

## **Review and Synthesis: operational prediction of ocean biogeochemistry in the Copernicus Marine Service. Main achievements and perspectives**

Dear Editor,

We thank the three reviewers for their detailed and constructive comments, which have led to a substantially revised manuscript. The reviewers converged on several points: the manuscript read too much like a technical report, its introduction and tables were overloaded, several acronyms and citations were missing or inaccurate, the SWOT framing was inappropriate for a scientific journal, and the future directions were not laid out consistently across the forecasting centres. We addressed each of these points as follows.

The introduction has been rewritten and shortened. Section 2 now presents the six models first and explains the modelling concepts in plain language, before describing what converged across the systems over the decade. The data assimilation section opens with a general definition, describes the algorithmic evolution of the operational systems, credits the assimilation software with the appropriate references and acknowledges the contribution of the SEAMLESS project. All tables have been redesigned: two remain in the main text, and the per-centre details have been moved to seven thematic supplementary tables, including a new table outlining the developments planned in each centre for 2025–2028. The SWOT section has been replaced by a discussion of achievements, limitations, prospects and risks, with an expanded treatment of artificial intelligence and of the algorithmic perspectives for data assimilation, and a Conclusions section has been added. The term "green ocean" has been removed and all acronyms are defined at first use and gathered in an appendix.

On two points we respectfully maintain our position and explain why in the responses: the review deliberately covers the first decade of the Service (2015–2024), with the subsequent developments summarised in a dedicated table (S7) and we consider that the Copernicus Marine Service collectively defines the state of the art of operational biogeochemical forecasting.

Our point-by-point responses to each reviewer is submitted and the revised version of our work takes into account all the comments raised by the three reviewers.

With kind regards, on behalf of all co-authors,

Marilaure Grégoire

### **Reviewer 2 (in bold our answer to the comments of reviewer 2)**

Review of "Review and Synthesis: operational prediction of ocean biogeochemistry in the Copernicus Marine Service. Main achievements and perspectives." By Grégoire et al.

This paper is a review and synthesis paper looking into the Copernicus Marine Environment Monitoring Service (CMEMS) operational biogeochemical ocean forecasting systems, covering the period 2015–2024 across seven regional forecasting centres spanning European seas, the Arctic, and the global ocean. It documents how the "green ocean" prediction systems have evolved in terms of model formulations, data assimilation, spatial resolution, and delivered products and demonstrates that overall forecast quality has improved over the decade. Using a Strengths-Weaknesses-

Opportunities-Threats (SWOT) framework, the paper concludes that while significant progress has been made, persistent challenges remain around the scarcity of biogeochemical observations, the quality of river forcing, and the computational demands of future ensemble and high-resolution systems, with artificial intelligence identified as a potential opportunity that is currently under testing but not yet operational. The detailed description of the CMEMS "green ocean" service and the changes that have occurred over the past decade (2015–2024) of operation represents a genuinely valuable community resource. As mentioned in the paper, this resource could be particularly useful to provide knowledge to a wide range of users, to policy-makers and to the general public.

However, also coming from the same end arise my biggest concerns about the paper. This paper reads too technical, the writing style of this paper is more like a project report to a funding agency rather than a review and synthesis trying to reach the broader audience, which is supposed to be the biggest strength of the paper. The current version of the review will for sure lose the broader audience, additional to the list of technical details, it contains too many unnecessary details that eventually bury the truly valuable information.

**Answer:** We have revised the manuscript to make it accessible for a broad audience but still relevant for a more technical reader. Concretely: the introduction has been rewritten and shortened, and now defines the two product lines (near-real-time and multi-year) for readers unfamiliar with the Service; Sect. 2.1 was restructured to present the six models first and then explain the modelling concepts, compartments, plankton functional types, stoichiometry, in plain language in order to answer a criticism made by Reviewer 1; the data assimilation section now opens with a short general definition of data assimilation and then with more details on each specific system with figures to highlight the effect of data assimilation. A conclusions section has been added. Service-specific jargon and single-use acronyms have been removed, the remaining acronyms are defined at first use and gathered in Appendix A, and the term "green ocean" has been replaced throughout by "ocean biogeochemistry". The detailed per-centre technical information has been moved to the supplementary material (Tables S1–S7), leaving only two tables in the main text, both rebuilt for readability. Finally, the SWOT framing, that has been perceived by Reviewer 3 as a management concept and that may contribute to the "project report" perception, has been replaced by a discussion of achievements, limitations and perspectives (Sect. 5).

