CHAPTER 22

Marine Biogeochemical Modelling and Data Assimilation for Operational Forecasting, Reanalysis, and Climate Research

David Ford[1], Susan Kay[1, 2], Robert McEwan[1], Ian Totterdell[1], and Marion Gehlen[3]

[1]Met Office, FitzRoy Road, Exeter, EX1 3PB, UK; [2]Plymouth Marine Laboratory, Prospect Place, The Hoe, Plymouth, PL1 3DH, UK; [3]Laboratoire des Sciences du Climat et de l'Environnement/Institut Pierre-Simon Laplace (LSCE/IPSL), Orme des Merisiers, 91191 Gif-sur-Yvette, France

Predictions of marine biogeochemistry are of importance for a range of applications, from operational forecasting of harmful algal blooms, to seasonal prediction of primary production, to understanding the influence of the marine carbon cycle on future climate change. Reanalyses, which include data assimilation in model hindcasts, are also required for the assessment of long-term environmental change. The inclusion of marine biogeochemistry in ocean forecasting and reanalysis systems is still in its early stages, but is already providing valuable insights. This chapter begins by giving an overview of biogeochemical modelling and data assimilation, and discussing challenges around physical-biogeochemical coupling and the use of observations. A summary of current applications to operational forecasting, reanalysis and climate studies is then given, before a vision is presented for a fully integrated prediction framework, linking five-day regional forecasting to global climate research.

Introduction

Marine biogeochemistry is the study of chemical elements in the ocean, and their interactions with marine life. Chief amongst these elements is carbon, the building block of life and a key influence on Earth's climate. Others of importance include nitrogen, oxygen, phosphorus, silicon, and iron. Biogeochemical cycling happens through physical transport, chemical reactions, and uptake and processing by plankton, which are organisms unable to swim against ocean currents. Phytoplankton, microscopic photosynthesising algae, form the base of the ocean food web and contribute about half of Earth's primary production. The zooplankton that consume them also process a significant quantity of carbon and nutrients. In addition, many plankton produce shells or skeletons mostly made of calcium carbonate or silicate (also referred to as biogenic opal). Higher trophic levels such as fish and marine mammals play a lesser role in elemental cycling, and so are generally considered separately.

Ford, D., et al., 2018: Marine biogeochemical modelling and data assimilation for operational forecasting, reanalysis and climate research. In "*New Frontiers in Operational Oceanography*", E. Chassignet, A. Pascual, J. Tintoré, and J. Verron, Eds., GODAE OceanView, 625-652, doi:10.17125/gov2018.ch22.

Broadly speaking, there are two main motivations for the modelling and prediction of marine biogeochemistry. The first is to study variability and trends in the uptake of carbon dioxide by the ocean, and the influence of this on global climate, as well as the process of ocean acidification. The second is to understand the impact of the physical-chemical marine environment on ecosystems and human activities. For instance:

- Algal blooms can be harmful to human health, either directly or via commercial fisheries.
- Visibility and turbidity is important for seafloor life and high biodiversity ecosystems such as tropical reefs, as well as for recreational, commercial, and naval diving activities.
- River runoff with increased nutrient contents, often with altered elemental ratios, can lead to excessive nutrient levels and increased algal growth, a process called eutrophication. This can be the result of fertiliser use or the discharge of domestic and agricultural waste, as well as natural processes. When these organisms die and decay the water can become depleted of oxygen, leading to conditions which are detrimental to many organisms, and can even result in widespread fish kills (Diaz et al., 2008).
- Different plankton species prefer different environmental conditions, and changes in temperature and nutrients can alter the composition of the whole marine ecosystem. The impacts of this can range from effects on commercial fish stocks, to large jellyfish blooms clogging the cooling vents of nuclear power stations.

Water quality and marine biodiversity are increasingly regulated by international law, so governments must monitor, anticipate and respond to environmental changes.

