Project	SEAMLESS No 101004032	Start / Duration	1 January 2021/ 36 Months
Dissemination	Public	Nature	Report
Date	31.12.2023	Version	2.0

Deliverable 7.3

Roadmap towards a transfer to CMEMS

Deliverable Contributors:	Name	Organisation	Role / Title
Deliverable Leader	Laurent Bertino	NERSC	WP4 leader
Contributing Author(s)	Anna Teruzzi	OGS	Researcher
	Lars Nerger	AWI	Researcher
	Pierre Brasseur	UGA	Researcher
	Tsuyoshi Wakamatsu	NERSC	Researcher
	Jozef Skákala	PML	Researcher
Reviewer(s)	Stefano Ciavatta	MOi	Advisory Board
	Jessica Heard	PML	Project manager
Final review and approval	Józef Skákala	PML	Project coordinator

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004032.

Project	SEAMLESS No 101004032	**Deliverable**	D7.3
Dissemination	Public	**Nature**	Report
Date	31.12.2023	**Version**	2.0

Document History:

Release	Date	Reason for Change	Status	Distribution
1.0	31.12.2023	Initial document, reviewed	Released	Public

To cite this document

This document is a public deliverable. It should be cited as "Bertino L., Teruzzi, A., Nerger L., Brasseur P., Wakamatsu T., Skakala J., Heard J., and Ciavatta S. 2023. Deliverable D7.4: Roadmap towards a transfer to CMEMS. EU H2020 SEAMLESS project, grant agreement No 101004032, doi: 10.5281/zenodo.10471589.

TABLE OF CONTENTS

1. Scope

The present deliverable maps the prospects to include results from SEAMLESS into the operation of the six Monitoring and Forecasting Centers (MFC) of the Copernicus Marine Services. We consider the three Tiers of research and development defined in the Copernicus Marine Services (CMEMS), listed by increasing technical readiness level (TRL):

- Tier 3: Research and Development of TRL 4-6 requiring more than 3 years of work to reach operations in CMEMS are in the scope of Horizon Europe, ERCs and nationally funded research projects.
- Tier 2 corresponding to TRL 6-8: Operations are impacted within 2 years of work: Tier 2 is the object of Service Evolution projects financed by CMEMS as separate small-size contracts (typically funding one full-time employee for two years). See for an overview: https://marine.copernicus.eu/about/research-development-projects
- Tier 1 corresponding to TRL 9: Implementation in operations within 1 year of work. Tier-1 uptake is usually taken care of within the CMEMS MFC contracts, presented during the Design Review meeting the year before its Entry Into Service (EIS).

SEAMLESS is a Tier 3 Research and Development (R&D) activity for CMEMS and the expected outcomes should in principle lead to Tier 2 activities before reaching operations. This however does not preclude some low-hanging fruits to reach operations earlier than expected or the more unexpected difficulties to require another Tier 3 project answering a possibly better formulated scientific question.

2. Introduction

All six regional MFCs of CMEMS provide biogeochemical products both in near-real-time (NRT) forecast and multi-year (MY) reanalysis modes. They also provide CMEMS with Ocean Monitoring Indicators (OMIs: https://marine.copernicus.eu/access-data/ocean-monitoring-indicators) as well as with climate indicators defined by the World Meteorological Organisation (WMO; https://marine.copernicus.eu/ocean-climate-portal). In simple terms, these indicators are regionally or globally averaged time series of an important variable for climate assessment or a map of a long-term trend). There is a running requirement from CMEMS that the forecasts and reanalysis systems should be as close as possible to each other to use reanalyses as reference to detect anomalous situations. An additional requirement is that each MFC should be able to submit one new Ocean Monitoring Indicator to the catalogue every year.

We recall the main outcomes from each WP in SEAMLESS:

- WP2 developed the Ensemble Assimilation Tool (EAT), by implementing and testing the PDAF data assimilation software in the GOTM 1-dimensional vertical ocean mixing model, coupled via the FABM interface to all the biogeochemical models used in the regional MFCs (https://www.seamlessproject.org/SEAMLESS_EAT).

