Project	SEAMLESS No 101004032	**Deliverable**	D3.4
Dissemination	Public	**Type**	Report
Date	31st July 2022	**Version**	1.0

Deliverable 3.4

Observability of the target indicators in the 3D CMEMS MFC systems

Deliverable Contributors:	Name	Organisation	Role / Title
Deliverable Leader	Brasseur P.	CNRS	WP3 leader
Contributing Author(s)	Brankart J.M.	CNRS	Contributor
	Popov M.	UGA	Contributor
	Skakala J.	PML	Contributor
	Teruzzi A.	OGS	Contributor
	Spada S.	OGS	Contributor
	Cossarini G.	OGS	Contributor
	Wakamatsu T.	NERSC	Contributor
	Bertino L.	NERSC	Contributor
	Sun Y.	AWI	Contributor
	Nerger L.	AWI	Contributor
Reviewer(s)	Ciavatta S.	Mercator Ocean international	Advisory Board
Final review and approval	Jozef Skakala	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**	D3.4
Dissemination	Public	**Type**	Report
Date	31st January 2023	**Version**	4.0

Document History:

Release	Date	Reason for Change	Status	Distribution
1.0	20-05-22	Initial document, reviewed		Internal
2.0	31-07-22	Updated document following internal review (without AWI contribution)		Internal
3.0	01-12-22	Updated document following internal review (including AWI contribution)		Internal / EC
4.0	31-01-22	Formatted document submitted to the EC		Public

To cite this document

Brasseur P. *et al.* (2022). D3.4 Observability of the target indicators in the 3D CMEMS MFC systems. Deliverable report of project H2020 SEAMLESS (grant 101004032.). doi: 10.5281/zenodo.7584865

TABLE OF CONTENTS

1. Scope

SEAMLESS has the strong ambition to develop novel ensemble data assimilation systems exploitable operationally in the CMEMS MFCs, in some cases as alternative of established variational methods. The rationale is to improve the estimation of key marine ecosystem indicators.

The objectives of this document are: (i) to report on the developments of ensemble assimilation methods and assimilation experiments performed in WP3 of SEAMLESS (Task 3.3), and (ii) to provide methodological guidelines to assess the observability and controllability of the target indicators across the CMEMS MFC 3D domains (Task 3.4) using common probabilistic evaluation tools such as those delivered in D3.1.

Each partner supporting a particular Monitoring and Forecasting Centre (MFC: AWI for BAL; NERSC for ARC; OGS for MED; PML for NWS; IGE for GLO and IBI) implemented assimilation experiments in 3D setups as compliant as possible with the ones used by the CMEMS systems in operation today, in order to facilitate the transfer of methods and results towards MFCs and to draw recommendations applicable to each MFC. This deliverable is complementary to deliverable D3.2 (based on 1D experiments), providing however less general but more directly exploitable results for each MFC.

The more generic developments made in WP3 (Task 3.1) to transition the NEMO and HYCOM modelling platforms (which are the building blocks of all MFC model systems) towards probabilistic systems are detailed in Annex 1 (for NEMO/PISCES, responsible: IGE) and Annex 2 (for HYCOM/ECOSMO, responsible: NERSC). The corresponding codes were delivered in February 2022 as part of D3.3. Annex 3 covers the upgrade of analysis tools based on anamorphic transformations, which also were included in the activity of Task 3.3.

2. Introduction

A key strategic evolution at Copernicus 2 horizon will be to consolidate the Copernicus Marine Service with more robust information about product uncertainties, whether in real time, in delayed mode (reanalyses) and in forecast mode with a few days of lead time. In that perspective, a necessary step undertaken by SEAMLESS is the transition to probabilistic methodologies for the “Green Ocean”, i.e. for ocean biogeochemistry, with the aim to provide more actionable information to help in decision-making and management of marine ecosystems.

In line with this strategy, the **first key objective of WP3** is to develop new ensemble modelling and data assimilation methods that maximize the flow of information from *in situ* observing networks and satellite constellations towards controllable ecosystem indicators. These new methods are designed as upgrades of the existing assimilation systems operated today by MFCs, which currently include variational and ensemble-based approaches.

The transition to ensemble assimilation systems including robust uncertainty estimation capabilities corresponds to a change of paradigm in most MFCs, as the deterministic models used today in MFCs need to be converted into probabilistic systems producing ensembles of finite size that sample the

probability density functions (pdfs) of the state variables. The methods developed by the partners in Task 3.3 demonstrate that such a transition is feasible in a near future since ensembles of model states used for the computation of error covariance matrices of Kalman filters can also be used in a similar way in variational assimilation algorithms. In addition, the exploitation of ensemble members derived from probabilistic systems can deliver uncertainty estimates using Gaussian or non-Gaussian analysis schemes (depending on the shape of the pdfs).

In section 3, we present the main ingredients of the new assimilation methods and associated algorithms that have been developed in Task 3.3 and tested by the partners in 3D setups, underlining the evolutions towards ensemble-based solutions.

The **second key objective** is to assess the observability/controllability/identifiability (OCI) of the SEAMLESS target indicators in the CMEMS models. The fundamental reason for studying these properties is to establish objective criteria for including a new product in the CMEMS catalogue, or enabling quality improvement for already-in-catalogue products. If the experiments show a sufficient level of controllability, it can be recommended to include them in the CMEMS catalogue. If the controllability is low, it means that the observations and the assimilation processing chains need to be improved to reach proven scientific value in the CMEMS catalogue products.

OCI properties can be assessed analytically for linear time invariant models, e.g., by computing local derivatives of observational equation with respect to the model states (see e.g., Di Stefano, 2015). These approaches are impractical to assess OCI of actual ocean ecosystem models which are highly-nonlinear. SEAMLESS therefore exploits the ensemble approaches developed in WP3 to assess OCI in practice.

The OCI properties of the SEAMLESS indicators were first explored in Task 3.2 using 1-D modelling setups. The pragmatic approach considered in Task 3.2 (see D3.2) was to perform sensitivity analysis with respect to the observed variables (by perturbing corresponding initial conditions) and biogeochemical parameters. However, this approach cannot be easily transitioned to 3D setups because of the high computing costs implied by model ensembles in 3D configurations. We have therefore adopted a different approach for Task 3.4, which is based on the measure of uncertainty reduction obtained by conditioning the 3D model ensembles on the available observations using the new methods developed in Task 3.3. The underlying assumption is that reduced uncertainty is an index of increased controllability of the system.

As part of Task 3.3, an internal workshop was organized on April 5th 2022, with the objectives of (i) reviewing the developments of the ensemble methods developed by the SEAMLESS partners, (ii) identifying a common methodology for setting up experiments in the different systems that replicate the MFCs environments, and (iii) specifying possible metrics for quantifying uncertainty reduction.

Section 4 reports the assimilation setup and the results obtained in Task 3.3 in the different MFC regions. Ensemble diagnostics are described in terms of (i) observed state variables, (ii) non-observed state variables and (iii) relevant SEAMLESS indicators (leveraging from the conclusions of Task 3.2). In section 5, the results obtained for indicators are discussed with a focus on phytoplankton phenology,

particulate organic carbon (POC) flux, primary production, zooplankton grazing efficiency and PFTs (pH and O_2 being of lesser relevance following the conclusions of Task 3.2: see D3.2 of SEAMLESS). Finally in section 6, we provide guidelines for its application to assess biogeochemical (BGC) products in the 5 MFCs and their integration in the CMEMS catalogue.

We stress that a new range of different assimilation systems for the CMEMS MFCs are presented here for the very first time, all based on ensemble methods (from the SEIK system for the Mediterranean Sea to the hybrid ensemble-variational system for the NWS MFC) demonstrating the innovation brought by SEAMLESS. Whilst the sections of the reports have similar structure for the different MFCs for the sake of readability, we allowed for some differences in the sections, to report on the unique features of the different novel assimilation systems (e.g. their specific hyperparameters). Also, some differences in the presentations of the results were allowed to focus on the most significative indicators in the different regions, as well as taking account of the availability of data for corroboration of the simulated indicators in the different MFCs.

3. Ensemble assimilation methods

This section describes the new assimilation methods developed and tested in Task 3.3 by the partners, and the assimilation gaps that SEAMLESS aims to bridge with respect to the approaches in operation today in the different centres. These new methods include one stochastic ensemble DA system applicable to the global (GLO) and Iberian-Biscay-Irish (IBI) MFC domain; one hybrid ensemble/particle filter for the Baltic (BAL) MFC; two hybrid ensemble/variational methods for the Mediterranean Sea (MED) MFC and North-West Shelf (NWS) MFC; one updated ensemble Kalman filter and smoother system in the Arctic (ARC) MFC. A schematic illustration of the transition to ensemble-based assimilation methods as achieved by SEAMLESS for BGC monitoring and forecasting in the different MFCs is given in Table 3.1 at the end of this section.

3.1 Assimilation methods for the BAL MFC domain

Assimilation system in operation in BAL system. The current operational systems for biogeochemistry implemented by the members of the CMEMS Baltic-Monitoring and Forecasting Center (BAL-MFC) are distinguished for the system for reanalysis and the system for producing the biogeochemistry forecast and analysis products. Both systems use the NEMO ocean model in the NORDIC configuration (Hordoir, et al., 2019). Here we describe the status of the operational system as of the beginning of December 2022.

The system for the reanalysis product uses NEMO in version 3.6 at a resolution of 4km with the biogeochemical model SCOBI (Swedish Coastal and Ocean Biogeochemical model) together with a LSEIK data assimilation scheme using PDAF (Parallel Data Assimilation Framework, Nerger & Hiller, 2013, Nerger et al., 2005). Assimilated are profile observations of nitrate, phosphate, ammonium, and dissolved oxygen in a univariate way. The data assimilation is operated with a covariance matrix generated from reading model snapshots around the date of the year from a multi-year simulation. Thus, only a single model state is integrated by the model. PDAF is coupled offline with the NEMO-SCOBI, model using restart and increment files to exchange information.

The system for producing the forecast/analysis data products uses NEMO in Version 4 at a resolution of 2 km with the biogeochemical model ERGOM (Neumann, 2000). Currently, no observations are assimilated.

Active developments are performed by the members of the BAL-MFC to enhance the systems. In the near future, both systems are planned to run using NEMO in Version 4 with a resolution of 2 km and ERGOM. Further the assimilation of profile observations is in progress for both systems. The data assimilation will still use a covariance matrix generated from reading model snapshots around the date of the year from a multi-year simulation so that only a single model state is integrated by the model.

Assimilation developments for the BAL system. In SEAMLESS, **AWI** developed further the BAL-MFC operational code of NEMO Version 4 coupled with ERGOM. The BAL-MFC consortium provided us access to the code and the current model setup with a resolution of 2 km. We particularly aim for a dynamic ensemble as this is expected to provide better covariance estimates than fixed ensemble perturbations derived from model snap shots.

The data assimilation developments by AWI consisted in the coupling of PDAF with the NEMO-ERGOM model system in the form of online coupling (see Nerger & Hiller, 2013). For this, NEMO is modified so that it is able to perform ensemble simulations. These code changes mainly consisted in adapting the parallelization of the model and reconfiguring the file output. This was done in analogy to different models to which PDAF was coupled before (e.g., for the FESOM ocean model (Nerger & Hiller, 2013), the AWI climate model (Nerger et al., 2020), MITgcm-REcoM (Pradhan et al., 2019), HBM-ERGOM (Goodliff et al., 2019)). Compared to the coupling of NEMO Version 3.3 with PDAF in Toedter et al. (2016) we use an updated structure that allows us more flexibility. In addition to modifying the parallelization, subroutines calls have been added to the higher-level routines of NEMO to first initialize the ensemble states as well as PDAF itself, and to conduct the actual data assimilation analysis step and the subsequent update of the fields of both NEMO and ERGOM. This approach was shown to be very efficient, because the exchange of model fields and grid information is performed in memory rather than via files on disks. Further, we can avoid frequent model restarts and can perform cycled data assimilation by starting the model system with its data-assimilation enhancement once in ensemble mode. This system now allows fully dynamic ensemble data assimilation, so that the ensemble represents at each time the dynamically-influence covariances. In addition, we implemented stochastic parameter perturbations for process parameters of ERGOM. Here a log-normal stochastic perturbation is used with a relative standard deviation of 0.125. The system supports both the localized ensemble square root Kalman filters LETKF and LESTKF (see Nerger et al., 2012), as well as nonlinear filters like the LNETF (Toedter & Ahrens, 2015) and the hybrid filter LKNETF (Nerger, 2022).

After augmenting the model code with data assimilation functionality by PDAF, we implemented the model-specific routines. In particular, the code now allows to include into the state vector (and hence the ensemble) both five physical ocean variables from NEMO as well as 16 prognostic and 3 diagnostic (pH, pCO2, total chlorophyll) variables from ERGOM. Using configuration operations, these variables can be activated or deactivated at run time. For example, we can select to only update the

concentrations of the three phytoplankton groups (diatoms, flagellates, cyanobacteria) when assimilated satellite chlorophyll observations. Further, an option to apply a log-transform to the variables of ERGOM was implemented.

For the assimilation the focus is on satellite observations. This is motivated by the higher availability of the observations, compared to the very sparse biogeochemical profile observations. So far, support for satellite surface temperature and chlorophyll observations was implemented. The observational data products are taken from CMEMS. The original data files are directly used and the observation operator interpolated from the model grid to the observation grid.

3.2 Assimilation methods for the GLO and IBI MFC domains

This section describes the assimilation method and associated algorithms developed by **IGE/UGA**, in the perspective of operational implementation in the GLO and IBI systems. In the systems operated at the time of writing this report, near real-time analyses and 10-day forecasts covering the GLO domain are delivered weekly using a 1/4° NEMO-PISCES coupled model that assimilates Level 4 ocean colour CMEMS products. A hindcast product is also delivered over the period 1993-2020, but without assimilation. In the IBI region, all products in real-time and delayed mode are generated without assimilation.

Assimilation system in operation in GLO and IBI systems. The assimilation system implemented by Mercator-Ocean in the NRT GLO domain is based on a tailored version of the SAM2v1 Mercator Assimilation System (Brasseur *et al*., 2005; Lellouche *et al*., 2013). The analysis increment is computed with a Reduced-Order analysis scheme derived from the SEEK formulation (Pham *et al*., 1998). The forecast error covariances of biogeochemical variables are estimated using a fixed-basis - but seasonally variable - ensemble of 3D biogeochemical state anomalies computed from a previous multi-year non-assimilative hindcast. The multivariate formulation is used to calculate 3D updates of selected state variables (e.g., nanophytoplankton, diatoms and nitrates model concentrations), using the surface total chlorophyll concentration data provided by satellite observations. The whole analysis is done in the log-space in order to ensure positiveness of the multivariate correction. Finally, an Incremental Analysis Updating (IAU) scheme is applied to update the upper ocean phytoplankton and nitrate model variables, and a vertical projection is applied to the water column following an exponential-based modulation function. This modulation ensures that the projected increments have a maximum at the surface and decreases towards 0 at the base of the model turbocline.

Assimilation developments for the GLO and IBI systems. Before SEAMLESS**,** in Santana-Falcon *et al*. (2020), a localized ensemble Kalman filter was developed and tested in a North Atlantic configuration, using the SEEK filter formulation for 3D multivariate analysis updates of biogeochemical variables, and a reduced basis that evolves with time according to the NEMO/PISCES model dynamics. This study showed that surface ocean colour data does not contain enough information to generate robust re-initialization of a complex model such as NEMO/PISCES which includes 24 BGC state variables in 3D. In SEAMLESS, this motivated the development of the new ensemble assimilation approach in WP3, that more closely corresponds to a Bayesian inversion scheme in 4D (space and time). The algorithm is built on a separation between the prior ensemble generation which relies on the full model

complexity, and the posterior pdf computation which only applies to a small subset of state variables in 4D. This approach makes the inverse problem simpler and avoids re-initialization and time-integration of the full set of state variables, while still enabling probabilistic forecasts (see Figure 4.1 here below). The method is also well suited to simultaneously constraining physical and biogeochemical state of the coupled model, as will be demonstrated in WP4.

Due to its 4D nature, the scheme has similarities with EnKS or 4DEnVar approaches that are currently developed for Numerical Weather Prediction. The 4D algorithm developed and tested in Task 3.3 includes the following steps:

Assessment of dominant uncertainty sources and related stochastic parameterizations in the probabilistic NEMO-PISCES model. As part of SEAMLESS, we have extended the generic method to simulate uncertainties in NEMO (Brankart *et al.*, 2015) to the coupled NEMO/PISCES modelling framework using autoregressive processes to generate randomness in the model equations. A broader range of stochastic parameterizations have been implemented and combined to simulate the dominant uncertainty sources associated to unresolved physical-biogeochemical processes, unresolved scales and unresolved biodiversity in marine ecosystems (see Annex 1 for more details on the formulation of the parameterizations). The code developments are integrated in the STOPAR module of NEMO4.0 (see detailed description in D3.3).

Prior ensemble production. A prior ensemble simulation covering the period of interest (i.e., the temporal window covered by observations, possibly extended to the forecast lead time as illustrated by Fig. 3.1) is performed using the parallel NEMO option coded in two routines mpp_start_ensemble and mpp_set_ensemble (this development required rewriting of the simulation code to adjust to the NEMO4.0 parallelization, which is very different from NEMO3.6, see detailed description in D3.3). Given the huge amount of data generated, an extraction of the ensemble output is performed in 4D (space-time) selected sub-domains where inversions are then calculated.

Anamorphic transformation. In order to account for the non-Gaussian nature of the prior ensemble, the ensemble-based anamorphic operator developed by Brankart *et al*. (2012) is applied to transform the model state variables into Gaussian variables. The SESAM analysis tool has been upgraded as described in Annex 3 (Task 3.3a) introducing a bias correction scheme to ensure compatibility between observation error statistics in the original and transformed spaces.

Localization in space / time. The implementation of the localization scheme described in Brankart *et al*. (2011) has been upgraded to generalize the inversion process to 4 dimensions (in space and time). New localization parameters are introduced (and need to be tuned) to regulate the temporal correlations between past, present and future states in compliance with the plausible predictability time scales of the system.

Posterior ensemble calculation in anamorphic space. This step is based on the LETKF analysis scheme implemented in SESAM, extended to 4D to account for space-time correlations between model state and observed variables. In practice, this operation requires concatenating the different 3D state vectors of the inversion time window into a single 4D state vector.

Reverse transformation to physical scale. This operation converts the posterior ensemble from the transformed (Gaussian) space into the physical (non-Gaussian) space using the reverse anamorphic operator applied in step c).

Verification and diagnostics: This final step is crucial to quantify the assimilation effectiveness and assess controllability. A variety of metrics can be calculated using the scores available in D3.1, which can be used in the physical or transformed space.

It should be noted that the same assimilation tools/codes apply for probabilistic hindcast inversions, analyses, forecasts and projections. Additional considerations on computational complexity and implications are necessary for transition to operational GLO and IBI systems. The algorithm has been applied in hindcast mode in WP3 as illustrated by the results described in section 4.2.

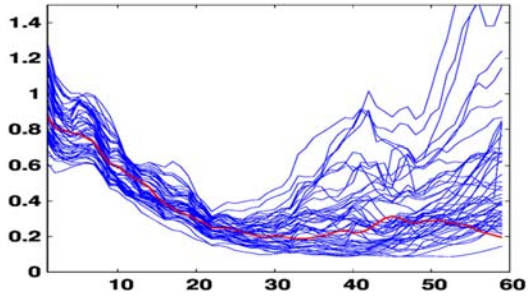


Figure 3.1. Illustration of the temporal evolution of surface chlorophyll from a 4D ensemble hindcast/nowcast/forecast in PAP station. The observations are assimilated during the first 30 days only, reducing the spread of the ensemble around the observed values (in red). The uncertainty (represented by the dispersion of the ensemble) increases again following the observation period.

3.3 Assimilation methods for the NWS MFC domain

Assimilation system in operation in NWS system. The current **NWS** system is based on 3D-variational assimilation using the NEMOVAR software (Mogensen *et al*., 2009, 2012; Waters *et al*., 2015). The background covariance matrices are supplied as monthly climatologies. The climatological variances originate from a decadal EnKF reanalysis (Ciavatta *et al*., 2016) and the system uses spatially varying climatological horizontal length scales. The vertical length scales are flow-dependent and calculated based on the simulated mixed layer depth. After the system performs univariate assimilation, it uses a balancing scheme to distribute the DA increments into some of the non-assimilated biogeochemical variables (e.g., spreading increments from total chlorophyll into all the nutrient components of the cells of the Phytoplankton Functional Types, PFTs, of the ERSEM model). The system is designed to assimilate on a daily basis satellite ocean colour (OC)-derived total chlorophyll, PFT chlorophyll and optical data, and has also the capacity to assimilate glider data, such as chlorophyll and oxygen (Skakala *et al*, 2021). In its operational version, the system assimilates only OC total chlorophyll.

Assimilation developments for the NWS system. In SEAMLESS, **PML** developed the ensemble-3DVAR (hybrid) capacity for the biogeochemical part of the coupled NWS NEMO-FABM-ERSEM model. The

new system uses existing physical ensembles based on 10-member ERA5 atmospheric forcing ensemble and observation perturbations. In SEAMLESS we added the perturbations of the 6 most sensitive ERSEM parameters to the existing physical ensemble, by exploiting the 1D sensitivity results from WP3.2. The six ERSEM parameters include 4 parameters of the PFTs and 2 parameters related to bacteria (see the SEAMLESS report D3.2). All these six parameters were perturbed by up-to 30% of their value and randomly drawn from a uniform distribution. The new assimilation developments were implemented in a version of the NWS NEMO-FABM-ERSEM model that includes also the OASIM bio-optical model (Gregg and Casey, 2009; Skakala *et al*., 2020) that drives both physics and biogeochemistry, and simulates biogeochemical feedback to physics (two-way coupling, Skakala *et al.*, 2022) through light attenuation by phytoplankton. The ensemble-3DVAR approach used in SEAMLESS is based on 30-member ensemble and assimilates satellite SST, profiles of temperature and salinity (provided by the EN4 Hadley centre), as well as OC-derived total chlorophyll. In hybrid systems, the information on covariances provided by the ensemble is typically combined with the climatological covariances and a weighted average is calculated between those. In our experiments we focused fully on the ensemble component and therefore the covariances were calculated purely from the ensemble. The ensemble provided information for the background variances, as well as for the horizontal and vertical length scales, but the assimilation remained univariate. Series of twin-experiments were performed to estimate some free parameters used in the system, such as the normalization length scales. Following the twin-experiments, a monthly-long simulation was run and compared between the hybrid ensemble-3DVAR and an equivalent ensemble of 3DVARs. After the 1-month simulation, a longer 4-month simulation was performed for the Spring bloom period of March-June 2018 and validated with the available glider data from the AlterEco mission, observations at the long-term monitoring station “L4” in the Western English Channel and NSBC climatologies. Further to that we used rank histograms to assess the spread of the ensemble.

3.4 Assimilation methods for the MED MFC domain

Assimilation system in operation in MED system. In the current NRT-**MED** system that operationally provides biogeochemical products for the Marine Copernicus Service, data assimilation is performed with a 3D-variational approach with decomposition on the error covariance matrix in three operators that account for the vertical, horizontal and biogeochemical covariances. Satellite and BGC-Argo observations are assimilated weekly and daily, respectively. The assimilated observations include chlorophyll concentration from satellite and floats as well as nitrate concentration from floats. In the current 3D-variational setup the updated variable are phytoplankton variables (the contents of nitrate, phosphate, silicon, carbon and chlorophyll of 4 the phytoplankton functional types), and dissolved nitrate and phosphate. The reanalysis MED system, which covers the 1999-2021 period, uses the same 3D variational assimilation scheme for satellite chlorophyll.

Assimilation developments for the MED system. In SEAMLESS, **OGS** developed a sequential ensemble assimilation approach (**SEIK-OGSTM**), which online couples the OGSTM-BFM model and a SEIK filter (Singular Evolutive Interpolated Kalman filter; Pham et al., 1998). **OGSTM–BFM** is the model in use for biogeochemistry in MED (Salon *et al*., 2019) forced offline by ocean dynamics provided by NEMO-OceanVar (that is not involved in the assimilation development within SEAMLESS). In the OGSTM-BFM system, OGSTM solves the transport of tracers and embeds the biogeochemical flux model BFM

(https://bfm-community.github.io/www.bfm-community.eu/) that solves biogeochemical evolution of 51 prognostic variables of the low trophic level (see the SEAMLESS deliverable D2.2 and reference cited therein). SEIK-OGSTM is designed to perform ensembles of OGSTM simulations and assimilations in a fully parallel framework, where the ensemble can be initialized on a set of different initial conditions (ICs) or different sets of parameters of the BFM model. Moreover, inflation and localization have been implemented as well as the option to choose additive and/or multiplicative observation errors.

