

RC2

Review report of Riom et al. “Three decades of marine bioregions evolution through analysis of phytoplankton phenology”

This manuscript addressed an important point in large-scale ocean monitoring: distinguishing interannual variability from persistent, phenology-driven shifts in marine bioregions. The approach used supervised clustering on satellite derived chlorophyll-a products by analysing the phenologies and suggested a poleward extension of less seasonal phytoplankton phenology. The manuscript is mainly clear in conception, methodology, and well-grounded in a previous literature (Baudena et al. 2025). Overall, the study is meaningful and provides new insights on interpretation of marine phytoplankton data sets. I

suggest a moderate to major revision for this manuscript. My comments are summarized in the following aspects:

- 1) the need for clearer methodological justification
- 2) a clearer connection between bioregions and underlying environmental drivers
- 3) some organization and wording improvements for readability
- 4) minor corrections in citations and phrasing

We thank the reviewer for his constructive evaluations of our manuscript. The comments have greatly contributed to improving the justification of the method and better specify the scope of the contribution. We have carefully addressed each point raised and revised the manuscript accordingly. Given the word-limit and figure-number constraints of the Ocean State Report, we have aimed to keep the main text concise while ensuring that all essential information is clearly presented. Additional methodological details, and illustrative material that would have excessively lengthened the manuscript are provided in a dedicated supplementary document attached to this response. A full rereading and language check have also been performed to improve the manuscript's readability and consistency. Attached, we provide a detailed, point-by-point response to all comments and describe the corresponding revisions.

Modifications of the manuscript are highlighted in blue.

Major Comments:

1. There are two satellite CHL products available on Copernicus Marine Service – GlobColour version and OCCI version both derived from the merged sensor products. The authors have apparently used the GlobColour version but L4 products where CHL is gap free by applying a gap filling method by the GlobColour consortium. This is not indicated in the study.

Using L4 gap-free product permits to maximize data availability in the annual ensembles thus maximizing the surface of the bioregion. In contrast, Baudena et al., 2025 identified the seasonal pattern on L3 data, with multi year averages compensating for data gaps. This reasoning behind the use of gap free L4 is now more clearly indicated within section 2.4, Line 78: “This product offers cloud free remotely sensed surface chlorophyll-a concentration at a 4 km spatial resolution and daily intervals. Using such a L4 product (an optimally interpolated, merged satellite dataset), as compared to L3 (merged product) in Baudena et al., (2025), compensates for the lack of observation inherent to satellite ocean color (Stock et al., 2020) and ensures a maximal spatialization of the bioregions at an annual temporal resolution.

2. Copernicus CHL product is with 4x4 km resolution but you had clusters with 1x1°, probably it was performed following Baudena et al. 2025 but a clear explanation or a methodology pipeline is necessary. The methodological pipeline is identical to the one of Baudena et al., (2025). This is now indicated in the supplementary material and referred to in the manuscript at Line 89: “ Following the method described in supplementary section S1, ”

3. Another major concern is that it is well recognized that CHL products from OCCI and GlobColour can generate contrast trends (Pauthenet et al. 2024) and OCCI has better inter-sensor consistency (Zhai et al. 2024). Xi et al. (2025) in OSR9 has tried to tackle this problem for Copernicus ocean color data sets of phytoplankton functional types. I believe this could also bring inconsistencies in the time series clustering in this study however the author did not cover this point and did not justify either why L4

GlobColour product was selected but not L3 or OC-CCI.

Pauthenet, E., Martinez, E., Gorgues, T., Roussillon, J., Drumetz, L., Fablet, R., & Roux, M. (2024).

Contrasted trends in chlorophyll-a satellite products. *Geophysical Research Letters*, 51, e2024GL108916.

<https://doi.org/10.1029/2024GL108916>

Zhai, D., Beaulieu, C., & Kudela, R. M. (2024). Long-term trends in the distribution of ocean chlorophyll.

Geophysical Research Letters, 51, e2023GL106577. <https://doi.org/10.1029/2023GL106577>

Xi et al: <https://doi.org/10.5194/sp-6-osr9-7-2025>

The use of gap-filling method is addressed above in major comment 1.

This study is a proof of concept of a proposed ocean monitoring indicator. Applying our approach to another ocean color product would give valuable insight in the reliability of our bioregionalization results and in the observed trends in bioregion surface areas but extend beyond the scope of this contribution. This is now clearly highlighted in the last paragraph of the Introduction section:

Line 65: “We introduce here a proof of concept for such an approach using the supervised clustering method introduced in Baudena et al., 2025. Annual partition of the global ocean (56°S–56°N) from 1998 to 2024 is obtained by classifying annual time series of remotely sensed chlorophyll-a concentration (Copernicus-GlobColor) into bioregions. [...] Finally, possible correlation with environmental changes and method limitations are evoked in the discussion section.”