- WP3 developed a stochastic modeling framework allowing to assess the observability and controllability of essential indicators of BGC models. WP3 has coordinated the experiments with the help of the EAT framework.
- WP4 studied the feasibility of coupled assimilation into physical-BGC models.
- WP5 assessed the impact of joined satellite and in situ measurements.
- WP6 performed parameter estimation. The PARSAC package was shared among the participants to perform parameter calibration, while EAT was used for joint state-parameter estimation via data assimilation in one-dimensional configurations

Table 1 gives an overall reminder of the initial project plan in respect to the status of CMEMS BGC systems at the time of writing of the proposal, slightly updated in December 2023. It still provides three years later a correct summary of the methodological state of play in CMEMS despite the regular upgrades of the service. The sections below will discuss the stepwise transformation of the CMEMS systems that we recommend in order to incorporate the results from SEAMLESS.

Table 1: Comparison of the CMEMS MFC systems before and after the SEAMLESS developments (copied from proposal in March 2020). The main difference with the CMEMS systems at the time of writing of this report is the NWS due to Brexit, See Section 9 for more details.

System	DA method	Assimilated BGC data	PHY-BGC coupling	SAT-IS coupling	Improved parameters
NWS CMEMS	3DVAR	OC: chl	No	No	Yes
NWS SEAMLESS	Hybrid Ensemble 3DVAR	OC: chl, ops IS: chl [g], N [g], P [g]	Yes	Yes	Yes
IBI CMEMS	None	No	No	No	
GLO CMEMS	SEEK (base climatology)	OC: chl	No	No	
N Atlantic SEAMLESS	SEEK stoch. error param.	OC: chl	Yes	No	Yes
MED CMEMS	3DVarBio	OC: chl	No	No	No
MED SEAMLESS	SEIK filter	OC: chl IS:chl [a,g], N [a,g], opt [a]	No*	Yes	Yes
BAL CMEMS	ESTKF in preparation	No	No	No	No
BAL SEAMLESS	Hybrid EnKF PF	OC: chl	Yes	No	Yes
ARC CMEMS	DEnKF/EnKS	OC: chl IS: N [b], P [b], S [b]	Yes	Yes	
ARC SEAMLESS	DEnKF/EnKS	OC: chl IS: chl [a], N [a],P [b], S [b]	Yes	Yes	Yes

OC=ocean colour; IS=*In situ*; chl=chlorophyll; ops=optical data; N=nitrogen; P=phosphate; S=silicate; [a] = BGC-Argo; [b] = bottle; [g] = glider; No* = implemented in 1D only

The Ocean Monitoring Indicators from CMEMS are presently under-representing the BGC models if one compares the "Ocean Health" OMIs to the physical OMIs. https://marine.copernicus.eu/access-data/ocean-monitoring-indicators. In contrast, WP3 has identified other BGC indicators in the perspective of optimal control: are these indicators observable and/or controllable? In case the answer is negative to either of them, there is no purpose in suggesting a new model-based OMI which cannot be relied upon. The list of indicators identified by SEAMLESS is the following:

- Phenology is a set of three indicators based on 3-dimensional Chlorophyll-a concentrations (see deliverable D3.2): the value of the maximum of chlorophyll concentration in the layer 0-5m (mgChl m^{-3}), the depth of the maximum of chlorophyll during the summer period (m) and the timing of the bloom, i.e. the time of the year when the two maxima occur (day). Chlorophyll-a is a standard output of all CMEMS BGC models, so phenology indicators can technically be provided, but not if they are unreliable.
- Net Primary Production is also a usual output of BGC models.
- Particulate Organic Carbon (POC) near-bottom concentration is available from all BGC models. It has been chosen as a proxy for the POC export, an important indicator of the ocean "Carbon pump".
- Plankton Functional Types (PFTs) are not presently provided by BGC models, but this can be expected to happen in the future as satellite-based PFT products are becoming available in CMEMS.
- Trophic Efficiency is also not provided by CMEMS BGC models, but can deduced with little efforts from the usual model variables at different trophic levels.
- pH is a necessary output from the BGC models with a Carbon cycle.
- Oxygen is also required from the ocean BGC Carbon cycle models.

Whenever the above indicators are both observed and produced by the models, SEAMLESS encourages their routine validation by the operational CMEMS systems to complete the product quality dashboard.

Table 2: Summary of the CMEMS Ocean Monitoring Indicators under the category "Ocean Health" as of December 2023. Note that all the OMIs are derived from observations and none from the models.

