



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

5 Marilaure Grégoire¹, Gianpiero Cossarini², Corinne Derval³, Elodie Gutknecht³, Susan Kay⁴, Julien Lamouroux³, Helen Morrison⁵, Coralie Perruche³, Annette Samuelsen⁶, Lena Spruch⁵, Anna Teruzzi², Luc Vandembulcke¹, Karina Von Schuckmann³, Tsuyoshi Wakamatsu⁶.

¹Department of Astrophysics, Geophysics and Oceanography, MAST-FOCUS research group, University of Liège, Belgium.

10 ² National Institute of Oceanography and Applied Geophysics - OGS, Italy

³ Mercator Ocean International

⁴ Met Office, UK.

⁵ Federal Maritime and Hydrographic Agency of Germany (BSH)

⁶ Nansen Environmental and Remote Sensing Center, Norway.

15

Correspondence to: Marilaure Grégoire (mgregoire@uliege.be)

Abstract. The operational prediction of marine biogeochemical cycles and ecosystems (the green ocean) has made significant progresses during the last decade. The green ocean is now routinely forecasted every day, and multi-decadal reanalyses are produced with an always increasing resolution and over longer periods. The quality of the green ocean products has increased thanks to the improved model formulations, resolution, data assimilation systems and increased availability of biogeochemical observations routinely delivered in near real time.

20

Here we review the advancements in our capabilities to predict the green ocean in the frame of the European Copernicus Marine Service (CMEMS) since its start in 2015 and for the five European seas, the Arctic and Global oceans. The evolutions of the prediction systems (e.g. model formulations, data assimilation, coupling with the physics and at the interfaces), delivered products (e.g. resolution, quality assessment, adequacy to support the development of indicators and the decision-making process), and computing resources requirements are reviewed. The predictability drivers and relevant time scales for marine biogeochemical cycles and ecosystems predictions are discussed. Recommendations for future developments are proposed based on a SWOT analysis of current CMEMS green ocean prediction systems and products.

25

30



1 Introduction

Operational oceanography combines model predictions and observations to deliver routine comprehensive information on the state of the ocean physics, biogeochemistry and sea-ice. This information is used for knowledge, decision making, downstream applications and societal well-being. It involves short-term forecasts (i.e. a few days), analysis and (multidecadal) reanalysis for the open, regional and coastal ocean. The 90's marks a paradigm shift in our capabilities of observing and forecasting the ocean. In this decade, remote and in-situ observing systems had developed sufficiently to make global, real-time observation possible. The ocean surface is observed from space at an unprecedented accuracy and resolution with, for instance, the launch of the Topex/Poseidon altimeter mission (Le Traon and Dibarboure, 1999). On the other hand, the increase in computing resources allows the development of models at global scale, solving processes at increasing resolution and assimilating data. These conditions create the ideal framework for the emergence of the international Global Ocean Data Assimilation Experiment (GODAE, 1998-2008, Bell et al., 2009) concept at the base of modern oceanography and inspired by the Global Research Atmospheric Program (Smith 2001, IGST, 2000). GODAE was a 10-year international initiative whose objectives were the integration of state-of-the-art ocean circulation models and (near real time) observations via the development of data assimilation to routinely forecast and provide global analyses of the ocean temperature, salinity and velocity components. GODAE demonstrated the feasibility and value of a long-term global ocean observing system supporting real-time ocean forecasting and (re)analysis. GODAE and CLIVAR supported the deployment of a global array of autonomous profiling floats, the Argos, that monitor the ocean temperature and salinity for climate and real time oceanography (Roemmich et al., 2019). The ARGO program was launched, complementing the data of the WOCE project by delivering in (near) real-time ((N)RT), temperature and salinity data. GODAE ended in 2008 and has allowed to achieve significant progress in our forecasting capabilities of ocean physics that has now reached a level of maturity comparable to that of the atmosphere.

The EU-funded MERSEA (Marine Environment and Security for the European Area, 2004-2008) project extended the forecasting of the physics to the biogeochemistry. MERSEA developed forecasting systems for the global and regional oceans, from days to months. During MERSEA, a prototype of a coupled physical-biological assimilation system based on the PISCES model (Aumont et al., 2003) were developed to routinely forecast biogeochemical variables over the globe with weekly delivery of biogeochemical products (Brasseur et al., 2009). MERSEA was followed by the EU MyOcean project (2009-2012) that aimed at setting up infrastructures and services in preparation for the Global Monitoring for Environmental Security (GMES) Marine Services.

GMES ended in 2013 and was followed by Copernicus, the Earth observation component of the European Union. Since its start, Copernicus offers information services streamlined through six thematic streams and derived from space and in-situ observation data. The Copernicus Marine Service (CMEMS) provides comprehensive monitoring and forecasting capabilities for the physics (blue ocean), biogeochemistry (green ocean) and sea ice (white ocean) for the five European seas, the Arctic and Global oceans. Seven Monitoring and Forecasting Centres (MFCs) integrate state of the art numerical ocean models with



65 the data provided by Thematic Assembly Centers (TACs) to provide daily forecasts and multi-decadal reanalyses for the blue, green and white ocean with an always increasing resolution (i.e. km scale) over longer periods (i.e. several decades).

GODAE transitioned to GODAE OceanView (GOV) to further mature the forecasting systems including new data streams and assimilation methods, sea-ice and biogeochemical modelling, and routinely providing an enhanced set of products of increased quality to support new user applications. In 2019 GOV was renamed OceanPredict, an international research and development network to accelerate, strengthen and increase the impact of ocean prediction (<https://oceanpredict.org/>). OceanPredict operates 70 six task teams, among which the Marine Ecosystem Analysis and Prediction Task Team (MEAP-TT). MEAP-TT works towards the improvement of biogeochemical forecasting systems in terms of model, data integration via data assimilation methods and products delivery to support user needs. OceanPredict initiates several programs and projects focused on ocean prediction endorsed by the UN Ocean decade.

75 Biogeochemical and biological ocean variables are more difficult to observe routinely using sensors on board satellite and in-situ autonomous platforms. The situation is particularly critical for biological data that still require field sampling and laboratory analyses. On the other hand, while ocean circulation models are based on a set of well-established differential equations that describe the conservation of mass, movement quantity, thermal energy, turbulent kinetic energy, the structure, formulation and parametrization of ocean biogeochemical models are more empirical and suffer from the lack of 80 biogeochemical observation for their calibration and validation. These last two decades, the deployment of biogeochemical Argo floats and the development of radiometric sensors with a refined spectral resolution have increased the delivery of NRT biogeochemical observations for assimilation, calibration and validation purposes and have opened the door to routine biogeochemical forecasting. After 10 years of CMEMS, driven by a common vision for service evolution towards users' needs, the CMEMS MFCs provide biogeochemical products that can be used for deoxygenation, acidification, and carbon 85 sequestration assessment.

This paper overviews the status of knowledge of the operational forecasting of the green ocean and reviews the progresses achieved in terms of model formulations, resolution, data assimilation systems, coupling with the physics, delivered products, quality assessment, service to the users. Based on that, a SWOT analysis of the biogeochemical forecasting system is established and future directions of development are outlined.

90 **2. The CMEMS biogeochemical prediction systems**

In this section, we overview the evolution of the CMEMS biogeochemical prediction systems over 2015-2024 including model structure and formulation, delivered products, resolution, coupling with the physics, data assimilation system. Details are listed in Tables A1 and A2 in Annex.



2.1 Model structure

95 2.1.1. Description

All the CMEMS models are state of the art biogeochemical models. They describe the cycling of essential elements, such as carbon (C), nitrogen (N), oxygen (O₂), phosphorus (P), and silicon (Si), within the lower trophic levels of pelagic marine ecosystems. They are based on the same conceptual approach that divides the ocean materials into the inorganic and organic compartments. The organic compartment is further divided into the dead and living part. Each compartment exchanges materials and energy between them and with their environment via biogeochemical processes and trophic interactions like photosynthesis, predation, mortality, excretion, respiration. All these processes are mathematically formulated. These models are NPZD-type models encompassing a wide range of complexities. The models differ by the levels of details in the description of the inorganic and organic (living and dead) components and interactions. ERSEM and BFM are the most complex with more than 50 state variables while the other are intermediate complexity models with ~25 state variables (Table 1).

105 CMEMS biogeochemical models are applied to a large range of marine environments extending from the global to the regional ocean, from the deep to the shallow ocean, from the open to the semi-enclosed seas, from the equatorial, subtropical to the polar ocean. Although CMEMS models share some commonalities, their conceptualization, formulation and parameterization are tailored to the specific environmental conditions, ecosystem characteristics and data availability of their regions of application. They are fine-tuned with regional observation and regularly evolve to incorporate the new knowledge revealed by field data and experiments. Additional processes and variables are incorporated to consider for instance, anaerobic degradation processes in anoxic conditions, nitrogen fixation and fixers in low N:P ratios, iron limitations in the high-nutrient, low-chlorophyll regions, gelatinous plankton in eutrophicated areas, benthic-pelagic coupling in shallow systems, and sea-ice biogeochemistry in ice-covered systems.

The six CMEMS models include ECOSMO (Daewel and Schrum, 2013; Yumruktepe et al. 2022) for the Arctic Ocean (ARC MFC), ERGOM (Neumann, 2000) for the Baltic Sea (BAL MFC), BAMHBI (Grégoire et al., 2026) for the Black Sea (BLK MFC), PISCES (Aumont et al., 2015) for the Global Ocean (GLO MFC) and the Atlantic Iberia-Biscay-Ireland area (IBI MFC), BFM (Vichi et al., 2015) for the Mediterranean Sea (MED MFC), and ERSEM (Butenschön et al., 2016) for the Atlantic North-Western European Shelf (NWS MFC) (Figure 1). Over the past decade, two MFCs changed their model, for a more suitable model or a change of partner. The ARC-MFC changed from NORWECOM to ECOSMO, ECOSMO has been more



120 developed with respect to the lower trophic level dynamics, which is the present focus of the forecast and reanalysis. The
BAL-MFC changed from SCOBI to ERGOM.

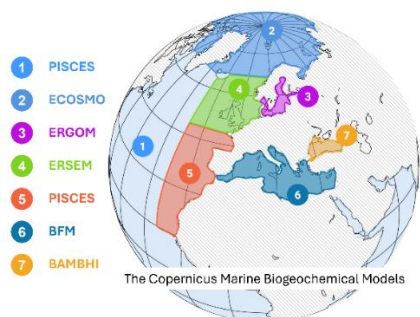


Figure 1: Biogeochemical models used in the CMEMS MFCs in 2024 in the seven MFCs: ARC, BAL, BLK, GLO, IBI, MED, NWS for producing NRT and MY products.

125 In each of these models, the plankton diversity is divided into groups (usually from two up to four) sharing common functional
characteristics (e.g. silicifiers, calcifiers, nitrogen fixers), the so-called Plankton Functional Types (PFTs). Each PFT
aggregates a large number of species into one compartment. For example, phytoplankton can be divided into diatoms, forming
silicate shells, and non-diatom groups. A two-phytoplankton group subdivision is typically used to simulate bloom conditions,
significantly contributing to primary production and carbon export versus, oligotrophic conditions. While the six models
130 explicitly simulate the diatom group, several approaches have emerged to include other groups of phytoplankton. They are
categorized either by size (nanophytoplankton, picophytoplankton), by ecological function (coccolithophores, cyanobacteria,
flagellates), by affinity to light spectra or combination of these factors (Table 1). Zooplankton is commonly separated into two
(microzooplankton and mesozooplankton) or more size classes (e.g. including nano heterotrophic plankton) and grazing
behavior (e.g. herbivorous versus carnivorous). The higher trophic levels (e.g. micronekton) are not explicitly resolved;
135 instead, their effects are parameterized as loss terms of the zooplankton that represent predation.

The CMEMS models differ in their treatment of elemental stoichiometry. PFTs are generally defined in terms of concentration
of one primary element (often nitrogen), with the other elements (e.g., carbon, phosphorus, silicon, or iron) derived using fixed
or variable stoichiometric ratios. The chlorophyll-a (chl-*a*) content of phytoplankton, critical for photosynthesis, is also
modelled as either a fixed or variable ratio, the latter allowing for acclimation to variable light conditions. Models adopt a
140 spectrum of approaches, ranging from fully fixed stoichiometric ratios (e.g., C:N, C:P, C:Si, C:Chl_a) to fully variable ones
(Table 1). Fully fixed ratios (used in ERGOM model) simplify the model formulations and reduce computational costs,



whereas fully variable ratios (in BFM and ERSEM) increase complexity, requiring additional state variables for each PFT but allow to represent multi-nutrient limitation and the uncoupling of element cycles. A compromised approach is also employed by some models, with autotrophs assigned variable ratios and heterotrophs fixed ratios. BAMHBI uses variable C:N and fixed N:P for autotrophs, and fixed C:N:P for heterotrophs. PISCES uses fixed Redfield ratios for C:N:P, while Fe:C and Si:C are predicted prognostically. ECOSMO, PISCES, BAMHBI, ERSEM, BFM use a variable C:Chl-ratio.

All CMEMS models include four essential macronutrients for phytoplankton growth: nitrate, ammonium, phosphate, and silicate, while PISCES and ERSEM also consider iron as a limiting nutrient. Oxygen dynamics is always included, given its critical role in biochemical and ecological processes. Oxygen concentrations regulate the transition between aerobic and anaerobic pathways, in which microbial communities break down organic matter into dissolved inorganic nutrients. Three out of six models (ECOSMO, ERSEM, PISCES) do not explicitly represent the bacterial biomass as a state variable but instead parameterize bacterial activity and its impact on the ecosystem. ERSEM, BAMHBI and BFM include explicit bacteria (Table 1). All models represent the recycling processes including at least one particulate and one dissolved non-living organic matter pool. More complex representations of the organic matter dynamics are included in ERSEM, BFM, BAMHBI and PISCES with different levels of dissolved pool lability, different rates of particle sinking, different affinity for nutrient utilization and different pathways for nutrient remineralization (e.g. denitrification, anaerobic ammonium oxidation). Improvements in the representation of organic particle dynamics, including vertical sinking velocities and remineralization rates, have further enhanced the accuracy of biogeochemical models.

For a comprehensive review of marine biogeochemical models, including their characteristics, operational predictive capabilities, applications and associated challenges, please refer to Gehlen et al (2015), Ford et al (2018), Fennel et al (2019), Gutknecht et al (2022), and Fennel et al (2022).

