

## Building the capacity for forecasting marine biogeochemistry and ecosystems: recent advances and future developments

M. Gehlen, R. Barciela, L. Bertino, P. Brasseur, M. Butenschön, F. Chai, A. Crise, Y. Drillet, D. Ford, D. Lavoie, P. Lehodey, C. Perruche, A. Samuelsen & E. Simon

To cite this article: M. Gehlen, R. Barciela, L. Bertino, P. Brasseur, M. Butenschön, F. Chai, A. Crise, Y. Drillet, D. Ford, D. Lavoie, P. Lehodey, C. Perruche, A. Samuelsen & E. Simon (2015) Building the capacity for forecasting marine biogeochemistry and ecosystems: recent advances and future developments, *Journal of Operational Oceanography*, 8:sup1, s168-s187, DOI: [10.1080/1755876X.2015.1022350](https://doi.org/10.1080/1755876X.2015.1022350)

To link to this article: <https://doi.org/10.1080/1755876X.2015.1022350>



© 2015 The Author(s). Published by Taylor & Francis.



[View supplementary material](#)



Published online: 09 Jun 2015.



[Submit your article to this journal](#)



Article views: 1665



[View related articles](#)



[View Crossmark data](#)



Citing articles: 32 [View citing articles](#)

## Building the capacity for forecasting marine biogeochemistry and ecosystems: recent advances and future developments

M. Gehlen<sup>a\*</sup>, R. Barciela<sup>b</sup>, L. Bertino<sup>c</sup>, P. Brasseur<sup>d</sup>, M. Butenschön<sup>e</sup>, F. Chai<sup>f</sup>, A. Crise<sup>g</sup>, Y. Drillet<sup>h</sup>, D. Ford<sup>b</sup>, D. Lavoie<sup>i</sup>, P. Lehodey<sup>j</sup>, C. Perruche<sup>h</sup>, A. Samuelsen<sup>c</sup> and E. Simon<sup>c</sup>

<sup>a</sup>Laboratoire des Sciences du Climat et de l'Environnement/Institut Pierre-Simon Laplace (LSCE/IPSL), Orme des Merisiers, 91191 Gif-sur-Yvette cedex, France; <sup>b</sup>Met Office, Hadley Centre, Exeter, UK; <sup>c</sup>Mohn-Sverdrup Center/NERSC, Bergen, Norway; <sup>d</sup>Laboratoire de Glaciologie et Géophysique de l'Environnement (LGGE), Grenoble, France; <sup>e</sup>Plymouth Marine Laboratory (PML), Plymouth, UK; <sup>f</sup>School of Marine Sciences, University of Maine, Orono, USA; <sup>g</sup>Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (OGS), Trieste, Italy; <sup>h</sup>Mercator Océan, Toulouse, France; <sup>i</sup>Maurice Lamontagne Institute, Fisheries and Oceans Canada, Mont-Joli, Canada; <sup>j</sup>Collecte Localisation Satellite (CLS), Toulouse, France

Building the capacity for monitoring and forecasting marine biogeochemistry and ecosystem dynamics is a scientific challenge of strategic importance in the context of rapid environmental change and growing public awareness of its potential impacts on marine ecosystems and resources. National Operational Oceanography centres have started to take up this challenge by integrating biogeochemistry in operational systems. Ongoing activities are illustrated in this paper by presenting examples of (pre-)operational biogeochemical systems active in Europe and North America for global to regional applications. First-order principles underlying biogeochemical modelling are briefly introduced along with the description of biogeochemical components implemented in these systems. Applications are illustrated with examples from the fields of hindcasting and monitoring ocean primary production, the assessment of the ocean carbon cycle and the management of living resources. Despite significant progress over the past 5 years in integrating biogeochemistry into (pre-)operational data-assimilation systems, a sustained research effort is still needed to assess these systems and their products with respect to their usefulness to the management of marine systems.

### Introduction

Global warming, in synergy with ocean acidification, eutrophication, deoxygenation, as well as the ongoing exploitation of living marine resources drive major changes in marine biogeochemistry and ecosystems (Bopp et al. 2005, 2013; Orr et al. 2005; Keeling et al. 2010; Gehlen et al. 2011; Stock et al. 2011). The awareness of ongoing and future changes prompted a considerable community effort over the past decade, resulting in the development of increasingly complex biogeochemical ocean general circulation models (BOGCM) (Moore et al. 2002; Le Quéré et al. 2005; Aumont & Bopp, 2006; Follows et al. 2007; Vichi et al. 2013). Today's BOGCMs capture the major features of large-scale distributions of dissolved inorganic carbon, alkalinity, nutrients, primary and export production (Schneider et al. 2008; Lazzari et al. 2010; Séférian et al. 2012). These models gradually developed into lower trophic level ecosystem (LTE) models by explicitly representing the first levels of the marine food-web from phytoplankton to zooplankton. They are increasingly used in a variety of studies encompassing the monitoring and short-term forecasting of ecosystem health up to the

assessment of potential impacts of climate change on higher trophic levels (Stock et al. 2011). Biogeochemical-LTE models are becoming essential tools for:

- (1) producing decadal reanalyses (Ford et al. 2012; Fontana et al. 2013) as well as projecting trends in ocean biogeochemistry and ecosystems in a changing global environment (Orr et al. 2005; Schneider et al. 2008; Bopp et al. 2013);
- (2) designing the observation system needed to monitor the current marine biogeochemical status (Lin et al. 2010; Xue et al. 2012);
- (3) quantifying anthropogenic impacts on biogeochemical cycles and marine ecosystems (Bopp et al. 2005; Orr et al. 2005);
- (4) assessing and forecasting changes on time-scales relevant to management purposes (Stumpf et al. 2009; Raine et al. 2010) (e.g. weekly to seasonal time-scales for the prediction of harmful algal blooms, the management of marine resources, episodic hypoxia/anoxia, oil spill or dispersal of a pollutant).

\*Corresponding author. Email: [marion.gehlen@lsce.ipsl.fr](mailto:marion.gehlen@lsce.ipsl.fr)

While there is an increasing demand for the last item, the development of a predictive capability for ocean biogeochemistry and LTEs on time-scales ranging from weeks to seasons remained, until recently, impeded by the lack of appropriate modelling tools and observational products.

Past decades saw the rapid expansion of global ocean observing capabilities building on remote sensing and *in situ* technologies [e.g. Argo floats, moored arrays such as Tropical Ocean-Global Atmosphere (TOGA)/Tropical Atmosphere Ocean (TAO) array, Prediction and Research Moored Array in the Tropical Atlantic (PIRATA) or Research Moored Array for African-Asian-Australian Monsoon Analysis and Prediction (RAMA)]. The improved data coverage, both in time and in space, allowed the successful implementation of ocean data-assimilation platforms in parallel with the development of operational oceanography applications. The development, implementation and demonstration of feasibility were at the centre of the ‘Global Ocean Data Assimilation Experiment (GODAE)’. This important 10-year international effort proved the feasibility and usefulness of integrated observation-model systems for physical variables. Over the recent years, biogeochemical variables were progressively added to global observing systems (e.g. Bio-Argo, ocean colour sensors and autonomous nutrient analysers). It was paralleled by the first ‘green’ applications of operational oceanography aiming at the near-real-time assessment of the ocean biogeochemical state (Brasseur et al. 2009).

Here, an overview of several (pre-)operational biogeochemical systems presently active in Europe and North America for global to regional and sub-regional applications is presented. Systems were selected such as to illustrate regional versus global scales of applications, complexity of biogeochemical models, various approaches to data assimilation, as well as status of implementation (e.g. hindcast simulation without biogeochemical data assimilation to fully coupled operational systems). Emphasis is put on developments over the past decade, and a status report highlighting new opportunities for applications is presented. A brief overview of physical–biogeochemical systems for operational applications is first discussed. The overview distinguishes between biogeochemical models and the forecasting systems in which these are embedded. Next, (pre-)operational applications for topics covering biogeochemistry to ecosystems are illustrated, and to conclude, major blocking points and prospects for future evolution are identified.

## Physical–biogeochemical systems for operational applications

### *Biogeochemical components*

Biogeochemical components embedded in large-scale physical operational systems are simplified representations

of the complex and poorly constrained network of biological interactions governing the cycling of matter between reservoirs of the Earth system. These models do not aim at a detailed representation of marine ecosystems, but rather focus on processes and stocks relevant to biogeochemical cycles up to the dynamics of the first trophic levels. Organisms are often grouped into ‘Plankton Functional Types’ (PFTs) according to their specific function in carbon and nutrient cycles (Le Quéré et al. 2005; Hood et al. 2006). Next to function, size is an important structuring element in these models. Size participates in modulating nutrient acquisition, as well as grazing pressure and hence both bottom-up and top-down control acting on LTEs. The recognition that fluxes within the marine foodweb are characterized by near-constant elemental ratios (C:N:P = 106:16:1) (Redfield et al. 1963) when averaged over space and time (e.g. seasonal cycle) allows models to use a single basic currency (e.g. carbon, nitrogen or phosphorus). Fluxes of the remaining elements are computed from fixed stoichiometric relationships. While this approach limits the number of state variables, it does not allow representation of the decoupling of elemental cycles observed during blooms and subsequent export events (Engel et al. 2002). Some models are thus based on plastic stoichiometry (Baretta et al. 1995; Vichi et al. 2013) which allows for flexible ratios between the major elements in response to environmental conditions and physiological requirements. The stoichiometric plasticity is generally implemented through a modified internal cell-quota model (Droop, 1973), which relates the specific growth rate and nutrient uptake to the intracellular nutrient to carbon ratio (Baretta-Bekker et al. 1997), by adding a Michaelis–Menten limitation term for the external nutrient availability. Flexible stoichiometry improves the representation of co-limited oligotrophic regimes in biogeochemical models.

The biogeochemical models (Table 1, supplementary material for a detailed presentation of individual models) considered here are of the NPZD type, where NPZD denotes nutrient, phytoplankton, zooplankton and detritus. While HadOCC, PISCES and BFM have been developed for open ocean applications, NORWECOM and GSBM are regional in scope. The European Regional Seas Ecosystem Model (ERSEM) ([www.shelfseasmodelling.org](http://www.shelfseasmodelling.org)) was originally developed and routinely applied in the shelf-seas context but has since been extended for the global ocean and basin scale applications. It is consequently applicable for shelf seas and open ocean. The contrast in complexity between these models (Table 1) reflects at first order the level of complexity required by the precise context of the application, e.g. biogeochemical cycling versus ecological application. The trade-off between complexity and computational cost is an additional important consideration. With the exception of HadOCC, all models distinguish between large and small plankton or, in other terms, between fast-

Table 1. Overview of biogeochemical models included in (pre-)operational systems.

	HadOCC	ERSEM	NORWECOM	PISCES	BFM	GSBM
Biogeochemical cycles	N (NO <sub>3</sub> , NH <sub>4</sub> ), Si, O <sub>2</sub> , P	N (NO <sub>3</sub> , NH <sub>4</sub> ), Si, O <sub>2</sub> , P	N (NO <sub>3</sub> , NH <sub>4</sub> ), Si, O <sub>2</sub> , P	N (NO <sub>3</sub> , NH <sub>4</sub> ), Si, Fe, P, C <sup>a</sup> , O <sub>2</sub>	N (NO <sub>3</sub> , NH <sub>4</sub> ), Si, Fe, P, C <sup>a</sup> , O <sub>2</sub>	N (NO <sub>3</sub> , NH <sub>4</sub> ), C <sup>a</sup> , O <sub>2</sub>
autotrophic PFTs	Phytoplankton	Picophytoplankton, flagellates, diatoms, dinoflagellates	Flagellates, diatoms	Nanophytoplankton, diatoms	Picophyto-, nanophytoplankton, diatoms	Flagellates, diatoms
Heterotrophic PFTs	Zooplankton	Microzoo-, mesozoo-plankton, heterotrophic nanoflagellates, bacterioplankton	Microzoo-, mesozoo-plankton	Microzoo-, mesozoo-plankton	Nanozoo-, microzoo-, mesozoo-, bacterioplankton	Microzoo-, mesozoo-plankton
BGC functions included without explicit PFT	CaCO <sub>3</sub> production			CaCO <sub>3</sub> production/dissolution, N <sub>2</sub> -fixation/denitrification	Biogenic Si dissolution, biogenic Fe dissolution	
External inputs		River nutrients, sediments, terrestrial sediments	River nutrients	River carbon and nutrients, aeolian Fe, Si and N, sedimentary Fe source	Aeolian Fe	River carbon and nutrients
References	<sup>b</sup>	<sup>c</sup>	<sup>d</sup>	<sup>e</sup>	<sup>f</sup>	<sup>g</sup>

<sup>a</sup>Dissolved and particulate, prognostic alkalinity, CaCO<sub>3</sub> production and dissolution, CO<sub>2</sub> chemistry fully resolved.

<sup>b</sup>HadOCC (Hadley Centre Ocean Carbon Cycle Model); Palmer and Totterdell (2001); Geider et al. (1997); Eppley (1972); Anderson (1993); Hemmings et al. (2008).

<sup>c</sup>ERSEM (European Regional Seas Ecosystem Model); Edwards et al. (2012); Blackford et al. (2004); Artoli et al. (2012); Sykes and Barciela (2012).