Global	Arctic	Baltic	Black Sea	IBI	Mediterranean	NWS
pH						
Chl-a	Chl-a	Chl-a	Chl-a	Chl-a	Chl-a	Chl-a
		Cod reproductive volume*				
		O2 Inflow				
			O2 trend			

** Cod reproductive volume is a proxy based on observed ocean Temperature, Salinity and Oxygen concentrations.*

The partners of SEAMLESS are involved at different levels in the CMEMS MFC hierarchy as indicated in Table 3 below. Apart from the Entrusted Entity MOi renewed for a 6-years phase (2021-2027), all the CMEMS contracts are awarded based on open competition on a 3-years basis. The SEAMLESS proposal was submitted during the 2018-2021 contract period and terminated during the following 2021-2024 period. Due to this 3-years competitive cycle, the SEAMLESS partners are not in control of the priorities post-2025 and the Roadmap below will line up suggested updates sorted by the expected efforts needed to implement them in each MFCs relying on the principle of continuity of the service (the quality of the service is equivalent to that obtained during the previous 3-years period).

Table 3: Summary of the institutes involved in the present CMEMS MFCs and their SEAMLESS counterparts.

MFC region	**MFC Leader**	**BGC Producer**	**SEAMLESS partner**
Arctic Ocean	NERSC, No	NERSC, No	NERSC, No
Baltic Sea	DMI, Dk	BSH, De	AWI, De
Black Sea	IO-BAS, Bg	U. Liège, Be	/
Global Ocean	MOi	MOi	UGA
Iberian-Biscaye-Irish Sea	MOi	MOi	UGA
Mediterranean Sea	CMCC	OGS	OGS
North West Shelf*	UK Met (pre-Brexit), MOi (post-Brexit)	PML (pre-Brexit), MOi (post-Brexit)	PML

** The Brexit contingency in CMEMS is such that the IBI MFC also provides the NWS products instead of the UK-based producers in the period 2023-2024.*

The present Roadmap is therefore organised with one section for each MFC and subsections by increasing efforts and therefore time horizons aligned along the three Tiers of R&D defined above. We have included the uptakes that have already taken place in some of the present CMEMS MFC systems under Tier 1 for discussion of other MFCs.

Table 4: (copied from Deliverable 6.1 "guidelines for delivery", table 5.1) Summary of performance of the final SEAMLESS system configurations across the different MFC domains. The Table shows the total readiness level (TRL), the cost of using the SEAMLESS configuration, qualitatively evaluated on a scale ``very high, high, medium, low, very low'', and the quality of the target indicator product evaluated on a scale ``very good, good, medium, poor, very poor''.

Indicator	**ARC**	**BAL**	**GLO/IBI**	**MED**	**NWS**
System TRL	TRL8*	TRL5	TRL4	TRL5	TRL5
Simulation cost	high	high	high	high	high
Phenology	good	good	very good	good	very good
Net primary production	medium	good	good	good	good
POC export	n/a	n/a	medium	good	medium
Trophic efficiency	medium	medium	medium	medium	medium

PFTs	good	medium	medium	good	poor
Bottom oxygen	n/a	very poor	n/a	n/a	good
pH	n/a	medium	n/a	n/a	poor

**: The TRL is as high as 8 for the Arctic MFC because the developments are done directly on the reanalysis system.*

3. Roadmap for the Arctic MFC

The Arctic MFC has a different strategy for NRT and reanalysis products than other MFCs. The reanalysis uses an advanced ensemble DA method, namely the one-lag EnKS. Differently, since the NRT forecast runs at higher resolution, it cannot afford the propagation of a dynamical ensemble yet, and it assimilates satellite retrievals of Chlorophyll-a using heuristic rules for the composition of the phytoplankton biomass and vertical interpolation. In SEAMLESS, our efforts have therefore targeted the reanalysis DA system as the worthiest of research and developments. In the following we specify how the outcomes from the WPs of SEAMLESS can underpin the future development of the Arctic MFCs through the different Tiers.

Tier 1

From WP2: FABM was already being used to couple HYCOM and ECOSMO-II in the Arctic MFC before the start of SEAMLESS. Crucially, in addition to FABM, the PARSAC framework used in SEAMLESS has now also been used in the Arctic MFC to optimise the static parameters of the latest NRT ECOSMO-II forecasting system. This has come in operations of the CMEMS NRT forecast product ARCTIC_ANALYSISFORECAST_BGC_002_004 (DOI: https://doi.org/10.48670/moi-00003) on the November 2023 Entry into Service and was not foreseen initially when drawing the plans from SEAMLESS.

