



Deliverable D6.3

WP6 metrics and process ranking

24 June 2025/version 1.0



About this document

Title	D6.3 WP6 Metrics and process ranking
Work Package	6 Determine significance of key processes in the evolving Ocean C Cycle
Lead Partner	GEOMAR, PML
Lead Author (Org)	Iris Kriest (GEOMAR), Jerry Blackford (PML)
Contributing Author(s)	Ken Haste Andersen (DTU), Giorgio Dall’Olmo (OGS), Volkmar Sauerland (GEOMAR)
Reviewers	NORCE Norwegian Research Center
Due Date	30.06.2025, M32
Submission Date	26.06.2025
Version	1.0

Dissemination Level

<input checked="" type="checkbox"/>	PU: Public – fully open (automatically posted online)
<input type="checkbox"/>	SEN: Sensitive – limited under the conditions of the Grant Agreement

OceanICU: Improving Carbon Understanding is a Research and Innovation action (RIA) funded by the Horizon Europe Work programme topics addressed: HORIZON-CL6-2022-CLIMATE-01-02. Start date: 01 November 2022. End date: 31 October 2027.



**Funded by
the European Union**



**UK Research
and Innovation**

This work was funded by the European Union under grant agreement no. 101083922 (OceanICU) and UK Research and Innovation (UKRI) under the UK government’s Horizon Europe funding guarantee [grant number 10054454, 10063673, 10064020, 10059241, 10079684, 10059012, 10048179]. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.



Contents

1. Key messages for the Ocean ICU stakeholders	4
2. Abstract.....	4
3. Work carried out.....	5
4. Main results achieved	6
5. Contribution to the overall objectives and relevant (KPIs)	13
6. Impact and progress beyond state of the art.....	13
7. Lessons learnt and links built	13
8. References.....	13



1. Key messages for the Ocean ICU stakeholders

Not applicable.

2. Abstract

WP6 incorporates the knowledge from WP3 to WP5 into OCEANICU modelling tools (regional, basin scale and global earth system models) to resolve and to quantify the importance of biological processes for regional and global C cycles and carbon storage under a range of climate pathways. As described in the grant agreement (DoA part B), to align modelling efforts with progress made in WP3-WP5 it was decided to provide the possibility to implement and explore new parameterisations first in light-weight global or regional models (1D iGOTM, GETM, TMM, UVic). Decisions on the transfer of these parameterisations to computationally expensive ESMs would then be made on metrics that describe the efficiency of the BCP with regard to long-term carbon storage. These metrics should allow a first estimate of the impact of each process on the marine carbon sink. The decision on "what is important" should be supported by information on the effects of physical and biogeochemical (parametric) uncertainty. This can be done, for example, by examining the effects of different model setups (1D vs. 3D; impact of different circulations; impact of different biogeochemical model setups) on metrics.

We here describe the progress made so far with respect to:

- Defining scalar metrics that describe and assess the impact of process parameterisations on the BCP and carbon storage (sections 3.1 and 4.1)
- Defining impact of physical and biogeochemical model uncertainty (sections 3.2 and 4.2)
- Describing progress on implementations of new parameterisations into models (sections 3.3 and 4.3). Note that a full description of process implementation will be provided by D6.4: Report on the integration of new parameterisations from T6.3 into shelf sea and climate models (due month 44, 30 June 2026)
- Describing metrics of those processes already implemented (section 3.4)
- Report on data delivered to WP7 (section 3.5)

As noted in the implementation plan for WP6 (D6.1) "depending on the availability of new process parameterisations (milestones 22, 24 and 24) and their implementation in the (light-weight) tools, the process ranking will be carried out in an iterative, step-wise manner, and may also include first-order estimates based on already published data (discussion during Kick-off meeting 21-23 February 2023), in collaboration with partners from WP3 and WP4." In particular, the process ranking (section 3.4) will progress during the project, as more parameterisations are being developed and/or supplied by WPs 3-5. We hence consider this report as a living document.