2.1.2 Evolution of CMEMS biogeochemical models over 2015-2024

Over 2015-2024, the structure and formulation of the CMEMS biogeochemical models have significantly evolved. Examples of major developments in the plankton description are the inclusion of dynamic C:Chl ratio in ECOSMO and BAMHBI, now considered in all CMEMS models. The addition of this feature increases the model complexity (additional state variables, including the prognostic chl-*a* content for each PFT) and allows to account for the physiological acclimation of phytoplankton to changing environmental conditions. ECOSMO introduces a new PFT (coccolithophores) and its associated chl-*a* dynamics, enabling better representation of ecological diversity, community structure, nutrient cycling, and carbon export. Another



170 significant evolution over the last 10 years is the explicit modelling, testing and operational production of the carbonate chemistry, necessary to simulate pH and the air-sea exchanges of CO₂.

Biogeochemical models describe the in-water photosynthetically active radiation (PAR; 400–700 nm), the portion of the visible light spectrum that phytoplankton requires for photosynthesis. To obtain the complete radiance distribution in time, spatial and spectral space requires solving a full radiative transfer model which is computationally expensive and will be limiting when coupled with an ecosystem model for long simulation periods. Given the efficiency requirements in an operational context and the current lack of information for certain processes, CMEMS models solve a simplified version of the radiative transfer equation assuming that the interaction between the light field and optically active particles does not depend on the direction of travel of the light. At the start of the service, the description of the in-water irradiance was mostly based on a simple analytical formulation involving an exponential decrease with depth of the scalar spectrally averaged irradiance, ignoring or roughly simplifying the spectral composition of light (i.e. 2-4 spectral bands are usually considered). The differentiation between the diffuse and direct parts was ignored and only a limited number of optically active components are considered (most of the time water molecules, chl-*a* and dead particulate organic matter). Such analytical models are computationally fast and are still used in PISCES, ERGOM, ECOSMO, and ERSEM. Some models (BAMHBI, BFM) have recently adopted spectrally resolved radiative transfer schemes (e.g. Terzić et al., 2021; Macé et al., 2025) to capture the complexity of light absorption and (back) scattering. These models include additional optically active components, such as coloured dissolved organic matter (CDOM) and detrital particles, which improves the realism of light dynamics and pioneer the possibility of assimilating direct light measurements (e.g. satellite reflectance).

The representation of benthic-pelagic interactions has been improved through either refined parameterizations (PISCES) or the (semi-)explicit modelling of benthic process (ERGOM, ERSEM, ECOSMO, BAMHBI, BFM). This improvement is a necessary requirement for the modelling of shallow systems where the sediment dynamics and benthic ecosystem significantly impact water column processes and quality.

2.2. Model products

At the start of the CMEMS, all the MFCs delivered chl-*a*, phytoplankton biomass in carbon, primary production, nitrate, phosphate and dissolved oxygen as NRT and MY products. These products constitute the baseline of the service and were delivered for the Global, Arctic ocean and the five European regional seas. A limited number of centres already delivered silicate, iron, and variables characterizing the light penetration (e.g. Secchi depth, euphotic depth, extinction coefficient). After 10 years of CMEMS, the NRT and MY baseline products have been extended with carbonate variables and more MFCs produce zooplankton, inorganic nutrients and light related variables. Most of the CMEMS biogeochemical models routinely provide maps of pCO₂, DIC and pH. Some are delivering the air-sea fluxes of CO₂ and alkalinity to support the estimation of

<https://doi.org/10.5194/egusphere-2026-813>

Preprint. Discussion started: 21 April 2026

© Author(s) 2026. CC BY 4.0 License.



the carbon budget. Information on PFTs, H₂S, POC are delivered for individual regions (Table A3). All these products are CF
200 compliant.



Table 1: Description of the biogeochemical models: In black: state of the system in 2015. In red: state of the system in 2024

Model	MFC	n. of state variables	Inorganic Matter	PFTs stoichiometry	Light Spectral bands	Zooplankton groups stoichiometry	Explicit Bacteria	Non-living particulate and dissolved organic pools (resp. POM, POM, DOM).	Particular Processes	Benthic
PISCES ¹	GLO, IBI	24	NO ₃ , PO ₄ , NH ₄ , SiO ₂ , Fe+ + Carbonate system (Ph, DIC, TA, pCO ₂ , CaCO ₃). O ₂ , CO ₂	Diatoms Nanophyto. Fixed : C:N:P Variable C:Si:Fe, C :Chla	3 bands (red, green, blue)	Micro. Meso. Fixed	no	2 POM (size classes; different sinking rates) 1 DOM small POM: variable C:Fe large POM: variable C:Fe:Si and CaCO ₃	Denitrif. Nitrogen fixation	No sediment module Oxic degradation or denitrification in the upper sediments is parameterized Permanent burial function of bottom shear
ERGOM ²	BAL	24	NO ₃ , NH ₄ , PO ₄ , SiO ₂ , Fe + Carbonate system (Ph, DIC, TA, pCO ₂). O ₂ , CO ₂ , N ₂	Diatoms Falgellates, Cyanobacteria Fixed	1 band	Micro. Meso. Zooplankton Fixed	no	1 POM 1 DOM CDOM	Denitrif. Sulfate reduction	2D Detritus pools for N and Si and Fe(III)-PO ₄
BAMHBI ³	BLK	30	NO ₃ , NH ₄ , PO ₄ , SiO ₂ + Carbonate system (Ph, DIC, TA, pCO ₂). O ₂ , CO ₂	Diatoms Dinoflagellates Nanoflagellates Fixed : N:P:Si Variable C:N C :Chla	multispectral	Micro. Meso. Noctiluca Jellyfish Fixed	yes	1 POM 2 DOM	Denitrif. ODU production Annamox	2D with meta-modelling
NORWECOM ⁴ ECOSMO ⁵	ARC	19	NO ₃ , NH ₄ , PO ₄ , SiO ₂ + Carbonate system (PH, DIC, alkalinity, pCO ₂ , CaCO ₃). O ₂ , CO ₂	Diatoms Flagellates Coccolithophores Sea-ice algae Fixed Variable C:Chla	1 band	Micro. Meso. Fixed	no	1 POM 1 DOM	Denitrif.	2D with 3 sediment pools, on for Si, N, and P respectively
BFM ⁶	MED	54	NO ₃ , NH ₄ , PO ₄ , SiO ₂ + Carbonate system. O ₂ , CO ₂ , PIC	Diatoms Dinoflagellates Nanophyto. Picophyto. Variable C:N:P C :Chla	multispectral	Nano. Micro. Meso. Variable	2 yes	1POM 3 DOM	Denitrif.	No sediment module. Burial and mineralization of sinking OM



ERSEM ⁷	NWS	54	NO ₃ , NH ₄ , PO ₄ , SiO ₂ + Carbonate system. O ₂ , CO ₂	Diatoms Dinoflagellates Nanophyto. Picophyto. Variable C:N :P C :Chla	1 band	Nano. Micro. Meso. Variable (except for Meso)	yes	2 POM 3 DOM	Denitrif.	Separate sediment model Fully coupled benthic model
---------------------------	-----	----	-----------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------	--------	-----------------------------------------------------------	-----	----------------	-----------	--------------------------------------------------------------

¹Aumont et al., (2015); ²Neumann et al., (2021); ³Grégoire et al., (2025); ⁴Samuelsen et al., (2015); ⁵Daewel, U., and Schrum, (2013); ⁶Vichi et al., (2015); ⁷Butenschön et al., (2016).



205 2.3 Coupling with the physics

The CMEMS MFCs describe the biogeochemical state of the ocean in the three directions of space and in time. The evolution in space and time of any biogeochemical state variable is governed by the following mass balance equation:

$$\frac{\partial C}{\partial t} + \underbrace{\nabla_H \cdot (\underline{u}C) + \frac{\partial}{\partial x_3}(wC)}_{\text{horizontal and vertical advection}} + \underbrace{\frac{\partial}{\partial x_3}(w_C^s C)}_{\text{sedimentation}} = \underbrace{Q^C}_{\text{production destruction}} + \underbrace{\frac{\partial}{\partial x_3}(\tilde{\lambda}^C \frac{\partial C}{\partial x_3}) + \lambda_H \nabla_H^2 C}_{\text{vertical and horizontal diffusion}} \quad (1)$$

210

Eq. (1) expresses that the rate of change of the concentration of each biogeochemical component C results from the three-dimensional transport by advection ($\nabla_H \cdot (\underline{u}C) + \frac{\partial}{\partial x_3}(wC)$) by the currents, sedimentation ($\frac{\partial}{\partial x_3}(w_C^s C)$ with a sinking velocity w_C^s , diffusion ($\frac{\partial}{\partial x_3}(\tilde{\lambda}^C \frac{\partial C}{\partial x_3}) + \lambda_H \nabla_H^2 C$) with vertical and horizontal diffusion coefficients of $\tilde{\lambda}^C$ and λ_H , respectively, and biogeochemical interactions (Q^C). Solving this equation requires information on the three-dimensional fields of temperature,

215 salinity, currents and diffusion coefficients. This information is provided by a hydrodynamical model that is run either at the same time as the biogeochemical model and provides the required data at each time step (i.e. online coupling) or beforehand and the data are regularly stored (i.e. offline coupling). The modelling systems of the CMEMS production centres differ in the physical models used, the mode of coupling (online vs. offline), the coupling method (e.g. via a coupler or through direct integration into the physical model) and the resolution (Table A4). The biogeochemical interactions (Q^C) are described by the
220 CMEMS biogeochemical models.

At the start of the CMEMS service in 2015, the seven production centres used five different hydrodynamic models in the production of the NRT and MY products. By 2024, this was considerably harmonized: the HYCOM (Hybrid Coordinate Ocean Model, <https://www.hycom.org/>; Bleck, 2002; Sakov, 2012) model is used in the ARC MFC and NEMO (Nucleus for
225 European Modelling of the Ocean, Madec, G. and the NEMO System Team, 2023) is used for all the other regions. Five systems are online coupled in NRT and MY mode (ARC, BAL, BLK, IBI, NWS) while two are offline coupled (GLO, MED).

In 2015 all biogeochemical models were directly coupled to the respective physical model, without the use of any specific coupling software or framework. By 2024, five modelling systems (2 NRT, 3 MY) are coupled using the Framework for
230 Aquatic Biogeochemical Models (FABM, Bruggeman and Bolding, 2014), which provides a Fortran based interface for the communication between hydrodynamic and biogeochemical models. The other systems are now in the process of transitioning to FABM. The use of a coupling framework makes it much easier to switch between physical models and model versions and provides the opportunity to exchange and combine modules of different biogeochemical models.



2.4 Spatial resolution

235 Over the 10 years of the CMEMS, four out of the seven NRT and three MY modelling systems increased in horizontal resolution and five in NRT (three in MY) reached a higher vertical resolution by a factor of almost 2 to 4. Four systems now have a similar spatial resolution for both the MY and the NRT product, while in 2015 the horizontal resolution of the MY systems was coarser than the respective NRT systems for four out of seven regions. This harmonisation between NRT and MY systems makes products more user friendly, as the data is more consistent and better comparable.

240

2.5 CMEMS biogeochemical models' data needs

In addition to the physical fields, the CMEMS MFC Biogeochemical systems need observations for calibration, validation and assimilation (e.g., satellite, Biogeochemical Argo, in-situ). The limited availability of biogeochemical data often leads to the fact that the data used for model calibration, validation and assimilation are the same. Very often, there is no independent data to validate the system prediction; the situation is even more critical in NRT. In addition, MFC Biogeochemical models need data for initialization, running (e.g. bathymetry) and for forcing at their boundaries with the atmosphere (wind speed, pCO₂, nutrients deposition), the continent (river and land inputs) and the open sea. The CMEMS biogeochemical forecasting systems need to consider the exchanges of materials at the ocean interfaces with the atmosphere, land and open sea. All model systems used in the CMEMS incorporate riverine inputs. This includes two components: freshwater inflow and river loads of nutrients (nitrate, phosphate, silicate, carbon). The evolution in the quality of the river forcing from 2015 to 2024 is small. The sources for the river forcing data, for the water discharges and nutrients, are still very heterogenous among the MFCs. We note that an increasing number of systems (4 over 7) use the data provided by a hydrological model, mainly E-hype and EFAS to estimate the water discharge. The other systems use climatologies (2) or reconstructed data series (2). For the nutrients, the situation is similar, with the use of river's modelling data from E-hype (1) and Globalnews (2), reconstructed datasets and climatologies or a combination of both (4). At the atmospheric boundary, data is needed for the physical component (e.g. wind speed), but also for the biogeochemical component (partial pressure of CO₂ and deposition of nutrients). While the spatial and temporal resolution of physical atmospheric forcing data increased significantly from 2015 to 2024 by a factor of ~10 in NRT (ECMWF) and ~6 in MY (ERA), the boundary data for nutrient input rely on climatological data in all systems in 2024 and pCO₂ is imposed considering an increasing trend. At ocean lateral boundaries, in systems connected by a Strait with the open sea, climatological data from various sources is used in 2015 as well as in 2024. In systems connected to the global ocean, results from GLO MFC are used (i.e. IBI NRT and MY, ARC NRT) while in the NWS MY systems BAL MFC products are used.

245

255

260

2.6 Data Assimilation

After 10 years of CMEMS biogeochemical predictions, the number of systems that assimilates biogeochemical data has increased moving from the assimilation of satellite chl *a* in one NRT and two MY systems in 2015 to five NRT and MY



265 systems in 2024. The data assimilation systems have been improved in terms of correction schemes, propagation of the
correction, number of assimilated and corrected variables, extending to biogeochemical parameters correction. (Ensemble)
Kalman filters and smoothers and 3Dvar are the DA methods of choice. When assimilating surface chl *a*, the modelled surface
chl *a* value is corrected and this correction is propagated at depth. In both the GLO and NWS systems, the surface increment
is applied uniformly throughout the mixed layer, with an exponential decrease down to zero at the base of the mixed layer
270 (imposing no correction below the mixed layer depth). The ARC MFC system, by contrast, uses an empirical projection
function based on Ardyna et al., (2013), a variant of the mixed layer depth-aware projection function (Uitz et al., 2006)
calibrated for Arctic Ocean chl *a* profiles. Unlike mixed layer depth-constrained correction, this approach directly adjusts
subsurface chl *a* maxima below the mixed layer, which are frequently observed in the western Arctic Ocean. The reconstructed
chl-*a* profiles are then assimilated in the ARC NRT system by nudging and this improves the quality of the prediction including
275 at depth as demonstrated by a comparison with Argo data (Figure 2). The MED MFC system propagates the information to
depth via the error covariance obtained from empirical orthogonal functions (EOFs) computed on the shelf and in the open
ocean as described in Teruzzi et al., (2018). In the Black Sea, 3D EOFs are computed and used to propagate the correction,
with along the vertical, an exponential decrease of the correction with depth that guarantees no correction below the mixed
layer depth.

