

Thanks to the anonymous reviewer for taking the time to do a thorough review of our revised submission. Below are the responses to their comments. Our responses to the comments are quoted in *italics* underneath each raised issue or question:

## General Comments

This manuscript aims to define HPLC replicate sample precision and characterize the associated uncertainties. For each available pigment, individual coefficient of variation values was computed for sample replicate sets, and the precision was thereafter assessed using five different sample parameters. Precision was found to differ between oceanic and coastal samples, from which oceanic samples had better precision in the full dataset compared to the censored dataset (samples with invariant replicates, CV%=0). Following various analyses, the authors concluded that no single variable could explain precision variability deviations across all pigments. However, it seemed that filtered volume and pigment concentrations were important for the full- and the censored dataset, respectively. Furthermore, in consensus with NASA's Ocean Biology and Biogeochemistry (OBB) program, the authors underpinned the importance of prioritizing minimum 5% of samples be collected in replicate in order to enable precision assessment.

The manuscript is highly relevant for the scientific community. Remote sensing is an emerging method from which Chl *a* and other environmental variables can be retrieved from - on a large and global scale. As more fields of science utilize remote sensing to establish time-series and to put findings into a broader context, this work benefits the broader scientific community. Quality insurance is therefore key to understanding and minimizing uncertainties. The scientific quality and presentation were excellent throughout the manuscript. The methodology and definitions were clear, and all the figures were useful tools for the reader.

## Detailed Comments

In its current form, the manuscript addresses relevant scientific questions within the scope of BG and only a few minor suggestions/notes for this manuscript were found (besides the ones already addressed by RC1 and RC2 in April):

**Comment 1:** It would be a great asset to add a table to the appendix that shows the % sample distribution for each major ocean basin as well as the distribution of coastal and oceanic samples.

*We appreciate the suggestion but respectfully disagree for the following reasons. Regarding the distribution of coastal versus oceanic samples, this information is already reported in the manuscript. Sample sizes are explicitly stated in Section 2.4 (884 oceanic and 1,909 coastal samples), and Figure 9 depicts the full CV% distribution for coastal and oceanic samples for each pigment via violin plots, providing a more informative representation of the distributional differences than a summary table would offer.*

*Regarding the distribution of samples across major ocean basins, we addressed this directly in Section 2.4 (~L254), where ocean basin was explicitly considered as an analytical variable but ultimately excluded because the severely imbalanced sample distribution across basins prevented robust statistical analysis. As shown in Figure 2, the dataset is heavily skewed toward the Eastern seaboard of North America, with 41% of all samples concentrated in the region bounded by 17–47°N, 60–100°W. Under such conditions of severe imbalance, any basin-level comparison of precision statistics would be statistically unreliable: the heavily sampled Eastern North American coastal region would dominate the within-basin estimates for the Atlantic, while basins such as the Southern Ocean, Indian Ocean, or Pacific would be represented by sample sizes too small to yield meaningful or comparable precision estimates. Presenting such a table without appropriate statistical caveats risks conveying false equivalence between well-sampled and poorly sampled regions, which would be misleading rather than informative. For these reasons, we believe a basin-level distribution table would not contribute meaningfully to the objectives of the manuscript and may in fact obscure rather than illuminate the geographic limitations of the dataset, which are already clearly communicated through Figure 2 in the context of coastal vs. oceanic sample distribution.*

**Comment 2:** Figure 2: Is panel d also global?

Yes, as stated in panel d's caption: "Histograms comparing satellite chlorophyll a concentration from panel (a) versus in situ total chlorophyll a measurements from HPLC analysis of samples **collected globally** from 2011–2022. Vertical lines mark the median concentration for each dataset." We will ensure the caption makes this explicit.

**Comment 3:** Under the section "Dataset" the number of samples could be clarified: 30,000 vs 19,125.

We have modified the text to further clarify the difference between these two sample totals. Briefly, the 30,000 represents the total number of samples processed by the facility since its inception, which includes sample types not directly relevant to ocean color validation activities such as size-fractionated samples, experimental samples from nutrient addition assays, productivity incubations, and phytoplankton cultures. The 19,125 samples analyzed here represent the subset that is amenable to satellite validation activities, meeting the inclusion criteria described in Section 2.3. We have restructured the opening of Section 2.3 to make this distinction prominent and unambiguous.

**Comment 4:** L265-275: Please specify what is meant by "the entire datasets."

We note that line numbers may differ slightly between manuscript versions. The phrase "entire datasets" refers to the non-censored dataset that retains invariant replicates ( $CV\% = 0$ ), as opposed to the censored dataset from which those samples are excluded. We have added a clarification at the relevant location in the manuscript to make this distinction explicit for the reader.

**Comment 5:** L435: The discussion could benefit by extending the section with the guidance of implementing best practices of POM collection. How could lack of POM guidelines introduce systematic variations in data quality across different ocean basins? I would suggest to add the laboratory specific LOD for all the available pigments in the appendix. This was a really exciting manuscript that establishes the importance of analytical uncertainties besides instrumental sensitivity and repeatability.

We thank the reviewer for these constructive suggestions and for the generous assessment of the manuscript. Regarding the extension of the best practices discussion, starting at L511 we have added a paragraph drawing on the IOCCG Protocols for the Collection and Analysis of Particulate Organic Matter (IOCCG, 2021) to provide concrete guidance on pre-analytical procedures that directly address the precision variability patterns identified in this study. Key recommendations highlighted include vacuum pressure monitoring during filtration, limits on filtration duration to minimize cell lysis and pigment degradation, accurate volume measurement using calibrated devices, immediate immersion of filters in liquid nitrogen following filtration, and maintenance of an uninterrupted cold chain from collection through analysis. We also note that in the absence of universally adopted field collection standards, variability in adherence to these practices across research groups, vessels, and ocean basins introduces systematic differences in pre-analytical sample quality that are difficult to distinguish from biogeochemical variability, and that systematic documentation of field sampling conditions alongside analytical data would substantially improve the ability to diagnose pre-analytical sources of precision variability in future assessments.

Regarding the suggestion to add laboratory-specific LOD values for all pigments, we have revised Section 2.3 to illustrate the volume-dependence of effective LOQ values explicitly in the manuscript narrative, which we believe communicates the relevant information more efficiently than a standalone appendix table. Specifically, a typical instrument-level LOQ of 0.06 ng per injection yields an effective LOQ of  $0.001 \mu\text{g L}^{-1}$  when 1000 mL is filtered and the sample is extracted in 2.5 mL, rising to  $0.006 \mu\text{g L}^{-1}$  at 200 mL. For the most common filtration volumes received at the GSFC facility, effective LOQ values are typically in the range of  $0.001\text{--}0.003 \mu\text{g L}^{-1}$ , making the per-sample effective LOQ approach essential for meaningful interpretation of near-threshold measurements.