3. Work carried out

3.1 Metrics definition

In order to define metrics not only within WP6, but with input from other partners of the consortium, WP6 organised three virtual workshops, where we discussed and collaboratively worked on a document that identifies several metrics to assess process impact on the BCP and carbon storage. In order to provide common criteria for process and model evaluation, the meetings and document also addressed the question of example stations (sites to run 1D iGOTM and/or to evaluate global or regional models) and time frames to run the model. The three-hour workshops were held on 21 April 2024, 15 October 2024, and 18 March 2025 and were attended by up to 23 participants from various WPs. Joint editing of the document continued between the meetings. The document was finally presented and discussed in a 1-hr workshop at the ASM in Sopot (21. May 2025; see report on this workshop).

3.2 Measures of model uncertainty

In addition to effects induced by the newly developed parameterisations, changes in spatial scale, resolution and forcing (as in the transition from light-weight models to coupled ESMs) might also affect the model outcome. Likewise, changes in biogeochemical rate constants, as often necessary to tune revised models for new processes or to different circulations, can have a considerable background on model performance and carbon storage. The effect of changing the physical model framework from the test environments to fully coupled ESMs has been addressed using ESM FOCI-MOPS, in direct comparison to experiments carried out with 1D GOTM-MOPS and 3D TMM-MOPS. The simple biogeochemical model has also been tested with different rate constants in different global configurations, and the results are summarised in section 4.

Parametric vs. circulation uncertainty: Earlier work has shown that the impact changes in model parameters (especially those related more or less directly to the biological carbon pump) on model performance and diagnostics can be as important as that of model structure (as defined by its equations and components; [Kriest, 2017](#)) or physics (e.g., [Kriest et al., 2020](#)). We therefore calculated uncertainty measures arising from model parameters and circulation with focus on the above-defined metrics. Figure 1 presents results from 25 global model experiments ([Kriest et al., 2020](#)) which were re-examined using selected metrics and stations defined above; additionally, we present results for the global scale. In these experiments, parameters (biogeochemical constants) of a simple biogeochemical model MOPS (also embedded in FABM and applied in OCEANICU), were calibrated at a global scale against observations of nutrients and oxygen. Calibration was carried out separately in 5 different circulations. In cross-validation experiments the model in each circulation was simulated with parameters optimised for a different circulation, resulting in a total of 25 model simulations.

From 1D-GOTM to TMM and/or FOCI-MOPS: In another set of experiments, that addresses three model tools available in OCEANICU, we examined the effects of transferring biogeochemical model MOPS from 1D GOTM (simulated at the 17 sites described above; referred to as GOTM-MOPS) to TMM (2.8°x2.8°, 15 layers, climatological; referred to as TMM-MOPS) and further to FOCI-MOPS (0.5°x0.5, 46 layers, piControl and/or historical/RCP). The 1D simulations were simulated over 20



years (2003-2023) with forcing as agreed in the metric workshops and documents, and provided by BB. MOPS in TMM was spun up over 3000 years with constant circulation and forcing (atmospheric CO₂ of 280 and 400ppm). FOCI-MOPS was spun up over 500 years with preindustrial forcing, then simulated until 2023 with historical (up until 2014) or SSP585 (until 2023) forcing (Guo/Kriest, unpublished). The configuration of MOPS - in particular, its rate constants - was the same in all simulations. To examine effects of biogeochemical model parameters in FOCI-MOPS we also compare it to the simulation by [Chien et al. \(2022\)](#), that applied different biogeochemical model parameters. We note that the simulations and comparisons are work-in-progress, being currently analysed and prepared for a publication.

3.3 Progress on process implementation

The Grant Agreement (DoA part B) in Table 1.8 identified 20 processes to be tested for their impact on carbon storage. These relate to the classical BCP, ecosystem controls and shunts (ECS) and external levers (fishing, deep-sea mining). We here only briefly report on the progress made with regard to their implementation in one of the model tools. Some of these process implementations are still underway, with many of them being carried out directly in global or regional models. To this end, we have set up a living document that keeps track of the progress ([see here](#)). Further update on this progress will be given in deliverable D6.4 (month 44, 30 June 2026).

3.4 Process metrics

As noted above, the implementation of processes is currently underway, often in large scale coupled models. To this end, we have prepared a common format and spreadsheet for the listing of most mandatory metrics defined in section 4.1 (see here) which will be updated as new model results become available. Sample scripts and tools to evaluate global models in 1D-GOTM, TMM (standalone and coupled to FABM) as well as NEMO have been prepared, and will be adapted to further configurations.

