toward higher values in coastal provinces and lower values in open ocean provinces, with some open ocean provinces displaying higher variability and irregular distributions (Figure 5e and Table S1). Overall, $S_{E275:295}$ displayed the broadest distribution of S_{CDOM} values, while the other spectral ranges clustered more tightly around a group mean in generally a normal distribution when considered across all biogeochemical provinces. Mean slope values for each spectral range were 0.036, 0.018, 0.021, 0.020, and 0.016 nm^{-1} for $S_{E275:295}$, $S_{E350:400}$, $S_{G240:700}$, $S_{G300:700}$, and $S_{G350:550}$, respectively, suggesting that S_{CDOM} values decrease as the spectral range considered shifts to longer wavelengths, an important bias to consider when choosing a spectral range for a given application.

While previous studies have shown a generally decreasing slope value with depth, $S_{E350:400}$ was the only spectral range that showed a consistent relationship between increasing depth zone and decreasing spectral slope value. $S_{G350:550}$ decreased with depth zone, $S_{E275:295}$ decreased below the first optical depth but did not display a trend between the second optical depth and below the photic zone for most provinces, while $S_{G240:700}$ and $S_{G300:700}$ displayed relatively consistent slope values across all depth zones for most provinces (Table S1). Overall, variability in the slope value decreased with depth for all spectral ranges considered.

3.1.2. Slope Values: Global Trends

The relationship between S_{CDOM} values and biogeochemical province was considered for $S_{E275:295}$, $S_{E350:400}$, $S_{G300:700}$, and $S_{G350:550}$. Provinces with statistically similar S_{CDOM} mean values are presented as clusters where each province presented displays similar mean S_{CDOM} values to at least one other province for that spectral range. $S_{E275:295}$ showed the clearest spatial relationships, with provinces dominated by high a_{CDOM} and riverine input displaying lower mean $S_{E275:295}$ values (0.024 nm^{-1}; Figures 6a and 6b) and open ocean regions, presumably dominated by photobleaching, displaying high mean $S_{E275:295}$ values (0.039 nm^{-1}, Figures 6c and 6d; 0.044 nm^{-1}, Figures 6e and 6f; 0.048 nm^{-1}, Figures 6g and 6h). Lower $S_{E275:295}$ values (0.026 nm^{-1}; Figures 6i and 6j) are found in provinces with a significant portion of samples taken in upwelling regions or regions where deep ocean waters pass over continental shelves (within the first optical depth). The remaining regions were either independent in their mean $S_{E275:295}$ value or displayed similarity with one other province

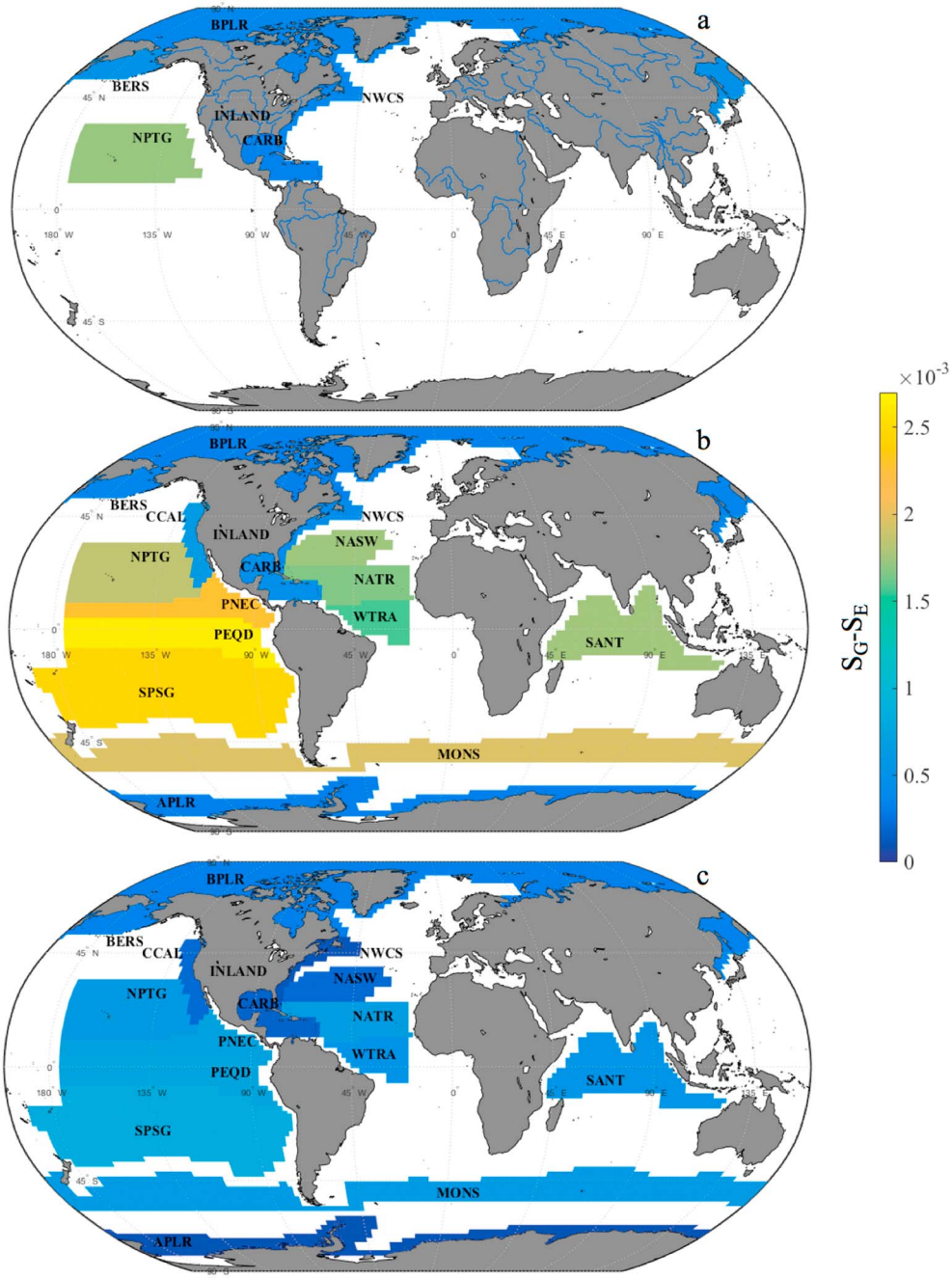

Figure 4. Mean difference between S_G and S_E for each province in the (a) 240–700 nm, (b) 300–700 nm, and (c) 350–550 nm spectral ranges.

(SANT and ANTA). INLAND displayed the lowest mean $S_{E275:295}$ value of 0.014 nm^{-1}. Spatial trends for $S_{E350:400}$ were not as clear, with high a_{CDOM} coastal, sub-Arctic, and Arctic provinces showing similar spectral slope values as all Atlantic provinces considered ($S_{E350:400}$ = 0.017 nm^{-1}, Figures 7a and 7b). Open ocean provinces from the Pacific, Indian, and Southern Oceans clustered with slope values either higher ($S_{E350:400}$ = 0.021 nm^{-1}, Figures 7c and 7d) or lower ($S_{E350:400}$ = 0.014 nm^{-1}, Figures 7e and 7f) than the $S_{E350:400}$ value from higher a_{CDOM} provinces, suggesting differences in the balance between photo- and microbial degradations relative to source signal. $S_{G300:700}$ showed little variation in mean slope value across all provinces while still producing three distinct clusters and INLAND as an independent province ($S_{G300:700}$ = 0.021 nm^{-1}, Figures 8a and 8b; $S_{G300:700}$ = 0.023 nm^{-1}, Figures 8c and 8d;

Figure 5. Spectral slope distribution for the first optical depth of each province with spectral range for each row as (a, f, k) $S_{E275:295}$, (b, g, l) $S_{E350:400}$, (c, h) $S_{G240:700}$, (d, i) $S_{G300:700}$, and (e, j) $S_{G350:550}$. Missing histograms in Figures 5c and 5h indicate provinces that did not have enough spectra measured down to 240 nm to be considered here.

