Authors Response to Reviewer 2 Comments for Manuscript egusphere-2024-4163: *The primacy of dissolved organic matter to aquatic light variability*

9 May 2025

The authors thanks Reviewer 2 for their valuable comments. The authors respond to all comments below and document resulting modifications to the manuscript, as appropriate. Reviewer 2's comments are shown in *slanted* typeface. The Authors Responses are shown in plain typeface. The numbering of the Authors Responses continues from the numbering of the Authors Responses to Reviewer 1 to minimize ambiguity when referencing across response documents. 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 2:

The manuscript proposes a change in paradigm in ocean color remote sensing, providing evidence that CDOM absorption—rather than Ca—primarily drives water-leaving light variability. They additionally make a strong case for exploiting information in the INV range of the spectra, clearly highlighting current limitations in both in situ and satellite datasets. The work has potentially great scientific implications, as it could change the way that aquatic light variability is thought, and ocean color algorithms are designed. However, I believe that the authors should address some comments detailed below to make their point stronger.

Authors Response 24: The authors thank the reviewer for their positive and constructive comments.

The authors acknowledge the important role of different OACs in the variability of aquatic light, with the purpose of ultimately demonstrating the primacy of CDOM over Ca. OACs in water are typically grouped in three different categories: CDOM, phytoplankton and non-algal particles (NAP) that can covary or not. However, the work does not assess the potential contribution of non-algal particles (NAP) to light variability (apart from briefly mentioning inorganic particles in lines 410-412). How relevant are NAPs in the study's datasets? Can their contribution be reasonably neglected? If so, this needs to be clearly justified. How does this omission in the datasets limit the findings of the work? How about the extrapolation of the findings to waters where NAPs contribution may be more relevant.

Authors Response 25: The authors agree with the reviewer regarding the importance of NAPs to aquatic light variability. Correlation coefficients relating the principal component time series to NAP variability were not presented because coincident observations of NAPs were not available. The challenge stems, in part, from the paradigm noted by the reviewer in their initial comment: Chlorophyll *a* (C_a) is generally considered the primary variable in ocean color. As a result, C_a is often the primary—and sometimes only—biogeochemical field measurement obtained to accompany observations of apparent optical properties (AOPs). The RSE2021 dataset was able to add routine measurements of $a_{\mathrm{CDOM}}(440)$, which, in part, made this study into the primacy of $a_{\mathrm{CDOM}}(440)$ possible.

However, despite the challenging deficiency in observations of NAPs for many applicable datasets, the RSE2021 dataset presented herein does capture substantial variability in NAP based on water bodies sampled. For example, many riverine systems frequently contain high concentrations of suspended sediments, and examples of sediment-rich rivers—namely the Colorado, Napa, and Columbia rivers—are represented in the RSE2021 eigenanalysis. Similarly, an expansive range in NAP is represented by lacustrine waters sampled spanning Crater Lake (OR) to Pinto Lake (CA), plus the sampling or estuaries such as Elkhorn Slough (CA) and San Francisco Bay (CA). This is not true for the RSE2022 or RSE2007 datasets, which did not include inland waters, and so the similarity in variance captured by the leading eigenfunction for all datasets (within $\pm 1\%$) suggests that NAPs are not strongly modifying the leading eigenfunction of RSE2021. This scenario is similar to the discussion of detrital absorption in Authors Response 4.

For a previous publication, the authors derived estimates of the coefficient for particulate backscattering at 443 nm, $b_{\mathrm{bp}}(443)$, based on an inversion scheme applicable to case-1 and case-2 waters (Matsuoka et al. 2013) and with updates for coastal waters (Hooker et al. 2021a). The estimation indicates an expansive range in $b_{\mathrm{bp}}(443)$ spanning 0.001 to $0.329\,\mathrm{m}^{-1}$. Combined with the representation of riverine, lacustrine, and estuarine water bodies wherein inorganic particle concentrations were highly variable, the $b_{\mathrm{bp}}(443)$ range indicates that the RSE2021 eigenanalysis is anticipated to capture a non-negligible amount of *in situ* variability associated with NAPs. The authors show the dataset quartile characteristic for the synthetic $b_{\mathrm{bp}}(443)$ coefficients in Table A1 (also discussed in Authors Response 29) to clarify elements of NAP representation by RSE2021. The authors do not test the $\boldsymbol{S}$ matrix values against the synthetic $b_{\mathrm{bp}}(443)$, because this could contribute to misinterpretation based on presently uncharacterized dynamics of the $b_{\mathrm{bp}}(443)$ algorithm performance, which is outside the scope of the submitted manuscript.