280 The correction of total chl *a* is propagated to other modelled variables. In some systems, the increment in chl *a* profile is
translated to the correction of PFTs chlorophyll and biomass based on chl *a* -biomass ratio in the background model state. In
the ARC MY system, the assimilated satellite chl *a* data updates both model parameters and a multivariate state that includes
PFTs and nutrients such as nitrate, phosphate, and silicate. The correction of PFTs is based on a covariance analysis (Simon
et al., 2015). In the MED, PFTs and nutrients (nitrate and phosphate) are updated using a monthly-dependent background
285 covariances linking chl *a* to the corrected biogeochemical variables based on a 20-year model simulation (Teruzzi et al., 2018).
In the NWS MY system, satellite derived PFT are assimilated and corrected considering in the data assimilation protocol the
large uncertainties in the PFT satellite product (Skákala et al., 2018); this brings the phytoplankton community structure in
line with that derived from satellite data (Figure 3). In the GLO NRT, satellite chl *a* has been assimilated since 2019 (cf. Figure
12 in Appendices) to correct the chl *a* content of both PFTs as well as nitrate, using model daily-climatological covariances
290 (Mignot et al., 2023).

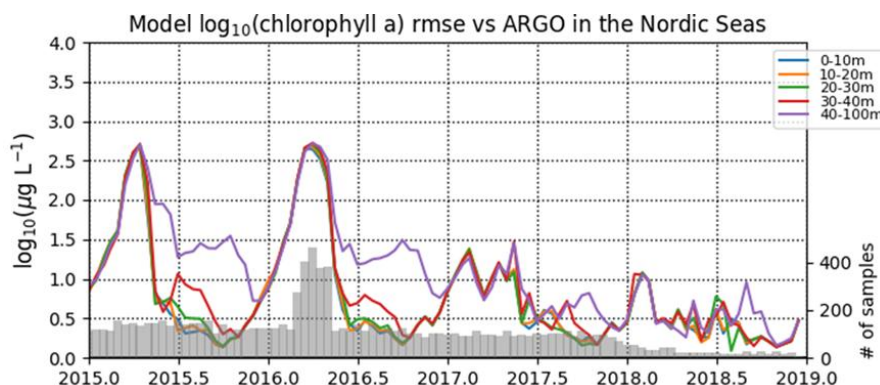
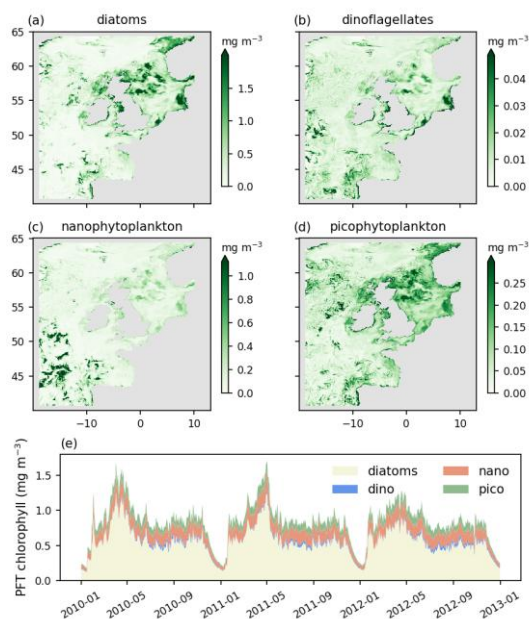


Figure 2: Impact of synthetic chl *a* profile assimilation, reconstructed from satellite observation extended at depth, in the ARC NRT system. The figure shows over the Nordic Seas, the RMSE computed between model results and ARGO observation, over the first 100m and 2015-2019 and for 5 depth ranges. Bar plots are number of ARGO profiles data used for calculating the RMSE. The nudging of the synthetic profiles in the ARC System starts in January 2017.

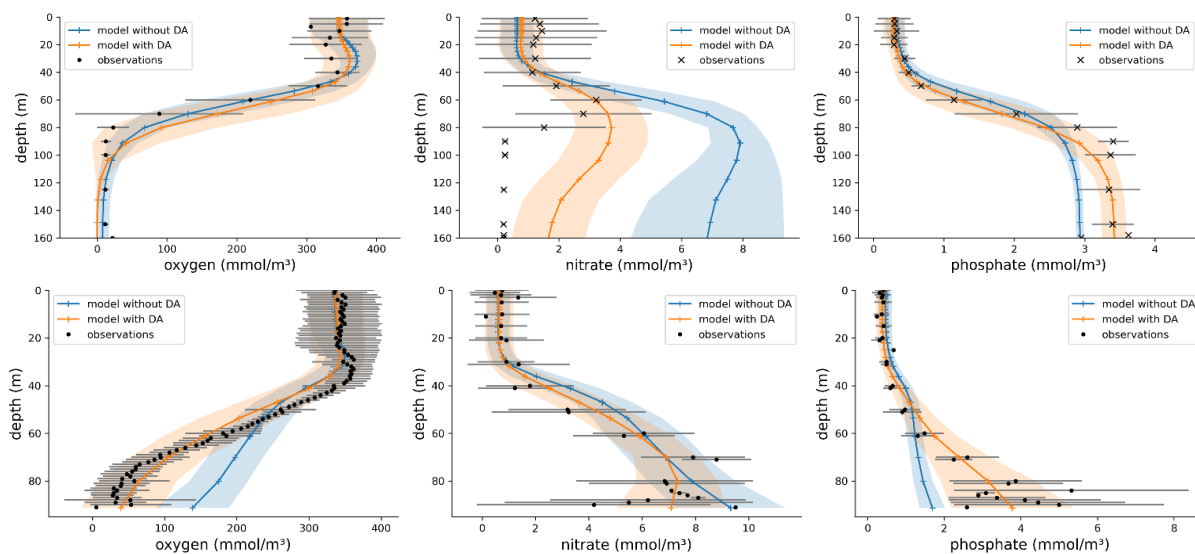
Three systems moved to a multi-observation data assimilation adding the assimilation of vertical biogeochemical profiles delivered by in-situ data from Biogeochemical Argo, gliders and ship to the satellite chl *a* (Table A7). The impact of these subsurface data is anticipated to be especially significant in oligotrophic conditions, where deep chl *a* maxima (DCM) contribute substantially to annual primary production. For instance, in the Baltic Sea, the assimilation of in-situ profiles of oxygen, nitrate and phosphate improves the quality of the prediction, particularly below 60m (Figure 4). In the NWS and MED systems, the joint assimilation of vertical profiles and satellite chl *a* improves the quality of subsurface Biogeochemical state reanalysis (Skákala et al., 2021; Teruzzi et al., 2021; Cossarini et al., 2021). In the MED, synthetic vertical profiles are reconstructed using convolutional neural network, the Predict Profiles Convolutional (PPCON) model, that predicts nitrate, chl *a*, and backscattering (bbp700) starting from the date and geolocation and from temperature, salinity, and oxygen profiles (Pietropolli et al., 2024). The 3DVar assimilation scheme operationally applied in the MED assimilates PPCON reconstructed nitrate profiles (Amadio et al., 2024). The assimilation of the large number of PPCON-augmented nitrate observations impacts both nitrate (Figures 5e and f) and phosphate (not shown) especially at the layer right under the DCM. The number of assimilated profiles is largely impacted by the PPCON augmentation, enhancing the possibility of reshaping the nutrient vertical distribution (e.g., the nutricline vertical position and intensity). The effects of PPCON-assimilation vary with the number of profiles during the length of the simulated period and across Mediterranean sub-basins.



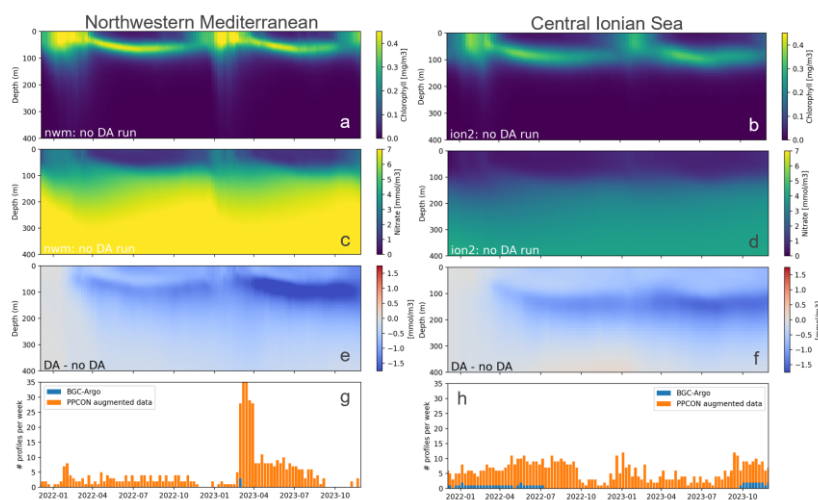
315 **Figure 3:** Surface level phytoplankton functional type (PFT) chl *a* concentration from the NWS multi-year product. (a-d) values for a sample day, 12 May 2012, for (a) diatoms (b) dinoflagellates (c) nanophytoplankton (d) picophytoplankton. (e) Time series of daily mean values for the southern North Sea, 2010-2012. This product was created using NEMO-ERSEM with assimilation of satellite PFT chlorophyll.

Four of the seven CMEMS biogeochemical models are forced by a physical model run with data assimilation. There are two ways for performing a joint data assimilation: weakly or strongly coupled assimilation. In weakly coupled assimilation both physical and biogeochemical variables are assimilated but are updated separately. In a strongly coupled assimilation, the updates of for example temperature would also update, for example, nitrate and chl *a*. Due to nonlinear processes strongly coupled assimilation has proven challenging and is not applied in any of the operational systems today, while four of the forecast centres use weakly coupled data assimilation. In some cases (e.g. in the equatorial region), the assimilation of physical variables caused spurious upwelling of nutrients and excessive primary production (Ford et al., 2018; Gasparin et al., 2021).

325 This problem is managed in the GLO systems by activating a 3D climatological relaxation.



330 **Figure 4:** Effect of in-situ profile data assimilation of dissolved oxygen, nitrate and phosphate on the respective product variables in the BAL MY at the mooring stations BMPH2 (top) and BMPK2 (bottom). The figures depict data for a timeframe of November 2019 to December 2020, with the blue line representing the temporal mean based on a model without any data assimilation, the orange line representing the temporal mean based on a model with oxygen, nitrate and phosphate data assimilation and the black dots denoting the temporal mean of the respective measured data. The highlighted blue and orange areas and the grey lines depict the standard deviation of the respective data over the considered period.



335 **Figure 5:** Hovmoeller diagrams of the years 2022 and 2023 of mean profiles in the northwestern Mediterranean (left column) and in the central Ionian Sea (eastern Mediterranean, right column) of chl *a* (first row), nitrate (second row) in a free run (i.e. no data assimilation), nitrate differences between an assimilated and a non-assimilated run (third row). The last row shows the number of BGC-Argo (blue) and PPCON (orange) assimilated profiles per week. Bulk nutrient levels and chlorophyll concentrations differ markedly between the western and eastern Mediterranean regions. Being a strong stratified and oligotrophic sea, in the Mediterranean Sea, phytoplankton alternates between surface bloom events, after winter mixing occurrences, and subsurface maxima in summer stratified conditions (first row), while nutrients (second row) are mostly confined in the subsurface layers with the exception of mixing periods. In the western area, the impact of PPCON-assimilation becomes more evident in the second year after a rapid change on the density of assimilated nitrate profiles (Figures 5e and g).



345 In the eastern almost permanently stratified Mediterranean areas, we observe a steady correction of the depth of nutricline (Figures 5f) related to the nearly uniform coverage of PPCON profiles.

2.7 COMPUTING resources

The increase in the number of products, resolution, length of simulations, frequency of outputs requires more storage space and computational resources. The computational resources in terms of number of cores used for the simulations has been
350 multiplied by 10-20 in 10 years. The required storage space for the MY has been multiplied by a factor of 5-10 and amounts to several tens of TB (e.g. 40 TB in the case of the MED) (Tables A5 and A6).

3. Quality and value of CMEMS biogeochemical products

3.1 Product quality assessment

The quality of the CMEMS products routinely delivered by its 7 MFCs is documented in Quality Information Documents
355 (QuIDs) produced by the MFCs in a coordinated way. These provide both qualitative and quantitative information to users. Since the start of the service the QuIDs have developed to include more comprehensive assessment of a greater number of variables, though assessment is still constrained by the limited availability of biogeochemical observations (Figure 6).

Where possible, a quantitative assessment of product quality is made by comparing observations and model values matched in space and time either in model or in the observation space (respectively GODAE Class 1 and 4; Hernandez et al, 2009). At
360 the start of the service this was mainly limited to chl-*a*, assessed against satellite values and, for the Baltic, nutrients and oxygen assessed against measurements at HELCOM monitoring stations. Now more variables are assessed by a greater number of MFCs, using databases of in situ measurements and new sources of observation data, notably BGC-Argo. In addition, the Class 4 assessments have more detail, often including breakdown by region and/or by depth. Mean bias and root mean square difference (RMSD) between the model and observations are the most common measure of accuracy used, but some MFCs also
365 include correlation and the use of summary plots such as Taylor (cf. Figure 12 in Appendices) and target diagrams is becoming more common.

Class 1 assessments are supplemented by qualitative and quantitative comparison to climatology of observations, which enable assessment of how well the products capture broad spatial and seasonal patterns and observed variability. Most MFCs are now able to provide some sort of assessment for most variables, though a few remain very limited by lack of observations, notably
370 iron, phytoplankton and zooplankton biomass.

The QuIDs use a wide variety of plots to add detail and nuance to the numerical accuracy metrics: surface and cross-section plots to show spatial variation compared to observations, depth profiles compared to BGC-Argo to show variability, time series and Hovmöller diagrams to show difference in bloom timings. Again, the number of variables and range of plots used has



375 developed greatly over the last few years as the MFCs gained experience and the QUIDs now provide a rich source of quality information for the user.