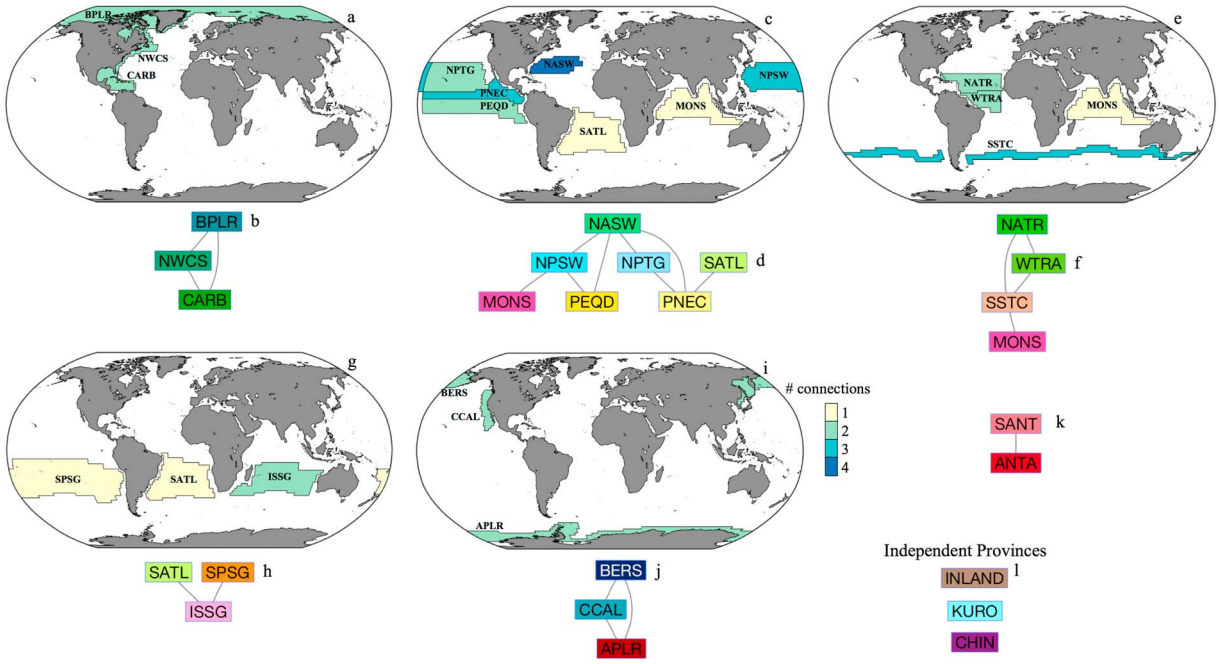

Figure 6. Global maps displaying provinces grouped based on results from the multiple comparison of mean test. The connecting lines indicate that provinces within the group are statistically similar. The province color indicates the number of provinces found to have a statistically similar mean slope based on the multiple comparison of mean test. The provinces that appear more than once were found to have a mean and uncertainty in the mean (described in section 2) overlapping with provinces in two different groups. Subplots display corresponding maps and bioplots as (a and b) Cluster 1 ($S_{E275:295}$ = 0.0244 nm^{-1}), (c and d)Cluster 2 ($S_{E275:295}$ = 0.0440 nm^{-1}), (e and f) Cluster 3 ($S_{E275:295}$ = 0.0399 nm^{-1}), (g and h) Cluster 4 ($S_{E275:295}$ = 0.0485 nm^{-1}), (i and j) Cluster 5 ($S_{E275:295}$ = 0.0267 nm^{-1}), and (k and bioplot only) SANT and ANTA ($S_{E275:295}$ = 0.0313 nm^{-1}) and (l) indicate provinces with unique mean $S_{275:295}$ values: INLAND ($S_{E275:295}$ = 0.0148 nm^{-1}), KURO ($S_{E275:295}$ = 0.0352 nm^{-1}), and CHIN ($S_{E275:295}$ = 0.0284 nm^{-1}).

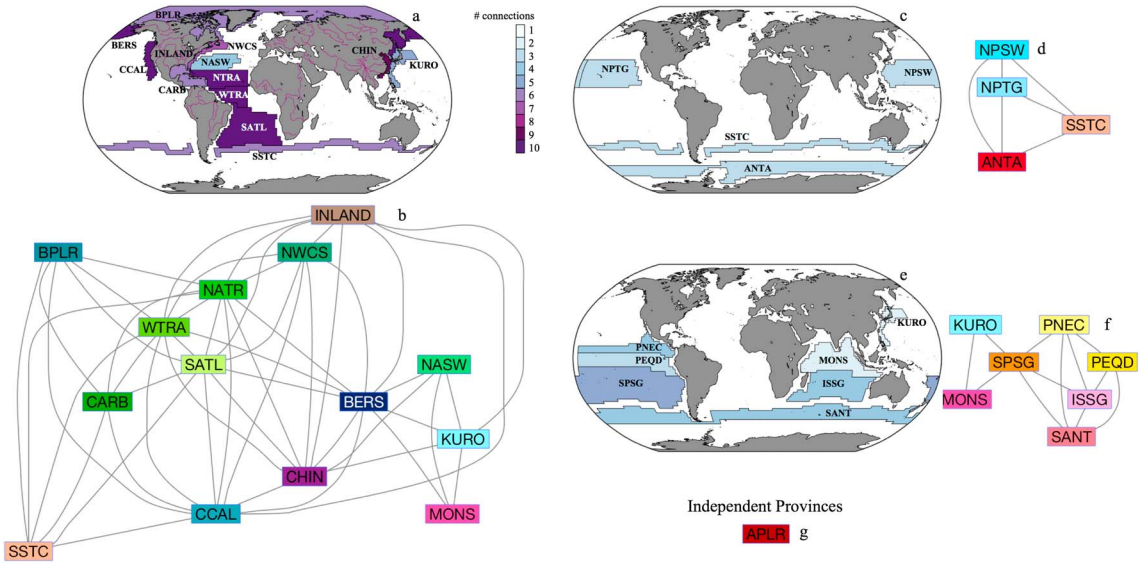

Figure 7. Global maps displaying provinces grouped based on results from the multiple comparison of means test. The connecting lines indicate that provinces within the group are statistically similar. The province color indicates the number of provinces found to have a statistically similar mean slope based on the multiple comparison of mean test. Provinces that appear more than once were found to have a mean and uncertainty in the mean (described in section 2) overlapping with provinces in two different groups. The subplots display corresponding maps and bioplots as (a and b) Cluster 1 ($S_{E350:400}$ = 0.0176 nm^{-1}), (c and d) Cluster 2 ($S_{E350:400}$ = 0.0210 nm^{-1}), and (e and f) Cluster 3 ($S_{E350:400}$ = 0.0230 nm^{-1}) to indicate groups with two or more statistically similar province means, while (g) APLR displayed a unique mean $S_{E350:400}$ value (0.0230 nm^{-1}).

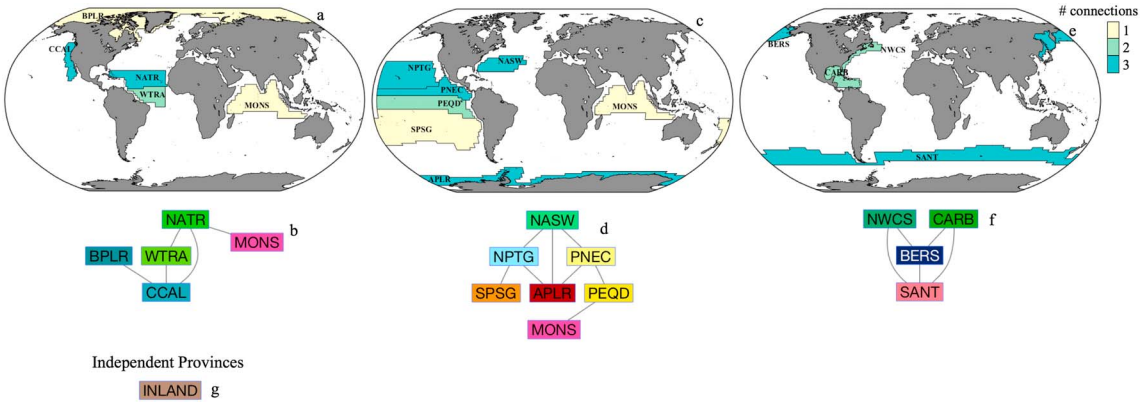

Figure 8. Global maps displaying provinces grouped based on results from the multiple comparison of mean test. The connecting lines indicate that provinces within the group are statistically similar. The province color indicates the number of provinces found to have a statistically similar mean slope based on the multiple comparison of mean test. Provinces that appear more than once were found to have a mean and uncertainty in the mean (described in section 2) overlapping with provinces in two different groups. The subplots display corresponding maps and bioplots as (a and b) Cluster 1 ($S_{G300:700}$ = 0.0212 nm^{-1}), (c and d) Cluster 2 ($S_{G300:700}$ = 0.0237 nm^{-1}), and (e and f) Cluster 3 ($S_{G300:700}$ = 0.0191 nm^{-1}), while (g) INLAND displayed a unique mean $S_{G300:700}$ (0.0166 nm^{-1}).

$S_{G300:700}$ = 0.019 nm^{-1}, Figures 8e and 8f; $S_{G300:700}$ = 0.016 nm^{-1}, Figure 8g). Clustering appears to be a mix of results from $S_{E275:295}$ and $S_{E350:400}$, suggesting that a lack of contribution from wavelengths below 300 nm and a broad spectral range lead to greater similarities between disparate CDOM pools. While also displaying a relatively narrow range of mean slope values, $S_{G350:550}$ showed the most connectivity between proximal provinces and region while also separating into four distinct clusters ($S_{G350:550}$ = 0.016 nm^{-1}, Figures 9a and 9b; $S_{G350:550}$ = 0.017 nm^{-1}, Figures 9c and 9d; $S_{G350:550}$ = 0.015 nm^{-1}, Figures 9e and 9f; $S_{G350:550}$ = 0.014 nm^{-1}, Figures 9g and 9h). APLR was the only independent province for this spectral range, displaying a significantly higher mean slope value ($S_{G350:550}$ = 0.023 nm^{-1}, Figure 9i).

3.2. Gaussian Components

The number of Gaussian components fit was highly dependent on the spectral range considered (Figure 10). $S_{G240:700}$ peak fitting was dominated by the lignin peak with the mean first Gaussian component location

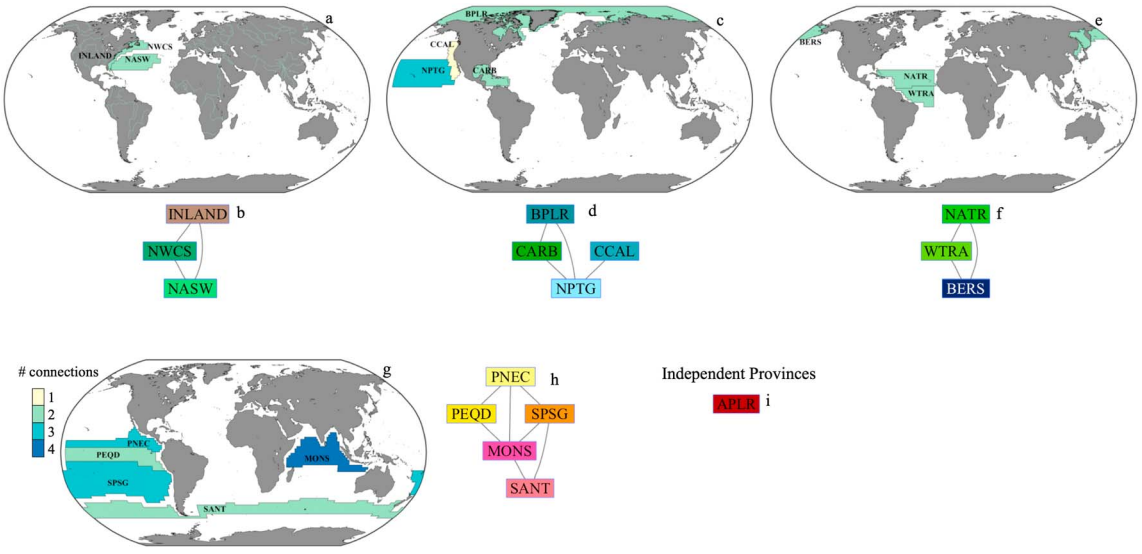

Figure 9. Global maps displaying provinces grouped based on results from the multiple comparison of mean test. The connecting lines indicate that provinces within the group are statistically similar. The province color indicates the number of provinces found to have a statistically similar mean slope based on the multiple comparison of mean test. Provinces that appear more than once were found to have a mean and uncertainty in the mean (described in section 2) overlapping with provinces in two different groups. The subplots display corresponding maps and bioplots as (a and b) Cluster 1 ($S_{G350:550}$ = 0.0167 nm^{-1}), (c and d) Cluster 2 ($S_{G350:550}$ = 0.0177 nm^{-1}), (e and f) Cluster 3 ($S_{G350:550}$ = 0.0156 nm^{-1}), and (g and h) Cluster 4 ($S_{G350:550}$ = 0.0140 nm^{-1}), while (i) APLR displayed a unique mean $S_{G350:550}$ value (0.0232 nm^{-1}).