Next, the authors note that the leading eigenfunction of the RSE2021 dataset explained 0.60 of the variance in the dataset. The leading eigenfunction produced a spectral shape similar to CDOM absorption and also indicated very high linear correlation to $a_{\mathrm{CDOM}}(440)$ with $r = 0.92$ (the somewhat lower value of 0.80 shown in Fig. 1 is a partial correlation value that includes accounting for effects of C_a, discussed in Authors Response 39), or an r^2 of 0.85. This suggests (approximately) that $a_{\mathrm{CDOM}}(440)$ captures on the order of 0.85 of the 0.60 captured by $\boldsymbol{\psi}_1$, which amounts to 0.51. Even if all other variance in the dataset was driven by a single other OAC (e.g., C_a, NAP, but also a_{phy} referencing Authors Response 26), the OAC could not achieve primacy because only 0.49 of the total variance remains. For the INV domain, the results are much stronger and suggest $a_{\mathrm{CDOM}}(440)$ captures greater than 0.70 of the variance. The results of these comparisons, therefore, do not suggest primacy of NAPs for the datasets tested herein.

The results of the datasets herein correspond to large-scale (e.g., global) dynamics. Targeted subsets of these datasets would *not* necessarily show similar importance of $a_{\mathrm{CDOM}}(440)$. For example, lacustrine datasets with a wide range in C_a—e.g., combining Crater Lake, (CA) and Clear Lake (CA)—might be anticipated to produce spectral variability most associated with C_a. Similarly, an onshore to offshore transect from the mouth of the Rio de la Plata—wherein very high suspended sediments injected into the coastal ocean by river discharge are mixed and diluted with oceanic water—might be anticipated to produce spectral variability most associated with changes in suspended sediment concentrations. The authors clarify that the results of the large-scale analyses do not apply to individual or regional scenarios by adding discussion after L515, as follows:

> "The findings presented herein correspond to datasets spanning globally representative or broad oceanic waters, and do not rank the importance of drivers in any specific region. For example, regional scenarios wherein high variability of non-algal particles (NAPs) is associated with riverine flux or resuspension would not be anticipated to reflect the findings herein supporting $a_{\mathrm{CDOM}}(440)$ primacy."

The authors also agree with the reviewer that not having field observations of b_{bp} or NAP (or detrital absorption, as described in Authors Response 7) is an important limitation to address. The authors clarify that the findings indicate high importance of CDOM absorption—despite the effects of C_a but also NAP and other OACs—but that adding field sampling of other OACs would help strengthen the analysis and provide more opportunities to explore additional relationships. The authors add discussion at L515, 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."

CDOM absorption coefficient (aCDOM) is an optical property itself, more directly related to Lw than Ca. Therefore, one could argue that it would be fairer to use phytoplankton absorption (aphy) rather than Ca in this study. They additionally have another important difference: while the spectral shape of CDOM absorption is smooth and can be roughly defined by two quantities (aCDOM at a reference wavelength, e.g., 440 nm, and an exponential decay exponent), aphy can present much more spectral variability in natural waters, not being as easily parameterized. Can these differences give an "advantage" to CDOM in the analyses presented in the manuscript, particularly when computing linear correlation coefficients with the eigenfunctions? Please include any pertinent discussion about this.

Authors Response 26: Similar to detrital absorption (Authors Response 4) and NAP (Authors Response 25), the authors are constrained by the sampling activities of globally representative radiometric datasets and their corresponding field observations—but the analyses nonetheless are informative regarding other OACs. For example, the likelihood of a scenario wherein variability in C_a to a_{phy} could alter the findings presented herein is challenged by the strength of the results concerning $a_{\text{CDOM}}(440)$—including the variance calculations discussed in Authors Response 25. Briefly, the results presented herein support very high correlation of the leading eigenfunction to log-transformed $a_{\text{CDOM}}(440)$, i.e., greater than 0.9 for the RSE2021, RSE2022, and RSE2007 datasets. The relationships to $a_{\text{CDOM}}(440)$ are also highly linear, shown in Fig. A2 (introduced in Authors Response 7), which also presents Pearson's correlation coefficient (ρ) and Spearman's rank correlation coefficient (herein $\tilde{\rho}$) side-by-side for comparison:

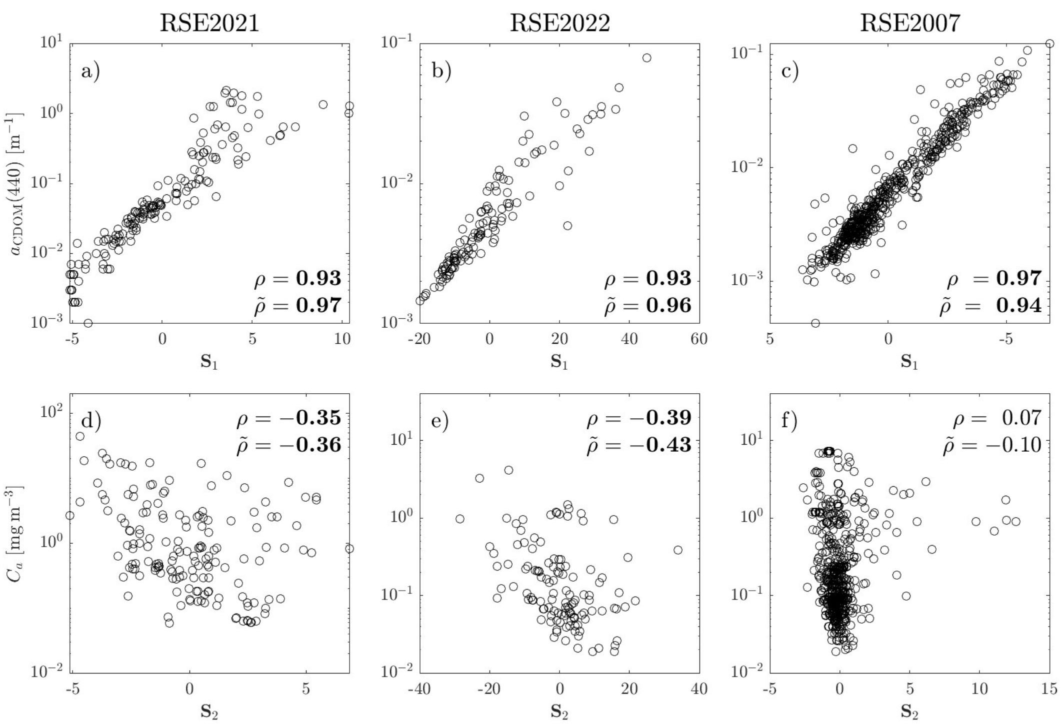

Fig. A2. Relationships between the $\boldsymbol{S}_1$ (top row) and $\boldsymbol{S}_2$ (bottom row) matrix predictors to CDOM absorption (top row) and C_a values (bottom row). Pearson's (ρ) and Spearman's

($\tilde{\rho}$) correlation coefficients are overlaid, with bold type indicating significance at $P < 0.01$.

The partial correlation coefficients presented in Fig. 1 for $a_{\mathrm{CDOM}}(440)$ and C_a as they relate to the leading eigenfunction are also quite disimilar. For example, RSE2021 indicates that correlation of $a_{\mathrm{CDOM}}(440)$ to the leading eigenfunction (and accounting for covariance with C_a) is 0.80. The opposite comparison—correlation of C_a to the leading eigenfunction and accounting for covariance with $a_{\mathrm{CDOM}}(440)$—is 0.11. The magnitude of these differences (0.80 versus 0.11) does not support the results being dependent on variability in C_a to a_{phy}, which are generally highly correlated. For example, as described in Roesler and Barnard (2013):

> "The line height absorption is shown to be significantly related to the extracted chlorophyll concentration over a large range of natural optical regimes and diverse phytoplankton cultures."

The spectral shape of a_{phy} is also routinely derived from C_a (Morel 2009), i.e., consistent with the focus herein on assessing the validity of C_a primacy. The authors clarify potential sources of variability in the relationship between C_a and phytoplankton absorption by adding discussion to L453, as follows:

> "Phytoplankton absorption—rather than C_a—was not tested due to the limitations of the datasets. Although phytoplankton absorption is generally considered to correspond to C_a, processes such as pigment packaging can add variability relevant to the correlations presented herein (Bricaud et al. 2004)."

Following the previous comments, why did the authors decide to use only Pearson correlation? I suggest they consider including (additionally) a non-linear correlation coefficient (e.g., Spearman's rank correlation). This could enrich the results and discussion.