Other results from SEAMLESS will primarily impact the Arctic reanalysis, which is using an ensemble data assimilation technique.

- The work done in WP3 to fine-tune the perturbations of forcings and ECOSMO-II parameters will come into service in the November 2024 release of the Arctic MFC reanalysis.
- From WP4: The upcoming BGC reanalysis planned for 2024 will use weakly coupled physical-to-biogeochemical assimilation with the EnKS (presently there is no physical DA in the BGC reanalysis).
- From WP6: NERSC will re-evaluate the regionalisation of parameter estimation with the EnKF/EnKS (See Table 4 in Introduction).

Tier 2

NERSC considers developing a Tier 2 follow-up project of WP5 on multiplatform assimilation, with an optimised combination of satellite and in situ BGC-Argo data.

The motivation is that the multiplatform assimilation experiment with the ARC MFC system conducted in WP5 assimilated both satellite and BGC Argo float data with the following observation error settings: additive noise with identical standard deviation (defined as 30% of the associated observed value of chlorophyll concentration). This experiment did not reveal synergies between the assimilation of satellite in situ data as the assimilation of in situ data worked against the effect of satellite data. It remains to be seen whether the lack of synergy is caused by inconsistencies between the two dataset or by deficiencies of the data assimilation setup.

We consider that the additional experiments would require a few months of full-time work and would therefore not fit within the Tier-1 yearly updates but could realistically be integrated in a Tier-2 type of project either as an adjustment of the input data or of the data assimilation settings.

A new Ocean Monitoring Indicator (OMI) on phytoplankton phenology can be provided based on model reanalysis and reprocessed observations. Reanalysis data of phytoplankton phenology together with its associated physical and biogeochemical environmental factors would benefit wider research and operational communities by extending our understanding mechanism of the phytoplankton seasonal cycle and inter annual variability further and by providing long term monitoring on the upper ocean biological carbon pump. In order to deliver phytoplankton phenology as an ecological indicator from reanalysis products, further discussion on its common definition and methodology would be required. For example, evaluation of approaches proposed for determining key parameters that defines phytoplankton phenology: (1) the time of initiation; (2) the time of maximum amplitude; (3) the time of termination; and (4) the duration, from satellite and float observations (e.g., Racault et al., 2012) would be useful as a first step.

Tier 3

The following methodological evolutions can be taken up at NERSC within 3 years of work in Tier 3 follow-up projects:

- The difference of **data assimilation methods** between the reanalysis (EnKS) and NRT forecasts (nudging) is likely to persist as long as the computational resources are insufficient to run ensembles of coupled PHY-BGC models in near real-time. Long-term efforts will therefore concentrate on super-resolution Machine Learning techniques to connect the low-resolution-advanced-DA with a high-resolution system with cheaper DA powered by machine learning (see Barthélémy et al. 2022 for a proof of concept). This research is presently pursued within the Tier 3 Horizon Europe NECCTON project led by S. Ciavatta at MOi.
- NERSC has implemented the PDAF system for the sake of the coordinated 1-dimensional experiments in WP2 and WP3. PDAF also allows the assimilation into physical and sea ice models, so an uptake of **PDAF as a data assimilation engine** for the whole Arctic MFC is in principle possible within a few years of work. However, NERSC does not consider it as a high-priority innovation, since the existing DA code developed at NERSC already includes the functionalities of PDAF necessary for that task.

