Two-decade satellite monitoring of surface phytoplankton functional types in the Atlantic Ocean

Hongyan Xi[1*], Marine Bretagnon[2], Svetlana N. Losa[1,3], Vanda Brotas[4], Mara Gomes[4], Ilka Peeken[1], Antoine Mangin[2], Astrid Bracher[1,5]

[1]Alfred Wegener Institute, Helmholtz-Centre for Polar and Marine Research, Bremerhaven, 27570, Germany
[2]ACRI-ST, Sophia Antipolis Cedex, France
[3] Shirshov Institute of Oceanology, Russian Academy of Sciences, Moscow, Russia
[4]MARE/ARNET - Marine and Environmental Sciences Centre, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016, Lisboa, Portugal
[5]Institute of Environmental Physics, University of Bremen, Bremen, 28359, Germany

Correspondence to: Hongyan Xi (hongyan.xi@awi.de)

Author Comments in response to Referee #1

Review of "Two-decade satellite monitoring of surface phytoplankton functional types in the Atlantic Ocean" by Hongyan Xi et al.

General comment:

The paper submitted by Hongyan Xi et al. focuses on PFTs in the Atlantic Ocean as seen by satellite in the last two decades. The paper is well written and easy to be read and represents a useful analysis and report to understand what is the Atlantic Ocean ecosystem status.

Overall, I have some specific comments that are preferable be included in the revised version of the manuscript. Especially, some sections need to be extended with more information: for instance, some more details about the EOFs technique in the section of time-series analysis, or the definition of the statistical indexes used to evaluate the satellite products in respect to the in-situ dataset. Two key points are: extend the discussion of the results found in the validation test of satellite vs. in situ data, specifically for what concern the statistics; and the inclusion xof a significance test (e.g. p-value) to better interpret the trends found in the different regions.

However, on my opinion, the paper needs only minor corrections before the publication. This is a useful work for the ocean colour community from both scientific (the understanding of the PFTs changes in the Atlantic Ocean ecosystem) and technical (e.g. products validation, long-term time-series) perspectives.

We thank very much the reviewer for the positive feedback and constructive comments on this manuscript. We have considered carefully all the suggestions to improve the manuscript. Before we address each of the comments individually, we would like to preface with a brief response to the above general comment.

Firstly, we would like to clarify that the EOF technique was not used in the time series analysis but is used in the satellite PFT estimation approach that was proposed and detailed in Xi et al. (2020, 2021). The CMEMS global PFT products used in this study are generated based on the

EOF-PFT approach therefore we only mentioned it briefly while describing CMEMS PFT products in the manuscript. Secondly, the discussion on the validation results and statistics has been updated and extended. Lastly, more details about the time series analysis have been added. However, we would like to point out that this manuscript was submitted as a contribution to an upcoming Ocean State Report, for which specific requirements in paper length and number of tables/figures have to be followed. The current paper length is just at this limit; therefore, we extended/added the necessary discussion/information in the manuscript as concise as possible, but more details are provided in the individual responses below.

Specific comment:

Section 2.1

I would include a table with the wavelengths/bands of the different sensor here used and that have been controlled. I would add a Figure that pointed out the life cycle of each sensor and the overlap of the different mission. Such a information give to the reader a clear, easy details about the last 20 years of satellite sensors and mission.

We agree with the reviewer that such information would give clearer details on the past/current satellite missions. We did not include the detailed wavelengths/bands of different sensors because of the length limit for such a report and also because they are detailed in our previous work (Xi et al. 2020; 2021). The period of PFT products from each sensor or sensor combinations has also been provided both in section 2.1 and Table1 in the manuscript (no figure was added for brevity).

However, in response to the reviewer we have still prepared the table and the figure (see below), as the review reports and responses are eventually also publicly available online for all readers.

