Authors Response to Reviewer 1 Comments for Manuscript egusphere-2024-4163: *The primacy of dissolved organic matter to aquatic light variability*

9 May 2025

The authors thank Reviewer 1 for their valuable comments. The authors respond to all comments below and document resulting modifications to the manuscript, as appropriate. Reviewer 1's comments are shown in *slanted* typeface. The Authors responses are shown in plain typeface. Revised or added text is indicated in red. All line numbers are indicated with a capital "L" and refer to the line numbers in the original submitted work. Citations within this response document correspond to the References section at the end of this document.

- **Comments from Reviewer 1:** *The manuscript by Houskeeper and Hooker presents a novel perspective on the primacy of colored dissolved organic matter (CDOM) over chlorophyll a (C_a) in driving aquatic light variability. This manuscript makes a significant contribution to the field of ocean optics by challenging long-standing assumptions about the primacy of C_a in driving aquatic light variability. Their comprehensive analysis of three independent bio-optical datasets provides compelling evidence that CDOM absorption—not C_a—represents the dominant factor influencing spectral variability in aquatic environments. This finding has profound implications for ocean color remote sensing, potentially reshaping algorithm development and improving our ability to monitor marine ecosystems from space. The authors' demonstration that expanded spectral ranges (including UV and NIR domains) improve the independent retrieval of optical constituents provides timely insights for utilizing data from new hyperspectral satellite missions like PACE. Their work effectively bridges historical perspectives on ocean optics with contemporary understanding of microbial loop dynamics, offering a more nuanced view of the biogeochemical processes influencing marine optical properties. Overall, the paper represents a valuable contribution that could significantly advance our ability to interpret and utilize ocean color observations in a changing climate.*

  Authors Response 1: The authors thank the reviewer for the positive comments.

  While their eigendecomposition analysis is thorough, I have concerns about terminology and methodology. The authors repeatedly refer to "EMA" (end-member analysis) without adequately explaining this approach or its relation to established methodologies. Have they invented it? If so, I am not sure it warrants a name nor acronym.

  Authors Response 2: End-member analysis (EMA) is an algorithmic approach wherein information is captured from the most spectrally separated (e.g., ultraviolet, UV, and near-infrared, NIR)—rather than from the internal (primarily VIS)—wavelengths. The concept was first described in Hooker et al. (2013) and the terminology was introduced in Hooker et al. (2020), which also noted that this approach provides continuity between historical multispectral observations and forthcoming hyperspectral observations:

> "...the simplified approach provided by end-member analysis can be used with both legacy and next-generation sensors, thereby providing continuity in space and time as well as a capability to generate high-quality in-water data with a simplified measurement approach..."

The EMA approach was subsequently leveraged to develop aquatic remote sensing algorithms or investigate optical relationships in Houskeeper et al. (2021,2022) and Hooker et al. (2021a,b). Greater community adoption—and therefore familiarity—with EMA is, perhaps, limited by the sparsity of compliant UV and NIR aquatic radiometric datasets.

The authors clarify this topic by better defining EMA at its first usage (L016–017), as follows:

> "...based on using ratios of the most spectrally separated optical data products (Hooker et al. 2013), especially those from the UV and NIR spectral domain, an approach hereafter termed end-member analysis (EMA) following Hooker et al. (2020). Conservative waters..."

After investigating their cited works and testing their CDOM prescription against the Loisel et al. (2023) dataset (see: `https://github.com/ocean-colour/bing/blob/ema/nb/EMA/Explore_EMA.ipynb`*), I found their approach exhibits sensitivity to algorithm variants that warrants further discussion. The authors appear to have developed this method across several publications, but it requires more explicit contextualization for readers unfamiliar with their previous work, particularly regarding how different implementations might affect results.*

Authors Response 3: The authors revise elements of the manuscript to provide context based on the reviewer's point. But first, some clarifications are warranted regarding the reviewer's referenced Python notebook.

Briefly, the reviewer has attached a notebook using data wherein synthetic apparent optical properties (AOPs), e.g., $\left[L_W(\lambda)\right]_N$, are synthesized from various datasets of inherent optical properties (IOPs), e.g., $a_{\text{CDOM}}(440)$. For individual AOP observations in this dataset, the suite of IOPs needed to derive the AOP do not represent the same set of *in situ* observations, i.e., they include dissimilar sampling times or locations. IOPs are randomly assembled using each IOP's distribution, and from these random assemblages, AOPs such as $\left[L_W(\lambda)\right]_N$ are derived. The synthetic dataset, therefore, is not constrained to the associations that exist among *in situ* IOPs in the ocean (Morel 2009) or aquatic environments.

Conversely, the applicable algorithm from Houskeeper et al. (2021), i.e., GLOBC, was derived using *in situ* observations of aquatic environments. GLOBC thereby includes *in situ* relationships in IOPs—generally discussed in terms of optically active constituents (OACs) in the manuscript under review.

The reviewer's notebook tests the performance of EMA algorithms using both the applicable tuning (GLOBC), as well as two inapplicable tunings (OCEAN and NOMAD). The latter correspond to observations derived using legacy above- and in-water instrumentation, with in-water data products derived in a deep extrapolation column with coarse vertical sampling resolution (VSR), and above-water data products derived using slow sampling speeds (on the order of 0.5 Hz) and low signal-to-noise ratio (SNR) observations (Kudela et al. 2024).

The reviewer's notebook graphic shows two characteristics, both of which the authors consider to be consistent (not inconsistent) with the published literature:

1. The reviewer's graphic shows that synthetic data products produce increased scatter and nonlinearity compared to *in situ* data products. This finding is anticipated and consistent with the literature, as follows: random re-organization of IOPs removes the OAC covariances observed within *in situ* water bodies; removing OAC covariances increases variability in OACs; and increasing variability in OAC relationships increases scatter in visible (VIS) algorithms.

   Algorithm sensitivity to variability in OACs was shown to be a function of algorithmic spectral range, which was the primary finding of Houskeeper et al. (2021). Non-visible—or invisible (INV)—algorithms are more robust to OAC variability than VIS algorithms. The reviewer's notebook implements VIS algorithms and has shown, as is expected, that removing IOP or OAC covariances increases the scatter for VIS algorithms.

   For example, consider this text from Houskeeper et al. (2020):

   > "...correlation to $a_{\mathrm{CDOM}}(440)$ increase[s] with greater spectral separation of the waveband ratio toward ultraviolet and near-infrared wavelengths."