<sup>d</sup>NORWECOM (Norwegian Ecological Model); Skogen et al. (1995); Skogen et al. (2007); Hansen and Samuelson (2009); Moll and Stegert (2007); Pätzsch et al. (2009); Stegert et al. (2009); Samuelson et al. (2009).

<sup>e</sup>PISCES (Pelagic Interaction Scheme for Carbon and Ecosystem Studies); Aumont and Bopp (2006); Aumont et al. (2008); Ludwig et al. (1996); Johnson et al. (1999); de Baar and de Jong (2001).

<sup>f</sup>BFM (Biogeochemical Flux Model); Lazzari et al. (2010); Vichi et al. (2007a); Vichi et al. (2007b); Baretta et al. (1995); Orr (1999); Geider et al. (1998).

<sup>g</sup>GSBM (Gulf of St Lawrence Biogeochemical Model); Le Fouest et al. (2005); Le Fouest et al. (2006); Martin et al. (1987); Jin et al. (2007); Signorini et al. (2001); Dickson et al. (2007).

growing small cells and slower-growing large cells. The former are favoured under low-nutrient conditions, while the latter tend to dominate under high-nutrient conditions, yielding high biomasses and efficiently driving carbon export. All models except NORWECOM include a formulation of the carbon cycle. In GSBM, alkalinity is not a prognostic tracer but is derived from the relationship between observed alkalinity and modelled salinity. The number of nutrients is another important between-model difference, in particular the inclusion or omission of iron (Fe). Fe limits primary production across large oceanic regions, i.e. the high-nutrient low-chlorophyll (HNLC) regions, such as the Southern Ocean, the Equatorial Pacific. The availability of Fe (along with phosphate) also modulates atmospheric N<sub>2</sub> fixation. Both NORWECOM and GSBM do not include Fe cycling, since Fe is not a limiting micro-nutrient in their target regions. BFM, as well as PISCES, consider Fe biogeochemistry along with corresponding sources of external input.

Despite all models being of the NPZD type and hence sharing similarities in their underlying conceptual framework, there is a progression in complexity from HadOCC to NORWECOM, respectively GSBM, and PISCES to BFM. With the widespread use of biogeochemical models and the rapid increase in state variables, the identification of optimal complexity has become an active area of research (Friedrichs et al. 2007; Ward et al. 2013; Kwiatkowski et al. 2014; Xiao & Friedrichs, 2014a, 2014b). In the context of this paper, we are more specifically concerned with how model complexity plays out in model forecast skill. The models presented here are coupled to different physical models, and any coordinated model inter-comparison exercise would highlight differences in skill resulting from differences in both the biogeochemical and physical components. It would fall short in identifying the link between the structure of the biogeochemical model and forecast skill of the system. Addressing the question of optimal complexity of biogeochemical models calls for an approach in which the complexity of the biogeochemical component is increased in a stepwise fashion within an unchanged physical environment (Kwiatkowski et al. 2014).

Published studies suggest that models with two phytoplankton and two zooplankton size classes would have the optimal complexity for representing biogeochemical dynamics across a variety of regimes together with a good predictive ability (Friedrichs et al. 2007; Ward et al. 2013; Xiao & Friedrichs, 2014a, 2014b). Increasing the number of state variables might increase model skill – in particular when compared with local 1D or regional data sets – but goes along with a decrease in portability and, possibly, reduced predictive skill. Finally, model complexity is a function not only of the number of PFTs, but also of the degree of sophistication of process formulation. Future studies are required to assess the benefits of the

inclusion of multiple nutrients and chemical speciation, of flexible stoichiometry and hence internal cell quota, etc. to model forecasting.

For operational applications, trade-offs between model complexity, computational costs and forecast skill have to be carefully evaluated, keeping in mind that models should be as simple as possible, but as complex as needed. Ultimately, the required level of complexity will depend on the scientific questions the models have been/are being designed to answer. Operational systems are deployed over a variety of spatial scales from regional to basin-wide to global applications. Models designed to address coastal processes often require a higher model complexity because key processes that govern plankton dynamics in the shallow waters (such as tidal processes, resuspension, sediment dynamics, benthic dynamics, etc.) are not of critical importance in the open ocean. Complexity in the number of plankton variables in coastal waters is also justified in order to capture, for example, occurrences of harmful algal blooms that affect fisheries, aquaculture, human health and local economies, whereas this is not an issue in the open ocean. There is thus no unique answer to the question of optimal model structure, and the preceding rather pleads in favour of a diversity of biogeochemical models.

### Forecasting systems

The major characteristics of the forecasting systems in which the biogeochemical components are embedded are summarized in Table 2.

#### FOAM-HadOCC

HadOCC has been coupled online to the global configuration of the Met Office's operational Forecasting Ocean Assimilation Model (FOAM) (Storkey et al. 2010; Blockley et al. 2014), which is based on the Nucleus for European Modelling of the Ocean (NEMO) hydrodynamic model (Madec, 2012) and the Los Alamos sea ice model (CICE) (Hunke & Lipscomb, 2010). A key feature of FOAM is the ability to assimilate remotely sensed and *in situ* observations of temperature, salinity, sea-level anomaly (SLA) and sea ice concentration. The data assimilation has recently been upgraded from the analysis correction scheme of Martin et al. (2007) to the 3D-Var NEMOVAR scheme (Waters et al. 2014).

The data-assimilation capability has been extended to also assimilate chlorophyll data derived from remotely sensed ocean colour and *in situ* observations of sea surface partial pressure of carbon dioxide ( $p\text{CO}_2$ ). This has been used for both reanalysis and near-real-time forecasting applications, with the aim of improving estimates of air-sea  $\text{CO}_2$  flux and other carbon-cycle variables, as well as biological variables. Both the chlorophyll (Ford

et al. 2012) and  $p\text{CO}_2$  (While et al. 2012) assimilation schemes follow a two-stage process. The first stage is to generate a set of sea-surface chlorophyll or  $p\text{CO}_2$  increments. This is currently done using the analysis correction scheme of Martin et al. (2007) but will be updated to use the same NEMOVAR scheme as the operational physics system. Both employ an observation operator that uses a first-guess-at-appropriate-time (FGAT) technique. The second stage is to generate a set of multivariate increments, which are then applied to the model. Surface chlorophyll increments are converted to a set of increments for all biogeochemical model state variables, at all depths, using the nitrogen-balancing scheme of Hemmings et al. (2008). This uses a principle of conservation of nitrogen and carbon to propagate the information from the chlorophyll observations as fully and realistically as possible to the rest of the model. The scheme is designed to be computationally efficient, making it suitable for operational applications. In the  $p\text{CO}_2$  assimilation scheme, the surface  $p\text{CO}_2$  increments are converted to dissolved inorganic carbon and alkalinity increments using the method described in While et al. (2012), and applied throughout the mixed layer.

The coupled FOAM-HadOCC system is currently pre-operational, and different configurations are used for different experiments. The open ocean physical FOAM system is fully operational, and details of its running are given in Blockley et al. (2014). There is also an operational FOAM configuration covering the northwest European shelf seas (Edwards et al. 2012; O'Dea et al. 2012). This is coupled online to ERSEM, and both physical and biogeochemical analysis and forecast products are provided operationally via the MyOcean project (<http://www.myocean.eu/>).

#### TOPAZ-NORWECOM

TOPAZ-NORWECOM is used in the Arctic Forecasting Center in the MyOcean project. The modelling system is the result of a collaboration between the IMR and NERSC for the developments and MET Norway, which exploits the system operationally. The coupled model is HYCOM-NORWECOM (Pätsch et al. 2009; Samuelsen et al. 2009). It uses the HYbrid Coordinate Ocean Model [HYCOM (Bleck, 2002)] as the physical model and NORWECOM (Oki & Sud, 1998; Skogen & Søiland, 1998) as the ecosystem model. The physical and biological models are coupled online. HYCOM uses a combination of isopycnal and z-coordinates, which allows for both good conservative properties in the deep ocean and high vertical resolution in the upper mixed layer. The present model configuration has 28 vertical layers, of which the five upper layers are in z-coordinates, and the lower 23 layers are hybrid layers.

The river forcing is generated using a hydrological model – total runoff integrating pathways (TRIP) (Oki &

Table 2. Synthetic overview of main characteristics of forecasting systems discussed in this paper.

System	Ocean model	Biogeochemical model	Configuration	Data-assimilation scheme		Assimilated data		System status
				PHYS	BGC	PHYS	BGC	
FOAM-HadOCC	NEMO3.2-CICE	HadOCC	Global, 1/4°cos(lat) resolution, 75 vertical layers	3D-Var	Analysis correction + multivariate balancing	Satellite SLA, SST, sea ice, <i>in situ</i> SST, T/S profiles	Chlorophyll-a or pCO <sub>2</sub>	Pre-operational (BGC), operational (PHYS)
FOAM-ERSEM	NEMO3.2	ERSEM	Atlantic Margin, 7 km resolution, 32 vertical layers	Analysis correction	No	SST	No	Operational
TOPAZ-NORWECOM	HYCOM	NORWECOM	North Atlantic and Arctic (Bering Strait), 50 km resolution 28 vertical layers	DEnKF	DEnKF and Gaussian anamorphosis	Satellite SLA, SST, sea ice	Chlorophyll-a	Pre-operational
TOPAZ-NORWECOM	HYCOM	NORWECOM	Same but 12 km resolution	DEnKF	No	Satellite SLA, SST, sea ice, <i>in situ</i> T/S profiles	No	Operational
MERCATOR-OCEAN/BIOMER	NEMO 3.1	PISCES (NEMO3.2) offline coupled <sup>a</sup>	Global, 1/4°cos(lat) resolution, 50 vertical layers	SAM2V1	No	Satellite SST, SSH, <i>in situ</i> T/S profiles	No	Operational
MFS	NEMO3.4 +waves+atm. pressure	BFM (OPATM) offline coupled <sup>a</sup>	Mediterranean Sea (1/16°), 72 vertical layers	3D-Var	3D-Var	Satellite SSH, <i>in situ</i> T/S profiles	Chlorophyll-a	Operational
CANOPA-GSBM	OPA9-LIM2	GSBM	East Canadian shelf (1/12°), 46 vertical layers	No	No	No	No	Non-assimilative hindcast

BGC: biogeochemistry; PHYS: physics.

<sup>a</sup>Biogeochemical model coupled offline (run sequential) to physical model.

Sud, 1998). The river outflow calculated by TRIP was combined with data from Global Nutrient Export from Watersheds [GlobalNEWS (Beusen et al., 2009; Seitzinger et al. 2005)] and used as nutrient river input to the model. Nutrients and oxygen are relaxed to climatology at the lateral boundaries. For initialization of NORWECOM, climatological values of nutrients and oxygen were used; all other variables were set to a constant low value.

The capacity to assimilate ocean colour data has been developed using a version of the Ensemble Kalman Filter that includes both a Gaussian anamorphosis in order to handle the non-Gaussian model variables (Simon & Bertino, 2009) and a state augmentation in order to estimate biological model parameters together with the model variables (Simon & Bertino, 2012). These assimilation procedures have been applied with the HYCOM-NORWECOM system coupled online at a coarser resolution than that of the real-time forecast system – 50 km instead of 12 km – for the period 2007–2010. An ensemble of 100 members has been integrated. The data are available as part of the MyOcean service as the first Arctic biological reanalysis that assimilates jointly physical (altimeter, SST, sea ice) and biological observations.

#### MERCATOR-OCEAN/BIOMER

Since 2012, the French centre for operational oceanography Mercator Océan has added near-real-time assessment of ocean biogeochemistry called BIOMER to its suite of products. BIOMER is based on the ecosystem model PISCES forced offline with ocean physical fields provided by the global operational system PSY3V3 (Lellouche et al. 2013) at a horizontal resolution of  $1/4^\circ \cos(\text{lat})$  (NEMO 3.1, 50 vertical levels, atmospheric forcings from ECMWF operational analysis at 3 h, CORE bulk formulation), with assimilation of temperature, salinity and sea-level data via the SAM2 assimilation scheme based on the SEEK method (Brasseur & Verron, 2006), an Incremental Analysis Update (Bloom et al. 1996; Ourmières et al. 2006) and a large-scale bias correction for temperature and salinity based on a 3D-var approach. To decrease the computational burden, the spatial resolution of physical fields was coarsened to  $1^\circ$ , and temporal output was averaged over one week (El Moussaoui et al. 2011) for BIOMER1. Assessment of biogeochemical tracer (El Moussaoui et al. 2011) fields revealed a degradation of modelled chlorophyll [Figure 1(b)] and nutrient distributions (not shown) in the equatorial band compared with a simulation without assimilation of physical data. The overestimation of chlorophyll and nutrients was attributed to spurious vertical fluxes in response to unbalanced physical data assimilation (Barciela et al. 2012) associated with biases in assimilated MSSH (mean sea surface height).

