The assimilation of phytoplankton functional types for operational forecasting in the North-West European Shelf

Jozef Skákala[1,2], David Ford[3], Robert J.W. Brewin[1,2], Robert McEwan[3], Susan Kay[1,3], Benjamin Taylor[1], Lee de Mora[1], and Stefano Ciavatta[1,2]

[1]Plymouth Marine Laboratory, The Hoe, Plymouth, PL1 3DH United Kingdom.
[2]National Centre for Earth Observation, Plymouth, PL1 3DH, UK.
[3]Met Office, FitzRoy Road, Exeter, EX1 3PB UK.

Key Points:

- We assess the forecasting skill of PFTs and total chlorophyll-a DA.
- PFTs chlorophyll-a DA performs best in 5 day forecasting.
- DA substantially improves the representation of CO_2 fugacity.

Corresponding author: Jozef Skákala, jos@pml.ac.uk

Abstract

This paper proposes the use of assimilation of phytoplankton functional types (PFTs) surface chlorophyll for operational forecasting of biogeochemistry on the North-West European (NWE) Shelf. We explicitly compare the 5 day forecasting skill of three runs of a physical-biogeochemical model: a) a free reference run, b) a run with daily Data Assimilation (DA) of total surface chlorophyll (ChlTot) and (c) a run with daily PFTs DA. We show that small total chlorophyll model bias hides comparatively large biases in PFTs chlorophyll, which ChlTot DA fails to correct. This is because in our study the ChlTot DA splits the assimilated total chlorophyll into PFTs by preserving their simulated ratios, rather than taking account of the observed PFT concentrations. Unlike ChlTot DA, PFTs DA substantially improves model representation of PFTs chlorophyll. During forecasting the DA reanalysis skill in representing PFTs chlorophyll degrades towards the free run skill, however PFTs DA outperforms free run within the whole 5-day forecasting period. We validated our results with in situ data and we demonstrated that (in both DA cases) the DA substantially improves the model representation of CO_2 fugacity (PFTs DA more than ChlTot DA). ChlTot DA has a positive impact on the representation of silicate, while the PFTs DA seems to have a negative impact. The impact of DA on nitrate and phosphate is not significant. The implications of using a univariate assimilation method which preserves the phytoplankton stochiometry, as well as the impact of model biases on the non-assimilated variables are discussed.

1 Introduction

Monitoring biogeochemistry in shelf seas is of great significance for the economy, ecosystems understanding and climate studies. The shelf seas contain 90% of world's fisheries and are responsible for 20% of marine primary production and 20% of atmospheric carbon dioxide uptake (*Pauly et al. [2002]; Borges et al. [2006]; Jahnke [2010]*). In the North-West European (NWE) Shelf ecosystem the need for more detailed information about marine ecosystem indicators and processes has been clearly pointed out by both users and policy makers (*Chassot et al. [2007]; Blauw et al. [2010]; Brandsma et al. [2013]; Skogen et al. [2014]; Kurekin et al. [2014]; Ford et al. [2017]*). Data Assimilation (DA) maximizes the use of information about processes in the shelf seas by methodically combining the available information from Earth Observations (EO) (satellite data), model simulations and sometimes also in situ measurements. The DA methods applied in ecosystem modelling have been successfully used in reanalysis simulations (i.e assimilation of time series in past observations of the system, e.g *Nerger and Gregg [2007]*) as well as operational forecasting (i.e the assimilation of recent observations to initialize model predictions of the future biogeochemical state, e.g *Teruzzi et al. [2014]*).

DA has its most well known application in numerical weather forecasting (*Kalnay [2003]*), but has also been applied for a long time in physical oceanography (for an overview see *Cummings et al. [2009]; Edwards et al. [2015]*). There are also a growing number of studies applying DA to ecosystem variables (*Gehlen et al. [2015]*). This is mostly focused on (ocean-color derived) chlorophyll-a (*Ishizaka [1990]*) using typically Kalman Filter methods (*Carmillet et al. [2001]; Natvik and Evensen [2003]; Hoteit et al. [2005]; Torres et al. [2006]; Nerger and Gregg [2007, 2008]; Fontana et al. [2010]; Ciavatta et al. [2011]; Simon and Bertino [2012]; Simon et al. [2015]; Ciavatta et al. [2016]*), but also Optimal Interpolation (*Gregg [2008]*) and variational methods (*Losa et al. [2004]*). There are also studies on biogeochemical DA of some optical fields: phytoplankton light absorbtion (*Shulman et al. [2013]*), diffuse light attenuation coefficient (*Ciavatta et al. [2014]*)), reflectance data (*Jones et al. [2016]*) and absorbtion by Colored Dissolved Organic Carbon (CDOC) (*Gregg and Rousseaux [2017]*). The variable most commonly used for DA is ocean-color derived total chlorophyll-a. Total chlorophyll-a relates to total phytoplankton, which contains species that vary in size by 9 orders of magnitude (*Finkel et al. [2009]*) and play very different roles within the ecosystem dynamics (*Lé Quére et al. [2005]*).

Many ecosystem models such as the European Regional Seas Ecosystem Model (ERSEM) (*Baretta et al. [1995]; Butenschön et al. [2016]*) therefore split phytoplankton into functional types (PFTs), largely based on the characteristic size and ecological niche of the functional group. It is acknowledged (*Gregg [2008]; Teruzzi et al. [2014]; Gehlen et al. [2015]*) that whilst DA of total chlorophyll-a improves the total chlorophyll representation, it often fails to improve the representation of other model variables (such as nutrients). This often results from the limitation imposed by univariate approaches, which update non-assimilated variables only through the model dynamics (see *Nerger and Gregg [2007]* for a discussion). However, the problem exists also for multivariate assimilation methods which can have limited, or even negative impacts on non-assimilated variables, in particular when the model has severe biases, for example because of the incomplete representation of the ecosystem processes, or deficiencies in specifying internal model parameters (see discussion in e.g *Ford et al. [2012]; Ciavatta et al. [2016]; Tsiaras et al. [2017]; Ciavatta et al. [2018]*). One might expect to improve the overall biogeochemical simulation through improvement in simulation of the phytoplankton community, which is the central component of the low trophic level models. The assimilation of total chlorophyll might not be sufficient for this purpose, because often this approach is not capable of correcting the relative ratios of the PFTs composing the community (*Ciavatta et al. [2011]*). This issue can be avoided by directly assimilating PFTs chlorophyll when the PFTs chlorophyll-a data are available. Such an approach was taken in an early 1D study by *Xiao and Friedrichs [2014]* and recently by *Ciavatta et al. [2018]* in a 3D model configuration of the NWE Shelf.

In the NWE Shelf *Brewin et al. [2017]* developed a novel phytoplankton size-class chlorophyll data-set for the Copernicus Marine Environment Monitoring Service (CMEMS, http://marine.copernicus.eu) project Towards Operational Size-Class Chlorophyll Assimilation (TOSCA), and this data-set can be directly associated with the PFTs used in the ERSEM model. These are (*Butenschön et al.* [2016]): picophytoplankton ($< 2\mu m$), nanophytoplankton ($2 - 20\mu m$) and microphytoplankton ($> 20\mu m$). Microphytoplankton is split into diatoms (having silicate cell walls) and dinoflagellates. The chlorophyll-a contained in the PFTs can be then directly assimilated into the ERSEM model (this is called PFTs DA in the rest of this article). It is expected that this would improve the representation of ecosystem dynamics compared to assimilation of total chlorophyll-a (ChlTot). The difference the PFTs DA makes to total chlorophyll (ChlTot) DA was shown to be significant in a 6-year reanalysis that assimilated monthly PFT data using Ensemble Kalman Filter (EnKF) and the pre-operational model Proudman Oceanographic Laboratory Coastal Ocean Modelling System (POLCOMS) - ERSEM (*Ciavatta et al. [2018]*). In this paper we focus on PFTs DA in the context of an operational system developed at the Met Office, based on the coupled model Nucleus for European Modelling of the Ocean (NEMO) - ERSEM and the variational DA system NEMOVAR (*Mogensen et al. [2009, 2012]; Waters et al. [2015]*). The differences to *Ciavatta et al. [2018]* are that we use daily DA (as opposed to monthly DA), different model (NEMO-ERSEM at 7 km resolution, as opposed to POLCOMS-ERSEM at 12 km resolution) and a different DA scheme (3DVAR, as opposed to EnKF). Most importantly, unlike *Ciavatta et al. [2018]* our objective is to assess the impact of PFTs DA on forecasting. This is because the NEMO-ERSEM model used here is run operationally at the Met Office, delivering daily analysis and forecast products to CMEMS, and it is planned to implement the assimilation scheme presented here as part of future upgrade (an outcome of the CMEMS TOSCA project). We compare PFTs DA forecasting skill with the ChlTot DA forecast and a free reference run. As with *Ciavatta et al. [2018]*, our analysis focuses on the NWE Shelf.

2 Methods

2.1 The physical component: NEMO

The Nucleus for European Modelling of the Ocean (NEMO) ocean physics component (OPA) is a finite difference, hydrostatic, primitive equation ocean general circulation model (*Madec et al. [2015]*). The version used in this work is CO6, based on NEMOv3.6, a development of the CO5 configuration described by *O'Dea et al. [2017]*. The model configuration was similar to *Ford et al. [2017]*. The model used the 7 km resolution grid on the Atlantic Meridional Margin (AMM7) domain with 51 vertical levels and a terrain-following $z^* - \sigma$ coordinate system. The river inputs were set using a climatology of daily discharge (*Edwards et al. [2012]*). The lateral boundary conditions for physical variables at the Atlantic boundary were taken from a reanalysis of the GloSea5 Seasonal Forecasting System (*MacLachlan et al. [2015]*); the Baltic boundary values were derived from a reanalysis produced by the Danish Meteorological Institute for CMEMS. The model was forced at the surface by atmospheric fluxes from the ERA-Interim reanalysis (*Dee et al. [2011]*). The same reanalysis data were used to force our 5-day model forecast experiments because suitable forecast fluxes were not available for the same period as the biogeochemical observation data used.

2.2 The ecosystem component: ERSEM

The European Regional Seas Ecosystem Model (ERSEM) (*Baretta et al. [1995]; Butenschön et al. [2016]*) is an ecosystem model for marine biogeochemistry, pelagic plankton, and benthic fauna (*Blackford [1997]*). It tracks carbon, chlorophyll, nitrate, phosphate and silicate separately, with variable stoichiometric ratios within the simulated plankton groups (*Geider et al. [1997]; Baretta-Bekker et al. [1997]*). The model splits phytoplankton into four functional types largely based on their size (*Baretta et al. [1995]*): picophytoplankton, nanophytoplankton, diatoms and dinoflagellates; only diatoms use silicate. Phytoplankton are a prey for three zooplankton types (mesozooplankton, microzooplankton and heterotrophic nanoflagellates) and organic material is decomposed by one functional type of heterotrophic bacteria (*Butenschön et al. [2016]*). The inorganic component is described in the form of nutrients (nitrate, phosphate, silicate, ammonium and carbon) and dissolved oxygen. The carbonate system is also included in the model (*Artioli et al. [2012]*). The ERSEM model has been validated in multiple studies using both point-wise and emergent skill metrics (*Allen and Somerfield [2009]; Edwards et al. [2012]; Saux Picart et al. [2012]; De Mora et al. [2013, 2016]*), and applied in many different contexts (e.g *Blackford and Gilbert [2007]; Holt et al. [2012]; Wakelin et al. [2012]; Polimene et al. [2012]; Artioli et al. [2014]*).

We used in this study a recent ERSEM parametrization described in *Butenschön et al. [2016]*. At the Atlantic boundary values for nitrate, phosphate and silicate were taken from World Ocean Atlas (*Garcia et al. [2014]*) and dissolved inorganic carbon from the GLODAP gridded dataset (*Key et al. [2015]; Lauvset et al. [2016]*).