SEIK-OGSTM works on log-transformed concentrations, thus negative concentrations cannot be produced by the assimilation. In order to limit the generation of spurious covariances due to the variability of low phytoplankton concentrations below the euphotic layer, SEIK-OGSTM includes a threshold on low concentrations applied before the assimilation.

SEIK-OGSTM parallel implementation takes advantage of MPI shared memory windows to maximize access speed to RAM data while reducing total RAM consumption and thus increasing the system scalability in multiple nodes configurations.

Being SEIK-OGSTM a sequential EnKF system, forecasts and assimilations are performed according to the following steps:

i) Initialisation of the ensemble based on ICs and/or different sets of parameters.
ii) Ensemble forecast phase that integrates the OGSTM-BFM for each ensemble member in parallel.
iii) Assimilation with update of error covariance matrix (in the error subspace spanned by the ensemble) and production of the analysis, using inflation and localization techniques.
iv) Analysis ensemble: resampling by second order exact sampling with random isometry matrix (as in Pham *et al*., 1998).

Steps from ii) to iv) are sequentially repeated through the whole simulation. In the SEIK-OGSTM BFM parameters in each member of the ensemble do not vary during the forecast simulations and are not affected by the assimilation.

3.5 Assimilation methods for the ARC MFC domain

Assimilation system in operation in ARC system. The present **ARC** MFC real-time operational system only includes a direct insertion of OC data using an empirical rule for the vertical projection of Chlorophyll-a to profiles of the different phytoplankton classes. This method was designed to add a minimal overhead on the cost of the forecast system at high horizontal resolution. The reanalysis system adopted on the contrary a more advanced ensemble data assimilation method with the Ensemble Kalman Smoother (EnKS). The EnKS offers more capabilities than the direct insertion method: the possibility to assimilate jointly satellite and in situ profiles, a flow-dependent multivariate update of non-observed variables, Gaussian anamorphosis, uncertainty estimates and the capability to estimate model parameters. However, it comes at a higher cost, with 80 dynamical ensembles and a lag-1 smoother. The computational costs are a factor of 160 times that of the free running model, so the biogeochemical reanalysis in the ARC MFC is provided at a resolution 4 times coarser than the operational forecast (25 km instead of 6.25 km).

Assimilation developments for the ARC system. In SEAMLESS, **NERSC** upgraded the current ARC MFC system by introducing the estimation of new biogeochemical parameters along with the biogeochemical states. The joint state-parameter estimation is based on the TOPAZ Deterministic Ensemble Kalman filter system (DEnKF, Sakov and Oke 2008) used in a one-lag smoother (EnKS) setting. The model is HYCOM-EVP-ECOSMO, coupled online. A joint model parameter-state estimation system is adopted using the state augmentation technique (Simon *et al*., 2015, El Gharamti *et al*., 2017) with parameter inflation by a first-order auto-regressive process. A Gaussian anamorphosis (log-normal transformation) is applied to the biogeochemical variables Chl-a, nitrate, silicate and phosphate and a bounded beta distribution is applied to the model parameters. Assimilation masks are defined by sea-ice concentration and mixed layer depth for mitigating analysis instability related to model bias in the timing of phytoplankton phenology.

Only biogeochemical data are assimilated (surface chlorophyll-a from satellite ocean colour sensor, in-situ Nitrate, Silicate and Phosphate) and only from March to October in each year, following the availability of light and in situ data in the Arctic. The assimilation cycles are defined by the frequency of the composite satellite ocean colour data OC-CCI v4.2 8-daily product. All in-situ nutrient data are binned to the nearest analysis date.

Our main achievements within WP3.3 are the following:

i) Analyze rank histograms for the surface Chl-a over the Norwegian Sea before and after DA from our BGC reanalysis. All members have been stored at each analysis step and this will shed some lights on how much the filter (EnKF) and smoother (EnKS) are contributing to improving our ensemble.
ii) Perform the analysis during a few cycles during mid-June in 2007 when the model forecast errors tend to grow large in general.
iii) Repeat the analysis with an ensemble of physical parameters as a new source of ensemble generation as a new prototype system and test its impact.

3.6 Summary of DA developments

A comparison of the main features of the present (operational) and targeted (developed by SEAMLESS) assimilation systems is shown in Table 3.1 for the different CMEMS MFCs. This table illustrates several convergent aspects of the methods used:

- all targeted data assimilation systems are based on ensemble methodologies, or combine ensemble and variational (hybrid) approaches;
- most of the new updating schemes rely on ensemble statistics, thus avoiding *ad hoc* adjustment schemes to spread increments to unobserved variables;
- the assimilation updating schemes include a broader range of BGC state variables, making the estimation of complex ecosystem indicators more straightforward;
- Uncertainty on products can be calculated more objectively, e.g., by considering the assimilation impact on ensemble spread.

System	DA method	Error covariances / Error subspace	Updated BGC variables	Uncertainty estimation
NWS CMEMS	NEMOVar	Prescribed stats (monthly climatol.)	Univariate DA (Chl) + balancing scheme	No
NWS SEAMLESS	Hybrid Ensemble/NEMOVar	3D ensemble-based	Univariate DA (Chl) + balancing scheme	Ensemble spread
IBI CMEMS	None			
GLO CMEMS	SEEK Filter Fixed basis	Prescribed stats (seasonal climatol.)	Bi-variate (PHY + N) + adjustment scheme	No
IBI/GLO SEAMLESS	Stochastic Ensemble Filter	4D ensemble-based (space + time)	Full state vector	Ensemble spread
MED CMEMS	3DVarBio	Prescribed stats	Multivariate (PHY+N+P)	No
MED SEAMLESS	SEIK	3D ensemble-based	Multivariate (PHY+N+P) or Full state vector	Ensemble spread
BAL CMEMS	None			
BAL SEAMLESS	LESKTF & Hybrid Filter LKNETF	3D ensemble-based	Multivariate (Chl+ 3 phytoplankton variables) or Full state vector	Ensemble spread
ARC CMEMS	DEnKF/EnKS	3D ensemble-based	Full state vector	Ensemble spread
ARC SEAMLESS	DEnKF/EnKS updated	3D ensemble-based	Full state vector + BGC model parameters	Ensemble spread

Table 3.1. Comparison of the CMEMS DA methods before (dark grey) and after WP3 developments (light grey) undertaken in the 3D MFC configurations, error covariance schemes, updated biogeochemical variables (PHY= phytoplankton, Chl = chlorophyll, N = nitrate, P = phosphate), and uncertainty estimation approaches. The features of the BAL system will be updated in a later version of this deliverable.

4. Assimilation experiments

This section describes the assimilation experiments performed in Task 3.3 by the partners, and the assimilation diagnostics computed for (i) observed model variables, (ii) non-observed model variables and (iii) derived quantities and relevant SEAMLESS indicators.

4.1 Assimilation results in the BAL MFC domain

Assimilation setup in BAL region. The assimilation is run over the three months March to May 2015 over the spring bloom. An ensemble of 30 members was generated. To initialize the ensemble for March 1, 2015, we first performed a spin up run for two months started at January 1, 2015 using the perturbed parameters. Initial ensemble perturbations were generated for January 1, 2015 from two months of daily model snapshots using second-order exact sampling (Pham, 2001) which is provided by PDAF. The initial central state of the ensemble is the model state on January 1, 2015 from the operational model run of the BAL-MFC. The data assimilation experiments are then started on March 1, 2015. We assimilate level-3 satellite chlorophyll data from CMEMS (data product OCEANCOLOUR_BAL_BGC_L3_MY_009_133). The assimilation is performed daily at model midnight. The model time step is 90 seconds so that each forecast phase includes 960 model time steps. Given that the operational systems for the Baltic do not use a dynamic ensemble, we performed two assimilation experiments: The first uses the localized error-subspace transform Kalman filter (LESTKF,

Nerger et al., 2012), which is also used by the BAL-MFC. This simulation represents the baseline applying a linear Kalman filter. The second experiment applies the hybrid Kalman-nonlinear ensemble transform filter (KNETF, Nerger, 2022). This filter combines the ensemble Kalman filter variant ETKF with the nonlinear ensemble transform filter (NETF, Toedter & Ahrens, 2015). The hybridization is done adaptively basing on the skewness and kurtosis of the observed ensemble. A free running ensemble simulation without assimilation is used to assess the impact of the assimilation. Both filters apply a localization with a radius of 30 km using a 5-th order polynomial to weight the observations.

Diagnostics on observed variables

To obtain an overview of the general effect of assimilating the satellite chlorophyll observations, Figures 4.1.1, 4.1.2, and 4.1.3 show the observations and the surface chlorophyll from the model for three different dates. The modelled surface chlorophyll is the ensemble mean and shown are the free run, 24-hour forecast from the assimilation, and the analysis. The assimilation was done here using the linear LESTKF method. Figure 4.1.1 shows the chlorophyll for March 3, 2015, thus at the third day of the data assimilation process. On this day, the observations are mainly available in the south-western Baltic and the Skagerrak and Kattegat. The free run shows higher concentrations in the southern Baltic and the central Baltic, while the concentrations are lower further north in the Bothnian Sea and Bothnian Bay. The Gulf of Finland shows low concentrations in the western part, but increased concentrations further to the east. The 24-hour forecast shows reduced concentrations in particular in the central Baltic. This difference from the free run is due to the assimilation on the first two days of March. The south-western part of the Baltic still shows higher concentrations. The assimilation reduces these higher concentrations to a wide part as is visible in Fig. 4.1.1(D). Due to the localization used in the LESTKF analysis step, this effect is restricted to the area where observations are available and 30km around these.

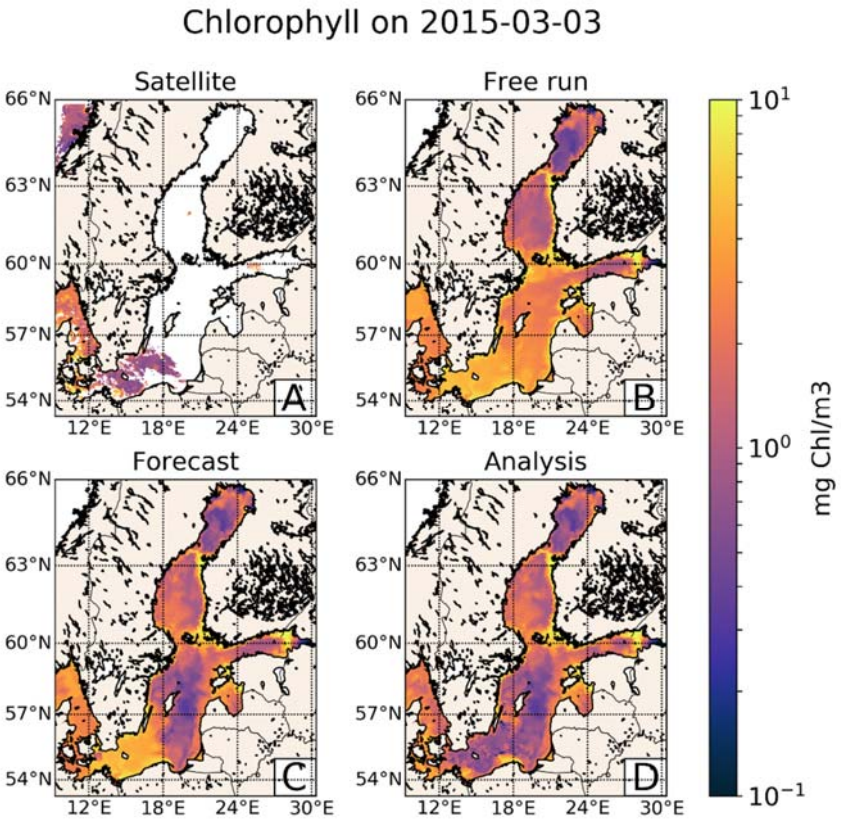


Fig. 4.1.1. Surface chlorophyll on March 3, 2015. Shown are the (A) satellite data, (B) free run, (C) 24-h forecast from the assimilation, (D) analysis.

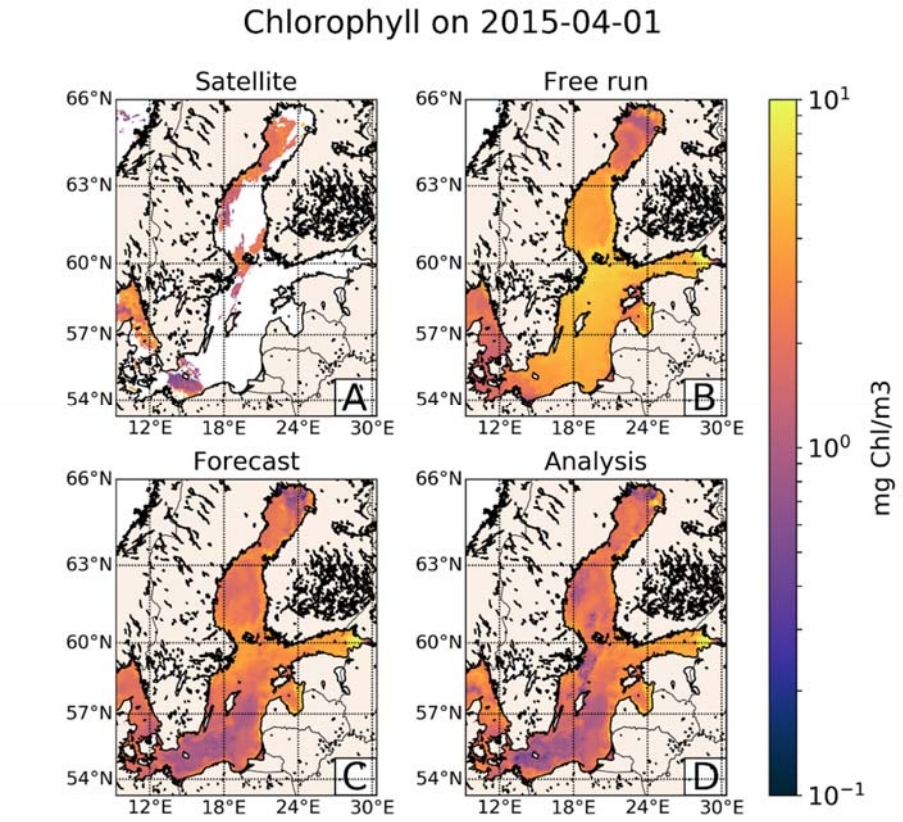


Fig. 4.1.2. Surface chlorophyll on April 1, 2015 analogous to Fig. 4.1.1

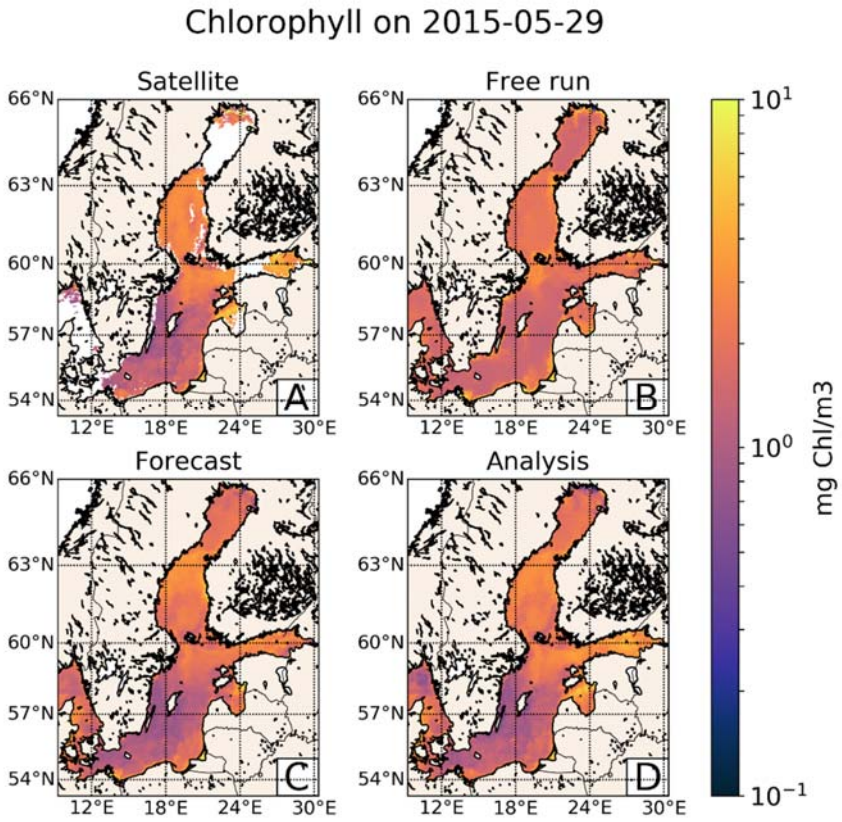


Fig. 4.1.3. Surface chlorophyll on May 29, 2015 analogous to Fig. 4.1.1

During March and April, the free-running model generally yields higher chlorophyll concentrations compared to the satellite data. Fig. 4.1.2 shows the situation on April 1, 2015. At this date, the free model run shows higher concentrations in the central Baltic, Gulf of Finland, and Bothnian Sea. In contrast, the concentration in the Southern Baltic around the station Arkona (13.87°E, 54.88°N) shows lower concentrations than the free run on March 3. The data assimilation process overall reduces the chlorophyll concentrations in accordance with the satellite data. The analysis step has a visible effect of reducing the concentration in particular between around 58°N and 60°N. The concentration in the free run decreases during May after the bloom and becomes more consistent with the satellite data. Fig. 4.1.3 shows the concentration on May 29, 2015 (this date is chosen because observations around the station Arkona are available on May 29, but not on May 31). Still, the forecast and analysis show lower concentrations than the free-running model ensemble in the central and southern Baltic and the Gulf of Finland. In contrast, the concentrations are slightly higher in the northern part of the Baltic (Bothnian Sea and Bothnian Bay). Overall, the data assimilation has a strong effect on the chlorophyll concentrations.

The data assimilation also directly influences the ensemble spread, hence the uncertainty estimate. Fig. 4.1.4 shows the ensemble standard deviation on April 1, 2015 for the free run and the assimilation

forecast. The free run exhibits a particularly high ensemble spread in the southern Baltic, the Gulf of Finland, and the northern coast in the Bothnian Bay. The Skagerrak and Kattegat and some coastal regions of Denmark, Germany, and Poland show the lowest standard deviation. The station Arkona is located at this date at the edge of lower and elevated standard deviations. The data assimilation strongly reduces the ensemble spread, hence the uncertainty, in most parts of the Baltic Sea. In the easternmost part of the Gulf of Finland and the north-eastern coast in the Bothnian Bay, a larger spread remains.

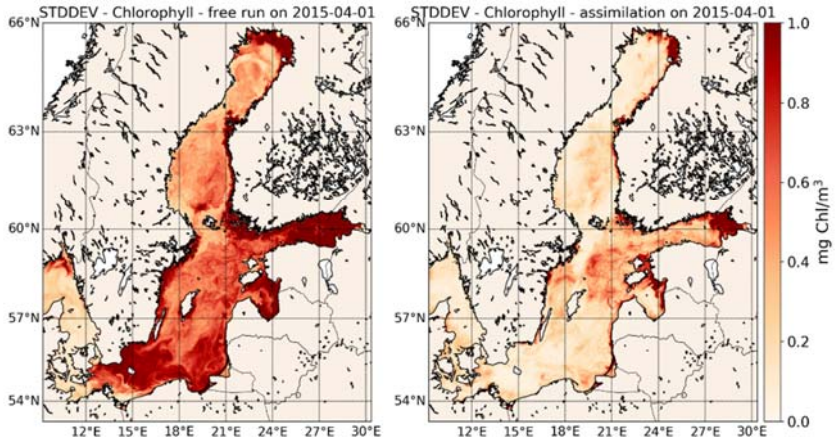


Fig. 4.1.4. Ensemble standard deviation of surface chlorophyll on April 1, 2015. Shown are the free run (left) and 24h forecast from the assimilation (right).

For the further discussion we now focus on the station Arkona. To assess the distribution of the ensemble states, Fig. 4.1.5 shows rank histograms at three different dates. The rank histograms are calculated using a quadratic region of 21x21 observation grid points centred at the station Arkona. The left-hand side of Fig. 4.1.5 shows the histogram for the forecast ensemble. For April 1, the histogram is skewed to the left, which is due to a bias in the forecast ensemble. This bias is even stronger on May 1 while after the bloom on May 29, the histogram is rather flat with a small tendency to the upper bins. The analysis strongly changes the ensemble distribution on April 1 and May 1. For all three dates, the histogram for the analysis is U-shaped. Thus, while a significant fraction of the bins is occupied, the ensemble has too little spread. Perhaps, this could be improved by using a smaller localization radius for the data assimilation. Since the data assimilation simulations are computationally very costly and time consuming, we were not yet able to perform tuning experiments for the localization.

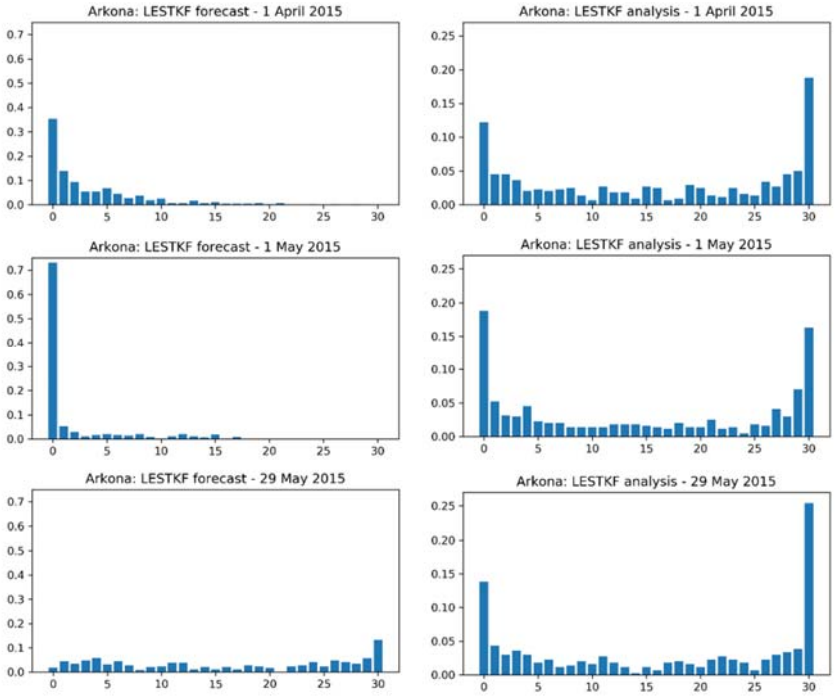


Fig. 4.1.5. Rank histograms for chlorophyll in a 21x21 grid point box centred at the station Arkona. Three different dates are shown. The left column shows the histograms for the 24-hour forecast, while the right column shows the analysis.

Figure 4.1.6 shows the time series of the surface chlorophyll concentration at the station Arkona. The free run (blue line) shows a bloom that starts in February and reaches its maximum of about 6 mg Chl/m^3 in March. In the second half of March, the bloom ends and the concentration decreases strongly to about 2.5 mg Chl/m^3. From the middle of April, the concentration decreases further to finally 1.5 mg Chl/m^3 at the end of May.

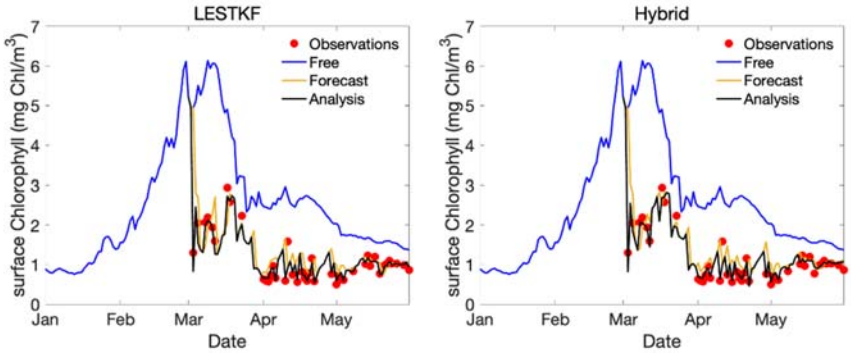


Fig. 4.1.6. Concentration of surface chlorophyll at the station Arkona over time. The blue line shows the free-running ensemble. The observations, when available, are shown as red dots. The assimilation is started on March 1, 2015. Shown are the 24-hour forecasts (orange) and the analysis (black).