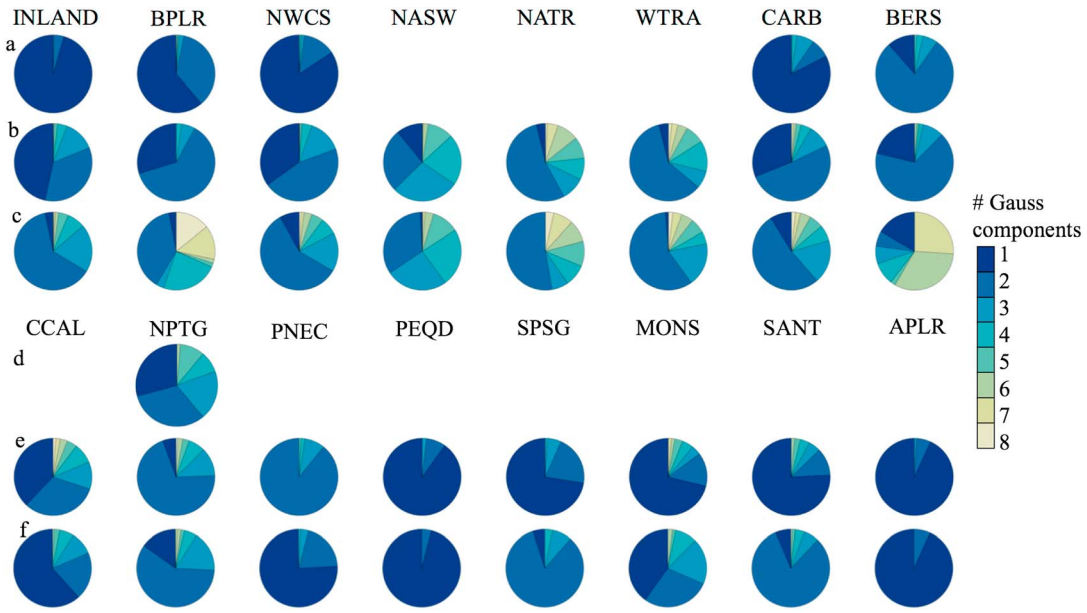

Figure 10. Number of Gaussian components indicated by color in the first optical depth for (a, d) 240–700 nm, (b, e) 300–700 nm, and (c, f) 350–550 nm.

occurring at 276 nm for all six biogeochemical provinces considered. Fewer Gaussian components were fit in INLAND and coastal provinces due to lignin peak fitting dramatically improving the goodness of fit (Figures 10a and 11a and Table S2). This is also evidenced by NPTG having the most Gaussian components fit relative to other provinces, while peak height was generally small for this province. Arctic and sub-Arctic provinces also saw an increase in fitted components relative to other provinces with relatively high a_{CDOM} (e.g., INLAND). For this spectral range, the location of the first component was strongly influenced by sample source and proximity to terrestrial material, as the first Gaussian component location was always below 300 nm for INLAND, NWCS, CARB, and BERS, while the first Gaussian component location ranged up to 385 and 381 nm for BPLR and NPTG, respectively. For $S_{G300:700}$, INLAND, sub-Arctic, Arctic, and Atlantic provinces were fit with more Gaussian components than Pacific and Southern Ocean provinces (Figures 10b and 11d and Table S2). This spectral range also observed the most fitted peaks across all provinces. $S_{G350:550}$ observed similar trends relative to ocean provinces as $S_{G300:700}$, but with fewer fitted peaks on average, likely due to peak height for components located in this fitting range being less prominent features of the absorption spectra (Figures 10c and 11g and Table S2). Interestingly, BPLR and BERS were fitted with significantly more peaks in this spectral range, potentially due to relatively high productivity and strong riverine influence in these regions leading to a complex CDOM pool. BPLR and BERS were fit with large Gaussian peaks below

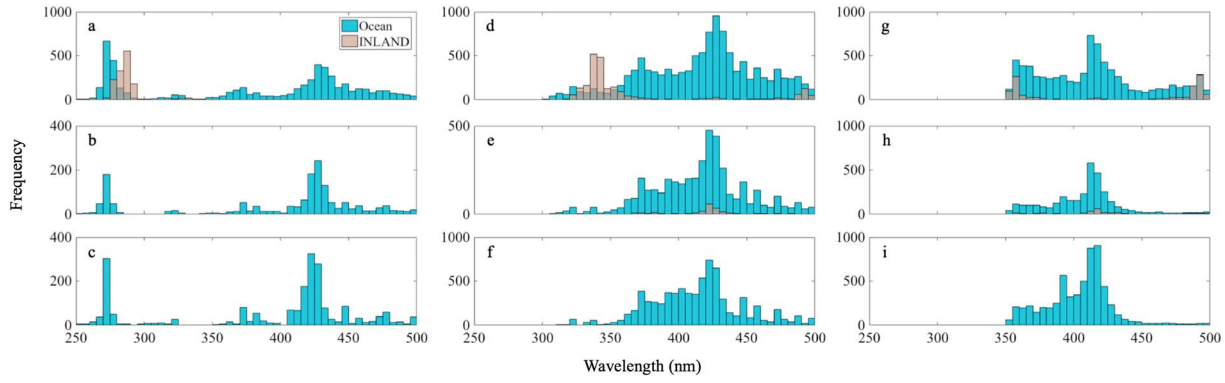

Figure 11. Distribution of Gaussian component spectral location between ocean and inland observations for (a–c) 240–700 nm, (d–f) 300–700 nm, and (g–i) 350–550 nm for the first (Figures 11a, 11d, and 11g) and second (Figures 11b, 11e, and 11h) optical depths and below the photic zone (Figures 11c, 11f, and 11i).

350 nm, also suggesting that fitting strictly above 350 nm captures smaller deviations in the absorption spectrum that are effectively masked with the improvement in fitting single, large deviations at shorter wavelengths.

Gaussian peak height, φ, was found to correlate quite well with $a_{CDOM}(350)$ (r^2 = 0.88), although the distribution appears bifurcated with a breakpoint between $a_{CDOM}(350) \leq 15$ m^{-1} (r^2 = 0.76) and $a_{CDOM}(350) > 15$ m^{-1} (r^2 = 0.90) (Figure S1 in the supporting information). Both relationships suggest a large dependency between φ and the magnitude of a_{CDOM} in the system. This effect was accounted for by normalizing φ with the modeled a_{CDOM} at the wavelength associated with the Gaussian peak location as described in section 2. Normalized φ showed that open ocean Gaussian components are more prominent relative to the magnitude of CDOM in the system when compared to terrestrial components, with INLAND mean normalized φ = 0.03, while ocean provinces had a mean normalized φ = 0.1 (data not shown). Overall, provinces with high a_{CDOM} tended to be fit with Gaussian peaks with larger peak height and width. Gaussian component location did not change significantly with depth for all spectral ranges considered (Figure 11 and Table S2), while Gaussian component location only displayed unique mean locations for INLAND ($S_{G300:700}$) and NASW ($S_{G300:700}$ and $S_{G350:550}$), suggesting that these features originate from a relatively small pool of chromophores that are likely unique to terrestrial and ocean regions. Across all spectral ranges, chromophores tended to be fit within spectral regions known to contain lignin species, photorefractory chromophores, and degradation products produced in situ (see section 4).

3.3. CDOM Metric Comparisons

3.3.1. Comparison of Broad Range Slopes to $S_{275:295}/S_{350:400}$

The influence of capturing absorption below 300 nm and the likely influence of lignin for relating broad range spectral slopes to absorption from $S_{275:295}$ is clear, as regions dominated by terrestrially sourced CDOM show tight relationships between $S_{E275:295}$ and $S_{G240:700}$ in the first optical depth (r^2 = 0.83–0.95; Table S3). Relationships between $S_{E275:295}$ and $S_{G300:700}$ were generally poor, while relationships between $S_{E275:295}$ and $S_{G350:550}$ entirely deteriorate for most regions and depths ($r^2 < 0.2$; Table S4), suggesting a distinct difference in the processes these spectral ranges describe. Notably, APLR a_{CDOM} spectra displayed consistent slope values across diverse spectral ranges.

As expected, $S_{E350:400}$ related inversely to broad slope ranges relative to $S_{E275:295}$. Most $S_{G240:700}$ correlations were relatively poor (r^2 = 0–0.34) with INLAND and CARB displaying relationships of 0.49 and 0.82, respectively, suggesting a broad influence of absorption by lignin or lignin derivatives in these regions. Overall, $S_{E350:400}$ showed distinct trends from spectral ranges capturing terrestrial features, suggesting a utility for monitoring in situ processes within this spectral range (Table S3).