Authors Response 27: The authors add Spearman's rank correlation in Fig. A2, which they also describe in Authors Response 7 of their responses to Referee 1. The authors find that the results from Spearman's rank correlation are extremely similar to those using Pearson's rank correlation. The authors show the Spearman's rank correlation values in Fig. A2, which also shows that the relationships between the $\boldsymbol{S}_1$ matrix and log-transformed $a_{\mathrm{CDOM}}(440)$ are highly linear. The authors thank the reviewer for the suggestion to also add Spearman's rank correlation.

I think the work could greatly benefit from exploiting the datasets a bit more, complementing the results of the principal components analysis. For example:

1. *For completeness and better contextualization, I suggest including a figure with the measured spectra when describing each dataset, so the reader can have a quick visualization of the spectral variability represented in the work.*

   Authors Response 28: Fig. A1, as described in Authors Response 6, presents the mean spectra from each dataset, plus the shapes of the changes associated with each eigenfunction. The authors believe that these examples are more informative given the high quantity and overlap of the spectra, combined with the eigenfunctions in Fig. A1 representing 97% of the variance in the datasets.

2. *It would also be beneficial to report the correlation coefficient between aCDOM and Ca in the considered datasets, as well as a table with the median, quartiles and ranges of aCDOM and Ca measurements. Could it be the case that aCDOM can "explain" more variance because it already had greater variability in the datasets, while Ca varied in a more limited range (maybe it is necessary to consider some relation between aphy and Ca to do this comparison)?*

   Authors Response 29: The reviewer makes multiple good points, and the authors address each element, as follows:

   - *Report the correlation coefficient between aCDOM and Ca in the considered datasets:*

     The authors add text stating the correlation coefficients for $a_{\mathrm{CDOM}}(440)$ and C_a to L287:

     "…captured by a remote sensing algorithm. Pearson's correlation coefficients comparing C_a and $a_{\mathrm{CDOM}}(440)$—measured or derived algorithmically—are 0.812, 0.912, and 0.909 for the RSE2021, RSE2022, and RSE2007 datasets, respectively."

   - *[Add] a table with the median, quartiles and ranges of aCDOM and Ca measurements:*

     Table A1 at the end of this Authors Response shows the requested information—plus synthetic $b_{\mathrm{bp}}(443)$ to also expand on Authors Response 25.

   - *Could it be the case that $a_{\mathrm{CDOM}}(440)$ can "explain" more variance because it already had greater variability in the datasets, while C_a varied in a more limited range?*

     The authors investigate this scenario and confirm it is not supported based on comparison of dynamic ranges expressed in the datasets. Briefly, $a_{\mathrm{CDOM}}(440)$ confers greater dynamic range than C_a in the RSE2007. C_a confers greater dynamic range than $a_{\mathrm{CDOM}}(440)$ in

RSE2022. $a_{\mathrm{CDOM}}(440)$ and C_a confer the same number of decades of dynamic range in RSE2021. Although scenarios wherein variability in $a_{\mathrm{CDOM}}(440)$ is greater, less than, or comparable to that of C_a arise, the results indicating primacy of $a_{\mathrm{CDOM}}(440)$ are consistent across datasets. The authors add discussion at 467 to discuss the dynamic range expressed in the biochemical quantities for the datasets, as follows:

> "The datasets assess herein represent $a_{\mathrm{CDOM}}(440)$ spanning greater, lesser, or comparable dynamic range compared to C_a, indicating that the result of $a_{\mathrm{CDOM}}(440)$ primacy is not associated with differences in the dataset ranges of the biochemical constituents."

Table. A1. The quartile summary of biogeochemical quantities from the RSE2021, RSE2022, and RSE2007 dataset field observations, plus estimated values for $b_{\mathrm{bp}}(443)$ derived following Matsuoka et al. (2013) and Hooker et al. (2021a). Out-of-bound values are indicated as limit of detection (LOD).