We started assimilating the chlorophyll observations on March 1, which is too late to cover the onset of the bloom. Consistent with Figs. 4.1.1-4.1.3, the assimilation strongly reduces the concentration to a level that is comparable to the observational data. During the bloom, the concentration from the simulation with data assimilation reaches only about 2.5 mg Chl/m^3, which drops down to between 0.5 and 1.0 mg Chl/m^3 after the bloom. The end of the bloom is at about March 27 and hence about 10 days later than in the free run.

So far, we have only discussed the assimilation effect for the LESTKF method. The right panel of Fig. 4.1.6 shows the time series of the surface chlorophyll concentration at the station Arkona for the hybrid nonlinear-Kalman filter. Here only small differences are visible. Thus, the application of the hybrid filter does not lead to a significant difference in the surface chlorophyll at the station Arkona. However, we have to mention that again, due to the high computing cost, in the time allotted, it was not possible to perform tuning of the parameters regulating the hybridization.

Diagnostics on "non-observed" variables

With regard to non-observed variables, we first assess the effect of the data assimilation with depth. Figure 4.1.7 shows the profile of chlorophyll at the station Arkona for the free run, the LESTKF 24h forecast, and the analysis for the three months period March to May 2015. The free run shows a clear bloom during the first half of March, which reaches down to about 30m. At the beginning of April, there is an increase of deep chlorophyll reaching down to the bottom. This is likely related to an event of enhanced mixing, e. g. due to a storm. The data assimilation reduces the magnitude of the bloom, as was discussed before. The bloom is also shallower with 20-25m when the data assimilation is applied. In addition, the assimilation reduces the concentrations close to the bottom. The assimilation results in a subsurface chlorophyll maximum at a depth of around 20m during April and May. Unfortunately, we do not have validation data for chlorophyll, so the existence of a subsurface chlorophyll maximum cannot be verified. With regard to the effect of the assimilation, it appears that the subsurface maximum occurs because the concentrations close to the ocean surface are more strongly reduced than at a depth of around 20m.

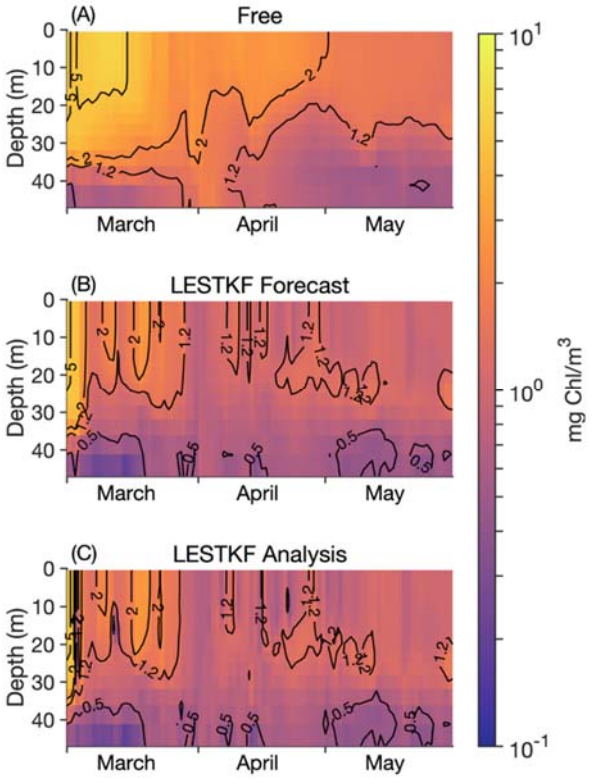


Fig. 4.1.7. Chlorophyll concentration over depth for the period March to May 2015. Shown are the (a) free run, (b) 24-hour forecast from LESTKF, and (c) LESTKF analysis.

The chlorophyll profile over time for the case of the hybrid filter LKNETF is shown in Figure 4.1.8. When the hybrid filter is applied, the bloom concentration at the surface is slightly larger in March compared to the LESTKF estimate. During April and May, the subsurface chlorophyll maximum is somewhat more pronounced. Further, while the concentrations close to the bottom are nearly equal for both data assimilation methods in March, the LKNETF results in higher bottom concentrations particularly in May. Thus, the hybrid filter appears to have a smaller effect in the lower layers compared to the LESTKF. Given the overall small differences of the LKNETF, we focus further discussion on the LESTKF.

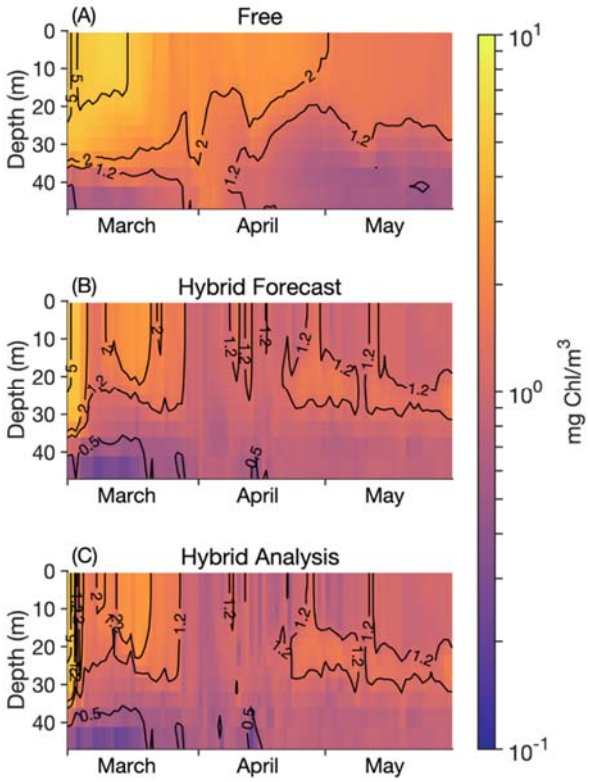


Fig. 4.1.8. Chlorophyll concentration over depth for the period March to May 2015. Shown are the (a) free run, (b) 24-hour forecast from the hybrid filter LKNETF, and (c) LKNETF analysis.

The changes in the chlorophyll concentration directly relate to changes in the concentrations of the three phytoplankton groups of ERGOM. These concentrations are changed in the analysis step through the ensemble-estimated cross covariances. Figure 4.1.9 shows on the left-hand side how the concentrations develop at the station Arkona over time. The concentration of diatoms, which are dominant at the beginning of March, is strongly reduced by the data assimilation. The concentration of the cyanobacteria is much lower and nearly unchanged. The flagellates initially show a tendency to increase, compensating partly for the reduced concentration of diatoms, but the concentration exhibits significant variations. The changed phytoplankton concentration induces a dynamic reaction of the zooplankton. As visible on the right-hand side of Fig. 4.1.9, the concentration of the two zooplankton groups decreases in the first half of March. Then, after increasing again in the second half of March, the concentrations further decrease during April and May. Overall, the zooplankton concentrations are lower in the case of the data assimilation compared to the free run. This behaviour is due to the fact that the assimilation reduces the phytoplankton and hence the food for the zooplankton. The change in the phytoplankton groups and the zooplankton changes both phytoplankton community structure and the trophic efficiency, as will be analysed for the SEAMLESS indicators further below.

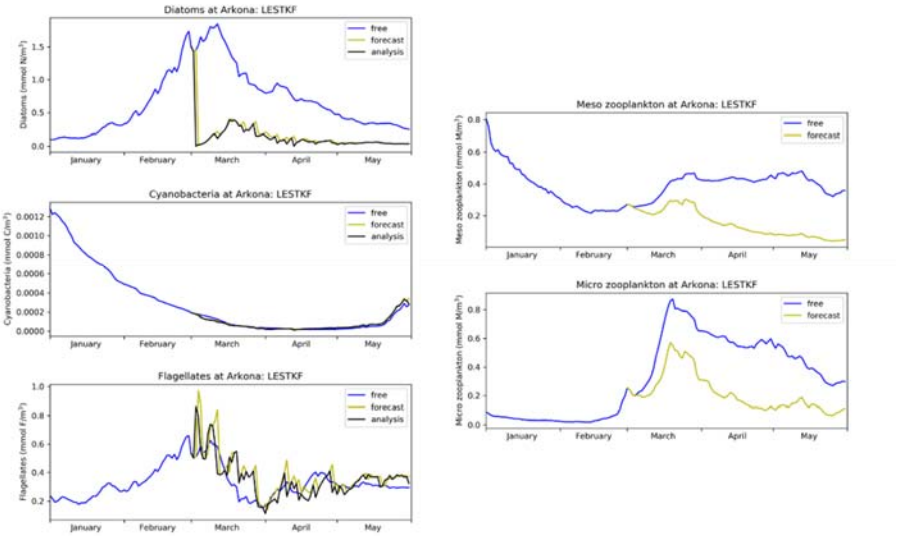


Fig. 4.1.9. Surface concentration of the three phytoplankton groups and two zooplankton groups around station Arkona.

Next to the phytoplankton and zooplankton, the nutrients react dynamically to the direct changes caused by the data assimilation. The effects are shown for the station Arkona in Fig. 4.1.10. The concentration of nitrate is slightly increased during the first half of March, which is likely due to the decreased uptake by phytoplankton. However, the overall change relative to the free run is very low. The change in phosphate is slightly larger with increased concentrations during the first half of March and lower concentrations afterwards. Significant changes are visible for ammonium and silicate. The dynamic reaction to the reduced concentration of diatoms is an increase of silicate, which is taken up now to a much smaller amount. In addition, the concentration of ammonium is significantly decreased due to the reduced abundance of zooplankton.

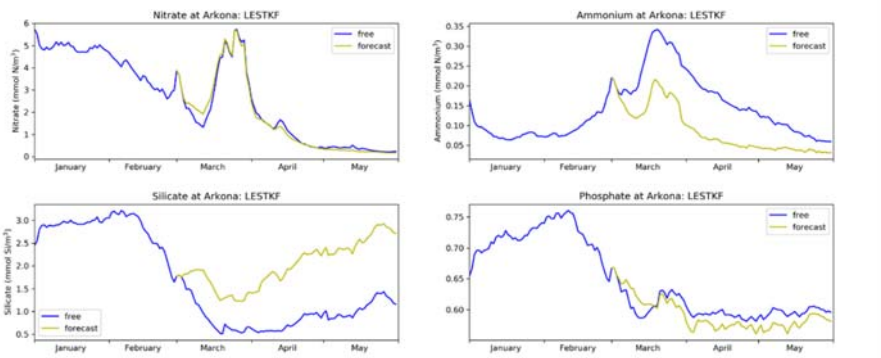


Fig. 4.1.10. Surface concentration of ERGOM nutrients at station Arkona. Shown are (top left) nitrate, (top_right) ammonium, (bottom left) silicate, and (bottom right) phosphate.

Diagnostics on derived quantities and indicators

Here we analyse different SEAMLESS indicators: phytoplankton phenology, plankton functional types, trophic efficiency, pH, dissolved oxygen, and primary production. Due to the shallowness of the station Arkona of only 46m, we omit the POC export from the analysis.

The impact of the data assimilation on phytoplankton phenology at the station Arkona can be analysed from Fig. 4.1.6. In particular, the data assimilation reduces the maximum amplitude of the bloom. Further, the bloom ends about 10 days later if the chlorophyll data is assimilated in comparison to the free run.

The time development of the other indicators is shown in Figure 4.1.11. While in the free run, the phytoplankton community is dominated by diatoms, the PFT ratio (Fig. 4.1.11 top left) is strongly changed by the data assimilation. Overall, the fraction of diatoms in the total phytoplankton is reduced from between 0.5-0.8 to 0.1-0.5 with the exception of the beginning data assimilation, where the PFT ratio is reduced to close to zero.

The trophic efficiency (TE, Fig. 4.1.11 top right) is influenced by both the reduction of the phytoplankton and the reduction of the zooplankton concentrations. Here, the data assimilation induces an increase in the TE from the beginning of the data assimilation process until April 7. After this date, the TE becomes lower than in the free run. The largest differences are in May, where the TE is reduced by more than 50%.

The change in pH is generally below 0.02 as is visible in the lower left panel of Fig. 4.1.11. Initially, the pH is reduced compared to the free run. However, from mid-March, the pH is larger in the assimilation than in the free run.

The effect of the data assimilation on the concentration of dissolved oxygen is negligible as the lower right panel of Fig. 4.1.11 shows. For the vertically integrated primary production (PP, Fig. 4.1.11 bottom centred panel), we only have daily data during the assimilation period, while during the spin-up, the values are only available on each 10th day. The PP shows fluctuations, but in general no systematic difference between the free run and the assimilation. The PP is computed by ERGOM as a flux during the model integration. Thus, the change in the phytoplankton concentrations by the data assimilation is not accounted for as a change in PP. The small change in the PP indicates that the biological production is not significantly changed by the assimilation, even though the assimilation reduces the chlorophyll and diatom concentrations.

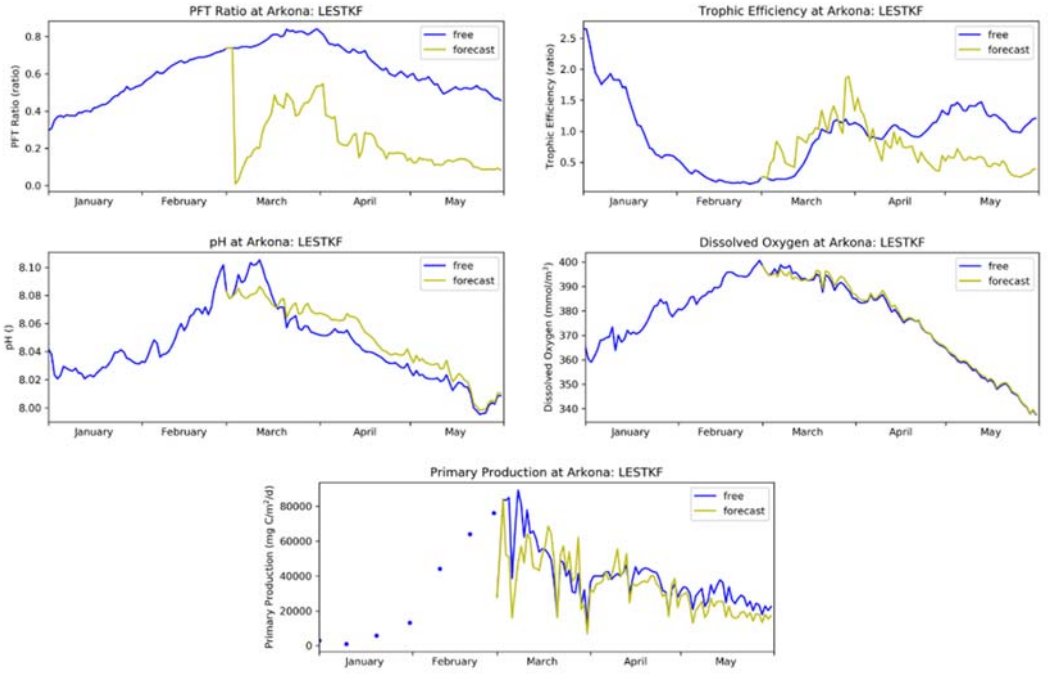


Fig. 4.1.11. SEAMLESS indicators at the station Arkona.

In order to obtain a wider view of the effects of the data assimilation, Fig. 4.1.12 shows the pH over the full Baltic Sea for April 1, 2015 (top) and May 29 (bottom). Both the free run and the 24h forecast from the assimilation are shown. Generally, the pH is lower in the northern area of the Baltic (Bothnian Sea and Bothnian Bay). Particularly high values of pH around 8.1 are visible in the Gulf of Finland. On April 1, such high values are also visible further west. Here, the data assimilation reduces the pH. On May 29, the pH is generally lower in most parts of the Baltic Sea compared to April. The data assimilation has little influence at the end of May, except for an increase of pH in the Gulf of Finland. Generally, the influence of the data assimilation at the station Arkona shown in Fig. 4.1.11 appears to be representative of pH for the southern Baltic Sea, while the Skagerrak and Kattegat as well as the Baltic Sea north of 57°N appear to be distinct.

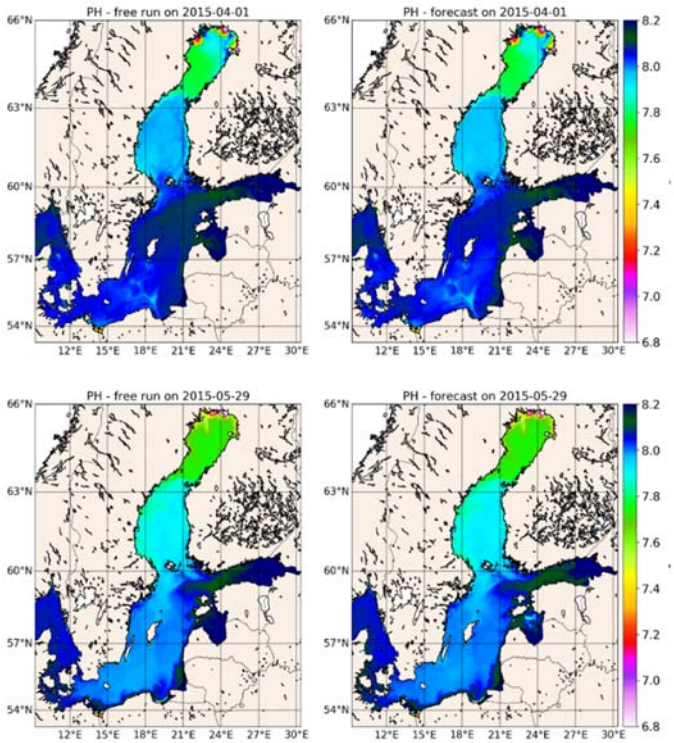


Fig. 4.1.12. pH at the ocean surface on April 1 (top) and May 29 (bottom). The left column shows the free run while the right shows the 24h forecast from the LESTKF assimilation.

The effects on the PFT ratio are on a finer scale as is visible from Fig. 4.1.13. On April 1 (upper row), the PFT ratio is above 0.8 in most of the large basins of the Baltic Sea. A particular situation is seen close to the coast in the northernmost part of the Baltic. Here the PFT ratio is strongly reduced, which is caused by the dominance of cyanobacteria. In contrast, the PFT ratio is smaller in the Skagerrak and Kattegat due to lower concentrations of diatoms in combination with higher concentrations of flagellates. The data assimilation overall reduces the PFT ratio. Here the effect is larger in the Southern and central Baltic than further north in the Bothnian Sea and Bay. An exception is the Gulf of Finland where the PFT ratio is increased by the data assimilation.

On May 29 (bottom row of Fig. 4.1.13), the PFT ratio is overall smaller than on April 1. In the Gulf of Finland, the ratio only reaches 0.2, which is caused by a dominance of flagellates. The data assimilation mainly reduces the PFT ratio as on April 1. However, the effect is small in the Gulf of Finland and locally an increase is visible, e.g., in the central part of the Bothnian Sea or in the southern Baltic around 18°E, 56°N. This change is caused by a local decrease in the flagellate concentration. In the southern Baltic west of 16°E, the data assimilation causes a particularly low PFT ratio. Here the assimilation leads to

higher concentrations of flagellates but a lower concentration of diatoms, which is consistent with Fig. 4.1.9.

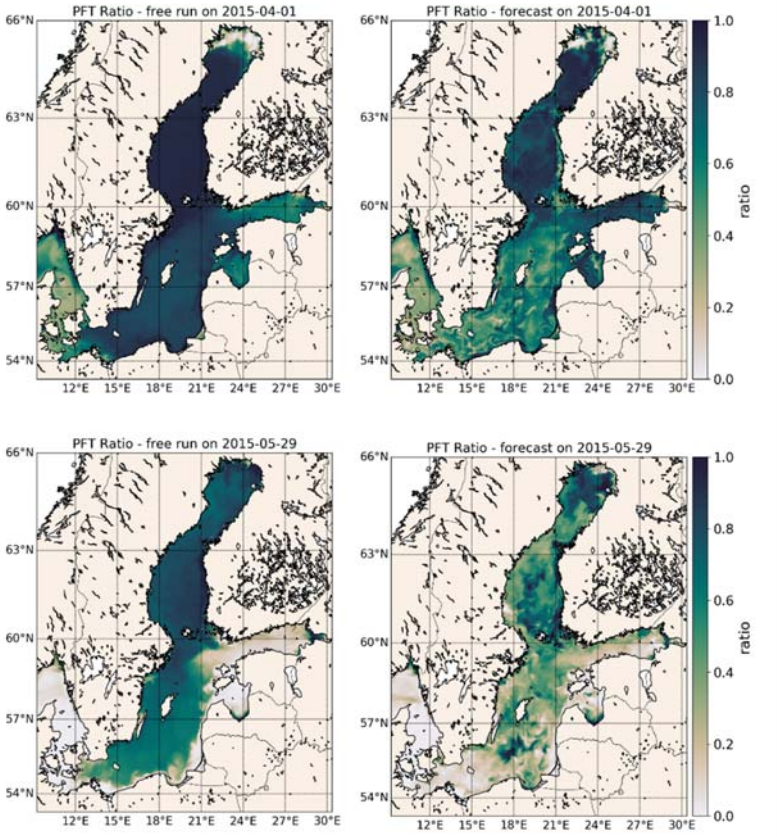


Fig. 4.1.13. PFT ratio at the ocean surface on April 1 (top) and May 29 (bottom). The left column shows the free run while the right shows the 24h forecast from the LESTKF assimilation.

Finally, we assess the trophic efficiency in the Baltic Sea. Figure 4.1.14 shows that on April 1, the TE is generally below 0.5 in most of the Baltic Sea. The TE increases west of 15°E, thus also at the station Arkona, and also shows elevated values in the transition zone toward the North Sea (Skagerrak and Kattegat). The data assimilation mainly increases the TE in Kattegat. On May 29, the relative fraction of zooplankton increased compared to April 1. Now, TE values between 2 and 3 are visible in the Gulf of Finland and the Gulf of Riga, while the southern Baltic exhibits a TE of around 1. The assimilation leads to lower TE in the southern Baltic including around the Arkona station (see Fig. 4.1.11). However, the TE is increased in the Gulf of Riga and the Gulf of Finland.

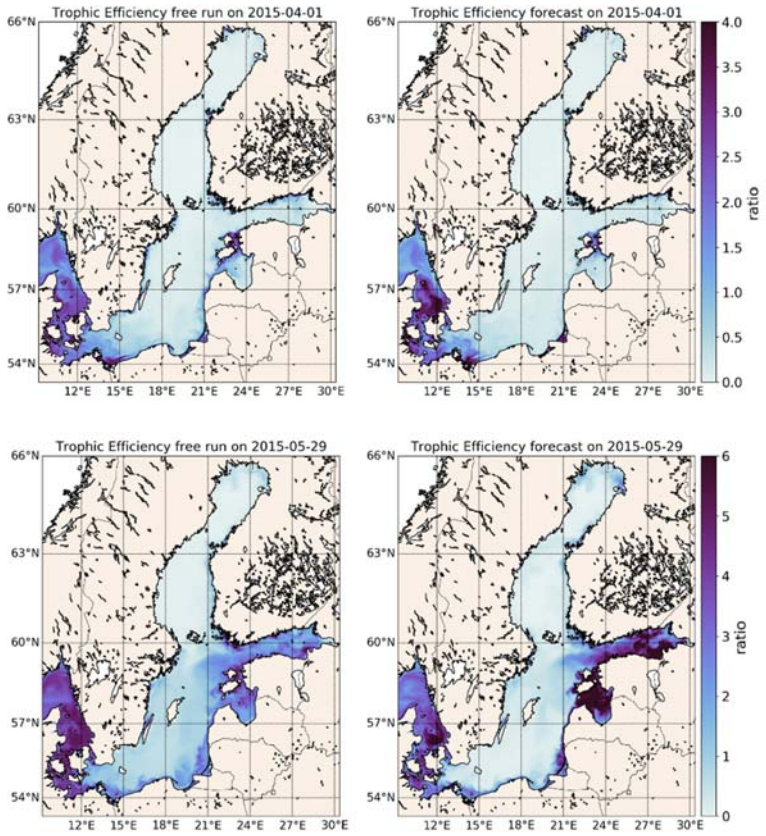


Fig. 4.1.13. Trophic efficiency at the ocean surface on April 1 (top) and May 29 (bottom). The left column shows the free run while the right shows the 24h forecast from the LESTKF assimilation.

4.2 Assimilation results in the GLO and IBI MFC domains

This section presents the results obtained by **IGE/UGA** as part of Task 3.3c using the assimilation method and related developments outlined in section 3.2.