3.5 Data transfer to WP7

Data transfer to WP7 included descriptions of the model suites and tools available within OCEANICU, and a common data format is being developed for further transfer on model output, if required.

4. Main results achieved

4.1 Metrics definition

The following **scalar (numerical) values** were identified as metrics:

- (1) **Mandatory:** Gravitational (sinking) particle flux at 100m, 200m, 500m, 1000m, 2000m, burial
- (2) **Mandatory:** Export efficiency (export flux at 100m/NPP) and transfer efficiency (flux at 1000m/flux at 100m).
- (3) **Mandatory:** POC, DOC, DIC and PIC between depth horizons of (1)
- (4) **Discretionary:** The amount of DIC released (or consumed) by new process per unit time (e.g., DVM, zooplankton respiration, POC and DOC remineralisation) between depth horizons mentioned in (1).



- (5) Discretionary: Total particulate and dissolved organic carbon flux (including physical transport) at 1000m (1D and 3D, transient and climatological)
- (6) **Mandatory:** AOU
- (7) Discretionary: For a process: the amount of carbon sequestered (following the definition of [Nowicki et al. \(2022\)](#)); and the turn-over rate of that carbon. This is specific to the application of transport matrices (TMs).
- (8) **Mandatory:** For an intervention: the transient - the amount of time - it takes for this intervention to change the air-sea CO₂ flux by 1, 2, 5 10 and 20%. "Change" is here defined as the branching off of the air-to-sea flux trajectory from the undisturbed run. We note that the higher criteria (10-20% change) may require long simulation times, which may not be feasible for every model setup. During the ASM workshop it was agreed that these criteria can be regarded as discretionary.

Metrics (1) to (3) can be quite easily diagnosed from any model, and scripts to do this are available for several model types and configurations. The metrics will be calculated for a reference model configuration (without the new process) and the newly developed model configurations (including the new process parameterisation). Metrics (4) and (5) involve a larger programming effort, and their evaluation may not be feasible for every model or process parameterisation. They are hence considered discretionary. To assess carbon storage, the optional metric (6) requires assumptions about the ratio between carbon and oxygen. Metric (7) includes either the application of TMs, or calculation of additional tracers such as remineralised DIC, as well as computation of a model's steady state. This cannot be guaranteed to be feasible for all models, and the metric is therefore considered discretionary. We note that metric (8) refers to an intervention, for example the release of sediment to the mesopelagic during deep-sea mining. This metric can serve the purpose of assessing the impact of an intervention, and will possibly be useful at a later stage of this project.

Regional shelf seas experience considerable horizontal advection, varying and shallow depths, and vertical mixing, and therefore metrics will be less clear-cut. Further there is more diversity in the processes that the regional models are developing, so there is unlikely to be a consistent metric approach across all. These are the key indicators, although some model setups may not be able to cover them all. Metrics detailing carbon pools and export are in particular:

- Change in total POC in model, in particular an increase in the refractory pool
- Change in sediment carbon content / burial
- Change in net air to sea CO₂ flux
- Change in export to off-shelf (for 3D set ups only)

A further interesting indicator of change for regional shelf-seas is the change in classical vs microbial food web fluxes, e.g., bacterial carbon uptake vs zooplankton carbon uptake, or bacterial productivity vs primary productivity.

To run FABM-GOTM (1D) and for local model evaluation we identified 17 **potential test sites** with sufficient data availability, covering a wide range of ecological and biogeochemical regimes, and/or being subject to different external pressures such as deep-sea mining or trawling. The test sites are roughly classified according to their potential use for different models and processes:

For North-West European Shelf and around:

- (1) **L4** (English Channel): 50.366°N 4.148°W
- (2) **Candyfloss** (Celtic Sea) 49.4°N, 8.6°W



- (3) **Oyster Grounds** (North Sea) 54.5°N 4.5°E
- (4) **PAP** (North Atlantic): 48.83°N 16.5°W
- (5) **M** (Norwegian Sea): 66°N 2°E

Sites more relevant to global models:

- (6) **BATS** (Bermuda): 31.67°N 64.17°W
- (7) **ESTOC** (Canary Islands): 29.04°N, 15.50°W
- (8) **HOT** (Hawaii): 22°45'N 158°W