The recent version of the system, BIOMER4, uses the same spatial resolution for the physical and biogeochemical

systems [ $1/4^\circ \cos(\text{lat})$ ]. Besides its higher spatial resolution, BIOMER4 [Figure 1(a)] relies on a daily forcing of the offline biogeochemical model with ocean physical and atmospheric fields, and assimilation of MSSH is changed to hybrid MSSH. Modifications of the biogeochemical component include the transport of passive tracers (coefficient of diffusion) and the tuning of biogeochemical parameters (specifically the photosynthetic response to changes in light intensity and the sinking speed of large particles). Figure 1 compares surface mean chlorophyll (*Chl*) distributions for 2013 for (a) BIOMER4, (b) BIOMER1 and (c) GlobColour data. The bias towards chlorophyll levels that are too high, especially in the western equatorial Pacific identified in BIOMER1 (El Moussaoui et al. 2011), is alleviated in the new version. Similarly, productive regions associated with the continental shelf (e.g. Indonesian Archipelago, China Sea and North Sea) and eastern boundary upwelling systems (e.g. Chile-Peru) are better represented in BIOMER4. Standard deviations computed from monthly surface nutrient, oxygen concentrations and chlorophyll fields (year 2013) for model output and data, as well as correlation coefficients, are presented on a Taylor diagram [Figure 1(d)]. Both scores are improved in the new model version. BIOMER4 proved skilful in reproducing the seasonal cycle and interannual variability of biogeochemical properties, as illustrated in Figure 2, for North Atlantic chlorophyll concentrations. For a complete evaluation, please refer to MyOcean quality information document MYO2-GLO-QUID-001014 (<http://www.myocean.eu/web/69-myocean-interactive-catalogue.php>). Improvements between BIOMER1 and BIOMER4 result from the combination of changes made to the physical–biogeochemical system: (1) increasing the spatial resolution from  $1^\circ$  to  $1/4^\circ$  improved the representation of productive shelf regions and Eastern Boundary Upwelling Systems; (2) shifting from assimilation of MSSH to hybrid-MSSH reduced spurious vertical velocities; and (3) adjusting selected parameters of PISCES reduced its productivity.

Model output will be made available in near-real-time with a lag of one day for nitrate, phosphate, dissolved oxygen, iron, dissolved silica, phytoplankton biomass, net primary production (NPP) and chlorophyll on the MyOcean portal in September 2014. The near-real-time assessment of ocean biogeochemistry is completed by an interannual hindcast simulation from 2007 to the present.

#### MFS-OPATM-BFM

The OPATM-BFM model system (Lazzari et al. 2010) has been developed at National Institute for Oceanography and Experimental Geophysics (OGS) and is the biogeochemical component of the MyOcean Mediterranean Monitoring and Forecasting System (MedMFC). Since the start of the pre-operational implementation in 2007 within the MERSEA

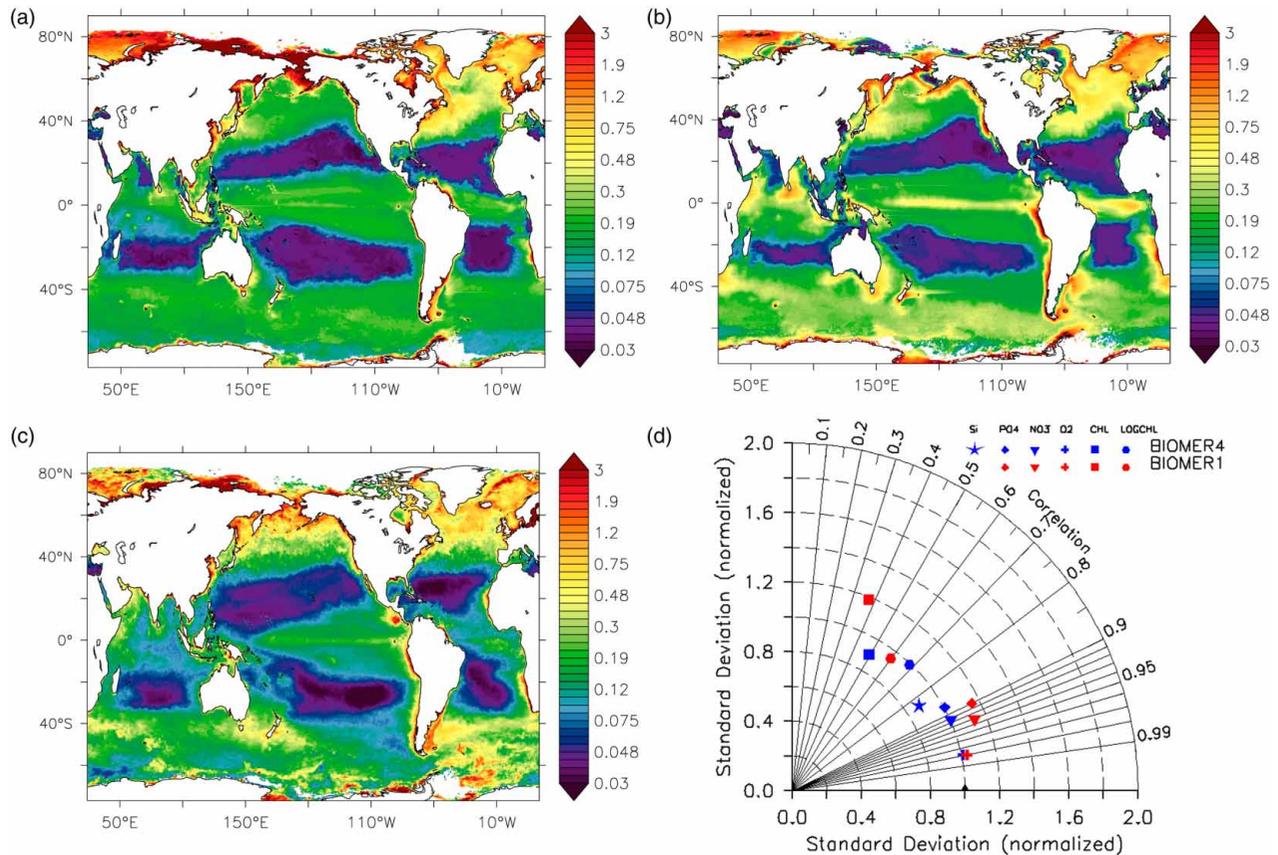


Figure 1. Mean surface chlorophyll-a concentrations ( $\text{mgChla}/\text{m}^3$ ) in 2013: (a) BIOMER4 (recent model version); (b) BIOMER1 (previous model version); (c) GlobColour observational product; (d) Taylor diagram of modelled surface concentrations of nutrients (dissolved silica, phosphate, nitrate), oxygen, chlorophyll-a and  $\log_{10}$ (chlorophyll-a). Standard deviations are normalised with respect to data (WOA2009 and GlobColour).

project, substantial improvements have been made. In the framework of MyOcean, the system was upgraded by adding a 3D variational assimilation scheme (Teruzzi et al. 2014a) for satellite (MODIS)-derived chlorophyll and apparent optical properties (AOP). While satellite-derived chlorophyll constrains the C:chlorophyll ratio of phytoplankton, AOP allow the estimation of the optical model. The transport model is based on OPA 8.1 (Foujols et al. 2000) including additional enhancements such as a flux-corrected advection scheme. The horizontal spatial resolution is  $1/16^\circ$ . The physical forcing for biogeochemical processes is pre-computed at the same resolution by the ocean general circulation model MFS based on NEMO3.4. MFS supplies the daily averaged fields of horizontal and vertical current velocities, vertical eddy diffusivity, potential temperature and salinity, in addition to surface data for solar shortwave irradiance and wind stress.

The Med-MFC operational chain runs unattended twice a week in an offline mode. It is composed of four steps: (1) downloading of forcing and daily MODIS chlorophyll data and pre-processing, (2) 3D-var satellite chlorophyll assimilation, (3) 7 day forecast of the biogeochemical state of the

sea and (4) MyOcean catalogue update. The biogeochemical run is composed of (1) 7 days of hindcast (the first day: analysis with data assimilation on Wednesday run only) and (2) 10 days of forecast for the Wednesday run. Operational output includes daily 3D fields of chlorophyll, oxygen, nitrate, phosphate, aggregated phytoplankton biomass and NPP.

#### CANOPA-GSBM

CANOPA-GSBM is being developed in collaboration with three Fisheries and Ocean Canada Institutes. The ocean circulation model, CANOPA, is described in Brickman & Drozdowski (2012). The modelling system is based on the ocean code OPA version 9.0 (Madec, 2012) coupled to the sea-ice model LIM2 (Madec et al. 1998). The physical and biochemical (GSBM) models are coupled online. The grid of the model covers the Gulf of St. Lawrence, the Scotian Shelf and the Gulf of Maine with a spatial resolution of  $1/12^\circ$  and 46 layers of variable thickness in the vertical. The model includes tidal and river forcing. It is a prognostic model, meaning that the temperature and

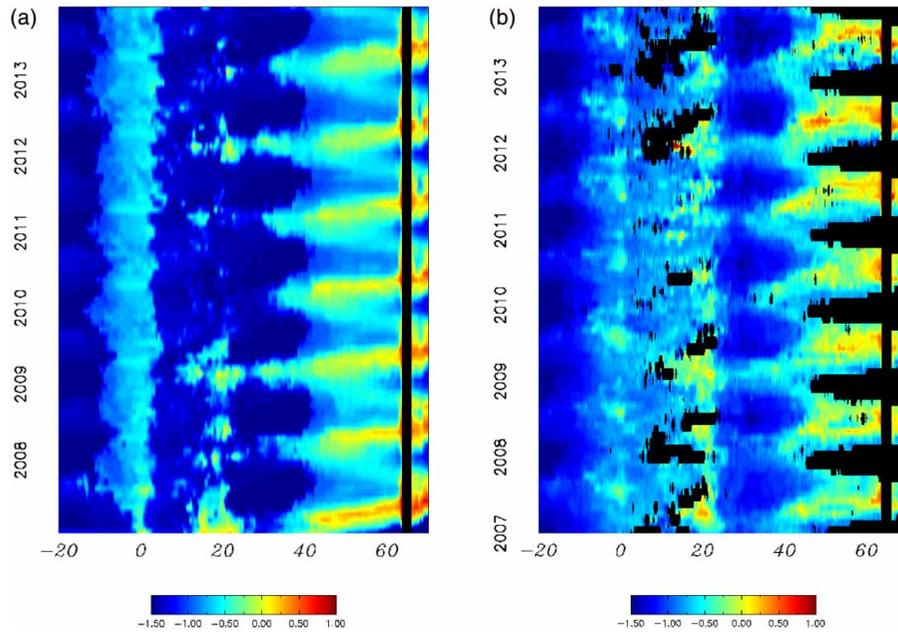


Figure 2. Hovmöller diagram of  $\log_{10}$  (chlorophyll-a) between 2007 and 2013 in the North Atlantic at 20W and 20S, 70N: (left panel) BIOMER4; (right panel) GlobColour.

salinity fields are free to evolve with time and are only constrained through open boundary conditions, freshwater runoff and surface forcing. Runoff from the large St. Lawrence river is estimated from water-level measurements, while runoff from the other rivers is obtained from a simple hydrological model (Lambert et al. 2013). Monthly climatologies for temperature, salinity, nitrate, oxygen and dissolved inorganic carbon are used to initialize the model and to provide lateral boundary conditions. Monthly climatologies of nitrate are also used for river input. Surface forcing for CANOPA is obtained from the Canadian Meteorological Centre Global Environmental Multiscale (CMC-GEM) atmospheric model (Pellerin et al. 2003).

CANOPA-GSBM is used to produce an annual state of the ecosystem for part of the Gulf of St. Lawrence. Within the Canadian National Approach under GODAE OceanView, CONCEPTS, EC and DFO plan over the next 4 years to couple GSBM to a subset of their Global Ice Ocean Prediction Systems (GIOPS) currently operational at Environment Canada as well as to a planned Regional Ocean Ice Prediction System (RIOPS) to be implemented in 2015 (Ryan et al. 2015; Tonani et al. 2015).

### ***Synthesis of present status of coupled physical–biogeochemical systems***

The review of (pre-)operational physical–biogeochemical systems presented here highlights the diversity of biogeochemical models and data-assimilation methods used across the community (Tables 1 and 2). CANOPA-

GSBM, in its present set-up, produces a yearly biogeochemical hindcast assessment at a regional scale. Two systems (MERCATOR-OCEAN/BIOMER and OPATM/BFM) take advantage of physical data-assimilation products to drive ocean biogeochemistry offline. This approach might be seen as an intermediate step towards online coupled physical–biogeochemical systems as exemplified by the FOAM-HadOCC and TOPAZ-NORWECOM model suites. Since the review by Brasseur et al. (2009), several GODAE systems are in the process of moving, or have moved, to operational mode (Table 2). Satellite ocean colour data are still the only biogeochemical data set meeting the requirements of spatial and temporal coverage for assimilation into operational systems. However, other sources of observations can be useful depending on the aims of the experiment – for instance, a  $p\text{CO}_2$  assimilation scheme is described in the following section of this paper.

A number of different techniques have been used to assimilate surface ocean chlorophyll, and these consistently improve modelled distributions compared with both the assimilated satellite data and independent *in situ* observations. However, very much as for the issue of model complexity discussed above, no unique recommendation emerges at this stage. The diversity of approaches guarantees the necessary degree of scientific diversity for the development of targeted applications and variable spatial scales. The increasing number of systems producing biogeochemical forecasts and reanalyses will enable coordinated evaluation and model intercomparison exercises in the near future.