Assimilation setup in GLO and IBI regions. The stochastic version of NEMO-PISCES has been developed and implemented in the global ocean configuration at ¼° inherited from the CMEMS GLO MFC. The generation of a 40-member ensemble was produced assuming uncertainty sources originating from (i) critical biogeochemical model parameters of the PISCES stochastic formulation, (ii) sub-grid scale effects associated to ¼°, eddy-permitting horizontal resolution, and (iii) location uncertainties of ocean mesoscale structures and associated advective/diffusive fluxes to anticipate coupled physical/biogeochemical data assimilation experiments in WP4 (see Annex I for more details). In the current setup, the perturbed PISCES parameters are linked to phytoplankton growth (i.e., to light and temperature forcing and maximum growth rates) and zooplankton processes, consistently with the main findings of the PISCES parameter sensitivity study of Task 3.2 that relies on BATS and L4

sites characteristics (see D3.2, part 2) as well as the prior study by Garnier *et al*. (2016) which was primarily guided by surface chlorophyll observability rationale .

Before activating the stochastic model, a spin-up was produced using the deterministic NEMO-PISCES model from 01.01.2017 to 22.12.2018 and initial conditions from a reanalysis produced by Mercator Ocean international (MOi) for the CMEMS (namely, the "BIOMER" run). The stochastic ensemble simulation was then initialized on 22 December 2018, and further integrated to cover the full 2019 year using unperturbed ERA5 atmospheric forcing. The resulting 40-member ensemble represents a probabilistic view of the 2019 seasonal cycle in the global and North Atlantic Ocean. The prior ensemble statistics in the North Atlantic basin are described in terms of spread, median, min and max distributions on Fig. 4.2.1, after 6 months of stochastic perturbations.

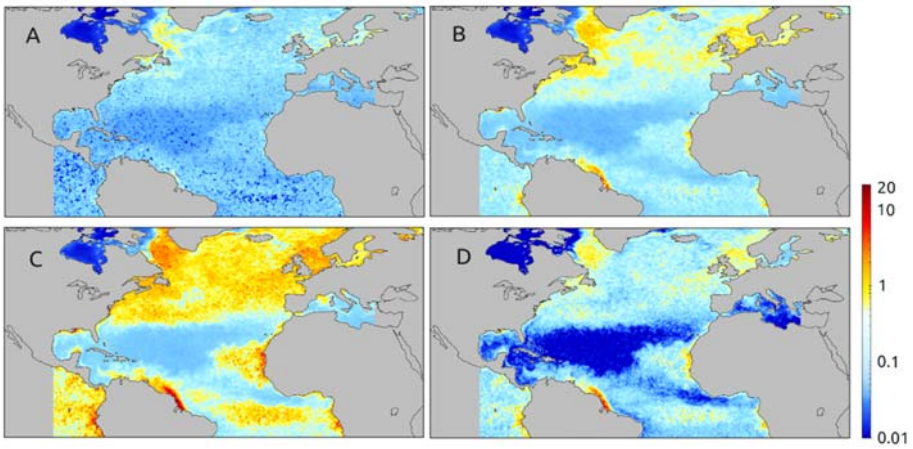


Fig. 4.2.1. Statistics of the prior ensemble, generated by the stochastic version of the NEMO-PISCES in the North Atlantic basin: spatial distributions of minimum (A), median (B), maximum (C) and standard deviation (D) of the ensemble of surface chlorophyll concentrations (in mg/m3).

The spatial variability of the ensemble statistics shows that the impact of the different sources of uncertainty considered in the stochastic model exhibits very different footprints from one region to another, reflecting the different productivity regimes in the North Atlantic biogeochemical provinces (i.e., the oligotrophic subtropical gyre, the West African upwelling, the Amazon discharge area and the large productive region of the North Eastern basin). When compared to L4 ocean colour (OC) data, the ensemble spread seems relevant to encapsulate OC observations (e.g., not reproduced here), however a more precise deciphering is needed on a regional scale.

The prior 40-member ensemble is then conditioned by CMEMS L3 ocean colour observations (daily product at 4km resolution combining MODIS, VIIRS and Sentinel-3A/B sensors) available over 2019 and analysed in a North Atlantic sector of the GLO and IBI MFC domain, around PAP station (48°50'N, 16°30'W, Hartman *et al*., 2021). The space-time localization scheme is applied using the local ensemble parameterization with a cutting length/time of 100 km/30 days and a correlation length/time of 40 km/10 days. The impact of data assimilation is then assessed by comparing statistics of the prior and posterior ensembles of model state variables (surface and sub-surface chlorophyll

concentrations), as well as on indicators directly related to plankton production (primary production, phenology, trophic efficiency, POC).

Diagnostics on observed variables

Figure 4.2.2 compares the priori and posterior ensemble reconstruction of a continuous time series describing the surface chlorophyll concentration as well as its uncertainty at PAP station. The comparison of the prior and posterior time series indicates that the assimilation of L3 data significantly improves the consistency of the ensemble mean with the equivalent time series derived from the unassimilated L4 CMEMS product collocated at PAP (in particular in terms of timing and amplitude of the spring bloom), and reduces the uncertainty of the estimated concentration. The ratio of the posterior/prior ensemble variance is used to quantify the reduction of uncertainty. Its value in the PAP region is 0.148 (i.e., a ratio of 38.5% of the ensemble standard deviation) in the anamorphic space when assuming a 30% observation error variance in the inversion scheme.

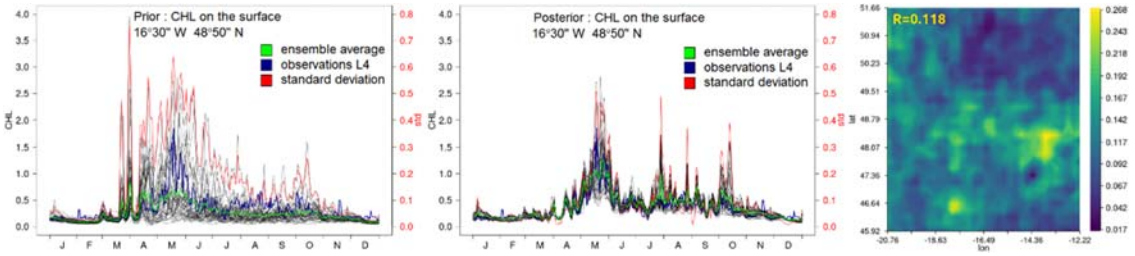


Fig. 4.2.2. Time series of the surface chlorophyll concentration (in mg/m3) at 16°30' W, 48°50'N (PAP station) for the prior (left panel) and posterior (central panel) ensembles. The black curves show the 40-member ensemble; the green curve is the ensemble mean; the blue curve is the L4 CMEMS OC product collocated at PAP; the red curve is the standard deviation. Right panel: ratio between posterior and prior ensemble variance in the region around PAP on May 15th (averaged value R = 0.118).

In order to assess how effective is the ensemble assimilation around PAP, additional probabilistic scores are computed using the EnsScores library (see Deliverable D3.1) to check the statistical consistency against L3 satellite products. The computed continuous rank probability score (CRPS; see D3.1) metrics (decomposed into reliability and resolution skill scores and rank histograms) are shown in Figs. 4.2.3 and 4.2.4, respectively. It shows that the assimilation improves the CRPS score throughout the year, except for a specific period in October and November where the ratio is slightly above 1. This could be related to the bias and non-Gaussianity of the prior ensemble pdf.

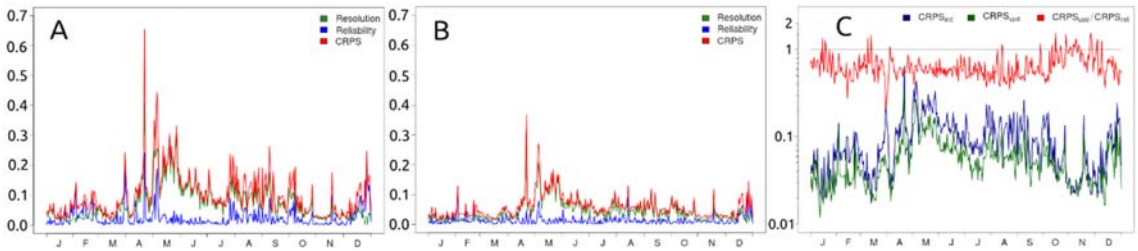


Fig. 4.2.3. CRPS scores for the PAP region (1100 km x 720 km centred on 16°30' W, 48°50'N). A: CRPS (red curve) score split into reliability (blue curve) and resolution (green curve) components for the prior ensemble. B: same statistics for the posterior ensemble. C: The CRPS score for the prior and posterior ensembles (blue and green curves correspondingly) and their ratio (red curve).

A necessary condition for expecting consistent inversion results is the reliability of the prior ensemble pdf. A consistency check of the prior/posterior ensemble pdfs against L3 CMEMS OC data is performed by computing rank histograms in the PAP region during 4 periods of the year (mid-February, mid-May, mid-August, mid-November). A flat histogram would represent a perfectly reliable ensemble (i.e., the verification data have the same probability to fall in every interval defined by the ensemble). Fig 4.2.4 shows that the assimilation tends to flatten histograms for the 4 considered periods, with nevertheless some tendency to transform a prior negative bias into a positive one after inversion.

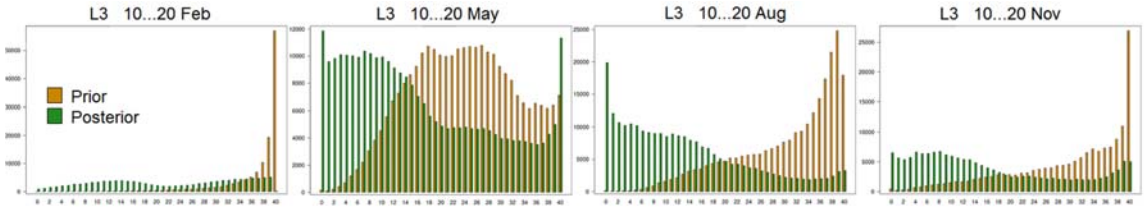


Fig. 4.2.4. Rank histograms for the PAP region (1100 km x 720 km centred on 16°30' W, 48°50'N) calculated on the basis of L3 CMEMS OC observations for short 11-days periods in February, May, August and November. The brown bars represent the prior ensemble, the green bars represent the posterior ensemble.

The above diagnostics indicate that, overall, the uncertainties accounted for in the stochastic model are a relatively good fit for the PAP region. In particular, one can expect a reasonable controllability level for some unobserved variables and indicators in that region. In other regions, the rank histograms obtained under the same conditions are sometimes less well conditioned (around BATS station in particular, not shown here), which suggests that the assumptions about the sources of uncertainty should be revised to generate a more relevant prior ensemble in these regions.

The analysis in PAP region is further extended to unobserved variables and indicators with a particular focus on the spring bloom period (around May 15), which is consistently represented in the ensemble pdf and of particular biogeochemical relevance.

Diagnostics on "non observed" variables

These diagnostics inform on the capacity of the 4D assimilation scheme to propagate information at local scale from observed to non-observed state variables, as well as in space (on the vertical) and time (backward and forward) given the incomplete coverage of L3 products (due to clouds). Figure 4.2.5 illustrate the spatial structures that emerge from the patchy data in a sequence of 5 days during the spring bloom period, out of the smooth "first guess" prior.

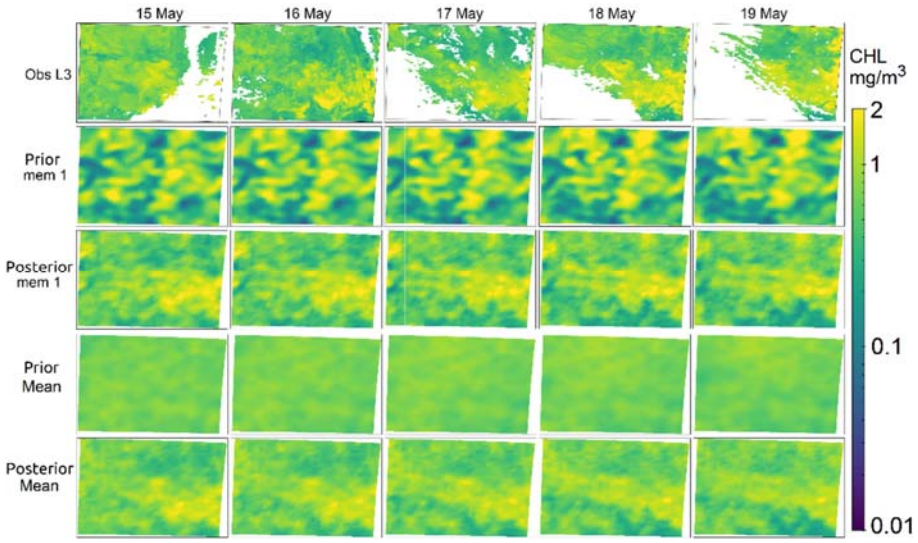


Fig. 4.2.5. Surface maps of chlorophyll concentration (in mg/m3) around PAP station during 5 successive days (from May 15 to May 19) impacted by clouds: (1st row) assimilated CMEMS L3 product at 4km resolution; (2nd row) first member of the prior ensemble; (3rd row) first member of the posterior ensemble; (4th row) prior ensemble mean; (5th row) posterior ensemble mean.

On the vertical, Figure 4.2.6 shows the impact of the surface observation on the ensemble mean and the associated reduction of uncertainty on the vertical, which mainly takes place in the top 30 meters (the averaged concentration increases in the top 40 meters). One can notice that while the prior was

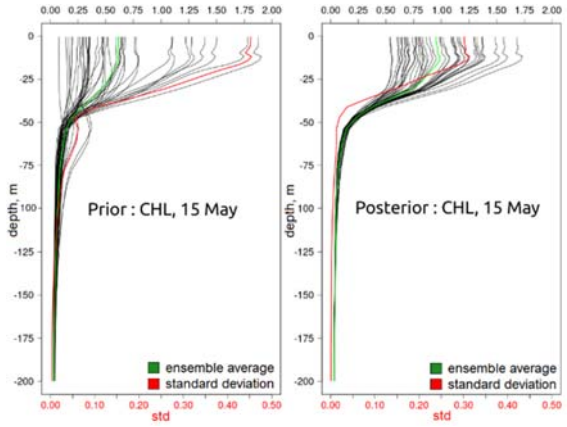


Fig. 4.2.6. Vertical distribution of the chlorophyll concentration 16°30' W, 48°50'N (PAP station). The black curves are the vertical profiles of the 40 ensemble members; the green curves are the ensemble mean; the red curves are the ensemble standard deviation. Prior and posterior distributions are presented for 15 May.

exhibiting a sub-surface maximum at ~ 60 meter for only a few members, the posterior ensemble doesn't represent such behaviour anymore. For this period of the year, a shallow mixed layer has been settled (~ 20 m depth) so that the vertical stratification seems to act as a "vertical barrier" on the control of -chlorophyll?. These results, together with the time series shown on Fig. 4.2.2, strongly

suggest that phenology indicators can be considered as faithfully controlled by the assimilation process.

Diagnostics on derived quantities and indicators

The assimilation experiment is now further assessed in terms of impact on other SEALMESS indicators related to phytoplankton dynamics: downward flux of particulate organic carbon (POC), Trophic Efficiency in the upper 200 meters and vertically integrated Primary Production.

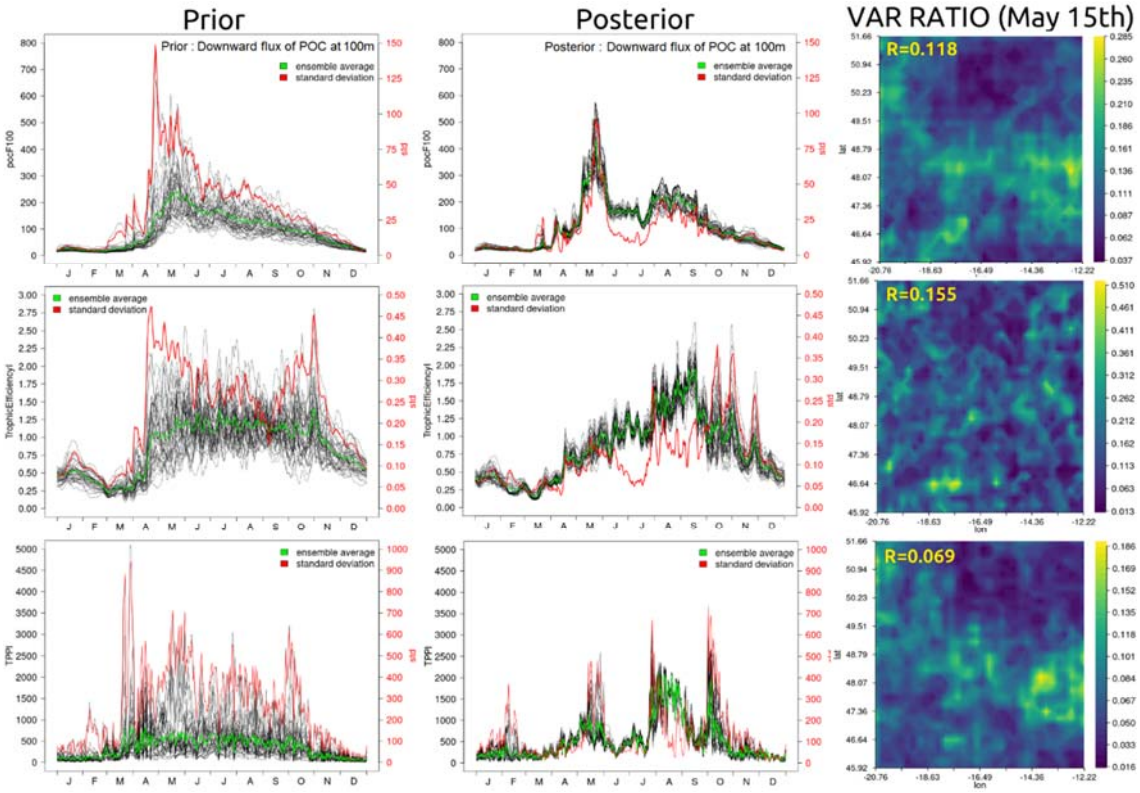


Fig. 4.2.7. Time series of (i) (upper row) downward flux of the particulate organic carbon at 100-meter depth (in mg.C/m2/day); (ii) (middle row) Trophic Efficiency (ratio of vertically integrated values of total zooplankton and phytoplankton biomass between 0- and 200- meter depth) and (iii) (lower row) vertically integrated total Primary production (in mg.C/m2/day) at 16°30' W, 48°50'N (PAP station), for the prior (left panel) and posterior (central panel) ensembles. The black curves show the 40-member ensemble; the green curve is the ensemble mean; the red curve is the standard deviation. Right panel: ratio between posterior and prior ensemble variance in the region around PAP on May 15th (R is the average ratio over the area).

The statistics shown in Fig. 4.2.7 suggest that, in the PAP region, the biogeochemical processes at depth are significantly influenced by the plankton dynamics in the upper layer constrained by the assimilated surface OC data. The variance reduction (computed in the anamorphic space) is significant for the 3 indicators throughout the year, with however, occasional events characterized by persisting dispersion of the members even after assimilation. At the time of the spring bloom (May 15th), the ratio between posterior and prior ensemble variance for POC is similar to the corresponding value obtained for Surface Chlorophyll (Fig 4.2.2) while the ratio is slightly larger for Trophic Efficiency.

We also note that some uncertainty maxima in the posterior pdfs do not occur at the same time for the different indicators, and do not systematically align with the uncertainty maxima of surface chlorophyll (Fig. 4.2.2), indicating that the model dynamics play its role in the propagation of uncertainties in space, time and across model variables. For example, the POC flux exhibits significant uncertainty at the time of the spring bloom, while uncertainty in trophic efficiency and primary production reaches maxima more towards the end of the year. However, some caution is needed in the comparison of statistics between different seasons given that the ensemble reliability against OC data is degraded towards the end of the year (see CRPS and rank histograms).

More generally, it should be recalled here that all interpretations of the statistics obtained on the derived quantities are based on the assumptions made about the sources of uncertainty and the way they are accounted for in the NEMO/PISCES stochastic model. In addition, we used the uncertainty reduction measured by the ensemble spread as a proxy for the assessment of the OCI properties. This is an approximation that is all the better as the simulations are consistent with our prior knowledge, and the probability distributions of the ensembles are consistent with those described by independent observations. The first evaluations of the observability/controllability properties obtained in the particular case of the experiments conducted here will need to be refined, ideally in situations where independent observations are available, but the approach provides a methodological guide that can be replicated with CMEMS systems operated in real conditions.

4.3 Assimilation results in the NWS MFC domain

Assimilation set-up on NWES domain

In this section are presented two 4-month experiments performed by PML for the Spring bloom period of March-June 2018, using the hybrid ensemble-3DVAR system described in section 3.3: (i) a simulation where only OC chlorophyll has been assimilated into the model and (ii) a simulation assimilating SST, EN4 Hadley profiles for temperature and salinity (we call all these further “physical data”), along with L3 product for OC chlorophyll, in a weakly coupled approach that corresponds to the present operational system. These two simulations have been compared with the current operational 3DVAR system of the CMEMS MFC NWS and have been validated using (i) glider data from the missions of the NERC AlterEco project, (ii) data of the monitoring station “L4” in the English Channel and (iii) North Sea Biogeochemistry Climatology (NSBC) data. Furthermore, rank histogram metrics have been used to assess the spread of the ensembles. It should be noted that the 3DVAR system assimilated all the physical data and OC chlorophyll. However, the difference in simulated chlorophyll is negligible between the 3DVAR weakly coupled physical-OC chlorophyll assimilation, and the 3DVAR assimilation of OC chlorophyll (e.g., Skakala *et al*., 2022).

Diagnostics on observed variables

Fig.4.3.1 demonstrates that the new hybrid ensemble-3DVAR produces different results from the operational 3DVAR set-up. The spread of the chlorophyll ensemble is quite substantial, and amounts in large parts of the domain to > 100% of the mean (Fig.4.3.2). The rank-histogram (Fig.4.3.3, right-hand panel) suggests that the posterior chlorophyll ensemble might have unrealistically large spread. We observed that the large spread of the ensemble is due to 1-2 sensitive parameters for the

phytoplankton functional types (PFTs) and reducing the interval of perturbations (currently at 30% around the parameter values) can be considered in future applications. The rank histogram in the left-hand panel of Fig.4.3.3 suggests that the analysis state is biased relative to the assimilated OC chlorophyll. This behaviour is consistent with the bias of the 3DVAR and hybrid DA chlorophyll shown in Fig.4.3.1, and the negligible bias between 3DVAR and the assimilated OC satellite data (e.g., Skakala *et al*., 2022).

Fig.4.3.2 and Fig.4.3.4 demonstrate that, in the hybrid system, physical DA has a major impact on the chlorophyll also in the situation when OC chlorophyll is assimilated to the model. Such impact of physical DA on chlorophyll was not evident in the previous 3DVAR system. In the hybrid system, the physical DA seems to increase the overall chlorophyll concentrations (Fig.4.3.4) and to drive chlorophyll much deeper into the ocean than either the 3DVAR, or the hybrid ensemble-3DVAR assimilation of only OC chlorophyll. The impact of temperature assimilation on chlorophyll is assumed to happen mainly through its influence on mixing, but the exact mechanism which produced this increase in primary productivity is not yet well understood and needs to be investigated further.

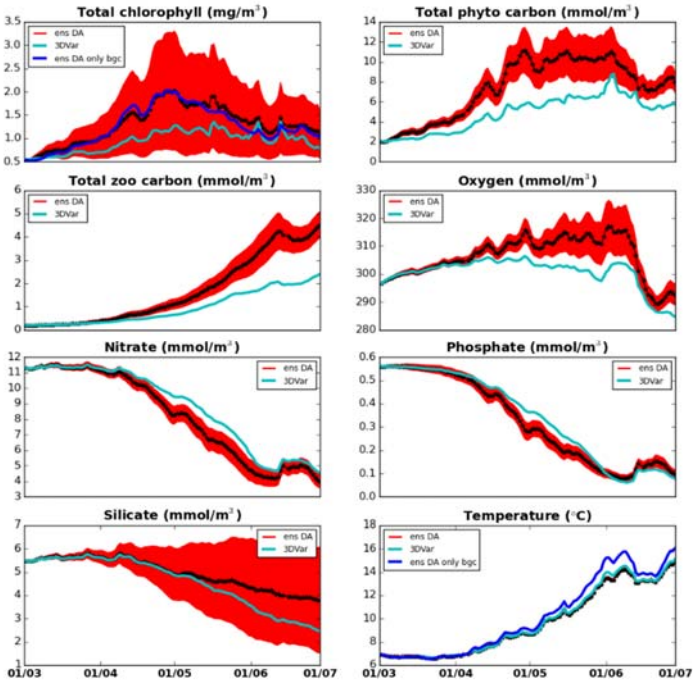


Fig.4.3.1. Comparing the hybrid ensemble-3DVAR simulation for the (i) weakly coupled physical DA – OC chlorophyll DA run ("ens DA", red band for the spread and black line for the mean) with (ii) the ensemble mean of the same DA system, but assimilating only OC chlorophyll ("ens DA only bgc", blue line), (iii) with the established 3DVAR assimilation ("3DVAR", turquoise line). Shown are surface values averaged throughout the NWS shelf.