   Another example, from Houskeeper et al. (2021):

   > "For all datasets, the dynamic range in [the waveband ratio] increases with increasing spectral separation between wavelength pairs, which is anticipated to increase the sensitivity and robustness of the algorithmic approach."

   And another, from the submitted manuscript (L445–447):

   > "Linearity and loglinearity of CDOM algorithms decreases with decreasing spectral separation of algorithm end members (Hooker et al. 2020, 2021a; Houskeeper et al. 2021) because increasing overlap in OAC absorption properties adds nonlinearity (Houskeeper et al. 2020a)."

Regarding relationships among OACs, the observed covariances within *in situ* water bodies have been fundamental to the development of ocean color algorithms, namely the Ocean Chlorophyll (OC) family of algorithms. The authors' manuscript does not argue against the existence of covariances among OACs within oceanic and global *in situ* data products, which are well documented (Morel et al. 2009; Hooker et al. 2021a). Rather, the authors have previously shown that EMA algorithms—especially INV but also VIS implementations—mitigate some of the algorithmic sensitivity to variability in the OAC relationships.

Finally, the scatter in the notebook graphic for the applicable algorithm (GLOBC) appears to be consistent with the uncertainty reported in Houskeeper et al. (2021) for the VIS algorithm, i.e., 45% (although uncertainty metrics are not presented in the graphic).

2. The reviewer's graphic shows that algorithms become biased if retuned using legacy data products (i.e., the OCEAN and NOMAD tunings), which are often nonphysical, e.g., negative or signal limited, and biased. The authors strongly agree and have discussed this challenge in many contexts (Hooker et al. 2020, 2021a, 2021b; Houskeeper et al. 2020, 2021, 2023, 2024; and Kudela et al. 2019, 2024). The reviewer's graphic provides additional support for this trajectory of literature by showing that algorithms derived using legacy AOP data products (most notably those of NOMAD) are strongly biased compared to AOP data products synthesized using radiative transfer. The point is consistent with the submitted manuscript, which did not use NOMAD data products based, in part, on the previously published findings indicating bias.

The authors revise the text to help clarify elements related to implementation of the EMA algorithms.

First, the authors clarify the tuning (to avoid readers implementing the OCEAN or NOMAD coefficients from the Python notebook) by specifying the applicable coefficients directly in Eq. 2, as follows:

$$a_{\text{CDOM}}(440) \; = \; 0.242 \left[\Lambda_{\lambda_2}^{\lambda_1} \right]^{-0.961}, \qquad\qquad [\text{m}^{-1}] \qquad (1)$$

The authors add discussion regarding decreased performance of VIS implementations compared to INV implementations at L113, as follows:

"...indicating the importance of accurately deriving UV data products for robust CDOM estimation. VIS algorithms, conversely, were shown to exhibit decreased robustness to variability in OAC relationships relative to INV algorithms (Hooker et al. 2020, 2021b; Houskeeper et al. 2020, 2021).

The authors revise text on L601–603 to clarify EMA is leveraged herein is improved robustness to OAC relationships, and the improvement is dependent on the spectral range (e.g., VIS compared to INV), as follows:

"Spectrally expansive data products have been shown to improve retrieval of $a_{\mathrm{CDOM}}(440)$ independent of C_a (Sathyendranath et al. 1987; Hooker et al. 2020; Houskeeper et al. 2021), attributed, in part, to improved robustness to variability in the relationships between OACs. The comparisons herein of eigenanalyses using the INV21 and VIS21 spectral subsets further support the hypothesis that an expansive spectral range for data products improves the separability of signals.

The authors improve discussion of the limitations by adding text regarding the efficacy of EMA when applied to legacy datasets (wherein spectral range is generally confined to VIS wavelengths). Most notably, the RSE2022 dataset corresponds to much coarser VSR plus depth aliasing (see L215–217 of the manuscript) and commercial-off-the-shelf (COTS) spectrometers limiting the SNR (Kudela et al. 2019; 2024). The authors add text to L265, as follows:

"The uncertainty of EMA is anticipated to increase when applied to optical data products obtained using spectrometers and rocket-shaped profilers (e.g., RSE2022) due to increasing extrapolation depths combined with the strong attenuation of signal, particularly within the end-member spectral domains (Kudela et al. 2019, 2024; Hooker et al. 2020; Houskeeper et al. 2021)."

Deficient data quality (most applicable to RSE2022 which corresponds to degraded SNR and VSR) is presently managed, in part, by applying quality control metrics to identify and remove spectra that are likely nonphysical. For example, oceanic spectra that are significantly darker than the darkest global waters (as observed using much higher SNR and finer VSR technologies) are associated with SNR-limited measurements or depth aliasing within vertical profiles. The authors add discussion of sensitivity testing regarding quality control to L238, as follows:

"...the threshold removal of nonphysical radiometric brightness was thus applied, albeit only using VIS data products (given that RSE2007 and RSE2022 are restricted to the VIS domain), following sensitivity testing. Briefly, more stringent quality control metrics (e.g., requiring observations to comply with the brightness ranges of the globally representative RSE2021 dataset) produced similar eigenanalyses results but reduced the size of RSE2022 by approximately 50%. Alternately, relaxed quality control thresholds permitting radiometric values up to 50 times darker than the global dataset did not alter the primary findings of RSE2022, and no threshold relaxation affected the primary findings of RSE2007. Considering acknowledged uncertainties relating to hardware, processing, data acquisition, and water mass differences, the quality control thresholds were performed within this insensitivity range by..."

The manuscript notably omits sufficient discussion of detrital absorption, which significantly contributes to absorption at wavelengths below 500nm. The authors acknowledge overlapping spectral characteristics between Ca and CDOM but don't adequately address how detrital absorption might influence their conclusions about CDOM primacy. Given that detritus can substantially affect spectral signatures in the same regions they analyze, this omission represents a gap in their analysis framework. A more comprehensive treatment of all optically active constituents would strengthen their argument regarding the relative importance of CDOM in aquatic environments.