2.3 The Data

The original data-set of total chlorophyll-a was obtained from the Ocean Colour - Climate Change Initiative (OC-CCI) project of the European Space Agency (ESA), Version 3.0 (*Sathyendranath et al. [2016]*). This total chlorophyll product was processed into a phytoplankton functional types chlorophyll data-set by *Brewin et al.* [2017] using a simple, conceptual model (*Brewin et al.* [2010, 2015]) designed to estimate the chlorophyll concentrations of three phytoplankton size classes (micro-, nano- and pico-phytoplankton) as a continuous function of the total chlorophyll provided from the OC-CCI data. In the implementation, the parameters of the model are varied according to the sea surface temperature (OISST version from *Reynolds et al. [2007]*), which is also used to split microphytoplankton chlorophyll concentration into the contributions from diatoms and dinoflagellates. The product of *Brewin et al.* [2017] is daily and has 4 km spatial resolution. The EO data validate well against in situ data (Pearson correlation coefficient 0.46-0.86, see *Brewin et al. [2017]*). The PFT EO errors were estimated in log-space, since chlorophyll is typically log-normally distributed *Campbell [1995]*. The PFT EO errors and biases were determined using both in situ and satellite data match-ups following the approach from *Jackson et al. [2017]* and fuzzy logic statistics (*Moore et al. [2009]*). The data (for both total chlorophyll and PFTs) were bias corrected and per pixel errors of the unbiased data were computed following the method of *Ciavatta et al. [2016]*. Because bias corrected EO products are supposed to be better than the original ones, it is reasonable to assimilate bias-corrected data. However, the sum of bias corrected PFTs chlorophyll may not be precisely equal to bias corrected total chlorophyll (for details see *Brewin et al. [2017]*). In fact the mean sum of bias corrected PFTs chlorophyll was approximately $0.07 mg/m^3$ lower than the mean value for bias corrected total chlorophyll, for 2010 data on the NWE Shelf. The bias-corrected EO data were upscaled to the model grid (wherever there were multiple EO data-points mapped to the nearest model grid point, the mean value of those data-points was taken). We also compared the 2010 OC-CCI chlorophyll data with the OC-CCI satellite data monthly climatology which was composed from bias-corrected OC-CCI products from 1998-2009.

The DA outputs were compared on the NWE Shelf with three in situ data-sets. The first was the Ecosystem Data Online Warehouse of the International Council for the Exploration of the Sea (ICES, http://www.ices.dk/marine-data/data-portals/Pages/), which contains measurements of three nutrients of specific interest (nitrate, phosphate and silicate) and also data for total chlorophyll. The ICES data-set contains measurements at a range of depths. We considered only ICES data from the section of the NWE Shelf not in the immediate vicinity of the coastline (bathymetry within the interval $10 - 200$ m). ICES data were available all over the North Sea and Irish Sea, however with a clear spatial bias towards nutrient- and chlorophyll-rich areas close to the coast of the Netherlands and western Denmark. The median depth of the measurement was around 10 m, but could vary from month to month. Also numbers of measurements varied from month to month between 20 and 300. The total number of ICES data-points for 2010 was well over 1000 for each nutrient and for total chlorophyll. The second data-set was from the Centre for Environment, Fisheries and Aquaculture Science (Cefas, https://www.cefas.co.uk/) and consisted of phytoplankton pigment data (nanophytoplankton, picophytoplankton and microphytoplankton) collected on International Bottom Trawl Surveys in the years 2010 and 2011 (*Ford et al. [2017]*, http://doi.org/10.14466/CefasDataHub.33). The Cefas data-set contained far less data than the ICES data-set (only around 60 data-points in the relevant area for 2010), but is one of the few available in situ data sets that can be used to perform an independent validation of PFT distributions. The third in situ comparison was for CO_2 fugacity (fCO_2) using the Surface Ocean CO_2 atlas (SOCAT, https://www.socat.info/, *Bakker et al. [2014]*). The SOCAT dataset was the most statistically robust of the three used, with around 10000 data-points. We also did a comparison for PFTs/total chlorophyll and nutrients (nitrate, silicate and phosphate) at the specific location L4 in the Western English Channel, with data obtained from the HPLC Western Channel Observatory pigments & nutrients data-set (*Airs and Martinez-Vicente [2014], https://www.bodc.ac.uk/-data/*). The in situ chlorophyll concentrations for the four PFTs at L4 (diatoms, dinoflagellates, nanophytoplankton and picophytoplankton) were estimated from HPLC pigment data following *Brewin et al. [2017]*. This essentially involves using accessory pigments as markers of the specific groups to help partition total chlorophyll into the chlorophyll concentrations of the four groups (see section 2.3.1 of *Brewin et al. [2017]* for additional details). All the in situ data were matched with the model outputs by finding the model grid point nearest to the in situ measurement.

2.4 The Data Assimilation (DA) set-up

We used the NEMOVAR (*Mogensen et al. [2009, 2012]; Waters et al. [2015]*) 3D-VAR variational DA system used for operational physical ocean DA at the Met Office. NEMOVAR is a computationally efficient DA system specifically adapted for the NEMO model, supporting both 4D-VAR and 3D-VAR algorithms. The 3D-VAR version applied in this study minimizes the cost function using the conjugate gradient method (*Mogensen et al. [2012]*). DA of chlorophyll into NEMO-ERSEM using NEMOVAR has been implemented at the Met Office for use in reanalysis and forecasting.

The PFTs and total chlorophyll DA has been adapted from the method used to assimilate total chlorophyll into the global NEMO-HadOCC model (*Ford et al. [2012]; Ford and Barciela [2017]*). The DA was run on a daily cycle, assimilating the daily merged OC-CCI chlorophyll products. Since chlorophyll is typically lognormally distributed (*Campbell [1995]*), $\log_{10}$(chlorophyll) was assimilated rather than chlorophyll. For total chlorophyll the procedure is described in the following steps.

Firstly, the model was run for the day in order to create innovations (observation minus background differences) using the NEMO observation operator. As in *Ford et al. [2012]*, the model surface total $\log_{10}$(chlorophyll) (i.e. the sum of the four PFTs in ERSEM) is bilinearly interpolated to each observation location at the nearest model time step to the validity time of the observation, providing background values in observation space. Since daily merged products were assimilated, with no per-pixel time information provided, all observations were assumed to be valid at 12:00 UTC. As the ocean color satellites used by OC-CCI are all heliosynchronous, this is a reasonable assumption for the AMM7 domain.

Secondly, these innovations were used by NEMOVAR to create a set of surface total $\log_{10}$(chlorophyll) increments, similarly to the DA of sea ice concentration described by *Waters et al. [2015]*. The model errors were specified by deriving the diagonal elements of the background error covariance matrix from a monthly climatology of log-transformed error variances obtained from the 100 member Ensemble Kalman Filter POLCOMS-ERSEM reanalysis of *Ciavatta et al. [2018]*. These variances were regularized and smoothed using the moving averages algorithm, and rescaled to the range 0.02-1.5 $\log_{10}(mg/m^3)$, so that the average ratio of background error to obervation error was similar to that calculated in the region when assimilating OC-CCI data into NEMO-HadOCC (*Ford and Barciela [2017]*). Experiments using different ratios demonstrated the results to be relatively insensitive to the average ratio. The off-diagonal elements of the background error covariance matrix were parametrised using correlation lengthscales set equal to the Rossby radius, as in *Waters et al. [2015]*. The diagonal elements of the observation error covariance matrix were set equal to the per-pixel observation uncertainties from the OC-CCI products (*Ciavatta et al. [2016]*), plus a constant of 0.01 $\log_{10}(mg/m^3)$ (*Ford and Barciela [2017]*), to take account of the remaining representation error (*Janjić et al. [2017]*) not included in the OC-CCI uncertainties, whilst maintaining the average ratios suggested by *Ford and Barciela [2017]*. The off-diagonal elements of the observation error covariance matrix were set to zero.

Thirdly, the model background was used to convert the total $\log_{10}$(chlorophyll) increments to total chlorophyll increments, and divide them into a set of chlorophyll increments for each PFT. At each grid point the total chlorophyll increments were split into PFT chlorophyll increments according to the ratios of the PFTs in the model background, so that the assimilation did not directly alter the phytoplankton community structure. Up to this stage the DA scheme updated only PFTs chlorophyll. The DA set-up was tested with this simplistic scheme (only updating PFTs chlorophyll) and the results are presented in the Supporting Information (1). However, it is important to maintain the phytoplankton physiological state adapted to the environmental conditions. To do this we used another scheme, where all the other phytoplankton cell variables (carbon, nitrogen, phosphorus and for diatoms silicon) were updated to preserve the existing stochiometric ratios. This means DA altered only concentrations of phytoplankton, but preserved its model physiology. Our approach is similar to the one used in *Teruzzi et al. [2014].*

Fourthly, the model was run again for the day to create the reanalysis state, with the increments applied using the incremental analysis update (IAU) technique (*Bloom et al. [1996]*), in which in an equal proportion of the increments is applied at each time step, in order to minimise initialisation shocks. The surface PFT (chlorophyll, carbon, nitrogen, phosphorus, silicon) increments were applied throughout the mixed layer. The reanalysis state was then used to initialize a 5-day "free" forecasting run.

The total chlorophyll assimilation has then been extended in this study to PFT chlorophyll assimilation, by considering each PFT separately at each step in the process. The observation operator step directly compared the satellite PFT data to the corresponding model PFTs, to create a set of innovations for each PFT. The background error variances used by NEMOVAR were calculated using the same method as for total chlorophyll DA from ensemble variances for the individual PFTs in the reanalysis of *Ciavatta et al. [2018].* The observational errors were obtained from the pixel errors provided by *Brewin et al. [2017]* with bias removed as per *Ciavatta et al. [2016]* and the representation error added as in case of total chlorophyll DA. NEMOVAR was used to calculate a set of $\log_{10}$(chlorophyll) increments for each PFT, which could be applied directly to the model, thereby directly updating the phytoplankton community structure.

2.5 The runs and the analysis

We performed three 1-year long simulations for 2010 on the Met Office and NERC Supercomputing Node (MONSooN). The first simulation was a free reference run (abbreviated as "noDA"), the second run was daily total chlorophyll DA (abbreviated as "ChlTot DA") and the third run PFTs chlorophyll daily DA (abbreviated as "PFTs DA"). In each DA run the assimilation step was followed by a 5 day forecast.

It is important to assess how DA impacts on the model representation of the true state of the simulated ecosystem (*Gregg et al. [2009]*). The DA (both reanalysis and forecasting) skill has to be evaluated using data-sets that are both statistically robust and at the same time reasonably independent of the assimilated EO data-set. For the 5 day chlorophyll forecasting skill we used the satellite OC-CCI data-set, since its robustness (number of data) seems to outweigh its inter-day correlation (dependence on the assimilated data). Although the dynamics of the satellite fields is slow (significant inter-day correlations between the same-pixel values), the rapid movement of atmospheric clouds means that the regions seen by the satellite in the successive days overlap by only 30% (we calculated this from the 2010 satellite data). We therefore considered the forecast validation EO dataset to be sufficiently independent of the assimilated data-set. The in situ observations are largely independent of the assimilated OC-CCI satellite data, but relatively sparse. The in situ chlorophyll measurements were used to evaluate the DA reanalysis skill (which is where the OC-CCI data-set cannot be used for validation, but just for verifying a correct implementation of the assimilation algorithm). This is relevant for the spatio-temporal regions with missing satellite EO data (such as cloudy regions, or regions below the $\sim$ 10 m surface layer measured by the satellite). Similarly to chlorophyll, we also used in situ data to evaluate the DA reanalysis skill to represent some of the relevant non-assimilated variables (such as nutrients, fCO_2). The DA reanalysis skill was considered sufficient for nutrients and fCO_2, because the impact of DA on the nutrient (or fCO_2) concentrations is slow compared to the short forecasting window. Consequently for non-assimilated variables there will be very little difference between DA reanalysis skill and DA forecasting skill. We confirmed this at the in situ locations by comparing the nutrient 5th forecasting day outputs with the reanalysis for the same day. The median difference for PFTs DA was of the order of 10^{-3} $mmol/m^3$ for silicate and phosphate; for nitrate the absolute value of the median difference was approximately 0.1 $mmol/m^3$.