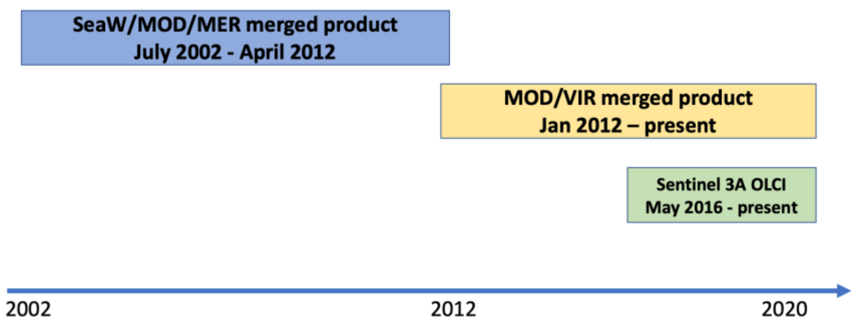

Figure R1. Lifespans of ocean colour sensor (combinations) where Rrs products were acquired.

Table R1: Wavebands of satellite sensor (combinations) involved in the PFT estimation approach.

Sensors involved	Center wavebands used in the EOF-PFT approach (nm)
SeaWiFS/MODIS/MERIS merged[a]	412 443 490 510 531 547 555 670 678
MODIS/VIIRS merged	412 443 490 531 547 551 555 670 678
Sentinel 3A OLCI	400 412 443 490 510 555[b] 560 620 665 674 681

[a] SeaWiFS terminated in December 2010, therefore from Jan 2011 to April 2012 only MODIS/MERIS merged data were available.
[b] There is no band at 555 nm for OLCI itself, but the GlobColour Team provides also the 555 nm band through an inter-spectral conversion from 560 nm (details see ACRI-ST GlobColour Team et al., 2017)

Section 2.3

Please extend the section with more details about the method you have applied (i.e. EOFs). Currently, it is not mentioned in the section about the time-series analysis.

The EOF-based approach is actually the theoretical basis of the CMEMS PFT products algorithm that were used in this study. Details about the algorithm are provided in Xi et al. (2020; 2021). Since this study focuses mainly on the PFT long term observations and time series extracted from the satellite operational PFT products but not on the EOF-based approach itself, we only cited the references but did not include the details in the manuscript.

Regarding the time series analysis, the reviewer was right that more information should be included. We have extended Section 2.3 by adding brief descriptions regarding the computations of the 20-year trend, phenology indicators and anomaly of 2021.

Text added in the revised manuscript (**section 2.3 Lines 140-147**): "We investigate the trends in the PFTs for the last 20 years using linear regression in the format of Y=SX+I, where Y is the monthly PFT Chla of either per-pixel or the regional log-based mean, X is the time on monthly basis, S is the slope of the regression and I is the intercept. Only trends with statistically significant correlations of the regression ($p<0.05$) are shown. Indicators of PFT phenology and the anomaly of 2021 (the last year of the considered time period) are also extracted in order to identify potential changes/shifts in PFTs. Abundance maxima time, as one of the phenology indicators, is identified for each pixel by finding the month when the maximum PFT Chla occurred during the year. Anomaly in percentage is determined by computing the relative difference between the PFT state of 2021 and the average state of the last two decades (i.e., climatology)."

Section 3.1

Lines 134: you mentioned that the same correction approach is then applied on Sentinel-3A OLCI derived PFTs. Please add some details about how it works for OLCI.

We have added more information about the correction on the OLCI derived PFTs.
Lines 157-161 "The same is applied to the Sentinel 3A OLCI derived PFTs by comparing them to the corrected MODIS/VIIRS derived PFTs for the overlapped period April – December 2016, so that all PFT data from both MODIS/VIIRS and OLCI are now referenced to SeaWiFS/MODIS/MERIS derived PFTs. Though R^2 is slightly weaker (R^2 between 0.77 and 0.83) compared to that from the MODIS/VIIRS versus SeaWiFS/MODIS/MERIS derived PFTs (R^2 between 0.82 and 0.98), OLCI derived PFTs still showed overall good correlations to the corrected MODIS/VIIRS data with regression slopes between 0.83 and 1.03 despite that prokaryote Chla retrievals from OLCI data are in general higher."