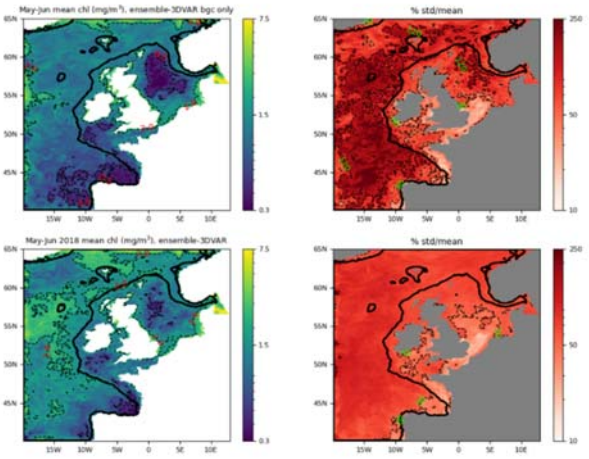


Fig.4.3.2. Surface chlorophyll averaged through May-June 2018 on the NWE Shelf (left hand panels) for the hybrid ensemble-3DVAR run assimilating OC chlorophyll only (upper panels) and the weakly coupled physical DA – OC chlorophyll DA run (bottom panels). Right hand panels show the ensemble spread (standard deviation) compared to the mean value (in %).

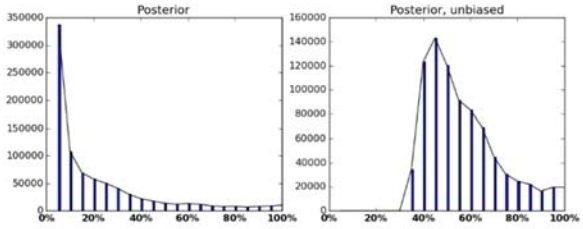


Fig.4.3.3. Rank histograms calculated using OC chlorophyll data for the first half of May 2018. The histogram on the right-hand side has been calculated from the unbiased observations. The histograms for other periods in May and June 2018 looked very similar to the plots shown here.

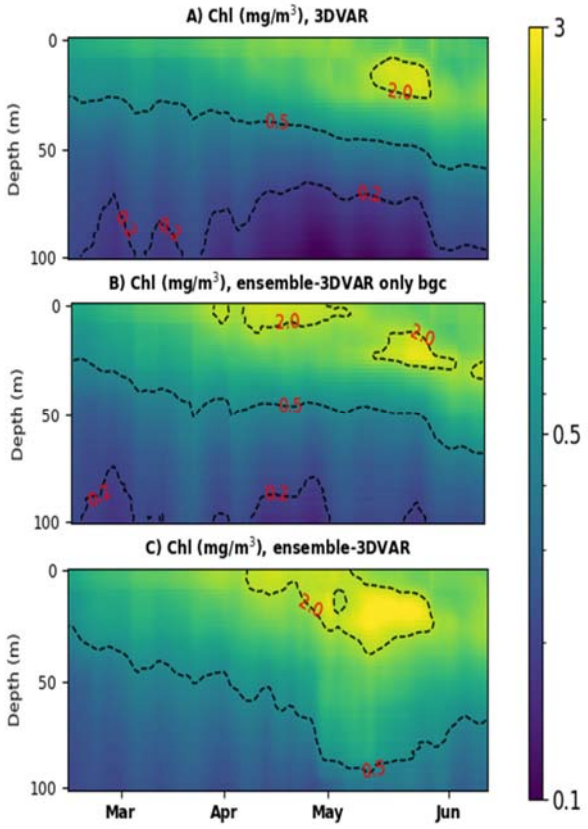


Fig.4.3.4. The Hovmoller plots comparing the horizontally averaged chlorophyll values throughout the NWE Shelf between the 3DVAR weakly coupled assimilation of physical data and OC chlorophyll (upper panel), the hybrid ensemble-3DVAR assimilation of OC chlorophyll (middle panel) and the weakly coupled hybrid ensemble-3DVAR assimilation of physical data and OC chlorophyll (bottom panel).

The different simulations have been validated with glider data of the UK NERC AlterEco campaign from May-June 2018 in the central North Sea (Fig. 4.3.5), as well as with data from station "L4" in the English Channel (Fig. 4.3.6). Fig. 4.3.5 demonstrates that OC chlorophyll assimilation using the hybrid system improves the model chlorophyll relative to the established 3DVAR assimilation (see also Tab.4.3.1). However, including physical data into the assimilation substantially degrades the model skill in simulated chlorophyll, also relative to the skill of the 3DVAR system. The comparison at the L4 station shows that the hybrid systems capture much more sensibly phytoplankton phenology (i.e., the bloom around the end of March) than the 3DVAR system. However, hybrid systems have larger overall bias than the 3DVAR system (Tab.4.3.1).

Finally, we compared the runs also using the NSBC in situ data (monthly climatology binned in 3D, https://www.cen.uni-hamburg.de/), which were collected over the 1960-2014 period. Such comparison has some limitation, due to temporal trends and potentially large interannual variability that are missed in climatological datasets, but in the absence of *in situ* time series, the climatology can be still considered as to some degree indicative of spatial variability. Tab.4.3.1 indicates that the hybrid systems perform in chlorophyll overall better when compared to NSBC data than the 3DVAR run.

Data type	**Run**	**Bias** (mg/m^3)	**BC-RMSD** (mg/m^3)
glider	3DVAR	0.266	1.223
	hybrid bgc only	0.096	1.308
	Hybrid	1.445	1.817
L4	3DVAR	0.363	1.883
	hybrid bgc only	1.613	2.378
	Hybrid	0.71	1.335
NSBC	3DVAR	-0.746	1.923
	hybrid bgc only	-0.599	2.393
	Hybrid	-0.251	2.478

Tab.4.3.1. Bias and Bias-Corrected Root Mean Square Error (BC-RMSD) metrics comparing the model skill in chlorophyll for different simulations: 3DVAR, hybrid DA assimilating only OC chlorophyll ("hybrid bgc only") and weakly coupled physical-OC chlorophyll hybrid DA ("hybrid"). The 3DVAR simulation assimilated both physical data and OC chlorophyll, but it has been known for some time that for 3DVAR, physical data assimilation leads to a negligible impact on chlorophyll, if also OC chlorophyll is included (Skakala et al., 2021, 2022).

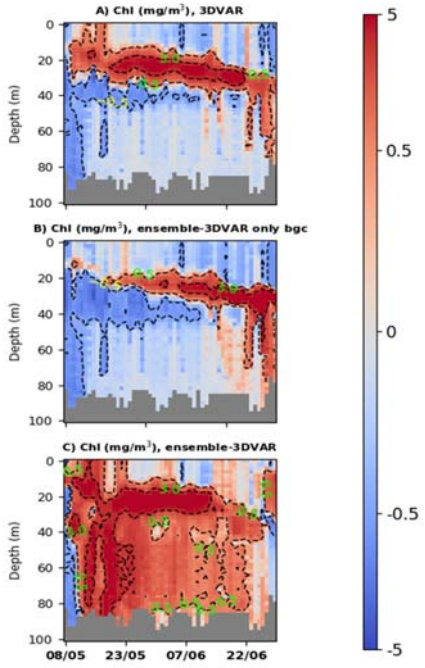


Fig 4.3.5. Validation of the different runs with AlterEco gliders from the central North Sea. The different panels show the difference in the total chlorophyll between the assimilative runs and the glider data. In the upper panel it is the 3DVAR system assimilating OC chlorophyll, in the middle panel it is hybrid 3DVAR-ensemble assimilation of OC chlorophyll and in the bottom panel the weakly coupled physical DA and OC total chlorophyll DA using the hybrid ensemble-3DVAR system.

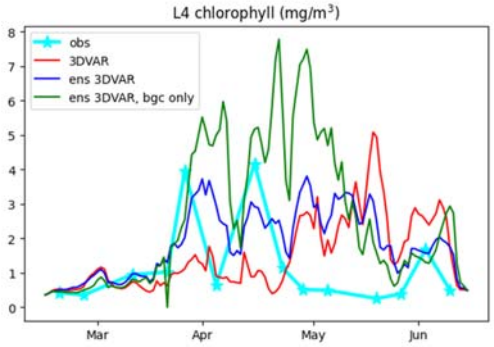


Fig 4.3.6. Validation of simulated chlorophyll at the monitoring station "L4" in the Western English Channel. The assimilative runs are compared with the L4 observations.

Diagnostics on "non-observed" variables

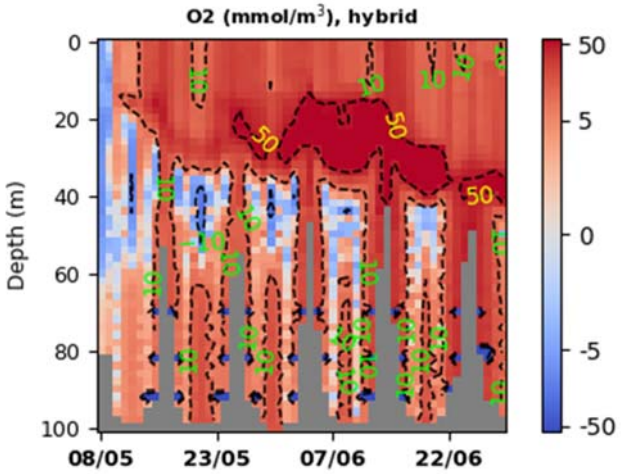


Fig.4.3.7 Validation of the oxygen from the reanalysis assimilating only OC chlorophyll using the ensemble-3DVAR system, by comparing it to the observed AlterEco glider data in the central North Sea. The Figure shows the reanalysis minus glider differences (in mmol/m^3).

One of the very few validation data available for the "non-observed" variables, are the oxygen data from the AlterEco 2018 mission. These observations are from gliders operating during the May-June period in the central North Sea. The comparison of oxygen from the ensemble-3DVar run, assimilating only the OC chlorophyll data, with the glider data is shown in Fig.4.3.7. It can be seen that the ensemble-3DVar oxygen skill is comparable to the skill of the 3DVar system that can be found in Fig.11 of Skakala et al, 2021.

Diagnostics on derived quantities and indicators

The outputs of our 4 month runs provided data for the following SEAMLESS indicators (see Fig.4.3.1): phytoplankton phenology, trophic efficiency, phytoplankton community structure and oxygen. In Fig.4.3.8 provides additional indicators (POC, net primary production, pH) simulated with a weakly coupled ensemble-3DVar that included assimilation of biogeochemical glider data. These results were produced in the framework of WP5, but were already available at the time of writing the present report. Since the results are relevant here, we decided to anticipate them into this discussion, and they will be discussed further in the report to WP5. The observability of the indicators was estimated from the differences between the ensemble-3DVar and the corresponding free run and 3DVar run (Fig.4.3.1 and Fig.4.3.8).

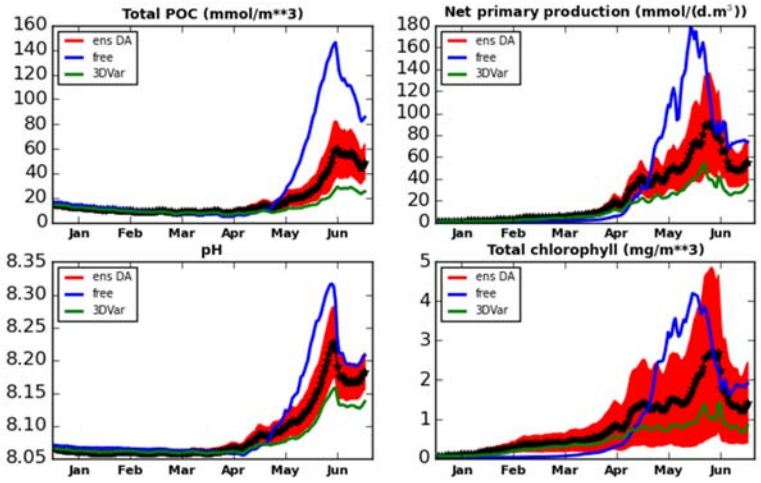


Fig.4.3.8. We compare 4 specific indicators between the ensemble-3DVar weakly coupled run that includes assimilation of glider data (better discussed in deliverable to WP5) with the 3DVar simulation (green) and the model free run (blue). The time series show surface values averaged across the NWES.

From the indicators, it is clear that assimilation of chlorophyll has a major impact on surface phenology (both Fig.4.3.1 and Fig.4.3.8). Consequently, since both chlorophyll and carbon are directly impacted by the OC chlorophyll assimilation, the DA has a major impact on the net primary production as well (Fig.4.3.8). Since assimilation of chlorophyll has impact both on the ERSEM total phytoplankton and zooplankton carbon biomass (Fig.4.3.1), it can be anticipated that it also substantially impacts their ratios, hence the trophic efficiency. It also impacts significantly oxygen (Fig.4.3.1) and POC (Fig.4.3.8, hence we can expect that also for POC fluxes). The indicators which are less significantly impacted by the chlorophyll assimilation are the PFT community structure and pH (not shown here). It should be noted that the community structure in NEMOVAR is not directly corrected by the total chlorophyll assimilation, to correct the PFT biomass ratios one needs to assimilate directly the PFT chlorophyll (Skakala et al, 2018).

4.4 Assimilation results in the MED MFC domain

This section describes the results for the Mediterranean MFC in Task 3.3e, which was delivered by OGS.

Assimilation setup in MED region. The novel SEIK-OGSTM system described in Section 3.4 was applied for a 9-week winter simulation with weekly assimilation of satellite chlorophyll starting from 1st February 2019. Since this is the first implementation of an ensemble framework of the MED biogeochemical system, it required a large amount of work in the code developing phase and in the tuning of several settings. Results are quite satisfactory even if preliminary. In particular, we tested different ensemble generation approaches and different forgetting factors as well as different combination of additive and multiplicative observation errors. Moreover, the calibration of the threshold on low concentrations allowed to further improve the assimilation results.

The results presented hereafter have been obtained using a 24 ensemble members simulation, where the initial conditions at 1st February were obtained from a one-month ensemble hindcast starting from 24 different IC sets extracted from a multi-annual simulation (i.e., the Mediterranean biogeochemistry reanalysis product provided in the Copernicus Marine Service). Furthermore, different sets of BFM parameters were adopted for the ensemble members. Seven parameters (out of the nearly 200 BFM parameters) were chosen as candidates for perturbation, based on the sensitivity results of the SEAMLESS WP3.2 (see Deliverable D3.2), excluding the ones that define the maximum chlorophyll to carbon ratio, since in the tuning we observed inconsistent trajectories in terms of primary production for some ensemble members when perturbing chlorophyll to carbon ratio. To generate the 24 sets of parameters (one for each ensemble member), the seven chosen parameters were randomly sampled in a range of 20% centred at their reference values.

The observation error was set as composite from an additive and a multiplicative part, where the last one is equal to 25% of the assimilated satellite observation while the additive error is provided as monthly varying standard deviation of the multi-annual satellite reprocessed product in the Copernicus Marine Service. Finally, inflation, as a forgetting factor of 0.8, has been adopted in the assimilation and a threshold of 1.e-5 was applied as lower limit for concentrations. Using the described settings, the two-month simulation with weekly assimilation takes nearly 22 hours running parallel on 2784 cores distributed on 58 nodes.

Diagnostics on observed and non-observed variables

The results of the SEIK-OGSTM assimilation were investigated considering the whole Mediterranean Sea and two areas that are representative of the documented differences among the western and eastern Mediterranean Sea (e.g., Teruzzi et al., 2021): the north-western Mediterranean (nwm; Fig. 4.4.1) and the eastern Levantine (lev1).

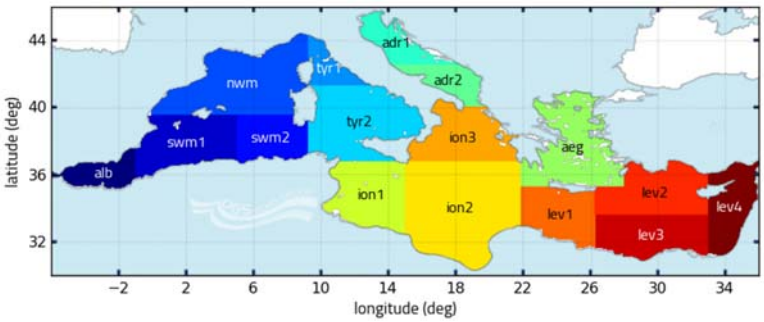


Fig. 4.4.1. Mediterranean Sea sub-basins.

The prescribed settings on inflation and on ensemble generation, based both on ICs and on parameters, provided that the ensemble spread is not degraded along the simulation. Figure 4.4.2 shows timeseries and mean profiles of the ensemble members and ensemble mean for phosphate (i.e., the limiting nutrient in the Mediterranean Sea) in nwm and lev1. For phosphate as well for other variables (not shown) the ensemble spread is spatially and temporal dependent. In the phosphate case, the spread is quite constant in time in the deeper layers, where the effects of the vertical mixing processes are relatively small and consequently the ICs spread is conserved along the simulation.

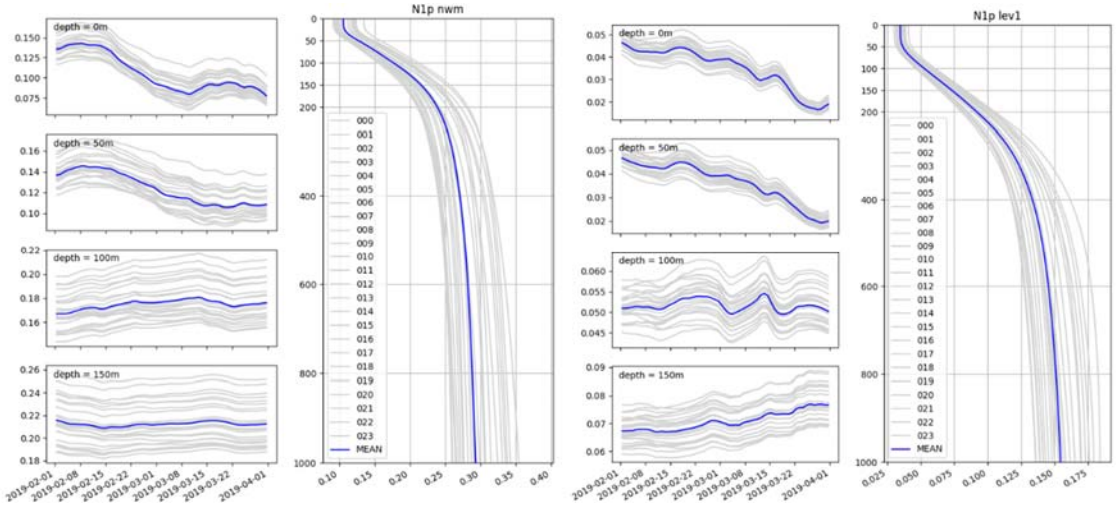


Fig 4.4.2. Phosphate [mmol/m^3] ensemble members (grey) and ensemble mean (blue) timeseries and mean profiles in nwm (left panels) and in lev1 (right panels). Timeseries (left for each sub-basin) are provided at 4 depths (from top to bottom: surface, 50 m, 100 m, and 150 m).

During the two-month simulation period, the assimilation showed non-uniform space and time impacts on the surface chlorophyll concentration (i.e., the observed variable) (Fig. 4.4.3). On average, in the nwm the assimilation impacted the chlorophyll concentrations with corrections nearly equal to 20% with exclusion of the last two assimilation steps when the assimilation increments were very small. On the other hand, in the eastern Mediterranean (lev1) on average the assimilation corrected the surface chlorophyll concentration by 10% at the beginning of the simulation reaching nearly 20% in the last simulation weeks. The assimilation increments propagate along depth (Fig. 4.4.3) nearly

uniformly in the surface layer consistently with the winter mixed conditions occurring in the simulation period. The negligible assimilation corrections in the last simulation period in nwm could be related to the relatively low impact of the set of BFM parameters used to build the ensemble during the post-bloom declining phase (i.e., the 7 parameters are more effective in generating ensemble variability during the bloom phase). Results of a longer simulation period could contribute to further investigate this aspect.

The ratio between posterior and prior standard deviation of BFM chlorophyll values shows that uncertainty reduction is higher at the beginning of the simulation and that the reduction in the posterior ensemble is not uniformly distributed in space and among the phytoplankton functional types (Fig. 4.4.4). On February 4 uncertainty reduction is similar in large phytoplankton (diatoms) and in one of the small phytoplankton groups (flagellates) with lowest ratios in the north-western and Sicily Channel areas (ratio lower than 0.8), while for the other small phytoplankton group (picophytoplankton) the posterior standard deviation mostly reduced in the eastern Mediterranean (ratio nearly equal to 0.83). At the end of the simulation, the standard deviation ratios are lower and more uniform among the two small phytoplankton types.

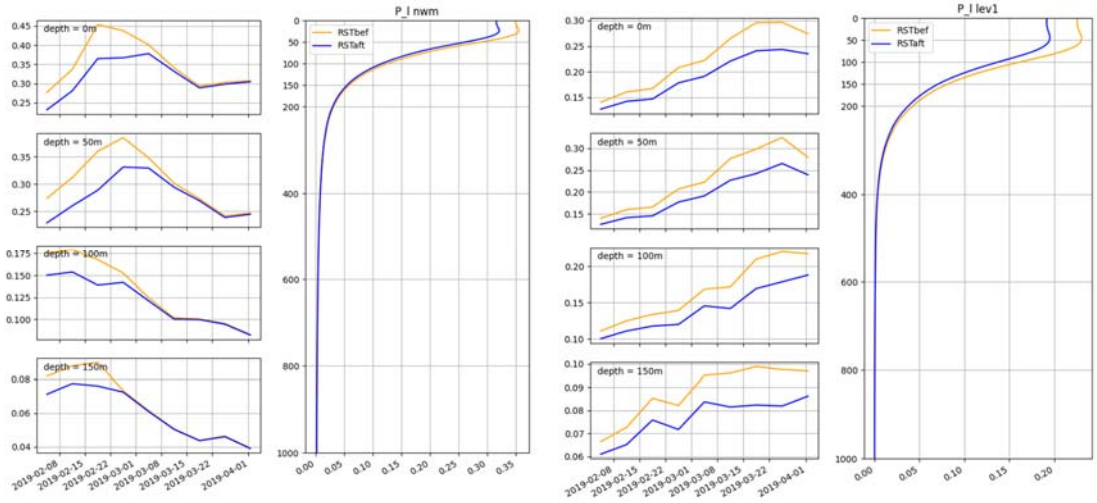


Fig 4.4.3. Prior (orange) and posterior (blue) chlorophyll [mg chl/m^3] timeseries and mean profiles in nwm (left panels) and in lev1 (right panels). Timeseries (left for each sub-basin) are provided at 4 depths (from top to bottom: surface, 50 m, 100 m, and 150 m).

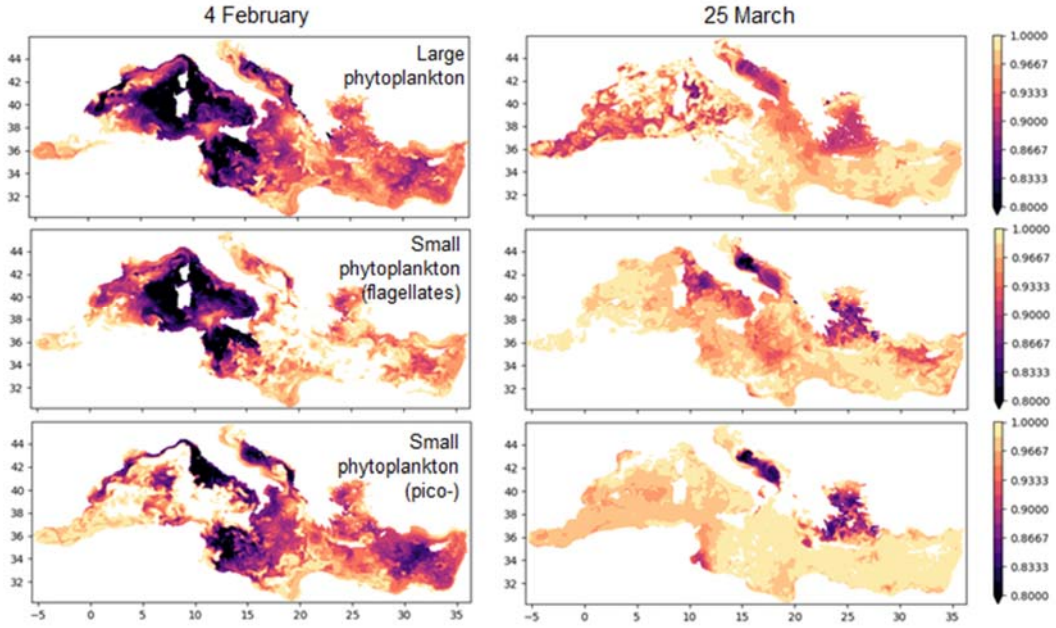


Fig 4.4.4. Ratio between posterior and prior standard deviation of surface chlorophyll in the large (top) and small (middle and bottom) BFM phytoplankton functional types on 4 February (left column) and 23 March (right column).