## Physical–biogeochemical model applications in operational oceanography

### Marine primary production

NPP is at the base of the marine food web. The reliable simulation of its spatial distribution and temporal variability is of importance for understanding and predicting ecosystem dynamics, as well as for the development of downstream applications such as the operational management of marine resources. The evaluation of modelled NPP is not straightforward. Estimates derived from remote sensing rely themselves on empirical models that derive NPP as a function of chlorophyll, temperature and light. *In situ* measurements are scarce and fail to resolve spatial and temporal variability. This explains why biogeochemical models are evaluated against chlorophyll and nutrient distributions. Three-dimensional nutrient distributions reflect the combination of physical transport and biological activity in surface waters and across the water column (Séférian et al. 2012). In addition, data from Eulerian time-series stations and cruises provide complementary information on plankton biomasses and biological rate estimates, including NPP. In order to illustrate the skill of (pre-)operational physical–biogeochemical model systems, we present hereafter two biogeochemical reanalyses: one at the global scale and one for the Mediterranean Basin.

FOAM-HadOCC has recently been used to produce a 1° global daily biogeochemical reanalysis covering the period September 1997 to July 2012, assimilating daily merged chlorophyll products from GlobColour (<http://www.globcolour.info>). For this particular work, only chlorophyll data were assimilated, not physics or pCO<sub>2</sub>. The data assimilation has been shown to improve the fit of modelled chlorophyll not only to the assimilated satellite data but also to independent *in situ* chlorophyll observations, both at and beneath the surface. The changes to model chlorophyll result in corresponding changes to the primary production fields. Comparisons with *in situ* observations of nitrate show no degradation of the model nutrient fields as a result of the assimilation. These results are similar to those detailed in Ford et al. (2012), to which the reader is referred for a more detailed validation of the chlorophyll-assimilation scheme. The reanalysis is being used to investigate seasonal and interannual variability in the model results, such as relationships between biogeochemical variables and known climate drivers such as the El Niño Southern Oscillation and North Atlantic Oscillation, and how the data assimilation impacts this. Work is also under way to assess reanalysis results using initial products from the European Space Agency's Climate Change Initiative project (<http://www.esa-oceancolour-cci.org>).

A 12-year (1999–2010) reanalysis of Mediterranean Sea biogeochemical tracer fields at 1/8° has been made available as part of the recent MyOcean release V3.1. It

provides monthly 3D mean fields of chlorophyll, phosphate, nitrate, dissolved oxygen, NPP and phytoplankton biomass. A database of satellite derived (SeaWiFS) surface chlorophyll concentrations delivered by the MyOcean Ocean Colour Thematic Assembly Center was used for data assimilation. Data assimilation of surface chlorophyll concentration is performed weekly through a 3D variational scheme (3D-var) following Dobricic and Pinardi (2008) and modified as in Teruzzi et al. (2014a). To statistically project surface information to a fully 3D chlorophyll field, EOF analysis has been applied to a multi-annual (1998–2004) OPATM-BFM run (Lazzari et al. 2012) identifying 12 months and nine sub-regions. The efficiency of the procedure was assessed by visual inspection (see Figure 3), as well as by comparing predicted surface chlorophyll fields to satellite-derived observations and *in situ* data (Teruzzi et al. 2014a).

For the purpose of model validation and assessment of the impact of monovariate chlorophyll assimilation on the overall model prediction skill, an extensive dataset was collated for dissolved oxygen, nitrate, phosphate, dissolved silica and NPP (Lazzari et al. 2012). It combines biogeochemical measurements available at the OGS National Oceanographic Data Centre with additional recent cruises. The metrics introduced are based on nearest-point statistics. Model bias and correlation were calculated according to Allen et al. (2007) and summarized in Table 3 for chlorophyll, dissolved oxygen and nutrient data for the reanalysis and the corresponding free simulation (hindcast). The mostly lower bias at depth reflects the realistic initial conditions. Relative model output/data mismatch is amplified when expressed as relative error for low to very low nutrient levels. This explains in part the larger bias obtained for nitrate and phosphate for oligotrophic surface waters compared with the ocean interior. Dissolved silica distributions are poorly represented. Basin-wide correlations between observed and modelled tracer distributions are good at depth (again reflecting initial conditions) and for surface phosphate and dissolved oxygen fields. It follows from the comparison between the two simulations that there is a substantial improvement in surface chlorophyll. The bias on nutrient estimates is not adversely affected by data assimilation with improvements for four out of nine fields, while reanalysis correlation outperforms hindcast in seven out of nine cases. For this reanalysis, chlorophyll was assimilated and projected on phytoplankton biomass only, that is, without correcting nutrient concentrations. While this approach improved surface chlorophyll distribution, 3D nutrient fields and oxygen distribution only partially benefit from univariate assimilation of surface chlorophyll. This example argues in favour of the development of multivariate approaches to the assimilation of biogeochemical data in coupled physical–biogeochemical models.

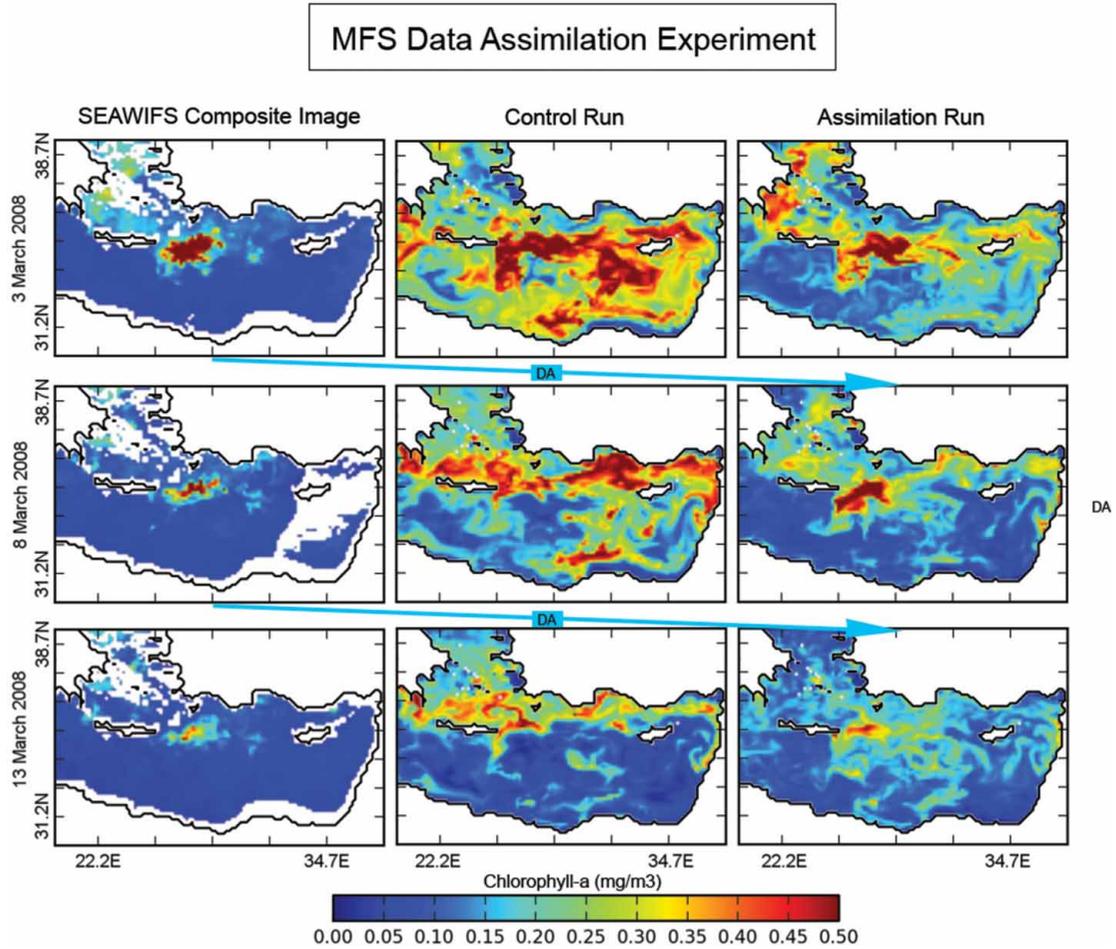


Figure 3. Comparison of surface chlorophyll as (1) observed by SeaWiFS (first column), (2) predicted by a hindcast (control run) (3) as in (2) but with surface chlorophyll-a data assimilation (Assimilation Run). Three model daily means (3, 8 and 13 March 2008) exhibit better predicting capabilities of the DA run than with the control run. The SeawiFS chlorophyll-a field is assimilated to produce the prediction of the next time (blue arrows), (modified after Brasseur & Verron, 2006).

Table 3. Uncertainty of Mediterranean biogeochemical reanalysis. Basin-wide bias (percentage) and correlation have been produced for the upper ocean (0–200 m, italics) and for the ocean interior (>200 m). The statistics are computed against *in-situ* profiles for dissolved oxygen, nitrate, phosphate, dissolved silica while surface Chl-a is compared with satellite observations. The best performances are indicated in bold [modified after Teruzzi et al (2014a)].

Variable	Number of match-ups	Model Bias (%)		Correlation	
		Hindcast	Reanalysis	Hindcast	Reanalysis
<i>Chlorophyll</i>	162368	<b>-0.85</b>	-337	0.60	<b>0.88</b>
<i>Dissolved Oxygen (0-200m)</i>	534	6.09	<b>5.46</b>	0.65	<b>0.67</b>
Dissolved Oxygen (200m-bottom)	1635	0.68	<b>0.22</b>	0.29	<b>0.33</b>
<i>Nitrate concentration (0-200m)</i>	2546	<b>-4.27</b>	27.37	0.21	<b>0.26</b>
Nitrate concentration (200m-bottom)	3593	-5.77	<b>5.17</b>	0.43	<b>0.46</b>
<i>Phosphate concentration (0-200m)</i>	4676	<b>39.78</b>	41.19	0.64	<b>0.65</b>
Phosphate concentration (200m-bottom)	6685	<b>6.00</b>	6.70	0.50	0.50
<i>Dissolved Silica concentration (0-200m)</i>	3536	-115.35	<b>-105.83</b>	0.21	<b>0.21</b>
Dissolved Silica concentration (200m-bottom)	6479	-38.36	<b>-36.93</b>	0.69	<b>0.70</b>

### Ocean carbon cycle

Accurate simulation of the marine carbon cycle, particularly air–sea CO<sub>2</sub> fluxes, is desirable for ocean carbon cycle monitoring and reanalysis purposes. As such, the aim of producing a reanalysis with FOAM-HadOCC based on the assimilation of ocean colour data is not just to simulate biological variables such as chlorophyll and primary production, but to assess the impact that assimilating satellite chlorophyll data have on carbon-cycle variables. When comparison is made with *in situ* observations of sea surface pCO<sub>2</sub> in the open ocean, the overall impact of the chlorophyll data assimilation on the model carbon cycle is small, with similar error statistics seen for both a free run and an assimilation run. This is due in part to pCO<sub>2</sub> being largely physically controlled. However, in regions where there is a strong biological component to the carbon cycle, for instance during the North Atlantic spring bloom, the chlorophyll assimilation can have a beneficial impact. An example is shown in Figure 4, which compares the June mean North Atlantic air–sea CO<sub>2</sub> flux from equivalent free and chlorophyll assimilation runs of FOAM-HadOCC with the climatology of Takahashi et al. (2009) for June. In the free run, the air–sea CO<sub>2</sub> flux is too weak across the northern part of the basin, with an area of spurious outgassing seen in the central North Atlantic. This error is much reduced in the chlorophyll assimilation run, which also shows increased CO<sub>2</sub> uptake in the northeast Atlantic, better matching the climatology. This demonstrates the potential of ocean colour assimilation to improve estimates of non-assimilated variables.

Next, a scheme to assimilate *in situ* pCO<sub>2</sub> observations has been developed by While et al. (2012) and implemented in FOAM-HadOCC. The assimilation was shown to reduce model pCO<sub>2</sub> errors, which has a direct impact on air–sea CO<sub>2</sub> fluxes. Impacts were also seen in the dissolved inorganic carbon and alkalinity fields. Where systematic pCO<sub>2</sub> errors were not reduced, it was concluded that this was due to sparse observational coverage in those regions. However, the assimilation had a long memory, with information from the

increments retained by the model for several months. This demonstrates that even a limited number of observations can have a substantial and lasting impact on the model results. In future, this method can be applied to larger pCO<sub>2</sub> observation databases such as the Surface Ocean Carbon Atlas (SOCAT) (Bakker et al. 2014) in order to produce long-term reanalyses. A development could be to assimilate physical, chlorophyll and pCO<sub>2</sub> data simultaneously, with cross-covariances or balance relationships between the physical and biogeochemical variables, thereby making maximum use of available observations.