- Over the long term, NERSC needs to follow the evolution of the Arctic **ocean modelling engine**, which should remain in hybrid vertical coordinates to maintain the continuity of the service. Whichever the ocean model of choice to succeed to HYCOM, the FABM framework will be kept for an efficient coupling to ECOSMO-II.
- The biogeochemical model developments are all planned to take place in ECOSMO-II, including those undertaken in NECCTON. Here again, the **tuning framework** of ECOSMO-II **parameters** from WP3 will be re-used along with the regionalization of the parameter values suggested after the analysis of the impact of the parameter tuning in WP6.
- The high sensitivity of the depth of the subsurface phytoplankton blooming to Photosynthetically Active Radiation (PAR) suggests that PAR-related BGC parameters are key to estimate properly the subsurface chlorophyll maximum (SCM) depth in the Arctic reanalysis system. The sensitivity to optical parameters also emerged in the other models and regions investigated in SEAMLESS. In order to assimilate satellite in-water inherent optical properties (IOPs) data, the implementation of an optical module to ECOSMO-II can be exploited in collaboration with PML. The SCM is a common feature in the Beaufort Sea and other margins of the Arctic Ocean, so the tuning of related parameters will be explored along with a weakly-coupled data assimilation of sea ice concentration data. The 1-dimensional tools developed in WP2-to-WP4 can be re-used effectively to that purpose (i.e., GOTM, FABM, EAT, PARSAC and PDAF). Another potential benefit of including an optical module in ECOSMO-II is a more efficient assimilation of satellite ocean colour data in areas such as the Kara Sea, where the satellite data are contaminated by coloured dissolved organic material (CDOM) and therefore removed from the assimilation.

4. Roadmap for the Baltic MFC

The Baltic MFC is using a coupled NEMO-ERGOM model to simulate the physical-biogeochemical processes in the Baltic Sea. This MFC is using a fixed-basis version of the PDAF software from AWI to assimilate physical data in NEMO, while no biogeochemical data are being assimilated in ERGOM. SEAMLESS used PDAF to assimilate biogeochemical data in ERGOM, by means of a dynamic members LESTKF method. This method could easily be transferred from SEAMLESS to the Baltic MFC by exploiting the common PDAF framework .

Note the AWI is not formally involved in the Baltic MFC. Hence the statements here on the transferability of methods are from the viewpoint of SEAMLESS, rather than from the Baltic MFC priorities. Nonetheless, SEAMLESS have been engaging with the Baltic MFC to promote such transfer during the development of the project (see Deliverable 7.4).

Tier 1

The Baltic MFC has already started to use the assimilation code including the new PDAF Observation Module Interface (PDAF-OMI) based on AWI in their test system (V. Huess, DMI, pers. comm., DR meeting Oct 2023), which is paving the way towards BGC assimilation with the same code. The code can also be used with a fixed ensemble that varies over time based on previous model snapshots, so that the computing requirements are not significantly increased.

Tier 2

Within the 2-year time frame of Tier 2, the coupled data assimilation of SST and Chl for the BGC products, as suggested by WP4 Task 2, should be feasible if fixed ensembles are used successfully. Further, an extension to assimilate more types of physics and BGC observations (e.g. oxygen and nutrients) should be possible. In addition, switching from offline-coupled DA (using separate programs for forecasting and the computation of the DA analysis and using output files for data transfer) to online coupled DA (integrating the DA functionality of PDAF with the model into a single program as done in SEAMLESS) should also be possible within this time frame. The online-coupled DA increases the computing efficiency as the number and volume of disk files is reduced.

Tier 3

On the long term of Tier 3 the results of WP 4 and WP 6 suggest the use of a dynamic ensemble because only this approach allows to estimate covariances for the environmental (physical and BGC) situation. Likewise, the use of the fully coupled assimilation is suggested. Given an ensemble of at least 30 members, as used in SEAMLESS, this approach would increase the required compute resources by a factor of 30 with the benefit of providing ensemble information.

Given the generic code structure of the NEMO-ERGOM-PDAF assimilation code developed in SEAMLESS, it would be possible to apply this assimilation code for all MFCs that use NEMO with one of the BGC models. The code developments for the physical part are not specific for the Baltic MFC, while the introduction of the BGC increments would need to be adapted to the BGC model. Utilizing one common community code would allow for direct synergies in the DA developments.

5. Roadmap for the Black Sea MFC

The Black Sea MFC has carried out developments with other funding sources than SEAMLESS. However, the University of Liège has been engaged by SEAMLESS partners routinely and the responsible of the biogeochemical component of the Black Sea MFC is member of the SEAMLESS Advisory Board (namely Marilaure Gregoire).

In addition, the University of Liege and a SEAMLESS partner (namely Pierre Brasseur, University of Grenoble) are collaborating on a joint CMEMS Service Evolution project (ODESSA), which is exploring the application of stochastic ensemble methods and the Ensemble Kalman filter in the Black Sea MFC. So the future collaboration and uptake of SEAMLESS tools is favoured although no concrete recommendations can be made on the Black Sea MFC from the SEAMLESS results.