To evaluate model skill we chose in situ and EO data only from the NWE Shelf. We matched both the EO and in situ spatio-temporal locations with the corresponding model data (i.e. the data closest in space and time). Both the EO and in situ data have different number of data points for different months. Furthermore the in situ (ICES and SOCAT data) spatial locations (geographic locations and depths) can vary substantially between months. We used two skill metrics: model bias and bias corrected median absolute difference. Under "bias" we mean median difference in model and EO (model minus EO) values. The biases were calculated for monthly binned data and the 2010 year bias was then taken to be the median of the monthly biases. The reason for binning data monthly was to correct for some of the spatio-temporal biases of the EO and in situ data. By "bias corrected median absolute difference" we mean median of absolute values of differences between model and EO, after subtracting the bias from the model outputs. This was again calculated for the monthly data (we subtracted monthly biases from absolute differences) with the annual value being the median of monthly values.

Both model and EO raw data can be (by definition) represented as a sum of climatology and anomalies from climatology. The model forecasting skill for both raw data and anomalies was also compared using a metric analogous to *Ryan et al. [2015]*:

$$F_S = 1 - \frac{\mathrm{AD}}{\mathrm{AD}_R}. \tag{1}$$

Here $\mathrm{AD}/\mathrm{AD}_R$ is the ratio between the annual median from monthly medians of absolute differences of the forecast and the reference outputs (both compared to the EO data). Positive values of F_S mean that forecast outperforms reference and vice versa. We considered here as reference the free run and persistence, where persistence means fixing the biogeochemical variables equal to the output of the reanalysis and using these constant values to forecast the biogeochemistry in the subsequent 1-5 days. AD from equation (1) is for raw data defined as $\mathrm{AD}_{raw} = Med(|\mathrm{Mod}_{raw} - \mathrm{EO}_{raw}|)$ and for anomalies as $\mathrm{AD}_{an} = Med(|\mathrm{Mod}_{an} - \mathrm{EO}_{an}|)$, ("$Med$" means median, "Mod" means "Model" and subscripts describe the type of data, with "an" standing for "anomaly"). Anomalies can be calculated by subtracting field climatology from the raw data. Twelve-year climatology was available only for the OC-CCI EO data-set. If we define the climatological model bias $B(x,t)$ as the difference between the model climatology (Mod_{clim}) and the climatology of the EO (EO_{clim}), the model climatology can be obtained as:

$$\mathrm{Mod}_{clim}(x,t) = \mathrm{EO}_{clim}(x,t) + B(x,t). \tag{2}$$

The bias $B(x,t)$ was estimated from the 2010 data as:

$$B(x,t) = \frac{B_A(x) + B_D(t)}{2}, \tag{3}$$

where $B_A(x)$ is annual median bias at the location x and $B_D(t)$ is spatial median bias on the NWE Shelf at the time t. The $B_A(x)$ and $B_D(t)$ functions were then calculated from the model and the EO 2010 data. The raw data and model bias are sufficient to calculate the AD_{an} value:

$$\mathrm{AD}_{an} = Med(|\mathrm{Mod}_{raw} - B - \mathrm{EO}_{raw}|), \tag{4}$$

and therefore they are sufficient to compute the anomaly forecast skill F_S.

Interpreting skill metrics (such as the one in equation (1)) needs some caution. The purpose of these skill metrics is merely to indicate: 1. whether reanalysis is closer to EO data than the reference run, 2. how rapidly forecast changes the match-ups between model outputs and EO data. The definition of these skill metrics assumes that the EO data represent the "true state". However, the EO data might also contain relatively large errors, although typically these errors are lower than the model errors. Therefore if DA moves

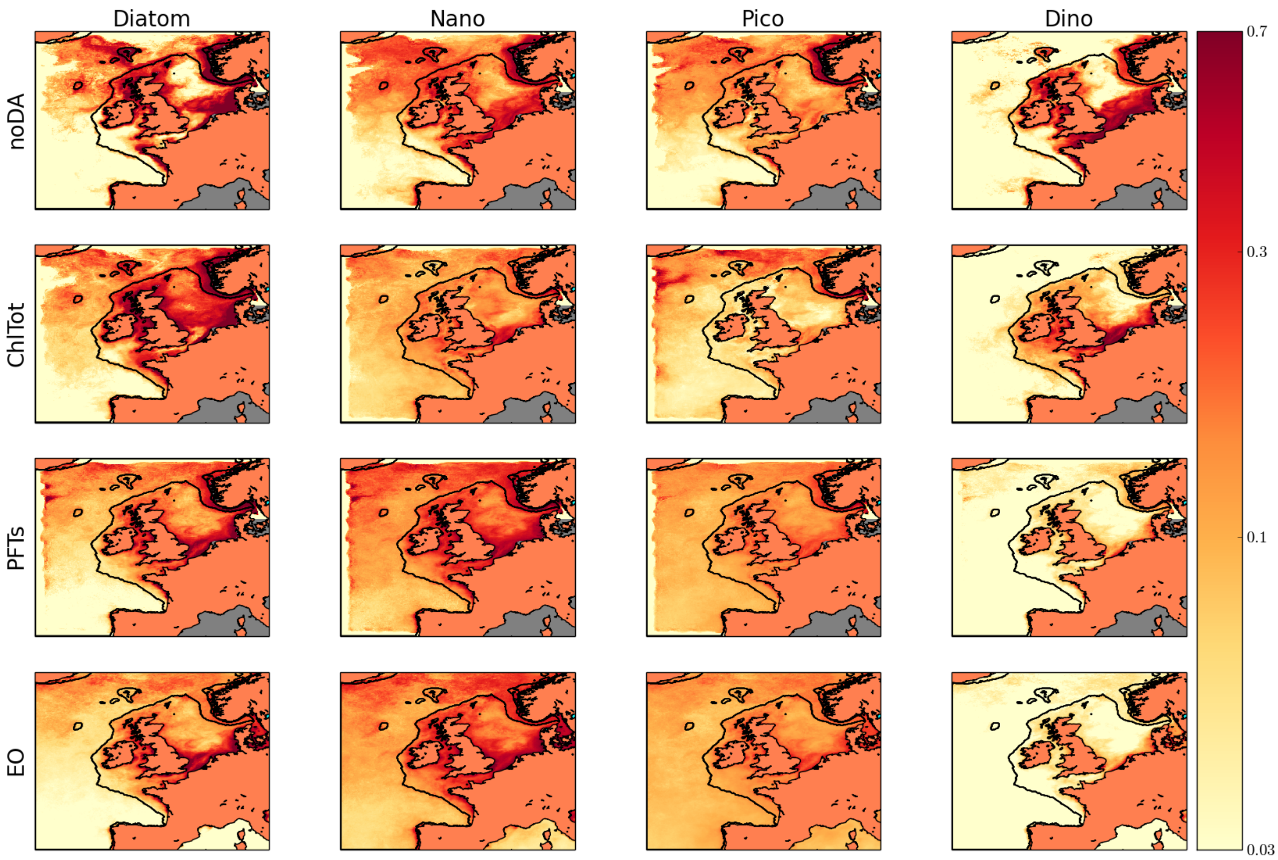


Figure 1. The Figure shows the 2010 annual median spatial distributions for the four PFTs chlorophyll (in mg/m^3) for the free run (first row), total chlorophyll DA (second row), PFTs DA (third row) and satellite EO data (fourth row). The shelf boundary (bathymetry < 200 m) is marked by the black line. The model data were masked whenever the EO data were missing.

the reference run outputs closer to the EO data, it typically moves it closer to the “true state" as well, but it can happen that a very close match with the EO data is not a very close match with the “true state". These metrics are therefore typically informative, but one has to keep in mind that there are situations in which they are misleading.

3 Results

3.1 Reanalysis

DA had a substantial impact on both PFTs and total chlorophyll distributions. In respect of the reference run, ChlTot DA does not improve the spatial match-ups with the EO PFTs chlorophyll. It does, however, substantially improve the match-ups with EO total chlorophyll. This can be seen in Figure 1, which shows the annual median chlorophyll distributions of the four phytoplankton functional types, and in Figure 2, which shows the same for the total chlorophyll. It is evident that the PFTs DA produced PFTs and to-

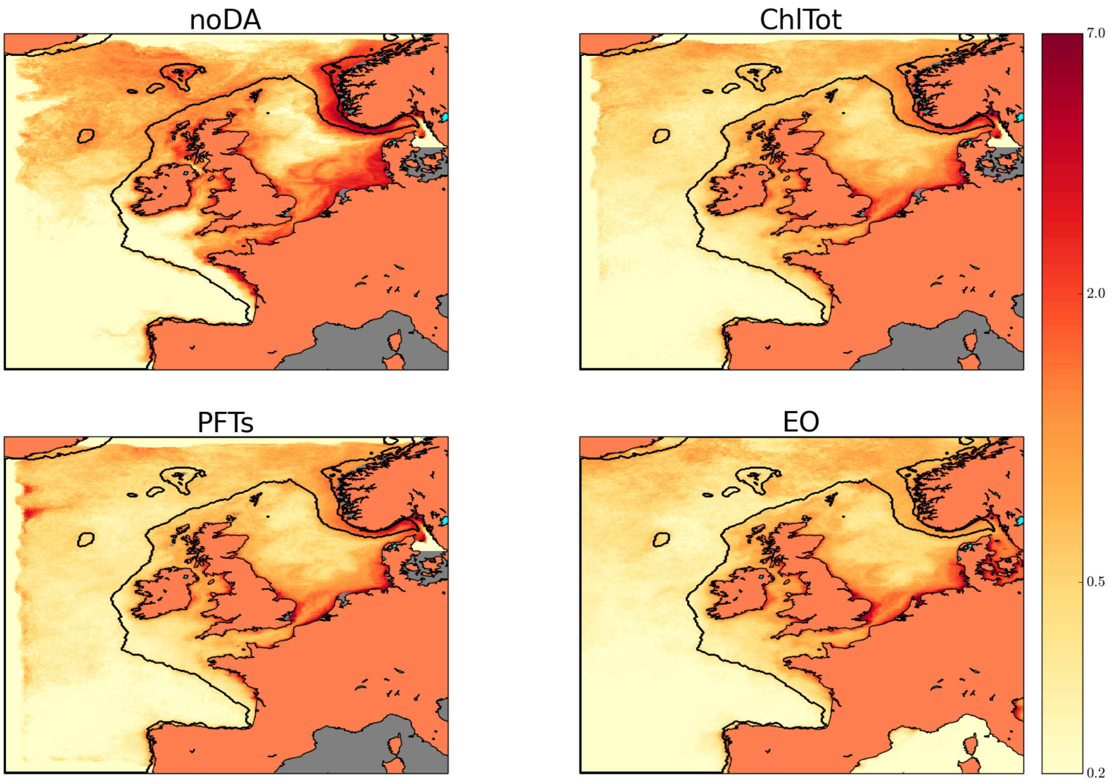
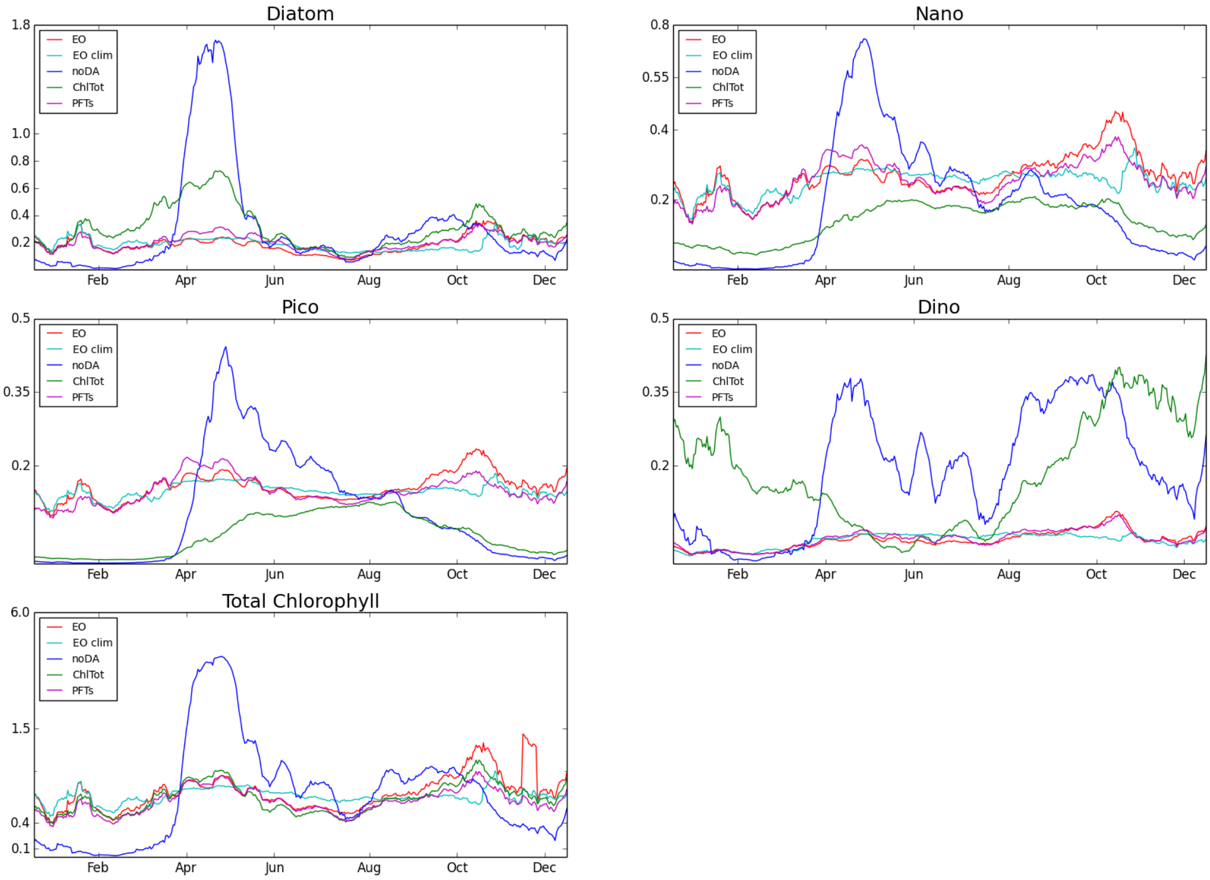