Sites for Deep Sea Mining:

- (9) **Loki's Castle** (Greenland/Norwegian Sea) 73.5°N 8°E
- (10) **Mohns Ridge** (Greenland/Norwegian Sea) 73.5°N 5°E
- (11) **Deep Insight** (Greenland/Norwegian Sea) 73°N 1°E
- (12) **GSR** (Eastern CCZ) 14.12°N 125.87°W
- (13) **BGR** (Eastern CCZ) 11.93°N 117.03°W

Other stations - use as necessary:

- (14) **DYFAMED** (Med Sea): 43.42°N 7.87°E
- (15) **K2** (North Pacific): 47°N 160°E
- (16) **S1** (North Pacific): 30°N 145°E
- (17) **SOTS** (Southern Ocean): 147°E 47°S

For all stations BB provided the necessary forcing to run 1D FABM-GOTM.

Protocols for **time frames and forcing** over which to run or evaluate the models were identified depending on model domain and setup. For the **1D model setups** we decided on the period between 2003-2023 forced by ERA5. Because 1D models can deteriorate considerably when run over a longer period, 3 potential pathways for model simulation are proposed, depending on site: (a) simulation of a full period of 20 years (2003 to 2023) without relaxation to observed values, (b) a limited simulation of 5 years (2019 to 2023) without relaxation or (c) run the model from 2003 to 2018 with relaxation to observed values, then restart from this model state to run 2019-2023 freely (without relaxation). Simulation times of the efficient 3D offline circulation **TMM** should be towards equilibrium, i.e., covering several millennia. Likewise, because the global model of intermediate complexity **UVic** can be run efficiently over long time frames with climatological/piControl/historical forcing, simulations towards equilibrium are within reach. The **3D ESMs and regional models** require different treatments, and typically come with their own spin-up strategy and setup: **NorESM** will be run in an ocean-only uncoupled framework forced with climatological atmospheric forcing to calibrate parameters toward quasi-equilibrium (100-1000s years) or (b) run in fully coupled preindustrial control (100-1000s years). For parameter calibration **CMCC-ESM2** will be run in forced ocean simulations with atmospheric climatological forcing; fully coupled (pi=control) will be carried out only if ocean-atmosphere feedbacks are important for the process parameterisation. **FOCI-MOPS** will be run for 500 years with climatological forcing, followed by parallel piControl and historical scenarios.

The full document on metric definitions, including specified space and time domains of the metrics, their units, auxiliary variables and their applicability to different model types, as well as any updates arising from new results is available [here](#).



4.2 Measures of model uncertainty

Parametric vs. circulation uncertainty: The maximum spread of model results with five fixed circulations describes the effect of parameters on the metric (left panel of Fig. 1). Keeping the parameter set fixed and exchanging the circulation reveals the maximum effect of circulation (middle panel of Fig. 1). The total spread across all simulations is shown in the right panel of Fig. 1. As could be expected, the magnitude of fluxes decreases from shallower waters to the sea floor. At a global scale circulation plays a larger role for the shallower fluxes, while in deeper waters the parametric effect is often larger, in line with findings by [Kriest et al. \(2020\)](#). Circulation also plays a large role for the distribution and deep storage of DOC, especially when the model is evaluated at specific, discrete sites in the ocean. Excluding primary production, the range due to changes in parameters (the effect of parametric uncertainty) can be as large as $466 \text{ mmol C m}^{-2} \text{ y}^{-1}$ (export production). Circulation uncertainty can lead to a variation of export production by up to $1357 \text{ mmol C m}^{-2} \text{ y}^{-1}$. Effects on organic carbon inventories are largest for DOC, with some local values being larger than 10 mol C m^{-2} .