Rank histograms for February 4 (Fig. 4.4.5) shows the bias among the model and the satellite chlorophyll observations, especially in the eastern Mediterranean (lev1). The model overestimates the surface chlorophyll concentration and the assimilation only partially reduces the bias. Further investigation is needed to improve the SEIK-OGSTM capability to impact the model-observations bias more effectively. To highlight the assimilation effect on modifying the histogram shape, unbiased rank histograms have been calculated by removing the sub-basin mean bias from the observations. Results show that the assimilation flatten only slightly the posterior histograms, increasing by a small amount the central histogram bars while non-univocally acts in reducing the most external ones. Similar rank histogram shapes and assimilation effects are observed for the other assimilation dates and for the other sub-basins.

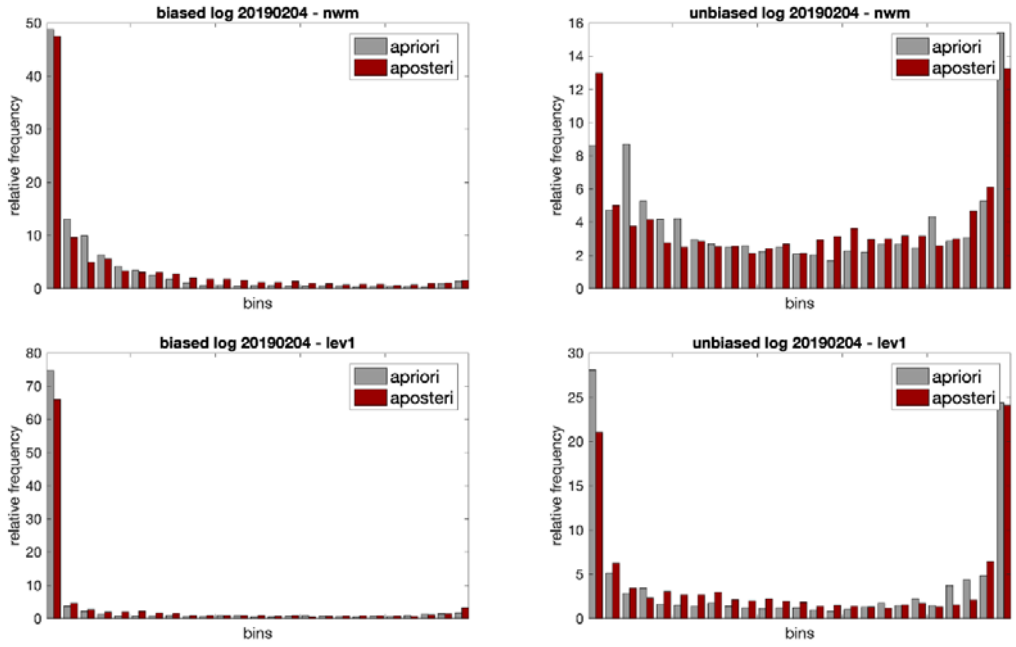


Fig 4.4.5. Rank histograms of the prior (grey bars) and posterior (red bars) ensemble for nwm (top) and lev1 (bottom) standard deviation of surface chlorophyll in the large (top) and small (middle and bottom) BFM phytoplankton functional types on 4 February (left column) and 23 March (right column). Right column histograms are calculated on unbiased observations.

Diagnostics on derived quantities and indicators

The results of the two-month winter SEIK-OGSTM simulation provide information on the effects of the assimilation of surface satellite concentration on derived quantities directly related to the SEAMLESS indicators. In nwm the assimilation affects the phytoplankton phenology moving the phytoplankton bloom peak one week later (Fig. 4.4.3 and 4.4.6), while in lev1 the assimilation reduces the intensity of the bloom without modifying its timing.

Simulation results on phytoplankton and zooplankton biomass indicate how the assimilation can affect the trophic efficiency (related to the phytoplankton-zooplankton biomass ratio). Effects on zooplankton biomass are negligible in the eastern Mediterranean (lev1) while they emerge more clearly in the western Mediterranean (nwm) in the lowest part of the euphotic layer (Fig. 4.4.6). Since the assimilation modifies the phytoplankton biomass and the zooplankton biomass to a less extent, the assimilation shows to have an impact on the trophic efficiency indicator. However, a longer simulation could allow to better estimate effects on trophic efficiency.

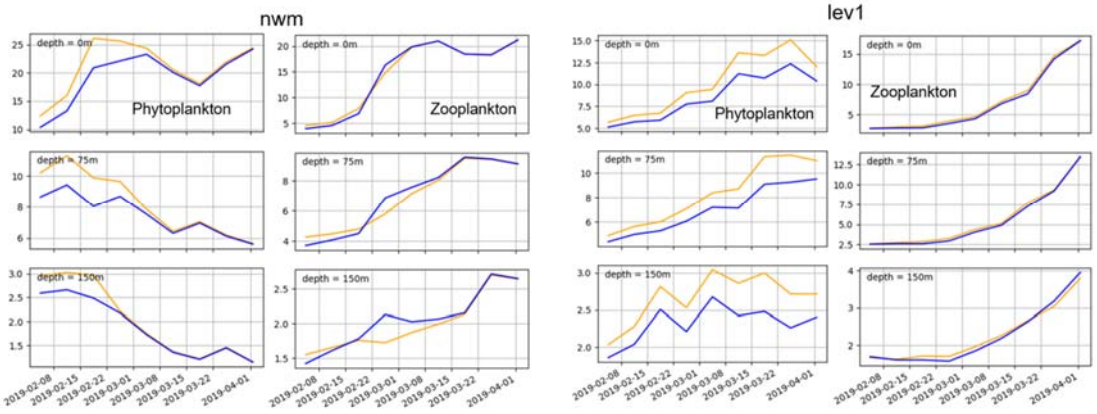


Fig 4.4.6. Prior (orange) and posterior (blue) phytoplankton and zooplankton biomass [mg C/m^3] timeseries in nwm (left panels) and in lev1 (right panels). Timeseries are provided at 3 depths (from top to bottom: surface, 75 m, and 150 m).

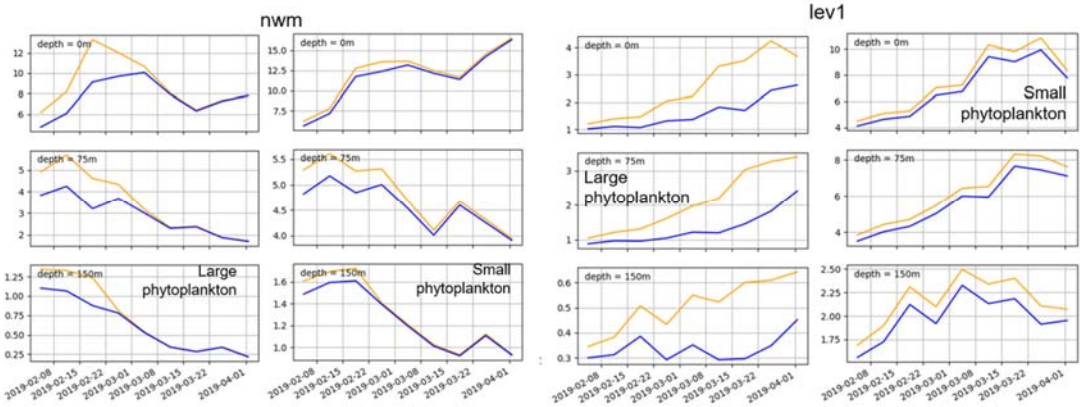


Fig 4.4.7. Prior (orange) and posterior (blue) large and small phytoplankton biomass [mg C/m^3] timeseries in nwm (left panels) and in lev1 (right panels). Timeseries are provided at 3 depths (from top to bottom: surface, 75 m, and 150 m).

The BFM phytoplankton functional types are differently affected by the assimilation in the simulation period. Large phytoplankton (diatoms) is more largely affected by the satellite chlorophyll assimilation with respect to the small phytoplankton functional types (flagellates and picophytoplankton) (Fig. 4.4.7). Assimilation increments on large phytoplankton are particularly relevant in the eastern Mediterranean (also in the subsurface layer). In other periods of the seasonal phytoplankton dynamics, effects on the phytoplankton functional types may differ from what emerges in the present winter simulation.

Differences between prior and posterior POC sink at 500 m (Fig. 4.4.8) indicate that in the simulation period the assimilation effects on POC sink are limited to the western Mediterranean (e.g., nwm) and

in the third simulation week (close to the end of the period when effects on phytoplankton chlorophyll and biomass are more intense, Fig. 4.4.3 and 4.4.7).

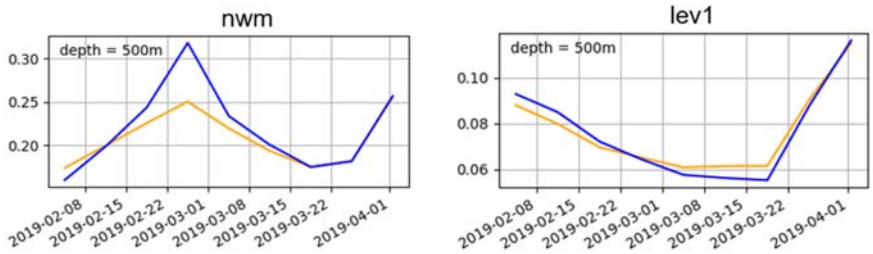


Fig 4.4.8. Prior (orange) and posterior (blue) POC sink at 500 m depth [mg C/ m^2/d] timeseries in nwm (left panel) and in lev1 (right panel).

4.5 Assimilation results in the ARC MFC domain

4.5.1 Assimilation setup in ARC region

The data assimilation system for the joint parameter-state estimate is based on the TOPAZ ensemble Kalman filter system (EnKF) used in a one-lag smoother (EnKS) setting (see Section 3.5). Here, only biogeochemical data (surface chlorophyll-a from satellite ocean colour sensor, in-situ Nitrate, Silicate and Phosphate) are assimilated from March to October in each year. Assimilation cycles are defined by the frequency of the composite satellite ocean-colour data OC-CCI v4.2 8-daily 4 km product. All in-situ nutrient data are binned to the nearest analysis date. Period of one assimilation cycle is at every 8 days when OC-CCI product is available. In this section, we call the model state right after the data assimilation analysis “analysis at day 0”, and the model state right before the analysis “analysis at day 8”.

The HYCOM-EVP-ECOSMO II ocean-biogeochemical model with NA2a0.80 configuration is initialized with a spun-up model state in January 2000, obtained from the SPONGES project. The model is forced by 6 hourly atmospheric forcing derived from ERA-5 reanalysis products. The ensemble spin-up run start son January 1st, 2006. The initial layer thickness is perturbed with 10% perturbation variance and correlation scales of 3 in layers and 10 in horizontal grid to create 80 ensemble members. Then the 80 ensemble members are further spun-up to 1 January 2007 with perturbed atmospheric forcing and perturbed ECOSMO II model parameters. The perturbed atmospheric forcing is generated internally in the HYCOM system with perturbation variance: 3 (hPa) for sea level pressure and 0.003 (N/m^2) for wind stress components and 20% for downward shortwave radiation in ECOSMO II with horizontal decorrelation scale 500 km and temporal decorrelation scale 5 day. The perturbation standard deviation of the model parameters is chosen to be 20% for all parameters.

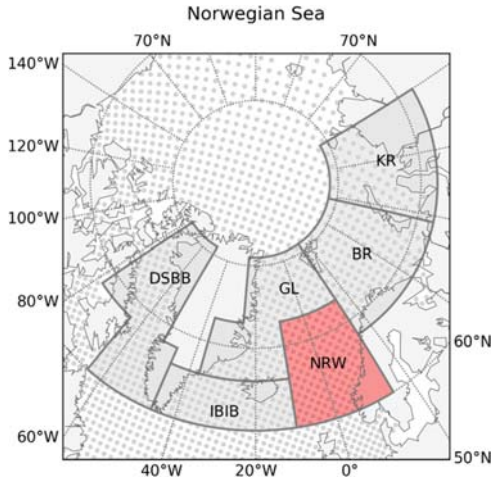


Figure 4.5.1. The Norwegian Sea (NRW) domain for the ensemble data assimilation evaluation.

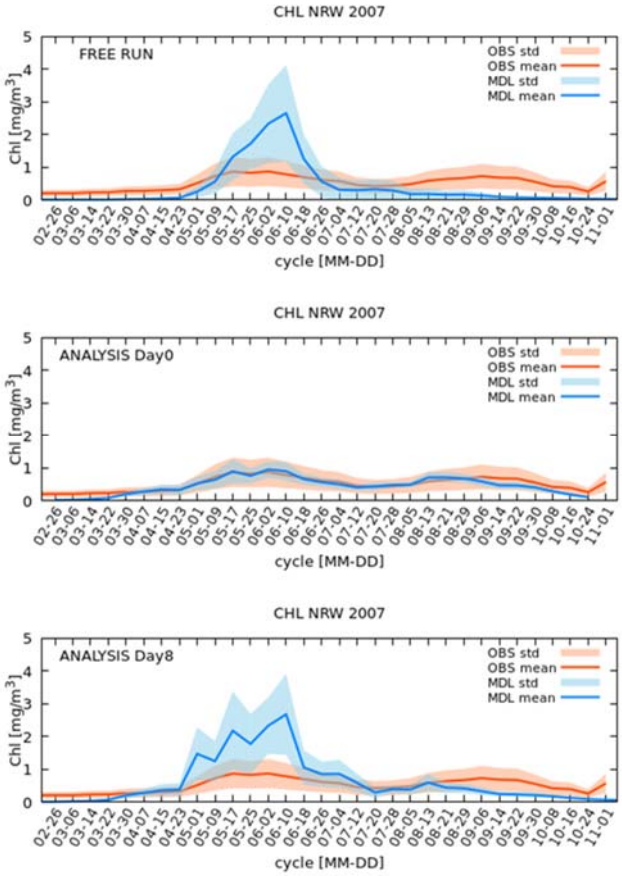


Figure 4.5.2. Time series of surface Chl-a from observation and reanalysis products at NRW. Top: Free run, middle: Reanalysis at day 0, bottom: Reanalysis at day 8. Solid lines are area mean and shades are area domain standard deviation. Blue colour is for model products (free run, reanalysis day 0 and day 8). Orange colour is for observation data.

4.5.2 Diagnostics on observed variables

In order to evaluate the quality of ensembles, we have performed rank histogram analysis (Hamill 2001) over the NRW domain (figure 4.5.1). The ranking of the satellite Chl-a data is counted at two data assimilations cycles. The first cycle is 20070501/20070509 and the second cycle is 20070821/20080929. The first cycle is attributed for the early stage of Spring bloom (see figure 4.5.2 and figure 4.5.3) and the second cycle is attributed to the early stage of Fall bloom (see figure 4.5.2 and figure 4.5.5) observed over NRW in 2007. The number of sample points for deriving a rank histogram is 552. At each cycle, three rank histogram plots are calculated. The first rank histogram is derived for ensemble generated by 8 days ensemble forecast from the previous cycle data assimilation analysis. The second rank histogram is for ensemble analysis generated by ensemble data assimilation and the third rank histogram is derived for another 8 days ensemble forecast from the ensemble analysis.

During the early stage of Spring bloom (20070501/20070509), rank histogram based on the ensemble forecast at 20070501 has a typical U-shape (figure 4.5.4, left panel) which is a sign of cases of either some low and some high biases in the ensemble, or when the ensemble doesn't spread out enough. After data assimilation, histogram becomes relatively flat (figure 4.5.4, middle panel) over the most range of the rank, but about 20 percent of observations fall to "too large compared to analysis ensemble" category. This character of analysis ensemble is identified for the first time from this rank histogram analysis. Rank histogram based on the ensemble forecast at 20070509 reveals development of model bias during 8 days forecast period. The "left stacked" rank histogram (figure 4.5.4, right panel) indicates many observations are commonly lower than ensemble members. This model bias development is known to our data assimilation system at the level of model free run (see figure 4.5.2, top panel). Our coupled model system tends to produce high surface Chl-a concentration during the Spring bloom and initial Chl-a corrected by data assimilation cannot suppress the development of model bias (see figure 4.5.2, bottom panel). We also found significant fraction of observations still lies between rank 25 and 80. This finding may help us understand what conditions cause the bias development by taking at close look at each ensemble member.

Rank histogram based on ensemble forecast at 20070821 (see figure 4.5.6, left panel) exhibits strong "right stacked" shape and indicates majority of observed Chl-a are commonly higher than the forecast ensemble members. The bias is removed from data assimilation (see figure 4.5.6, middle panel), but comes back quickly within 8 days forecast period (see figure 4.5.6, right panel). Again, correcting phytoplankton biomass through assimilation Chl-a at the surface does not have strong enough impact to sustain Fall bloom as is the case in the model free run experiment (see figure 4.5.2, top panel). Recent study indicates this Fall bloom over the Norwegian Sea can be accurately predictable from the atmospheric forcing, which are overlapped with the external atmospheric forcing that drives our coupled ocean model. This suggests the existence of intrinsic model bias in the current settings of ECOSMO biogeochemical module. Rank histogram based on ensemble forecast at 20070829 (see figure 4.5.6, right panel) indicates more than 30 percent of observations are outside of the forecast ensemble members. Close look at the development of the surface Chl-a and its associated biomass

and nutrients at these observation points may shed some light on the dynamical reason of the model bias in future work.

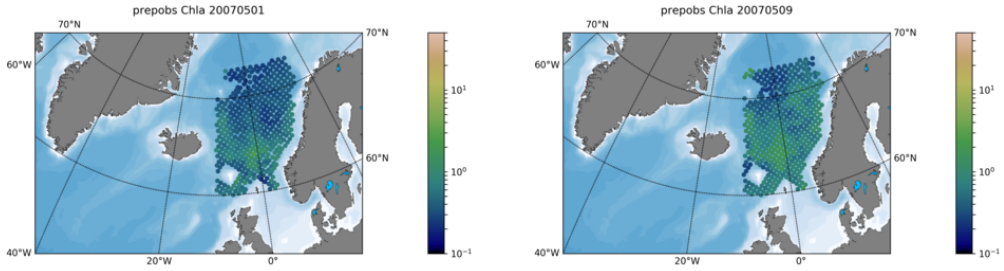


Figure 4.5.3. Satellite Chl-a concentration [log(mg/m3)] over the NRW domain at cycle 20070501 (left) and cycle 20070509 (right).

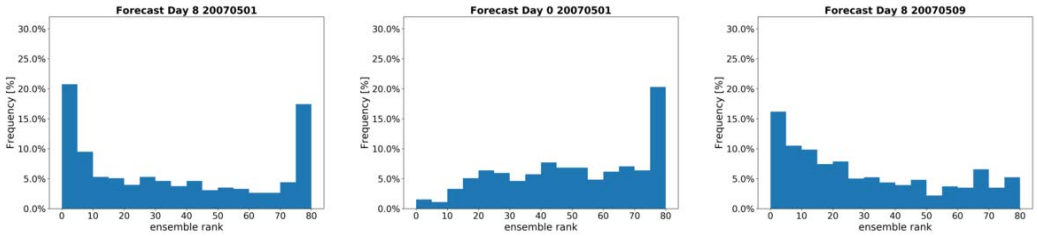


Figure 4.5.4. Rank histogram of the surface Chl-a at cycle 20070501 (left and middle) and at cycle 20070509. Note middle panel is the ensemble states right after data assimilation analysis.

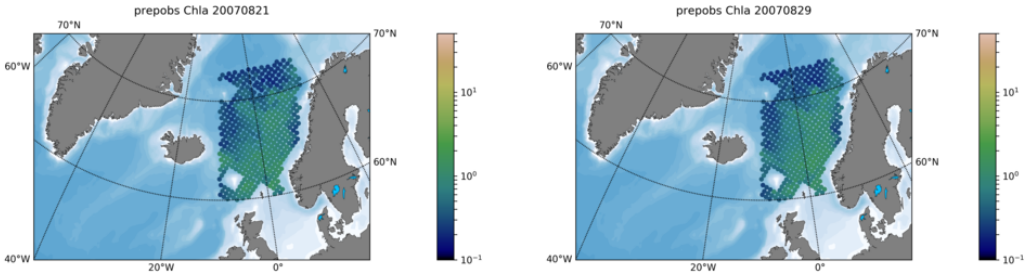


Figure 4.5.5. Satellite Chl-a over the NRW domain at cycle 20070821 (left) and cycle 20070829 (right).

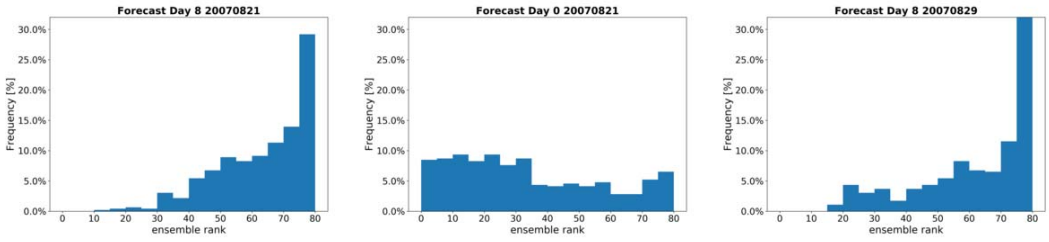


Figure 4.5.6. Rank histogram of the surface Chl-a at cycle 20070821 (left and middle) and at cycle 20070829. Note middle panel is the ensemble states right after data assimilation analysis.

4.5.3 Diagnostics on "non-observed variables"

In the current settings of ECOMSO biogeochemical model, we have two phytoplankton groups: large phytoplankton (diatom), small phytoplankton (flagellate) and two zooplankton groups: large (meso) zooplankton and small (micro) zooplankton as plankton functional type. In the ensemble free run experiment, two phytoplankton groups show the same timing and length in spring bloom that starts at cycle 20070423 and lasts about 8 analysis cycles (64 days) (Figure 4.5.7, top panel on the left column). Associated zooplankton development in both micro and meso zooplanktons starts from cycle 20070517 and lasts for about 15 cycles (120 days) (Figure 4.5.7, top panel on the right column). On the other hand, data assimilated products at day 0 and at day 8 (Figure 4.5.7, middle and bottom panels on the left column) show different timings in onset of spring bloom for diatom and flagellate. Diatom bloom starts at cycle 20070423 and lasts about 7 cycles (56 days) while flagellate bloom starts at cycle 20070525 and lasts about 8 cycles (64 days). This indicates that we have clear difference in peak and duration of spring bloom among two phytoplankton groups after data assimilation. This lags in phytoplankton phenology among two groups of phytoplankton is typical to a high latitude North Atlantic spring bloom system (Dale et al., 1999) and data assimilation has positive impact to correct the bias in the free run experiment.