Authors Response 4: The authors have not chosen to omit detritus or any other OAC from their study, but rather the analyses reflect the reality that detrital measurements are more sparse for globally applicable bio-optical datasets. Nonetheless, elements of the results describe detrital absorption, as follows: the leading eigenfunction for each dataset—the eigenfunction with a spectral shape most similar to that of detrital (but also CDOM) absorption—captures approximately 60% of the variance in each dataset. The datasets represent dissimilar water body types—ranging from purely oceanic waters (RSE2007) to oceanic, coastal, and inland waters in nearly equal representation (RSE2021). Contributions from detritus would therefore be anticipated to vary greatly between datasets. Differences in the variance captured between the datasets, however, are negligible (within $\pm 1\%$), thereby not supporting strong differences in detrital effects. This means that although detritus certainly contributes to blue absorption in the dataset, the spectral modifications captured by the leading eigenfunction do not support primacy of detrital absorption. This point is also supported by the leading eigenfunction indicating strong linearly correlation to $a_{\text{CDOM}}(440)$. Detrital absorption and CDOM absorption exhibit variable correlation for different natural waters (Babin et al. 2003; Twardowski et al. 2004), which would degrade the linear correlation if detritus was playing a major role in the analysis.

Therefore, the sparsity of detrital absorption data is unfortunate (and not due to omission by the authors), but the data and analyses nonetheless still support the findings presented. The authors improve discussion of detrital absorption by adding a new paragraph discussing detrital absorption following L467, as follows:

"Observations of detrital absorption were not available, and detritus and CDOM share similarities in the spectral-shape of absorption, i.e., absorption increases for both at shorter wavelengths. Some key factors relevant to the focus herein on $a_{\text{CDOM}}(440)$ over detritus in driving variability in aquatic spectra include that the leading eigenfunction captures similarly high proportions of the variance for each dataset (within 1% to the mean), whereas each dataset corresponds to dissimilar contributions from open-ocean waters. Absorption by CDOM and detritus also exhibit variable correlation patterns (Babin et al. 2003; Twardowski et al. 2004), which would degrade the correlation between the leading eigenfunction and $a_{\text{CDOM}}(440)$, if the leading eigenfunction was strongly associated with detritus."

In addition, the authors add discussion of the benefits conferred by expanded field sampling, namely for detrital absorption and non-algal particles, at L515 (further discussed in Authors Response 25), as follows:

> "The findings demonstrate high variance captured by the leading eigenfunction and the eigenfunction's strong association with $a_{\mathrm{CDOM}}(440)$ variability, but the findings would be strengthened by adding field sampling for additional biochemical or bio-optical parameters. For example, RSE2021 demonstrates the importance of combining spectrally expansive radiometry and observations of C_a with observations of $a_{\mathrm{CDOM}}(440)$. Also adding observations of non-algal particles (NAPs) and detrital absorption would help to further improve clarity on the eigenfunction interpretations, and, perhaps, contribute to a better understanding of the subsequent modes."

Throughout the manuscript, the authors consistently refer to their analysis as an "eigendecomposition" rather than using more widely recognized terms like Principal Component Analysis (PCA) or Empirical Orthogonal Functions (EOF). While technically correct, this terminological choice may unnecessarily distance their work from the broader scientific literature, potentially impeding recognition of their method's relation to established techniques.

Authors Response 5: The authors improve clarity based on the reviewer's comment in two ways. First, when the eigenanalysis is introduced on L13, the authors add the terms "Empirical Orthogonal Function (EOF)" and "Principal Component Analysis (PCA)," as follows:

> "...is tested herein using eigenanalysis—e.g., an Empirical Orthogonal Function (EOF) analysis, Principal Component Analysis (PCA), or other eigendecomposition depending on the literature—on three independent bio-optical datasets ..."

Second, the authors replace the term "eigendecomposition" with the term "eigenanalysis" when it appears in the methodological description on L267–269 and the results text on L301.

Additionally, the manuscript would benefit significantly from showing the mean spectra alongside their eigenfunctions, as well as demonstrating reconstructed spectra with varying CDOM and Ca concentrations to illustrate how these constituents individually contribute to spectral variability.

Authors Response 6: Fig. A1 shows mean spectra plus reconstructed spectra based on varying influences of the eigenfunctions (the eigenfunctions are already shown in Fig. 1). The magnitude to which the reconstructed spectra are modified by the eigenfunctions (i.e., represented by the scalar in the legend) is selected based on

