We thank reviewer 1 for the thorough evaluation of our manuscript contribution. Below are our responses to the general and specific comments (quoted in **bold**). All other minor spelling and grammar revisions noted will be incorporated in the revised manuscript with tracked changes, unless otherwise noted.

Please provide information on the reporting practices, particularly for concentrations near LOD/LOQ, and on the number of significant digits reported. I am also missing an evaluation of how many datasets actually had CV=0%, and how many of these cases were associated with measurements at or below LOD/LOQ. Furthermore, please clarify how such numbers were reported: were concentrations replaced by LOD/LOQ values, e.g., 0.0001, or were zeros reported when pigments were not detected? This information is important for understanding the procedure used to exclude samples with invariant replicates (CV=0%).

The NASA GSFC facility issues a comprehensive analytical report to each sample submitter that includes pigment concentrations in μg L^{-1} (i.e., mg m^{-3}) alongside sample-specific effective limits of quantitation (LOQ). The effective LOQ is calculated from instrument-level LOQ values (expressed in ng per injection) carried through to concentration units (μg L^{-1}) using the extraction volume and sample-specific filtration volume, recognizing that the same instrumental LOQ translates to substantially different concentration detection thresholds depending on the volume of seawater filtered. Concentrations are typically reported to three decimal places (0.001 μg L^{-1}).

Pigments not detected in a given sample are flagged with a numeric replacement value rather than reported as zero or substituted with LOD/LOQ values. Prior to March 2016, the flag -111 was used; from March 2016 onward, this was updated to -8888 to align with conventions used by other oceanographic data repositories. In both cases, the flag explicitly distinguishes "not detected" (investigated and determined to be below detection) from missing values (measurement not performed). All flagged non-detect values were excluded from analysis prior to any precision calculations. Consequently, invariant replicate sets (CV% = 0) in our dataset represent cases where two or more replicates yielded identical detectable concentrations above the LOD, not artifacts of below-detection reporting conventions.

The frequency of invariant replicates per pigment is reported in Table 2 and ranges from 3.5% for TChl *a* to 71.3% for DVChl *b*. As discussed in the manuscript (Discussion, L451), this range reflects the interaction between concentration range and analytical resolution rather than analytical quality: pigments with narrow, low concentration ranges such as DVChl *b* have a much higher probability of yielding identical quantified values in replicate sets --particularly duplicates-- because the limited degrees of freedom near the analytical lower limit constrain the number of possible reported values at 0.001 μg L^{-1} resolution. Pigments spanning wider concentration ranges, such as TChl *a*, have far more possible quantified values, reducing the probability of invariant replicates occurring by chance alone.

Chlide *a* generally results from artificial degradation of Chl *a* during extraction and should therefore be added to Chl *a*. Since most Chlide *a* quantified by HPLC is not natural occurring it should not be interpreted independently and I suggest removing Chlide *a* from the figures and most of the text (see a later comment on it).

We agree that Chlide *a* largely results from chlorophyll *a* degradation during extraction due to water content in acetone, and we will acknowledge this in the manuscript. However, we disagree with the suggestion to remove it from the analysis and figures for the following reasons.

As indicated in Table 1, Chlide *a* is already incorporated into the reported TChl *a* value (TChl a = [Chlide *a*] + [DVChl *a*] + [MVChl *a*]), so its contribution to the primary quantity of interest for ocean color validation is fully accounted for. Nevertheless, the GSFC facility reports Chlide *a* as a discrete value precisely because its availability to data end users serves purposes beyond TChl *a* computation. Its concentration provides direct insight into sample handling and extraction quality --elevated Chlide *a* relative to TChl *a* can flag potential degradation issues in individual samples. Furthermore, the facility's role is to provide the most complete and transparent data record possible; restricting reported values based on assumed end-user objectives would be

inconsistent with that mandate. We therefore retain Chlide *a* in the figures and analysis, while ensuring the manuscript clearly describes its artifactual origin and its inclusion in TChl *a*, so that readers have the full context needed to interpret its precision characteristics appropriately.

L.80: "In the absence of standardized reference materials… intercalibrations are a necessary substitute"?: The use of standardized reference materials and participation in intercalibrations are complementary activities and should both be applied, where possible, to ensure robust quality assurance.

We agree that intercalibrations with standard reference materials and natural samples are complementary activities and that both should be applied where possible. We will revise the manuscript language to clarify that there is an absence of certified reference materials (CRMs, such as those available from the NIST in the US) and intercalibrations that include both pigment standards and samples are complementary activities. However, we maintain that the situation for phytoplankton pigments, and marine particulate matter in general, is distinct from other oceanographic measurements in that true matrix-matched CRMs for such analytes does not currently exist.