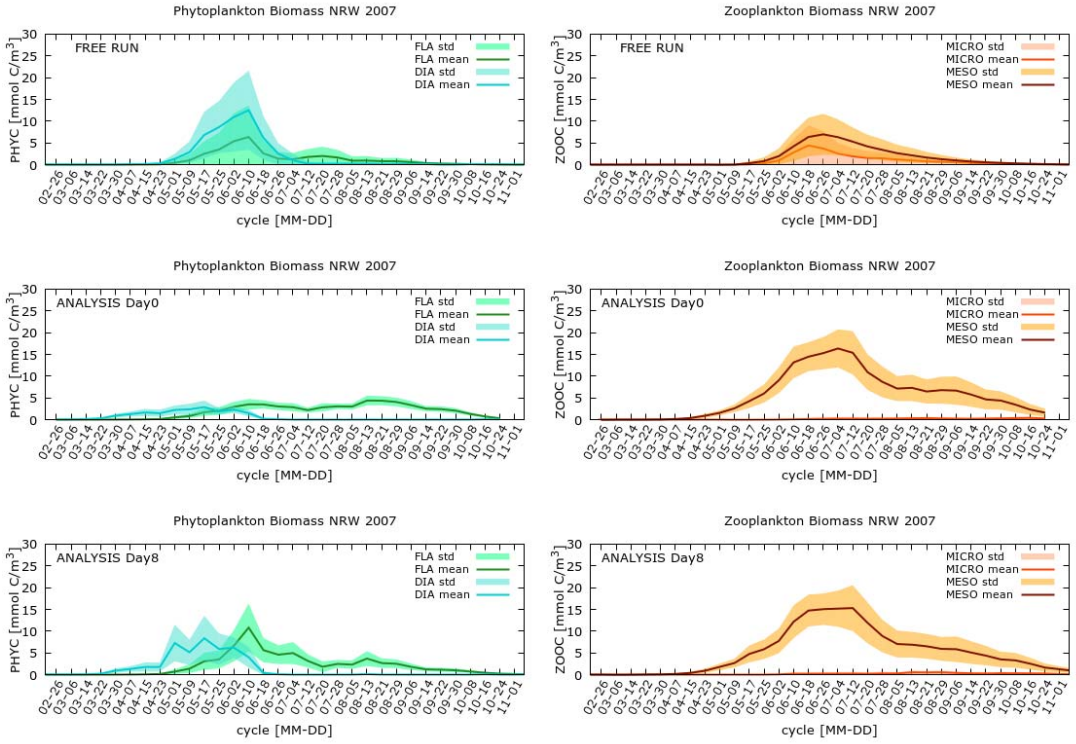


Figure *4.5.7. Time series of surface plankton biomass (left column: phytoplankton, right column: zooplankton) from reanalysis products at NRW. Top: Free run, middle: Reanalysis at day 0, bottom: Reanalysis at day 8. Solid lines are area and depth (upper 200m) mean of ensemble model states and shades are area and depth (upper 200m) averaged ensemble standard deviation. DIA: Diatom (large phytoplankton), FLA: Flagellate (small phytoplankton), MICRO: Micro zooplankton (small zooplankton), MESO: Meso zooplankton (large zooplankton). Solid line is area averaged ensemble mean, shade is area averaged ensemble standard deviation.*

The composition of zooplankton also is impacted by data assimilation. In the free run experiment, both meso and micro zooplankton have similar timing of peak and duration (Figure 4.5.7, top panel in the left column). After data assimilation, zooplankton biomass is dominated by meso zooplankton and micro zooplankton disappears through entire assimilation cycles. Due to lack of seasonal cycle data of zooplankton composition in the Norwegian Sea, it is difficult evaluate if data assimilation has positive or negative impact on the zooplankton composition.

4.5.4 Diagnostics on derived quantity and indicators

Here we analyse the impact of ensemble data assimilation to three SEAMLESS indicators of particular interest in the ARC MFC: phytoplankton phenology, plankton functional type and trophic efficiency. Diagnostics on plankton functional type is reported in 4.5.3.

a. Phytoplankton phenology

Phytoplankton phenology in NRW domain (Figure 4.5.1) based on area averaged satellite Chl-a data (Figure 4.5.2, red line and orange shade) is characterised by two bloom peaks in 2007. The spring (first) bloom starts from cycle 20070501, reaches its peak at cycle 20070517 and lasts 8 cycles (64 days) and the fall (second) bloom starts from cycle 20070813, reaches at cycle 20070906 and lasts 7 cycles (56 days). However, free run has only one peak during the spring bloom which starts from cycle 20070509, reaches its peak at cycle 20070610 and lasts 5 cycles (40 days). By assimilating the surface Chl-a, analysis right after the data assimilation (day 0) recovers the observed spring bloom phenology, but only early stage of the fall bloom is represented in the analysis (Figure 4.5.3, middle panel). We can find degradation of the retrieved phytoplankton phenology at the end of 8 days forecast after data assimilation (analysis at day 8). Onset of the spring bloom (Figure 4.5.2, bottom panel) is at cycle 20070501 as is in observation, but its peak is now shifted to cycle 20070610 which is the same timing as the one in the free run. Onset of the second bloom can be seen at cycle 20070813, but the bloom quickly decays by the next cycle.

b. Plankton Functional Type (PFT)

See 4.5.3.

c. Trophic efficiency (TE)

Trophic efficiency (TE) is defined for this assessment as :

TE = (total zooplankton biomass) / (total phytoplankton biomass + detritus + DOM)

where, total zooplankton biomass is calculated from micro-zooplankton and meso-zooplankton biomass and total phytoplankton is calculated from ensemble average of small phytoplankton (flagellate) and large phytoplankton (diatom) biomass. Note detritus and dissolved organic matter (DOM) are part of denominator in TE definition since zooplankton grazes on detritus and DOM in the ECOSMO food web settings. As is observed in the time series of zooplankton biomass (Figure 4.5.7,

right column), total zooplankton biomass development in data assimilated product (analysis) starts at cycle 20070423 which is about 4 cycles (32 days) earlier compared to free run. This earlier onset of zooplankton biomass development has the strongest impact to seasonal cycle of TE in analysis. Peak time of TE shifted from cycle 20070704 in free run to 20070712 in analysis.

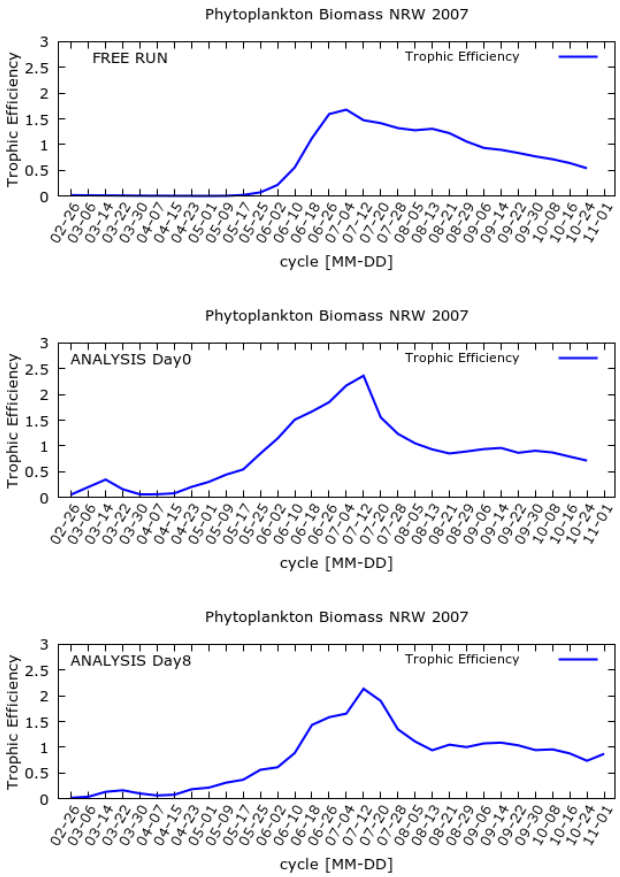


Figure 4.5.8. Time series of trophic efficiency (TE) from reanalysis products at NRW. Top: Free run, middle: Reanalysis at day 0, bottom: Reanalysis at day 8. Solid lines are area and depth (upper 200m) mean of ensemble model states.

5. Observability of the ecological indicators

In this section, a cross-cutting assessment of observability properties of selected SEAMLESS indicators is made, relying on the 3D assimilation experiments results and related diagnostics obtained in the different regions. The underlying observability/controllability question can be rephrased as follows: “Which ecosystem indicators can be faithfully estimated using the assimilation systems implemented in the 3D setups?”. A subsidiary question is: “Does the complexity of the considered BGC models (in terms of diversity of modelled PFTs) influence the controllability in 3D?”.

This assessment is an extension of the comprehensive observability analysis undertaken in 1D (see Deliverable 3.2) for the 5 targeted indicators discussed here below. However, in terms of numerics a straightforward implementation of the sensitivity analysis approach is unfeasible with 3D systems because of the computational burden implied by ensemble simulations (~ 1000 in 1D). As a result, it is only practical to assess the observability of selected indicators against selected assimilated variables of interest in the considered CMEMS regions. Table 5.1 provides a summary of the assimilation experiments featured in section 4. Due to the use of ensemble methods by all partners, ensemble diagnostics provide useful metrics to assess the assimilation impact and their interpretation in terms of observability properties for the selected indicators, in particular phenology and POC which are two indicators proposed by SEAMLESS as new products in the CMEMS portfolio.

	GLO/IBI	NWS	ARC	MED	BAL
Uncertainty sources	BGC parameters sub-grid scales location of features	atmo. forcings BGC parameters	atmo. forcings init. conditions BGC parameters	init. conditions BGC parameters	Init. conditions BGC parameters
Assimilated data	OC	OC SST T/S profiles	OC in situ nutrients	OC	OC
Assimilation time	12 months	4 months	8 months	2,5 months	3 months
Ensemble size	40	30	80	24	30
Diagnostics	rank histograms ens. spread, CRPS	rank histograms ens. spread	rank histograms ens. Spread	rank histograms ens. spread	Ens. Spread
Informed indicators	PHE, POC, TE, PP	PHE, (TE, PFT), OX	PHE, TE, PFT	PHE, TE, PFT, POC	PHE, TE, PFT, PP, OX

Table 5.1. Key features of the assimilation experiments preformed to assess controllability of ecosystem indicators in the 5 CMEMS MFC regions. Informed indicators are phenology (PHE), POC, Trophic Efficiency (TE), primary production (PP), phytoplankton functional types (PFT), oxygen (OX). See details in section 4.

5.1 Surface Chlorophyll, phenology (PHE)

Phytoplankton phenology can be defined as the timing, duration and amplitude of major blooms. The mapping of this indicator provides information on the degree of temporal matching between predator and prey in marine ecosystems, which is known to have a key influence on the food web, fish spawning and fisheries recruitment. Hence, it is a key indicator for monitoring the state of the pelagic ecosystem and for detecting changes triggered by perturbation of the environmental conditions.

A variety of metrics can be used to characterize phenology (depending on regions and trophic regimes), such as the timing (initiation, peak, amplitude, termination and duration) and the value of maximum chlorophyll concentration in the surface layer (see Deliverable D3.2). In practice, time series of surface phytoplankton contain all the required information to derive these metrics, whatever they may be. The main challenges for the MFC systems are therefore: (i) the reconstruction of continuous time series of surface chlorophyll concentrations by filling gaps (due to clouds and uneven satellite coverage) of observed ocean colour L3 products, and (ii) the reduction of uncertainty with respect to the equivalent observation-based L4 products (despite the fact that current observation-based CMEMS products still describe their uncertainties in a rather simple way).

Satellite OC data being assimilated in all systems, the skill assessment of the 3D assimilative systems in terms of phenology is quite straightforward. The impact of OC is generally very significant on the chlorophyll reconstruction in the upper ocean in all experiments, corroborating the conclusions of D3.2. In the open ocean but also the coastal station Arkona, the assimilation modifies both the timing and amplitude of the spring bloom (see Figure 4.1.6 for BAL; Figure 4.2.5 for GLO/IBI; Figure 4.4.3 and 4.4.6 for MED; Figure 4.5.2 for ARC) while reducing the ensemble spread. In the ARC, the assimilation is also able to restore the secondary (fall) bloom, while it was not present in the free run. On the northwest shelf, Figure 4.3.5 and related error statistics demonstrate that the new hybrid system captures much more sensibly the phytoplankton phenology (i.e., the bloom around the end of March) than the 3DVAR system by improving the chlorophyll distribution not only at the surface but also at mid-depth. It should be noted that the reconstruction of a deep chlorophyll maximum in some regions of the open ocean (e.g. at BATS station, see D3.2) is likely much more difficult to achieve using satellite OC only, though this would require further investigation.

5.2 Primary Production (PP)

Marine Primary Production is a direct by-product of the BGC models as they all rely on the mechanistic modelling of photosynthesis processes. In the CMEMS context, the challenge is to estimate PP as accurately as possible to refine the quantification of the oceanic contribution to the carbon cycle and the transfer of organic matter and energy to marine ecosystems. PP estimates can be derived from satellite OC using empirical assumptions and related schemes to project surface information at depth. In SEAMLESS, the purpose is to demonstrate that 3D assimilative models can bring added value by using more relevant vertical projection schemes (in spite of persistent modelling uncertainties in biogeochemical processes, as discussed by Tagliabue *et al*., 2021). Hence, the controllability of PP will depend on the assimilation capacity to dynamically extrapolate the OC surface information (with possibly additional observation constraints from *in situ* profiles) in the euphotic depth.

The PP indicator has been estimated in the North Atlantic experiment during 2019 (see Figure 4.2.7 showing the total PP vertically integrated over the water column in the PAP region). The impact of surface OC assimilation is reflected by the occurrence of several PP peaks during the year (in February, May, August and October, successively) showing a reduced ensemble spread compared to the prior ensemble run. Moreover, the comparison with the corresponding time series in Figure 4.2.2 indicates that the succession of PP peaks does not univocally match with those of Chl concentration, suggesting that nonlinear effects are at play. These diagnostics demonstrate *stricto sensu* that (i) PP can be estimated from the 4D inverse system output, and (ii) PP is sensitive to updated surface Chl values. However, it does not demonstrate the level of reliability of these posterior ensemble estimates. This would require further verification with independent data which remains quite challenging as PP cannot be observed with the suitable coverage and resolution for modelling purposes.

5.3 Phytoplankton Functional Types (PFT)

The PFT indicator is defined here as the ratio between large phytoplankton biomass and total phytoplankton biomass. Its purpose is to provide information on the relative importance of the large vs. small phytoplankton pathways in the functioning of the herbivorous food chain (see D3.2). The BGC model complexity level, characterized in SEAMLESS by the diversity of modelled PTFs, is expected

to have some implications on which ecosystem indicators are controllable. In the CMEMS context, the challenge is the choice and implementation of appropriate models that can make the best compromise between the complexity required to generate the targeted indicators, and the availability of independent data sets for calibrating the simulated ensembles.

The assimilation experiments documented in section 4 allow a first exploration of PFT indicators, comparing the "high-complexity" BFM model (with 4 size-based functional types) in the Mediterranean domain and the "intermediate complexity" ECOSMO model (with 2 phytoplankton groups) in the Arctic region. In the MED region, the winter experiment suggests that large phytoplankton (diatoms) is more largely affected by the satellite chlorophyll assimilation than small phytoplankton functional types (flagellates and picophytoplankton), especially in the Eastern Mediterranean (Fig. 4.4.7). In the Arctic (Figure 4.5.7), the assimilation results in lags in phytoplankton phenology among two groups of phytoplankton (as typically observed in a high latitude North Atlantic spring bloom system). Both cases demonstrate that the resulting PFT indicator is sensitive to updated surface chlorophyll values, though with different manifestations. These preliminary results suggest that the contrasted assimilation impacts are more directly related to the specific model implementations in the considered regions than on the model complexities themselves. This confirms one of the conclusions of the 1-D sensitivity experiments reported in Deliverable D3.2, i.e. that controllability is not linked clearly to the complexity of the BGC models.

5.4 Particulate Organic Carbon (POC) flux

The POC flux designates here the fraction of non-living particulate organic matter, i.e. the detritus, escaping from the productive upper ocean and contributing to the sequestered carbon in the deep ocean. Detritus are represented as one state variable in BFM, ERGOM and ECOSMO, and two state variables (big and small particles, to be summed) in ERSEM and PISCES. The computation of the flux depends on the detritus concentration and the vertical sedimentation speed prescribed in each model. While data assimilation is expected to affect the stock of detritus, the sedimentation speed is left unchanged in all experiments, and this may be a source of significant unconstrained uncertainty.

Section 4 presents the estimates of the POC flux indicator at 500 meters for the MED region (Figure 4.4.8), and at 100 meters the GLO/IBI region (Figure 4.2.7). The depths have been chosen to be the most significant for the regions under consideration. In the North Atlantic, it appears that the maximum POC flux increases w.r.t. the prior ensemble, replicating the evolution of the chlorophyll concentration (though with higher uncertainty range). In the MED region, the assimilation impact in the Western basin also results in an increased flux, though limited in time duration.

It is difficult to raise clear conclusions about controllability for POC flux estimates since the number of available assimilation experiments is limited, the dependency on vertical sedimentation is uncontrolled, and no independent data are available for faithful verification. It is therefore suggested to pursue investigations about this indicator before transitioning the production in the CMEMS catalogue.

5.5 Trophic Efficiency (TE)

Trophic Efficiency is an ecological metric estimating the energy ratio that consumers in one trophic level gain from the previous trophic level and convert into biomass. The TE indicator is defined in D3.2 as the ratio between the zooplankton biomass and the phytoplankton biomass, computed as the vertical integral of the corresponding state variables between 0 and 200 meters.

In the North Atlantic experiment (Figure 4.2.7), the posterior ensemble reflects substantial modifications of the trend of the zoo/phyto ratio during the entire productive period. The ratio peaks in late summer, following a period of increased phytoplankton biomass as a result of OC assimilation (Figure 4.2.2). This suggests an increased accumulation of energy and biomass in the herbivorous trophic level throughout the year. A similar trend seems to occur during the productive period in the NWS (Figure 4.3.1), the ARC (Figure 4.5.7), and the BAL experiment (Figure 4.1.13). In the Western Mediterranean region, it is interesting to note that the assimilation modifies the phytoplankton biomass and at a less extent the zooplankton biomass, suggesting an apparent impact on the trophic efficiency indicator. However, it is not possible to conclude whether this impact is realistic, or the result of poor controllability of zooplankton.

It should be noted that the indicator, expressed in terms of instantaneous biomass instead of production rates, may be misleading especially in case of low phytoplankton biomass (in the denominator). Nevertheless, the TE definition we adopted is applicable to all systems allowing for generalisation.

A summary of the observability/controllability empirical assessment of indicators is indicatively given in Table 5.2 here below.

Indicator	ARC	BAL	NWES	MED	GLO/IBI
Phenology	High	High	High	Medium	High
PP	n/a	Low	High	Medium	Medium
POC flux	n/a	n/a	High	Low	Medium
PFT	Medium	Medium	Medium	Medium	n/a
Trophic efficiency	Medium	Medium	High	Low	Low

Table 5.2. Assessment of observability/controllability levels of ecosystem indicators in the 5 CMEMS regions.

6. Guidelines and conclusions

Based on the results of Section 4 and the assessment in Section 5, we can now draft a number of preliminary recommendations to assist in the implementation strategy of a future ensemble assimilation framework in CMEMS, and to adopt more rigorous methodologies for enriching the CMEMS portfolio with new indicators. These preliminary recommendations are intended to address questions that frequently arise in the design stages of versions of the new analysis and forecasting systems. They could be revised after project completion.

1) Strategy for dealing with uncertainty sources. The aim of the experiments conducted in WP3 was to provide a baseline assessment of the overall performance of the ensemble methods, including their ability to estimate uncertainties of selected SEAMLESS indicators. This led us to consider a limited number of sources of uncertainty, based on common sense (see Table 5.1). The introduction of uncertain biogeochemical parameters in the models, selected according to the results of Task 3.2, was however unanimously adopted in all simulations since BGC parameters typically have high uncertainty and are rarely calibrated. It seems logical to further this direction in SEAMLESS, and to increase the number of uncertain parameters in future experiments. At CMEMS MFC level, this option is probably the easiest to implement, both algorithmically and in terms of parameterization settings, and can be recommended as a priority, while incrementing in conjunction other sources of uncertainty.

2) Transition to ensemble monitoring and forecasting systems. The methods developed by the partners in Task 3.3 demonstrate that such a transition is feasible in a near future since ensembles of model states used for the computation of error covariance matrices of Kalman filters can also be used in a similar way in variational assimilation algorithms. Therefore, this transition can be envisaged in an incremental way, with a first step consisting in simulating ensembles, and a second step dedicated to the upgrade of assimilation kernels facilitated by the use of the SEAMLESS prototype. The initial priority effort should therefore be to produce ensemble simulations and to verify their statistical consistency. The use of ensembles for probabilistic forecasts could be a parallel track for MFCs, not requiring the redesign of the whole assimilation chains.

3) Optimal complexity level of BGC models. The complexification of BGC models is a strong trend observed for decades in the scientific community, linked to the increased knowledge of BGC processes. In the operational context, the increase in the number of state variables can be motivated by user's needs about some high-level ecological indicators (e.g. PFT), but it implies in parallel an increase in the numerical cost (especially with ensemble systems) and in the amount of verification data to be processed for calibration. The assessment of our assimilation experiments (in Tasks 3.2 and 3.3) suggest that the controllability of PFT is not clearly related to the complexity of the model, so that this criterion does not help to guide the model choice. This question cannot be addressed without taking into account the comparison with? observational data and their uncertainty, and must be placed in the context of a "system" approach in which uncertainties are evaluated for all components (model, assimilation algorithms, data, couplings). From a probabilistic point of view, there is no reason to believe that simplifying the models (which is out of the SEAMLESS scope) would decrease the reliability of, at least, some estimated ecological indicators investigated here. Certainly, the "realism"

and objectives of the model in relation to the different MFCs need to be taken into account to select the relevant complexity level, but this is out of the scope of SEAMLESS.

4) Choice of numerical settings and assimilation methods. The variety of assimilation methods implemented in WP3 is quite diverse but converge all to the ensemble paradigm. It was not the purpose of Task 3.3 to inter-compare different assimilation algorithms in a common 3D setting, or to test the impact of numerical choices such as vertical or horizontal resolution on observability/controllability. It is therefore out of scope to provide guidance on related choices. However, ensemble approaches imply to reconsider some modelling choices (e.g. resolution) made previously with respect to assimilation cost constraints and ensemble reliability, especially for coupled physical-BGC systems. Consistently with the 2021 document on CMEMS R&D priorities (v6), a general recommendation could be to stimulate more inclusive design studies at MFC level to identify the relevant modelling and assimilation choices according to the observational information available to constrain the quantities to be estimated. The need for integrated design studies, however, is to be distinguished from the benefits derived from more modular software functionality (separation of modelling and assimilation modules), as experienced in SEAMLESS with PDAF which allows for greater flexibility and agility to adapt and test methods.

5) Assessment of Ocean Colour data products. In most of the experiments conducted in WP3, the assimilated data came from L3 chlorophyll composite products estimated from satellite reflectance measurements and dedicated processing algorithms. The data coverage is therefore reduced due to the presence of clouds, depending on the season and the region considered. The ensemble assimilation methods implemented in SEAMLESS return an uncertainty analysis over the whole domain, overcoming the problem of cloud coverage of L3 products. Thus, the SEAMLESS approach propagates information (from observed to cloud-covered areas) based on the dynamic evolution of uncertainty. In some cases, L4 products have been used for verification purposes, but the description of L4 product uncertainties over cloudy regions remains imperfect. A more systematic analysis of the ability of MFC assimilation systems to fill cloud gaps, based on a rigorous error budget analysis of MFC assimilated products compared to those processed by TACs, would be useful to conduct as part of a future dedicated data challenge.

6) Diagnostics and observational data sets. Rank histograms are the simplest and most intuitive probabilistic diagnostics informing on (i) the consistency of ensembles in general, and (ii) the impact of data assimilation more specifically. These diagnostics are standard in the NWP community, but innovative for marine BGC applications. Such diagnostics have been produced by all partners (see Figures 4.1.5, 4.2.4; 4.3.4; 4.4.5; 4.5.4, 4.5.6) to look at the behaviour of their assimilation system and identify possible flaws. Some character of the ARC analysis ensemble was identified for the first time from this rank histogram analysis. For instance, rank histograms can help detect the development of model bias during the forecast period. Such finding may help us understand what conditions cause the bias development by taking at close look at each ensemble member. The use of such diagnostics needs to be encouraged to exploit split between verification and assimilation.

7. References

Bessières L., Leroux S., Brankart J.-M., Molines J.-M., Moine M.-P., Bouttier P.-A., Penduff T., Terray L., Barnier B. and Sérazin G., 2017: Development of a probabilistic ocean modelling system based on NEMO 3.5: application at eddying resolution. Geoscientific Model Development, 10(3), 1091–1106.

Brankart J.-M., Cosme E., Testut C.-E., Brasseur P. and Verron J., 2011: Efficient local error parameterizations for square root or ensemble Kalman filters: application to a basin-scale ocean turbulent flow, Mon. Weather Rev., 139(2), 474-493.

Brankart J.-M., Candille G., Garnier F., Calone Ch., Melet A., Bouttier P.-A., Brasseur P. and Verron J., 2015: A generic approach to explicit simulation of uncertainty in the NEMO ocean model, Geophysical Model Development, 8, 1285–1297, https://doi.org/10.5194/gmd-8-1285-2015 .

Brankart J.-M., Testut C.-E., Béal D., Doron M., Fontana C. Meinvielle M. and Brasseur P., 2012: Towards an improved description of oceanographic uncertainties: effect of local anamorphic transformations on spatial correlations, Ocean Sci., https://doi.org/10.5194/os-8-121-2012, 8, 121–142.

Brasseur P., Bahurel P., Bertino L., Birol F., Brankart J.-M., Ferry N., Losa S., Rémy E., Schröter J., Skachko S., Testut C.-E., Tranchant B., van Leeuwen P.J. and Verron J., 2005: Data Assimilation for marine monitoring and prediction: The MERCATOR operational assimilation systems and the MERSEA developments, *Q. J. R. Met. Soc*., 131, 3561-3582.