Figure R2 shows scatterplots and regression relationships between the two products. However, due to the paper length limit, we decided not to include the figure in the revised manuscript.

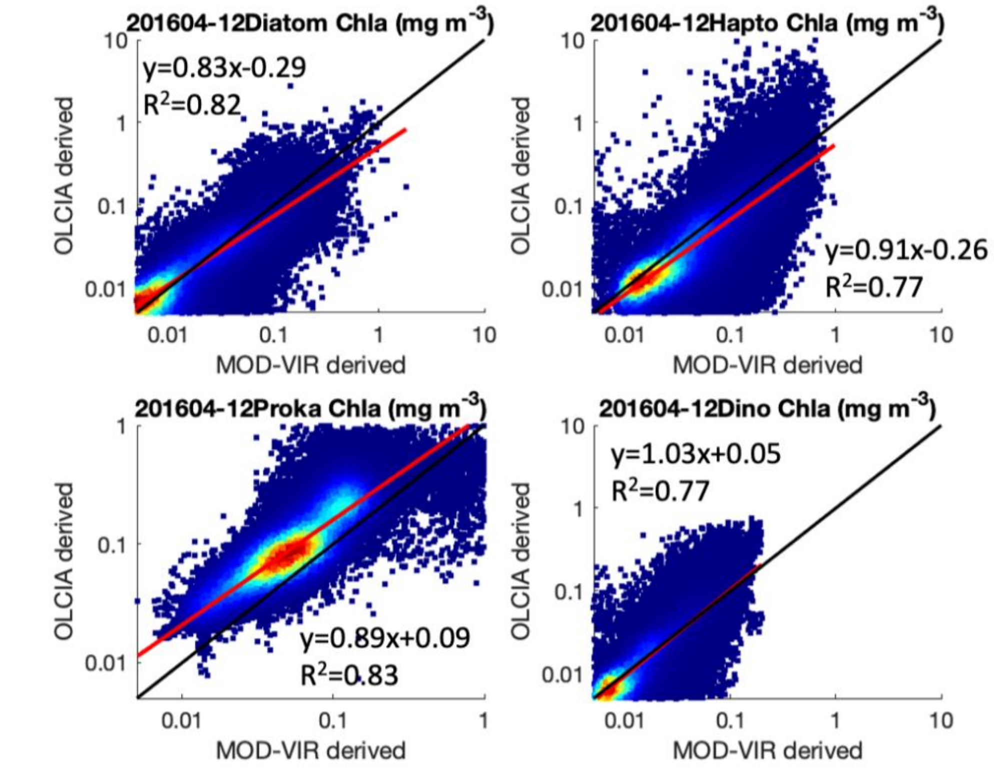

Figure R2: Scatterplots of monthly PFTs derived from OLCI Rrs and corrected MODIS/VIIRS merged Rrs data for the overlapping period April – December 2016. The 1:1 line is shown in black and the linear regression line (using type II regression with per-pixel uncertainty) in red. R^2, slopes and offsets determined in log-10 scale are also presented.

Lines 139: Please add some more details about the type of errors reported in the table 1. How is defined the MDPD? and the RMSD? What means a MDPD of 89.6%? It indicates a good agreement between satellite and in-situ data or vice versa. Which is the unit of median uncertainty, it is expressed in %.

We agree with the reviewer that the statistics should be described in more detail, therefore we have updated the caption of the statistics table (Table 2) with more details

Lines 530-535 "Table 2. Statistical validation results of satellite derived PFT Chla (after inter-mission correction) as a function of in situ PFT Chla using least square fit in logarithmic scale. N: number of matchups; R: Pearson correlation coefficient; MDPD: median percent difference; RMSD: root-mean-square difference; definition equations of these terms were referred to Xi et al. 2020. Note that Slope, Intercept and R were calculated based on logarithmic scale. Median uncertainties calculated based on satellite per-pixel PFT uncertainty (equivalent to relative error in %) are also shown in the last column."

Regarding the statistics indicating how good the agreement between the satellite and in-situ data is, please see our response below to the next relevant comment.