Regarding the contrasting trends in ocean color products, we would like to recall that our proposed indicator mostly emphasized changes in the seasonal cycle of chlorophyll-a (i.e., timing) and, therefore, should be less sensitive to long term-trend in the magnitude of surface chlorophyll-a concentrations. Nonetheless, such uncertainties need to be acknowledged, which is now clearly mentioned in the discussion section (with the two papers mentioned by the reviewer) and could lead to potential future works.

Line 272: “Furthermore, identifying long term trends in the distribution of chlorophyll-a concentration was shown to be dependent on the type of ocean color product (Zhai et al., 2024, Pauthenet et al., 2024).

Investigating the presence and variability of the bioregions based on ESA's Ocean Color Climate Change Initiative (Sathyendranath et al., 2019) would further consolidate (invalidate) the highlighted trends. In this paper, we mainly propose an approach, a proof-of-concept, that could be further applied to other ocean colors time-series.”

Clustering methodology needs clearer justification

4. Line 74 ‘To dampen sub-seasonal variations, time series were all smoothed with a 5 weeks centered rolling average method..’ it is not clear why 5-week smoothing window has been chosen. How does this choice influence detection of narrow bloom events (e.g., subpolar spring bloom)? Moreover, there are also mathematically more accurate ways to deseasonalise time series such as LOESS (STL; Cleveland et al., 1990) which is commonly used to decompose trend and seasonal variations.

The time series were created to be representative of seasonal forcing (of event period ~3 months). Smoothing with a five week rolling average, follows Baudena *et al.* (2025) procedure, and dampens presence of peak or noise of monthly (40 days) scale. Nevertheless, narrow bloom events (e.g. subpolar spring bloom) are found to be successfully detected and assigned into cluster 6.

Our approach relies on the clustering of seasonal cycles of chlorophyll-a concentration. Therefore, it cannot be applied to de-seasonalised time series. Removing the long-term trend on the chlorophyll-a concentration data could be a solution to apply our proposed approach to other datasets presenting different long-term trends in chlorophyll-a concentration (see our response to major comment 3). However, for this study, we preferred to closely follow the procedure of Baudena *et al.* (2025).

5. You vary acceptance thresholds (26 thresholds) to capture methodological uncertainty, however, several aspects remain unclear: Why 26 thresholds? A brief methodological rationale (e.g., spacing, range definition, prior sensitivity tests) would strengthen confidence. How sensitive is the partitioning to threshold choices? You mention $\pm 4\%$ classification uncertainty, but no examples of spatial sensitivity are

shown. A supplemental figure could show how cluster assignment changes across extreme thresholds.

To keep the manuscript concise the rationale behind the use of the set of thresholds have been reformulated in the manuscript and detailed in a supplementary section (see supplementary section “Section S2: threshold distance”) with sensitivity tests (see figure S.2). The manuscript now reads as: Line 104: “The assignment of a given time serie into one of the clusters was done according to its minimal euclidean distance (maximal similarity) with the central time series of Baudena et al., 2025 clusters. To prevent assignment of too dissimilar seasonal patterns, a threshold distance was considered to accept (distance below the threshold) or refuse (distance above the threshold) an assignment. To account for the method’s sensitivity to this hyperparameter, we discretized a range of possible values, defined by the minimum and maximum classification distances found among the most deviating elements of any cluster, from 1.33 (cluster 1) to 2.60 (cluster 7), based on the unsupervised clustering of Baudena et al., (2025). Each annual ensemble was partitioned several times from a low (strict definition of the clusters) to high threshold values (lax definition of the clusters) selected within this range with a regular step of 0.05 (see sensitivity analysis in supplementary section 2). At least 88 % of the time series were always successfully sorted within the seven clusters (see Fig. S2b). Unclassified time series, mostly located in regions of low satellite data availability region (see Fig. 2a), were all excluded from ulterior analysis.”

L79 how this 111 km has been determined? Is it related to the coarse resolution of the CHL Product?

111 km is approximately 1° degree over the equator. This distance limit is applied to all coastlines and ensures that none of the considered pixels are contaminated by the coast or coastal waters. An explicit reference to the grid was added in the text. Line 96: “Coastal pixels located within 111 km (~1°) from the coast were also discarded...”