3.3.2. Comparison Between a_{CDOM} and S_{CDOM}

We considered relationships between a_{CDOM} at 350, 412, and 443 nm and the calculated slope values to determine the feasibility of using the magnitude of a_{CDOM} at a reference wavelength to estimate S_{CDOM}. Wavelengths at or near 412 and 443 nm have applications to heritage satellite sensors, while 350 nm is the shortest wavelength anticipated to be available for NASA's Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) sensor, providing maximum signal to noise for retrieving a_{CDOM} while also avoiding overlap with phytoplankton pigments, including mycosporine-like amino acid absorption peaks around 330 and 360 nm (Pavlov et al., 2014). Relationships were best fit using a nonlinear least squares exponential fit. The goodness of fit did not change significantly between different reference wavelengths; subsequently, we only focus on $a_{CDOM}(350)$ in anticipation of future sensor capabilities. $S_{E275:295}$ had the strongest relationships with $a_{CDOM}(350)$, $S_{G300:700}$ relationships were highly variable, while $S_{E350:400}$ and $S_{G350:550}$ did not correlate with $a_{CDOM}(350)$ for almost all regions and depths considered (Table S3).

We also assessed the error introduced in satellite-derived estimates of $a_{CDOM}(412)$ by using an assumed S_{CDOM} value. CDOM has traditionally been considered alongside nonalgal particulate (NAP) material, as both have absorption spectra that follow an approximately exponentially increasing with decreasing wavelength relationship. These absorption terms, a_{CDOM} and a_{NAP}, are combined into a single term, colored detrital matter (CDM) absorption such that $a_{CDM} = a_{CDOM} + a_{NAP}$ and a_{CDM} have an average spectral slope (S_{CDM}) representative of the slope and percent contribution of each component. The GSM algorithm and the Quasi-Analytical Algorithm assume an S_{CDM} of 0.015 and 0.013–0.017 nm^{-1}, respectively (Lee et al., 2002;

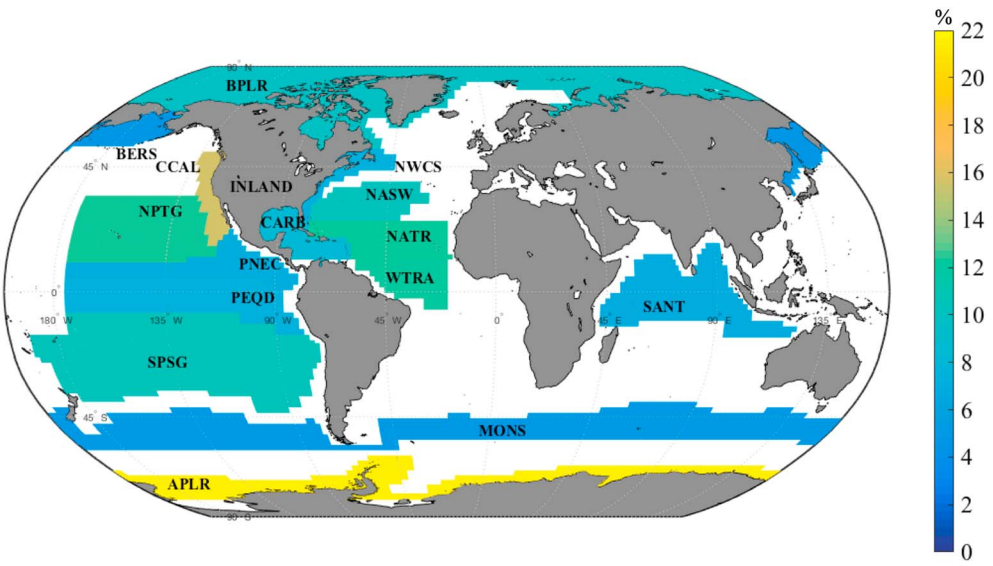

Figure 12. Percent error between mean observed $a_{CDOM}(412)$ for each province and calculated $a_{CDOM}(412)$ using mean observed $a_{CDOM}(443)$ and an assumed S_{CDOM} of 0.0154 nm^{-1}.

Maritorena et al., 2002). Global observations of a_{NAP} suggest it accounts for 10–20% of the a_{CDOM} signal, and a typical spectral slope for a_{NAP} is 0.011 nm^{-1} with a range of values much smaller than S_{CDOM} (Dong et al., 2013; Hoepffner & Sathyendranath, 1993; Roesler et al., 1989). Assuming an S_{CDOM} of 0.015 nm^{-1}, a contribution of 10% and 20% of a_{NAP} with spectral slope of 0.011 nm^{-1} results in an assumed S_{CDOM} of 0.016 and 0.0154 nm^{-1}, respectively. We assessed the percent error between mean $a_{CDOM}(412)$ observed for each province and $a_{CDOM}(412)$ calculated for the province using the mean $a_{CDOM}(443)$ and these assumed S_{CDOM} values. Error rates were similar when assuming 10% and 20% contributions from a_{NAP} to a_{CDOM}. When assuming a 20% contribution and an S_{CDOM} of 0.0154 nm^{-1}, estimated $a_{CDOM}(412)$ error varied from 4 to 22% of the observed mean $a_{CDOM}(412)$ (Figure 12).

4. Discussion

4.1. CDOM Models

Bias introduced in the S_{CDOM} parameter from varying the spectral range has been widely recognized for some time (Twardowski et al., 2004). Despite this understanding, only a few narrow range spectral slope parameters have been adopted consistently, $S_{275:295}$ and $S_{350:400}$ (Helms et al., 2008), with highly variable broad range slopes reported throughout the literature. Regardless of the CDOM model used, exponential or Gaussian decomposition, the bias due to spectral range considered remains an issue. S_{CDOM} differences between the exponential CDOM and Gaussian decomposition models are relatively minor for most spectra, with S_{CDOM} values generally higher for the Gaussian decomposition model. However, differences between the models suggest that removing deviations from the exponential model and modeling these features with Gaussian curves allows for a better characterization of the underlying exponential signal and a better fit of the measured CDOM spectra through improved r^2 and RMSE. This, in turn, will allow more consistent comparisons of S_{CDOM} from the same spectral range and a stronger basis for estimating CDOM composition optically (Del Vecchio & Blough, 2004a). In particular, we found the Gaussian decomposition method to be crucial for accurately modeling CDOM spectra strongly influenced by terrestrial material.

4.2. S_{CDOM}

Various spectral ranges have been used to characterize S_{CDOM}, with narrow range slopes typically focusing on specific attributes related to CDOM source, diagenetic state, and molecular weight (Helms et al., 2008; Spencer et al., 2008). The basis for linking optical properties to estimates of CDOM composition has strong backing in the literature (Boyle et al., 2009; Coble, 2007; Hernes & Benner, 2003; Kowalczuk et al., 2005;

Wünsch et al., 2015). Here we assessed how broad spectral range S_{CDOM} relates to common optical indicators of CDOM composition (e.g., $a_{CDOM}(\lambda)$ and $S_{275:295}$) and varies across spectral ranges, ocean biogeochemical provinces, and depth ranges.

CDOM absorption at a reference wavelength is often used as an indicator of CDOM composition as it relates well with $S_{E275:295}$ and lignin concentration (Fichot et al., 2016; Mannino et al., 2014). Past studies have shown strong relationships between a_{CDOM} and S_{CDOM} along transects from lower salinity coastal waters to higher salinity offshore waters where the range in $a_{CDOM}(\lambda)$ typically varies by orders of magnitude (Kowalczuk et al., 2006; Pavlov et al., 2016; Stedmon & Markager, 2003), while a study in the Arctic Ocean found that CDOM absorption and S_{CDOM} do not relate well between marine and terrestrially derived CDOM pools (Granskog et al., 2012). For our study, $S_{E275:295}$ displayed relatively strong relationships with $a_{CDOM}(350)$ for most provinces ($r^2 > 0.5$). Notably, APLR displayed a poor relationship between $a_{CDOM}(350)$ and S_{CDOM} across all spectral ranges, while most provinces displayed highly variable or poor relationships between $a_{CDOM}(350)$ and $S_{350:400}$. Relationships between broad range S_{CDOM} and a_{CDOM} were highly variable and were generally poor, even within the INLAND province where variability in $a_{CDOM}(350)$ is high across similar S_{CDOM} values, an observation also seen by Meler et al. (2016). We consider these inconsistencies within the literature and our findings as representative of a decoupling between $a_{CDOM}(\lambda)$ and S_{CDOM} in systems that display higher variability in one parameter over the other or when data sets are not taken within a single sampling period. Open ocean CDOM has been characterized as a mix of degraded terrestrial material and contributions from in situ production (Andrew et al., 2013), suggesting that spectral shape due to CDOM composition is reflected by varying contributions from these processes within the context of low $a_{CDOM}(350)$ values. We hypothesize that within oceanic regions with little terrestrial influence, process-specific S_{CDOM} variability (e.g., photodegradation and alteration by the microbial community) is not necessarily reflected in $a_{CDOM}(350)$ values at different times. These degradation processes occur regardless of the amount of CDOM; thus, an $a_{CDOM}(350)$ value of 0.1 m^{-1} could be affiliated with a spectral shape consistent with microbial alteration of the CDOM pool or with a spectral shape consistent with photodegraded terrestrial material depending on the time and/or location sampled. This is a potential bias of ship-based sampling that should be accounted for when attempting to accurately estimate a_{CDOM} from satellite-based remote sensing.

Spatial trends across all spectral ranges indicate that S_G is quite variable across the global ocean (Figure 5). Typically, global distributions of CDOM are presented as satellite-derived climatologies of colored dissolved and detrital matter (CDM) absorption (a_{CDM}), where nonalgal particulate (NAP) absorption (a_{NAP}) and CDOM are retrieved together and a_{NAP} typically accounts for $<$20% of CDM in the global ocean (Nelson, Siegel, & Michaels, 1998). These climatologies typically display low $a_{CDM}(\lambda)$ in subtropical ocean gyres and higher $a_{CDM}(\lambda)$ along continental margins, the equator, and high latitudes (Siegel et al., 2005). S_G did not follow spatial patterns observed in satellite observations of a_{CDM} for any spectral range considered, reinforcing a disconnect between the magnitude of $a_{CDOM}(350)$ and S_{CDOM}.