The authors need to better describe the results found here, since this is the first step behind the successive important and impactful time-series analysis: for instance, an R coefficient of 0.6 is equal to an R2 of 0.36 that implies not a strong correlation between in-situ and satellite-derived PFT. This is consistent or a better results in respect to the previous works and literature?

During the "under-review" stage we have obtained more in situ data for the validation also with a more thorough matchup extraction, therefore, we have been able to extend the matchup data set and update the statistics in Table 2. We have added Figure R3 in this response document to show the scatterplots between the inter-sensor corrected satellite PFTs versus the in situ PFT data. This figure is not included in the revised documents as Table 2 summarizes adequately the statistics.

The discussion regarding the validation has been also updated in Section 3.1 and the Discussion (Section 4).

Section 3.1 Lines 162-168: "Validation was carried out by comparing the collocated satellite PFTs with the in situ PFTs using the extracted matchup data. Statistical results of the validation in Table 2 show in general acceptable agreement between the in situ and satellite derived PFTs. Median percent differences (MDPD) are consistent with the median satellite PFT uncertainties (relative error in %) estimated through Monte Carlo simulation and error propagation in Xi et al. (2021), and for dinoflagellates, notably lower. Higher MDPD is found for prokaryotes due to a systematic overestimation of the picophytoplankton in the retrieval algorithms for all the three sets of satellite OC sensors, however, no significant bias of satellite prokaryote products is detected between different sensors, therefore the overestimation should have minor influence on the time series data of prokaryotes."

Section 4 Lines 279-287: "Validation using in situ data shows no significant biases of PFTs derived from different sensors, indicating that the inter-mission offset was effectively corrected. Chla of different PFTs are more upscaled retrievals compared to bulk satellite OC products such as total chlorophyll a, coloured dissolved organic matter (CDOM) and absorption properties. Especially, it is still challenging to retrieve accurately prokaryotic phytoplankton because in the open ocean these are dominating in the low Chla areas for which the satellite signals are weaker. Therefore higher uncertainties exist in these products (e.g., Brewin et al. 2017; Losa et al. 2017; Xi et al. 2021) as compared to uncertainties for other PFTs (see Table 2). In summary, our statistical results of PFT validation are comparable to the evaluations of satellite PFT products derived from different approaches, according to the Quality Information Documents (QUID) that have been published on CMEMS (Garnesson et al., 2022; Pardo et al., 2022)."

Table 2: Statistical validation results of satellite derived PFT Chla (after inter-mission correction) as a function of in situ PFT Chla using least square fit in logarithmic scale. N: number of matchups; R: Pearson correlation coefficient; MDPD: median percent difference; RMSD: root-mean-square difference; definition equations of these terms were referred to Xi et al. 2020. Note that Slope, Intercept and R were calculated based on logarithmic scale. Median uncertainties calculated based on satellite per-pixel PFT uncertainty (equivalent to relative error in %) are also shown in the last column.

	N	Slope	Intercept	R	MDPD (%)	RMSD (mg m^{-3})	Median satellite PFT uncertainty (%)
Diatoms	192	0.71	-0.27	0.87	60.5	0.30	57.3
Haptophytes	191	0.95	-0.007	0.64	58.9	0.18	41.5
Prokaryotes	187	0.71	0.12	0.60	185	0.06	86.5
Dinoflagellates	144	1.07	0.04	0.81	59.1	0.07	74.3