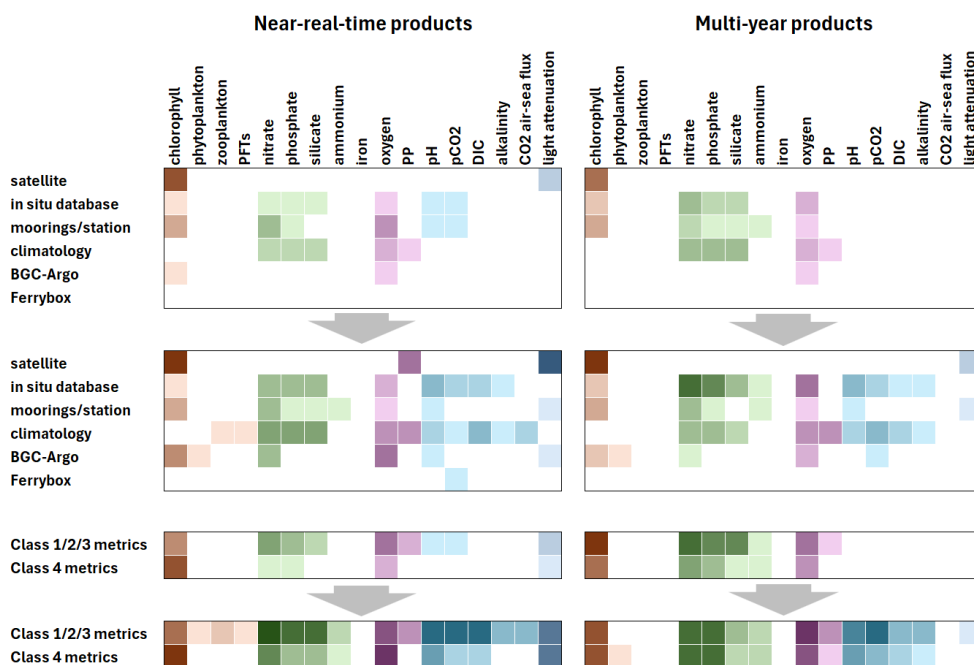


Figure 6: Observations and metrics used in product Quality Information Documents. Darker colours show more MFCs using that type of observation or metric. 1st and 3rd rows are for 2015 while 2nd and 4th rows are for 2024.

380

3.2 Evolution of Product quality

As the number of metrics and the number and types of observations used to evaluate CMEMS product quality have increased, the accuracy of these products has also improved. Because of the increase and modification of metrics, the comparison can be made only at an aggregated scale rather than for individual CMEMS MFCs products, but it remains informative regarding changes in product quality.

385

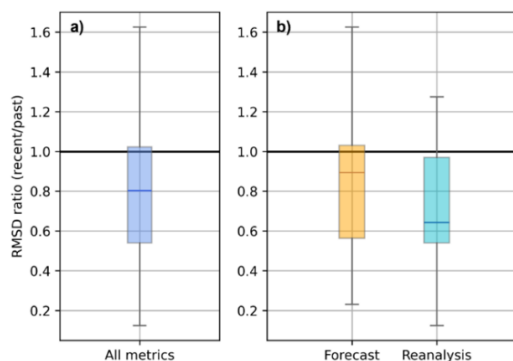
Considering RMSD of NRT and MY products with respect to both satellite and in situ observations, the ratio between present-day and past RMSDs provides an indication of quality evolution. Such ratios can be calculated when and where the RMSDs are currently calculated similarly to the beginning of CMEMS. When these ratios are smaller than 1, they indicate an improvement in quality (i.e., a reduction in RMSD). RMSD ratios have been calculated for the different MFCs, for different observation sources (satellite and in-situ), and for NRT and MY products. The RMSDs are extracted from the CMEMS documentation dedicated to product quality, and, in the case of short-term forecast products, are based on pre-operational runs

390

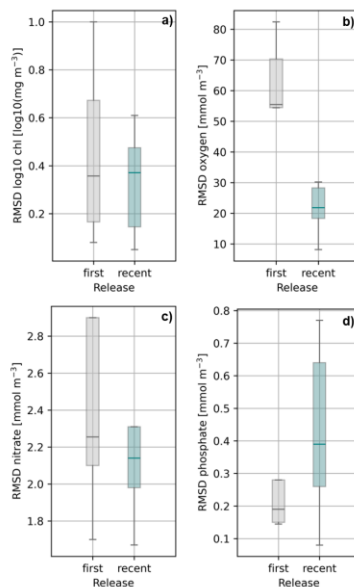


of at least one year. The RMSDs with respect to satellite are based on chl *a* observations, while in situ measurements included mainly chl *a*, oxygen, nitrate and phosphate (and only marginally silicate).

Statistics calculated on the resulting RMSD ratios (shown as boxplots in Figure 7) indicate a general improvement in product quality, with a median close to 0.8 when considering all RMSD ratios together. The inclusion of specific developments of data assimilation are among the elements that potentially affected the performance of the CMEMS biogeochemical modelling systems; however, in many cases it is not easy to disentangle the effects of assimilation from those of other system developments (e.g., model parametrisation, changes in physical forcings, improved boundary or initial conditions). Examining the 25th and 75th percentiles together with the whisker values, a relatively large variability in the RMSDs is detectable, indicating that improvement can depend on the area, the type of observation, and the type of product. While the relatively small number of RMSD ratios available in single MFC limits an in-depth analysis of variability across different areas, the statistics on RMSD ratios can be further detailed by differentiating RMSD ratios based on NRT and MY products, and on satellite or in-situ observations. In particular, the median of the RMSD ratios in MY is smaller than in NRT products (close to 0.65 and 0.9, respectively), also showing smaller variability (Figure 7). Statistically larger improvements and greater uniformity in reanalyses can be attributed to the fact that reanalysis assessments are usually carried out on long-term datasets that have changed relatively little since the beginning of CMEMS, whilst the forecast RMSDs are usually calculated considering a relatively short simulation (from one to three years) on a varying time period (that usually shortly precedes the moment when the forecast is made operational in CMEMS). In addition to the RMSDs, other metrics (not shown) available in the CMEMS documentation demonstrate the improvement in the capability of the CMEMS forecast and reanalysis products to simulate a range of biogeochemical processes, including the seasonality of surface phytoplankton blooms and vertical dynamics such as the DCM, the nutricline, and the oxygen penetration depth. The RMSD has decreased over the last 10 years for chl *a*, nitrate and, in particular, for oxygen while for phosphate it has increased (Figure 8). It is worth to note that in-situ observations of nitrate and phosphate are generally scarcer and unevenly distributed over time with respect to oxygen. In particular, phosphate RMSD values in Figure 8 are limited to two Copernicus Marine MFCs, and the RMSD increase occurs in one of the two MFCs and for the NRT product only. This limited and irregular data coverage limits the direct comparison of quality assessments based on in-situ phosphate and nitrate between the ten-year distant validation (i.e., 2015 and 2024).



420 **Figure 7:** Box plots of the ratio between RMSDs at a recent CMEMS product release (2024) and at the start of the release of CMEMS operational products in 2015. Each box plot shows the median (line inside the box), the 25th and 75th percentiles (lower and upper limits of the box), and the most extreme data points within 1.5 times the interquartile range (i.e., the distance between the 75th and 25th percentiles) from the box (whiskers). Box plots are shown for: all RMSD ratios together (a), RMSD ratios separated between forecast and reanalysis (b).



425 **Figure 8:** Box plots of the RMSDs at a recent CMEMS product release (2024) and at the start of the release of CMEMS operational products. Each box plot shows the median (line inside the box), the 25th and 75th percentiles (lower and upper limits of the box), and the most extreme data points within 1.5 times the interquartile range (i.e., the distance between the 75th and 25th percentiles) from the box (whiskers). Box plots are shown for: satellite chl *a* (logarithmic) (a), oxygen (b), nitrate (c), and phosphate (d).

430



4. Use of CMEMS products for application

4.1. Indicators

To raise public awareness and to support sustainable decision-making, there is a need for reliable and science-based knowledge transfer on the state, variability, and change of the ocean and their socio-economic implications. To this end, the construction of environmental indicators from raw model outputs is a useful tool that is already performing multiple functions in science and policy-making, in particular in the international dialogue on climate and biodiversity change (e.g., IPCC, 2021; WMO, 2024; IPBES, 2020). In response to this need, the CMEMS has initiated the Copernicus Ocean State Report (OSR; <https://marine.copernicus.eu/access-data/ocean-state-report>) that provides the scientific foundation to support the delivery of Ocean Monitoring Indicator (OMI; <https://marine.copernicus.eu/access-data/ocean-monitoring-indicators>) and to communicate ocean knowledge to a wide range of users, from scientific community, to policy-makers, decision-makers, and the wider public. Today, the CMEMS OMI framework comprises a total of 23 operational indicators relevant for monitoring changes in the ocean’s biogeochemical state and health (Table 2). In addition, the CMEMS is in direct contact with the Statistical Office of the European Union, EuroStat4, for the annual provision of ocean acidification and eutrophication indicators to support European reporting on the UN Sustainable Development Goals (SDG 14 « Life Below Water »). The CMEMS also aims to better serve the Regional Seas Conventions (RSC) such as OSPAR, HELCOM, UNEP-MAP. Regional Seas Conventions are agreements between countries sharing a common body of marine water, establishing rules for the management of marine resources in that region. These conventions address topics such as pollution prevention, climate change, pressures from human activities, and the protection and conservation of biodiversity and ecosystems. Their aim is to foster cooperation and coordination among coastal states to combat pollution and ensure sustainable development by establishing strategies and regional plans for their respective regions.

Table 2: Overview on biogeochemical indicators currently included in European and international environmental indicator frameworks. All these indicators can be estimated by CMEMS biogeochemical models.

Indicator framework	Indicator type
Copernicus Ocean Monitoring Indicators	Chlorophyll & primary production Ocean acidification Ocean deoxygenation Eutrophication & bloom Oligotrophication
WMO Global Climate Indicator framework	Ocean acidification
GOOS Ocean Indicator framework	Ocean acidification Ocean deoxygenation Net community Production



455 4.2 Coastal downscaling in support to policy implementation

Many of the users of biogeochemical model-based products primarily use the data to conduct scientific research and to provide the knowledge needed on marine ecosystems health to inform policy and decision-making, and then to implement these policies and to develop conservation and management strategies. Several use cases featuring Biogeochemical model products illustrate the Service's user-driven approach and highlight the stakeholder engagement that the CMEMS supports. Among these, 460 CONNECT Tagus (https://connect-portal.lnec.pt/connect/indicators?estuary_id=1) - launched in December 2023 under the Copernicus Marine National Collaborative programme - serves as a high-resolution coastal monitoring service in the Lisbon Metropolitan Area. Its objectives are to monitor and manage water quality and extreme water levels in the Tagus Estuary, a crucial ecosystem in Portugal. Underpinned by the IBI NRT system, CONNECT Tagus provides valuable data on key water quality parameters to monitor eutrophication processes and to identify hypoxic zones, both of which are essential for assessing 465 the impact of nutrient overload and for maintaining water quality standards. Within this context, the essential ocean variables (EOVs) provided by the CMEMS include chl *a* concentration, dissolved oxygen, and nutrient concentrations (e.g. nitrogen and phosphorus). This service directly contributes to the objectives of the MSFD (Marine Strategy Framework Directive) which requires European Union Member States to achieve a "good environmental status" of their marine waters. By facilitating the systematic monitoring of water quality and extreme water levels, this CONNECT Tagus service indeed supports informed 470 decision-making and sustainable management of coastal ecosystems. The CONNECT Tagus use case also benefits environmental agencies and coastal managers by improving forecasting capabilities and providing effective tools for better early warning systems.

5. The CMEMS biogeochemical prediction systems: strengths, weaknesses, opportunities and threats.

We overview the main strengths and weaknesses of the current CMEMS biogeochemical prediction systems and opportunities 475 and threats for the next 5-10 years. Table 3 summarizes the main points.

Strengths

The CMEMS biogeochemical modelling systems are grounded on state-of-the-art prediction capabilities and are continuously updated with the latest developments to improve the quality of the service and to meet user's needs. The evolution of the modelling systems, the quality assessment protocols and documentation are coordinated across the seven forecasting centres 480 following community agreed best practices. During the last 10 years, our capabilities to forecast the green ocean has significantly improved with the routine delivery of quality-controlled information at mesoscale in the global ocean and submesoscale in European regional seas and Arctic Ocean. The baseline CMEMS biogeochemical products have been extended from inorganic nutrients, oxygen and phytoplankton related variables to include carbonates, light and zooplankton



delivered with an enhanced quality. This improvement results from the continuous development of biogeochemical models
485 that assimilate new and improved data streams. Ocean colour sensors sample visible light radiation at an increased spatial,
temporal and spectral resolution and autonomous platforms like BGC Argos and gliders deliver NRT information that allows
the reconstruction of ocean oxygen, chl *a*, spectral radiation, pH, nitrate. The CMEMS biogeochemical models, except two,
assimilate satellite chl *a* and a limited number of them (2 in NRT and 1 in MY), assimilate in-situ data. CMEMS systems are
now extended for simulating variables more directly linked to observations; avoiding the use of an imperfect observation
490 model. The best example of this model extension is the coupling of the biogeochemical model with a spectral radiative transfer
model able to simulate the sea surface reflectance, better connected with what is observed by the satellite.

Weaknesses

The abundance of biogeochemical and, more critically, biological data is still lagging well behind physical data. This limits
495 model development and the delivery of a sound quality assessment of biogeochemical products to fit user's purposes. The
assessment of the impact of ocean observations on the quality of the prediction is still very limited, and this is particularly
striking for biogeochemistry. It remains very difficult to robustly highlight the benefit of the assimilation of physical and
biogeochemical variables on the quality of the prediction. The assimilation of physical variables may deteriorate the solution
by generating spurious circulation (e.g. vertical velocities) while the lack of biogeochemical data very often prevents a sound
500 impact assessment with independent data (i.e. data that are not assimilated). Another important issue is the lack of consistency
between the biogeochemical data collected with different sensors (e.g. reflectance/fluorescence/pigments derived chl *a*).

The CMEMS MFCs currently provide 10-day forecast at km scale, which is still limiting for some users' applications requiring
seasonal forecasts at a scale below the km. These requirements question the capabilities of our system to forecast the
505 biogeochemical ocean at several weeks and at scales below a km. A coordinated predictability analysis as such done in Leroux
et al., (2022) would allow to set a limit to the minimum length scales and the time horizon that are predictable considering the
different sources of error, with among them, some that cannot be reduced, at least currently. For instance, the quality of the
boundary conditions remains critical especially those from the rivers systems. Most of the CMEMS models use climatologies
for forcing their model at the rivers' mouths and this hampers the quality of model predicting capabilities (Rose et al., 2017)
510 and the assessment of biogeochemical budgets. The connection between the different Copernicus services is expected to
improve the situation.