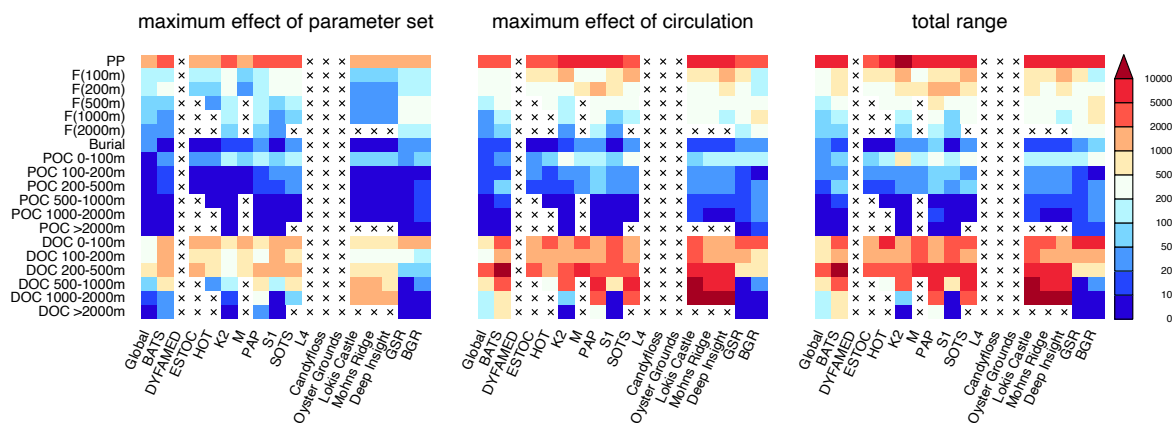


Figure 1: Maximum effect of exchanging biogeochemical model parameters (left), circulation (middle), and both (right) on metrics defined in section (4.1). Results from 25 global model simulations with TMM-MOPS ([Kriest et al., 2020](#)), after model optimisation for 5 different Transport Matrices (circulations). Following optimisation, parameter sets and circulation were swapped in 20 cross-validation experiments. Evaluation of effects follows that of Kriest et al. (2020; their Figure 11). The three panels present the metrics in different rows. Columns indicate global effect (leftmost column) followed by local effect at the different stations defined in section (3.1). Colour code shows the magnitude of the effect in $\text{mmol C m}^{-2} \text{ y}^{-1}$ (fluxes, i.e., top seven rows) or mmol C m^{-2} (following rows). An "x" denotes a station or vertical domain that is missing because at least one global model circulation does not resolve this location.

From 1D-GOTM to TMM and/or FOCI-MOPS: When comparing the 1D simulations with GOTM-MOPS to those with FOCI-MOPS we find large differences, partly with different signs for the different stations (Figure 2). At many stations FOCI-MOPS simulates larger biogeochemical fluxes and hence organic carbon storage in the deep ocean. This may be due to the considerably longer simulation time of FOCI-MOPS (750 years vs decades in GOTM-MOPS), different forcing, horizontal processes, as well as due to a higher numerical diffusivity (given the coarser vertical resolution of FOCI-MOPS).