the approximate inverse of the y-axis range in Fig. 1. That range is a mathematical consequence of the spectral resolution of each dataset and is not prescribed by the authors.
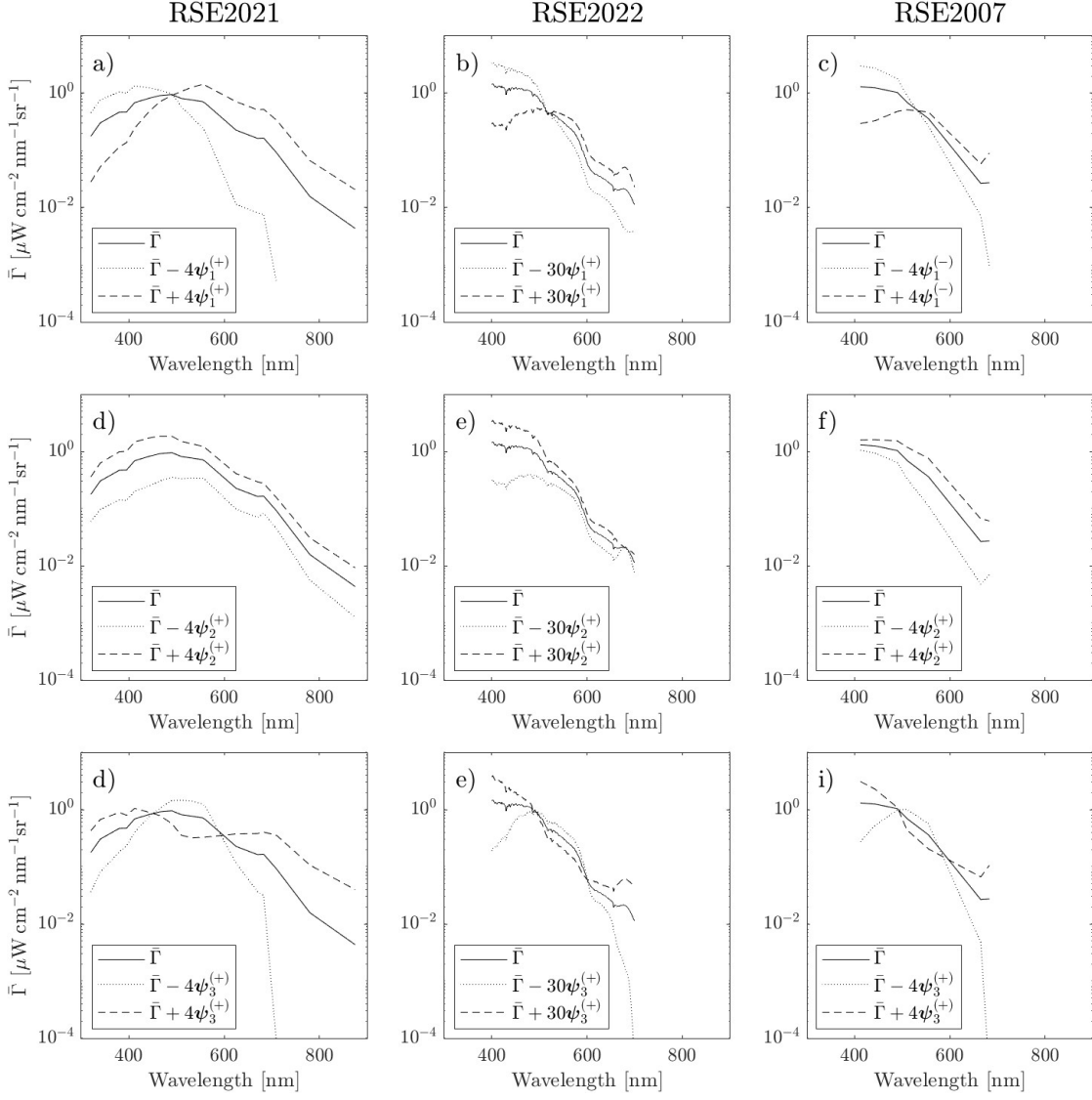
Fig. A1. Mean $\left[L_W(\lambda)\right]_{\mathrm{N}}$ spectra—represented as $\bar{\Gamma}$ for brevity following Hooker et al. 2021a—are shown for each dataset as solid black lines. Reconstructed spectra are shown in dashed plus dotted lines and represent spectral modifications associated with adding or subtracting, respectively, scalar quantities derived from the eigenfunctions. Modifications associated with $\boldsymbol{\psi}_1$, $\boldsymbol{\psi}_2$, and $\boldsymbol{\psi}_3$ are shown in the top, middle, and bottom rows, respectively.

The authors should acknowledge that their eigendecomposition approach assumes linear relationships among spectral variables, which may not fully capture nonlinear aspects of aquatic optical properties. Ocean color properties often exhibit complex nonlinear relationships, particularly in optically complex waters. This impacts some of the conclusions drawn. In particular, on Line 503-505, the authors write: "...the leading eigenfunction would be anticipated to capture nearly all of the variance of the dataset." This only holds if the relationships are linear.

Authors Response 7: The authors agree with the reviewer that eigenanalyses capture variability expressed in a linear framework. The authors also agree with the reviewer that ocean color properties can exhibit nonlinear relationships, while simultaneously acknowledging that quasi-linear approximations have been applied for decades. For example, Gordon et al. (1988) established a fundamental framework for relating IOPs to the remote-sensing reflectance, R_{rs}, wherein R_{rs} is proportional to the ratio of backscattering to absorption. In waters where absorption greatly exceeds backscattering (which include most global waters if considering VIS optical properties and nearly all global waters if considering INV optical properties), this fundamental framework would invoke approximately log-linear relationships, e.g., $\ln(R_{rs}) \propto -\ln(a)$. The authors also agree that nonlinearities, e.g., those commonly associated within case-1 algorithms wherein optical properties are empirically tuned to C_a (Morel et al. 2009), are applicable to the discussion of the idealized case-1 dataset at L503–505.

The analyses conducted herein do not negate nonlinear processes, but the datasets also do not suggest the eigenanalysis results are strongly swayed by nonlinearity. For example, the leading eigenfunction—which captures most of the variance for each dataset—produces $\boldsymbol{S}_1$ values with a distribution similar to that of $a_{\mathrm{CDOM}}(440)$. The $\boldsymbol{S}_1$ matrix predictors produce linear correlation to log-transformed $a_{\mathrm{CDOM}}(440)$ with Pearson's correlation coefficient indicated as ρ, plus Spearman's rank correlation coefficient included as $\tilde{\rho}$. High variance captured by the leading eigenfunction combined with loglinear relationships produced between $\boldsymbol{S}_1$ and a physically interpretable environmental parameter, support that the eigenanalysis is uncovering a dominant parameter influence, as intended. The relationships between the $\boldsymbol{S}_1$ plus $\boldsymbol{S}_2$ matrix predictors and $a_{\mathrm{CDOM}}(440)$ plus C_a, respectively, are shown in Fig. A2, as follows:
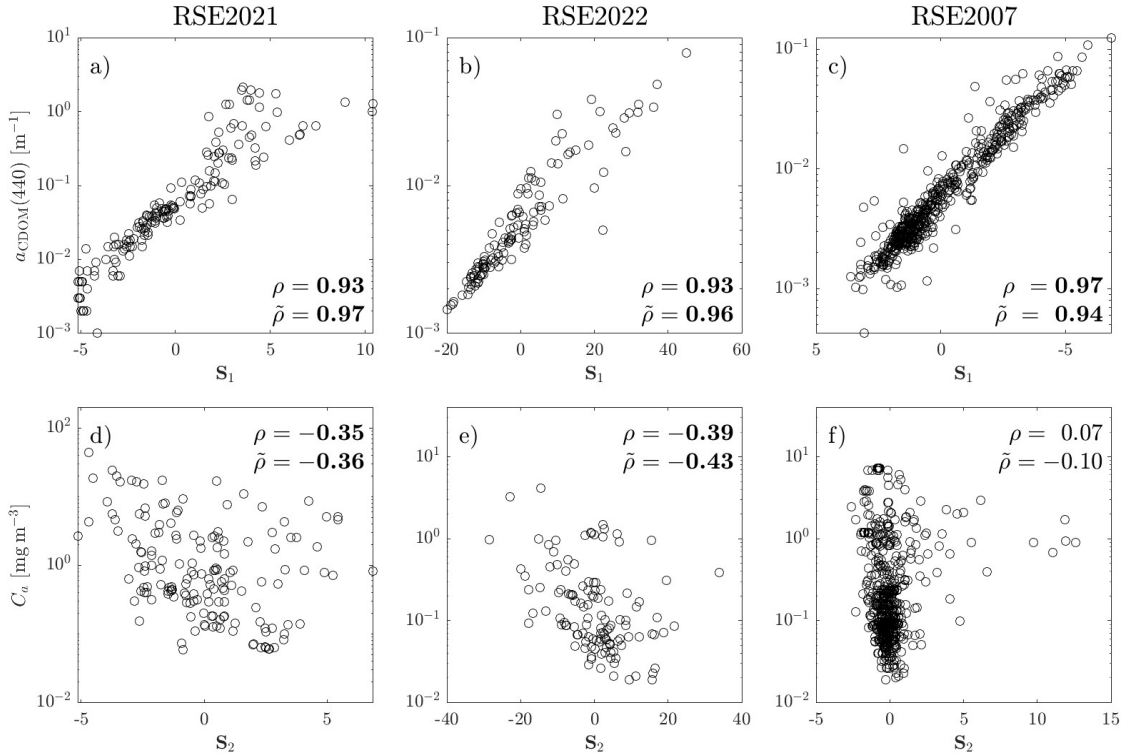
Fig. A2. Relationships between the S_1 (top row) and S_2 (bottom row) matrix predictors to CDOM absorption (top row) and C_a values (bottom row). Pearson's (ρ) and Spearman's ($\tilde{\rho}$) correlation coefficients are overlaid, with bold type indicating significance at $P < 0.01$.