Opportunities and perspectives

The needs for biogeochemical products are exploding to support a diversified set of applications at different time and spatial
515 scales like climate change detection (deoxygenation, acidification), carbon budget estimation, ocean biodiversity conservation,



water quality assessment, extreme events forecasts (e.g. hypoxia, HABs, sulphur emission). Today, the CMEMS OMI are mainly based on observations, but there is a growing interest in model data. For example, in the context of climate change, calculating trends and predicting the three-dimensional multivariate state of the biogeochemical ocean is becoming an issue. The advances in model development, parameterization and data assimilation have significantly improved the accuracy of long-term biogeochemical trend simulations, particularly for critical global indicators such as pCO₂, pH and oxygen. An enhanced contribution of modelling products to indicators estimation is then expected. As an illustration, GLO-MFC is currently evaluating the performance of its MY system in representing surface pCO₂ (spCO₂) — and, by extension, air-sea CO₂ fluxes — over the standard 30-year MY period. This assessment involves extending the MY system to assimilate surface maps of DIC, Total Alkalinity, and pH, derived from reconstructions provided by the CMEMS Multi-Observations Thematic Assembly Center. These reconstructions are generated using a Neural Network trained on the SOCAT database (Chau et al., 2022; Bakker et al., 2016). Preliminary results, illustrated in Figure 9, compare the model's DIC and spCO₂ outputs to the corresponding SOCAT-based reconstructed products. These results are promising as the assimilation effectively corrects both the amplitude and trends of DIC and spCO₂, aligning them more closely with observation-based reconstructions (Perruche et al., 2026, *in prep.*).

530

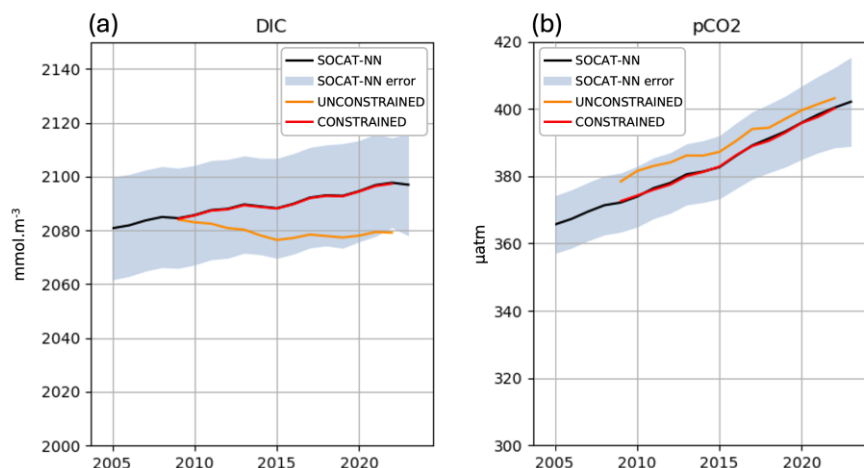


Figure 9: Global surface average of (a) DIC ($\text{mmol}\cdot\text{m}^{-3}$) and (b) pCO₂ (μatm). Black lines denote the reconstructed product from the SOCAT-based neural network. The orange line represents the 3D GLO-MY model with assimilated dynamical forcing, without climatological relaxation or data assimilation. The red line corresponds to the 3D GLO-MY model with assimilated dynamical forcing, including climatological relaxation and assimilation of DIC, Total Alkalinity, and pH. (Preliminary results, Perruche et al., 2026, *in prep.*).

535

The limited availability of biogeochemical data hampers the estimation of the quality of products, a quantity of high priority for users. Further improvements in model uncertainties estimation of marine indicators and products are expected from the adoption of ensemble modelling systems. Ensemble modelling systems offer information on uncertainty that complements that provided by a classic model-data comparison exercise. Yet these ensemble systems are still costly to run as the computing

540



resources or time, depending on the implementation priorities (e.g. to keep unchanged the time of simulations or the computing cores) are roughly multiplied by the number of members, usually at least 10. The disk storage is also multiplied by a factor that depends on what is saved, at least the mean and spread of the ensemble. Transitioning towards ensemble modelling would have a huge impact and would require a vision for the development of computing resources. In some MFCs, an ensemble forecasting system is under development and testing (Figure 10).

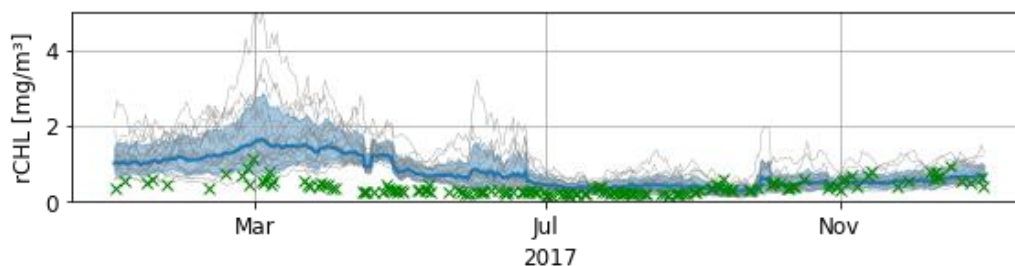
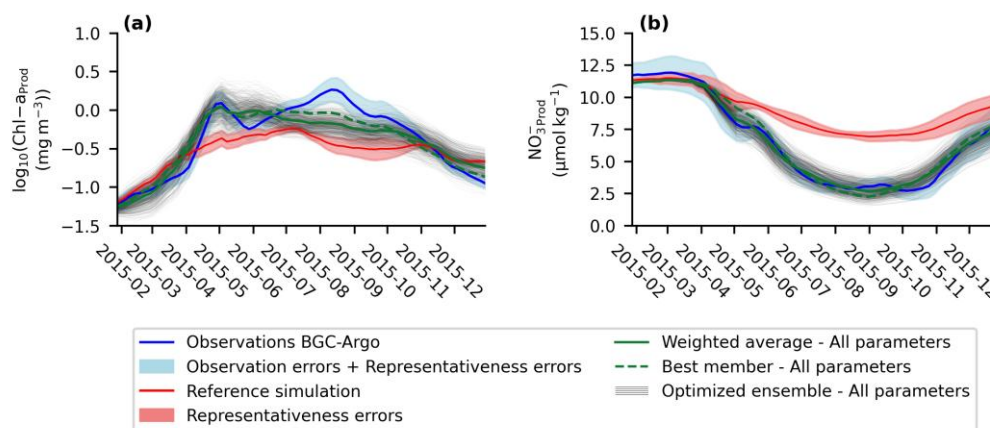


Figure 10: Seasonal cycle of the surface chl *a* (in mg m^{-3}) from satellite (green crosses) and simulated by an ensemble 3D version of the BLK MFC system obtained by perturbing inherent optical parameters. The mean and standard deviation of the ensemble are represented by respectively the bold blue line and the shaded blue area, and each thin grey line corresponds to a member of the ensemble. Series are taken in the eastern gyre (43.22° N , 36.63° E). Figure reproduced from Macé et al., (2025).

So far, the calibration of the hundredths of biogeochemical parameters is still mostly done by hand focusing on the most sensitive one. Some MFCs (IBI, GLO) have started an automated calibration procedure that fits model prediction with biogeochemical Argo data (Hyvernat et al., 2025). The calibration is done on a one-dimensional version of the biogeochemical model that simulates the ARGO biogeochemical data along the Argo track. This approach is still under testing and is not yet implemented in the current CMEMS prediction systems. The conducted experiments have shown that this improves the seasonal cycle in the productive layer, particularly the magnitude and timing of the spring phytoplankton bloom (Chl-*a*) and the seasonal drawdown of nutrients (here, NO_3^-) (Figure 11).

560



565 **Figure 11:** Seasonal cycle of (a) $\log_{10}(\text{Chl } a)$ and (b) NO_3^- in the productive layer. The blue curve represents observations from BGC-Argo float (#5904479) in the IBI area, with associated error (blue shading). The red curve corresponds to the reference PISCES in 1D configuration, with associated error (red shading). The green line indicates the ensemble weighted mean, while the black curves represent individual ensemble members. The ensemble was generated by optimizing all parameters of the PISCES model. Figure from Hyvernat et al. (2025).

The current CMEMS green ocean products are mainly related to the biogeochemistry. Biological products are very limited and only related to the pelagic low trophic levels. Progress in our capabilities to model the biology, pelagic and benthic, are expected in the next years. The Horizon Europe NECCTON project (Grant Number 101081273, <https://neccton.eu/>) aims at enhancing our prediction capabilities of the biology in an operational framework. NECCTON addresses the modelling of the high trophic levels, the benthic and pelagic systems and pollutants. New modelling capabilities are being developed and recommendations about the optimal protocol of delivery of biological products and assessment of their uncertainty will be delivered. In particular, the spatial (e.g. bioregions, EEZ, 3D/2D) and temporal (e.g. annual/seasonal/monthly averages) scales of biological products like fishes' biomass, benthic flora and fauna that can be tested and fit the purpose of users' applications are defined. Yet the extension of the CMEMS service to the biology and higher trophic levels is currently limited by the availability of adequate quality-controlled reference datasets related to biology.

580 With increasing pressures on marine ecosystems, "What-if scenarios", and the forecasting models that underpin them, are becoming essential tools for improving coastal management and ensuring more informed and proactive decision-making in response to environmental change. "What-if scenarios" using marine biogeochemical models enable researchers to explore the potential impacts of various environmental changes on marine ecosystems and global biogeochemical cycles. The most prominent example is provided by the IPCC consortium's simulations which assess the consequences of increased carbon dioxide emissions on phenomena such as ocean acidification and deoxygenation. These scenarios can serve a wide range of additional applications such as predicting the consequences of altered nutrient inputs from terrestrial sources - resulting from climate change or shifts in land-use practices - on coastal eutrophication. They are also valuable in examining the implications



of large-scale ocean interventions, including geoengineering strategies aimed at mitigating global warming, or evaluating changes in marine biodiversity associated with rising ocean temperatures. In the context of biodiversity conservation, they can help assess the impact of different levels of protection in Marine Protected Areas (MPA). By altering input variables or climate forcing factors, these model-based “What-if scenarios” provide valuable insights into the potential risks, vulnerability, and resilience of marine ecosystems under different environmental stressors. This information is expected to play a critical role in helping decision-makers to better anticipate risk, optimize resource management and prepare conservation strategies.

Artificial intelligence has the potential to drastically change the way we routinely predict the green ocean. In 2025, the European Centre for Medium-Range Weather Forecasts (ECMWF) launched its Artificial Intelligence Forecasting System (AIFS) for the atmosphere and for the ocean, Mercator Ocean International launched the Ocean’s global ocean neural network-based forecasting system for the physics (GLONET). A prediction system of the green ocean based on artificial intelligence is now under testing in different MFCs but is not yet operational. Machine learning methods are also expected to improve our data assimilation systems by for instance, improving the link between observed and modelled variables.

Threats

The availability of computing resources may not be sufficient to support the development of (ensemble) (high-resolution) prediction systems and their efficient integration in the European digital twin ocean for offering the users the possibility to run “What-if scenarios”. Increasing the resolution, extending the forecasts and reanalysis length, running ensemble simulations, and offering the users the possibility to run “What-if scenarios” will rapidly push our modelling systems to their computing resources limits. Currently, the NRT and MY prediction systems are run on ~200-600 CPUs. The NRT system requires 1-3 hours to perform 10-day forecasts and a few GB for data storage. The MY systems are run for at least 15 years up to 43 years and may require several months of computation and TB for storage (Table A6). An anticipation of computing resources requirements at European scale for running operational ocean systems for the next 10 years is required to avoid limitations of the potential developments. For instance, the lonely transitioning to ensemble prediction systems roughly multiplies by the number of members (~at least 10) the cumulated time of use of CPUs (i.e. #CPU*Running time). The development of artificial intelligence-based prediction system extends our computing resources needs to GPUs.

The continuous development of biogeochemical forecasting systems, including the integration of the full potential of artificial intelligence, has to be grounded on the new knowledge gained from observing system. This continuous evolution requires the support of a high-quality observing system whose long-term maintenance is increasingly challenged by the recurrent lack of fundings. The organization of the ocean community to demonstrate the value of observation for ocean prediction is by far less mature than the atmospheric community. The atmospheric community has implemented the rolling review of requirements and observation’s impact experiments to deliver guidelines for the evolution of observing systems to support WMO applications. A coordinated approach across biogeochemical forecasting centres, satellite missions and in-situ platforms implementers would help to design and demonstrate the value of biogeochemical observing systems. The UN decade ForeSea



620 (The Ocean Prediction Capacity of the Future) program and its SynObs (Synergistic Observing Network for Ocean Prediction) project are steps in that direction.

Table 3: SWOT analysis of the CMEMS biogeochemical prediction system.

Strengths	Weaknesses	Opportunities	Threats
<ul style="list-style-type: none"> • State of the art methodology • Coordinated development • Common best practices • Quality information • Documented procedures • User driven evolution • Open Access 	<ul style="list-style-type: none"> • Lack of biogeochemical observations. • Empirical formulations of biogeochemical models. • The quality of boundary conditions especially from the land-river system is still insufficient. 	<ul style="list-style-type: none"> • Increased requirements for biogeochemical and biological products. • New data streams (e.g. hyperspectral, acoustic, image analysis). • Enhancing the connection with biology (e.g. benthos, fish). • Advent of new methods (AI, ensemble) to improve the quality of the service to users. 	<ul style="list-style-type: none"> • Risk of limitation of computing resources to support systems developments (ensemble, resolution, AI forecast extension). • Decline funding to support observing systems, in particular, for biogeochemical variables.

625

630



Appendices

Table A1: Characteristics of the NRT MFC-BIO. In black state of the system in 2015 and in red EIS2024.