Operational oceanography primarily focuses on short-term impacts on the marine environment, and climate research largely focuses on long-term impacts on ecosystems, the carbon cycle and ocean acidification. The main focus of this chapter is on the applications and challenges of operational biogeochemical forecasting, but for completeness and context, a brief overview of reanalysis and climate research is also presented, focussing on how these relate to operational applications. In most cases, operational biogeochemical forecasting systems have developed by extending existing physical forecasting systems, as described elsewhere in this book, by combining them with existing biogeochemical models developed either for climate research or for ecological modelling. As such, operational biogeochemistry is less mature than operational physical forecasting, and often fundamental model development remains driven by research priorities, with reanalyses and hindcasts providing a natural link between different applications.

Numerous textbooks have already been dedicated to the field of marine biogeochemistry (e.g. Sarmiento and Gruber, 2006), as well as reviews of modelling (Heinze and Gehlen, 2013), and operational forecasting (Gehlen et al., 2015), and this chapter will not try to replicate them. A high-level overview is given in order to introduce physical oceanographers to the concepts and methods of biogeochemical modelling and data assimilation. The focus is on how biogeochemistry interacts with the various systems described in the other chapters of this book, and the motivations for its inclusion. A summary of existing applications is presented, and future challenges and ambitions are explored, with a vision given of a fully integrated physical-biogeochemical prediction framework. Examples are drawn primarily from work performed at the Met Office, but also more broadly from international initiatives such as the Copernicus Marine Environment Monitoring Service (CMEMS), the GODAE OceanView Marine Ecosystem Analysis and Prediction Task Team (GOV MEAP-TT), and the Climate Model Intercomparison Projects (CMIP) which inform the reports of the Intergovernmental Panel on Climate Change (IPCC).

Biogeochemical Models

General formulation

Most marine biogeochemical models are based on an "NPZD" approach (Fasham et al., 1990), standing for nutrient-phytoplankton-zooplankton-detritus (see Fig. 22.1). One or more state variables are used to represent each of these compartments, which together are used to calculate the evolution of the lower trophic levels of the pelagic (water column) ecosystem. Additionally, models may contain variables describing the carbon and oxygen cycles, sediments, benthic (seafloor) ecosystem, bacteria, and viruses. The general principles of ocean modelling are described elsewhere in this book (Chapter 2 by Fox-Kemper), and the same considerations of numerical methods apply to biogeochemistry, so are not repeated here. Biogeochemical models are reliant on physical models, and either the two are coupled and run together online, or physical model output is used to force the biogeochemistry offline (Heinze and Gehlen, 2013). The state variables are advected and diffused by the physical model, in the same way as temperature and salinity.