6. Roadmap for the Global MFC

The assimilation system implemented by Mercator-Ocean in the GLO MFC is currently based on a tailored version of the SAM2v1 Mercator Assimilation System using a Reduced-Order analysis scheme derived from the SEEK formulation. The forecast error covariances of biogeochemical variables of the coupled NEMO-PISCES system are estimated using a fixed-basis ensemble of 3D biogeochemical state anomalies computed from a previous multi-year non-assimilative hindcast. The NRT system assimilates routinely ocean colour products, as listed in Table 1.

According to the presentation at the 2023 Design Review meeting, a consistent ensemble assimilation system will be targeted by MOi for the physical (NEMO) and biogeochemistry component (PISCES). The stochastic coupled NEMO-PISCES model which has been developed by UGA in WP3 is using exactly the same global setup as the one implemented in the GLO MFC system (version r4.0-HEAD.r13720, provided by MOi, see Deliverable 3.3). The SEAMLESS recommendation #2 in D3.4 is to envisage an incremental transition, with a first step (Tier-1) consisting in simulating ensembles, and a second step (Tier-2) dedicated to the upgrade of assimilation kernels facilitated by the use of the SEAMLESS prototype.

Tier 1

The uptake of the SEAMLESS results should first impact the forecast mode of the GLO MFC operations by:

- Implementing an ensemble forecasting system based on the stochastic ensemble ¼° NEMO-PISCES developed in WP3;
- Verifying the model ensemble forecast statistics using the EnsDAM packaged developed in WP3 (Deliverable 3.3, recently modified to improve the compatibility of the code using Python interfaces) and satellite / in situ chlorophyll observations;
- Extending the verification to SEAMLESS indicators (phenology, POC concentrations) under the condition that observations are available.

One condition of feasibility is the availability of CPU resources, which could initially limit the size of the ensemble to 10 members, which is the ensemble size selected by the GLO MFC for its global 1/4th degree resolution system.

The most influential PISCES parameters identified in the SEAMLESS Deliverable 3.2 have been identified by MOi as candidates for perturbation and generation of the biogeochemical model ensembles.

Tier 2

We estimate that the uptake of the stochastic ensemble inversion scheme developed in WP3 (Popov et al. 2023) is feasible within 2 years to upgrade the GLO MFC analysis scheme from 3D to 4D but only in selected regions of interest (while keeping unchanged the current global system in operation). This will require further R&D to

- select the regions, model variables and indicators of interest;
- fine-tune the parameterization of the 4D inversion scheme (space-time window, parameters of anamorphic transformations).

Tier 3

MOi has organised a working group on Data Assimilation involving its international shareholders as a think tank to assist its long-term DA software strategy. PDAF has been presented to them (by L. Nerger) and is noted as a possible way forward to assimilate data in their ocean BGC, physical and sea ice models, even though the working group has not concluded firmly on a selection of software. NERSC (L. Bertino) was a member of that working group.

The transition to assimilate satellite spectral reflectance data is foreseeable in a 3-5 years time. R&D has started on this subject in the Tier-3 NECCTON project in collaboration between UGA and ULG to implement a first demonstrator in the Black Sea using either the RADTRANS radiative transfer model or the equivalent OASIM model within the FABM framework and assess the benefit wrt assimilation of ocean colour products.

7. Roadmap for the Iberian-Biscay-Irish MFC

As of the present state of the IBI MFC, there is no assimilation of BGC data (Table 1). The IBI plans towards BGC data assimilation have been laid out during the November 2023 CMEMS Design Review meeting by MOi (S. Cailleau):

- 1st step: OC assimilation in a similar fashion as the Global MFC (no intervention of SEAMLESS for that).
- 2nd step: Adding a bio-optical module as in PML's ERSEM model and developing an assimilation scheme for optical properties obtained from satellite ocean colour data.
- Testing the above 2-way PHY-BGC coupling.
- PHY-BGC coupled assimilation.

In the longer term, IBI will implement an ensemble assimilation system based on the LETKF, coherently with the developments in the GLO MFC.

SEAMLESS proposes to adopt a similar roadmap as the one suggested for the GLO MFC as the basic modelling and assimilation components can be shared by both MFCs.