Existing methods that estimate a_{CDOM} from satellite radiometry requires assuming $S_{CDOM/CDM}$ or estimating a_{CDOM} by parameterizing nonalgal particulate absorption (a_{NAP}) through empirical relationships and treating a_{CDOM} as the residual (Matsuoka et al., 2013). Attempts to estimate S_{CDOM} directly have relied on removal of the a_{NAP} signal and adjusting an initially assumed S_{CDOM} using ratios of remotely sensed reflectance (Dong et al., 2013). We found that assuming an S_{CDOM} of 0.0154 nm^{-1}, equivalent to a 90% contribution of a_{CDOM} to the a_{CDM} signal (Nelson et al., 2010), and using $a_{CDOM}(443)$ result in errors in $a_{CDOM}(412)$ of 4–22% relative to the average $a_{CDOM}(412)$ observed in the provinces, suggesting that the assumed values or starting points used for S_{CDOM} should be regionalized to reduce uncertainty (Figure 12).

Retrieving S_{CDOM} independently without assumptions would be a first step toward estimating CDOM composition through a direct, rather than parameterized, observation. $S_{E350:400}$ can potentially be retrieved using NASA's PACE sensor. However, we found clearer separation between regions using $S_{G350:550}$, suggesting that accounting for a broader spectral range and Gaussian components within that range draws sharper contrasts between distinct CDOM pools in the global ocean. Neither $S_{E350:400}$ or $S_{G350:550}$ related well with $S_{E275:295}$; thus, both will likely be poor proxies for evaluating terrestrial contribution and, potentially, molecular weight (Helms et al., 2008).

Biogeochemical models assume an S_{CDOM} slope (Xiu & Chai, 2014) with some models accounting for microbial and photodegradation of CDOM over appropriate time scales (Dutkiewicz et al., 2015). While the inclusion of optical parameters in global biogeochemical ocean models is a recent development, further partitioning the spectral properties of CDOM by region would enhance the accuracy when modeling the underwater light field. Past studies have found changes in CDOM spectral characteristics with depth to be significant (Hickman et al., 2010; Pérez et al., 2016). We found changes in mean CDOM spectral characteristics to change up to 0.034, 0.02, 0.008, and 0.009 nm^{-1} between biogeochemical provinces, while changes across depth ranges within a given biogeochemical province varied up to 0.016, 0.003, 0.005, and 0.004 nm^{-1} for $S_{E275:295}$, $S_{G240:700}$, $S_{G300:700}$, and $S_{G350:550}$, respectively. While spatial variability in S_{CDOM} was greater between provinces than variability by depth within provinces, $S_{E350:400}$ displayed ranges up to 0.007 nm^{-1} between provinces and by depth within provinces. From this, we suggest that in situ production pathways for CDOM are more variable by depth than across global ocean provinces. To aid in future modeling and satellite remote sensing efforts, we have presented slope values for each spectral range, province, and depth range (Table S3).

We observed the differences between slope values measured across the spectral ranges for each province, suggesting that each spectral range does not convey the same information about CDOM composition. To understand what could be determined about CDOM composition from each spectral range, we considered how $S_{G240:700}$, $S_{G300:700}$, and $S_{E350:550}$ related to $S_{E275:295}$, which is regarded as an indicator of source, molecular weight, and photobleaching of CDOM (D'Sa et al., 2014; Helms et al., 2008), and $S_{E350:400}$, a slope range less entrenched in the literature but indicative of photobleaching, molecular weight, and microbial processing and production of CDOM (Helms et al., 2008, 2013, 2015; Matsuoka et al., 2015). We found that $S_{E350:400}$ likely characterizes in situ production and degradation of CDOM with potential links to microbial processes (Matsuoka et al., 2015; Nelson, Carlson, & Steinberg, 2004; Seidel et al., 2015) as this parameter varied throughout the global ocean, showed consistency across presumably diverse CDOM pools (e.g., INLAND and ocean provinces; Figure 7), and consistently increased with depth. $S_{G240:700}$ is strongly correlated with $S_{E275:295}$ in most regions (r^2 = 0.48–0.95, excluding BERS), suggesting that wavelengths below 300 nm strongly influence S_{CDOM} in this spectral range. $S_{G300:700}$ displayed variable relationships with $S_{E275:295}$ (r^2 = 0–0.90) and $S_{E350:400}$ (r^2 = 0.01–0.86). Typically, a region characterized by a poor relationship between $S_{G300:700}$ and $S_{E275:295}$ displayed a strong relationship between $S_{G300:700}$ and $S_{E350:400}$ (e.g., BPLR). A strong relationship with each parameter would be expected if $S_{G300:700}$ represents CDOM composition in a similar manner to $S_{E275:295}$ or $S_{E350:400}$. However, considering that a consistent trend with one over the other was not observed, it is likely that $S_{G300:700}$ blurs characteristics of each into a single parameter that may not be particularly effective at characterizing the CDOM pool except under ideal circumstances such as a single, dominant process contributing to the CDOM pool. Thus, past work that has found potential relationships between molecular weight and $S_{G300:600}$, a parameter we found to behave quite similarly to $S_{G300:700}$, may be contingent on the environment (e.g., Stedmon & Nelson, 2014).

Considering the variability and trends in slope for the different spectral ranges and relationships with $S_{E275:295}$, it is likely that $S_{G350:550}$ will be sensitive to a smaller suite of processes than spectral ranges that extend down to or below 300 nm. While $S_{G350:550}$ could be useful for accurately modeling a_{CDOM} above 350 nm, it appears to be limited for extending a_{CDOM} lower than 350 nm. $S_{G350:550}$ showed greater differentiation between biogeochemical provinces while maintaining a strong relationship with $S_{E350:400}$ in most provinces. Few studies have reported spectral slope from 350 to 550 nm to our knowledge (Hancke et al., 2014; Kowalczuk et al., 2006), making it difficult to directly assess what $S_{G350:550}$ estimates about CDOM composition. We propose that $S_{E350:400}$ is a better metric for tracking compositional changes in the CDOM pool affiliated with photodegradation and in situ production of CDOM as suggested by Helms et al. (2015, 2013). However, $S_{G350:550}$ displayed greater uniqueness between provinces, while $S_{E350:400}$ displayed more consistent trends with depth across all spectral ranges. If these trends are consistent in future data sets, we propose that $S_{G350:550}$ will be an ideal parameter for directly estimating compositional changes in autochthonous CDOM from hyperspectral ocean color data, while changes in $S_{E350:400}$ will be indicative of vertical transport of unique CDOM or distinct in situ production pathways. Based on previous studies, $a_{CDOM}(\lambda)$ will likely remain a useful parameter for estimating terrestrial CDOM contributions from hyperspectral satellite observations (Fichot, Lohrenz, & Benner, 2014; Mannino et al., 2014). These parameters appear to relate

well with oceanic, in situ processes with the potential for tracking vertical movement of the CDOM pool and informing estimates of CDOM composition.

4.3. Gaussian Components

The spectral range used to fit a_{CDOM} strongly impacted the number of Gaussian components fitted. When the first absorption peak was below 300 nm, presumed to be lignin (McKnight & Aiken, 1998), the goodness of fit increased so significantly that smaller peaks at longer wavelengths that were fitted in the 300–700 nm spectral range were no longer fitted. Fitting from 300 to 700 nm resulted in the most peaks fitted for all provinces, while 350–550 nm fitted fewer peaks than 300–700 nm suggesting that chromophores between 300 and 350 nm can be fit while still fitting features at wavelengths >350 nm. In waters where lignin is a strong or moderate contributor to a_{CDOM} below 300 nm, fitting from 240 to 500 nm, then fitting from 300 to 500 nm, may be a better approach, although blending models could become an issue. The majority of waters sampled to 240 nm in the SeaBASS data set occurred in waters that historically have a strong terrestrial component (Benner et al., 2005; D'Sa & DiMarco, 2009; Del Vecchio & Blough, 2004b; D'Sa et al., 2014), precluding a thorough analysis of the 240–700 nm spectral range across the full range of oceanic conditions observed in the entire SeaBASS data set. However, $S_{E275:295}$ suggests that the trend in spectral slope values that account for a_{CDOM} below 300 nm prevails in the global oceans. For spectra fitted with a peak below 300 nm, mean peak location was 286 nm, suggesting that $S_{E275:295}$ is strongly influenced by the shape of the Gaussian component, if present, rather than the underlying exponential curve. It is likely that the large, complex molecular structure of lignin and the absorption peak associated with lignin drive the relationship between $S_{E275:295}$ and CDOM molecular weight (McKnight & Aiken, 1998). Additionally, we observed a shift in the location of the lignin absorption peak from terrestrial waters to oceanic waters (Figure 11a) consistent with photodegradation of this component (Del Vecchio & Blough, 2004a).

Absorption peaks at wavelengths less than 300 nm can extend beyond these wavelengths through a complex process of charge-transfer interactions in the CDOM pool (Del Vecchio & Blough, 2004a). For waters strongly impacted by terrestrial material and displaying a low $S_{275:295}$ spectral signature, the first Gaussian component occurred at a much smaller wavelength in the spectra. Past studies have shown that terrestrial material absorption is dominated by lignin, which absorbs below 300 nm (McKnight & Aiken, 1998; Spencer et al., 2008); however, the deviation from the baseline associated with this peak extends beyond 300 nm (Fichot et al., 2016), resulting in a distortion of $S_{300:700}$ in these waters. Fitting Gaussian peaks provides a method to pick out unique components within the CDOM pool, similar to fitting fluorescence peaks in excitation-emission matrix spectroscopy and accounts for deviations that impact S_{CDOM} in the spectral range considered. More work is required to determine the significance of these absorption-based features to particular groups of molecules in the CDOM pool, including whether features fitted between 300 and 325 nm in the 300–700 nm spectral range are unique chromophores or a residual effect from lignin absorption extending above 300 nm.