Figure 2. The Figure shows the 2010 annual median spatial distributions for the total chlorophyll (in mg/m^3) for the free run, total chlorophyll DA, PFTs DA and satellite EO data. The model data were masked whenever the EO data were missing.



372 **Figure 3.** The Figure compares daily time series of PFTs chlorophyll and total chlorophyll spatial median
373 values (in mg/m^3, for the NWE Shelf) for free run (noDA), ChlTot DA, PFTs DA, satellite EO data (EO) and
374 satellite EO data climatology (EO clim). The time series were smoothed on a 10 day time scale using moving
375 averages. The model data were masked wherever the EO data were missing.

tal chlorophyll distributions that look very similar to the EO satellite products (Figure 1 and Figure 2). The DA impact is largest in the Southern North Sea, which is the area with the largest chlorophyll concentrations. Figure 1 demonstrates the major impact of PFTs DA, especially on dinoflagellates where the difference between model and EO data is most significant. The improved match-up between the model output and the EO data (as one moves from the free run to DA in Figures 1 and 2) can be understood as a basic self-consistency test for the DA algorithms.

Figure 3 displays a daily time series for 2010 of spatial median PFT chlorophyll values (for the NWE Shelf). Figure 3 shows that bias between free run and EO data depends largely on the season. The model tends to underestimate PFTs chlorophyll in the Autumn and Winter, and greatly overestimate PFTs chlorophyll during Spring bloom and Summer (especially diatoms in Spring). This implies that the model has much larger seasonal variability than the EO data. Consistently with Figures 1 and 2, Figure 3 shows that: 1. The PFTs DA moves the annual time series very close to the EO data. The same is true for ChlTot DA and total chlorophyll time series. 2. The largest impact of PFTs DA is on dinoflagellates, where there is the poorest match between the model free run and the EO data. 3. ChlTot DA slightly improves the time series of nanophytoplankton and diatoms, however in Winter it considerably degrades dinoflagellates. 4. The model shows a dominant PFTs bloom in Spring (with huge concentrations of diatoms), whereas the EO PFTs data (and PFTs in PFT DA run) have an Autumn peak in chlorophyll concentrations. 5. Satellite EO data anomalies are relatively small when compared to the satellite monthly climatology.

The PFT chlorophyll-to-total chlorophyll ratios represent the composition of the phytoplankton community structure which can be seen as an emergent property of the ecosystem model and it can be used as a tool for model skill assessment (*De Mora et al. [2016]*). Figure 4 shows that PFTs DA improved the model representation of the plankton community structure (as represented by the assimilated data of *Brewin et al. [2017]*), when compared to both the model free run and the assimilation of total chlorophyll.

3.2 Forecasting

Figure 5 demonstrates model skill to predict the satellite EO observations for each PFT and total chlorophyll. For all PFTs, PFTs DA substantially outperforms both ChlTot DA and the free run over the whole 5 day forecasting period. The PFTs DA and ChlTot DA total chlorophyll have biases with opposite signs (except for the last forecasting day). The reason why there is difference between PFT and ChlTot DA total chlorophyll distributions is that, as previously mentioned, the bias corrected EO total chlorophyll concentrations assimilated in ChlTot DA are approximately $0.07 mg/m^3$ larger than the sum of bias corrected PFT chlorophyll EO assimilated in PFTs DA. Figure 5 further shows that the model (free run) accurately represents total chlorophyll levels (bias close to zero), however this hides large biases in PFTs concentrations (except for diatoms).

Figure 6 compares ChlTot DA and PFTs DA forecasting skill using the metric from equation (1), with the free run and the persistence as references. The upper row (plots A and B) shows model skill to predict the total and PFTs chlorophyll raw values (sum of climatology and anomaly). The bottom row (plots C and D) shows model skill to predict anomalies. In both cases (plots A and C) PFTs DA substantially outperforms the free run on the 5 day time scale (this is consistent with Figure 5). In the case of raw values (plot A) PFTs DA substantially outperforms ChlTot DA in PFTs chlorophyll and performs similarly than ChlTot DA in total chlorophyll forecasting. However, it is interesting that persistence outperforms the dynamical forecast from the PFTs DA on the 5 day forecasting time scale, which suggests that (PFTs) reanalysis plays an essential role in forecasting skill. The fact that persistence outperforms the model forecast simulation implies that the model degrades chlorophyll faster than the chlorophyll dynamics observed in the EO data.

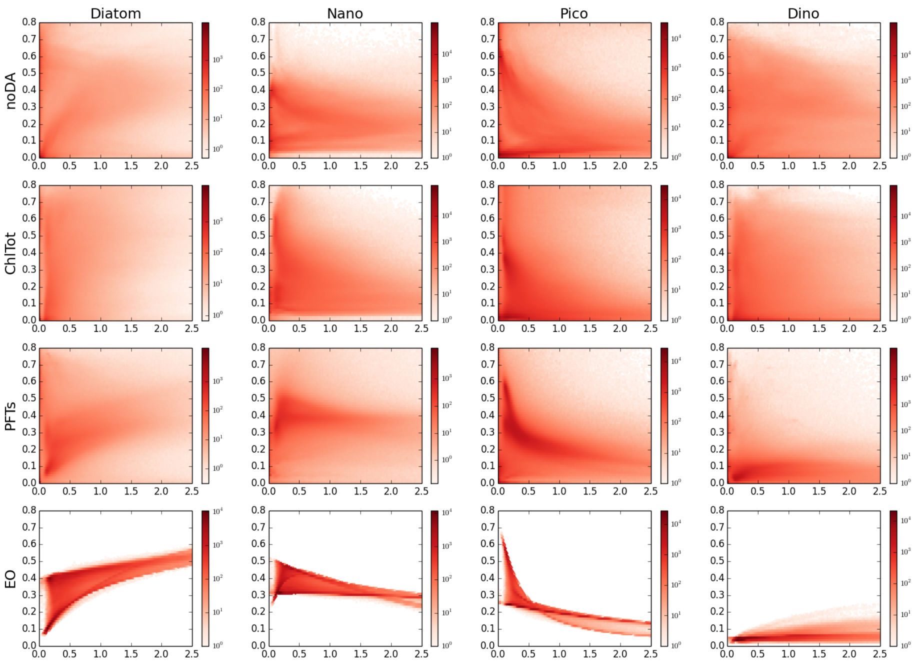


Figure 4. The Figure compares the 2010 PFTs to total chlorophyll ratios. The x-axis shows the total chlorophyll concentrations (in mg/m^3) and the y-axis shows PFT to total chlorophyll ratio. The EO data ratios are split based on the model of *Brewin et al. [2010, 2015]*. The shades of the red color mark the number of overlapping datapoints.

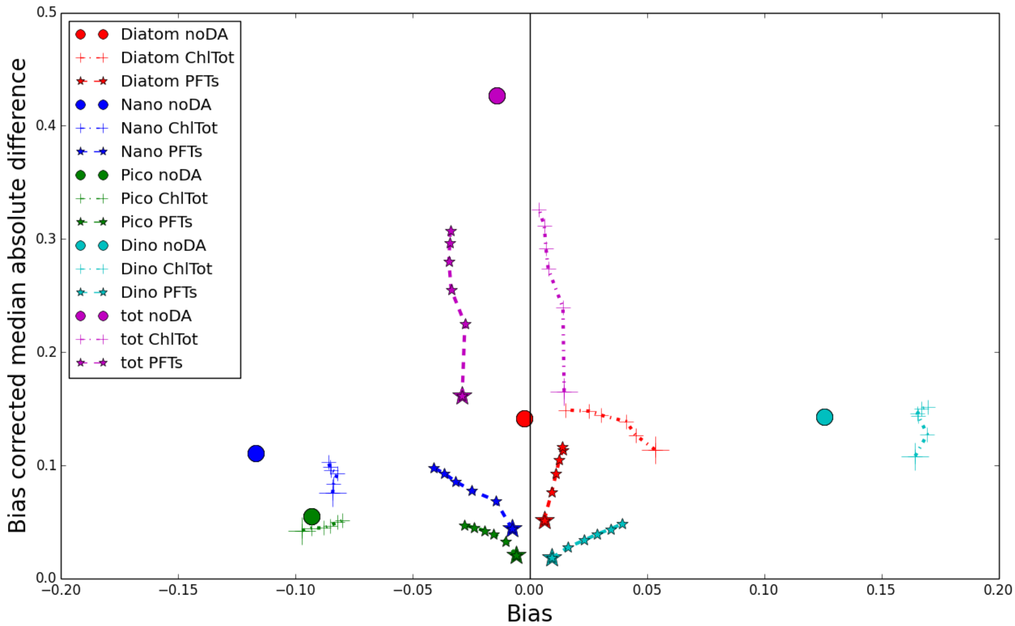


Figure 5. The Figure compares the reanalysis and forecasting of the assimilative runs with the reference for the data of the four PFTs (diatoms, dinoflagellates, nano-, picophytoplankton and of total chlorophyll). The bullet point is the free run, for the DA runs the first point on each line (with larger marker size) is reanalysis and the other five points are the five forecasting days. The x axis shows bias (in mg/m^3) and the y axis shows bias corrected median absolute difference (mg/m^3). The Figure shows that PFTs DA outperforms on the 5-day forecasting scale both free run and ChlTot DA in how it represents PFTs concentrations. From the lines on the plot one can see (for each PFT as well as total chlorophyll) that in the forecasting run model skill moves from the reanalysis skill towards the free run skill.