The challenge of developing CRMs for marine biogeochemical measurements is well recognized in the oceanographic community. As highlighted in the National Research Council's review of chemical reference materials for ocean science, compositional or matrix reference materials derived from natural substances such as phytoplankton offer advantages over primary standards by better matching sample composition and minimizing matrix effects, yet such materials remain unavailable for phytoplankton pigments (National Research Council, 2002). The pigment standards used for HPLC calibration are pure compounds in solution, which do not capture the matrix complexity of whole phytoplankton cells processed through filtration, storage, and extraction. In this context, intercalibration exercises such as SeaHARRE, which distribute natural field samples among participating laboratories, remain the most appropriate available substitute for accuracy assessment against a matrix-matched reference. We will clarify this distinction in the revised manuscript.

L. 110: one ancillary pigment? Which pigment? It is evident from Table 1. Also, the text states "two sets of pigment sums and ratios routinely reported"; however, Table 1 shows more than two pigment sums and no ratios. Please correct the text/table.

L. 114: "Tertiary pigments are a set of less frequently analyzed pigments" -less frequently reported pigments?

We agree that the ambivalent allusion to gyroxanthin diester in two separate sentences created unnecessary redundancy, and the description of pigment sums and ratios was inconsistent with Table 1; the table will be revised accordingly. We have revised Section 2.1 as shown below:

L. 114: Why specifically mention gyroxanthin diester (which is often present in certain regions) while other omitted pigments are not mentioned (dinoxanthin, myxoxanthophyll, C2-MGDG, etc.)?

We appreciate the reviewer raising this point, as it allows us to clarify the rationale for specifically mentioning gyroxanthin diester. The GSFC facility chromatographically resolves several pigments beyond those reported, including dinoxanthin, astaxanthin, myxoxanthophyll, canthaxanthin, and crocoxanthin, but does not calibrate for or report them, as demand from the ocean color validation community for these data products has been minimal. Dinoxanthin, for example, is well resolved chromatographically and its separation is important precisely because co-elution with diadinoxanthin would compromise the latter's quantitation; however, dinoxanthin itself is not routinely requested by investigators nor does it carry independent diagnostic value beyond what peridinin already provides for dinoflagellate detection.

Gyroxanthin diester warrants specific mention because it occupies a unique position among less-reported pigments: it is the diagnostic marker for bloom-forming toxic dinoflagellates such as *Karenia brevis* that lack peridinin, the standard dinoflagellate indicator used in pigment-based phytoplankton functional type

algorithms. In regions prone to *Karenia* harmful algal blooms, its absence from a validation dataset risks misidentifying a toxic dinoflagellate bloom as a diatom or haptophyte assemblage due to shared fucoxanthin dominance. Additionally, gyroxanthin diester absorbs strongly in the blue-green spectral region (~430–470 nm), overlapping with primary satellite sensor bands (e.g., 443 nm, 490 nm on MODIS and VIIRS), meaning its presence measurably affects in situ phytoplankton absorption coefficients used in ocean color validation matchups. These combined factors, such as taxonomic misidentification risk, harmful algal bloom relevance, and direct spectral impact on validation quantities, distinguish gyroxanthin diester from other unreported pigments and justify its specific mention.

We will revise the manuscript to briefly acknowledge that other pigments are resolved but not reported, and to clarify the specific diagnostic and ocean color validation rationale for singling out gyroxanthin diester.

L. 131: Please specify the extraction solvent (e.g. 90% acetone).

The reviewer is correct that the extraction solvent should be specified more explicitly. This information, along with a more complete description of the extraction procedure, will be added to Section 2.2 in response to a related and more detailed comment from Reviewer 2, which we address on our response to their comment. The revised manuscript will clearly specify the extraction solvent as approximately 90% acetone.

L. 144: 30,000 samples mentioned in the introduction?

The 30,000+ figure cited in the introduction refers to the total number of samples processed by the NASA GSFC HPLC facility since its inception, which includes samples not directly relevant to ocean color validation activities, such as size-fractionated samples, phytoplankton cultures, experimental incubations, and damaged samples. The 19,125 figure at L. 144 reflects the subset retained after applying the inclusion criteria described in Section 2.3. We will revise the language at L. 144 to make this distinction explicit.