Dataset	Parameter	Min	Lower Quartile	Median	Upper Quartile	Max
$a_{\mathrm{CDOM}}(440)$ $[\mathrm{m}^{-1}]$	RSE2021	0.0010	0.0175	0.4550	0.2185	2.1460
	RSE2022	0.0015	0.0028	0.0044	0.0088	0.0795
	RSE2007	0.0004	0.0027	0.0040	0.0124	0.1242
C_a $[\mathrm{mg\,m}^{-3}]$	RSE2021	0.0590	0.2925	0.7440	2.690	44.1960
	RSE2022	0.0190	0.0580	0.1015	0.2520	4.1510
	RSE2007	0.0190	0.0780	0.1450	0.4815	7.3180
$b_{\mathrm{bp}}(443)$ $[\mathrm{m}^{-1}]$	RSE2021	LOD	0.0010	0.0050	0.0290	0.3290

3. *Finally, I suggest the authors include an example of improved Ca retrieval after determining aCDOM first. I know this is a non-trivial and challenging request, but I think it would be a very convincing proof that the change of paradigm is necessary.*

Authors Response 30: The authors acknowledge the reviewer's point, but respectfully do not perceive that adding an example of C_a retrieval would improve the manuscript. The authors summarize four main reasons for their perspective, as follows:

First, improving phytoplankton biomass estimation (e.g., C_a) via simultaneous or prior estimation of dissolved organic absorption is very well documented in the literature. Briefly, the topic supports the development of so-called semi- and quasi-analytical algorithms—including the Generalized Inherent Optical Property (GIOP), Garver Siegel Maritorena (GSM), and Quasi-Analytical Approach (QAA) algorithms (e.g., Maritorena et al. 2022; Lee et al. 2002; Werdell et al. 2013)—and is inherent in machine-learning approaches (e.g., O'Shea et al. 2023). In these and many other instances, advances in quantifying phytoplankton biomass is shown via corresponding retrieval of dissolved organic (and other) optical properties.

The authors will improve clarity of the historical context, by modifying L560–561, 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)."

Second, demonstrating that C_a retrieval is improved by first determining $a_{\mathrm{CDOM}}(440)$ is documented by Fig. 2. In this figure, all three datasets indicate higher R^2 values for C_a estimation using an $a_{\mathrm{CDOM}}(440)$ algorithm (panels d–f) compared to that derived using the OC algorithm (panels a–c). In other words, Fig. 2 proves that estimation of C_a improves when $a_{\mathrm{CDOM}}(440)$ is derived first, because Fig. 2 shows that a loglinear fit of algorithmic $a_{\mathrm{CDOM}}(440)$ improves C_a estimation. The manuscript states on L384–386:

> "Although the differences observed are slight for RSE2007 (indicating the dataset adheres well to case-1 waters), all datasets nonetheless produce stronger R^2 statistics when comparing the OC algorithm values to $a_{\mathrm{CDOM}}(440)$ versus C_a."

Third, the suggestion would unnecessarily expand the scope of the submitted manuscript, which already approaches 20 pages without including figures or other revisions described herein. Responsibly producing a C_a algorithm would also require much more analysis to ensure that the observations being fit—as well as the fitting procedure—produce C_a estimations that are representative of natural waters. The C_a algorithm would also require addition of substantial new text to describe the quantification of uncertainties and define the limitations for the new C_a approach.

Comments from another reviewer also provide an example as to why the authors do not anticipate that adding a new approach for estimation of C_a would be beneficial: algorithms produced to investigate a topic or provide an example can also reduce clarity of a manuscript or cause confusion later. For example, a notebook analysis discussed in Authors Response 3 invokes algorithms (e.g., the blue and green dots from the link in Authors Response 3) that were previously produced for comparison purposes to investigate whether dissimilar datasets—corresponding to significant differences in data quality and bias—produce similar degradations in algorithmic performance when the spectral range of the algorithmic relationship is decreased.

Fourth, the authors perceive that adding a new method for estimating C_a would decrease the clarity of the manuscript by redirecting the focus towards C_a estimation and ocean color algorithm development.

Other minor and more specific comments are:

- *The references throughout the manuscript to "end-member analysis (EMA)" are somewhat confusing. At first I thought that it was a stablished methodology (the term is actually used in other disciplines), but then I realized that it is unrelated to the methodologies used in other disciplines, and it rather seems to refer to a type of algorithm developed by the authors in previous works, referring to the use of wavebands in the extremes of the spectra to estimate CDOM absorption. I think this should be clearly stated in the manuscript to avoid any confusion.*

  Authors Response 31: The authors clarify on this point by defining end-member analysis at its first usage. The authors revise L106–107 (also described in Authors Response 2), 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..."