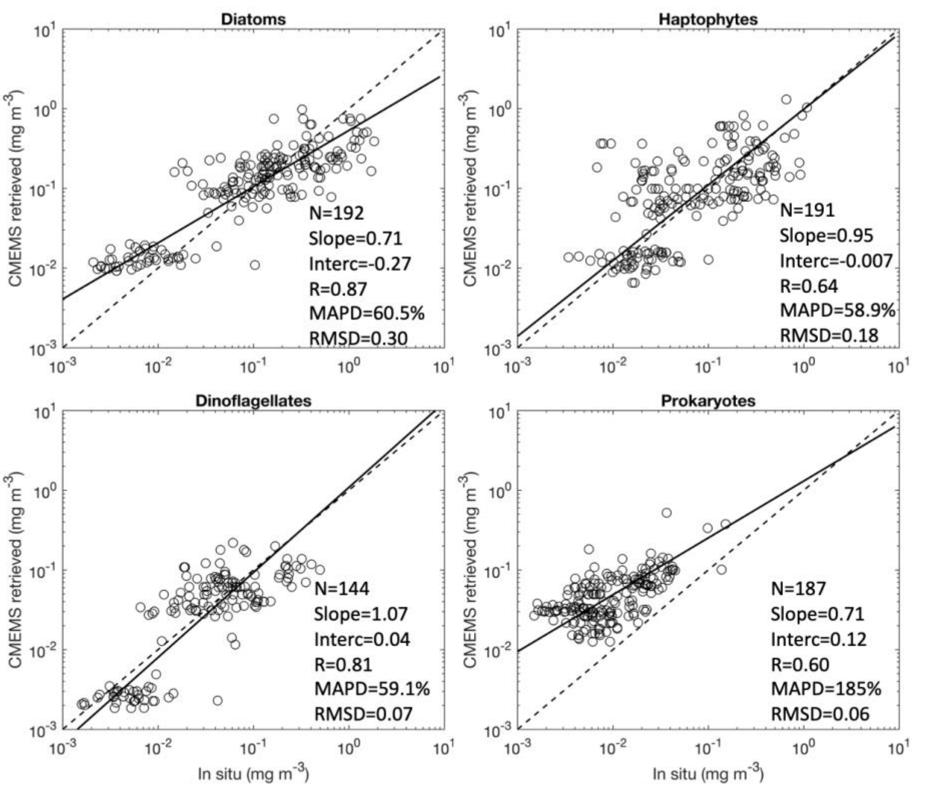

Figure R3: Scatterplots of matched satellite derived PFT Chla after inter-mission correction versus in situ PFT Chla including statistical results.

Section 3.3.

Discussion of trend (increasing/decreasing) needs to be also coupled to a statistical significance test as the p-value. This coefficient could give to the reader a tool to understand the inter-annual variability and if it has a statistical significance in the last 20 years.

We would like to clarify that only areas with significant trends ($p<0.05$) were shown in the per-pixel trend maps and were discussed in the manuscript. We have mentioned the significant level in section 2 and also in the figure captions. For the overall trends of the four PFTs shown in Figure 3b, significant trend was found for prokaryotes only and the slope and p-value were indicated.

For Figure 4 trendlines with slopes and correlation coefficients are now shown in the time series plots for provinces with significant trends ($p<0.05$) only. These trends in different provinces correspond well to the descriptions in the results (Section 3.3 Lines 218-229).