Ciavatta, S., Kay, S., Saux-Picart, S., Butenschön, M., & Allen, J. I. (2016). Decadal reanalysis of biogeochemical indicators and fluxes in the North West European shelf-sea ecosystem. *Journal of Geophysical Research: Oceans*, *121*(3), 1824-1845.

Dale T., Rey F. and Heimdal B. R., 1999. Seasonal development of phytoplankton at a high latitude oceanic site. Sarsia 84:419-435. https://doi.org/10.1080/00364827.1999.10807347.

Di Stefano J., 2015. Dynamic Systems Biology Modeling and Simulation. London, UK: Academic Press/ Elsevier, 884 pp.

Garnier F., Brankart J.-M., Brasseur P. and Cosme E., 2016: Stochastic parameterizations of biogeochemical uncertainties in a 1/4° NEMO/PISCES model for probabilistic comparisons with ocean color data, J. Mar. Systems, 155, 59-72, https://doi.org/10.1016/j.jmarsys.2015.10.012.

Gharamti, M.E., Tjiputra J., Bethke I., Samuelsen A. , Skjelvan I., Bentsen M. and Bertino L. , 2017. Ensemble data assimilation for ocean biogeochemical state and parameter estimation at different sites. Ocean Modelling, 112, 65–89.

Goodliff, M., Bruening, T., Schwichtenberg, F., Li, X., Lindenthal, A., Lorkowski, I., Nerger, L. (2019) Temperature assimilation into a coastal ocean-biogeochemical model: Assessment of weakly and strongly-coupled data assimilation, Oce. Dyn., 69, 1217-1237, https://doi.org/10.1007/s10236-019-01299-7

Gregg, W. and Casey, N. ,2009. Skill assessment of a spectral ocean–atmosphere radiative model. Journal of Marine Systems, 76(1-2), 49–63.

Hartman S., Bett B., Durden J., Henson S., Iversen M., Jeffreys R., Horton T., Lampitt R. and Gates A., 2021. Enduring science: Three decades of observing the Northeast Atlantic from the Porcupine Abyssal Plain Sustained Observatory (PAP-SO), Progress Oceanogr., 191, 102508, https://doi.org/10.1016/j.pocean.2020.102508.

Hamill, T. M., 2001. Interpretation of Rank Histograms for Verifying Ensemble Forecasts, Monthly Weather Review, 129(3), 550-560, DOI: https://doi.org/10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2

Hordoir, R., et al., 2019, Nemo-Nordic 1.0: a NEMO-based ocean model for the Baltic and North seas – research and operational applications, Geoscientific Model Development, 12, 363-386, https://doi.org/10.5194/gmd-12-363-2019

Lellouche, J.-M., Le Galloudec, O., Drévillon, M., Régnier, C., Greiner, E., Garric, G., et al. (2013). Evaluation of Global Monitoring and Forecasting Systems at Mercator Océan. Ocean Sci. 9, 57–81.

Leroux S., Brankart J.M., Albert A., Brodeau L., Molines J.M., Jamet Q., Le Sommer J., Penduff T., and Brasseur P., 2022: Ensemble quantification of short-term predictability of the ocean dynamics at kilometric-scale resolution: A Western Mediterranean test-case, Ocean Science Discussions, https://doi.org/10.5194/os-2022-11.

Mogensen, K., M. A. Balmaseda, A. T. Weaver, M. Martin, and A. Vidard. "NEMOVAR: A variational data assimilation system for the NEMO ocean model." ECMWF newsletter 120 (2009): 17-22., http://dx.doi.org/10.21957/3yj3mh16iq

Mogensen, K., M. A. Balmaseda, and A. Weaver. "The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4, ECMWF." Tech Memo 668 (2012): 6222-6226, http://dx.doi.org/10.21957/x5y9yrtm

Neumann T (2000) Towards a 3D-ecosystem model in the Baltic Sea. J Mar Syst 25:405–419

Nerger, L., Janjić, T., Schröter, J., Hiller, W. (2012). A unification of ensemble square root Kalman filters. Monthly Weather Review, 140, 2335-2345. https://doi.org/10.1175/MWR-D-11-00102.1

Nerger, L., Hiller, W., Schröter, J.(2005). PDAF - The Parallel Data Assimilation Framework: Experiences with Kalman Filtering, Use of high-performance computing in meteorology: proceedings of the Eleventh ECMWF Workshop on the Use of High-Performance Computing in Meteorology, Reading, UK, 25 - 29 October 2004 / Eds.: Walter Zwieflhofer; George Mozdzynski, Singapore: World Scientific, 63-83. https://doi.org/10.1142/9789812701831_0006

Nerger, L., Hiller, W. (2013). Software for Ensemble-based Data Assimilation Systems - Implementation Strategies and Scalability. Computers and Geosciences, 55, 110-118. https://doi.org/10.1016/j.cageo.2012.03.026

Nerger, L., Tang, Q., Mu, L. (2020). Efficient ensemble data assimilation for coupled models with the Parallel Data Assimilation Framework: Example of AWI-CM. Geoscientific Model Development, 13, 4305–4321, https://doi.org/10.5194/gmd-13-4305-2020

Nerger, L. (2022) Data assimilation for nonlinear systems with a hybrid nonlinear-Kalman ensemble transform filter. Q. J. Meteorol. Soc., 148, 620-640 https://doi.org/10.1002/qj.4221

Pham D.T., Verron J., Roubaud M.C., 1998. A singular evolutive extended Kalman filter for data assimilation in oceanography. Journal of Marine Systems 16, 323–340. https://doi.org/10.1016/S0924-7963(97)00109-7

Pham, D. T., 2001: Stochastic methods for sequential data assimilation in strongly nonlinear systems. Mon. Wea. Rev., 129, 1194–1207.

Pradhan, H.K., Voelker, C., Losa, S.N., Bracher, A., Nerger, L. (2019) Assimilation of global total chlorophyll OC-CCI data and its impact on individual phytoplankton fields. J. Geophys. Res. Oceans, 124, 470-490, https://doi.org/10.1029/2018JC014329

Sakov P. and Oke P., 2008. A deterministic formulation of the ensemble Kalman filter: An alternative to ensemble square root filters. Tellus 60(2):361 – 371, https://doi.org/0.1111/j.1600-0870.2007.00299.x.

Salon S., Cossarini G., Bolzon G., Feudale L., Lazzari P., Teruzzi A., Solidoro C., Crise A., 2019. Novel metrics based on Biogeochemical Argo data to improve the model uncertainty evaluation of the CMEMS Mediterranean marine ecosystem forecasts. Ocean Science 15, 997–1022. https://doi.org/10.5194/os-15-997-2019

Santana-Falcon Y., Brasseur P., Brankart J.M. and Garnier F., 2020: Assimilation of chlorophyll data into a stochastic ensemble simulation for the North Atlantic Ocean, Ocean Sci., 16, 1297–1315, 2020, https://doi.org/10.5194/os-16-1297-2020

Simon E., Samuelsen A., Bertino L. and Mouysset S., 2015. Experiences in multiyear combined state–parameter estimation with an ecosystem model of the North Atlantic and Arctic Oceans using the Ensemble Kalman Filter, J. Marine Syst., 152, 1-17, https://doi.org/10.1016/j.jmarsys.2015.07.004.

Skákala, J., Bruggeman, J., Brewin, R. J. W., Ford, D. A., & Ciavatta, S. (2020). Improved representation of underwater light field and its impact on ecosystem dynamics: A study in the North Sea. Journal of Geophysical Research: Oceans, 125, e2020JC016122. https://doi.org/10.1029/2020JC016122

Skákala, J; Ford, D; Bruggeman, J; Hull, T; Kaiser, J; King, RR; Loveday, BR; Palmer, MR; Smyth, TJ; Williams, CAJ; Ciavatta, S; 2021. Towards a multi-platform assimilative system for North Sea biogeochemistry. Journal of Geophysical Research: Oceans.

Skákala, J., Bruggeman, J; Ford, D; Wakelin, S; Akpınar, A; Hull, T; Kaiser, J; Loveday, BR; O'Dea, E; Williams, CAJ; Ciavatta, S; 2022. The impact of ocean biogeochemistry on physics and its consequences for modelling shelf seas. Ocean Modelling.

Tagliabue, A., Kwiatkowski, L., Bopp, L., Butenschön, M., Cheung, W., Lengaigne, M. and Vialard, J. 2021. Persistent Uncertainties in Ocean Net Primary Production Climate Change Projections at Regional Scales Raise Challenges for Assessing Impacts on Ecosystem Services. Frontiers in Climate, 3. doi:10.3389/fclim.2021.738224.

Teruzzi, A., Bolzon, G., Feudale, L., Cossarini, G., 2021. Deep chlorophyll maximum and nutricline in the Mediterranean Sea: emerging properties from a multi-platform assimilated biogeochemical model experiment. Biogeosciences 18, 6147–6166. https://doi.org/10.5194/bg-18-6147-2021

Toedter, J., Kirchgessner, P., Nerger, L., Ahrens, B. (2016) Assessment of a nonlinear ensemble transform filter for high-dimensional data assimilation. Monthly Weather Review, 144, 409-427 https://doi.org/10.1175/MWR-D-15-0073.1

Waters, J., Lea, D.J., Martin, M.J., Mirouze, I., Weaver, A. and While, J. (2015), Implementing a variational data assimilation system in an operational 1/4 degree global ocean model. Q.J.R. Meteorol. Soc., 141: 333-349. https://doi.org/10.1002/qj.2388

Annex 1 - Probabilistic NEMO-PISCES developments

In this annex are described the scientific and technical developments carried out by IGE as part of Task 3.1 to generate the NEMO-PISCES probabilistic modelling tool used for the assimilation experiments of WP3.3 and subsequent activities. While the probabilistic version of NEMO already existed to simulate ocean-ice dynamic (Bessières *et al*., 2017), the extension to a coupled NEMO-PISCES global model is needed in the context of SEAMLESS (and further GLO and IBI implementations).

Consistently with the work done in WP2, the NEMO-PISCES version considered for this transformation is inherited from the r4.0-HEAD.r13720 NEMO code used in the CMEMS GLO MFC and ported into FABM (see D2.2). The probabilistic NEMO-PISCES code developed in Task 3.1 relies on independent Gaussian autoregressive processes as described in Brankart *et al*. (2015) and Garnier *et al.* (2016). The stochastic routines of the stochastic code (delivered earlier with D3.3) can be easily used to simulate uncertainties in physical model parameterizations, external forcing, or even initial conditions (Leroux *et al*., 2022).

The prior ensemble is generated by introducing stochasticity in the evolution equations used to simulate the time evolution of all ensemble members. Stochastic perturbations are calculated assuming probability distributions at every time step, taking into account previous values within the predefined correlation time interval. The stochastic processes simulate a variety of uncertainty sources such as uncertain biogeochemical model parameters, unresolved subgrid scale processes and location uncertainties. The 3 following classes of stochastic parameterizations have been implemented in the NEMO-PISCES code:

- parameterization of uncertainties in biogeochemical parameters;
- parametrization of unresolved processes at sub-grid scales;
- perturbations of the horizontal mesh to simulate location uncertainties.

A1.1 Parameterization of uncertain biogeochemical parameters

This type of parameterization is implemented by introducing additional exponential factors for selected parameters in the evolution equations of the concentration C of biogeochemical variables:

$$\left.\frac{\partial C}{\partial t}\right|_{bio} = \mathrm{SMS}\left(C, u, pe^{\xi(t)}, t\right),$$

where u represents the external forces and p is a perturbed parameter. The perturbations are introduced with the help of Gaussian distributions

$$\xi(t) \sim \mathcal{N}(\mu = 0, \sigma = 0.3).$$

The quantity $p' = pe^{\xi(t)}$ can be considered as a horizontal field of white noise with log-Normal distribution $\log\mathcal{N}(\mu = 0, \sigma = 0.3)$, which is very close to the Gaussian $\mathcal{N}(\mu = 1, \sigma = 0.3)$.

The distributions $\xi(t)$ are defined as 2D horizontal fields by using autoregressive processes, where the values at a given instant t_{n+1} are linearly dependent on previous values of the same time series. In such a way, the temporal correlation of the considered state is taken into account. An autoregressive process is calculated using the recurrence relation

$$\xi(t_{n+1}) = a\xi(t_n) + bw + c\ ,$$

where w is a Gaussian white noise, a, b and c are free parameters, calculated on the basis of the desired mean value, standard deviation, and correlation interval in time. In addition, a horizontal 5-points Laplace filter is applied to $\xi(t)$ on the discrete mesh. The noise w is bounded by a given threshold to avoid inconsistent perturbations.

In practice, we replace the predefined values of the parameters p that are considered uncertain by the stochastic fields which have a log-Normal distribution around the original deterministic values.

In the experiments reported in section 4.2, we have chosen seven parameters of the PISCES model, which are subject to significant uncertainties (Garnier *et al.*, 2016) and generate perturbations in the simulation results in the North Atlantic configuration, consistently with the main conclusions of the sensitivity analysis documented in D3.2. These parameters are:

- photosynthetic efficiency of nanophytoplankton;
- photosynthetic efficiency of diatoms;
- phytoplankton growth rate at 0°C;
- sensitivity of phytoplankton growth to temperature;
- sensitivity of zooplankton grazing to temperature;
- dependence of nanophytoplankton growth on day length;
- dependence of diatoms growth on day length.

A1.2 Parameterization of unresolved scales

A second type of uncertainty associated to unresolved sub-grid scale processes is taken into account by introducing stochastic perturbations of the concentrations themselves C, still using autoregressive processes as described above. This is achieved by computing the right hand side of the evolution equation of the BGC variables as an average of two source terms, where the perturbations of the variables are added with opposite signs:

$$\left.\frac{\partial C}{\partial t}\right|_{bio} = \frac{1}{2}[\mathrm{SMS}(C + C\xi(t), u, p, t) + \mathrm{SMS}(C - C\xi(t), u, p, t)]$$

The perturbations $C\xi(t) \equiv \delta C(t)$ are considered as fluctuations, not resolved by the mesh. The non-linearities of the SMS terms are in this case a source of uncertainty, which can be significant especially

for coarse resolution models. In total there are 24 biogeochemical state variables in PISCES, and this procedure is applied to 20 of them. The procedure is not applied to dissolved inorganic concentration, alkalinity, oxygen and calcite variables since these variables are directly diagnosed from the 20 remaining PISCES variables.

A1.3 Horizontal perturbations of the mesh

A third type of uncertainties has been implemented in the coupled NEMO/PISCES code in order to account for errors in the location and spatial structure of the physical and biological features (essentially at the mesoscale) simulated by the coupled model. The concept of location uncertainties has been previously developed in NEMO (physics only) within the framework of the H2020 IMMERSE project, to assess the predictability of the sub-mesoscale circulation in the Mediterranean Sea. They are adapted here to represent location uncertainties, which can affect both physical and biogeochemical patterns.

The idea is to perturb the NEMO metrics as described in Leroux *et al*. (2022), by adding random fluctuations to the Lagrangian displacement of fluid particles. This is achieved by transforming the arrays describing the cell sizes of the horizontal mesh in the coupled NEMO/PISCES code into time-dependent stochastic processes. The stochastic perturbations are set with the help of autoregressive processes, as in the cases of the other types of uncertainties as described above.

Annex 2 - Probabilistic HYCOM-ECOSMO developments

In this annex are reported the scientific and technical developments carried out by NERSC as part of task 3.1 to generate the HYCOM-ECOSMO probabilistic modelling tool (D3.1).

In addition to ensemble generation system based on ECOSMO biogeochemical model parameters ensemble perturbation system in the current ARC MFC biogeochemical reanalysis system, we have newly implemented parameter ensemble perturbation in HYCOM ocean model mixing parameterization module. The HYCOM ocean model parameters selected for the ensemble generation are listed on Table A2.1. 6 parameters are chosen from KPP vertical mixing parameterisation module aiming at probabilistic description of mixed layer depth development. Given default values, we added Gaussian random error to each parameter. The size of standard error is defined by 20% of prescribed parameter range (see Table A2.1). Note that the parameter values outside the prescribed parameter range are pushed back to edge of the range after ensemble parameter members are generated.

Table A2.1 HYCOM model parameters for ensemble generation. qrinfyl: maximum gradient Richardson number (shear inst.). ricr: critical bulk Richardson number. cs: constant 1 for nonlocal flux term. cstar: constant 2 for nonlocal flux term. cekman: scale factor for Ekman depth. cmonob: scale factor for Monin-Obukov depth.

Parameter	qrinfy	ricr	cs	cstar	cekman	cmonob
Error [%]	20	20	20	20	20	20
Min value	0.5	0.25	98.0	9.75	0.6	0.95
Max value	1.0	0.5	100.0	10.25	0.8	1.05
Default value	0.7	0.45	98.96	10.0	0.7	1.0

Annex 3 - Analysis tools based on anamorphic transformations

In this annex are described the developments carried out by IGE as part of task 3.3a to upgrade the analysis tools based on anamorphic transformations and to facilitate the porting to the SEAMLESS prototype or to CMEMS MFC operational systems.

Anamorphosis is a nonlinear transformation that is applied to a model state variable x to transform its marginal probability distribution into a Gaussian distribution (with zero mean and unit variance). It is useful because many data assimilation method (like Ensemble Kalman Filter or other similar ensemble methods developed in SEAMLESS) make the assumption of Gaussian distributions. In this annex, section 1 provides a brief reminder of the original algorithm that was used as a starting point, section 2 presents the two main limitations of this algorithm, and section 3 and 4 describe the solution that have been developed to address these two limitations. The algorithm described below are all available as part of the EnsDAM library, more specifically in the directory: https://github.com/brankart/ensdam/tree/master/src/EnsAnam.

1. Description of the original algorithm

Our starting point is the simple anamorphosis algorithm described in Brankart et al. (2012), which consists in remapping the quantiles of the marginal distribution of x on the quantiles of the target Gaussian distribution (with zero mean and unit variance), using a piecewise linear transformation (interpolating between the quantiles). The transformed variable $x'=A(x)$ is then approximately Gaussian.

The quality of the approximation depends on the size of the ensemble that is used to describe the probability distribution of x, which imposes a limit on the number of quantiles that can be used to describe the transformation. With a larger ensemble, more quantiles can be used and the transformation is more accurate. However, even if the transformation is only approximate, x' is still closer to Gaussian than x (especially if x is subjected to inequality constraints such as positiveness of concentration variables), and the problem is usually simplified.

Many data assimilation methods (like the Ensemble Kalman Filter) also require that the observation operator (linking state variables to observations) is linear, and that the observation error is Gaussian. With such methods, anamorphosis must also be applied to observations to keep the observation operator linear for the transformed variables, and the observation error on this transformed observation must be assumed Gaussian. A simple approach to do this is just to apply the same anamorphosis operation to the observation themselves and multiply the observation error standard deviation by the slope of the transformation, but this simple formulation is not without causing potential problems as explained below.

Other data assimilation methods (as the MCMC sampler proposed in Brankart, 2019, but not used yet in WP3 experiments) do not require that the observation operator is linear, so that observations do not need to be transformed. The original observation operator needs only to be complemented by an inverse anamorphosis transformation: $H(x)$ is just replaced by $H[A^{-1}(x')]$. This makes the use of

anamorphosis much easier since only the state variables need to be transformed, not the observations.

2. Limitations of the algorithm

A first important difficulty with the algorithm described above is related to the transformation of observations. If the observation error is large, transforming the standard deviation using the local slope of the transformation becomes a very rough approximation, which can lead to substantial biases in the transformed observations.

Second, a key assumption of the algorithm is that the cdf of x is continuous, so that it can be inverted to obtain the anamorphosis transformation. If there is a finite probability concentrating on a particular value of x, the cdf becomes stepwise (several quantiles are equal), it cannot be inverted, and the original algorithm does not apply.

3. Transformation of the observations

The following algorithms have been implemented to address the problem of transforming the observations. We first consider a simplified particular case allowing a more efficient algorithm, before addressing the general problem. In both cases, the objective is to produce an unbiased transformed observation, with a consistent observation error standard deviation.

a. The observation error pdf is symmetric and does not dependent on the true state

In this particular case, the transformation of the probability distribution for observation error is straightforward, and can be obtained using the following algorithm for each observation (assuming independent observation errors):

1. Compute the anamorphosis transformation A for the observed quantity from the ensemble equivalent of the observation: Hx_i, i =1,…,m.
2. Produce a sample of perturbed observations: $y^o_j = y^o + e_j$, where the perturbations e_j are sampled from the observation error probability distribution.
3. Transform the sample using the anamorphosis transformation A.
4. Use the mean and covariance of the sample as parameters for the transformed observation error probability distribution (if assumed Gaussian).

b. General observation error probability distribution

In the general case, when the observation error probability distribution depends on the true state of the system (e.g., ocean colour data errors are usually represented by multiplicative noise) or when it is not symmetric, perturbations cannot be added to the observations. They can only be added to a model equivalent to the observations Hx. In this case, applying the simplified algorithm above can lead to substantial biases in the transformed observations, especially for bounded variables (when observations are close to the bounds). It is then important to use a more general algorithm to obtain the transformed observation and observation error:

1. Sample a rank r for the observation error (uniformly between 0 and 1).

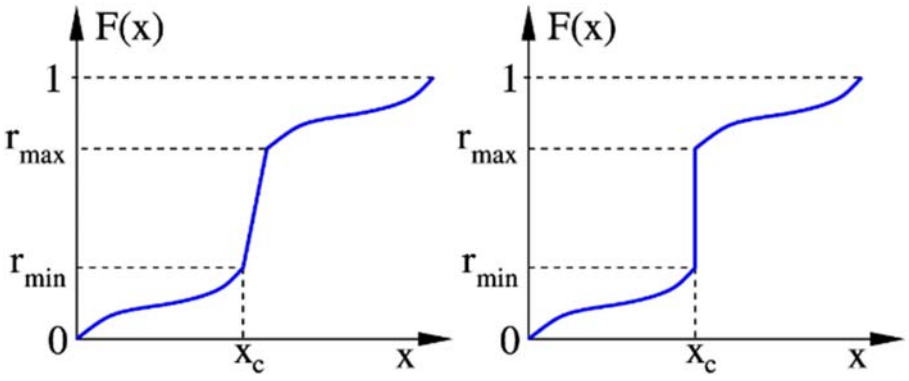
2. Perturb ensemble members with an observation error with the given rank: $y_i = pert(Hx_i) = F^{-1}(r)$, where F is the cdf of the observation error probability distribution $p(y_i|Hx_i)$.
3. Compute the anamorphosis transformation A from this transformed ensemble.
4. Transform the observation with A to obtain a transformed perturbed observation.
5. Repeat the above steps for a sample of ranks to obtain a sample of transformed observations.
6. Use the mean and covariance of the sample as parameters for the transformed observation error probability distribution (if assumed Gaussian).

It is easy to see that this more general algorithm is equivalent to the simplified algorithm above in the particular case of observation errors that have a symmetric probability distribution, independent of the state of the system (for instance if it is Gaussian with a variance independent of x).

4. Discrete events

In many practical applications, there can be situations in which a finite probability concentrates on some critical value x_c of the state variable (for instance zero for concentrations). In this case the cdf F(x) is discontinuous and the standard anamorphosis algorithm described above does not apply.

To generalize the algorithm, we can imagine the discontinuity in F(x) as the limit of a very steep slope (as illustrated in the figure below). As long as there is a slope (left panel), we know which value of the rank r=F(x) corresponds to every value of x: a small uncertainty in x just produces a larger uncertainty in r when the slope is steeper. As soon as the slope becomes a step (right panel), we do not know anymore which rank r, between r_{min} and r_{max}, should correspond to x_c.



The solution is then to make the transformation stochastic and transform x to a random rank (with uniform distribution) between r_{min} and r_{max}. In this way, the forward transformation will transform the marginal distribution of all variables to the target distribution as required, the discrete events being transformed into a continuous variable by the stochastic transformation; and the backward transformation will transform it back to a discrete event, by transforming all ranks between r_{min} and r_{max} to x_c.

In the above scheme, it is important that the ranks are sampled independently for different members, but not necessarily for different variables of the state vector. We have thus the freedom to introduce spatial correlation in the sampling of the ranks r. If the transformed ensemble is meant to be updated with the assumption of joint Gaussianity (as in the Ensemble Kalman Filter), a reasonable option is to

avoid destroying the ensemble correlation structure where part of the members displays the discrete event $x=x_c$. This can be done by using the same random rank for all variables from the same member. In this way, decorrelation can only be amplified where members move from a critical value to a non-critical value.

The EnsDAM library has been updated as part of Task 3.3.a to incorporate the solutions presented above.