- *Lines 19-20: "Spectral subset eigen analyses indicate expansive spectral range observing improves the independence in retrieving CDOM absorption and Ca." I cannot understand this sentence, please rephrase it.*

  Authors Response 32: The authors revise L19–L20, as follows:

  > "Eigenanalyses applied to spectral subsets of the data indicate expansive spectral range observing improves the independence in retrieving CDOM absorption and C_a."

- *Lines 58-84: Although I find the information summarized in these paragraphs very interesting, I do not think that the level of details included here is necessary, particularly because it does not flow with the reading of the previous and following paragraphs. I would recommend shortening this part and reducing the number of citations to those only necessary to support the current work.*

Authors Response 33: The authors perceive this material as necessary to understand the differences in bias between the *in situ* datasets leveraged herein. To improve flow based on the reviewer's comment, the authors shorten and re-organize this material so that the discussion occurs nearer to existing methodological discussion of the *in situ* datasets. The authors remove L71–84 and add a portion of this text at L189, as follows:

> "…approximately the length of the downward-pointing radiance radiometer (Hooker et al. 2020). Improvements in sampling rates support management of high-frequency, non-Gaussian variability in flux observed by an above- or in-water instrument (due to glint and wave focusing, respectively), thereby expanding the spectral range of optical data products to preserve INV information and retaining information associated with spectral signal amplitudes or brightness (Houskeeper et al. 2023; 2024).
>
> SOTA advances in hyperspectral instrumentation include hybridspectral sensing configurations (Hooker et al. 2022), wherein microradiometer and spectrograph observations are obtained in concert, with the multispectral microradiometer providing necessary quality control for the hyperspectral spectrograph. The quality control is desirable, in part, because COTS spectrographs suffer from slower integration times, narrower dynamic range, and a degraded signal-to-noise ratio (SNR) relative to COTS microradiometers (Houskeeper et al. 2024; Kudela et al. 2019; 2024). In addition, a radiance control arm positions the $L_u(z, \lambda)$ aperture near the water surface and approximately aligned with the upward-pointing irradiance radiometer (Hooker et al. 2018a). Improvements in the number of spectrograph pixels to as high as 2,048…"

This also requires minor revisions to L120 to preserve presentation of acronyms, as follows:

> "…preclude most non-visible—or invisible (INV)—data products…"

As well as to and L101:

> "…spectral range, resolution, and signal-to-noise ratio (SNR) of legacy ocean color datasets…"

- *Line 84: The reference to Ruddick et al. (2023) is missing.*

Authors Response 34: The authors thank the reviewer for catching the missing reference. The Ruddick et al. 2023 citation is removed in Authors Response 33.

- *Line 104, 154: I am not familiar with the expression "spectrally expansive", if this is well-known terminology, please disregard this comment, otherwise, I suggest to briefly define what the authors mean by this.*

  Authors Response 35: The authors add a sentence defining spectrally expansive following its first usage on L104, as follows:

  > "Spectral *expansivity* herein refers to the spectral range of (compliant) data products, with the most spectrally expansive observations corresponding to those representing the greatest spectral range."

- *Lines 151, 199, 204, 210, 221, etc. What are contemporaneous measurements? Do you mean simultaneous?*

  Authors Response 36: The authors revise the text to define contemporaneous—following the Merriam-Webster dictionary—at its first usage on L63, as follows:

  > "... This era of development also included contemporaneous—i.e., occurring during the same time—advances in. . . ."

- *Line 268: "the square root transformation improves normality" Why? Is there a reference to support this assertion?*

  Authors Response 37: Transforming a dataset using, e.g., the natural logarithm can be helpful for improving normality so that the assumptions of some parametric statistical tests are approximately satisfied. Normality may be assessed, e.g., by visually inspecting histograms of the data products, as well as using metrics such as the Shapiro-Wilk test statistic, W, wherein values closer to 1 indicate normality.

  In their study, the authors tested various transformations of the different datasets using W and found that no one transformation type was optimal across all wavelength domains. Overall, the authors found that W was maximal in the longer wavelength domain (> 500nm) if logarithmic transformation were applied, and maximal in the shorter wavelength domain (< 500nm) if root transformations were applied. The authors applied a square-root transformation because most absorption variability in aquatic optics is associated with the blue (< 500nm) domain (e.g., Chase et al. 2013). In addition, the authors performed sensitivity testing to test whether the results were insensitive to the transformation type. Briefly, natural-log transformation, square-root transformation, and no transformation all produced eigenfunctions with the same approximate spectral shapes.