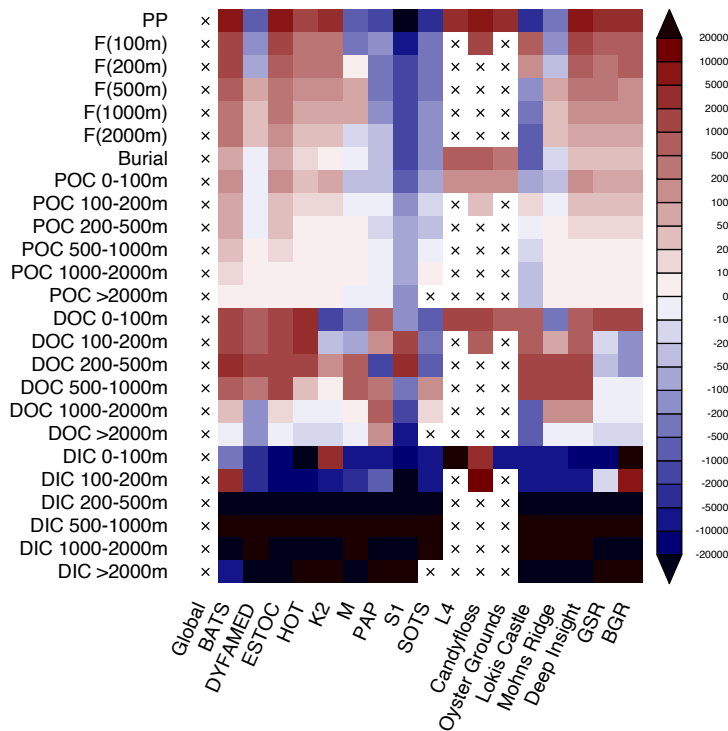


Figure 2: Metrics (1) and (3) defined in section 4.1 for the difference between FOCI-MOPS (Guo/Kriest, unpubl.; mean of 2004-2023 of piCtrl) minus 1D-GOTM (preliminary results). Columns indicate local effect at the different stations defined in section 4.1. Colour code shows the magnitude of the effect in $\text{mmol C m}^{-2} \text{y}^{-1}$ (fluxes, i.e., top seven rows) or mmol C m^{-2} (following rows). Note the logarithmic colour scale. A warm colour indicates a higher flux or concentration in FOCI-MOPS. An "x" denotes a station or vertical domain that is missing because at least one global model does not resolve this location.

Differences between FOCI-MOPS and TMM-MOPS are also considerable (Figure 3). Globally, FOCI-MOPS simulates less production and vertical mesopelagic fluxes. Likewise, DOC concentrations are less in FOCI-MOPS, despite the same biogeochemical model setup.

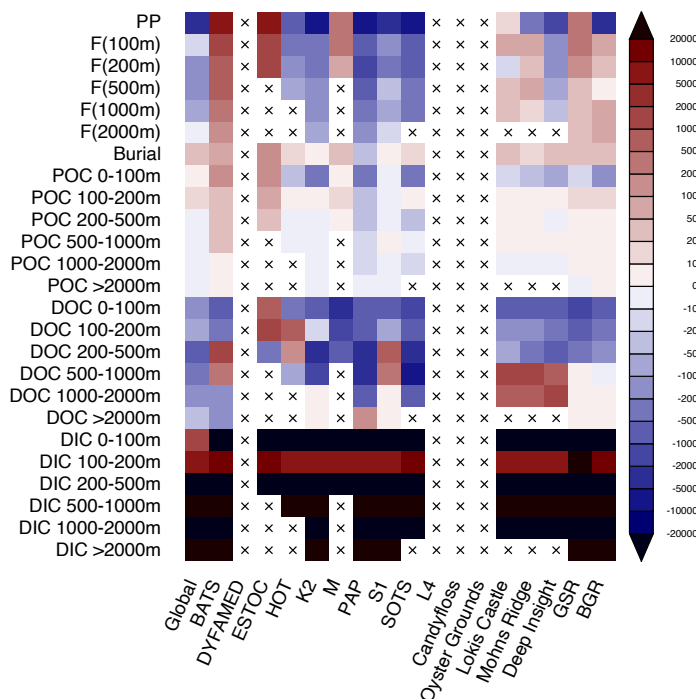


Figure 3: As Figure 2, but for the difference between FOCI-MOPS (Guo/Kriest, unpubl.; mean of 2004-2023 of piCtrl) minus TMM-MOPS (MIT2.8; climatological, forced with with 280ppm atmospheric CO₂; preliminary results). Columns indicate global effect (leftmost column) followed by local effect at the different stations defined in section 4.1.

Very large differences (on an absolute scale) in both comparisons occur for DIC (note that its background concentration is orders of magnitude larger than that of the organic tracer). These changes can also be likely attributed to different forcing by wind, atmospheric pressure, freshwater fluxes etc. in the two models. Most importantly, different spatial resolutions and resulting topography of the two models



lead to the large discrepancy in the inventories of deep domains below ~1000m. Hence, comparisons of average concentrations are more suited to compare output resulting from different model geometries. Scripts for metric evaluation provided to OCEANICU will therefore also provide the option to calculate domain averages.

To get a deeper insight into the biogeochemical vs. circulation and scale, we also examined the difference between two biogeochemical configurations of FOCI-MOPS (Figure 4). In particular, a more recent configuration of FOCI-MOPS simulates a particle flux that propagates deeper into the ocean, and a far lower production and decay rate of DOC. This leads to larger deep particle flux globally and at most sites, as well as to lower DOC concentration. While the differences in metrics for the organic tracers and fluxes are around the same order of magnitude as in Figures 2 and 3, the magnitude of changes in DIC is now comparable to those of DOC. This highlights the impact of physics, forcing and topography especially for deep inorganic tracers, but also the importance of biogeochemical parameterisations for organic tracers. The detailed analysis of this comparison is still underway, currently prepared for publication, and will be contrasted with the changes arising from a changing climate and circulation.