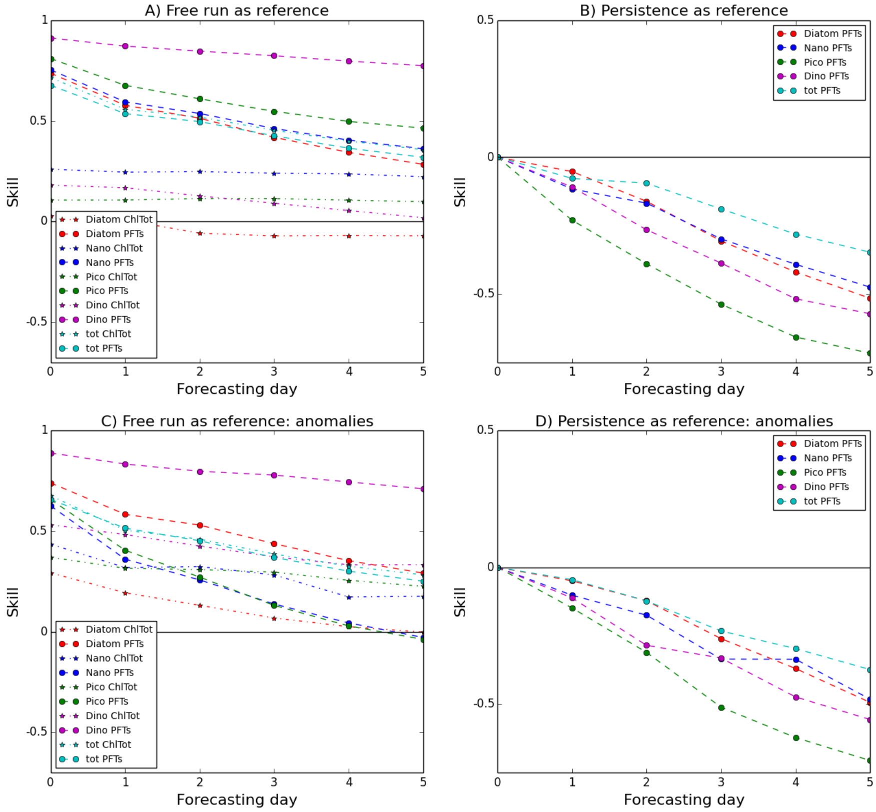


Figure 6. The Figure compares the reanalysis and 5 day forecasting skill (the first point on the line is reanalysis and the other five are the five forecasting days) using the skill metrics defined in equation (1). The left-hand plots (A and C) use as reference the free run and the right-hand plots (B and D) use persistence. The upper plots (A and B) are predictions of raw data (sum of climatology and anomalies) and the bottom plots (C and D) are only predictions of anomalies. Positive values mean that the evaluated model forecast outperforms the reference, whereas negative values mean that reference outperforms the model forecast.

3.3 Validation using in situ data

The validation using in situ data is summarized in Table 1 and Table 2. The two tables present annual values of the bias and the bias corrected absolute difference. Table 1 shows that for most of the year the model overestimates observed nitrate (the biases are almost 200% of in situ nitrate values). This is moderately improved by DA, where the bias decreases by 5% (of its value). The model overestimates observed silicate by approximately 50% and the bias can be reduced (ChlTot), or increased (PFTs) by the DA quite substantially (by about 40%). Unlike nitrate and silicate, the model has very low (positive) phosphate bias and even though this is to some extent degraded by the DA, the bias is always between $1 and 3.5$% of the observed value. Table 1 demonstrates that DA has substantial positive impact on the fCO_2 representation reducing the model negative bias by 50% (PFTs DA more than ChlTot DA). This reduces model relative error from 11.3% to 5.6% (PFTs DA).

Table 1. The annual bias (model minus in situ data) for the three nutrients (nitrate, phosphate and silicate) in $mmol/m^3$, CO_2 fugacity (fCO_2) in $\mu bars$ (SOCAT data), total chlorophyll (ICES data) and three phytoplankton size classes (Cefas data) in mg/m^3. The columns show free run, ChlTot DA and PFTs DA.

variable	noDA	ChlTot	PFTs
nitrate	8.82	8.4	8.65
phosphate	0.007	0.012	0.019
silicate	2.47	1.87	3.41
fCO2	-45.3	-28.7	-22.4
total chlorophyll	-0.2	-0.23	-0.35
microphytoplankton	-0.15	-0.14	-0.16
nanophytoplankton	-0.19	-0.14	-0.15
picophytoplankton	-0.06	-0.04	-0.05

The DA increases the negative bias of total chlorophyll with respect to the in situ data (Table 1). This can be explained by the larger (relative to the free run) negative bias of satellite data with respect to the in situ data (the satellite data are on average 0.45 mg/m^3 lower than the in situ data). This suggests the ICES and OC-CCI data-sets are not entirely consistent and the DA drives chlorophyll away from the in situ distributions. The evaluation of the impact of DA on the three phytoplankton size-classes using the in situ observations from the Cefas dataset is ambiguous (see Tables 1 and 2). In this case DA seems to improve the representation of both nanophytoplankton and picophytoplankton (in general ChlTot DA more than PFTs DA), but it increases the bias of microphytoplankton.

The L4 data (see Figure 7) demonstrate a very good total chlorophyll match between satellite and in situ data in Spring-Summer season (the annual mean absolute difference between in situ and satellite data was 0.4 mg/m^3, compared to the larger 0.7 mg/m^3 mean absolute difference between PFTs DA and the in situ data). In the same season there is a good match between satellite and in situ nanophytoplankton and dinoflagellates, but not a good consistency in diatoms and picophytoplankton (Figure 7). From

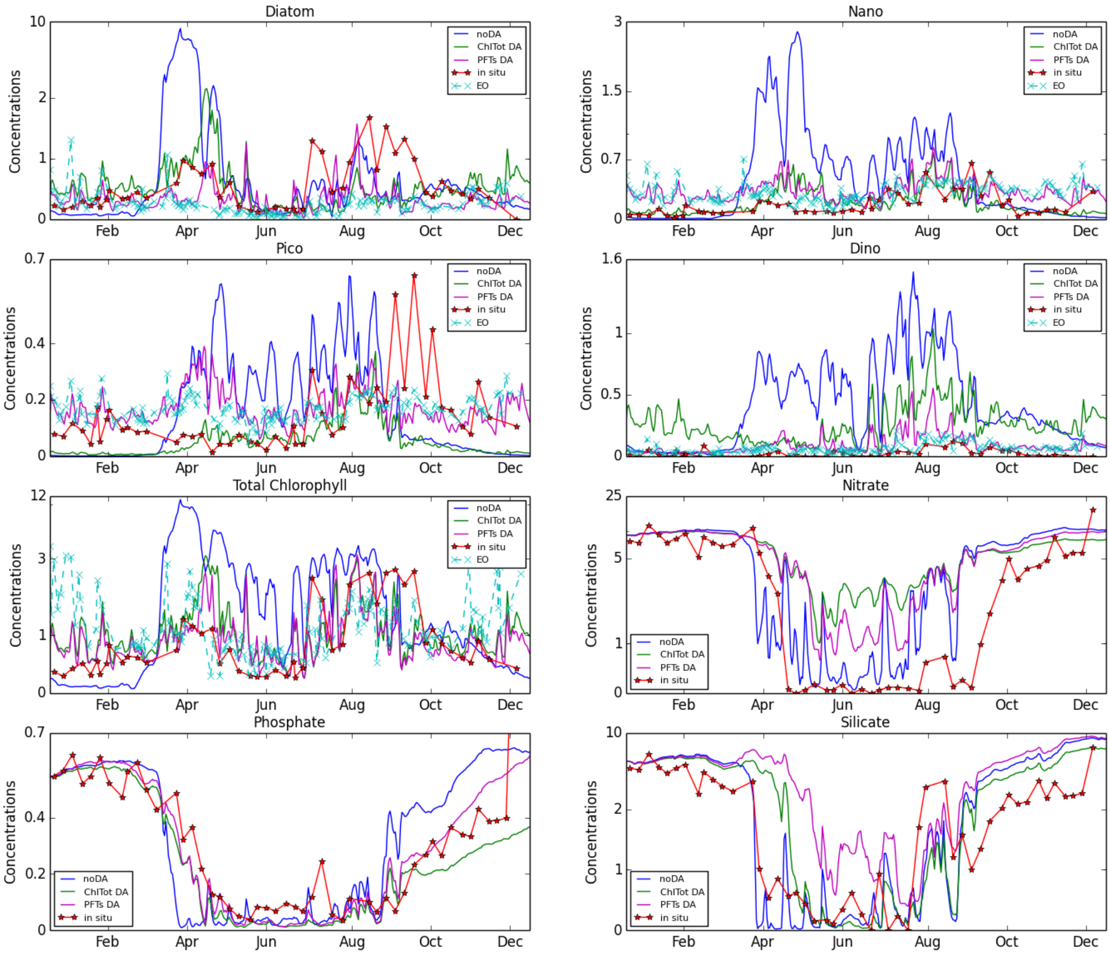


Figure 7. The Figure shows PFTs chlorophyll-a and nutrients (nitrate, phosphate and silicate) annual time series (noDA, ChlTot DA, PFTs DA and in situ data) at the L4 location in 2010. The nutrient concentrations are in $mmol/m^3$ and the chlorophyll concentrations in mg/m^3.

Table 2. The bias corrected median absolute difference for the three nutrients (nitrate, phosphate and silicate) in $mmol/m^3$, CO_2 fugacity (fCO_2) in $\mu bars$ (SOCAT data), total chlorophyll (ICES data) and three phytoplankton size classes (Cefas data) in mg/m^3. The columns show free run, ChlTot DA and PFTs DA.

variable	noDA	ChlTot	PFTs
nitrate	3.61	3.63	4.24
phosphate	0.13	0.13	0.13
silicate	2.27	2.19	1.97
fCO2	21.3	23.5	23.1
total chlorophyll	0.94	0.81	0.9
microphytoplankton	0.39	0.33	0.32
nanophytoplankton	0.18	0.16	0.17
picophytoplankton	0.07	0.06	0.05

Figure 7 one can draw similar conclusions as from Figure 3 (showing time evolution on the whole NWE Shelf): 1. The model overestimates Spring blooms and underestimates Autumn blooms of the PFTs. 2. There is less seasonal variability in the in situ data than in the model data. 3. PFTs DA drives the model PFTs chlorophyll-a towards the EO data. Since the EO data are much closer to the in situ data than the model, DA also improves the match up with the PFT and total chlorophyll in situ data. It is interesting that there are large similarities between the annual patterns of the satellite data time series on the whole NWE Shelf (Figure 3) and in situ data time series at L4, except that: 1. Satellite data have the bloom peak in Autumn slightly later (1 month). This discrepancy between EO and in situ data has been observed for the L4 site by *Smyth et al. [2009]*, but is not clearly visible for 2010 (Figure 7). It can be potentially explained by the L4 satellite data errors caused by terrestrial CDOM and sediments (*Smyth et al. [2009]; Groom et al. [2009]*). 2. The Autumn peak is more dominant at L4 for the situ data (see especially picophytoplankton in Figure 7). There is a good match between the model and the in situ nutrients at L4 (Figure 7), where the main difference seems to be that the nitrate and phosphate minima are phase-shifted in the model by roughly 1 month. The 2010 chlorophyll and nutrients in situ data have seasonal behavior similar to the L4 2004-2008 time series analysis from *Widdicombe et al. [2010]*. The L4 data also suggest that PFTs DA degrades silicate with respect to the reference run (the last panel in Figure 7).

4 Summary and discussion

This work demonstrates that both PFTs DA and ChlTot DA have substantial impact on the simulation of phytoplankton size-class chlorophyll, as well as of total chlorophyll distributions (Figures 1 - 4), when applied with an operational model in 5 day forecasting. Figures 1 - 3 demonstrate that the DA assimilated variables are very close to the EO satellite data. This is not because of large model-to-observational error ratio. The model errors used were typically around three times higher than observational errors, which is similar to the NEMO-HadOCC DA set-up (*Ford et al. [2012]*) and our own tests showed that the DA results have been relatively insensitive to the errors. We ran the same DA set-up (with the same background and observational errors), but without keeping the phytoplankton internal stochiometric ratios fixed (only phytoplankton chlorophyll was updated). The scheme still substantially improved the assimilated fields, however the final distributions were much further from the assimilated satellite data (see Supporting Information (1)). For example, the total chlorophyll bias (with respect to the EO) was nearly five times higher for ChlTot DA when DA changed the stochiometric ratios than when it did not. This is because the changed stochiometric ratios create internal imbalances and in the period between two assimilation steps these imbalances drive the assimilated state away from the EO. By preserving the model background stochiometry during the analysis update we stabilize the model dynamics and the DA gradually drives the analysis state close to the EO data.