Across all regions, spectral ranges that were commonly fitted were associated with spectral locations (~350 nm and 375 nm) of chromophores that are likely photorefactory (Helms et al., 2013) or chromophores that have been found to be produced from photobleaching of autochthonously produced CDOM (Swan et al., 2012). The latter, a feature observed between 410 and 420 nm and noted in previous studies (Bricaud et al., 2010; Swan et al., 2012), was the most commonly fitted peak across all provinces in the 300–700 nm spectral range. This feature was noticeably present across all spectral ranges, typically representing the second peak fitted in the 240–700 nm spectral range when more than one peak was fitted to the spectrum in this spectral range.

Two provinces that stood out regarding the average number of Gaussian peaks fitted per spectra were BPLR and BERS (Figure 10). BPLR was dominated by measurements in the Chukchi Sea, a region with a greater magnitude of CDOM than the global ocean but less than most Arctic shelf regions due to predominantly autochthonously produced CDOM (Dainard & Guéguen, 2013). However, both BPLR and BERS have elevated levels of CDOM, likely a contributing factor to significantly more fitted peaks on average than other provinces. Both provinces were predominantly sampled during or shortly after the intense spring phytoplankton blooms associated with ice melt in these regions (Arrigo et al., 2014; Goes et al., 2013). Phytoplankton absorption spectra in this region frequently display absorption in UV wavelengths due to the presence of mycosporine-like amino acids, a feature previously observed in the CDOM pool in Arctic habitats

(Pavlov et al., 2014). This, in conjunction with an active microbial community (Matsuoka et al., 2015), is a potential explanation for why these regions were dominated by spectra with an average of more than four Gaussian components. The dramatic reduction in number of Gaussian peaks fitted to a_{CDOM} spectra in BERS and BPLR from 300–700 nm to 350–550 nm despite a relatively constant mean location of peaks (when fitted) suggests that further consideration can be given to the weighting factor in environments with a diverse CDOM pool if spectral range is a limiting factor. In these environments, reducing the weighting factor removes more residuals, allowing for more, smaller peaks to potentially be fit if the data present an appropriate signal-to-noise (SNR) ratio while not distorting the BIC analysis. Locating and observing changes in these peaks can potentially provide insight into links between the phytoplankton community and CDOM as well as the degradative state of CDOM.

APLR spectra were typically fitted with few features and stood out as a unique province in most analyses (Figures 6–9). This is potentially due to low a_{CDOM} for some samples, but many samples contained $a_{CDOM}(350)$ greater than 0.1 m^{-1}. This province is characterized by autochthonously produced CDOM, with a distinct $S_{E275:295}$ signal and a high correlation of $a_{CDOM}(325)$ with chlorophyll concentrations and upwelled waters transporting subsurface water with elevated levels of CDOM into the photic zone (D'Sa & Kim, 2017; Ortega-Retuerta et al., 2009, 2010). For many of these spectra, the lack of components is likely due to old, upwelled CDOM that behaves remarkably consistent across all spectral ranges, evidenced as an approximate one-to-one line in slope comparisons across different spectral ranges (not shown). This feature was also seen in CCAL, another province seasonally driven by upwelling and displaying different CDOM signatures between upwelled waters and waters dominated by phytoplankton blooms (Day & Faloona, 2009). Considering that the number of Gaussian components decreased with depth for most provinces, zones of upwelling will likely display unique CDOM characteristics that include relatively uniform spectra with deviations primarily resulting from recent biological contributions.

The utility of Gaussian peak height, φ, and width, σ, are less certain from our analysis. When we normalized for peak height by modeled a_{CDOM} at the location of μ, oceanic waters presented a broader range of normalized φ including larger values, suggesting that these peaks are more prominent features relative to other chromophores contributing to a_{CDOM} in these regions. This suggests that these regions contain chromophores that are consistently produced amid a background of relatively degraded CDOM or are photorefractory in nature, consistent with the spectral locations of the peaks relative to previous studies (Helms et al., 2013; Swan et al., 2012).

We did not find any significant trends in σ for any spectral ranges considered. While μ and φ are relatively intuitive features, the parameterization of σ carries more uncertainty related to the methodology. While peak location and height can shift due to changes in C (weighting factor for residual removal) and spectral range used for fitting, they are largely grounded in features of the a_{CDOM} spectra as evidenced by similarity in location and height across spectral ranges used. Peak width can change dramatically based on C and, to a lesser extent, spectral range used, suggesting less interpretability. For our analysis, C was held constant at one for the entire data set. Thus, while our results are not impacted by changes in C, it is an important consideration for the community as the method becomes utilized more broadly.

4.4. Applications to Remote Sensing

Previous studies have found reliable relationships between a_{CDOM} at specific wavelengths, including $a_{CDOM}(412)$ and $a_{CDOM}(443)$, and spectral slope values, including $S_{E275:295}$ (e.g., Mannino et al., 2014). For our analysis, $S_{E275:295}$ could potentially be retrieved with reasonable accuracy in most biogeochemical provinces assuming an accurate retrieval of $a_{CDOM}(\lambda)$ and a predetermined relationship between these two parameters. However, the INLAND and BPLR regions displayed particularly poor relationships, suggesting high variability in CDOM pools and significant differences in CDOM characteristics with similar $a_{CDOM}(\lambda)$ values. Thus, some regions would be precluded from this methodology. Considering that most regions displayed unique trends and distributions of slope, a global relationship is anticipated to poorly predict $S_{E275:295}$.

NASA's PACE sensor is anticipated for launch in the near future and is expected to have hyperspectral (every 5 nm) capabilities down to 350 nm. Considering this, it is prudent to advance the knowledge of what can be determined regarding CDOM composition using a_{CDOM} in this spectral range. To this end, we

applied the methodology of Massicotte and Markager (2016) to identify spectral regions frequented by deviations that can impact satellite-derived estimates of chlorophyll *a* and phytoplankton pigments, particularly methods utilizing band ratios as the most common region of deviations occurred from 410 to 440 nm. Hyperspectral sensors may allow for a baseline exponential absorption spectra, such as $S_{E350:550}$, to be fitted to a_{CDOM} spectra, allowing for these features to be ignored if the SNR of the sensor does not allow for confident fitting of these features or if methodologies cannot separate out diverse inherent optical property (IOP) signals. An accurate $S_{E350:550}$ directly estimated from hyperspectral satellite data may also aid in accounting for deviations in a_{CDOM} that are currently centered on or near multispectral bands. Our approach allows for these spectral regions, and their frequency to be assessed to determine if regional accounting for these features can decrease error propagated through the spectrum by an ill-defined $a_{CDOM}(\lambda_0)$.

We also investigated the impact of inaccurate S_{CDOM} values for estimating $a_{CDOM}(412)$ by comparing calculated $a_{CDOM}(412)$ to mean $a_{CDOM}(412)$ for each province. The resultant errors, from 4 to 22%, suggest that poorly parameterizing S_{CDOM} lends significant uncertainty to satellite estimates of IOP's strictly introduced through S_{CDOM} or S_{CDM}. For multispectral algorithms, we suggest accounting for differences in S_{CDOM} between geographical regions, such as the provinces presented here, and continued consideration of the mean ratio of a_{NAP}:a_{CDOM} across distinct biogeochemical regions for methods that utilize S_{CDM}.

5. Conclusions

In an attempt to close the knowledge gap and move toward a common methodology, we have presented S_G values for three broad spectral ranges, 240–700 nm, 300–700 nm, and 350–550 nm as well as S_E for 275–295 nm and 350–400 nm. We also explored the ability of the Gaussian decomposition method to expand our optical understanding of CDOM composition from a global database. Ideally, S_{CDOM} can both adequately model the a_{CDOM} spectra and relate to CDOM characteristics. We presented S_{CDOM} as a diagnostic tool that can provide insights into CDOM composition with the potential to be applied to hyperspectral ocean color applications for optical estimates of CDOM composition.

Of the slope ranges considered, we found that $S_{E275:295}$ and $S_{E350:550}$ display clear, unique spatial trends that can be exploited for optically estimating changes in CDOM across diverse open ocean environments. $S_{E350:400}$ displayed potential for tracking changes to in situ production of CDOM, particularly with depth at a given location. $S_{G300:700}$ is a useful metric for particular regions if the underlying CDOM pool is known; however, this metric displayed good relationships with both $S_{E275:295}$ and $S_{E350:400}$, implying that the metric itself does not clearly delineate specific compositional traits that impact the optical signature of CDOM.

Hyperspectral capabilities may allow for direct estimates of S_{CDOM}, providing insight into CDOM degradative state and in situ production pathways. However, S_{CDOM} calculated using wavelengths anticipated to be available from NASA's PACE sensor (>~350 nm) differs significantly from $S_{E275:295}$, requiring alternative methods for estimating terrestrial contribution, lignin content, and molecular weight. The divergence in optical properties of $S_{E350:400}$ and $S_{E275:295}$ observed here and documented in the literature suggests that $S_{E350:400}$ can provide insight into in situ production pathways; however, further consideration should be given to this parameter in future studies along with whether a broader spectral range such as $S_{G350:550}$ provides more insight into these processes or clearer regional distinctions.

Beyond tracking changes in CDOM and presenting a way to estimate CDOM composition, S_{CDOM} is also important for accurately modeling a_{CDOM}. We considered the mean $a_{CDOM}(443)$ for each province and assessed how much error is introduced to the a_{CDOM} parameter using an assumed S_{CDOM} of 0.0154 nm^{-1} and propagating a_{CDOM} to 412 nm. We found that this assumed slope introduced errors in $a_{CDOM}(412)$ of 4–22% across all provinces relative to the mean $a_{CDOM}(412)$ observed (Figure 12). Thus, poorly parameterizing S_{CDOM} with a constant global slope can introduce a similar, if not greater, level of uncertainty in retrieved a_{CDOM} to mischaracterizing the percent contribution of a_{NAP} to $a_{CDM}(412)$.