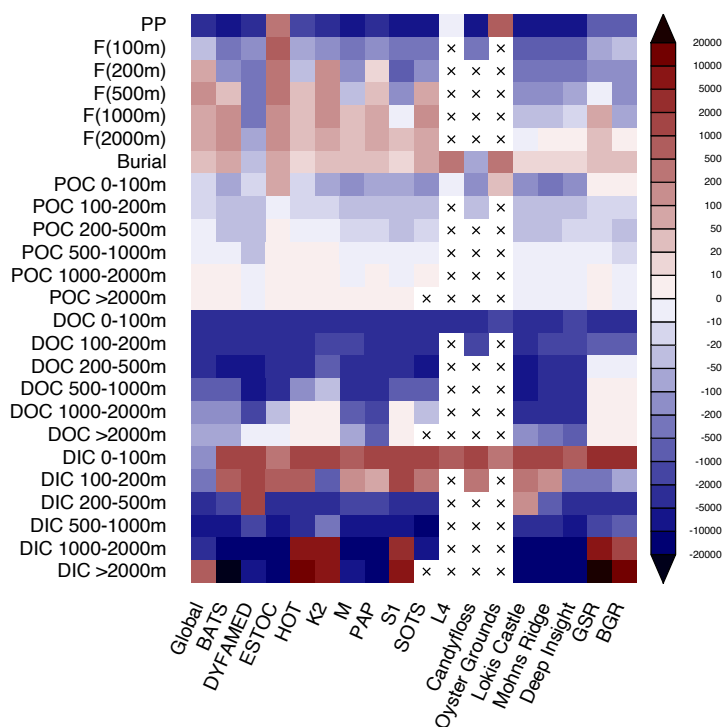


Figure 4: As Figure 2, but for differences between two different biogeochemical configurations of FOCI-MOPS (Chien et al., 2022 and Guo/Kriest, unpubl.; both mean of 2004-2023 of piCtrl; preliminary results).

To summarise, changes in model outcome may be considerable (especially for DIC) when moving a biogeochemical model from one circulation or geometry to the other (Figures 2 and 3), and differ among the sites. A large fraction of these differences can be explained by changes in forcing, spin-up time and especially spatial resolution and topography. Changes in biogeochemical model setup - especially when considering changes in particle flux - often affect the organic tracers, and are more similar among different sites (Figure 4).

The results highlight the importance of biogeochemical setup for the deep storage of organic carbon. To address the problems associated with transferring a biogeochemical model between different architectures, circulations and computational demands (see above), we are currently examining how to combine



recent innovations to accelerate model spin-up (Khatiwala, 2024) with efficient tools that allow for biogeochemical model calibration in a specific circulation (as applied in Kriest et al., 2020). This can also help to mediate between parametric uncertainty and long simulation times required, e.g., for refractory DOC (see above).

4.3 Progress on process implementation

We here only briefly report on the progress made with regard to their implementation in one of the model tools. Some of these process implementations are still underway, and many of them are carried out directly in global or regional models. To this end, we have set up a living document that keeps track of the progress ([see here](#)). Further update on this progress will be given in deliverable D6.4 (month 44, 30 June 2026).

Classical biological carbon pump: To investigate the effects of nutrient control over plankton production MOPS in TMM (potentially NEMO) has been extended by an iron cycle, and the results are currently prepared for publication. So far, data analysis showed no effect of external variables **T, pH and O₂** on particle flux (data transfer from WP3), so this has not been pursued further for implementation into large-scale models. The effect of external variables of remineralisation or **respiration** is investigated with a range of models. Various pools of DOM have been implemented directly in iHAMOCC (BLOM), in close collaboration with WP4. Variable degradability of semi-labile DOC based on age and/or bacteria is investigated with ERSEM (GOTM/NEMO). CMCC addresses the vertical production-respiration balance in BFM (CMCC). We further pursue a collaborative analysis of data (in-vitro; BGC-Argo) on ecosystem and microbial respiration together with WP3 and WP4, which will support model assessment and parameterisation. The link between phytoplankton **size** structure and particle flux profile (data transfer from WP4) is being investigated in MOPS-TMM, with potential further implementations in ERSEM (NEMO). Likewise, two size-classes of phytoplankton have been implemented in iHAMOCC (BLOM). To gain insight into the impact of **calcification** MOPS (TMM/NEMO) has been extended by a (simplified) calcite cycle, and the parameterisation of calcifiers in PISCES (NEMO) has been revised. Calcification and its dependence on various external variables (T, O₂, pH) is being investigated in detail in ERSEM (GOTM/NEMO).