In order to achieve successful pCO<sub>2</sub> forecasts, the estimation of poorly known static and globally constant biological model parameters is a necessary way ahead. However, affordable parameter estimation methods still need to be proven for non-linear models with positive-valued variables. An attempt was therefore made at NERSC to estimate two parameters of the MICOM-HAMOCC model (Assmann et al. 2010) using a local Ensemble Kalman Filter with state augmentation and Gaussian anamorphosis, similarly to the aforementioned HYCOM-NORWECOM system. The two model parameter values (namely phytoplankton growth rate and air–sea gas transfer) were also estimated locally. The surface pCO<sub>2</sub> observations from SOCAT were assimilated for 4 years 1999–2003 and did improve marginally, but systematically, the forecasts of pCO<sub>2</sub> (Simon & Bertino, 2013) (Figure 5). The proof of concept was only a partial success: a perfect twin experiment showed that the air–sea gas transfer coefficient could be retrieved everywhere, but the estimates of phytoplankton growth rate diverged in extra-tropical areas such as the North Pacific, where the nonlinearities seemed more severe than elsewhere, and denser observations would be necessary. For real-time forecasting of pCO<sub>2</sub>, the ultimate system should combine assimilation of physical data, chlorophyll data and surface pCO<sub>2</sub> as well as online estimation of a few model parameters, for which the scientific basis is being progressively established.

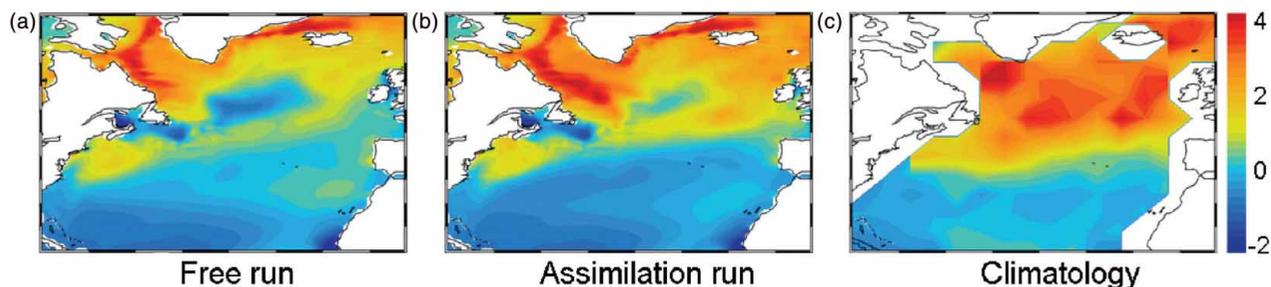


Figure 4. Air–sea CO<sub>2</sub> flux (mol m<sup>-2</sup> yr<sup>-1</sup>) in the North Atlantic from (a) a free run of FOAM-HadOCC (June mean, 1998–2012), (b) a chlorophyll-a assimilation run of FOAM-HadOCC (June mean, 1998–2012) and (c) the climatology of Takahashi et al. (2009) for June. Positive values denote a flux into the ocean.

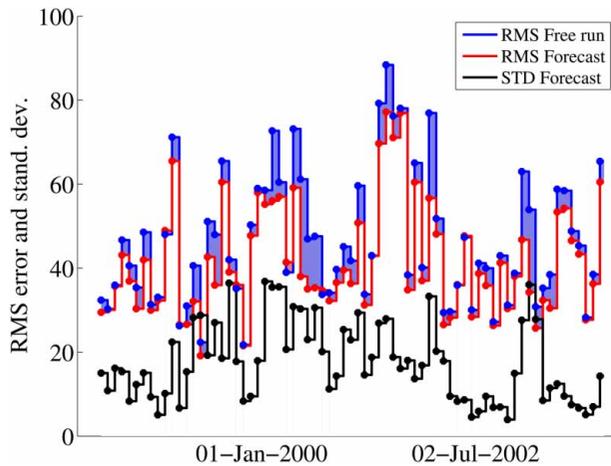


Figure 5. Time series of  $p\text{CO}_2$  RMS errors for a free run of MICOM-HAMOCC and the corresponding EnKF forecast, before assimilation of the SOCAT data. The shaded area shows the improvement gained by data assimilation. STD forecast refers to the EnKF uncertainty estimates.

### **Operational management of fisheries**

As illustrated throughout this paper, coupled physical–biogeochemical model systems are becoming increasingly skilful and starting to provide useful information to real-world applications, such as the operational management of fisheries. Preliminary steps towards this goal have been initiated, for example, in the case of the high-seas pelagic tuna fisheries. Operational physical model data are used to drive the spatial ecosystem and populations dynamics model SEAPODYM, to simulate functional groups of organisms at the intermediate trophic levels (Lehodey et al. 2010) and the dynamics of their predators, i.e. tuna (Lehodey et al. 2008). The model also includes a quantitative estimation approach (Senina et al. 2008) allowing the use of this model for fish stock assessment and testing management scenarios (Sibert et al. 2012).

The INDES0 (Infrastructure Development of Space Oceanography) project, which includes the development of an operational ecosystem and tuna model (SEAPODYM) in a regional Indonesian configuration, provides a proof of concept. A global operational configuration driven by Mercator–Ocean PSY3V3 model outputs (temperature and currents), and global primary production derived from satellite ocean colour data (Behrenfeld & Falkowski, 1997) provide the boundary conditions for the regional configuration. The first version of the model chain uses dynamical output from a regional configuration of the ocean general circulation model NEMO/OPA at  $1/12^\circ$  resolution and satellite derived NPP. In the near future, the regional coupled physical–biogeochemical model based on NEMO/PISCES will provide physical and biogeochemical forcings for SEAPODYM (e.g. temperature, salinity, euphotic depth, oxygen and

NPP). Environmental constraints on tuna population dynamics are complemented by best estimates of fishing mortality. The chain of production runs every week and delivers one week of hindcast, one week of nowcast and 10 days of forecast. These outputs will be used by the Fishing Authority to implement a better collection and verification of fishing data. The real-time and high-resolution configuration should help to quickly improve the model calibration and consequently abundance estimates needed to establish the optimal level of exploitation (total allowed catch) and the conservation measures (e.g. identification and protection of spawning grounds and nurseries) required for the sustainable exploitation of this essential resource.

Another example is given by Christensen et al. (2013), who, within the first MyOcean project, developed a short-term (seasonal up to a couple of years) forecasting system for sand-eel population based on larval dispersal and population analysis covering several life cycles, in order to support fisheries management providing information on the maximum sustainable yield and the effects of alternative fishing scenarios.

### **Future challenges, perspectives and blocking points**

#### **Biogeochemical data assimilation**

Today, ocean colour observations and derived products (e.g. surface estimates of chlorophyll concentration, NPP, particulate organic carbon) are the only source of data collected routinely at a global scale that provide a quantitative description of the surface signature of marine primary production at relevant space and time-scales. However, when the models reviewed in the previous section are used as deterministic tools to hindcast/forecast the biogeochemical properties of the ocean without data to constrain the simulations, they quite often display systematic biases as compared with ocean colour patterns observed from satellites. Misfits between ocean colour observational products and simulated field properties may have many different origins, e.g. inaccurate forcing fields (heat/momentum surface fluxes, photosynthetically active radiation fluxes, precipitations) used to drive the simulations, unrealistic physical model response in terms of advection/diffusion by ocean currents or unresolved spatial scales, oversimplified complexity of biogeochemical model structure and process formulation (e.g. no day–night cycle for photosynthesis) and uncertain parameterizations of interactions between biological species. Satellite-derived chlorophyll fields in turn suffer from a low signal-to-noise ratio, sensitivity to cloud cover and aerosols, as well as water transparency. They depend on (regional) empirical algorithms that transform the water-leaving radiance into average chlorophyll concentration of surface waters. As both models and observation products are uncertain, the assimilation

of data into coupled physical-biological models provides a useful means to reduce uncertainty in modelled properties and help identify the source of errors in modelling and data processing chains. Data assimilation into coupled models can be used to (1) refine the forcing functions of physical models (Skandrani et al. 2009) that drive biology, (2) improve the representation of mesoscale features into eddy-permitting or eddy-resolving models (e.g. Martin et al. 2015), (3) reduce the range of uncertainty of model parameters in biological interaction terms (Doron et al. 2013) or (4) produce statistical estimations of the biological fields (Fontana et al. 2013). Data assimilation, however, cannot be a substitute for model improvement, since data assimilation itself becomes more effective with model skill.

In essence, data assimilation into coupled models should be treated as a fully multivariate estimation problem, i.e. one piece of data related to physical or biological properties should impact the whole coupled state vector. However, as most coupled models do not include significant dynamical feedback from biology to physics (i.e. biology can be considered as mostly ‘driven’ by physics to a large extent), the assimilation can be treated as a decoupled problem. In this case, however, spatial mismatches between assimilated physical and biogeochemical fields could arise (i.e. front position), owing to the different data sets used in the assimilation processes. In addition, owing to the limited amount of information available to data assimilation, controlling many different error sources (in the physical or biological modelling components) at once is often impossible (Solidoro et al. 2003). The vast majority of operational systems provide physical fields with reduced uncertainty regarding horizontal advective fields, which is already very useful to reduce errors from the physics into coupled systems. However, the consistency of assimilated vertical advective or diffusive fluxes (associated with sub-grid scale parameterizations) may be unchanged or even degraded by data assimilation (Berline et al. 2006). This is recognised as a major problem with present assimilative systems that still requires dedicated research effort from the community.

In parallel, the assimilation of ocean colour data has received a growing attention by the research community during the past years. A variety of approaches have been tested [see the review by Gregg et al. (2009)] with some of them expanded today to (pre-)operational status (Ford et al. 2012; Fontana et al. 2013; Teruzzi et al. 2014a, 2014b). A key issue for future operational applications will be to build appropriate observation operators for data-assimilation systems to efficiently link remotely sensed ocean colour radiances with phytoplankton biomass concentrations from models by using optical modules as described in Fujii et al. (2007) and Ciavatta et al. (2014). Another key development will be to ensure consistent merging of ocean colour maps at the surface

with subsurface vertical profiles from a sparse Argo network equipped with nitrate, oxygen and fluorescence sensors.

Finally, it is worth noting that exploratory work has been initiated recently by applying the concept of image assimilation to ocean colour data (Gaultier et al. 2013). This could be a promising way to refine the control of eddies and associated mesoscale circulation from the physical space using ocean colour data assimilation into operational systems. It might also help to mitigate the frequent problem of the low persistence of data assimilation innovation impact. As a matter of fact, if the modification of the biogeochemical fields is not supported by a concomitant adjustment of the governing physical fields, often the biological signal quickly fades away (e.g. blooms driven by winter mixing).

### ***Global ocean biogeochemical reanalyses of the ocean colour era***

A key challenge for the future will be to develop a capacity to routinely deliver reanalyses of the biogeochemical state of oceans and regional seas, combining ocean colour and other relevant observations from satellites (e.g. SST, altimetry, surface winds, cloud coverage), *in situ* measurements (e.g. T/S profiles, fluorescence, oxygen, pCO<sub>2</sub>, nutrient concentrations) and coupled physical–biogeochemical models. Such products will be unique to monitor the impact of global change on marine primary production, regional ecosystems, the role of oceans in anthropogenic carbon storage, the occurrence statistics and trends of harmful algal blooms etc.

The ocean colour era started with the CZCS satellite mission in the late seventies, and since 1997 ocean colour has been monitored almost continuously by SeaWiFS, Envisat/MERIS and MODIS missions after an over 10-year gap between 1985 and 1997 (Wilson, 2010). In spite of significant progress in sensor technologies, algorithmic tools and spatial resolutions, the production of long-term time series needed to capture interannual to decadal variability still requires a sustained effort to develop and maintain the required ocean colour constellation and inter-calibration between instruments (Antoine et al. 2005). Additional opportunities for placing sensors on geostationary platforms in synergy with polar orbiting missions should be promoted (Ruddick et al. 2014) to enable the observation of the fine, mesoscale to submesoscale signature of surface phytoplankton and its associated diel variability. A key scientific challenge is indeed to incorporate the effect of those very fine space and time-scales on the local, basin and large-scale biogeochemical budgets estimated by integrative models (Lévy et al. 2012). In parallel to the space component, a denser network of *in situ* data from advanced autonomous platforms (e.g. BioArgo) will be required to extrapolate vertically the surface information

collected from ocean colour satellites into the entire euphotic layer through assimilative models (Claustre et al. 2010).

The sustained effort on biogeochemical monitoring has to be paralleled by a targeted development of biogeochemical models in order to allow the best possible utilisation of novel data. For example, current operational biogeochemical and lower trophic ecosystem models use very sophisticated treatments of the hydrodynamics and fairly sophisticated biology, but use grossly oversimplified treatments of the optics. Most models used for global and regional predictions are often driven by daily-averaged or hourly above-surface irradiances and very simple analytical formulations of light penetration through the water column. Some recent models have begun to include feedbacks between biology and hydrodynamics (Danabasoglu et al. 2006; Jochum et al. 2010), but still rely on very simple treatments of the optics. Fujii et al. (2007) incorporated spectrally resolved inherent optical properties (IOPs) into a coupled physical–biogeochemical model (Chai et al. 2002) connected with a radiative transfer model (Eco-Light-S) (Mobley, 2011). This fully coupled physical–biogeochemical–optical model simulates detailed and spectrally resolved underwater light field that not only can improve biological simulations but also can model explicitly biological feedbacks to ocean temperatures. The fully integrated model can be improved with direct comparison of model-predicted and satellite-measured remote-sensing reflectance, without the intermediate step of converting satellite data to chlorophyll. The model output also can be compared and improved with *in situ* optical measurements on various platforms, such as buoys, Argo floats, gliders or AUVs. Finally, all satellite and *in situ* measurements of optical variables can be assimilated into a fully coupled physical–biogeochemical–optical model to constrain ecosystem conditions and improve forecasts.