Figure 3 shows that the model chlorophyll has distinctive maxima during the Spring bloom, whereas the EO data (and similarly the DA outputs) have lower seasonal variability with the maxima in the Autumn. The model bias has a seasonal signature (Figure 3), with the model underestimating EO chlorophyll values in the Autumn and Winter and overestimating them in the Spring-Summer period. We have shown that the EO satellite data seasonality is largely consistent with the in situ data seasonality in the L4 region (Figure 7, see also *Smyth et al. [2009]*). The DA impact on PFTs and total chlorophyll values is spatially most substantial in the Southern North Sea (Figures 1 - 2).

The model (free run) has a very small negative total chlorophyll bias, which hides much larger biases in PFTs concentrations (see Figure 5). This immediately points out the need to correct PFTs chlorophyll. ChlTot DA impacts PFTs chlorophyll, but it fails to improve the model skill in PFTs (Figure 5). This is because ChlTot DA redistributes the total chlorophyll-a increments into functional types using the model functional type-to-total chlorophyll ratio at a specific spatio-temporal point. Unlike ChlTot DA, PFTs DA substantially improves the model representation of PFTs chlorophyll. The forecasting run degrades the PFTs DA reanalysis bias and absolute differences by moving their values towards the values of the free run. However within the 5-day forecasting period PFTs DA always outperforms the free run (see Figure 5). PFTs DA increases the total chlorophyll negative bias of the free run (Figure 5). This is because the sum of bias corrected EO PFTs chlorophyll-a gives smaller values of total chlorophyll (for 2010 on average by 0.073 mg/m^3) than the bias corrected EO total chlorophyll (which is assimilated by ChlTot DA). The most substantial impact of PFTs DA is the large decrease in dinoflagellates concentrations. This is a consequence of a large mismatch in the EO and the model concentrations of dinoflagellates, mentioned already in *Ciavatta et al. [2018]*. Improving dinoflagellate estimates, their representation and their associated errors by both model and the satellite algorithms (*Brewin et al. [2017]*), is a major challenge which needs to be addressed in the future.

Similarly to Figure 5, Figure 6 shows that the PFTs DA substantially improves the model 5 day forecasting skill (on the NWE Shelf) for all the phytoplankton functional types, as well as for the total phytoplankton chlorophyll-a. Plot A in Figure 6 shows that PFTs DA outperforms both ChlTot DA and the free run in forecasting the raw data (sum of climatology and anomaly) of all the PFTs chlorophyll within the 5 day forecasting period. The PFTs DA and ChlTot DA total chlorophyll forecasting skills are comparable. Surface chlorophyll has relatively small anomalies compared to the chlorophyll monthly climatology (see Figure 3). This means most of the model skill in forecasting the raw data (see Figure 6 plot A) depends on its skill to represent the PFTs chlorophyll climatology. However, PFTs DA also outperforms the free run for all the assimilated variables in forecasting the anomalies (see Figure 6, plot C). The comparison with the skill of the persistence has negative values (see Figure 6, plots B and D) for both raw data and anomalies, which means the PFTs DA forecast skill mostly originates from the persistence of the reanalysis. Negative persistence skill means it is more useful to predict future chlorophyll distributions by assuming the status quo (based on reanalysis), than running the model. This might be a consequence of the fact that the univariate DA scheme changes phytoplankton concentrations, while keeping the other variables (especially nutrients) intact. The model is therefore "off-balance" and the forecasting simulation moves away from

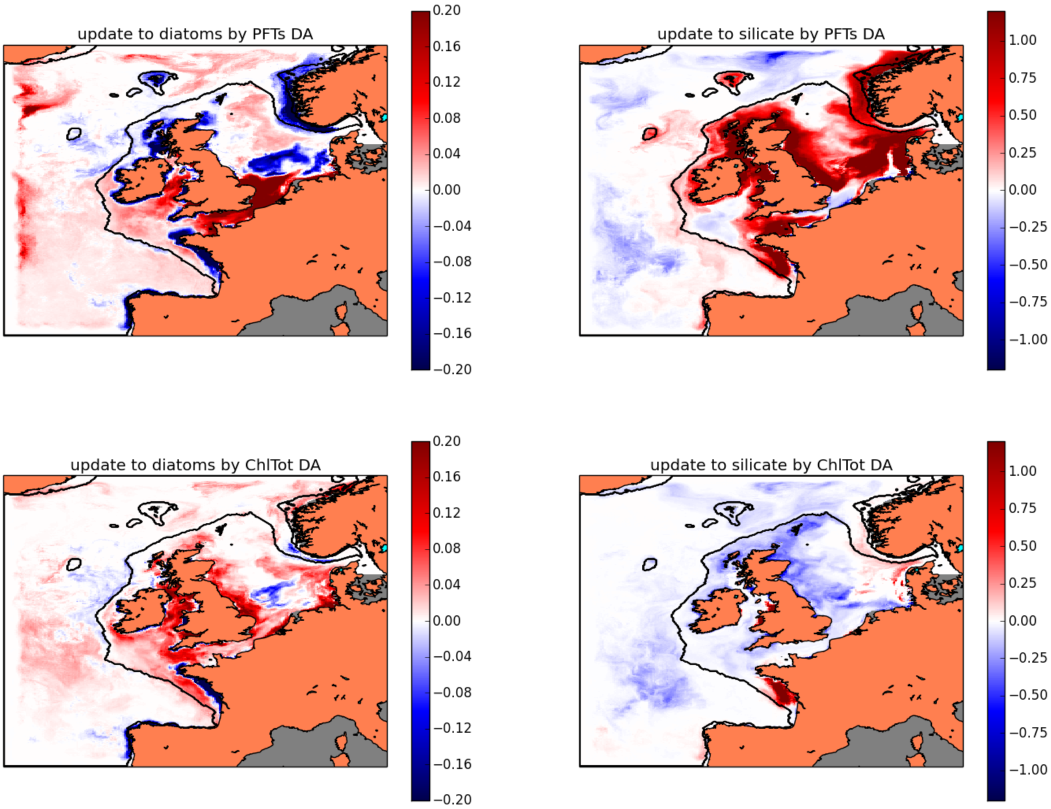


Figure 8. The DA updates to the diatom (mg/m^3) and silicate ($mmol/m^3$) concentrations. The Figure shows (upper panels) the annual spatial median concentration of the PFTs DA minus the free run, and the same differences between ChlTot DA and the free run (bottom panels). In most of the regions the updates to silicate are visibly anti-correlated with the updates to diatoms.

the reanalysis state faster than the chlorophyll dynamics. The model simulation degrades fields slowly compared to the reference run skill (as discussed before), however it still degrades them faster than the observed field dynamics (at least within the 5 day forecasting period). To conclude, the reanalysis can be a better predictor of the 5-day future state than the reinitialized model simulation. However, both the reanalysis and the 5-day forecast substantially outperform the skill of the reference simulation. This proves that using PFTs DA for operational applications is of substantial value.

The most regularly distributed validation in situ data with the largest statistical significance were fCO_2 SOCAT data (around 10000 data-points). The comparison with the SOCAT data has shown that the model underestimates CO_2 fugacity (having 11.3% lower value than in situ data). The DA has a large positive impact on CO_2 fugacity and improves the CO_2 fugacity bias by more than 50% (more PFTs DA than ChlTot DA). It is possible that this is because correcting phytoplankton biomass has an impact on the primary production and consequently affects the model representation of the carbon cycle (e.g *Ciavatta et al. [2018]*). Based on the ICES data it was shown (see Table 1) that the model typically overestimates nutrients, in particular it overestimates nitrate by almost 200%. The DA moderately lowers nitrate bias by 5%. Given the spatio-temporal biases of the in situ data it is hard to estimate the confidence intervals, but a simplified analysis based on calculating the 95% confidence interval for a representative sample of the same size (than the size of the nitrate ICES data-set) suggested the effect of DA on nitrate is not statistically significant. The same is true for phosphate, where the model bias fluctuates between 1.3-3.5% of the phosphate value. This means that model represents phosphate levels with a very good accuracy, possibly within the systematic error of the measurements. (Note that Table 2 suggests that the model does not represent equally well phosphate spatio-temporal distributions.) Interestingly the ChlTot DA and PFTs DA have very different impact on silicate. The model free run overestimates silicate values by roughly 50%. The bias is substantially improved by ChlTot DA (lowered by 25%), but degraded by PFTs DA (increased by 40%). Since diatoms are silicate users, the impact of DA on silicate is mainly related to how DA changes the concentrations of diatoms. We calculated the differences in diatoms concentrations between each of the assimilative runs (i.e PFTs DA and ChlTot DA) and the free run at the in situ data locations. At the same locations we calculated the same differences in silicate concentrations. The impact of DA on diatoms was anti-correlated with its impact on silicate at the in situ locations, with Spearman coefficients equal to -0.44 (Pfts DA) and -0.27 (ChlTot DA). Since there were around 1300 in situ data-points the result is statistically significant, with p values less than 10^{-20}. The silicate and diatoms are anti-correlated because diatoms are controlling the concentration of silicate (top-down control). Figure 8 shows that ChlTot DA substantially increases concentrations of diatoms (see also Figure 5) and the increased concentrations of diatoms then take up more silicate and lower its concentrations. The model dynamics in response to PFTs DA increased the silicate bias at the in situ locations because it lowered diatoms concentrations on those sites (-0.44 Spearman coefficient). However, Figures 5 and 8 show that overall PFTs DA did not lower the diatoms concentrations on the NWE shelf. This suggests that the increase in silicate bias by PFTs DA could be specific to the in situ spatio-temporal locations. However, this still points out an issue of the model. The model is overestimating silicate (Table 1), while it is representing accurately the levels of diatoms (see Figure 5). Under such conditions the model representation of silicate cannot be improved by correcting diatoms. There is a reason other than diatoms for why the model overestimates silicate and the problem needs to be better understood in the future.

Perhaps unexpectedly, the in situ ICES data showed that DA increases the total chlorophyll bias (more substantial for PFTs DA than for ChlTot DA). The effective overlap between the in situ total chlorophyll data and the OC-CCI EO data (considered up to the optical depth of 10 m) was roughly 20% (however, over 50% in situ measurements were from less than 10 m deep). The observed match-ups (Table 1) between satellite and in situ total chlorophyll have shown that the satellite data have negative bias with respect to the in situ data; in situ data are larger by 0.45 mg/m^3, on average. This suggests that the two total chlorophyll datasets are not entirely consistent. This is quite possibly a consequence of the spatio-temporal difference between the highly localized in situ measurements and the 7-km resolution of the EO composites. The larger EO negative bias towards in situ data ($-0.45\ mg/m^3$) then possibly degraded the relatively smaller negative model bias ($-0.2mg/m^3$) towards in situ data.

The comparison of model phytoplankton functional types pigments with Cefas (in situ) data-set was inconclusive (Table 1 and Table 2). The model concentrations showed negative biases with respect to in situ data (consistent with the total chlorophyll ICES data-set), but no clear impact of DA on model skill was observed. However, it needs to be emphasized that the Cefas in situ data-set had only 56 relevant data-points and it only contained relevant data from August 2010. The evidence is therefore too limited to justify any broader conclusions.

We also analysed DA skill in the specific L4 location. There was a good match during Spring-Summer period between in situ data and the EO for total chlorophyll, chlorophyll in nanophytoplankton and dinoflagellates (see Figure 7). The comparison with satellite data showed a worse match for chlorophyll in picophytoplankton and diatoms. The 2010 in situ time series presented in Figure 7 have more similarity with the median NWE Shelf EO time series (see Figure 3) than with the L4 satellite data. This could be explained by large satellite errors at the L4 location (especially in the Autumn-Winter season). Interestingly, Figure 7 shows that in the L4 location the model represents nutrients with no significant biases. The main difference between model and in situ nutrient data is a 1 month shift in the seasonal dynamics (for nitrate and phosphate). This is probably linked to the large Spring bloom in the model time series. Interestingly also L4 data suggest that the PFTs DA degrades to some degree silicate (the bottom left panel of Figure 7).