  The authors revise L268 to clarify this topic, as follows:

"…observations of the square root of $\left[L_W(\lambda)\right]_{\mathrm{N}}$, hereafter $\left[L_W(\lambda)\right]_{\mathrm{N}}^{0.5}$, because the square-root transformation was found to improve normality in the blue spectral domain based on the Shapiro-Wilk test statistic, W. Sensitivity testing was performed by reproducing the eigenanalyses using natural-log transformation, as well as no transformation, to assess whether the spectral shapes of the eigenfunctions were approximately insensitive to the transformation treatment."

- *Lines 270-276: I must admit that I am not very familiar with PCA applied to spectral datasets, and it would have greatly helped if the authors included matrices dimensions in equations (3) and (4).*

Authors Response 38: The authors revise the text to add dimension information based on the reviewer's comment. First, the authors revise L271, as follows:

"…in which $\boldsymbol{C}$ is the covariance matrix of the $\left[L_W(\lambda)\right]_{\mathrm{N}}^{0.5}$ values for each dataset, with square dimensions, each corresponding to the number of wavelengths in each dataset. The $\boldsymbol{\psi}$ value denotes…"

Next, the authors revise L278, as follows:

"…(where i is the eigenfunction index). The length of each $\boldsymbol{S}_i$ column is the number of observations in the underlying dataset."

- *Line 283-284: "after adjusting for covariance with the biogeochemical quantity y." How was this done? Please describe or add a reference.*

Authors Response 39: Adjusting the correlation coefficient for covariance with another quantity is a statistical technique termed the partial correlation coefficient. The authors add a reference, plus re-organize L281–284 to add clarity, as follows:

"Because biogeochemical parameters covary, e.g., C_a and $a_{\mathrm{CDOM}}(440)$ are strongly correlated (Morel and Prieur 1977), partial correlation coefficients (Fisher 1924) are derived using notation wherein $\rho_{i,x|y}$ indicates correlation of $\boldsymbol{S}_i$ with biogeochemical quantity x after adjusting for covariance with the biogeochemical quantity y."

- *Lines 335: "or fluorescence properties" instead of "of fluorescence properties"?*

Authors Response 40: The authors thank the reviewer for catching this typo. The authors revise the text on L335, as follows:

> "…the spectral domains of C_a absorption or fluorescence properties."

- *Lines 438-440: I do not understand the sentence that begins with: "The variability may, perhaps, correspond in part to…" I do not follow the connection with the previous sentence. Please rephrase it.*

Authors Response 41: The authors agree that this sentence should be revised to improve clarity, and revisions to the sentence (and paragraph in which this sentence occurs) are described in Authors Response 18. The revision is, 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."

- *Line 580: "aCDOM and Ca absorption" instead of "aCDOM and CDOM absorption"?*

Authors Response 42: The authors thank the reviewer for catching this mistake. The authors revise the text, as follows:

"…confounding effects of spectral overlap in C_a and CDOM absorption…"

The authors thank Reviewer 2 for useful comments and suggestions, which helped the authors to improve the manuscript.

References

Bricaud, Annick, Hervé Claustre, Josephine Ras, and Kadija Oubelkheir. "Natural variability of phytoplanktonic absorption in oceanic waters: Influence of the size structure of algal populations." *Journal of Geophysical Research: Oceans* 109, no. C11 (2004).

Chase, Alison, Emmanuel Boss, Ronald Zaneveld, Annick Bricaud, Herve Claustre, Josephine Ras, Giorgio Dall'Olmo, and Toby K. Westberry. "Decomposition of in situ particulate absorption spectra." *Methods in Oceanography* 7 (2013): 110-124.

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.

Fisher, Ronald Aylmer. "The distribution of the partial correlation coefficient." *Metron* 3 (1924): 329-332.

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

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.

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, 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.

Roesler, Collin S., and Andrew H. Barnard. "Optical proxy for phytoplankton biomass in the absence of photophysiology: Rethinking the absorption line height." *Methods in Oceanography* 7 (2013): 79-94.

Ruddick, Kevin G., Pieter De Vis, Clémence Goyens, Joel Kuusk, Héloise Lavigne, and Quinten Vanhellemont. "Second derivative water reflectance spectra for phytoplankton species detection: origin, impact, and removal of spectral wiggles." In *Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions* 2023, vol. 12728, pp. 81-100. *SPIE*, 2023.

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

Uitz 2006

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.