There are way too many abbreviations in the paper, which are definitely not reader friendly. Some of the important abbreviations are even used without explanations, e.g., SWOT in the abstract and later in the paper; NRT and MY, throughout the paper, etc.

**Answer:** All abbreviations have been reviewed as follows. Acronyms used only once or twice are spelled out and removed, and the remaining ones are defined at first use and gathered in a list at the end of the manuscript (Appendix A). The two examples raised by the reviewer are treated specifically: NRT and MY are now properly defined in the introduction, where a new paragraph describes the two product lines; SWOT no longer appears at all, the framing has been abandoned (see also our response to the general comment), and Sect. 5 is now a discussion of achievements, limitations and perspectives.

Unnecessary details in the paper, for example, Section 2.7 Computing resources. Number of cores, storage in TB, simulation times are operational logistics, not relevant to the general audience.

**Answer:** The former Sect. 2.7 has been merged with the forcings into a short Sect. 2.5 and all details about cores, storage, simulation times, have been moved to the supplementary material (Tables S3 and S6). We prefer to keep one aggregate statement, about the evolution of computing resources as it is a critical parameter that may limit system evolution for the next decades. It quantifies the rapidly growing computational cost of operational biogeochemistry, on which the discussion (in section 5) of the risks for future evolutions (ensembles, resolution, AI-based systems) is based.

The writing style and level of details are not consistent. For example, in tables A1 and A2, each row corresponds to one centre and reads as a self-contained technical description. The formatting, level of detail, and writing style vary noticeably between centres.

**Answer:** The former Tables A1 and A2 with the heterogeneity the reviewer rightly points out have been removed. Their content has been redistributed into thematic tables (products, resolution and coupling, data assimilation, observations, computing resources; Tables S1–S6). By construction, this format enforces a homogeneous level of detail and style across centres and tables. The same comment was made by Reviewer 3, and the two revisions are aligned.

The general text reads more like a list of enormous amount of information. This could be useful as technical documentation, but it is very difficult to extract scientific insight from, especially given the need to jump between pages in the main text and the appendix tables.

**Answer:** All the technical detail given across the 7 MFCs have been moved to the supplementary material, and each section of the main text now leads with its take-home message rather than with an enumeration: Sect. 2 states what converged across the systems over the decade (hydrodynamic models reduced from five to two, carbonate modules everywhere, biogeochemical assimilation extended from two to five systems) before pointing to the tables; Sect. 2.4 summarizes the evolution of the product catalogue (from a nutrients–chlorophyll–oxygen baseline in 2015 to carbonate, zooplankton and radiative variables in 2024); Sect. 3 opens with the overall result (a general improvement of product quality, larger in the reanalyses) before the detailed statistics. The main text can now be read on its own; the supplementary tables serve the reader who wants the per-centre specifics.

Content-wise, I'm hoping to get some information about inter-model consistency and disagreement across the seven systems. Considering the review is about an overview and updates on the 7 models, it's natural to include what aspects the models behave the same and what aspects are different and ideally also why and let downstream users know which products are most reliable for which applications.

**Answer:** we have added in Sect. 5.1: *"Across the MFCs, Chl-a and dissolved oxygen are the most extensively validated products (against satellite ocean colour and BGC-Argo) and those whose quality most clearly improved over 2015–2024: they have the highest readiness level for applications. Inorganic nutrients come next: delivered by all systems, but validated against sparse and uneven in-situ data, with contrasted skill between variables (the RMSD decreased for nitrate but not consistently for phosphate; Sect. 3). Carbonate variables, introduced in most catalogues after 2015, are promising for acidification monitoring but their validation record is still short. Biological products (zooplankton, PFTs) remain the least constrained, and iron, phytoplankton and zooplankton biomass are essentially unvalidated for lack of observations. A systematic*

*intercomparison of the systems is only possible where their domains overlap, i.e. between the global system and each regional one and would be a valuable coordinated exercise for the coming years. Whatever the variable, fitness-for-purpose for a given application (climate-change detection, hypoxia occurrence, seasonal-cycle studies) needs to be assessed case by case using the quality documentation of each product”.*

I do like section 4 and section 5, which are what I originally expected from this review. I would expect the AI part to particularly attract the public attention and would deserve more text than a short paragraph now. I think this part can be better received by the broader audience, potential users and policy makers. However, due to the issue above, it would be fairly difficult for a non-specialist reader to make it this far.