5 Concluding remarks

This work shows that assimilating PFTs chlorophyll substantially improves operational model forecasting on the NWE Shelf. The model represents accurately the total chlorophyll levels. However, the small total chlorophyll bias hides large biases in PFTs chlorophyll, which cannot be corrected through ChlTot DA. The representation of PFTs chlorophyll is substantially improved by PFTs DA. The PFTs DA reanalysis skill is degraded by the forecasting run, but it remains much better than the skill of the free run within the 5-day forecast period. DA substantially improves representation of pCO2. It does not have significant impact on nutrients, but work is being carried out on developing a suitable multivariate balancing algorithm between phytoplankton functional types and the ERSEM variables of interest. Such a balancing scheme is expected to improve the coherence between phytoplankton biomass and dissolved nutrient concentrations to further slow down the model skill deterioration in the forecasting run.

Despite the advantages of the method, we stress that further research is needed to improve the understanding and representation of plankton functional types and related biogeochemical process in marine models (*Shimoda and Arhonditsis [2016]*). For example, our application does not account for calcification within the nano-plankton group (e.g coccolitophores), or mixotrophy by dinoflagellates, which are certainly relevant processes in the North Atlantic (e.g. *Gregg and Casey [2004]*), but remain open challenges in current operational models (e.g. *Anderson* [2005]; *Flynn et al.* [2012]; *Yool et al.* [2013]; *Aumont et al.* [2015]).

Acknowledgments

This work was funded by the Copernicus Marine Environment Monitoring Service (CMEMS) project Towards Operational Size-Class Chlorophyll Assimilation (TOSCA) and NOWMAPS. CMEMS is implemented by MERCATOR OCEAN in the framework of a delegation agreement with the European Union. JS, RB and SC were also supported by the UK NERC National Centre for Earth Observation (NCEO) and JS and SC by the UK NERC through the projects CAMPUS and ABC. This work also contributes to the project TAPAS, which received funding from the European Union Horizon 2020 research and innovation programme under grant agreement No 678396. We thank the European Space Agency Climate Initiative “Ocean Color” (http://www.esa-oceancolor-cci.org) for providing the ocean color data. We acknowledge use of the MONSooN system, a collaborative facility supplied under the Joint Weather and Climate Research Programme, a strategic partnership between the Met Office and the Natural Environment Research Council. The in situ measured data can be freely downloaded as follows: for nutrients and total chlorophyll (ICES data-set) from http://ices.dk/marine-data, for CO_2 fugacity (SOCAT data-set) from http://www.socat.info/ and for phytoplankton pigments (Cefas data-set) from https://www.cefas.co.uk/, http://doi.org/10.14466/CefasDataHub.33. The L4 data-set can be downloaded from the Western Channel Observatory as https://www.bodc.ac.uk/data/. The Surface Ocean CO2 Atlas (SOCAT) is an international effort, endorsed by the International Ocean Carbon Coordination Project (IOCCP), the Surface Ocean Lower Atmosphere Study (SOLAS) and the Integrated Marine Biosphere Research (IMBeR) program, to deliver a uniformly quality-controlled surface ocean CO2 database. The many researchers and funding agencies responsible for the collection of data and quality control are thanked for their contributions to SOCAT.

References

Airs, R., and V. Martinez-Vicente (2014), Amt18 (jr20081003) hplc pigment measurements from ctd bottle samples, *British Oceanographic Data CentreâĂŤNatural Environment Research Council, UK (doi: 10/tk2)*.

Allen, J., and P. Somerfield (2009), A multivariate approach to model skill assessment, *Journal of Marine Systems*, *76*(1), 83–94.

Anderson, T. R. (2005), Plankton functional type modelling: running before we can walk?, *Journal of Plankton Research*, *27*(11), 1073–1081.

Artioli, Y., J. C. Blackford, M. Butenschön, J. T. Holt, S. L. Wakelin, H. Thomas, A. V. Borges, and J. I. Allen (2012), The carbonate system in the north sea: Sensitivity and model validation, *Journal of Marine Systems*, *102*, 1–13.

Artioli, Y., J. C. Blackford, G. Nondal, R. Bellerby, S. L. Wakelin, J. T. Holt, M. Butenschön, and J. I. Allen (2014), Heterogeneity of impacts of high co2 on the north western european shelf.

Aumont, O., C. Ethé, A. Tagliabue, L. Bopp, and M. Gehlen (2015), Pisces-v2: an ocean biogeochemical model for carbon and ecosystem studies, geosci. model dev., 8, 2465–2513, doi: 10.5194.

Bakker, D., B. Pfeil, K. Smith, S. Hankin, A. Olsen, S. Alin, C. Cosca, S. Harasawa, A. Kozyr, Y. Nojiri, et al. (2014), An update to the surface ocean co2 atlas (socat version 2), earth syst. sci. data, 6, 69–90.

Baretta, J., W. Ebenhöh, and P. Ruardij (1995), The european regional seas ecosystem model, a complex marine ecosystem model, *Netherlands Journal of Sea Research*, *33*(3-4), 233–246.

Baretta-Bekker, J., J. Baretta, and W. Ebenhöh (1997), Microbial dynamics in the marine ecosystem model ersem ii with decoupled carbon assimilation and nutrient uptake, *Journal of Sea Research*, *38*(3-4), 195–211.

Blackford, J. (1997), An analysis of benthic biological dynamics in a north sea ecosystem model, *Journal of Sea Research*, *38*(3-4), 213–230.

Blackford, J., and F. Gilbert (2007), ph variability and co 2 induced acidification in the north sea, *Journal of Marine Systems*, *64*(1), 229–241.

Blauw, A., F. Los, J. Huisman, and L. Peperzak (2010), Nuisance foam events and phaeocystis globosa blooms in dutch coastal waters analyzed with fuzzy logic, *Journal of Marine Systems*, *83*(3), 115–126.

Bloom, S., L. Takacs, A. Da Silva, and D. Ledvina (1996), Data assimilation using incremental analysis updates, *Monthly Weather Review*, *124*(6), 1256–1271.

Borges, A., L.-S. Schiettecatte, G. Abril, B. Delille, and F. Gazeau (2006), Carbon dioxide in european coastal waters, *Estuarine, Coastal and Shelf Science*, *70*(3), 375–387.

Brandsma, J., J. M. Martínez, H. A. Slagter, C. Evans, and C. P. Brussaard (2013), Microbial biogeography of the north sea during summer, *Biogeochemistry*, *113*(1-3), 119–136.

Brewin, R. J., S. Sathyendranath, T. Hirata, S. J. Lavender, R. M. Barciela, and N. J. Hardman-Mountford (2010), A three-component model of phytoplankton size class for the atlantic ocean, *Ecological Modelling*, *221*(11), 1472–1483.

Brewin, R. J., S. Sathyendranath, D. Müller, C. Brockmann, P.-Y. Deschamps, E. Devred, R. Doerffer, N. Fomferra, B. Franz, M. Grant, et al. (2015), The ocean colour climate change initiative: Iii. a round-robin comparison on in-water bio-optical algorithms, *Remote Sensing of Environment*, *162*, 271–294.

Brewin, R. J., S. Ciavatta, S. Sathyendranath, T. Jackson, G. Tilstone, K. Curran, R. L. Airs, D. Cummings, V. Brotas, E. Organelli, et al. (2017), Uncertainty in ocean-color estimates of chlorophyll for phytoplankton groups, *Frontiers in Marine Science*, *4*, 104.

Butenschön, M., J. Clark, J. N. Aldridge, J. I. Allen, Y. Artioli, J. Blackford, J. Bruggeman, P. Cazenave, S. Ciavatta, S. Kay, et al. (2016), Ersem 15.06: a generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels, *Geoscientific Model Development*, *9*(4), 1293.

Campbell, J. W. (1995), The lognormal distribution as a model for bio-optical variability in the sea, *Journal of Geophysical Research: Oceans*, *100*(C7), 13,237–13,254.

Carmillet, V., J.-M. Brankart, P. Brasseur, H. Drange, G. Evensen, and J. Verron (2001), A singular evolutive extended kalman filter to assimilate ocean color data in a coupled physical–biochemical model of the north atlantic ocean, *Ocean Modelling*, *3*(3), 167–192.

Chassot, E., F. Mélin, O. Le Pape, and D. Gascuel (2007), Bottom-up control regulates fisheries production at the scale of eco-regions in european seas, *Marine Ecology Progress Series*, *343*, 45–55.

Ciavatta, S., R. J. W. Brewin, J. Skákala, L. Polimene, L. de Mora, Y. Artioli, and J. I. Allen (2018), Assimilation of ocean-color plankton functional types to improve marine ecosystem simulations, *Journal of Geophysical Research: Oceans*, pp. n/a–n/a, doi: 10.1002/2017JC013490.

Ciavatta, S., R. Torres, S. Saux-Picart, and J. I. Allen (2011), Can ocean color assimilation improve biogeochemical hindcasts in shelf seas?, *Journal of Geophysical Research: Oceans*, *116*(C12).

Ciavatta, S., R. Torres, V. Martinez-Vicente, T. Smyth, G. Dall'Olmo, L. Polimene, and J. I. Allen (2014), Assimilation of remotely-sensed optical properties to improve marine biogeochemistry modelling, *Progress in Oceanography*, *127*, 74–95.

Ciavatta, S., S. Kay, S. Saux-Picart, M. Butenschön, and J. Allen (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.

Cummings, J., L. Bertino, P. Brasseur, I. Fukumori, M. Kamachi, M. J. Martin, K. Mogensen, P. Oke, C. E. Testut, J. Verron, et al. (2009), Ocean data assimilation systems for godae, *Oceanography*, *22*(3), 96–109.

De Mora, L., M. Butenschön, and J. Allen (2013), How should sparse marine in situ measurements be compared to a continuous model: an example, *Geoscientific Model Development*, *6*(2), 533.

De Mora, L., M. Butenschön, and J. Allen (2016), The assessment of a global marine ecosystem model on the basis of emergent properties and ecosystem function: a case study with ersem, *Geoscientific Model Development*, *9*(1), 59.

Dee, D. P., S. Uppala, A. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. Balmaseda, G. Balsamo, P. Bauer, et al. (2011), The era-interim reanalysis: Configuration and performance of the data assimilation system, *Quarterly Journal of the royal meteorological society*, *137*(656), 553–597.

Edwards, C. A., A. M. Moore, I. Hoteit, and B. D. Cornuelle (2015), Regional ocean data assimilation, *Annual review of marine science*, *7*, 21–42.

Edwards, K., R. Barciela, and M. Butenschon (2012), Validation of the nemo-ersem operational ecosystem model for the north west european continental shelf, *Ocean Science*, *8*, 983–1000.

Finkel, Z. V., J. Beardall, K. J. Flynn, A. Quigg, T. A. V. Rees, and J. A. Raven (2009), Phytoplankton in a changing world: cell size and elemental stoichiometry, *Journal of plankton research*, *32*(1), 119–137.

Flynn, K. J., D. K. Stoecker, A. Mitra, J. A. Raven, P. M. Glibert, P. J. Hansen, E. Granéli, and J. M. Burkholder (2012), Misuse of the phytoplankton–zooplankton dichotomy: the need to assign organisms as mixotrophs within plankton functional types, *Journal of Plankton Research*, *35*(1), 3–11.

Fontana, C., C. Grenz, and C. Pinazo (2010), Sequential assimilation of a year-long time-series of seawifs chlorophyll data into a 3d biogeochemical model on the french mediterranean coast, *Continental Shelf Research*, *30*(16), 1761–1771.

Ford, D., and R. Barciela (2017), Global marine biogeochemical reanalyses assimilating two different sets of merged ocean colour products, *Remote Sensing of Environment*.

Ford, D., K. Edwards, D. Lea, R. Barciela, M. Martin, and J. Demaria (2012), Assimilating globcolour ocean colour data into a pre-operational physical-biogeochemical model, *Ocean Science*, *8*(5), 751–771.

Ford, D. A., J. van der Molen, K. Hyder, J. Bacon, R. Barciela, V. Creach, R. McEwan, P. Ruardij, and R. Forster (2017), Observing and modelling phytoplankton community structure in the north sea, *Biogeosciences*, *14*(6), 1419.