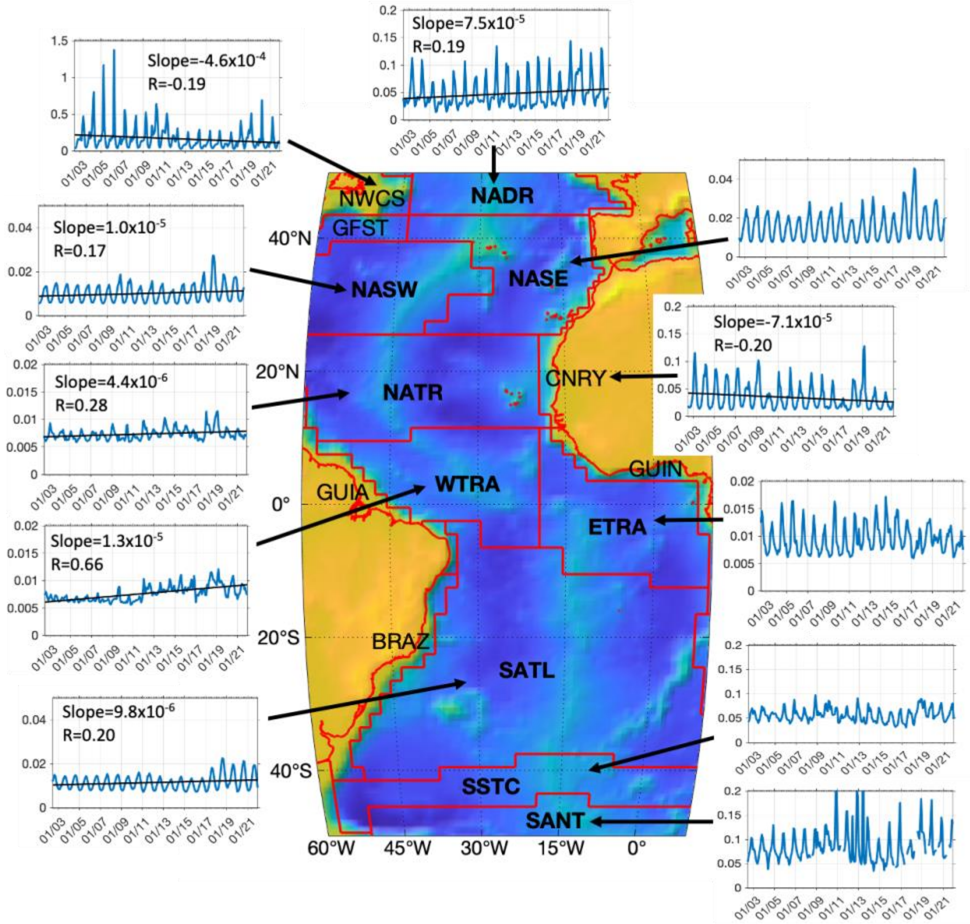

Figure 4: Time series of diatom Chla (unit: mg m^{-3}) in 11 Longhurst provinces in the Atlantic Ocean with bathymetric information based on ETOPO1 bathymetry (Amante & Eakins, 2009). Provinces according to Longhurst (2007) are: NADR for North Atlantic Drift Province, NWCS for Northwest Atlantic Shelves Province, NASW for North Atlantic Subtropical Gyral Province (West), NASE for North Atlantic Subtropical Gyral Province (East), NATR for North Atlantic Tropical Gyral Province, CNRY for Canary Current Coastal Province, WTRA for Western Tropical Atlantic Province, ETRA for Eastern Tropical Atlantic Province, SATL for South Atlantic Gyral Province, SSTC for South Subtropical Convergence Province, SANT for Subantarctic Water Ring Province, respectively. Trendlines with slopes (unit: mg m^{-3} month^{-1}) and correlation coefficients are shown for provinces with significant trends ($p<0.05$).

Section 3.4

How did you compute the abundance maxima time? How it is defined? Please add more details about the computation in the Section 2.3 (time-series analysis).

As monthly data were used in this study, the abundance maxima for the year of 2021 was identified by finding the month for each pixel when the maximal PFT Chla occurred during the year.

More details (but have also to be concise due to length limit) on time series analysis have been added in **section 2.3 Lines 140-147**: "We investigate the trends in the PFTs for the last 20 years using linear regression in the format of Y=SX+I, where Y is the monthly PFT Chla of either per-pixel or the regional log-based mean, X is the time on monthly basis, S is the slope of the regression and I is the intercept. Only trends with statistically significant correlations of the regression ($p<0.05$) are shown. Knowledge of PFT phenology and anomaly of 2021 (the last year of the considered time period) are also gained through to help identify potential changes/shifts in the PFTs. Abundance maxima time, as one of the phenology indicators, is identified for each pixel by finding the month when the maximal PFT Chla occurred during the year. Anomaly in percentage is determined by computing the relative difference between the PFT state of 2021 and the average state of the last two decades (i.e., climatology)."

Technical Corrections (some examples):

Lines 35-38: please rephrase the sentence.

The sentence was rephrased to "Climate induced changes causing rising temperatures, ocean acidification and ocean deoxygenation, stress the ocean's contemporary biogeochemical cycles and ecosystems, thereby impact the phytoplankton communities (Gruber, 2011; 2021; Bindoff et al. 2019)." **(Lines 36-38)**

Line 42: I'd remove "as a whole" à you can leave the sentence as: "phytoplankton biomass does not provide a full description of the complex nature of phytoplankton community and function".