L. 190: Correct "pigmen" to pigment. ✅

Table 1: Correct Pheophtyin a to Pheophytin a. ✅

L. 128: "a reference wavelength of 700 nm was added" Do you mean: "…a reference wavelength at 700 nm (±10 nm) was included and used for baseline correction of the 665 nm signal"?

The reviewer is correct and the wording will be changed to the reviewer's suggestion.

L. 250, Fig. 2d: What is the rationale for comparing satellite data from 2002-2023 to TChl a from 2011-2022? Could the offset observed in Fig. 2d partly reflect the difference in time periods?

The comparison was not intended as a temporal matchup between the two datasets but rather as an assessment of how well the concentration range represented in the GSFC validation dataset covers the full dynamic range of TChl *a* observed by the longest continuous ocean color satellite record to date (MODIS-Aqua, 2002–present). The intent is to characterize whether the in situ dataset adequately samples the global TChl *a* distribution relevant to ocean color validation, particularly at the lower concentration ranges typical of oligotrophic open ocean waters. The observed offset toward higher concentrations in the in situ dataset reflects the coastal bias in sample geographic distribution discussed earlier, not a temporal mismatch

between the two records. We will revise the manuscript language to make this rationale explicit and avoid the impression of a direct temporal comparison.

L. 257: Can you elaborate on why Diato and Perid in particular exceed the 8 % benchmark? Possible factors could include broad peaks, coelution, or generally small peak areas. Please expand on this in the discussion.

Two distinct mechanisms are likely responsible for the slightly elevated CV% for these pigments. For Diato, its role in the reversible xanthophyll cycle, where it is interconverted with diadinoxanthin as a photoprotective response, means that Diato is typically present in relatively small amounts compared to its precursor Diadino. This cycling relationship results in Diato frequently occurring at low concentrations, where quantitation uncertainty could be higher, while Diadino, being the dominant form is less affected. It is worth noting that combining Diato and Diadino into a single sum consistently often improves precision, though this is not standard reporting practice. For Perid, its elution position in the chromatogram is generally free from significant interferences; its slightly elevated CV% is therefore more likely attributable to the predominance of small pigment amounts (see Fig. 1). The manuscript will be amended to include this discussion on possible reasons why these two carotenoids exceed the 8% benchmark.

This interpretation is consistent with our overall finding that concentration does not explain precision variability across the full dataset, as the effect here is a within-pigment phenomenon, reflecting the characteristically low concentrations of Diato and Perid relative to the other primary pigments, rather than a systematic across-pigment driver.

L. 293: (CV % >0 (excluding invariant replicates). One ")" is missing. ✅

Fig. 5, 6, and 7: It is difficult to see the details, especially the white lines. Can the figures be enlarged?

We agree on this issue. This is a consequence of the number of figures and the constraints imposed by keeping figures proximate to their relevant text within the submission manuscript format in MS Word. We are confident that the journal's typesetting process during final production will resolve this issue, as the production team has considerably more flexibility in figure placement and sizing than is available in the preprint format. All figures have been prepared at high resolution and will render clearly at larger sizes in the final typeset article.

L. 383: As mentioned earlier, Chlide a is (mostly) an artifactual product from chlorophyll a formed during extraction due to water content in acetone. I suggest briefly describing this with references and then remove chlorophyllide *a* from the article (it would still be included in TChl a).

Please see our response to this issue above

L. 400: "Volumes below 200 mL showed precision degradation in our dataset": refer to Fig. 7? Where can it be seen that there will be minimal additional precision benefit above 1 L in high biomass waters?

Indeed, a figure reference is lacking in that sentence. We will add an explicit reference to Fig. 7 at L. 400 for the statement regarding precision degradation below 200 mL.

Regarding the observation of minimal additional precision benefit above 1 L, we concede the reviewer's point that the manuscript language specifically attributing this pattern to high biomass waters is not supported by the figure data. Indeed, samples with filtration volumes exceeding 1 L are more likely to represent meso- to oligotrophic waters where larger volumes are typically required to collect sufficient biomass, rather than high biomass conditions. We will remove the specific reference to high biomass waters accordingly.

The broader observation that there is minimal additional precision benefit above 1L remains supported by the LOESS fits for the full dataset in Fig. 7. With the exception of Allo (Fig 7.g), Diato (Fig 7.i), Zea (Fig 7.l), Lut (Fig 7.t), and the tertiary pigments (Fig 7.u–y), for which the LOESS fits show noticeable continued improvement in precision above 1000 mL, the trend for all other pigments is nearly flat or shows a slight increase in CV% above that filtration volume. We will revise the manuscript language at L. 400 to remove the high biomass characterization while retaining the general precision plateau observation and identifying the pigment-specific exceptions explicitly.