Garcia, H., R. Locarnini, T. Boyer, J. Antonov, O. Baranova, M. Zweng, J. Reagan, and D. Johnson (2014), World ocean atlas 2013, volume 4: Dissolved inorganic nutrients (phosphate, nitrate, silicate), noaa atlas nesdis, vol. 76, edited by s. levitus, 25 pp, *US Gov. Print. Off., Washington, DC*.

Gehlen, M., R. Barciela, L. Bertino, P. Brasseur, M. Butenschön, F. Chai, A. Crise, Y. Drillet, D. Ford, D. Lavoie, et al. (2015), Building the capacity for forecasting marine biogeochemistry and ecosystems: recent advances and future developments, *Journal of Operational Oceanography*, *8*(sup1), s168–s187.

Geider, R., H. MacIntyre, and T. Kana (1997), Dynamic model of phytoplankton growth and acclimation: responses of the balanced growth rate and the chlorophyll a: carbon ratio to light, nutrient-limitation and temperature, *Marine Ecology Progress Series*, pp. 187–200.

Gregg, W. W. (2008), Assimilation of seawifs ocean chlorophyll data into a three-dimensional global ocean model, *Journal of Marine Systems*, *69*(3), 205–225.

Gregg, W. W., and N. W. Casey (2004), Global and regional evaluation of the seawifs chlorophyll data set, *Remote Sensing of Environment*, *93*(4), 463–479.

Gregg, W. W., and C. S. Rousseaux (2017), Simulating pace global ocean radiances, *Frontiers in Marine Science*, *4*, 60.

Gregg, W. W., M. A. Friedrichs, A. R. Robinson, K. A. Rose, R. Schlitzer, K. R. Thompson, and S. C. Doney (2009), Skill assessment in ocean biological data assimilation, *Journal of Marine Systems*, *76*(1-2), 16–33.

Groom, S., V. Martinez-Vicente, J. Fishwick, G. Tilstone, G. Moore, T. Smyth, and D. Harbour (2009), The western english channel observatory: Optical characteristics of station l4, *Journal of Marine Systems*, *77*(3), 278–295.

Holt, J., M. Butenschon, S. Wakelin, Y. Artioli, and J. Allen (2012), Oceanic controls on the primary production of the northwest european continental shelf: model experiments under recent past conditions and a potential future scenario, *Biogeosciences*, *9*, 97–117.

Hoteit, I., G. Triantafyllou, and G. Petihakis (2005), Efficient data assimilation into a complex, 3-d physical-biogeochemical model using partially-local kalman filters, in *Annales Geophysicae*, vol. 23, pp. 3171–3185.

Ishizaka, J. (1990), Coupling of coastal zone color scanner data to a physical-biological model of the southeastern us continental shelf ecosystem: 1. czcs data description and lagrangian particle tracing experiments, *Journal of Geophysical Research: Oceans*, *95*(C11), 20,167–20,181.

Jackson, T., S. Sathyendranath, and F. Mélin (2017), An improved optical classification scheme for the ocean colour essential climate variable and its applications, *Remote Sensing of Environment*.

Jahnke, R. A. (2010), Global synthesis, *Carbon and nutrient fluxes in continental margins*, pp. 597–615.

Janjić, T., N. Bormann, M. Bocquet, J. Carton, S. Cohn, S. Dance, S. Losa, N. Nichols, R. Potthast, J. Waller, et al. (2017), On the representation error in data assimilation, *Quarterly Journal of the Royal Meteorological Society*.

Jones, E. M., M. E. Baird, M. Mongin, J. Parslow, J. Skerratt, J. Lovell, N. Margvelashvili, R. J. Matear, K. Wild-Allen, B. Robson, et al. (2016), Use of remote-sensing reflectance to constrain a data assimilating marine biogeochemical model of the great barrier reef, *Biogeosciences*, *13*(23), 6441.

Kalnay, E. (2003), *Atmospheric modeling, data assimilation and predictability*, Chapter 5: Data Assimilation, Cambridge university press.

Key, R. M., A. Olsen, S. van Heuven, S. K. Lauvset, A. Velo, X. Lin, C. Schirnick, A. Kozyr, T. Tanhua, M. Hoppema, et al. (2015), Global ocean data analysis project, version 2 (glodapv2).

Kurekin, A., P. Miller, and H. Van der Woerd (2014), Satellite discrimination of karenia mikimotoi and phaeocystis harmful algal blooms in european coastal waters: Merged classification of ocean colour data, *Harmful Algae*, *31*, 163–176.

Lauvset, S. K., R. M. Key, A. Olsen, S. van Heuven, A. Velo, X. Lin, C. Schirnick, A. Kozyr, T. Tanhua, M. Hoppema, et al. (2016), A new global interior ocean mapped climatology: the 1°x1° glodap version 2, *Earth System Science Data Discussions*, pp. 1–30.

Losa, S. N., G. A. Kivman, and V. A. Ryabchenko (2004), Weak constraint parameter estimation for a simple ocean ecosystem model: what can we learn about the model and data?, *Journal of marine systems*, *45*(1-2), 1–20.

MacLachlan, C., A. Arribas, K. Peterson, A. Maidens, D. Fereday, A. Scaife, M. Gordon, M. Vellinga, A. Williams, R. Comer, et al. (2015), Global seasonal forecast system version 5 (glosea5): a high-resolution seasonal forecast system, *Quarterly Journal of the Royal Meteorological Society*, *141*(689), 1072–1084.

Madec, G., et al. (2015), Nemo ocean engine.

Mogensen, K., M. Balmaseda, A. Weaver, M. Martin, and A. Vidard (2009), Nemovar: A variational data assimilation system for the nemo ocean model, *ECMWF newsletter*, *120*, 17–22.

Mogensen, K., M. A. Balmaseda, and A. Weaver (2012), *The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4*, European Centre for Medium-Range Weather Forecasts.

Moore, T. S., J. W. Campbell, and M. D. Dowell (2009), A class-based approach to characterizing and mapping the uncertainty of the modis ocean chlorophyll product, *Remote Sensing of Environment*, *113*(11), 2424–2430.

Natvik, L.-J., and G. Evensen (2003), Assimilation of ocean colour data into a biochemical model of the north atlantic: Part 1. data assimilation experiments, *Journal of Marine Systems*, *40*, 127–153.

Nerger, L., and W. W. Gregg (2007), Assimilation of seawifs data into a global ocean-biogeochemical model using a local seik filter, *Journal of Marine Systems*, *68*(1-2), 237–254.

Nerger, L., and W. W. Gregg (2008), Improving assimilation of seawifs data by the application of bias correction with a local seik filter, *Journal of marine systems*, *73*(1-2), 87–102.

O'Dea, E., R. Furner, S. Wakelin, J. Siddorn, J. While, P. Sykes, R. King, J. Holt, and H. Hewitt (2017), The co5 configuration of the 7 km atlantic margin model: large-scale biases and sensitivity to forcing, physics options and vertical resolution, *Geoscientific Model Development*, *10*(8), 2947.

Pauly, D., V. Christensen, S. Guénette, T. J. Pitcher, et al. (2002), Towards sustainability in world fisheries, *Nature*, *418*(6898), 689.

Polimene, L., S. D. Archer, M. Butenschön, and J. I. Allen (2012), A mechanistic explanation of the sargasso sea dms âĂIJsummer paradoxâĂİ, *Biogeochemistry*, *110*(1-3), 243–255.

Lé Quére, C., S. P. Harrison, I. Colin Prentice, E. T. Buitenhuis, O. Aumont, L. Bopp, H. Claustre, L. Cotrim Da Cunha, R. Geider, X. Giraud, et al. (2005), Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models, *Global Change Biology*, *11*(11), 2016–2040.

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax (2007), Daily high-resolution-blended analyses for sea surface temperature, *Journal of Climate*, *20*(22), 5473–5496.

Ryan, A., C. Regnier, P. Divakaran, T. Spindler, A. Mehra, G. Smith, F. Davidson, F. Hernandez, J. Maksymczuk, and Y. Liu (2015), Godae oceanview class 4 forecast verification framework: Global ocean inter-comparison, *Journal of Operational Oceanography*, *8*(sup1), s98–s111.

Sathyendranath, S., R. Brewin, C. Brockmann, V. Brotas, S. Ciavatta, A. Chuprin, A. Couto, R. Doerffer, M. Dowell, M. Grant, et al. (2016), Creating an ocean-colour time series for use in climate studies: The experience of the ocean-colour climate change initiative, *Remote Sens. Environ.*

Saux Picart, S., M. Butenschön, and J. Shutler (2012), Wavelet-based spatial comparison technique for analysing and evaluating two-dimensional geophysical model fields, *Geoscientific Model Development*, *5*(1), 223–230.

Shimoda, Y., and G. B. Arhonditsis (2016), Phytoplankton functional type modelling: running before we can walk? a critical evaluation of the current state of knowledge, *Ecological modelling*, *320*, 29–43.

Shulman, I., S. Frolov, S. Anderson, B. Penta, R. Gould, P. Sakalaukus, and S. Ladner (2013), Impact of bio-optical data assimilation on short-term coupled physical, bio-optical model predictions, *Journal of Geophysical Research: Oceans*, *118*(4), 2215–2230.

Simon, E., and L. Bertino (2012), Gaussian anamorphosis extension of the denkf for combined state parameter estimation: Application to a 1d ocean ecosystem model, *Journal of Marine Systems*, *89*(1), 1–18.

Simon, E., A. Samuelsen, L. Bertino, and S. Mouysset (2015), Experiences in multiyear combined state–parameter estimation with an ecosystem model of the north atlantic and arctic oceans using the ensemble kalman filter, *Journal of Marine Systems*, *152*, 1–17.

Skogen, M. D., K. Eilola, J. L. Hansen, H. M. Meier, M. S. Molchanov, and V. A. Ryabchenko (2014), Eutrophication status of the north sea, skagerrak, kattegat and the baltic sea in present and future climates: A model study, *Journal of Marine Systems*, *132*, 174–184.

Smyth, T. J., J. R. Fishwick, L. Al-Moosawi, D. G. Cummings, C. Harris, V. Kitidis, A. Rees, V. Martinez-Vicente, and E. M. Woodward (2009), A broad spatio-temporal view of the western english channel observatory, *Journal of Plankton Research*, *32*(5), 585–601.

Teruzzi, A., S. Dobricic, C. Solidoro, and G. Cossarini (2014), A 3-d variational assimilation scheme in coupled transport-biogeochemical models: Forecast of mediterranean biogeochemical properties, *Journal of Geophysical Research: Oceans*, *119*(1), 200–217.

Torres, R., J. Allen, and F. Figueiras (2006), Sequential data assimilation in an upwelling influenced estuary, *Journal of Marine Systems*, *60*(3), 317–329.

Tsiaras, K. P., I. Hoteit, S. Kalaroni, G. Petihakis, and G. Triantafyllou (2017), A hybrid ensemble-oi kalman filter for efficient data assimilation into a 3-d biogeochemical model of the mediterranean, *Ocean Dynamics*, *67*(6), 673–690.

Wakelin, S., J. Holt, J. Blackford, J. Allen, M. Butenschön, and Y. Artioli (2012), Modeling the carbon fluxes of the northwest european continental shelf: Validation and budgets, *Journal of Geophysical Research: Oceans*, *117*(C5).

Waters, J., D. J. Lea, M. J. Martin, I. Mirouze, A. Weaver, and J. While (2015), Implementing a variational data assimilation system in an operational 1/4 degree global ocean model, *Quarterly Journal of the Royal Meteorological Society*, *141*(687), 333–349.

Widdicombe, C., D. Eloire, D. Harbour, R. Harris, and P. Somerfield (2010), Long-term phytoplankton community dynamics in the western english channel, *Journal of Plankton Research*, *32*(5), 643–655.

Xiao, Y., and M. Friedrichs (2014), Using biogeochemical data assimilation to assess the relative skill of multiple ecosystem models in the mid-atlantic bight: effects of increasing the complexity of the planktonic food web, *Biogeosciences*, *11*(11), 3015–3030.

Yool, A., E. Popova, and T. Anderson (2013), Medusa-2.0: an intermediate complexity biogeochemical model of the marine carbon cycle for climate change and ocean acidification studies, *Geoscientific Model Development*, *6*(5), 1767–1811.