The authors add text in the methods section at L269 to clarify that an eigenanalysis characterizes linear patterns of variability, as follows:

> "The eigenanalysis captures linear patterns of variability in the datasets, and is performed, as follows:"

Next, the authors revise L503–505 to incorporate the reviewers point, as follows:

> "For example, considering an idealized case-1 dataset wherein all OACs are accurately derived from C_a, the leading eigenfunction would be anticipated to produce rank correlation coefficients for each OAC. In contrast..."

While their analysis provides valuable insights into the primary modes of variability, explicitly discussing these limitations would provide readers with important context for interpreting their findings. This is especially relevant given their conclusion that CDOM, rather than Ca, drives the primary mode of variability in aquatic light fields—a finding that challenges conventional assumptions in ocean color remote sensing.

Authors Response 8: The authors improve the discussions of limitations, as well as clarify elements regarding linearity and algorithm implementation. In Authors Response 7, the authors describe revisions to the text at L269 and L503 to add clarity regarding linearity assumptions of eigenanalyses. The authors also clarify that the eigenfunctions are capturing most of the variance in the dataset, that Spearman's $\tilde{\rho}$ and Pearson's ρ produce similar metrics, and the authors agreed to add the new Fig. A2, which demonstrates linearity between log-transformed $a_{\mathrm{CDOM}}(440)$ and S_1. These elements, particularly the relationships shown in Fig. A2, do not support strong nonlinear effects. In addition to the authors' described revisions—including those described at L269 and L503—the authors also add additional discussion of potential limitations within the Conclusion section, described in Authors Response 23.

A last, potentially challenging request: I am going to insist that the authors make the RSE2021 and RSE2022 datasets public, not merely by request. These are too valuable to leave to the chance that the 2 authors become unavailable, etc.

Authors Response 9: The authors acknowledge the reviewer's comment regarding data availability and appreciate the recognition that the RSE2021 and RSE2022 datasets are valuable. Most importantly, the authors fully comply with the data policy of *Biogeosciences*. This policy specifically acknowledges that scenarios arise

when "data cannot be deposited publicly." In these circumstances, the policy requires a Data Availability statement to be produced to clearly explains this limitation. As it pertains to the reviewer's comment, the authors are constrained in the full release of the RSE2021 dataset because of a prior data acquisition agreement, and the RSE2022 dataset is already released (and already adequately referenced in the manuscript).

The authors will update the Data Availability statement. To comply with all journal requirements, the authors will also provide a supplementary data file containing the specific data used to generate all figures in the manuscript. The authors believe that this approach is in keeping with both the letter and spirit of the *Biogeosciences* data policy: the authors agree that open data is the ideal, while also acknowledging that there are legitimate situations where complete data sharing is not immediately possible.

Here are some additional, more minor comments for the authors to consider. In order of appearance, not importance:

1. *Include numbers in the Abstract, i.e. be quantitative.*

   Authors Response 10: The authors revise L17–19 of the Abstract, as follows:

   "...Blue and green band-ratio algorithms routinely used for remote sensing of C_a are found to be maximally sensitive to CDOM—rather than C_a—variability based on validation tests of OC algorithm performance (e.g., R^2 of 0.85 versus 0.78), plus partial correlation coefficients relating eigenfunction scalar amplitude functions to field or derived observations."

2. *I encourage the authors to reference Cael+2020 near line 95.*

   Authors Response 11: The authors add a citation for Cael et al. 2020 at L130, as follows:

   "...of the variance—similar information constraint was likewise previously shown for hyperspectral observations of particulate absorption (Cael et al. 2020)."

3. *Line 137: maybe specify that the "optical signatures" are spatial not spectral.*

   Authors Response 12: The authors revise L137, as follows:

   "...airborne investigations leveraged spatially cohesive optical signatures associated with..."

4. *Lines 139-142: The leading components are not always (maybe not even typically) dominated by broadband features. Those are taken out by the mean. One example I know where the first modes are *very* informative are galaxy spectra. Please reword these lines accordingly.*

> Authors Response 13: The authors agree with the reviewer's comment about the first modes being informative—the implication of the first mode is fundamental to the results herein. This statement instead describes a potential pitfall in quantifying information content based on eigenanalyses: Processes or constituents that modify spectra across an expansive wavelength range mathematically receive more weight in an eigenanalysis than those that modify only a narrow spectral region. Therefore, the results of an eigenanalysis may sometimes not adequately reflect information associated with processes that modify a narrow spectral region. The authors clarify this by revising L139–141, as follows:
>
> > "Quantifying information content using eigenanalyses can likewise be challenging: dimensionality bias can produce an incomplete perspective of information content by increasing the variance captured by broadband features relative to that of narrow features. The latter—when not spectrally diluted—often provide informative and exploitable information (Houskeeper et al. 2020b)."