	PRODUCTS	Horizontal and vertical resolution	Coupling with the physics (online/offline, one-way, two-way, code version)	Bathymetry (source and resolution)	Atmospheric forcing Product, frequency, resolution	Lateral boundaries open sea	Rivers (number, dataset, frequency, other)	Data Assimilation: Code, Scheme, Assimilated Obs, Corrected Obs, assim. Cycle, Independent data for validation	Model Formulation (new variables, new processes)	Products Frequency	Validation Metrics, Observation
ARCTIC (ARC-MFC)	Chla, Phyto, PP, ZOO, NO ₃ , PO ₄ , POC, O ₂ , SiO ₂ , Kd, pH, DIC, spCO ₂ .	12 km, 12 vertical layers 6.25* 6.25 km, 40 vertical levels	Online 1-way coupling with HYCOM 2.2 Online 1-way coupling with HYCOM 2.2 via FABM	GEBCO	ECMWF 6h Climatological atmospheric deposition of N and P	Climatological nutrients, oxygen, DIC and Alkalinity GLO-MFC nutrients	Arctic-HYPE and runoff from Greenland (monthly climatology), Nutrients from Global NEWS (annual climatology)	None Nudging of synthetic chlorophyll profiles (based on satellite chlorophyll)	Model: NORWECOM Model: ECOSMO. Added variable Chl:N ratio, phytoplankton growth formulation, coccolithophores, carbon chemistry	Daily Daily and monthly 2019-2024	NRT:Satellite chlorophyll bias and RMSE Class 4 in-situ and BGC Argo nutrients and oxygen from model calibration/validation before EIS
Iberian (IBI-MFC)	Chla, Phyto biomass in C, PP, NO ₃ , PO ₄ , NH ₄ , O ₂ , SiO ₂ , Fe, Zeuph, pH, pCO ₂ , DIC, ZOO	1/36° 50 vertical levels	PISCESv2 Online coupling with NEMOV3.6	GEBCO 08 + local databases	ECMWF 3h ECMWF 1h Monthly and annual Climatology for atmospheric Dust and Nitrogen deposition	Daily fields from GLO NRT for physics, Weekly fields from GLO NRT for the 24 PISCES variables	Climatology from Ludwig et al. (1996), GlobalNews2 (Mayorga et al., 2010), Additional NO ₃ , PO ₄ from EEA	None	None	Daily and monthly 2020-2024 7-day forecast 10-day forecast	Bias, RMSE, Correlation Satellite WOA climatology GLODAP BGC-Argo ICES, EMODnet MAREDAT



BALTIC (BAL-MFC)	Chla, Phyto, PP, ZOO, NO ₃ , PO ₄ , NH ₄ , O ₂ , SiO ₂ , pH, pCO ₂ , Secchi depth, H ₂ S, Fe	2*2 km, 56 vertical levels 2*2 km, 25 vertical levels	Online coupled to Nemo4.2 via FABM Online coupled to HBM	GEBCO	DMI-HIRLAM So5 MetCOO P-HARMON IE mixed with ECMWF Nutrient deposition: EMEP	Climatology, Interpolated ICES data	Ehype for runoff and nutrients HBV model data	-	ERGOM update: Carbon cycle, H ₂ S and CDOM	hourly Daily, monthly	ICES in-situ data (HELCOM monitoring stations) L3 Satellite data (CHL, KD)
BLACK (BLK-MFC)	Chla, Phyto, NO ₃ , PO ₄ , O ₂ , PP, pH, Alk, DIC, pCO ₂ , air-sea flux, Zd, ZOO, ODU, SiO ₂	5km, 31 levels 2.5 km, 59 levels	Online coupled with GHER, one way Online coupled with NEMO 4.2.		ECMWF (physics) Nutrients deposition (averaged field from atmospheric model)	Bosporus	Reconstructed time series climatology during gaps periods	Satellite chla + Argo O ₂ Kalman fixed based Assimilated variables =Corrected variables	Radiative transfer model Chlorophyll II	Daily Hourly for air-sea fluxes	Class 1: L3 satellite (averaged per region, no spatial maps) Class2: BG ARGO for Oxygen Class 4
Global (GLO-MFC)	Total Chla, Chla from 2 PFTs (diatoms, Nanophytoplankton), PP, NO ₃ , PO ₄ , O ₂ , SiO ₂ , Fe, pH, pCO ₂ DIC, Alk, Kd, total biomass of zooplankton	0.25° x .25, 50 vertical levels Same in 2024	2015: NEMOV3.2- PISCESv2 2024: NEMOV3.6 PISCESv2 Offline coupling	Combination of ETOPO1 and GEBCO 8	ECMWF (daily) Dust and N atmospheric deposition (monthly/annual climatologies)		Climatology from Ludwig et al. (1996), GlobalNews2 (Mayorga et al., 2010)	2015: None 2024: satellite chla Reduce d-Order Kalman Filter, based on a SEEK formulation.		Daily and monthly 2022-2024 7-day forecast 10-day forecast	Bias, RMSE, Correlation Obs: Satellite, WOA climatology, BGC-Argo, BATS, MAREDAT climatology (zooplankton)
Mediterranean (MED-MFC)	Chla, Phyto, NO ₃ , PO ₄ , O ₂ , pH, pCO ₂ , primary production, NH ₄ , SiO ₂ , Alk, DIC, CO ₂ air-sea flux, Zd, PFTs, ZOO,	0.0625° x 0.0625° x 73 vertical levels 0.042° x 0.042° x 125 vertical levels	BFM2 BFM5.3 Off-line coupled with NEMO	GEBCO	ECMWF (physics) pCO ₂ air conc Nutrients deposition: clim; dry/wet dep (CAMs and literature);	Atlantic Ocean: WOA GLODAP Dardanelles strait: as a river (clim); OLBC (literature)	8 rivers from Ludwig et al., 1996 39 rivers from EFAS and PERSEUS	3DVarBio: OC Chla + BGC-Argo NO ₃ , O ₂ , PO ₄ + AlNO ₃	Radiative transfer model + CDOM	Daily and monthly 2020-2024 7-day forecast 10-day forecast	Class 1: L3 sat for chl, reflectance, Kd, SOCAT for pCO ₂ , Emodnet for NO ₃ , PO ₄ , pH, Alk, DIC, L4 sat for primary production, TARA/MARE MED for PFTs



					ECMWF (optical variables)	e)				st	Class4: L3 satellite for chl, PFTs, Kd; BG ARGO for O ₂ , NO ₃ , Chl, pH, PhytoC,
NWS (NWS-MFC) 2024-2025	Chla, Phyto, NO ₃ , PO ₄ , O ₂ , pH, pCO ₂ , Kd,	0.027° x 0.027° , 50 vertical levels	Online coupling to NEMO 3.4 / 3.6 via FABM	GEBCO + additional data from NOOS partners	Met Office Numerical Weather Prediction model, 1/3 h	Met Office operational FOAM 1/12° deep ocean model for physics, WOA for BGC.	E-HYPE Climatology of daily values (multiple data sources)	None	Carbonate system (delivering pH, pCO ₂)	2021-2024	Satellite chla ICES nitrate, phosphate, oxygen, SOCAT pCO ₂ .
2015-2023 and from December 2025	Chla, Phyto, NO ₃ , PO ₄ , O ₂ , Kd, net PP pH, pCO ₂	0.111° x 0.067° (approx 7 km square), 24 vertical levels						None		daily	

635

Table A2: Characteristics of the MY MFC-BIO. In black: state of the system in 2015 and in red: EIS2024.

PRODUCTS	Horizontal and vertical resolution	Coupling with the physics (online/offline, one-way, two-way, code version)	Bathymetry (source and resolution)	Atmospheric forcing Product, frequency, resolution	Lateral boundaries	Rivers (number, dataset, frequency, other)	Data Assimilation : Code, Scheme, Assimilated Obs, corrected Obs, assim. cycle	Model Formulation (new variables, new processes)	Products, Frequency of outputs	Validation Metrics OBS	
ARCTIC (ARC-MFC)	Chla, Phyto, PP, ZOO, NO ₃ , PO ₄ , O ₂ , SiO ₂ Kd	25*25 km ² vertical levels	Online 1-way coupling with HYCO M2.2-Online 1-way coupling	GEBCO	ERA-Interim 6h (https://doi.org/10.1002/qj.828) 0.25 degree (31km) ERA-5 6h (https://doi.org/10.1002/qj.3803)	WOA Climatology for temperature, salinity and nutrients (https://doi.org/10.7289/v5f769)	TRIP runoff (https://hydro.iis.u-tokyo.ac.jp/~taikan/TRIPDATA/TRIPDATA.html) combined with nutrients flux from Global NEWS (annual climatology: https://doi.org/10.1016/j.envsoft.2010.01.006)	DEnKF with parameter estimation. Satellite chlorophyll, temper	NORWECOM ECOSMO II Variable Chl:C ratio, phytoplankt	2007-2010 One-time production, monthly output 2007-	Class 1 Satellite chl Class 2 In-situ time series stations Regional bias, RMSD for Chl, and



		vertical levels	with HYCO M2.2 via FABM		0.25 degree (31km)		Z https://doi.org/10.1029/2009GB003587	ature, sea level anomaly. Weekly DEnKF in a one-lag smoother. setting with parameter estimation. Satellite chlorophyll, In-situ nutrients, (NO3, PO4, SiO3) 8 daily	on growth formulation	2022 Extended annually Daily, monthly, yearly output	Nutrients. Regional Chl-time-series
BALTIC (BAL-MFC)	Chla, PHY, NO3, PO4, NH4, O2, pH, pCO2, Secchi depth	3.7* 3.7 km, 83 vertical levels, 2*2km, 56 vertical levels	Scobionline coupled with the Rossby Center Ocean Circulation Model. ERGO M Online coupled with NEMO4.0	GEB CO	ERA-40 (25km resolution) ERAS	Climatology	E-Hype for runoff and nutrients Monthly climatologies	None Oxygen, nitrate and phosphate profiles (SHARK database)	Carbon cycle	1970-1999 1 Jan 1993 to 31 Dec 2021	ICES in-situ data (HELCOM monitoring stations) L3 Satellite data (CHL, KD)
BLACK (BLK-MFC)	Chla, Phyto, NO3, PO4, O2, PP, pH,	15 km, 31 levels, 2.5 km,	Online coupled with GHER Online coupled		ERA Nutrients deposition (averaged field from atmospheric ERAS	Bosporus	6 Sesame, PERSEUS Reconstructed data sets (annual mean values+ typical seasonal cycle)	None Assimilation of multisensors L3		1992-2015, weekly averaged 1992-	Class2: • In-situ: for nutrients, oxyg



	Alk, DIC, pCO2, air-sea flux, Zd, SiO2, ZOO, ODU	59 levels	with NEMO 4.2.					chlorophyll-a products (but no assimilation in the climate MY)		2022, daily averaged 1950-2023, daily averaged	<ul style="list-style-type: none"> en BGC ARGO (oxygen)
Global (GLO-MFC)	Chla, Phyto, PP, NO3, PO4, O2, SiO2, Fe, pH, pCO2	0.25 * 0.25, 75 vertical levels Same in 2024	2015: NEMO V3.1- PISCE Sv2 2024: NEMO v3.6 PISCE Sv2 Offline coupling		ECMWF (daily) Dust and N atmospheric deposition (monthly/annual climatologies)		climato	No DA (neither BGC, nor PHY)		1993-2023 (31/12/2022) Daily and monthly	Bias, RMS E, Correlation Obs: Satellite, WOA/GLODAP climato, BATS, SOCAT-based NN product
Iberian (IBI-MFC)	Chla, Phyto, PP, NO3, PO4, NH4, O2, SiO2, Fe, Zeuph, pH, pCO2, DIC, ZOO C, Kd	1/12° 75 vertical levels	PISCE Sv2 Online coupling with NEMO v3.4 NEMO 3.6	GEB CO	Monthly and annual climatology for atmospheric Dust and Nitrogen depositon	GLO MY for physics and the 24 PISCES variables	Climato from Ludwig et al. (1996), GlobalNews2, Additional NO3, PO4 from EEA	DA of physical variables; No BGC DA		2002-2014 1993-2023 Daily and monthly	Satellite WOA climato GLODAP BGC-Argo ICES, EMODnet MAREDA T
Mediterranean (MED-MFC)	Chla, Phyto, NO3, PO4, O2, pH, pCO2, primary production, CO2	0.06 25° x 0.06 25° x 73 vertical levels 0.04 2° x 0.04	BFM2 BFM5.0 Off-line coupled with NEMO	GEB CO	ERA-interim (physics) ERA5 (physics) pCO2 air conc Nutrients deposition: clim;	Atlantic Ocean: WOA GLODAP Dardanelles strait: as a river (clim); OLBC (literature)	8 rivers from Ludwig et al., 1996 39 rivers from PERSEUS	3DVar Bio: OC Chla		1999-2024 Daily and monthly	Class 1: L3 sat for chl, SOCAT for pCO2, Emodnet for NO3, PO4, pH,, Alk, DIC, Class4: L3 satellite for chl, BG ARGO for



	air-sea flux NH ₄ , DIC, Alk	2° 125 verti cal level s																	O ₂ , NO ₃ , Chl
NWS (NWS- MFC)	Chla, Phyto, PP, NO ₃ , PO ₄ , O ₂ , pH, pCO ₂ , Kd, PFT chl	0.11 1° × 0.06 7° (appr ox 7 km squar e), 24 verti cal level s (no chan ge for 2024)	Online coupli ng to NEMO 3.4 NEMO 3.6 via FABM	GEB CO + additi onal data from NOO S partn ers	ERA5	Atlantic: Met Office GloSea for physics, WOA for BGC Baltic: BAL reanalysis	Daily time series, multiple data sources	None Satelli te PFT chla NEM OVA R, daily 3DVA R.	Carbo nate syste m (delive ring pH, pCO ₂)	1993- 2024 daily	Satellite chla Satellite chla and Kd490, ICES and North Sea Biogeoche mical Climatolo gy nitrate, phosphate , oxygen, SOCAT pCO ₂ , GLODAP pH								

Table A3: Number of MFC delivering the mentioned products in 2015 and 2024.

		Chlorophyll	Phytoplankton	Zooplankton	PFTs	Nitrate	Phosphate	Silicate	Ammonium	Iron	Oxygen	PP	pH	pCO ₂	DIC	Alkalinity	CO ₂ air-sea flux	Light	POC	H ₂ S/ODU
NRT	2015	8	8	2	1	8	8	3	2	4	8	8	3	3	0	0	0	4	1	1
	2024	8	8	6	2	8	8	6	1	4	8	8	8	7	5	3	2	8	1	2
MY	2015	8	8	1	0	8	8	4	3	3	8	8	2	2	0	0	0	3	0	0
	2024	8	8	4	1	8	8	5	4	3	8	8	5	7	4	2	2	7	0	1

640

Table A4: Characteristics of the MFC-BIO: horizontal and vertical resolutions (resp. H. Res and V. levels), coupling with the physics, coupler in NRT (left) and MY (right). All the models are coupled with NEMO, except ARC that is coupled with HYCOM and in 2015, BAL AND BLK that were coupled with respectively HBM(NRT)/RCO(MY) and GHER.