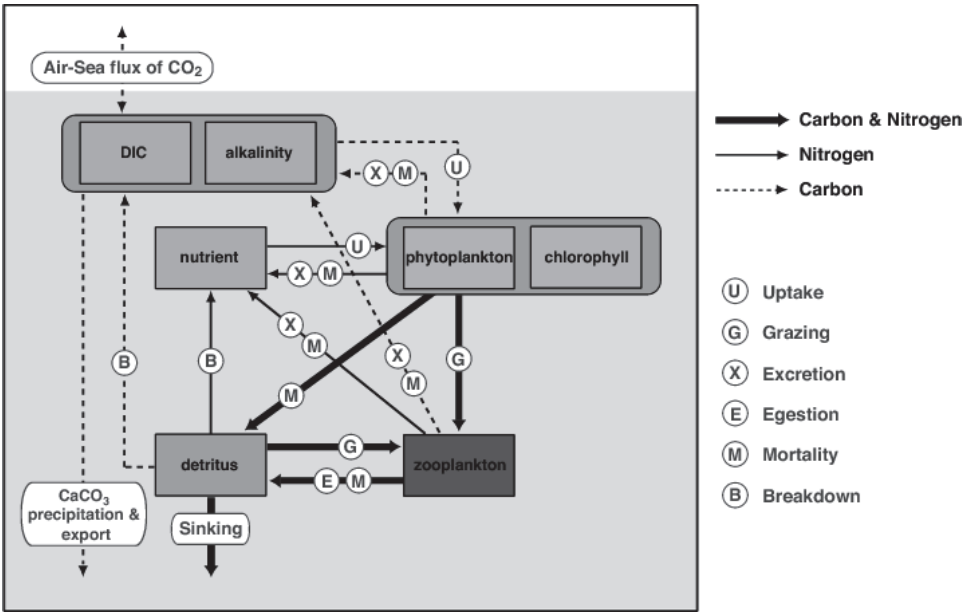


Figure 22.1. Schematic of the Hadley Centre Ocean Carbon Cycle model (HadOCC): an NPZD model with a coupled carbon cycle (Palmer and Totterdell, 2001). The NPZD component is represented by the four boxes labelled nutrient, phytoplankton, zooplankton, and detritus, with the carbon cycle components to the left and top of the diagram. Processes linking the components are shown in circles, and are represented in the model by equations.

Each of the NPZD variables are modelled as tracers, and are formulated as concentrations of (typically) nitrogen or carbon, rather than individual organisms. The nutrient compartment usually includes a variable for dissolved inorganic nitrogen (DIN), sometimes separated into nitrate and ammonium, as DIN is the main limiting nutrient throughout much of the ocean, particularly coastal regions. Phosphate, silicate, and iron may also be included. Plankton are split into phytoplankton, organisms which obtain their energy from sunlight through photosynthesis (autotrophy), and zooplankton, which obtain their energy by consuming other organisms (heterotrophy). Classical NPZD models just have a single variable for each of these, making no consideration of differences between species. A more complex approach is to split each up into two or more plankton functional types (PFTs), which group together species based on their function within the ecosystem. Phytoplankton grow by synthesising organic matter through photosynthesis, a process referred to as primary production. This may be limited by availability of light or nutrients, and is further regulated by temperature. Phytoplankton losses occur through grazing by zooplankton, natural mortality, and respiration. Zooplankton experience similar losses. Some of the organic matter lost from the plankton variables returns directly to the nutrient pool, and the remainder forms detritus. This is either grazed on by zooplankton, or broken down by bacterial and other processes and converted back to dissolved nutrients as it sinks to the deep ocean, in a process called remineralisation. Nutrient concentrations are typically high at depth, but become depleted at the surface through biological utilisation.

Models which represent the carbon cycle typically include additional state variables for dissolved inorganic carbon (DIC) and alkalinity. DIC accounts for the vast majority of carbon in the ocean, with the remainder stored as either dissolved or particulate organic carbon, including the carbon in plankton and detritus. Alkalinity is a measure of the capacity of seawater to neutralise an acid. This is distinct from pH, which is not a state variable, but can be diagnosed from other biogeochemical and physical variables (Orr et al., 2015). At the ocean surface, carbon dioxide (CO_2) is exchanged with the atmosphere, altering surface DIC concentrations. Whether the CO_2 flux is into or out of the ocean varies regionally. In the pre-industrial period the ocean and atmosphere CO_2 concentrations were in approximate balance, with the ocean a weak source of CO_2 to the atmosphere to balance riverine inputs of DIC (although in open ocean climate models riverine DIC inputs are not usually considered, so in pre-industrial model simulations the atmosphere and ocean are in equilibrium through the sea surface). However, in the present day, due to increasing anthropogenic CO_2 emissions, the global ocean is a net sink for CO_2. This helps mitigate the impact of climate change. However, CO_2 is a weak acid, and when it dissolves in water it reacts with it, and in so doing lowers the pH, leading to ocean acidification which impacts marine chemistry and living organisms. The air-sea flux is dependent on the surface partial pressure of CO_2 (pCO_2), which is a function of DIC, alkalinity, temperature, salinity, and pressure (with biological activity acting to modulate DIC and alkalinity). As with nutrients, the vertical distribution of DIC is such that the surface concentration is much lower than at depth. Physical processes typically act to bring carbon to the surface, whilst two mechanisms allow carbon to be transported to the deep ocean: the solubility pump and the biological pump. The solubility pump acts in regions of deep water formation such as the North Atlantic and Southern Ocean, where a strong uptake of atmospheric CO_2 is mixed to the deep ocean by the physically-driven thermohaline circulation. The biological pump acts through the gravitational sinking and recycling of detritus. The deeper the remineralisation takes place, the longer the carbon is shielded from exchange with the atmosphere. A small fraction of detritus will escape remineralisation and be buried into sediments, where it will be stored on geological timescales. Systems with a strong seasonal variability, such as high latitude bloom systems, are believed to be particularly efficient in exporting particulate carbon to depth. The carbon brought to the deep ocean by these mechanisms may be stored for centuries or longer, so studying their strength and variability is important for understanding climate change.