5. *Somewhere in the last paragraph of the Introduction, I encourage the authors to cite the recent papers by Z. Erickson (2020 and 2023).*

> Authors Response 14: The authors revise L141–143, as follows:
>
> > "Mixture density networks (MDNs) and inverse models that incorporate prior knowledge (i.e., leveraging a Bayesian framework) improve the management of degeneracy in radiative transfer, and are forthcoming for PACE science objectives (O'Shea et al. 2021; Erickson et al. 2020,2023)."

6. *Somewhere, consider citing Siegel+2013 as a previous reference showing/asserting that CDOM dominates absorption at bluer wavelengths.*

> Authors Response 15: The authors previously cited Jerlov (1968) on L320–321. Following the reviewer's preference, the authors add Siegel et al. (2013):
>
> > "Conversely, spectral darkening of shorter wavelengths with minimal spectral features is most consistent with CDOM absorption (Jerlov 1968; Siegel et al. 2013)."

7. *Line 313: Please cite a reference supporting the assertion that "scattering processes confer less spectral dependencies."*

Authors Response 16: The authors add a citation for Kirk (2011), which provides a useful review on this topic, and also add text to reference absorption band effects for completeness, as follows:

> "The spectral shapes of the eigenfunctions are primarily considered as a function of absorption processes, which generally—notwithstanding absorption band effects (Zaneveld and Kitchen 1995)—confer stronger spectral dependencies than scattering processes (Kirk 2011)."

8. *Line 373: I have to admit I am not familiar with the term "R2 statistics."*

   Authors Response 17: The authors replace "R^2 statistics" with "R^2" following the reviewer's preference.

9. *Line 436: I struggle to parse this paragraph. Found myself mainly confused reading it. Maybe provide additional context?*

   Authors Response 18: The discussion is related, in part, to the reviewer's earlier questions regarding nonlinearities in aquatic optics. The authors revise the paragraph to improve clarity and reduce length plus complexity, as follows:

   > Bio-optical formulations to derive optical properties as a function of C_a do not require primacy of C_a. Rather, algorithms parameterize evolution in optical properties—notwithstanding observational challenges—based on empirical OAC relationships (Morel and Prieur 1977; Morel 2009). Nonetheless, variability in OAC relationships exists for *in situ* water bodies (Hansell and Orellana 2021). Empirical approaches such as the OC family of algorithms must mitigate observational artifacts (Uitz et al. 2006; Kudela et al. 2019, 2024), plus regional variability and nonlinearity in OAC relationships (Morel 2009). The latter manifests with the need to regionally tune VIS algorithms for specific waters. For example, tunings for arctic waters account for higher $a_{\mathrm{CDOM}}(440)$ relative to C_a content (Matsuoka et al. 2013; Lewis and Arrigo 2020), and tunings for antarctic waters account for lower $a_{\mathrm{CDOM}}(440)$ relative to C_a (Dierssen and Smith 2000). The latter is also in keeping with highly nonlinear formulations of OC algorithms, e.g., the number of power terms used in the polynomial model to fit C_a to ratios of R_{rs} is routinely four (Morel 2009; O'Reilly and Werdell 2019). Linearity and loglinearity of CDOM algorithms increases with increasing spectral separation of wavelengths (Hooker et al. 2020, 2021a; Houskeeper et al. 2021). The results herein suggest that differences in linearity correspond, in part, to differences in the separability of signals associated with spectral range.

10. *Line 447: "are stable" –> "are [un]stable" ?*

Authors Response 19: The original usage, i.e., stable, was correct, but the authors revise this paragraph to improve clarity and reduce length, as described in Authors Response 18.

11. *Line 460: I was not persuaded that item c) was actual, independent evidence*

Authors Response 20: The authors revise this text to help clarify their meaning, i.e., that their findings are supported by both the variance captured by the leading and secondary eigenfunctions, as well as the correlations observed between the S matrixes and the biochemical parameters. The latter is also further clarified by Fig. A2. The authors reorganize L455–465, as follows:

"The similarities favor primacy of $a_{\mathrm{CDOM}}(440)$—rather than C_a—as the predominant driver of optical variability for waters, including those of the open ocean, summarized briefly, as follows: a) the leading eigenvector from each dataset—capturing approximately 60% of the variance—indicates opposing anomalies for longer versus shorter wavelengths with minimal amplitude in the blue-green transition domain. The secondary eigenvector from each dataset—capturing approximately 32% of the variance—indicates internal VIS spectral dependencies characteristic of C_a absorption, including a maximum near the blue-green transition; b) the S_1 term representing the stretching and compressing necessary to best represent the data using the leading eigenfunction is always more strongly associated with variability in $a_{\mathrm{CDOM}}(440)$ than C_a values. The S_2 term more strongly associates with field observations of C_a than $a_{\mathrm{CDOM}}(440)$ for two of the three datasets..."

12. *Lines 489: I encourage authors to include equations describing the processes written about here.*

Authors Response 21: The reviewer has requested that the authors include an equation regarding L489, which describes the negative buoyancy of particles—which is parameterized following Stokes law (1851). While the authors are generally supportive of including equations, the authors perceive sinking rate parameterizations as only indirectly relevant to the manuscripts findings. The authors will therefore add clarity regarding the reviewer's comment by adding text, plus a citation regarding the following line which the reviewer also may have been intending. The authors perceive that this revision helps clarify the material without adding unnecessary complexity. The authors revise L489–490, as follows:

"Spatial differences occur, perhaps, due to negative buoyancy of particles—e.g., sinking velocity increases proportionally with density differences and the squared particle radius (Bach et al., 2012)—driving particulate organic distributions away from the surface where remote sensing signals are weighted (e.g., Morel and Berthon 1989)."

13. *Line 560: Consider citing Prochaska & Frouin (2025) here.*

> Authors Response 22: The authors add Prochaska and Frouin (2025) at L560. Based on additional revisions described in Authors Response 30, the full revision at L569 is as follows:

>> "Accurate retrieval of $a_{\mathrm{CDOM}}(440)$ is, therefore, prerequisite to—not separate from—accurate retrieval of phytoplankton properties, in keeping with the trajectory of research into simultaneous inherent optical property (IOP) inversions (e.g., Maritorena et al. 2022; Lee et al. 2002; Werdell et al. 2013), machine learning approaches (O'Shea et al. 2021), and degeneracy (Prochaska and Frouin 2025)."