Ultimately, CDOM will be best considered using a suite of metrics applied to ocean color imagery. Past work focusing on relationships between a_{CDOM} at a given reference wavelength to estimate $S_{E275:295}$ and lignin content can continue to be improved using in situ data and are anticipated to provide additional

information not directly available from NASA's anticipated PACE sensor. The emphasis should be on continuing to relate quality, in situ measurements with $S_{E350:400}$ and $S_{G350:550}$, two metrics anticipated to be directly available via NASA's PACE sensor in the near future, to maximize data potential from remotely sensed imagery. It is our view that, prior to mission launch, the community will be well served with a better understanding of what information is directly retrievable with S_{CDOM} and which spectral range is best suited for discrimination between distinct CDOM pools within the spectral capability of the mission.

Acknowledgments
We gratefully acknowledge the contributors to the SeaBASS data set (https://seabass.gsfc.nasa.gov/). Piotr Kowalczuk and an anonymous reviewer were crucial to shaping the final version of the manuscript. Philippe Massicotte was integral for clarifying the Gaussian decomposition method and sharing source code. Karl Meingast provided comments that shaped the final version of this manuscript. The research conducted here was funded by a NASA Earth and Space Science Fellowship awarded to Grunert.

References

Andrew, A. A., Del Vecchio, R., Subramaniam, A., & Blough, N. V. (2013). Chromophoric dissolved organic matter (CDOM) in the Equatorial Atlantic Ocean: Optical properties and their relation to CDOM structure and source. *Marine Chemistry*, *148*, 33–43.

Arrigo, K. R., Perovich, D. K., Pickart, R. S., Brown, Z. W., van Dijken, G. L., Lowry, K. E., … Swift, J. H. (2014). Phytoplankton blooms beneath the sea ice in the Chukchi sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, *105*, 1–16.

Benner, R., Louchouarn, P., & Amon, R. M. W. (2005). Terrigenous dissolved organic matter in the Arctic Ocean and its transport to surface and deep waters of the North Atlantic. *Global Biogeochemical Cycles*, *19*, GB2025. https://doi.org/10.1029/2004GB002398

Boyle, E. S., Guerriero, N., Thiallet, A., Vecchio, R. D., & Blough, N. V. (2009). Optical properties of humic substances and CDOM: Relation to structure. *Environmental Science & Technology*, *43*(7), 2262–2268.

Bricaud, A., Babin, M., Claustre, H., Ras, J., & Tièche, F. (2010). Light absorption properties and absorption budget of Southeast Pacific waters. *Journal of Geophysical Research*, *115*, C08009. https://doi.org/10.1029/2009JC005517

Carder, K. L., Steward, R. G., Harvey, G. R., & Ortner, P. B. (1989). Marine humic and fulvic acids: Their effects on remote sensing of ocean chlorophyll. *Limnology and Oceanography*, *34*(1), 68–81.

Coble, P. G. (2007). Marine optical biogeochemistry: The chemistry of ocean color. *Chemical Reviews*, *107*(2), 402–418.

D'Sa, E. J., & DiMarco, S. F. (2009). Seasonal variability and controls on chromophoric dissolved organic matter in a large river-dominated coastal margin. *Limnology and Oceanography*, *54*(6), 2233–2242.

Dainard, P. G., & Guéguen, C. (2013). Distribution of PARAFAC modeled CDOM components in the North Pacific Ocean, Bering, Chukchi and Beaufort Seas. *Marine Chemistry*, *157*, 216–223.

Day, D. A., & Faloona, I. (2009). Carbon monoxide and chromophoric dissolved organic matter cycles in the shelf waters of the northern California upwelling system. *Journal of Geophysical Research*, *114*, C01006. https://doi.org/10.1029/2007JC004590

Del Vecchio, R., & Blough, N. V. (2002). Photobleaching of chromophoric dissolved organic matter in natural waters: Kinetics and modeling. *Marine Chemistry*, *78*, 231–253.

Del Vecchio, R., & Blough, N. V. (2004a). On the origin of the optical properties of humic substances. *Environmental Science & Technology*, *38*(14), 3885–3891.

Del Vecchio, R., & Blough, N. V. (2004b). Spatial and seasonal distribution of chromophoric dissolved organic matter and dissolved organic carbon in the Middle Atlantic Bight. *Marine Chemistry*, *89*(1-4), 169–187.

Dong, Q., Shang, S., & Lee, Z. (2013). An algorithm to retrieve absorption coefficient of chromophoric dissolved organic matter from ocean color. *Remote Sensing of Environment*, *128*, 259–267.

D'Sa, E. J., Goes, J. I., Gomes, H., & Mouw, C. (2014). Absorption and fluorescence properties of chromophoric dissolved organic matter of the eastern Bering Sea in the summer with special reference to the influence of a cold pool. *Biogeosciences*, *11*(12), 3225–3244.

D'Sa, E. J., & Kim, H.-c. (2017). Surface gradients in dissolved organic matter absorption and fluorescence properties along the New Zealand sector of the Southern Ocean. *Frontiers in Marine Science*, *4*.

Dutkiewicz, S., Hickman, A. E., Jahn, O., Gregg, W. W., Mouw, C. B., & Follows, M. J. (2015). Capturing optically important constituents and properties in a marine biogeochemical and ecosystem model. *Biogeosciences*, *12*(14), 4447–4481.

Fichot, C. G., Benner, R., Kaiser, K., Shen, Y., Amon, R. M. W., Ogawa, H., & Lu, C.-J. (2016). Predicting dissolved lignin phenol concentrations in the coastal ocean from chromophoric dissolved organic matter (CDOM) absorption coefficients. *Frontiers in Marine Science*, *3*.

Fichot, C. G., Lohrenz, S. E., & Benner, R. (2014). Pulsed, cross-shelf export of terrigenous dissolved organic carbon to the Gulf of Mexico. *Journal of Geophysical Research: Oceans*, *119*, 1176–1194. https://doi.org/10.1002/2013JC009424

Goes, J. I., do Rosario Gomes, H., Haugen, E., McKee, K., D'Sa, E., Chekalyuk, A. M., … Sambrotto, R. (2013). Fluorescence, pigment and microscopic characterization of Bering Sea phytoplankton community structure and photosynthetic competency in the presence of a Cold Pool during summer. *Deep Sea Research Part II: Topical Studies in Oceanography*, *109*, 84–99.

Granskog, M. A., Stedmon, C. A., Dodd, P. A., Amon, R. M. W., Pavlov, A. K., de Steur, L., & Hansen, E. (2012). Characteristics of colored dissolved organic matter (CDOM) in the Arctic outflow in the Fram Strait: Assessing the changes and fate of terrigenous CDOM in the Arctic Ocean. *Journal of Geophysical Research*, *117*, C12021. https://doi.org/10.1029/2012JC008075

Hancke, K., Hovland, E. K., Volent, Z., Pettersen, R., Johnsen, G., Moline, M., & Sakshaug, E. (2014). Optical properties of CDOM across the polar front in the Barents Sea: Origin, distribution and significance. *Journal of Marine Systems*, *130*, 219–227.

Hansell, D. A., Carlson, C. A., Repeta, D. J., & Schlitzer, R. (2009). Dissolved organic matter in the ocean: A controversy stimulates new insights. *Oceanography*, *22*(4), 202–211.

Hansell, D. A. (2013). Recalcitrant dissolved organic carbon fractions. *Annual Review of Marine Science*, *5*, 421–445.

Hansen, A. M., Kraus, T. E. C., Pellerin, B. A., Fleck, J. A., Downing, B. D., & Bergamaschi, B. A. (2016). Optical properties of dissolved organic matter (DOM): Effects of biological and photolytic degradation. *Limnology and Oceanography*, *61*(3), 1015–1032.

Hayter, A. (1986). The maximum familywise error rate of Fisher's least significant difference test. *Journal of the American Statistical Association*, *81*(396), 1000–1004.

Helms, J. R., Mao, J., Chen, H., Perdue, E. M., Green, N. W., Hatcher, P. G., … Stubbins, A. (2015). Spectroscopic characterization of oceanic dissolved organic matter isolated by reverse osmosis coupled with electrodialysis. *Marine Chemistry*, *177*, 278–287.

Helms, J. R., Stubbins, A., Perdue, E. M., Green, N. W., Chen, H., & Mopper, K. (2013). Photochemical bleaching of oceanic dissolved organic matter and its effect on absorption spectral slope and fluorescence. *Marine Chemistry*, *155*, 81–91.

Helms, J. R., Stubbins, A., Ritchie, J. D., Minor, E. C., Kieber, D. J., & Mopper, K. (2008). Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. *Limnology and Oceanography*, *53*(3), 955–969.

Hernes, P. J., & Benner, R. (2003). Photochemical and microbial degradation of dissolved lignin phenols: Implications for the fate of terrigenous dissolved organic matter in marine environments. *Journal of Geophysical Research*, *108*(C9), 3291. https://doi.org/10.1029/2002JC001421