NRT	H. Res. (km)		V. levels		Coupling with physics		Coupler
	2015	2024	2015	2024	2015	2024	2024
ARC	12	6.25	12	40	Online	Online	FABM
BAL	2	2	25	56	Online	Online	FABM
BLK	5	2.5	31	59	Online	Online	
GLO	27.5	27.5	50	50	Offline	Offline	
IBI	3.1	3.1	50	50	Online	Online	
MED	6.9	4.6	73	125	Offline	Offline	
NWS ¹	7	2.9	24	50 ¹	Online	Online	FABM

MY	H. Res. (km)		V. Levels		Coupling with physics		Coupler
	2015	2024	2015	2024	2015	2024	2024
ARC	25	25	12	40	Online	Online	FABM
BAL	3.7	2	56	56	Online	Online	
BLK	15	2.5	31	59	Online	Online	
GLO	27.5	27.5	75	75	Offline	Offline	
IBI	9.25	9.25	75	75	Online	Online	
MED	6.9	4.6	73	125	Offline	Offline	
NWS	7	7	24	24	Online	Online	FABM

645

Table A5: Characteristics of the NRT MFC-BIO: storage space, simulation time, computational resources used. (NA= Not Available. For the NWS-NRT, 2015 and 2024 cannot be compared due to a change of producer.)

NRT	Storage space Daily uploads in GB		Simulation time for one day [minutes]		Computational resources (no. of cores).	
	2015	2024	2015	2024	2015	2024
ARC	NA	5.8	4.5	4	512	470
BAL	NA	1.3	NA	12	NA	608
BLK	0.015	1.5	6	6	16	192
GLO	1.2	1.4	1	2	NA	890
IBI	2017: 4	6	19	10	1032	1032
MED	1	7	8	8	48	240
NWS	NA	NA	NA	NA	NA	NA

¹ Since November 2025, the NWS system is again 7 km horizontal resolution with 24 vertical levels.



650

Table A6: Characteristics of the MY MFC-BIO: number of years simulated, storage space, production time, computational resources used. (NA= Not Available).

MY	Years simulated		Storage space (in TB)		Production time per year simulated (days)		Computational resources (no. of cores)	
	2015	2024	2015	2024	2015	2024	2015	2024
ARC	2007-2014	2007-2022	0.2	0.5	1	3	320	512
BAL	1970-1999	1993-2024	NA	10.4	NA	3.5		630
BLK	1992-2014	1980-2023	0.3	12	6	1.5	8	192
GLO	1998-2014	1993 - 2023	23	8.4	NA	0.7	NA	600
IBI	2002-2014	1993-2023	4	10	4	1	9	330
MED	1999-2012	1999-2021	1	40	2	3	48	480
NWS	1998-2013	1993-2024	NA	9	NA	5	NA	432

655

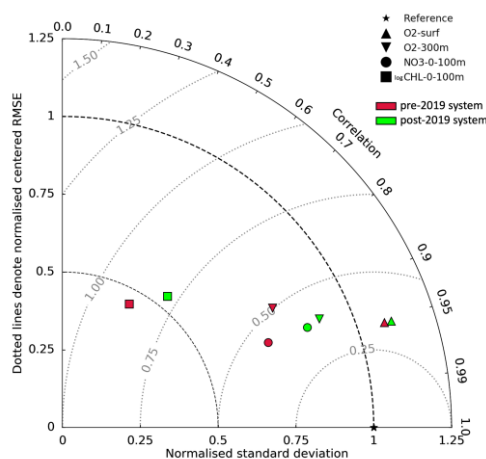


Figure 12: Taylor diagram showing the impact of the introduction in 2019 of satellite chlorophyll assimilation in the GLO NRT system for 4 metrics: O2 at sea surface (up triangles), O2 at 300m depth (down triangles), NO3 in the 0-100m surface layer (circles) and log(chl) in the 0-100m surface layer (squares).

660



665 **Table A7: Link with observations for DA, validation, differentiate 2015 to 2024**

	Satellite	BGC ARGO	Mooring	Ship	Gliders	Others	Metric used for the validation	Independent data (non-assimilated data) to validate the system
	<ul style="list-style-type: none"> Product NRT(n)/DT(n) Use (assimilation, validation) Validation Class 	<ul style="list-style-type: none"> Product NRT(n)/DT(n) Use (assimilation, validation) Validation Class 	<ul style="list-style-type: none"> Product NRT(n)/DT(n) Use (assimilation, validation) Validation Class 	<ul style="list-style-type: none"> Product NRT(n)/DT(n) Use (assimilation, validation) Validation Class 	<ul style="list-style-type: none"> Product NRT(n)/DT(n) Use (assimilation, validation) Validation Class 	<ul style="list-style-type: none"> Product NRT(n)/DT(n) Use (assimilation, validation) Validation Class 	Bias, RMS, NS, ...	
A R C	Surface chlorophyll (GlobColour 8-day averaged), Ensemble data assimilation in DT Class 1 validation in DT ESA OC-CCI v6 8-day composites in DT for assimilation and validation	None Chlorophyll NRT Class 4 validation	None None	Nutrient profiles. Assimilation in DT Class 4 validation in NRT and DT	None None	Synthetic chlorophyll profiles Assimilation (nudging) in NRT Synthetic surface DIC and CO2 flux ((https://doi.org/10.48670/moi-00047) Validation in NRT	Bias, RMSE	BGCA rgo, Ship-based nutrient profiles
G L O	Chlorophyll L4 (4km) multi-satellites	“Native” Chla, NO3, O2, pH + “NN-Extended”:	BATS => Validation DT	pCO2 and pH from SOCAT-based Neural Network product from Copernicus MOB-TAC (MULTIOBS_GLO_BIO_CARBON_SURFACE_REP_015_008)	None	None	Bias, RMSE, Correlation	BATS BGCA rgo, SOCAT-



	2014: Validation for NRT and DT 2024: data assimilation , NRT (coarsened 4km=>24k m)	NO3, PO4, Si, DIC, TAlk, pH => Validation NRT		=>Validation NRT and DT MAREDAT climato (Moriarty and O’Brian, 2013) for zooplankton => Validation NRT				based NN produc ts, MARE DAT zoopla nkton climat ology
B A L	L3 satellite data Chlorophyll and KD for validation	-	-	HELCOM monitoring stations for validation Ferrybox data (SOCAT) for pCO2 validation	-	SHARK data for profile assimilation	Bias, RMS D. correl ation	
BL K	Chlorophyll L3 (1km), multi- satellites Data assimilation in NRT and MY Correction in the surface layers of chloro and phytoplankt on groups	Dissolved Oxygen • as si mi lat io n (N R T) Chla • va lid ati on PAR and Kd • va lid ati on	None	WOD profiles MY Validation	none	NATO database Nutrients (nitrate, phosphate) • validat ion	Bias RMS Correl ation standa rd- deviati on- ratio (SDR) Nash- Suttlif fe efficie ncy (N-S)	BGC- Argo, in situ databa ses
IB I	Chlorophyll , Zeuph NPP Kd	Chla, O2, NO3		ICES bottle data		WOA GLODAP MULTIOBS MAREDAT		



M E D	Chlorophyll : L3 (1km), multi- sensors assimilation and validation in NRT and MY Kd and reflectance: validaiton in NRT	Chla, NO3, O2: Assimilation and validation in NRT, validatioin in MY PhytoC, pH: validation in NRT NN- reconstructio n NO3: assimilation in NRT Classs4 and process oriented metrics		Bottle data in EmodNet and other datasets for Alk, DIC, NO3, SiO2, NH4, Chl, pCO2, pH SOCAT for pCO2 Class 4 metrics		WOA, SOCAT, GLODAP, MAREDAT, CAFE and TARA	Bias, RMS D, C correl ation	
N W S	Chlorophyll and sediment L3 for validation (Ifremer OC5, Globcolour) Chlorophyll L3 (4km), multi- satellite; total chl for NRT, PFT for MY; assimilation and validation		Chlorophyll from 3 buoys, nitrate from 2 buoys (validation, class 2) Chlorophyll and nitrate from 1 buoy (validation, class 2)	WOA climatology, validation, class 1 North Sea Biogeochemical Climatology, GLODAP climatology, SOCAT fCO2 measurements and ICES bottle data for nutrients, oxygen; validation class 1 and class 4			Bias, RMSE Bias, RMSE , MAD, Spear man correl ation, target diagra ms	

<https://doi.org/10.5194/egusphere-2026-813>

Preprint. Discussion started: 21 April 2026

© Author(s) 2026. CC BY 4.0 License.



	Kd490, validation							
--	----------------------	--	--	--	--	--	--	--



Code, data, or code and data availability

The CMEMS biogeochemical products for the 7 regions predicted in NRT and MY are openly available at the marine
670 Copernicus Service website (<https://marine.copernicus.eu>). The modelling systems and quality assessment of the delivered
products are available on the same website in dedicated documents, respectively, the Product User Manual (PUM) and
Quality Information Document (QUID).

Author contributions

675 MG wrote the manuscript with contributions from all the co-authors. The coauthors are running the biogeochemical part of the
MFCs: MG, LV for BLK, GC, for MED, EG for IBI, SK for NWS, JL, CP for GLO, AS, TW for ARC, HM, LM for BAL.
CV is responsible for the downstream service organisation, KV for the Ocean State Report in Mercator Ocean International.
All co-authors have reviewed, discussed, and agreed on the submitted version of the manuscript.

Competing interests

680 The co-authors have no potential conflicts of interest.

Disclaimer

Copernicus Publications adds a standard disclaimer: “Copernicus Publications remains neutral with regard to jurisdictional
claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While
Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.
685 Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.”
Please feel free to add disclaimer text at your choice, if applicable.

Acknowledgements

For running BLK-MFC MG and LV benefit from the computational resources have been provided by the Consortium des
Équipements de Calcul Intensif en Fédération Wallonie Bruxelles (CECI) funded by the Fonds de la Recherche Scientifique
690 (F.R.S.-FNRS) under convention 2.5020.11 . For running ARC-MFC TW and AS acknowledges the computational resources
from (nn9481k and nn2993k) and storage (NS9481K and NS2993K) from the Norwegian Sigma2 infrastructure.



Financial support

All the co-authors acknowledge the support from the Copernicus Marine Service CMEMS, the Horizon Europe NECCTON project under grant agreement no. 101081273, MG acknowledge the support from the Horizon Europe BioGeoSea project
695 (under grant agreement no. 101216427) and RIVIERADE (under grant agreement no. 101181983).

Review statement

The review statement will be added by Copernicus Publications listing the handling editor as well as all contributing referees according to their status anonymous or identified.

700 References

- Ardyna, M., Babin, M., Gosselin, M., Devred, E., Bélanger, S., Matsuoka, A., and Tremblay, J.-É.: Parameterization of vertical chlorophyll *a* in the Arctic Ocean: impact of the subsurface chlorophyll maximum on regional, seasonal, and annual primary production estimates, *Biogeosciences*, 10, 4383–4404, <https://doi.org/10.5194/bg-10-4383-2013>, 2013.
- Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., and Gehlen, M.: PISCES-v2: an ocean biogeochemical model for carbon and
705 ecosystem studies, *Geosci. Model Dev.*, 8, 2465–2513, <https://doi.org/10.5194/gmd-8-2465-2015>, 2015.
- Aumont, O., E. Maier-Reimer, S. Blain, and P. Monfray, 2003. An ecosystem model of the global ocean including Fe, Si, Pcolimitations, *Global Biogeochem. Cycles*, 17(2), 1060, doi:10.1029/2001GB001745.
- Bakker, D. C. E., Pfeil, B. Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., Smith, K., Cosca, C., Harasawa, S., Jones, S. D., Nakaoka, S., Nojiri, Y., Schuster, U., Steinhoff, T., Sweeney, C., Takahashi, T., Tilbrook, B., Wada, C., Wanninkhof, R.,
710 Alin, S. R., Balestrini, C. F., Barbero, L., Bates, N. R., Bianchi, A. A., Bonou, F., Boutin, J., Bozec, Y., Burger, E. F., Cai, W.-J., Castle, R. D., Chen, L., Chierici, M., Currie, K., Evans, W., Featherstone, C., Feely, R. A., Fransson, A., Goyet, C., Greenwood, N., Gregor, L., Hankin, S., Hardman-Mountford, N. J., Harlay, J., Hauck, J., Hoppema, M., Humphreys, M. P., Hunt, C. W., Huss, B., Ibáñez, J. S. P., Johannessen, T., Keeling, R., Kitidis, V., Körtzinger, A., Kozyr, A., Krasakopoulou, E., Kuwata, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lo Monaco, C., Manke, A., Mathis, J. T., Merlivat, L., Millero,
715 F. J., Monteiro, P. M. S., Munro, D. R., Murata, A., Newberger, T., Omar, A. M., Ono, T., Paterson, K., Pearce, D., Pierrot, D., Robbins, L. L., Saito, S., Salisbury, J., Schlitzer, R., Schneider, B., Schweitzer, R., Sieger, R., Skjelvan, I., Sullivan, K. F., Sutherland, S. C., Sutton, A. J., Tadokoro, K., Telszewski, M., Tuma, M., Van Heuven, S. M. A. C., Vandemark, D., Ward, B., Watson, A. J., Xu, S. (2016) A multi-decade record of high quality fCO₂ data in version 3 of the Surface Ocean CO₂ Atlas (SOCAT). *Earth System Science Data* 8: 383-413. [doi:10.5194/essd-8-383-2016](https://doi.org/10.5194/essd-8-383-2016).
- 720 Bell, M.J., M. Lefèvre, P.-Y. Le Traon, N. Smith, and K. Wilmer-Becker. 2009. GODAE: The Global Ocean Data Assimilation Experiment. *Oceanography* 22(3):14–21, doi:10.5670/oceanog.2009.62.