14. *The conclusions are largely summary + self-promoting. If that is the standard for this Journal, no problem.*

> Authors Response 23: While the *Biogeosciences* submission guidelines do not provide any specialized requirements relating to the composition of the Conclusion section (https://www.biogeosciences.net/submission.html), the authors appreciate this feedback on the Conclusion text. The authors revise the Conclusions section, as follows:

>> "~~Spectral observations of aquatic environments produce leading eigenfunctions of AOPs more strongly correlated to variability in CDOM absorption than C_a.~~ Greater independence of OACs and the elevated importance of CDOM variability in governing aquatic light variability are consistent with advancing knowledge of microbial loop dynamics and an increasing diversity of trophic pathways represented therein (Azam 1998). The consistency of results using three independent datasets strengthens support for ~~and are based on consistent results for the eigenanalyses—including the spectral shapes of the eigenfunctions and the partial correlation coefficients relating the eigenfunctions to biogeochemical variables—plus performance metrics of OC algorithms. The eigenanalyses indicate~~ accurate and independent estimation of CDOM as prerequisite to retrieval of C_a and other phytoplankton parameters. However, not every bio-optical dataset, quality control protocol is assessed, not all applicable OACs are present in the datasets, and the eigen analyses capture linear—but not nonlinear—patterns of variability. ~~and satellite OC algorithms are found herein to produce values more highly correlated with CDOM than C_a.~~ Confounding signals identified for CDOM and C_a are consistent with early investigations into the drivers of color variability (Yentsch 1960), as well as subsequent work assessing vulnerabilities in band-ratio algorithms for characterizing C_a (Dierssen 2010; Sauer et al. 2012). Spectrally expansive data products have been shown to improve retrieval of $a_{\mathrm{CDOM}}(440)$ independent of C_a (Sathyendranath

et al. 1987; Hooker et al. 2020; Houskeeper et al. 2021), and comparisons herein of eigenanalyses using ~~the INV21 and VIS21~~ complete spectra versus spectral subsets further support potential opportunities associated with spectral range expansions. Spectrally expansive data products have been demonstrated *in situ* spanning the UV to short-wave infrared (SWIR) wavelength domain, with the latter formerly ascribed as null ~~in the ocean color community of practice~~ (Houskeeper and Hooker 2023). The recently launched OCI sensor of the PACE mission provides hardware capabilities to support expansive spectral range observing of the global ocean surface. The findings herein support opportunities associated with leveraging the spectrally expansive capabilities of OCI, although further advances in image processing, atmospheric correction, algorithm design, and ~~applicability of~~ *in situ* datasets would be necessary."

The authors thank Reviewer 1 for their comments and suggestions, which helped the authors to improve the manuscript.

References

Azam, F. "Microbial control of oceanic carbon flux: the plot thickens." *Science* 280, no. 5364 (1998): 694-696.

Babin, Marcel, Dariusz Stramski, Giovanni M. Ferrari, Herve Claustre, Annick Bricaud, Grigor Obolensky, and Nicolas Hoepffner. "Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe." *Journal of Geophysical Research: Oceans* 108, no. C7 (2003).

Bach, Lennart Thomas, Ulf Riebesell, Scarlett Sett, Sarah Febiri, Paul Rzepka, and Kai Georg Schulz. "An approach for particle sinking velocity measurements in the 3–400 μm size range and considerations on the effect of temperature on sinking rates." Marine biology 159, no. 8 (2012): 1853-1864.

Cael, B. B., Alison Chase, and Emmanuel Boss. "Information content of absorption spectra and implications for ocean color inversion." Applied Optics 59, no. 13 (2020): 3971-3984.

Dierssen, H.M. and R.C. Smith. "Bio-optical properties and remote sensing ocean color algorithms for Antarctic Peninsula waters." *Journal of Geophysical Research: Oceans* 105, no. C11 (2000): 26301-26312.

Dierssen, H.M. "Perspectives on empirical approaches for ocean color remote sensing of chlorophyll in a changing climate." *Proceedings of the National Academy of Sciences* 107, no. 40 (2010): 17073-17078.

Erickson, Zachary K., P. Jeremy Werdell, and Ivona Cetinić. "Bayesian retrieval of optically relevant properties from hyperspectral water-leaving reflectances." *Applied Optics* 59, no. 23 (2020): 6902-6917.

Erickson, Zachary K., Lachlan McKinna, P. Jeremy Werdell, and Ivona Cetinić. "Bayesian approach to a generalized inherent optical property model." *Optics Express* 31, no. 14 (2023): 22790-22801.

Gordon, H.R., Brown, O.B., Evans, R.H., Brown, J.W., Smith, R.C., Baker, K.S., Clark, D.K. A semianalytic radiance model of ocean color. *J. Geophys. Res.* 1988;93:10909–10924.

Hooker, S.B., J.H. Morrow, and A. Matsuoka. "Apparent optical properties of the Canadian Beaufort Sea–Part 2: The 1% and 1 cm perspective in deriving and validating AOP data products." *Biogeosciences* 10, no. 7 (2013): 4511-4527.

Hooker, S.B., R.N. Lind, J.H. Morrow, J.W. Brown, K. Suzuki, H.F. Houskeeper, T. Hirawake, and E.R. Maure, 2018a: Advances in Above- and In-Water Radiometry, Vol. 1: Enhanced Legacy and State-of-the-Art Instrument Suites. TP–2018–219033/Vol. 1, NASA Goddard Space Flight Center, Greenbelt, Maryland, 60 pp.

Hooker, S.B., A. Matsuoka, R.M. Kudela, Y. Yamashita, K. Suzuki, and H.F. Houskeeper. "A global end-member approach to derive $a_{\mathrm{CDOM}}(440)$ from near-surface optical measurements." *Biogeosciences* 17, no. 2 (2020): 475-497.

Hooker, S.B., H.F. Houskeeper, R.M. Kudela, A. Matsuoka, K. Suzuki, and T. Isada. "Spectral modes of radiometric measurements in optically complex waters." *Continental Shelf Research* 219 (2021a): 104357.

Hooker, S.B., H.F. Houskeeper, R.N. Lind, and K. Suzuki. "One-and two-band sensors and algorithms to derive $a_{\text{CDOM}}(440)$ from global above- and in-water optical observations." *Sensors* 21, no. 16 (2021b): 5384.