**Answer:** We agree with the reviewer. We have substantially expanded the AI paragraph (Sect. 5) highlighting the potential of AI emulators as already demonstrated for the atmosphere and physics, the expected gain of time for the delivery of what-if-scenario, and how data-driven techniques are currently embedded in our MFCs. The paragraph reads as follows: *“AI has the potential to profoundly change how the biogeochemical ocean is routinely predicted. In 2025, ECMWF launched its AI Forecasting System for the atmosphere, and Mercator Ocean International its neural-network global ocean forecasting system for the physics (GLONET; El Aouni et al., 2025). These systems are emulators: neural networks trained on decades of (re)analyses that, once trained, produce, in inference mode, a forecast in seconds on a single GPU, several orders of magnitude faster than the physics-based systems. This speed changes what is affordable: large forecast ensembles, systematic “what-if scenarios”, and interactive services become tractable, directly addressing the computing-resource limitations discussed above. For biogeochemistry, machine learning is in fact already embedded in the operational chain, neural networks reconstruct unobserved nitrate profiles that are assimilated in the Mediterranean system (PPCon; Sect. 2.3), and automated calibration of model parameters against BGC-Argo data is under testing. Full AI-based biogeochemical forecasting is currently under development in several MFCs. For example, Mercator Ocean International has recently developed BG4Sea, the first global data-driven system dedicated to marine biogeochemistry (Martinez Balbontin et al., 2026). Trained and evaluated using GLO-MFC MY product, BG4Sea provides seasonal forecasts of the 24 state variables of PISCES (nutrients, oxygen, biological, and carbon variables), at 1/4° spatial resolution. Such systems face specific challenges: the training record is shorter and sparser than for physics, biogeochemical variables are strongly non-Gaussian, and an emulator trained on a model inherits that model's biases (which may be larger for biogeochemical models). Hybrid approaches, combining physics-based models with machine-learning components, are therefore the most likely near-term pathway”.*

Due to the above-mentioned issues. I don't think the paper should be published in its current form. I would suggest the authors to choose a clear direction, either developing it as a technical paper or a broader-audience-reaching review and synthesis paper. If the authors choose the technical-paper route, the paper should include more comprehensive model comparisons, and the appendix tables should be moved into the main text. However, in this case, I wonder whether Biogeosciences is the best choice, as the strong technical emphasis of the paper may not fit the typical scope of the journal. If the authors decided to go with more of a review and synthesis paper, I would recommend removing a substantial portion of the technical details in Section 2 or relocating them to the appendix. The

authors should focus on synthesizing the insights from the observed changes, including their potential causes and consequences, rather than describing each change individually on a model-by-model basis. At the same time, Sections 4 and 5 should be expanded and given greater emphasis to better engage a broader audience.

As the manuscript suffers from significant structural issues, I therefore focus my review on general comments. Detailed line-by-line comments are not provided at this stage, as substantial reorganising of the manuscript would likely be required before such comments would be useful.

**Answer:** The manuscript has been significantly updated considering the three main lines of comments of the reviewer. More specifically, the changes can be overviewed as follows: 1) the technical detail of Sect. 2 has been moved to the supplementary material (Tables S1–S7), leaving two tables in the main text; Sect. 2 now reads as a synthesis: each subsection states what changed across the systems and why and convergence achieved over the decade (hydrodynamic models reduced from five to two, carbonate modules in all seven systems, biogeochemical assimilation extended from two to five systems); 2) Sect. 5 has been substantially expanded: the artificial-intelligence paragraph now explains what AI-based forecasting is and why it matters for the service and its users; the algorithmic perspectives for data assimilation are discussed; a consistent per-MFC outlook of the planned developments has been added (Table S7); and Section 5.1 includes a new user-guidance paragraph ranking the product classes by validation maturity; 3) the accessibility work described in our answer to the general comment, rewritten introduction, definitions of the product lines, removal of Service jargon and of most acronyms, is intended to carry the non-specialist reader through to Sects. 4 and 5.