### ***Ocean carbon cycle***

Operational oceanography systems expanded to biogeochemistry have the potential to contribute to ocean carbon cycle research by providing (1) near-real-time monitoring of surface ocean pCO<sub>2</sub>, pH and CO<sub>2</sub> fluxes and (2) multi-year reanalyses of the ocean carbon cycle. These products will ultimately serve a variety of purposes, such as environmental monitoring (e.g. of pH), yearly global carbon cycle assessment (Global Carbon Project, <http://www.globalcarbonproject.org/>), and academic carbon cycle research. Since the first review of potential contributions of operational systems to carbon cycle research by Brasseur et al. (2009), the overall volume of carbon system observations (ALK, DIC, pCO<sub>2</sub>, CO<sub>2</sub> flux, pH) has increased, and new data compilations have been released. Despite this positive trend, data coverage is still

too sparse to consider operational carbon system data assimilation. Since pCO<sub>2</sub> is a function of salinity and temperature, improved distributions of these variables and of their temporal–spatial variability should translate into increased model skill for pCO<sub>2</sub>. Coupled biogeochemical–physical forecast systems could contribute to ocean carbon-cycle research by providing ocean physical-state estimates as forcing for the biogeochemical model component. Modelled pCO<sub>2</sub> is, however, very sensitive to errors in salinity and temperature. Errors introduced by data assimilation into ocean physics will be amplified in carbon-related variables such as pCO<sub>2</sub> (While et al. 2012). Shortcomings in current biogeochemical models are an additional source of error of modelled pCO<sub>2</sub>. Despite current limitations, the potential of assimilation of pCO<sub>2</sub> data into coupled physical–biogeochemical models for improving modelled pCO<sub>2</sub> distributions needs to be emphasized. Recent efforts to produce maps of surface ocean pCO<sub>2</sub> by means of statistical techniques (Landschützer et al. 2013) hold promise of an improved data availability in the years to come.

### ***Coupling up the food chain***

The current set of modelling systems for the higher trophic levels is described by a wide range of conceptual approaches and purposes. One strong focal point here is the development of systems built around the population of single species that are commercially important, such as the above-mentioned models for tuna and sand eel. On the other side of the spectrum, there are the whole-system approaches that can be largely divided into food-web and size-spectrum approaches (Blanchard et al. 2012) assessing ecosystem services in the wider sense. The food-web-based models (Shin et al. 2004; Christensen et al. 2009) consider extensive and complex food-web structures well beyond the complexity level of current LTE models but are consequently operating at coarser time and space scales than the current operational systems offer and that may be required for some important management applications. Size-spectrum models synthesize the energy transfer across the primary producers to top predators in a much lighter framework but lose information on taxonomical and functional diversity beyond size. A step forward here may be the combination of some of these approaches, similar to Fernandes et al. (2013), who combined a bio-climate envelope model with a size-spectrum approach, or Travers-Trolet et al. (2014), who coupled a regional physical model to an NPZD representation of the LTE and an individual-based multi-species higher trophic ecosystem model to assess the long-term impact of climate change on marine fish distributions.

Some major issues concerning the end-to-end modelling from biogeochemistry to top predators are largely still unaddressed (Rose et al. 2010). A main concern,

particularly in the operational context, is that all of the above approaches suffer from a substantial lack of data suitable for independent validation of their outputs in most areas of the global ocean, an essential requirement in order to demonstrate maturity for operational systems. Another key point is the identification of the zooplankton modules as the weak link in the coupled systems, as these modules were often developed with emphasis on the adequate description of the biogeochemical cycles, but do not give a good representation of their role as prey to higher-level predators.

Future research will consequently have to provide an adequate assessment of these frameworks and approaches as to which is most adequate for each operational context in terms of forecast precision, relevance of the information provided and the feasibility of implementation into operational systems in terms of computational requirements. A significant step in this direction is currently being pursued in the EU-FP7 Operational Ecology project (OpEc, [www.marine-opec.eu](http://www.marine-opec.eu)), where a series of near-operational forecasting systems for the physical, biogeochemical and ecological state of the European Seas are implemented and compared, bringing together the variety of the above-mentioned approaches in one framework that takes a strong focus on the operational delivery of relevant and reliable ecological indicators.

### Recommendations

Several demonstrations have been made recently, with Lazzari et al. (2010), Ford et al. (2012) and Fontana et al. (2013) showing that concepts of biogeochemical data assimilation using state-of-the-art BOGCMs are getting mature enough to be transitioned to operational centres. The capacity for producing biogeochemical forecasts, along with reanalyses, is a key requirement for the development of biogeochemical/ecological applications and services. However, several blocking points still need to be considered which require further research:

- the identification of the appropriate level of complexity that biogeochemical models need to include in their representation of biological processes to properly ‘ingest’ the observed data;
- the objective characterization of the uncertainty of BOGCMs associated with physical atmospheric and ocean circulation forcings, unresolved space and time-scales, and parameterizations of biogeochemical processes and fluxes;
- the coupling of BOGCMs with bio-optical models to assimilate the radiances observed by satellites in the context of ‘the situation of the day’ instead of derived products that often rely on climatologies (e.g. chlorophyll);
- the development of coupled physical–biogeochemical data-assimilation schemes;
- the development of ocean colour data-assimilation methodologies that will be able to tackle highly non-linear model responses and non-Gaussian probability distributions associated with ensemble coupled physical–biological models, as well as, in the case of multivariate assimilation schemes, to ensure appropriate projection of observed fields onto unobserved state variables using physically balanced increments (While et al. 2012);
- the specification of relevant sets of metrics to assess the scientific value of operational products, the development of intercomparison exercises between operational groups and the identification of weaknesses in the observing systems;
- the use of Observing System Simulation Experiments to inform on the optimal architecture of an observing system designed for biogeochemical data assimilation and to identify biogeochemical products that can be robustly provided;
- improvement in biogeochemical datasets for better initial conditions and model validation.

### Supplemental data and research materials

Supplemental data for this article will be hosted alongside the article on [tandfonline.com](http://tandfonline.com) and can be accessed at <http://dx.doi.org/10.1080/1755876X.2015.1022350>.

### Disclosure statement

No potential conflict of interest was reported by the author(s).

### Funding

M.G., P.B., C.P., Y.D. and P.L. acknowledge support by the French research programme LEFE-GMMC under the grant agreement no. 452 PPR-GREEN-MERCATOR-II. R.B. and D.F. would like to thank ACRI for providing the GlobColour chlorophyll data and acknowledge funding support from the UK Ministry of Defence and from the European Space Agency for funding the Climate Modelling User Group (Contract 4000100222/10/I-AM) as part of its Climate Change Initiative. The European Community’s Seventh Framework Programme FP7 (2007–2013) provided support under grant agreement no. 283367 (MyOcean2) to P.B., R.B., D.F., L.B., A.S., C.P., Y.D. and A.C., and under grant agreement no. 264879 (CarboChange) to L.B., E.S. and M.G.

### References

- Allen IJ, Somerfield PJ, Gilbert FJ. 2007. Quantify uncertainty in high-resolution coupled hydrodynamic-ecosystem models. *J Mar Sys.* 60:3–14.
- Anderson TR. 1993. A spectrally averaged model of light penetration and photosynthesis. *Limnol Oceanogr.* 38:1403–1419.

- Antoine D, Morel A, Gordon H, Banzon V, Evans R. 2005. Bridging ocean color observations of the 1980s and 2000s in search of long-term trends. *J Geophys Res.* 110:C06009. doi:10.1029/2004JC002620
- Artioli Y, Blackford JC, Butenschön M, Holt J, Wakelin SL, Thomas H, Borges AV, Allen JI. 2012. The carbonate system in the North Sea: Sensitivity and model validation. *J Mar Syst.* 102–104:1–13.
- Assmann KM, Bentsen M, Segsneider J, Heinze C. 2010. An isopycnic ocean carbon cycle model. *Geosci Model Dev.* 2:142–167.
- Aumont O, Bopp L. 2006. Globalizing results from ocean in situ iron fertilization studies. *Global Biogeochem Cycles.* 20:GB2017 doi:10.1029/2005GB002591.
- Aumont O, Bopp L, Schulz M. 2008. What does temporal variability in aeolian dust deposition contribute to sea-surface iron and chlorophyll distributions? *Geophys Res Lett.* 35:L07607. doi: 10.1029/2007GL031131
- de Baar HJW, de Jong JTM. 2001. Distributions, sources and sinks of iron in seawater. In: Turner D, Hunter K, editor. *The biogeochemistry of iron in seawater.* Hoboken, N.J.: John Wiley; p. 85–121.
- Bakker DCE, Pfeil B, Smith K, Hankin S, Olsen A, Alin SR, Cosca C, Harasawa S, Kozyr A, Nojiri Y, O'Brien KM, Schuster U, Telszewski M, Tilbrook B, Wada C, Akl J, Barbero L, Bates NR, Boutin J, Bozec Y, Cai W-J, Castle RD, Chavez FP, Chen L, Chierici M, Currie K, De Baar HJ, Evans W, Feely RA, Fransson A, Gao Z, Hales B, Hardman-Mountford NJ, Hoppema M, Huang W-J, Hunt CW, Huss B, Ichikawa T, Johannessen T, Jones EM, Jones S, Jutterstrøm S, Kitidis V, Körtzinger A, Landschützer P, Lauvset SK, Lefèvre N, Manke AB, Mathis JT, Merlivat L, Metzl N, Murata A, Newberger T, Omar AM, Ono T, Park G-H, Paterson K, Pierrot D, Rios AF, Sabine CL, Saito S, Salisbury J, Sarma VVSS, Schlitzer R, Sieger R, Skjelvan I, Steinhoff T, Sullivan KF, Sun H, Sutton AJ, Suzuki T, Sweeney C, Takahashi T, Tjiputra J, Tsurushima N, Van Heuven SMAC, Vandemark D, Vlaho P, Wallace DWR, Wanninkhof R, Watson AJ. 2014. An update to the Surface Ocean CO<sub>2</sub> Atlas (SOCAT version 2). *E Sys Sci Data.* 6:69–90. doi:10.5194/essd-6-69-2014
- Barciela R, Brasseur P, Wilmer-Becker K, Marine Ecosystem Prediction Task Team (MEAP-TT) report. 2012. Available from: <https://www.godac-oceanview.org/>.
- Baretta JW, Ebenhöf W, Ruardij P. 1995. The European regional seas ecosystem model, a complex marine ecosystem model. *Netherlands J Sea Res.* 33:233–246.
- Baretta-Bekker JG, Baretta JW, Ebenhöf W. 1997. Microbial dynamics in the marine ecosystem model ERSEM II with decoupled carbon assimilation and nutrient uptake. *J Sea Res.* 38:195–211.
- Behrenfeld MJ, Falkowski PG. 1997. A consumer's guide to phytoplankton primary productivity models. *Limnol. Oceanogr.* 42:1479–1491.
- Berline L, Brankart JM, Brasseur P, Ourmières Y, Verron J. 2006. Improving the dynamics of a coupled physical–biogeochemical model of the North Atlantic basin through data assimilation: Impact on biological tracers. *J Mar Syst.* 64:153–172. doi:10.1016/j.marsys.2006.03.007
- Beusen AHW, Bouwman AF, Durr HH, Dekkers ALM, Hartmann J. 2009. Global patterns of dissolved silica export to the coastal zone: Results from a spatially explicit global model. *Glob Biogeochem Cycles.* 23:GB0A02. doi:10.1029/2008GB003281
- Blackford JC, Allen JI, Gilbert FJ. 2004. Ecosystem dynamics at six contrasting sites: A generic modelling study. *J Marine Syst.* 52:191–215. doi:10.1016/j.jmarsys.2004.02.004
- Blanchard JL, Jennings S, Holmes R, et al. 2012. Potential consequences of climate change for primary production and fish production in large marine ecosystems. *Philosophical Trans of the Royal Society B: Biological Sciences.* 367:2979–2989.
- Bleck R. 2002. An oceanic general circulation model framed in hybrid isopycnic-cartesian coordinates. *O Model* 4:55–88.
- Blockley EW, Martin MJ, McLaren AJ, Ryan AG, Waters J, Lea DJ, Mirouze I, Peterson KA, Sellar A, Storkey D. 2014. Recent development of the Met Office operational ocean forecasting system: An overview and assessment of the new Global FOAM forecasts. *Geosci Model Dev.* 7:2613–2638, doi:10.5194/gmd-7-2613-2014.
- Bloom SC, Takacs LL, Da Silva AM, Ledvina D. 1996. Data assimilation using incremental analysis updates. *Monthly Weat Rev.* 124:1265–1271.
- Bopp L, Aumont O, Cadule P, Alvain S, Gehlen M. 2005. Response of diatoms distribution to global warming and potential implications: A global model study. *Geophys Res Lett.* 32:L19606. doi:10.1029/2005GL023653
- Bopp L, Resplandy L, Orr JC, Doney SC, Dunne JP, Gehlen M, Halloran P, Heinze C, Ilyina T, Séférian R, Tjiputra J, Vichi M. 2013. Multiple stressors of ocean ecosystems in the 21st century: Projections with CMIP5 models. *Biogeosciences.* 10:6225–6245.
- Brasseur P, Gruber N, Barciela R, Brander K, Doron M, El Moussaoui A, Hobday AJ, Huret M, Kremer A-S, Lehodey P, Matear R, Moulin C, Murtugudd R, Senina I, Svendsen E. 2009. Integrating biogeochemistry and ecology into ocean data assimilation systems. *Oceanography.* 22:206–215.
- Brasseur P, Verron J. 2006. The SEEK filter method for data assimilation in oceanography: A synthesis. *Ocean Dynamics.* 56:650–661.
- Brickman D, Drozdowski A. 2012. Development and validation of a regional shelf model for Maritime Canada based on the NEMO-OPA circulation model. *Can Tech Rep Hydrogr Ocean Sci.* 278:vii+57pp.
- Chai F, Dugdale RC, Peng T-H, Wilkerson FP, Barber RT. 2002. One dimensional ecosystem model of the equatorial Pacific upwelling system, part I: Model development and Silicon and Nitrogen cycle. *Deep-Sea Res. II.* 49:2713–2745.
- Christensen A, Butenschön M, Gürkan Z, Allen JI. 2013. Towards an integrated forecasting system for fisheries on habitat-bound stocks. *Ocean Sci.* 9:261–279.
- Christensen A, Jensen H, Mosegaard M. 2009. Spatially resolved fish population analysis for designing MPAs: Influence on inside and neighbouring habitats. *Ices J Mar. Sci.* 66:56–63.
- Ciavatta S, Torres R, Martinez-Vicente V, Smyth T, Dall'Olmo G, Polimene L, Allen JI. 2014. Assimilation of remotely-sensed optical properties to improve marine biogeochemistry modelling. *Prog Oceanogr.* Available from: <http://dx.doi.org/10.1016/j.pocean.2014.06.002>, in press.
- Claustre H, et al. 2010. Bio-optical profiling floats as new observational tools for biogeochemical and ecosystem studies: potential synergies with ocean color remote sensing. in *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2)*, Venice, Italy, 21–25 September 2009, Hall, J., Harrison, D.E. & Stammer, D., Eds., ESA Publication WPP-306. doi:10.5270/OceanObs09.cwp.17