Ecosystem shunts and controls: The effects of **diel vertical migration** are investigated with many models, among them ERSEM (GOTM/NEMO) and UVic. A revised version of ERSEM also addresses the effects of a stoichiometric modulation of predation. There is still ongoing data analysis on **dark C fixation**, and upcoming parameterisations will be implemented in MOPS or, if necessary, more complex BGC models. The **Whale pump** is investigated with model IMR/NEMO, as also **marine vertebrate mediated carbon** through whales. Effects of carbon flux through **fish** is being addressed by the coupling of UVic with the Feisty fish community model. This will cover carbon fluxes from respiration (including vertical migrations), fecal pellets, and **deadfalls**.

Ecosystem lever fishing and mining, trawling, dredging, drilling lever: Most of the effects of fishing are being investigated in WP5. The effect of **sediment** additions from mining has been implemented in MOPS/GOTM in collaboration between WP5 and WP6; it will likely be transferred to an extended biogeochemical model that includes higher trophic levels. Effects of **trawling** are investigated in ERSEM (GETM/GOTM/NEMO). The effects of the removal of biomass by fishing on the carbon fluxes and sequestration is explored in WP5, mainly on shelf ecosystems, but also in the open oceans.



As can be seen from the above collection, a considerable number of processes is already being implemented in either light-weight 1D (GOTM) or 3D models (TMM, UVic), but also often directly in the more time-consuming global and regional models such as BFM/NEMO, ERSEM/GETM and/or ERSEM/NEMO and iHAMOCC/BLOM. Testing and tuning the latter requires somewhat more time - however, it saves a later transition of the parameterisation between different (physical and spatial) model types, thereby avoiding disturbance from scale- and forcing-transitions noted in section 4.2. A large number of parameterisations relate directly or indirectly to the remineralisation of dissolved particulate organic matter (DOC). This addresses the large impact of some processes on the metrics for DOC noted above.

5. Contribution to the overall objectives and relevant (KPIs)

The deliverable contributes to ST4 by supporting the development and definition of consistent and coherent criteria to quantify the importance of new processes built into models. It also helps to assess the importance of processes before the background of variability induced by the (necessary) diversity of tools (model frameworks) applied in this project. Close and ongoing collaboration with several partners from other WPs, in terms of model development as well as metrics development, supports ST3 and ST4.

6. Impact and progress beyond state of the art

Not applicable.

7. Lessons learnt and links built

The close collaboration between modellers and partners from WP3-5 helps to define the model details required from the observing side, as well as the type and form of information required by modelers. There are now several (sub)working groups or task teams including oceanographers and modellers addressing specific topics such as respiration, DOC, effects of sediments, etc. This has, for example, been documented in very constructive and engaged discussions during the ASM 2025 in Sopot, and will be continued. Updates on the links built and their outcome will be given in D6.4.

8. References

- Chien et al. (2022). FOCI-MOPS v1- integration of marine biogeochemistry within the Flexible Ocean and Climate Infrastructure version 1 (FOCI 1) Earth system model. *Geosci. Model Dev.*, 15, doi:10.5194/gmd-15-5987-2022
- Khatiwala (2024). Efficient spin-up of Earth system models using sequence acceleration. *Scie. Adv.*, 10, doi:10.1126/sciadv.adn2839
- Kriest (2017). Calibration of a simple and a complex model of global marine biogeochemistry. *Biogeosciences*, 14, doi:10.5194/bg-14-4965-2017
- Kriest et al. (2020). One size fits all? Calibrating an ocean biogeochemistry model for different circulations. *Biogeosciences*, 17, doi:10.5194/bg-17-3057-2020
- Nowicki et al. (2022) Quantifying the Carbon Export and Sequestration Pathways of the Ocean's Biological Carbon Pump. *Glob. Biogeochem. Cyc.*, 36(3), doi:10.1029/2021GB007083