Hickman, A. E., Dutkiewicz, S., Williams, R. G., & Follows, M. J. (2010). Modelling the effects of chromatic adaptation on phytoplankton community structure in the oligotrophic ocean. *Marine Ecology Progress Series*, *406*, 1–17.
Hoepffner, N., & Sathyendranath, S. (1993). Determination of the major groups of phytoplankton pigments from the absorption spectra of total particulate matter. *Journal of Geophysical Research*, *98*(C12), 22,789–22,803.
Kim, G. E., Gnanadesikan, A., & Pradal, M.-A. (2016). Increased surface ocean heating by colored detrital matter (CDM) linked to greater Northern Hemisphere ice formation in the GFDL CM2Mc ESM. *Journal of Climate*, *29*(24), 9063–9076.
Kirk, J. T. (1994). Light and photosynthesis in aquatic ecosystems. Cambridge University Press.
Kowalczuk, P., Stedmon, C. A., & Markager, S. (2006). Modeling absorption by CDOM in the Baltic Sea from season, salinity and chlorophyll. *Marine Chemistry*, *101*(1-2), 1–11.
Kowalczuk, P., Stoń-Egiert, J., Cooper, W. J., Whitehead, R. F., & Durako, M. J. (2005). Characterization of chromophoric dissolved organic matter (CDOM) in the Baltic Sea by excitation emission matrix fluorescence spectroscopy. *Marine Chemistry*, *96*(3-4), 273–292.
Lee, Z., Carder, K. L., & Arnone, R. A. (2002). Deriving inherent optical properties from water color: A multiband quasi-analytical algorithm for optically deep waters. *Applied Optics*, *41*(27), 5755–5772.
Longhurst, A. R. (2006). *Ecological Geography of the Sea* (2nd ed., 560 pp.). San Diego, CA: Academic Press.
Mannino, A., Novak, M. G., Hooker, S. B., Hyde, K., & Aurin, D. (2014). Algorithm development and validation of CDOM properties for estuarine and continental shelf waters along the northeastern U.S. coast. *Remote Sensing of Environment*, *152*, 576–602.
Maritorena, S., Siegel, D. A., & Peterson, A. R. (2002). Optimization of a semianalytical ocean color model for global-scale applications. *Applied Optics*, *41*(15), 2705–2714.
Massicotte, P., & Frenette, J.-J. (2011). Spatial connectivity in a large river system: Resolving the sources and fate of dissolved organic matter. *Ecological Applications*, *21*(7), 2600–2617.
Massicotte, P., & Markager, S. (2016). Using a Gaussian decomposition approach to model absorption spectra of chromophoric dissolved organic matter. *Marine Chemistry*, *180*, 24–32.
Matsuoka, A., Hooker, S. B., Bricaud, A., Gentili, B., & Babin, M. (2013). Estimating absorption coefficients of colored dissolved organic matter (CDOM) using a semi-analytical algorithm for southern Beaufort Sea waters: Application to deriving concentrations of dissolved organic carbon from space. *Biogeosciences*, *10*(2), 917–927.
Matsuoka, A., Ortega-Retuerta, E., Bricaud, A., Arrigo, K. R., & Babin, M. (2015). Characteristics of colored dissolved organic matter (CDOM) in the Western Arctic Ocean: Relationships with microbial activities. *Deep Sea Research Part II: Topical Studies in Oceanography*, *118*, 44–52.
McKnight, D. M., & Aiken, G. R. (1998). Sources and age of aquatic humus. In D. O. Hessen & L. J. Tranvik (Eds.), *Aquatic Humic Substances: Ecology and Biogeochemistry* (pp. 9–39). Berlin, Germany: Springer.
Meler, J., Kowalczuk, P., Ostrowska, M., Ficek, D., Zabłocka, M., & Zdun, A. (2016). Parameterization of the light absorption properties of chromophoric dissolved organic matter in the Baltic Sea and Pomeranian lakes. *Ocean Science*, *12*(4), 1013–1032.
Nelson, N. B., Carlson, C. A., & Steinberg, D. K. (2004). Production of chromophoric dissolved organic matter by Sargasso Sea microbes. *Marine Chemistry*, *89*, 273–287.
Nelson, N. B., Siegel, D. A., Carlson, C. A., & Swan, C. M. (2010). Tracing global biogeochemical cycles and meridional overturning circulation using chromophoric dissolved organic matter. *Geophysical Research Letters*, *37*, L03610. https://doi.org/10.1029/2009GL042325
Nelson, N. B., Siegel, D. A., & Michaels, A. F. (1998). Seasonal dynamics of colored dissolved material in the Sargasso Sea. *Deep Sea Research, Part I*, *45*, 931–957.
Organelli, E., Bricaud, A., Antoine, D., & Matsuoka, A. (2014). Seasonal dynamics of light absorption by chromophoric dissolved organic matter (CDOM) in the NW Mediterranean Sea (BOUSSOLE site). *Deep Sea Research Part I: Oceanographic Research Papers*, *91*, 72–85.
Ortega-Retuerta, E., Frazer, T., Duarte, C. M., Ruiz-Halpern, S., Tovar-Sanchez, A., Arrieta, J. M., & Reche, I. (2009). Biogeneration of chromophoric dissolved organic matter by bacteria and kirll in the Southern Ocea. *Limnology and Oceanography*, *54*(6), 1941–1950.
Ortega-Retuerta, E., Reche, I., Pulido-Villena, E., Agustí, S., & Duarte, C. M. (2010). Distribution and photoreactivity of chromophoric dissolved organic matter in the Antarctic Peninsula (Southern Ocean). *Marine Chemistry*, *118*, 129–139.
Pavlov, A. K., Silyakova, A., Granskog, M. A., Bellerby, R. G. J., Engel, A., Schulz, K. G., & Brussaard, C. P. D. (2014). Marine CDOM accumulation during a coastal Arctic mesocosm experiment: No response to elevated pCO2levels. *Journal of Geophysical Research: Biogeosciences*, *119*, 1216–1230. https://doi.org/10.1002/2013JG002587
Pavlov, A. K., Stedmon, C. A., Semushin, A. V., Martma, T., Ivanov, B. V., Kowalczuk, P., & Granskog, M. A. (2016). Linkages between the circulation and distribution of dissolved organic matter in the White Sea, Arctic Ocean. *Continental Shelf Research*, *119*, 1–13.
Pérez, G. L., Galí, M., Royer, S.-J., Sarmento, H., Gasol, J. M., Marrasé, C., & Simó, R. (2016). Bio-optical characterization of offshore NW Mediterranean waters: CDOM contribution to the absorption budget and diffuse attenuation of downwelling irradiance. *Deep Sea Research Part I: Oceanographic Research Papers*, *114*, 111–127.
Rafter, J. A., Abell, M. L., & Braselton, J. P. (2002). Multiple comparison methods for means. *SIAM Review*, *44*(2), 259–278.
Reche, I., Pace, M. L., & Cole, J. J. (2000). Modeled effects of dissolved organic carbon and solar spectra on photobleaching in lake ecosystems. *Ecosystems*, *3*(5), 419–432.
Roesler, C. S., Perry, M. J., & Carder, K. L. (1989). Modeling in situ phytoplankton absorption from total absorption spectra in productive inland marine waters. *Limnology and Oceanography*, *34*(8), 1510–1523.
Schwarz, J. N., Kowalczuk, P., Kaczmarek, S., Cota, G. F., Mitchell, B. G., Kahru, M., … Raine, R. (2002). Two models for absorption by coloured dissolved organic matter (CDOM). *Oceanologia*, *44*(2), 209–241.
Seidel, M., Yager, P. L., Ward, N. D., Carpenter, E. J., Gomes, H. R., Krusche, A. V., … Medeiros, P. M. (2015). Molecular-level changes of dissolved organic matter along the Amazon River-to-ocean continuum. *Marine Chemistry*, *177*, 218–231.
Sharpless, C. M., Aeschbacher, M., Page, S. E., Wenk, J., Sander, M., & McNeill, K. (2014). Photooxidation-induced changes in optical, electrochemical, and photochemical properties of humic substances. *Environmental Science & Technology*, *48*(5), 2688–2696.
Siegel, D. A., Maritorena, S., Nelson, N. B., Behrenfeld, M., & McClain, C. R. (2005). Colored dissolved organic matter and its influence on the satellite-based characterization of the ocean biosphere. *Geophysical Research Letters*, *32*, L20605. https://doi.org/10.1029/2005GL024310
Spencer, R. G. M., Aiken, G. R., Wickland, K. P., Striegl, R. G., & Hernes, P. J. (2008). Seasonal and spatial variability in dissolved organic matter quantity and composition from the Yukon River basin, Alaska. *Global Biogeochemical Cycles*, *22*, GB4002. https://doi.org/10.1029/2008GB003231
Stedmon, C. A., & Nelson, N. B. (2014). The optical properties of DOM in the ocean. In *Biogeochemistry of Marine Dissolved Organic Matter* (2nd ed., pp. 481–508).
Stedmon, C. A., Amon, R. M. W., Rinehart, A. J., & Walker, S. A. (2011). The supply and characteristics of colored dissolved organic matter (CDOM) in the Arctic Ocean: Pan Arctic trends and differences. *Marine Chemistry*, *124*(1-4), 108–118.

Stedmon, C. A., & Markager, S. (2003). Behaviour of the optical properties of coloured dissolved organic matter under conservative mixing. *Estuarine, Coastal and Shelf Science*, *57*(5-6), 973–979.
Stedmon, C. A., & Markager, S. (2005). Resolving the variability in dissolved organic matter fluorescence in a temperate estuary and its catchment using PARAFAC analysis. *Limnology and Oceanography*, *50*(2), 686–697.
Swan, C. M., Nelson, N. B., Siegel, D. A., & Kostadinov, T. S. (2012). The effect of surface irradiance on the absorption spectrum of chromophoric dissolved organic matter in the global ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, *63*, 52–64.
Twardowski, M. S., Boss, E., Sullivan, J. M., & Donaghay, P. L. (2004). Modeling the spectral shape of absorption by chromophoric dissolved organic matter. *Marine Chemistry*, *89*(1-4), 69–88.
Visser, S. A. (1983). Fluorescence phenomena of humic matter of aquatic origin and microbial cultures. *Aquatic and terrestrial humic materials*, 183–202.
VLIZ (2009). Longhurst biogeographical provinces. Retrieved from http://www.marineregions.org
Walker, S. A., Amon, R. M. W., & Stedmon, C. A. (2013). Variations in high-latitude riverine fluorescent dissolved organic matter: A comparison of large Arctic rivers. *Journal of Geophysical Research: Biogeosciences*, *118*, 1689–1702. https://doi.org/10.1002/2013JG002320
Werdell, P. J., Bailey, S. W., Fargion, G. S., Pietras, C., Knobelspiesse, K. D., Feldman, G. C., & McClain, C. R. (2003). Unique data repository facilitates ocean color satellite validation. *Eos, Transactions American Geophysical Union*, *84*(38), 377.
Wünsch, U. J., Murphy, K. R., & Stedmon, C. A. (2015). Fluorescence quantum yields of natural organic matter and organic compounds: Implications for the fluorescence-based interpretation of organic matter composition. *Frontiers in Marine Science*, *2*.
Xiu, P., & Chai, F. (2014). Connections between physical, optical and biogeochemical processes in the Pacific Ocean. *Progress in Oceanography*, *122*, 30–53.