- Danabasoglu G, Large WG, Tribbia JJ, Gent PR, Briegleb BP, McWilliams JC. 2006. Diurnal coupling in the tropical oceans of CCSM3. *J Climate*. 19:2347–2365.
- Dickson AG, Sabine CL, Christian JR, editors. 2007. Guide to best practices for ocean CO<sub>2</sub> measurements. PICES Special Publication 3, 191pp.
- Dobricic S, Pinardi N. 2008. An oceanographic three-dimensional variational data assimilation scheme. *Ocean Modelling*. 22:89–105. doi:10.1016/j.ocemod.2008.01.004
- Doron M, Brasseur P, Brankart J-M, Losa S, Melet A. 2013. Stochastic estimation of biogeochemical parameters from Globcolour ocean colour satellite data in a North Atlantic 3D coupled physical-biogeochemical model. *J Marine Syst*. 117–118:81–95.
- Droop MR. 1973. Some thoughts on nutrient limitation in algae. *J Phycol*. 9:264–272.
- Edwards KP, Barciela R, Butenschön M. 2012. Validation of the NEMO-ERSEM operational ecosystem model for the North West European Continental Shelf. *Ocean Sci*. 8:983–1000.
- El Moussaoui A, Perruche C, Greiner E, Ethé C, Gehlen M. 2011. Integration of biogeochemistry into Mercator Ocean systems. *M Oce newslet*. 40:3–14.
- Engel A, Goldthwait S, Passow U, Alldredge A. 2002. Temporal decoupling of carbon and nitrogen dynamics in a mesocosm diatom bloom. *Limnol and Oceanogra*. 47:753–761.
- Eppley RW. 1972. Temperature and phytoplankton growth in the sea. *Fish. B.-NOAA*. 70:1063–1085.
- Fernandes JA, Cheung WW, Jennings S, Butenschön M, Mora L, Frölicher TL, Barange M, Grant A. 2013. Modelling the effects of climate change on the distribution and production of marine fishes: Accounting for trophic interactions in a dynamic bioclimate envelope model. *Glo Change Biology*. 19:2596–2607.
- Follows MJ, Dutkiewicz S, Grant S, Chisholm SW. 2007. Emergent bio-geography of microbial communities in a model ocean. *Science*. 315:1843–1846.
- Fontana C, Brasseur P, Brankart J-M. 2013. Toward a multivariate reanalysis of the North Atlantic ocean biogeochemistry between 1998–2006 based on the assimilation of SeaWiFS data. *O Sci*. 9:1–21.
- Ford DA, Edwards KP, Lea D, Barciela RM, Martin MJ, Demaria J. 2012. Assimilating GlobColour ocean colour data into a pre-operational physical-biogeochemical model. *Ocean Sci*. 8:751–771.
- Foujols M, Lévy M, Aumont O, Madec G. 2000. OPA8.1 tracer model reference manual, inst. Paris: Pierre Simon Laplace.
- Friedrichs MAM, Dusenberry JA, Anderson LA, Armstrong RA, Chai F, Christian JR, Doney SC, Dunne J, Fujii M, Hood R, McGillicuddy DJ Jr., Moore JK, Schartau M, Spitz YH, Wiggert JD. 2007. Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups. *J Geophys Res*. 112:C08001. doi:10.1029/2006JC003852
- Fujii M, Boss E, Chai F. 2007. The value of adding optics to ecosystem models: A case study. *Biogeosci Disc*. 4:1585–1631.
- Gaultier L, Verron J, Brankart J-M, Titau O, Brasseur P. 2013. On the inversion of submesoscale tracer fields to estimate the surface ocean circulation. *J Marine Systems*. 126:33–42. doi:10.1016/j.jmarsys.2012.02.014
- Gehlen M, Gruber N, Gangstø R, Bopp L, Oschlies A. 2011. Biogeochemical consequences of ocean acidification and feedback to the Earth system. In: Gattuso J.-P., Hansson L, editor. *Ocean acidification* Oxford: Oxford University Press; p. 230–248.
- Geider RJ, MacIntyre HL, Kana TM. 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. *Mar Ecol.-Prog. Ser*. 148:187–200.
- Geider RJ, MacIntyre HL, Kana TM. 1998. A dynamic regulatory model of phytoplankton acclimation to light, nutrients and temperature. *Limnol. Oceanogr*. 43:679–694.
- Gregg WW, Friedrichs MA, Robinson AR, Rose KA, Schlitzer R, Thompson KR, Doney SC. 2009. Skill assessment in ocean biological data assimilation. *J Mar Syst*. 76:16–33. doi:10.1016/j.jmarsys.2008.05.006
- Hansen C, Samuelson A. 2009. Influence of horizontal model grid resolution on the simulated primary production in an embedded primary production model in the Norwegian Sea. *J Mar Syst*. 75:236–244. doi:10.1016/j.jmarsys.2008.10.004
- Hemmings JCP, Barciela RM, Bell MJ. 2008. Ocean color data assimilation with material conservation for improving model estimates of air-sea CO<sub>2</sub> flux. *J Mar. Res*. 66:87–126.
- Hood RR, Laws EA, Armstrong RA, Bates NR, Brown CW, Carlson CA, Chai F, Doney SC, Falkowski PG, Feely RA, Friedrichs MAM, Landry MR, Moore JK, Nelson DM, Richardson TL, Salihoglu B, Schartau M, Toole DA, Wiggert JD. 2006. Pelagic functional group modelling: progress, challenges and prospects. *Deep-Sea Res. II*. 53:459–512.
- Hunke EC, Lipscomb WH. 2010. CICE: The Los Alamos sea ice model. Documentation and software user's manual, Version 4.1 (LA-CC-06-012), T-3 Fluid Dynamics Group, Los Alamos National Laboratory, Los Alamos.
- Jin X, Najjar RG, Louanchi F, Doney SC. 2007. A modeling study of the seasonal oxygen budget of the global ocean. *J Geophys Res*. 112:C05017. doi: 10.1029/2006JC003731
- Jochum M, Yeager S, Lindsay K, Moore K, Murtugudde R. 2010. Quantification of the feedback between phytoplankton and ENSO in the community climate system model. *J Clim*. 23:2916–2925.
- Johnson KS, Chavez FP, Friederich GE. 1999. Continental-shelf sediment as a primary source of iron for coastal phytoplankton. *Nature*. 398:697–700.
- Keeling RF, Körtzinger A, Gruber N. 2010. Ocean deoxygenation in a warming world. *Annu Rev Mar Sci*. 2:199–229.
- Kwiatkowski L, Yool A, Allen JJ, Anderson TR, Barciela R, Buitenhuis ET, Butenschön M, Enright C, Halloran PR, Le Quéré C, de Mora L, Racault M-F, Sinha B, Totterdell IJ, Cox PM. 2014. iMarNet: An ocean biogeochemistry model inter-comparison project within a common physical ocean modelling framework. *Biogeosciences* 11:7291–7304. doi:10.5194/bg-11-7291-2014
- Lambert N, Chassé J, Perrie W, Long Z, Guo L, Morrison J. 2013. Projection of future river runoffs in Eastern Atlantic Canada from global and regional climate models. *Can Tech Rep Hydrogr Ocean Sci*. 288:viii+34.
- Landschützer P, Gruber N, Bakker DCE, Schuster U, Nakaoka S, Payne MR, Sasse TP, Zeng J. 2013. A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink. *Biogeosci*. 10:7793–7815. doi:10.5194/bg-10-7793-2013
- Lazzari P, Solidoro C, Ibello V, Salon S, Teruzzi A, Béranger K, Colella S, Crise A. 2012. Seasonal and inter-annual variability of plankton chlorophyll and primary production in the