Removed as suggested.

Line 44: please rephrase the sentence "Phytoplankton composition structure varies across ocean biomes and different phytoplankton groups influence marine ecosystem and biogeochemical processes differently (Bracher et al., 2017) à I would change in "Phytoplankton composition varies across ocean biomes and the different phytoplankton groups influence marine ecosystem and biogeochemical processes differently (Bracher et al., 2017)"

Revised as suggested.

Line 51: I would remove "as well", it is not necessary here.

Removed as suggested.

Lines 72-74: I do not understand to what referred such sentence, from "Previously" to "measured pigments".

This sentence referred to the previous studies of Xi et al. (2020; 2021) and the algorithms developed therein. We moved the citations right after this sentence for a better clarification.

Line 78: chlorophyll a can be modified in Chla since you have defined in the first part of the paragraph.

Removed as suggested.

Line 80: maybe a refuse of something, maybe you can change in in this work.

Thanks for the comment. By "refuse" we suppose the reviewer means data fusion? So far we would only focus on merging the PFT data sets from different sensors for a consistent long-term data set to enable trustworthy time series analysis of the phytoplankton groups on either global or regional scales.

References

ACRI-ST GlobColour Team, Mangin, A., & Hembise Fanton d'Andon, O. (2017). GlobColour product user Guide (GC-UM-ACR-PUG-01,version 4.1). ACRI-ST

Bindoff, N.L., Cheung, W.W.L., Kairo, J.G., Arístegui, J., Guinder, V.A., Hallberg, R., … & Williamson, P. (2019). Changing Ocean, Marine Ecosystems, and Dependent Communities. In: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA, pp. 447–587. doi:10.1017/9781009157964.007

Garnesson, P., & Bretagnon, M. (2022). QUID for OC TAC Products OCEANCOLOUR OBSERVATIONS GlobColour, Issue 2.0, Ref: CMEMS-OC-QUID-009-101to104-116-118, Date: 16 June 2022.

Gruber, N. (2011). Warming up, turning sour, losing breath: Ocean biogeochemistry under global change. Phil. Trans. R. Soc. A, 369, 1980-1996.

Gruber, N., Boyd, P.W., Frölicher T.L., & Vogt, M. (2021). Biogeochemical extremes and compound events in the ocean. Nature, 600, 395-407. doi:10.1038/s41586-021-03981-7

Losa, S.N., Soppa, M.A., Dinter, T., Wolanin, A., Brewin, R.J.W., Bricaud, A., … & Bracher, A. (2017). Synergistic Exploitation of Hyper- and Multi-Spectral Precursor Sentinel Measurements to Determine Phytoplankton Functional Types (SynSenPFT). Frontiers in Marine Science, 4, 1–22. doi:10.3389/fmars.2017.00203

Pardo, S., Jackson, T., Netting, J., Calton, B., & Howey, B. (2022). QUID for OC TAC Products Atlantic and Arctic Observation Products. Issue 2.0, Ref: CMEMS-OC-QUID-009-111 to114-121to124, Date: 16 June 2022.

Xi, H., Losa, S. N., Mangin, A., Garnesson, P., Bretagnon, M., Demaria, J., … & Bracher, A. (2021). Global chlorophyll a concentrations of phytoplankton functional types with detailed uncertainty assessment using multi-sensor ocean color and sea surface temperature satellite products. Journal of Geophysical Research: Oceans, 126, e2020JC017127. doi:10.1029/2020JC017127

Xi, H., Losa, S.N., Mangin, A., Soppa, M.A., Garnesson, P., Demaria, J., … & Bracher, A. (2020). A global retrieval algorithm of phytoplankton functional types: Towards the applications to CMEMS GlobColour merged products and OLCI data. Remote Sensing of Environment, 240, 111704. doi:10.1016/j.rse.2020.111704