L .430: ".The divergent behavior at high filtration volumes, where precision degrades despite greater sample volume, supports the interpretation that physical stresses during extended filtration (e.g., cell lysis, filter overloading) compound the natural heterogeneity already present". When filtering large volumes in low Chl a concentration (oligotrophic) regions, e.g. 4-5000 mL, many pigments (e.g., prasinoxanthin, neoxanthin, lutein, alloxanthin) from algal species not abundant in these areas will still be detected near their LOD/LOQ, which means that there will be increased uncertainty on the determination of the peak areas. While natural pigment concentrations span ~4 orders of magnitude, the filtered volumes do not.

This is an insightful observation that complements the interpretation presented in the manuscript. In oligotrophic waters, large volumes (>4 L) are filtered precisely because pigment concentrations are low; yet for rare pigments from low-abundance taxa, the amount of material collected may still fall near the LOD/LOQ regardless of the volume filtered. This can introduce increased uncertainty in peak area quantitation at low signal-to-noise levels, representing an additional contributing mechanism to precision degradation at high filtration volumes for this subset of pigments, alongside the physical stress interpretation already discussed. We will revise the manuscript to acknowledge this detection limit mechanism as a complementary potential factor contributing to the observed precision degradation pattern.

L. 447: At least part of the explanation may again relate to low cell numbers and the higher uncertainty associated with very small pigment peaks from sparsely represented taxa.

We agree with the reviewer's observation. As with the high filtration volume discussion, the precision degradation observed when microplankton fraction decreases likely reflects at least in part the low cell numbers and consequently small peak areas for pigments diagnostic of sparsely represented taxa, introducing quantitation uncertainty at low signal-to-noise levels. We will revise the manuscript to acknowledge this as a contributing mechanism alongside the filtration-induced fragility interpretation.

L. 452-455: Please provide further details on reporting practices near the LOD/LOQ (see general comments).

Please see our response to this issue above

L. 463: The statement that larger filtration volumes reduce relative measurement error in oligotrophic environments is not clearly supported for all pigments in Fig. 7, although it may hold if even larger volumes were filtered (acknowledging practical limitations).

The statement is not primarily based on the LOESS trends in Fig. 7 but rather on a fundamental metrological principle: when the same volume measurement device is used across a range of filtration volumes, the absolute measurement uncertainty is approximately constant, meaning the *relative* uncertainty in the measured volume is inherently larger for smaller volumes. For example, a fixed absolute uncertainty of ±5 mL represents a 5% relative error at 100 mL but only 0.5% at 1000 mL. This relative volume measurement error propagates directly into the calculated pigment concentration, contributing to precision degradation at smaller filtration volumes independently of any pigment-specific analytical factors. We will revise the manuscript language to make this metrological rationale explicit, so the statement is clearly grounded in volume measurement uncertainty rather than solely in the Fig. 7 LOESS trends.

L. 479: iSeaHARRE-5 – should "in" be added?

Yes, that's a typo.

L. 487-which methodological factors are referred to here (e.g., calibration, peak-area integration)?

The "other methodological factors" referred to at L. 487 include pre-analytical and field-side procedures that may differ between coastal and oceanic sampling operations, such as filtration volumes, as well as differences in the pigment selectivity of analytical methods chosen by participating laboratories in the SeaHARRE exercises, the impact of which can vary depending on water source. This latter factor is particularly relevant in the SeaHARRE context: coastal waters harbor a more diverse and complex pigment composition than oceanic waters, meaning that differences in chromatographic resolution among laboratory methods may manifest more prominently in coastal samples where a broader suite of pigments must be resolved simultaneously. We invoke the SeaHARRE context specifically to support our interpretation: since SeaHARRE demonstrated that properly validated laboratory analytical procedures yield consistent precision regardless of sample origin, the coastal vs. oceanic precision differences we observe in our dataset, where all samples are processed by a single standardized facility, eliminating inter-laboratory method selectivity as a factor, are more plausibly attributable to pre-analytical and field-side factors. We acknowledge that we cannot fully enumerate or verify which specific factors are responsible, as the necessary metadata to directly test these hypotheses are not available, and we will revise the manuscript to reflect this more carefully.

L. 503. Chlorophyllide *a* should not be mentioned here, as it is an extraction artefact (see comments above).

Please see our response to this issue above