- Mediterranean Sea: A modelling approach. *Biogeosciences*. 9:217–233.
- Lazzari P, Teruzzi P, Salon A. 2010. Pre-operational short-term forecasts for Mediterranean Sea biogeochemistry. *O Sci*. 6:25–39.
- Le Fouest V, Zakardjian B, Saucier FJ, Çizmeli SA. 2006. Application of SeaWiFS- and AVHRR-derived data for mesoscale and regional validation of a 3-D high-resolution physical-biological model of the Gulf of St. Lawrence (Canada). *J Mar Syst*. 60:30–50.
- Le Fouest V, Zakardjian B, Saucier FJ, Starr M. 2005. Seasonal versus synoptic variability in planktonic production in a high-latitude marginal sea: The Gulf of St. Lawrence (Canada). *J Geophys Res*. 110:C09012. doi: 10.1029/2004JC002423
- Lehodey P, Murtugudde R, Senina I. 2010. Bridging the gap from ocean models to population dynamics of large marine predators: a model of mid-trophic functional groups. *Prog in Oceanogra*. 84:69–84.
- Lehodey P, Senina I, Murtugudde R. 2008. A spatial ecosystem and populations dynamics model (SEAPODYM) – modelling of tuna and tuna-like populations. *Prog in Oceanogra*. 78:304–318.
- Lellouche J-M, Le Galloudec O, Drevillon, Regnier C, Greiner E, Garric G, Ferry N, Desportes C, Testut C-E, Bricaud C, Bourdalle-Badie R, Tranchant B, Benkiran M, Drillet Y, Daudin A, De Nicola C. 2013. Evaluation of global monitoring and forecasting systems at Mercator Océan. *Ocean Sci*. 9:57–81. doi:10.5194/os-9-57-2013
- Le Quéré C, Harrison SP, Prentice IC, Buitenhuis ET, Aumont O, Bopp L, Claustre H, Da Cunha LC, Geider R, Giraud X, Klaas C, Kohfeld KE, Legendre L, Manizza M, Platt T, Rivkin RB, Sathyendranath S, Uitz J, Watson AJ, Wolf-Gladrow D. 2005. Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Glo Change Biology*. 11:2016–2040.
- Lévy M Ferrari R, Franks P, Martin A, Rivière P. 2012. Bringing physics to life at the submesoscale. *Geophys. Res. Lett*. 39. doi:10.1029/2012GL052756
- Lin P, Rubao Ji R, Davis CS, McGillicuddy DJ Jr. 2010. Optimizing plankton survey strategies using observing system simulation experiments. *J Mar Syst*. 82:187–194.
- Ludwig W, Probst J-L, Kempe S. 1996. Predicting the oceanic input of organic carbon by continental erosion. *Global Biogeochem Cycles*. 10:23–41.
- Madec G. 2012. NEMO ocean engine. Note du Pole de modélisation de l'Institut Pierre-Simon Laplace. 27:367pp.
- Madec G, Delecluse P, Imbard M, Lévy C. 1998. LIM2 ice model, in OPA 8.1 Ocean general circulation model reference manual. Institut Pierre-Simon Laplace (IPSL), 91pp, France.
- Martin M, Balmaseda M, Bertino L, Brasseur P, Brassington G, Cummings J, Fujii Y, Lea DJ, Lellouche J-M, Morgensen K, Oke P, Smith GC, Testut C-E, Waagbo GA, Waters J, Weaver AT. 2015. Status and future of data assimilation in operational oceanography. *J Opera Oceanography*, doi:10.1080/1755876X.2015.1022055.
- Martin MJ, Hines A, Bell MJ. 2007. Data assimilation in the FOAM operational short-range ocean forecasting system: a description of the scheme and its impact. *Q.J.R. Meteorol. Soc*. 133:981–995. doi:10.1002/qj.74
- Martin JH, Knauer GA, Karl DM, Broenkow WW. 1987. VERTEX: Carbon cycling in the Northeast Pacific. *Deep Sea Research*. 34:267–285.
- Mobley CD. 2011. Fast light calculations for ocean ecosystem and inverse models. *Optics Express*. 19(20):18927–18944. doi:10.1364/OE.19.018927.
- Moll A, Stegert C. 2007. Modelling Pseudocalanus elongatus stage-structured population dynamics embedded in a water column ecosystem model for the northern North Sea. *J Mar Syst*. 64:35–46.
- Moore JK, Doney SC, Kleypas JA, Glover DM, Fung IY. 2002. An intermediate complexity marine ecosystem model for the global domain. *Deep-Sea Res. II*. 49:403–462. doi:10.1016/S0967-0645(01)00108-4
- O'Dea EJ, Arnold AK, Edwards KP, Furner R, Hyder P, Martin MJ, Siddom JR, Storkey D, While J, Holt JT, Liu H. 2012. An operational ocean forecast system incorporating NEMO and SST data assimilation for the tidally driven European North-West shelf. *J Opera Oceanography*. 5:3–17.
- Oki T, Sud YC. 1998. Design of total runoff integrating pathways (TRIP) – a global river channel network. *Earth Interact*. 2:1–37.
- Orr JC. 1999. On ocean carbon-cycle model comparison. *Tellus B*. 51:509–530.
- Orr JC, Fabry VJ, Aumont O, et al. 2005. Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature*. 437:681–686.
- Ourmieres Y, Brankart JM, Berline L, Brasseur P, Verron J. 2006. Incremental analysis update implementation into a sequential ocean data assimilation system. *Journal of Atmospheric and Oceanic Techno*. 23:1729–1744.
- Palmer JR, Totterdell IJ. 2001. Production and export in a global ocean ecosystem model. *Deep-Sea Res. I*. 48:1169–1198.
- Pätsch J, Kühn W, Moll A, Lenhar H. 2009. ECOHAM4 user guide – ecosystem model, hamburg, Version 4. Technical Reports 01-2009. Hamburg, Germany: Institut für Meereskunde.
- Pellerin G, Lefavre L, Houtekamer P, Girard C. 2003. Increasing the horizontal resolution of ensemble forecasts at CMC. *Nonlin Proc in Geophy*. 10:463–468.
- Raine R, McDermott G, Silke J, Lyons K, Nolan G, Cusack C. 2010. A simple short range model for the prediction of harmful algal events in the bays of southwestern Ireland. *J Mar Syst*. 83:150–157.
- Redfield AC, Ketchum H, Richards FA. 1963. The influence of organisms on the composition of seawater. In: Hill MN, editor. *The sea*. vol. 2. New York: Wiley Interscience; p. 26–77.
- Rose KA, Allen JI, Artioli Y, Barange M, Blackford J, Carlotti F, Cropp R, Daewel U, Edwards K, Flynn K, Hill SL, Lambers HR, Huse G, Mackinson S, Megrey B, Moll A, Rivkin R, Salihoglu B, Schrum C, Shannon L, Shin Y-J, Smith SL, Smith C, Solidoro C, St. John M, Zhou M. 2010. End-to-end models for the analysis of marine ecosystems: Challenges, issues, and next steps. *Mar. Coast. Fish. Dyn. Manag. Ecosyst Sci*. 2:115–130. doi:10.1577/C09-059.1
- Ruddick K, Neukermans G, Vanhellefont Q, Jolivet D. 2014. Challenges and opportunities for geostationary ocean colour remote sensing of regional seas: A review of recent results. *Remote Sensing Env*. 146:63–76.
- Ryan AG, Regnier C, Divakaran P, Spindler T, Mehra A, Smith GC, Davidson F, Hernandez F, Liu Y. 2015. GODAE OceanView Class 4 forecast verification framework: Global ocean intercomparison. doi:10.1080/1755876X.2015.1022330.
- Samuelson A, Bertino L, Hansen C. 2009. Impact of data assimilation of physical variables on the spring bloom from TOPAZ operational runs in the North Atlantic. *Ocean Sci*. 5:635–647.

- Schneider B, Bopp L, Gehlen M, et al. 2008. Climate-induced interannual variability of marine primary and export production in three global coupled climate carbon cycle models. *Biogeosciences*. 5:597–614.
- S  ferian R, Bopp L, Gehlen M, et al. 2012. Skill assessment of three earth system model with common marine biogeochemistry. *Clim Dyn.*: 1–25. doi:10.1007/s00382-012-1362-8
- Seitzinger SP, Harrison JA, Dumont E, Beusen AHW, Bouwman AF. 2005. Sources and delivery of carbon, nitrogen and phosphorus to the coastal zone: An overview of Global Nutrient Export from Watersheds (NEWS) models and their application. *Global Biogeochem. Cycles*. 19:GB4S01. doi:10.1029/2005GB002606
- Senina I, Sibert J, Lehodey P. 2008. Parameter estimation for basin-scale ecosystem-linked population models of large pelagic predators: Application to skipjack tuna. *Prog in Oceanogra*. 78:319–335.
- Shin Y-J, Shannon LJ, Curt PM. 2004. Simulations of fishing effects on the southern Benguela fish community using an individual-based model: Learning from a comparison with ECOSIM. *Afri J Mar Sci*. 26:95–114.
- Sibert J, Senina I, Lehodey P, Hampton J. 2012. Shifting from marine reserves to maritime zoning for conservation of Pacific bigeye tuna (*Thunnus obesus*). *Proce of the Nat Acad of Sci*. 109:18221–18225.
- Signorini SR, McClain CR, Christian JR, Wong CS. 2001. Seasonal and interannual variability of phytoplankton, nutrients, TCO<sub>2</sub>, pCO<sub>2</sub>, and O<sub>2</sub> in the eastern subarctic Pacific (ocean weather station Papa). *J Geophys Res*. 106:31197–31215.
- Simon E, Bertino L. 2009. Application of the Gaussian anamorphosis to assimilation in a 3-D coupled physical-ecosystem model of the North Atlantic with the EnKF: A twin experiment. *Ocean Sci*. 5:495–510.
- Simon E, Bertino L. 2012. Gaussian anamorphosis extension of the DEnKF for combined state and parameter estimation: Application to a 1-D ecosystem model. *J Mar Syst*. 89:1–18.
- Simon E, Bertino L. 2013. Optimized parameters of mathematical formulation with respect to observations. First attempt to optimize biogeochemical-governing parameters through EnKF. SKD Biofeedback project deliverable D2.3. Bjerknes Center, Bergen, Norway.
- Skandrani S, Brankart J-M, Ferry N, Verron J, Brasseur P, Barnier B. 2009. Controlling atmospheric forcing parameters of global ocean models: Sequential assimilation of sea surface Mercator-Ocean reanalysis data. *O Sci*. 5:403–419.
- Skogen MD, Budgell WP, Rey F. 2007. Interannual variability in Nordic seas primary production. *ICES J Mar. Sci*. 64:889–898.
- Skogen M, S  iland H. 1998. A user's guide to NORWECOM v2.0. The NORwegian ECOlogical model system. Bergen: Ins Mar Res.
- Skogen MD, Svendsen E, Berntsen J, Aksne D, Ulvestad KB. 1995. Modelling the primary production in the North Sea using a coupled three-dimensional physical-chemical-biological ocean model. *Estuar. Coast. Shelf Sci*. 41:545–565. doi:10.1016/0272-7714(95)90026-8
- Solidoro C, Crise A, Crispi G, Pastres R. 2003. An a priori approach to assimilation of ecological data in marine ecosystem models. *J Mar Syst*. 40–41:79–97.
- Stegert C, Moll A, Kreis M. 2009. Validation of the three-dimensional ECOHAM model in the German Bight for 2004 including population dynamics of *Pseudocalanus elongatus*. *J Sea Res*. 62:1–15.
- Stock CA, Alexander MA, Bond NA, et al. 2011. On the use of IPCC-class models to assess the impact of climate on living marine resources. *Prog in Oceanogra*. 88:1–27.
- Storkey D, Blockley EW, Furner R, Guiavarc'h C, Lea D, Martin MJ, Barciela RM, Hines A, Hyder P, Siddorn JR. 2010. Forecasting the ocean state using NEMO: The new FOAM system. *J Opera Oceanography*. 3:3–15.
- Stumpf RP, Tomlinson MC, Calkins JA, Kirkpatrick B, Fisher K, et al. 2009. Skill assessment for an operational algal bloom forecast system. *J Mar Syst*. 76:151–161.
- Sykes PA, Barciela RM. 2012. Assessment and development of a sediment model within an operational system. *J Geophys Res*. 117:C04036. doi:10.1029/2011JC007420
- Takahashi T, Sutherland SC, Wanninkhof R, Sweeney C, Feely RA, Chipman DW, Hales B, Friederich G, Chavez F, Watson A, Bakker DCE, Schuster U, Metzl N, Yoshikawa-Inoue H, Ishii M, Midorikawa T, Nojiri Y, Sabine C, Olafsson J, Arnarson TS, Tilbrook B, Johannessen T, Olsen A, Bellerby R, K  rtzinger A, Steinhoff T, Hoppema M, de Baar HJW, Wong CS, Delille B, Bates NR. 2009. Climatological mean and decadal changes in surface ocean pCO<sub>2</sub>, and net sea-air CO<sub>2</sub> flux over the global oceans. *Deep-Sea Res. II*. 56:554–577.
- Teruzzi A, Cossarini G, Salon S, Bolzon G. 2014b. Quality information document on Mediterranean Sea biogeochemistry reanalysis product MYO2-MED-QUID-006-007. Issue: 4.0    MyOcean.
- Teruzzi A, Dobricic S, Solidoro C, Cossarini G. 2014a. A 3-D variational assimilation scheme in coupled transport-biogeochemical models: Forecast of Mediterranean biogeochemical properties. *J Geophy Res, Oceans*. 119:200–217, doi:10.1002/2013JC009277
- Tonani M, Balmaseda M, Bertino L, Blockley E, Brassington G, Davidosn F, Drillet Y, Hogan P, Kurano T, Lee T, Mehra A, Paranathara F, Tanajura CAS, Wang H. 2015. Status and future of global and regional prediction systems. *J Opera Oceanography*. forthcoming.
- Travers-Trolet M, Shin Y-J, Shannon LJ, Moloney CL, Field JG. 2014. Combined fishing and climate forcing in the Southern Benguela upwelling ecosystem: An end-to-end modelling approach reveals dampened effects. *PLoS ONE*. 9:e94286. doi:10.1371/journal.pone.0094286
- Vichi M, Cossarini G, Gutierrez Mlot E, Lazzari P, Lovato T, Mattia G, Masina S, McKiver W, Pinardi N, Solidoro C, Zavatarelli M. 2013. The biogeochemical flux model (BFM): equation description and user manual, BFM version 5 (BFM-V5). Release 1.0; BFM Report series N. 1, Bologna, Italy (2013), p. 87. Available from: <http://bfm-community.eu>
- Vichi M, Masina S, Navarra A. 2007b. A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part II: Numerical simulations. *J Mar Sys*. 64:110–134.
- Vichi M, Masina S, Pinardi N. 2007a. A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part I. Theory. *J Mar Syst*. 64:89–109.
- Ward BA, Schartau M, Andreas Oeschlies A, Martin AP, Michael JF, Anderson TR. 2013. When is a biogeochemical model too complex? Objective model reduction and selection for North Atlantic time-series sites. *Prog in Oceanogra*. 116:49–65. ISSN 0079-6611. Available from: <http://dx.doi.org/10.1016/j.pocean.2013.06.002>
- Waters J, Lea DJ, Martin MJ, Mirouze I, Weaver AT, While J. 2014. Implementing a variational data assimilation system in an operational 1/4 degree global ocean model. *Q.J.R. Meteorol Soc*. doi:10.1002/qj.2388

- While J, Totterdell IJ, Martin MJ. 2012. Assimilation of pCO<sub>2</sub> data into a global coupled physical–biogeochemical ocean model. *J Geophys Res.* 117:C03037. doi:10.1029/2010JC006815
- Wilson C. 2010. The rocky road from research to operations for satellite ocean-colour data in fishery management. *ICES J of Mar Science.* doi:10.1093/icesjms/fsq168
- Xiao Y, Friedrichs MAM. 2014a. 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.* 11:3015–3030. doi:10.5194/bg-11-3015-2014, 2014
- Xiao Y, Friedrichs MAM. 2014b. The assimilation of satellite-derived data into a one-dimensional lower trophic level marine ecosystem model. *J Geophys Res Oceans.* 119:2691–2712. doi:10.1002/2013JC009433.
- Xue P, Chen C, Beardsley RC. 2012. Observing system simulation experiments of dissolved oxygen monitoring in Massachusetts Bay. *J Geophys Res.* 117:doi:10.1364/OE.19.018927.