Hooker, S.B., H.F. Houskeeper, R.N. Lind, R.M. Kudela, and K. Suzuki. "Verification and validation of hybridspectral radiometry obtained from an unmanned surface vessel (USV) in the open and coastal oceans." *Remote Sensing* 14, no. 5 (2022): 1084.

Houskeeper, Henry F. "Advances in bio-optics for observing aquatic ecosystems." University of California, Santa Cruz, 2020a.

Houskeeper, H.F., D. Draper, R.M. Kudela, and E. Boss. "Chlorophyll absorption and phytoplankton size information inferred from hyperspectral particulate beam attenuation." *Applied Optics* 59, no. 22 (2020b): 6765-6773.

Houskeeper, H.F., S.B. Hooker, and R.M. Kudela. "Spectral range within global $a_{\text{CDOM}}(440)$ algorithms for oceanic, coastal, and inland waters with application to airborne measurements." *Remote Sensing of Environment* 253 (2021): 112155.

Houskeeper, Henry F., Stanford B. Hooker, and Kyle C. Cavanaugh. "Spectrally simplified approach for leveraging legacy geostationary oceanic observations." *Applied Optics* 61, no. 27 (2022): 7966-7977.

Houskeeper, H.F., and S.B. Hooker. "Extending aquatic spectral information with the first radiometric IR-B field observations." *PNAS Nexus* 2, no. 11 (2023): pgad340.

Houskeeper, H.F., S.B. Hooker, and R.N. Lind. "Expanded linear responsivity for Earth and planetary radiometry." *J. Atm. Ocea. Tech.* 41, no. 11 (2024): 1,093–1,105.

Jerlov, N.G. "Optical Oceanography," *Elsevier Oceanography*, Series 5 (1968).

Kirk, John TO. Light and photosynthesis in aquatic ecosystems. Cambridge university press, 2011.

Kudela, R.M., S.B. Hooker, H.F. Houskeeper, and M. McPherson. "The influence of signal to noise ratio of legacy airborne and satellite sensors for simulating next-generation coastal and inland water products." *Remote Sensing* 11, no. 18 (2019): 2071.

Kudela, R.M., S.B. Hooker, L.S. Guild, H.F. Houskeeper, and N. Taylor. "Expanded signal to noise ratio estimate for validating next-generation satellite sensors in oceanic, coastal, and inland waters." *Remote Sensing* 16 (2024): 1238.

Lee, ZhongPing, Kendall L. Carder, and Robert A. Arnone. "Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters." *Applied optics* 41, no. 27 (2002): 5755-5772.

Lewis, K.M., and K.R. Arrigo. "Ocean color algorithms for estimating chlorophyll a, CDOM absorption, and particle backscattering in the Arctic Ocean." *Journal of Geophysical Research: Oceans* 125, no. 6 (2020): e2019JC015706.

Maritorena, Stephane, David A. Siegel, and Alan R. Peterson. "Optimization of a semianalytical ocean color model for global-scale applications." *Applied optics* 41, no. 15 (2002): 2705-2714.

Matsuoka, A., M. Babin, D. Doxaran, S.B. Hooker, B.G. Mitchell, S. Bélanger, and A. Bricaud. "A synthesis of light absorption properties of the Pan-Arctic Ocean: application to semi-analytical estimates of dissolved organic carbon concentrations from space." *Biogeosciences Discussions* 10, no. 11 (2013).

Morel, A., and L. Prieur. "Analysis of variations in ocean color 1." *Limnology and Oceanography* 22, no. 4 (1977): 709-722.

Morel, André, and Jean-François Berthon. "Surface pigments, algal biomass profiles, and potential production of the euphotic layer: Relationships reinvestigated in view of remote-sensing applications." *Limnology and oceanography* 34, no. 8 (1989): 1545-1562.

Morel, A. "Are the empirical relationships describing the bio-optical properties of case 1 waters consistent and internally compatible?" *Journal of Geophysical Research* 114, no. C01016 (2009).

O'Reilly, J.E., and P.J. Werdell. "Chlorophyll algorithms for ocean color sensors-OC4, OC5 & OC6." *Remote Sensing of Environment* 229 (2019): 32-47.

O'Shea, R.E., N. Pahlevan, B. Smith, M. Bresciani, T. Egerton, C. Giardino, L. Li, et al. "Advancing cyanobacteria biomass estimation from hyperspectral observations: Demonstrations with HICO and PRISMA imagery." *Remote Sensing of Environment* 266 (2021): 112693.

Prochaska, J. Xavier, and Robert J. Frouin. "On the Challenges of Retrieving Phytoplankton Properties from Remote-Sensing Observations." *EGUsphere* 2025 (2025): 1-39.

Sathyendranath, S., L. Prieur, and A. Morel. "An evaluation of the problems of chlorophyll retrieval from ocean colour, for case 2 waters." *Advances in Space Research* 7, no. 2 (1987): 27-30.

Sauer, M.J., C.S. Roesler, P.J. Werdell, and A. Barnard. "Under the hood of satellite empirical chlorophyll a algorithms: revealing the dependencies of maximum band ratio algorithms on inherent optical properties." *Optics Express* 20, no. 19 (2012): 20920-20933.

Siegel, David A., Michael J. Behrenfeld, Stéphane Maritorena, Charles R. McClain, David Antoine, Sean W. Bailey, Paula S. Bontempi et al. "Regional to global assessments of phytoplankton dynamics from the SeaWiFS mission." *Remote Sensing of Environment* 135 (2013): 77-91.

Twardowski, Michael S., Emmanuel Boss, James M. Sullivan, and Percy L. Donaghay. "Modeling the spectral shape of absorption by chromophoric dissolved organic matter." *Marine Chemistry* 89, no. 1-4 (2004): 69-88.

Werdell, P. Jeremy, Bryan A. Franz, Sean W. Bailey, Gene C. Feldman, Emmanuel Boss, Vittorio E. Brando, Mark Dowell et al. "Generalized ocean color inversion model for retrieving marine inherent optical properties." *Applied optics* 52, no. 10 (2013): 2019-2037.

Yentsch, C.S. "The influence of phytoplankton pigments on the colour of sea water." *Deep Sea Research* 7, no. 1 (1960): 1-9.

Zaneveld, J.R.V. and J.C. Kitchen, "The variation in the inherent optical properties of phytoplankton near an absorption peak as determined by various models of cell structure," *J. Geophys. Res. Oceans* 100, 13309–13320 (1995).