- Bleck, R., 2002. An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates, *Ocean Modelling*, Volume 4, Issue 1, 55-88, doi.org/10.1016/S1463-5003(01)00012-9.
- Brasseur, P., N. Gruber, R. Barciela, K. Brander, M. Doron, A. El Moussaoui, A.J. Hobday, M. Huret, A.-S. Kremer, P. 725
Lehodey, R. Matear, C. Moulin, R. Murtugudde, I. Senina, and E. Svendsen. 2009. Integrating biogeochemistry and ecology into ocean data assimilation systems. *Oceanography* 22(3):206–215, doi:10.5670/oceanog.2009.80.
- Bruggeman, J. and Bolding, K., 2014. A general framework for aquatic biogeochemical models. *Environmental Modelling and Software* 61: 249–265. DOI: [10.1016/j.envsoft.2014.04.002](https://doi.org/10.1016/j.envsoft.2014.04.002).
- Butenschön, M., Clark, J., Aldridge, J. N., Allen, J. I., Artioli, Y., Blackford, J., Bruggeman, J., Cazenave, P., Ciavatta, S., 730
Kay, S., Lessin, G., van Leeuwen, S., van der Molen, J., de Mora, L., Polimene, L., Saille, S., Stephens, N., and Torres, R.: ERSEM 15.06: a generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels, *Geosci. Model Dev.*, 9, 1293–1339, <https://doi.org/10.5194/gmd-9-1293-2016>, 2016.
- Chau T., Gehlen M., and Chevallier F. (2022) A seamless ensemble-based reconstruction of surface ocean pCO₂ and air–sea CO₂ fluxes over the global coastal and open oceans. *Biogeosciences*, 19, 1087–1109, 2022. <https://doi.org/10.5194/bg-19-1087-2022>. 735
- Cossarini, G., Feudale, L., Teruzzi, A., Bolzon, G., Coidessa, G., Solidoro, C., et al. (2021). High-resolution reanalysis of the Mediterranean Sea biogeochemistry, (1999–2019). *Front. Mar. Sci.* 8. doi:10.3389/fmars.2021.741486
- Daewel, U., and Schrum, C., Simulating long-term dynamics of the coupled North Sea and Baltic Sea ecosystem with ECOSMO II: Model description and validation, *Journal of Marine Systems*, Vol 119–120, pp 30-49, 740
<https://doi.org/10.5194/gmd-15-3901-2022>, 2013.
- Fennel, K., Gehlen, M., Brasseur, P., Brown, C. W., Ciavatta, S., Cossarini, G., Crise, A., Edwards, C. A., Ford, D., Friedrichs, M. A. M., Grégoire, M., Jones, E., Kim, H.-C., Lamouroux, J., Murtugudde, R., Perruche, C. and the GODAE OceanView Marine Ecosystem Analysis and Prediction Task Team (2019) Advancing Marine Biogeochemical and Ecosystem Reanalyses and Forecasts as Tools for Monitoring and Managing Ecosystem Health. *Front. Mar. Sci.* 6:89. doi: 10.3389/fmars.2019.00089
- 745 Fennel, K., Mattern, J. P., Doney, S. C, Bopp, L., Moore, A. M, Wang, B. and Yu, L. (2022), Ocean biogeochemical modelling, *Nature Reviews Methods Primers*, 2(1), doi: 10.1038/s43586-022-00154-2.
- Ford, D., Key, S., McEwan, R., Totterdell, I. and Gehlen, M., 2018. Marine biogeochemical modelling and data assimilation for operational forecasting, reanalysis, and climate research. *New Frontiers in Operational Oceanography*, pp.625-652.
- Gehlen, M., R. Barciela, L. Bertino, P. Brasseur, M. Butenschön, F. Chai, A. Crise, Y. Drillet, D. Ford, D. Lavoie, P. Lehodey, 750
C. Perruche, A. Samuelsen & E. Simon (2015) Building the capacity for forecasting marine biogeochemistry and ecosystems: recent advances and future developments, *Journal of Operational Oceanography*, 8:sup1, s168-s187, DOI: 10.1080/1755876X.2015.1022350.



- GODAE OceanView Marine Ecosystem Analysis and Prediction Task Team(2019). Advancing marine biogeochemical and ecosystem reanalyses and forecasts as tools for monitoring and managing ecosystem health. *Frontiers in Marine Science*, 6, 89, doi: 10.3389/fmars.2019.00089.
- Grégoire, M., Vandenbulcke, L., Chevalier, S., Choblet, M., Drozd, I., Grailet, J.-F., Ivanov, E., Macé, L., Verezhenskaya, P., Yu, H., Alaerts, L., Randresihaja, N. R., Mangeleer, V., Maertens de Noordhout, G., Capet, A., Meulders, C., Mouchet, A., Munhoven, G., and Soetaert, K.: The Biogeochemical Model for Hypoxic and Benthic Influenced areas: BAMHBI v1.0, EGUsphere [preprint], <https://doi.org/10.5194/egusphere-2025-4196>, 2025.
- Gutknecht E, L. Bertino, P. Brasseur, S. Ciavatta, G. Cossarini, K. Fennel, D. Ford, M. Grégoire, D. Lavoie, and P. Lehodey (2022), Chapter 9. Biogeochemical modeling, in the Guide on Implementing Operational Ocean Monitoring and Forecasting Systems, <https://www.mercator-ocean.eu/wpcontent/uploads/2022/10/22-09-26-ETOOFS-BOOK-DOI-UPDATED-AB.pdf>.
- Hernandez, F., Bertino, L., Brassington, G.B., Chassignet, E., Cummings, J., Davidson, F., Drevillon, M., Garric, G., Kamachi, M., Lellouche, J.-M., et al. (2009). Validation and intercomparison studies within GODAE. *Oceanography*, 22(3), 128-143, <https://doi.org/10.5670/oceanog.2009.71>
- Hood, R. R., Laws, E. A., Armstrong, R. A., Bates, N. R., Brown, C. W., Carlson, C. A., ... and Wiggert, J. D.: Pelagic functional group modeling: Progress, challenges and prospects. *Deep Sea Research Part II: Topical Studies in Oceanography*, 53(5-7), 459-512. <https://doi.org/10.1016/j.dsr2.2006.01.025>, 2006.
- IGST (International GODAE Steering Team). 2000. The Global Ocean Data Assimilation Experiment Strategic Plan. GODAE Report No. 6, December, 2000.
- Hyvernat, Q., Mignot, A., Gutknecht, E., Ruggiero, G., Perruche, C., Samson, G., Sauzède, R., Aumont, O., Claustre, H., and D'Ortenzio, F.: Optimizing a large number of parameters in a Biogeochemical Model: A Multi-Variable BGC-Argo Data Assimilation Approach, EGUsphere [preprint], <https://doi.org/10.5194/egusphere-2025-4369>, 2025.
- IPBES (2020) Workshop Report on Biodiversity and Pandemics of the Intergovernmental Platform on Biodiversity and Ecosystem Services. Daszak, P., Amuasi, J., das Neves, C. G., Hayman, D., Kuiken, T., Roche, B., Zambrana-Torrel, C., Buss, P., Dundarova, H., Feferholtz, Y., Földvári, G., Igbinsola, E., Junglen, S., Liu, Q., Suzan, G., Uhart, M., Wannous, C., Woolaston, K., Mosig Reidl, P., O'Brien, K., Pascual, U., Stoett, P., Li, H., Ngo, H. T., IPBES secretariat, Bonn, Germany, DOI:10.5281/zenodo.4147317
- IPCC, 2021: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*[Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, In press, doi:[10.1017/9781009157896](https://doi.org/10.1017/9781009157896).
- Leroux, S., Brankart, J.-M., Albert, A., Brodeau, L., Molines, J.-M., Jamet, Q., Le Sommer, J., Penduff, T., and Brasseur, P.: Ensemble quantification of short-term predictability of the ocean dynamics at a kilometric-scale resolution: a Western Mediterranean test case, *Ocean Sci.*, 18, 1619–1644, <https://doi.org/10.5194/os-18-1619-2022>, 2022.



- Le Traon and Dibarboure 1999. Mesoscale Mapping Capabilities of Multiple-Satellite Altimeter Missions. *Journal of Atmospheric and Oceanic Technology*, 16(9), 1208-1223.
- Madec, G. and the NEMO System Team, 2023. NEMO Ocean Engine Reference Manual, Zenodo, 790 <https://doi.org/10.5281/zenodo.8167700>.
- Macé, L., Vandenbulcke, L., Brankart, J.-M., Brasseur, P., and Grégoire, M.: Characterisation of uncertainties in an ocean radiative transfer model for the Black Sea through ensemble simulations, *Biogeosciences*, 22, 3747–3768, <https://doi.org/10.5194/bg-22-3747-2025>, 2025.
- Martin, J., D. Dumont, and J.-E. Tremblay (2013), Contribution of subsurface chlorophyll maxima to primary production in thecoastal Beaufort Sea (Canadian Arctic): A model assessment, *J. Geophys. Res. Oceans*, 118, 5873–5886, doi:10.1002/2013JC008843. 795
- Mignot, A., Claustre, H., Cossarini, G., D'Ortenzio, F., Gutknecht, E., Lamouroux, J., Lazzari, P., Perruche, C., Salon, S., Sauzède, R., Taillandier, V., and Teruzzi, A.: Using machine learning and Biogeochemical-Argo (BGC-Argo) floats to assess biogeochemical models and optimize observing system design, *Biogeosciences*, 20, 1405–1422, [https://doi.org/10.5194/bg-](https://doi.org/10.5194/bg-20-1405-2023) 800 [20-1405-2023](https://doi.org/10.5194/bg-20-1405-2023), 2023.
- Neumann, T. 2000. Towards a 3D-ecosystem model of the Baltic Sea, *Journal of Marine Systems*, Volume 25, Issues 3–4, 2000, Pages 405–419, doi.org/10.1016/S0924-7963(00)00030-0.
- Perruche C., Lamouroux J., Mamnun N., Gutknecht E., Mignot A., Bourdalle-Badie R. (2026, *in preparation*). The Copernicus global 1/4° 30-years reanalysis of the biogeochemical ocean, aka BIORYS4 (*working title*).
- Pietropolli, G., Manzoni, L., and Cossarini, G.: PPCon 1.0: Biogeochemical-Argo profile prediction with 1D convolutional networks, *Geosci. Model Dev.*, 17, 7347–7364, <https://doi.org/10.5194/gmd-17-7347-2024>, 2024. 805
- Roemmich D, Alford MH, Claustre H, Johnson K, King B, Moum J, Oke P, Owens WB, Pouliquen S, Purkey S, Scanderbeg M, Suga T, Wijffels S, Zilberman N, Bakker D, Baringer M, Belbeoch M, Bittig HC, Boss E, Calil P, Carse F, Carval T, Chai F, Conchubhair DÓ, d'Ortenzio F, Dall'Olmo G, Desbruyeres D, Fennel K, Fer I, Ferrari R, Forget G, Freeland H, Fujiki T, 810 Gehlen M, Greenan B, Hallberg R, Hibiya T, Hosoda S, Jayne S, Jochum M, Johnson GC, Kang K, Kolodziejczyk N, Körtzinger A, Le Traon P-Y, Lenn Y-D, Maze G, Mork KA, Morris T, Nagai T, Nash J, Naveira Garabato A, Olsen A, Patabhi RR, Prakash S, Riser S, Schmechtig C, Schmid C, Shroyer E, Sterl A, Sutton P, Talley L, Tanhua T, Thierry V, Thomalla S, Toole J, Troisi A, Trull TW, Turton J, Velez-Belchi PJ, Walczowski W, Wang H, Wanninkhof R, Waterhouse AF, Waterman S, Watson A, Wilson C, Wong APS, Xu J and Yasuda I (2019) On the Future of Argo: A Global, Full-Depth, Multi- 815 Disciplinary Array. *Front. Mar. Sci.* 6:439. doi: 10.3389/fmars.2019.00439.
- Sakov, P., Counillon, F., Bertino, L., Lisæter, K. A., Oke, P. R., and Korablev, A.: TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic, *Ocean Sci.*, 8, 633–656, <https://doi.org/10.5194/os-8-633-2012>, 2012.
- Samuelson, A., Hansen, C., and Wehde, H.: Tuning and assessment of the HYCOM-NORWECOM V2.1 biogeochemical modeling system for the North Atlantic and Arctic oceans, *Geosci. Model Dev.*, 8, 2187–2202, [https://doi.org/10.5194/gmd-](https://doi.org/10.5194/gmd-8-2187-2015) 820 [8-2187-2015](https://doi.org/10.5194/gmd-8-2187-2015), 2015.



- Schartau, M., Wallhead, P., Hemmings, J., Löptien, U., Kriest, I., Krishna, S., and Oschlies, A. (2017). Reviews and syntheses: parameter identification in marine planktonic ecosystem modelling. *Biogeosciences*, *14*, 1647-1701.
- Simon, E., Samuelson, A., Bertino, L., & Mouysset, S., 2015. Experiences in multiyear combined state–parameter estimation with anecosystem model of the north atlantic and arctic oceans using the ensemble kalman filter. *Journal of Marine Systems*, 825 152, 1–17.
- Skákala, J., Ford, D. A., Brewin, R. J., McEwan, R., Kay, S., Taylor, B. H., de Mora, L., & Ciavatta, S. (2018). The assimilation of phytoplankton functional types for operational forecasting in the North- West European Shelf. *Journal of Geophysical Research: Oceans*.
- Smith, N. 2001. *Report of the GODAE High Resolution SST Workshop*. October 30–November 1, 2000, available from the 830 International GODAE Project Office, Bureau of Meteorology, Melbourne, Australia (<http://www.godae.org/Project-Office.html>), 64 pp.
- Terzić, E., Miró, A., Organelli, E., Kowalczyk, P., D'Ortenzio, F., and Lazzari, P.: Radiative transfer modeling with biogeochemical Argo float data in the Mediterranean Sea, *J. Geophys. Res.-Oceans*, *126*, e2021JC017690, <https://doi.org/10.1029/2021JC017690>, 2021.
- 835 Teruzzi, A., Bolzon, G., Salon, S., Lazzari, P., Solidoro, C. and Cossarini, G., 2018. Assimilation of coastal and open sea biogeochemical data to improve phytoplankton simulation in the Mediterranean Sea, *Ocean Modelling*, Volume 132, 2018, Pages 46-60.
- Uitz, J., H. Claustre, A. Morel, and S. B. Hooker (2006), Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll, *J. Geophys. Res.*, *111*, C08005, doi:10.1029/2005JC003207.
- 840 Vichi M., Lovato T., Lazzari P., Cossarini G., Gutierrez Mlot E., Mattia G., Masina S., McKiver W. J., Pinardi N., Solidoro C., Tedesco L., Zavatarelli M. (2015). The Biogeochemical Flux Model (BFM): Equation Description and User Manual. BFM version 5.1. BFM Report series N. 1, Release 1.1, August 2015, Bologna, Italy, <http://bfm-community.eu>, pp. 104.
- Ward, B. A., Friedrichs, M. A., Anderson, T. R., & Oschlies, A. (2010). Parameter optimisation techniques and the problem of underdetermination in marine biogeochemical models. *Journal of Marine Systems*, *81*, 34-43.
- 845 WMO, 2024. State of the Global Climate, WMO-No. 136B, 42pp. <https://library.wmo.int/idurl/4/69455>
- Yumruktepe, V. Ç., Samuelson, A., and Daewel, U.: ECOSMO II(CHL): a marine biogeochemical model for the North Atlantic and the Arctic, *Geosci. Model Dev.*, *15*, 3901–3921, <https://doi.org/10.5194/gmd-15-3901-2022>, 2022.