Why seven clusters specifically? Although inherited from Baudena et al. (2025), briefly re- emphasizing ecological or statistical reasoning would improve clarity. Again, as mentioned in my above comment 2, reading Section 2.2 it is uneasy to get a really clear schematic about the whole methodology, I suggest a flowchart illustrating better the analysis pipeline.

Seven clusters were found in Baudena et al. (2025) to partition the phytoplankton phenology over the global ocean. The ecological interpretation of each bioregion is now recalled in the manuscript and a flowchart is added as a supplementary figure S1.

Line 181: “Baudena et al. (2025) offered a biogeochemical interpretation of the bioregions, summarized as follows. Bioregions 6 and 7 correspond to Longhurst’s (1995) model 1 (polar, irradiance-limited). Bioregion 5 exhibits features of both model 1 and model 2 (mid-latitude, nutrient-limited spring production peak). Bioregion 4 closely aligns with the Bloom bioregion identified by D’Ortenzio et al. (2012). Bioregion 3 matches Longhurst’s (1995) model 3 (subtropical, winter nutrient-limited), whereas Bioregion 2 closely resembles the Tropical bioregion described by D’Ortenzio et al. (2012). Finally, Bioregion 1 corresponds to Longhurst’s model 4 (tropical). The additional subdivision introduced here produces spatially coherent domains that distinguish between oligotrophic regions in the centers of subtropical gyres and mesotrophic regions primarily distributed around the equator.”

6. Section 2.4 (uncertainties) could be reorganized to first explain the conceptual approach, then list statistical tools – for now it is not very much clear.

The conceptual approach has been clarified in section 2.2 of the manuscript and more exhaustively detailed in the supplementary material including a flow chart as suggested, by the reviewer. Section 2.4 has also been rephrased.

Section 2.2 (From Line 91) : “Each annual ensemble was then partitioned into the seven clusters of Baudena et al., 2025 based on their seasonality patterns. These seven clusters were found to optimally partition open ocean seasonal cycles of surface chlorophyll-a concentration in respect of multiple statistical tests discussed in the companion study. The assignment of a given time serie into one of the clusters is done according to its minimal euclidean distance (maximal similarity) with the central time series of Baudena et al., 2025 clusters. To prevent assignment of too dissimilar seasonal patterns, a threshold distance is considered to accept (distance below the threshold) or refuse (distance above the threshold) an assignment. To account for the method sensitivity to this hyperparameter, we discretize the set of possible values from the classification distance of the most deviating element of cluster 1 (1.33) to the the classification distance of the most deviating element of cluster 7 (2.60) from the unsupervised clustering of Baudena et al., (2025) . Each annual ensemble is partitioned several times from a low (strict definition of the clusters) to high threshold values (lax definition of the clusters) with a regular step of 0.05 (see sensitivity analysis in supplementary section 2). At least 88 % of the annual time series were

always successfully sorted within the seven clusters (see Fig. S2b). Unclassified time series, mostly located in regions of low satellite data availability region (see Fig. 2a), were all excluded from ulterior analysis.”

And Section 2.4: “2.4 Uncertainties estimation

Following the set of threshold distances, each annual ensemble is partitioned twenty six times (see section 2.2). In each of these partitions, the distribution of the bioregion surface area relative to the set of threshold distances is characterised hereafter by its central value (the median) within its interquartile range (Q1–Q3). The interquartile range of surface area estimates accounts for the bioregionalization method uncertainty. All surface areas are reported relative to the total global ocean area. To identify long-term changes, we fitted a linear model to the central surface area estimates in respect of time. The statistical significance of these linear trends was determined using a Spearman rank correlation, with a trend considered significant at $p < 0.1$.”

7. Treatment of uncertainty in area trends: The use of interquartile range and Spearman correlation is reasonable, but some aspects need elaboration: The area uncertainty of bioregions 6 and 7 is “ $\approx 50\%$ ”; this raises the question of whether area trends are meaningful at all for these regions. Clarify whether the trend significance accounts for serial correlation in time series or the uncertainty distribution across thresholds. If not, these limitations should be acknowledged.

We do agree with the reviewer's observation. The trend significance (p value) accounts for the linearity model between the central surface area (medium value of the bioregion's surface area distribution across threshold) in respect of years. We hope that this is now clarified with the rewriting of paragraph 3 in section 2.2 and of section 2.4 (see above point 6). High method uncertainty in identifying the surface area of bioregion 6 and 7 (wide distributions across threshold) do not directly hinder the detection of statistical trends, it rather question the representativeness of the bioregions at these latitudes and therefore indirectly question the meaning of the identified trends. The sentence has been rephrased accordingly.