Tier 1

As for the Global MFC, the priority for the uptake of the SEAMLESS results should first address the forecast mode of the IBI MFC operations by transitioning to ensemble forecast. This should also be complemented by the uptake of the Service Evolution SCRUM and SCRUM-2 (as well as MULTICAST) since both projects shared the use of the Stochastic Parameters (STOPAR) routines developed for the stochastic version of NEMO-PISCES. The most influential PISCES parameters identified in the

SEAMLESS Deliverable 3.2 have been identified by Mercator as candidates for perturbation and generation of the biogeochemical model ensembles.

Tier 2

A similar roadmap as for GLO would apply here, with a specific focus on ecosystem indicators of interest to coastal users.

Tier 3

The coupled physical-BGC data assimilation is expected to follow the recommendations from the WP4 reports which have been presented to MOi experts. This follow-up does not depend strictly on adopting the same software tools but as a recipe to monitor the correct implementation and functioning of coupled DA in the IBI MFC tools.

The Tier-3 FOCUS project that will start in early 2024 will include R&D activities that will extend the effort made in SEAMLESS to transition from deterministic to stochastic approaches for coastal applications (in particular the transition of the stochastic methodology from NEMO to CROCO by SHOM).

8. Roadmap for the Mediterranean MFC

The starting point for NRT forecast system of the MFC of the Mediterranean Sea (MED) in 2023 is the 3DvarBio 3.4 variational scheme (Teruzzi et al., 2021; Cossarini et al., 2019). This assimilates Copernicus ocean colour (OC) surface chlorophyll, as well as chlorophyll, nitrate and oxygen from BGC-Argo profiles.

In reanalysis mode, the MED is using the 3DVarBio 3.3 system and assimilates surface chlorophyll data from ocean colour time series starting from year 1999 . The assimilation of BGC-Argo floats chlorophyll will be included consistently with the NRT forecasting system at the next reanalysis release (expected in 2025).

Tier 1

In SEAMLESS, the FABM framework in 1D setups has been used together with the PARSAC tool (https://zenodo.org/records/4280520) for parameter optimization of the bio-optical component of the BFM model (Alvarez et al., 2023). The parameters optimized in SEAMLESS have been taken-up by the CMEMS MFC in their 3D OGSTM-BFM NRT system for the November 2023 Entry into Service.

Tier 2

The MED MFC plans to further test the joint physical-biogeochemical data assimilation (1-way) through a balancing scheme developed for the 1D EAT framework in the Task 4.1 of SEAMLESS. The development requires additional investigation before it can be deployed on the 3D operational system i.e., transferring the calculated innovation and covariances from 1D to 3D is not yet satisfactory. This relevant development would be addressable in a tailored Tier 2 project.

Tier 3

The long-term evolution of the MED MFC includes the development of ensemble data assimilation methods following the successful preliminary implementation of a SEIK filter in the OGSTM-BFM model for the assimilation of OC and BGC-Argo by SEAMLESS WP3, 5 and 6.

As a first step, the implementation of an ensemble forecasting framework for the OGSTM-BFM would be based on perturbations on forcing fields, BGC model parameters and initial conditions up-taking results from WP3 of SEAMLESS and SE ODESSA. Such development would require a dedicated Tier 3 project.

9. Roadmap for the Northwest Shelf MFC

The NWS MFC has changed leadership during SEAMLESS because of Brexit, as of the submission of this deliverable, the IBI MFC is assuming the continuity of the operations of the NWS MFC products albeit without BGC data assimilation (See Table 1). After the re-entry of the UK in the Copernicus programme in 2023, the SEAMLESS partner PML can legitimately put a competitive bid to the next phase of CMEMS activities to be awarded in December 2024.

In the meantime, the multiyear NWS BGC product will be provided at the same resolution as the physical and wave reanalyses: at 1/36th degree resolution in November 2024.

Tier 1

The developments in SEAMLESS are an extension of a global physics ensemble-based system developed at the UK Met Office (Lea et al, 2022). The code developed in SEAMLESS for biogeochemical data assimilation is already integrated in the research and development (R&D) version of the Met Office operational system for the NWS and being already used for such purposes (e.g. Skakala et al, 2023). To turn the ensemble techniques developed in SEAMLESS into operations we mostly recommend to

- obtain a much more detailed validation across many years and multiple observational sources and variables.
- detailed assessment of the computational cost, which might easily constrain these methods to reanalysis applications.