Line 208: “While surface areas of bioregions 6 and 7 exhibit large changes from one year to the other, their surface area estimates are associated with large method uncertainties ($\sim 50\%$), precluding the

consideration of the trend significance.

8. Interpretation of phenological transitions can be strengthened by adding some environmental context: eg. show coincident trends in stratification indices, SST, or nutrients from independent datasets such as model simulations or literature. This strengthens the claim that phenological changes are climate-driven rather than methodological artifacts.

We do think that adding such quantitative estimates would be a relevant addition, but extend beyond the scope of this study focused on proofing a monitoring method that tracks significant changes of phytoplankton phenology. A contextual interpretation of the changes origin is addressed qualitatively in the first paragraph of the discussion based on existing literature. With this discussion, our intention is not to claim that the observed changes are surely climate driven. To dampen the assignment of observed changes toward climate change the word “likely” was changed to “possibly” at line 219.

Line 245: “The cascading changes (i.e., transitions from bioregion 4 to 3, bioregion 3 to 2, and 2 to 1) identified above, are possibly driven by an intensification of upper-ocean stratification in relation to climate warming (Capotondi et al., 2012)”.

We believe that all the transitioning regions we identified, should be investigated in search of the underlying mechanisms (environmental or methodological artifacts). We considered this option, however, the correlation would reveal itself as not trivial: several environmental data-sets should be considered as raised by the reviewer. Each of these data-sets have a specific spatio-temporal resolution introducing further numerical bias. Lastly, the choice of a comparison or correlation method would be complex. E.g a spatial comparison would require to define spatial averages of the environmental variables inside the already defined bioregions, or from independent unsupervised partition of the environmental variables.

A recent study, published in the OSR9 and focused on micronekton habitat provinces (mean temperature in the epipelagic layer, index of stratification and integrated net primary production, Alberhne et al., 2025), finds similar results further comforting the use of surface chlorophyll-a seasonality as a marker of environmental changes. This study is now explicitly cited in the discussion.

9. The “stable change” criterion could be better justified. You choose ≥ 6 years per decade to identify a stable phenology shift. This choice is reasonable, but: Why 6 years and not 5 or 7? Could year-to-year autocorrelation bias persistence estimates? A brief sensitivity analysis (even qualitative) would enhance confidence.

This threshold period needs to be strictly above five years (half of the decade) to ensure the sorting of pixels of equal assignment. The closer to ten years, the least permissive this criterion is. We decided to set it on six years which is the most permissive threshold allowing modulation related to inter annual climate variability. This qualitative sensitivity analysis has been added in the text.

Line 135: “This temporal criterion was chosen strictly above half a decade to grant a reasonable and robust indicator of the bioregion persistence while remaining permissive in respect of inter annual climate variability.”

10. Bioregion 1 includes oligotrophic/mesotrophic subdivisions. This is a meaningful addition however I suggest to describe the two sub-regions more clearly (Fig. 1b only briefly mentions a bimodality) and state whether the split was used purely for interpretation or also within the clustering procedure.

The split is used solely in the interpretation as the clustering is performed on normalized time series. This is now more clearly highlighted in the text (Line 143) and Line 186: “Its additional subdivision, introduced here, produces spatially coherent domains that distinguish between oligotrophic regions in the centers of subtropical gyres and mesotrophic regions primarily distributed around the equator.”

11. Limitations section is strong but could expand on satellite products bias as mentioned above in comments 1-3.

Following the reviewer comment, the section discussing method limitation now mentions possible bias

linked to merged satellite ocean color products (see response to comments 1-3).

Minor Comments

Language check throughout the manuscript is necessary, such as:

L8 “these evolution” → “these evolutions” [ok](#)

L9, Rephrase “monitoring indicator approach”. [ok](#)

L13 “bioregions’ extent appear” → “bioregions’ extents appear”.[ok](#)

L36 “limits growth” → “limit growth”. [ok](#)

L49-50 “Which are often used” → “which are often used”.[ok](#)

L108 “This temporal criteria” → “This temporal criterion”. [ok](#)

L61: “(1998–2007 and 2025–2014)” is likely a typo for 2015–2024. [ok](#)

L97-98: rephrase. [ok](#)

Fig 2b 2a and Fig 4a were first mentioned other than Fig.1? This is not commonly seen in articles. [ok](#)

Some citations have small inconsistencies (e.g., Flawosky à Falkowski?). [Ok](#)