The above two tasks are addressable in the framework of a Tier 1 project.

It is noteworthy that the ensemble development and implementation of ensemble-variational methods are an essential part of Met Office future marine data assimilation strategy and so SEAMLESS is perfectly aligned with the Met Office future plans.

Tier 2

In the intermediate run we will suggest accounting for more sources of uncertainty than done by SEAMLESS, e.g. including riverine perturbations into the ensemble, or spatio-temporally varying perturbations of parameters. These will further enhance the realism of the ensembles and make them more attractive for the operational applications. It is also noteworthy that SEAMLESS has developed methods very well suited to model such new perturbations. There are also efforts to further assess the ensembles through comparison with other ensemble methods developed at NWS, directly linked to PDAF. We also suggest to focus on optimization of computational resources to reduce the computational cost of the ensembles, e.g. through reducing the time resolution of the biogeochemistry model. Such work could be developed in a Tier 2 project.

Tier 3

In a long run the key interest of MFCs on NWS is ensemble-based multi-variate data assimilation, including strongly coupled DA between physics and biogeochemistry. This requires some further technical developments in the code, but more importantly much finer assessment of ensemble skill, as the potential for instability in model dynamics is much larger. The strongly coupled DA based on multi-variate ensemble methods is also a strong focus of Met Office future marine DA strategy. Although SEAMLESS has not yet delivered a skilled strongly coupled scheme, it provides crucial source of experience for future developments in this field.

10. Conclusions

There are several results from SEAMLESS that have already been taken up in CMEMS MFC even before the anticipated schedule. The most immediate has been the EAT tool for 1-dimensional model calibration in the ARC and MED MFCs.

The FABM tool has also gained popularity in the framework of CMEMS thanks to SEAMLESS. Developments initiated in SEAMLESS will be continued and expanded in the current CMEMS Tier 3 project NECCTON, as well as in the Horizon Europe OCEAN-ICU project. The adoption of FABM allows harvesting efforts across different teams and different regions, even though the BGC models remain specific for each region. This modular philosophy is more inclusive than would have been a monolithic common model.

A common data assimilation system PDAF has been used in SEAMLESS, at least for all the 1-dimensional experiments. Here again, the MFCs are not required to implement PDAF to benefit from the SEAMLESS results. The Baltic MFC does use PDAF for its 3D assimilation system and most ensemble data assimilation techniques (SEIK, LESTKF, DEnKF) can be applied in PDAF across different MFCs with a few years of efforts although most MFCs will prioritize developments of their own data assimilation software. Variational data assimilation has been included in PDAF for the purpose of SEAMLESS but in a rudimentary level, not aiming at the level of complexity of dedicated variational DA software.

Stochastic forecasts are not presently part of CMEMS products. The Ensemble Approach Working Group (EAWG) has been setup by MOi to initiate the groundwork and demonstrate its utility, but the BGC models have not participated so far by lack of maturity. SEAMLESS, as explained in the Global MFC section, does provide software for generating and validating ensemble forecasts that can be adopted by various MFCs (deemed Tier 1 for the Global MFC, more likely Tier 2 or 3 for other MFCs). SEAMLESS has tested an ensemble validation methodology in WP3 that can contribute to the EAWG once the MFCs start producing ensembles of BGC models.

A few new Ocean Health indicators can be proposed that would be reliable enough to be included into the CMEMS portal and published on a yearly basis in the Ocean State Report (OSR, von Schuckmann et al. 2023). See Table 4 above for details per MFC. However, these OMIs must be validated by forecast and reanalysis systems before they are submitted as new OMIs to the OSR, which we recommend as a first step below.

Recommended requirements (generic for all MFCs, not detailed by MFC)

1. To include in their products: POC
2. To include as indicator: Phytoplankton phenology, trophic efficiency
3. to adopt FABM
4. to perform an analysis of controllability of the biogeochemical variables with respect the BGC data they assimilate
5. to generate coupled physical-biogeochemical ensembles, by also perturbing the biogeochemical parameters that are proven to be the highest ranking for sensitivity
6. To implement coupled physical->biogeochemical DA in accordance to the cost-benefit analysis from WP4
7. To implement joint satellite-in situ BGC DA