How much complexity?

Biogeochemical models can be very simple, with just a single variable for each of the NPZD compartments (or even omitting detritus) (Edwards, 2001), or much more complex, with dozens of state variables representing different species and processes (Le Quéré et al, 2005; Butenschön et al., 2016). The desirable level of complexity is a matter of great contention within the scientific community (Anderson, 2005; Flynn, 2005), and fundamentally boils down to a simple trade-off: more complex models include processes which are known to be important but are not fully understood, whereas simpler models only include processes which are much better understood, but neglect or amalgamate key aspects of the real world. It is vital to choose a model that is fit for answering the question under consideration.

A common difference between biogeochemical models is the number of PFTs included. Simple models such as HadOCC (Palmer and Totterdell, 2001) include a single phytoplankton and a single zooplankton functional type, as shown schematically in Fig. 22.1. Other models commonly used for global-scale forecasting and climate research, such as MEDUSA (Yool et al., 2013) and PISCES (Aumont et al., 2015), extend this to two phytoplankton and two zooplankton. Zooplankton are split into microzooplankton and mesozooplankton, which have differing sizes and diets, with mesozooplankton preferentially grazing on larger phytoplankton and microzooplankton. Meanwhile, phytoplankton are split between diatoms, which are relatively large (2-200 μm) plankton that form silicate shells and play a particularly important role in the sinking of carbon, and the remaining non-diatom species. The more complex model ERSEM (Baretta et al., 1995; Butenschön et al., 2016) introduces further PFTs, splitting non-diatoms into picophytoplankton (< 2 μm), nanophytoplankton (2-20 μm), and microphytoplankton (> 20 μm), and adding heterotrophic nanoflagellates to zooplankton. The PFTs are distinguished in the model by differing parameters for traits such as growth rates, grazing, and nutrient affinity. PFTs are typically grouped according to cell size, as different-sized organisms play different biogeochemical roles, and their relative distributions can define the entire food chain (Finkel et al., 2010). An alternative approach, taken by the DARWIN model (Follows et al., 2007), is to include tens or hundreds of PFTs with randomly prescribed parameters, allowing the fittest to emerge in the resulting ecosystem. Despite the relative complexity of some models, they still do not generally consider processes such as day-night cycles, buoyancy adjustment, or diurnal migration of zooplankton.

As well as the number of PFTs, the number of independently varying elements is a key difference between models. In 1934, Alfred Redfield reported that the ratios of carbon (C), nitrogen (N), and phosphorus (P) within phytoplankton and the deep oceans were remarkably constant at C:N:P = 106:16:1, which has become known as the Redfield ratio (Redfield, 1934). This ratio allows models to be formulated with fixed stoichiometry (elemental ratios), meaning PFTs can be represented in terms of a single element, with the quantity of other elements derived using the Redfield ratio. However, it has become clear that while the Redfield ratio may hold on average, stoichiometry varies between species and environmental conditions, and plays an important role in phytoplankton growth and diversity (Finkel et al., 2010). Therefore, some models include variable stoichiometry (e.g. Vichi et al., 2007), necessitating the inclusion of multiple state variables for each PFT, and additional nutrients: for instance, the nitrogen biomass and silicon biomass of diatoms would be separate state variables, with both DIN and silicate required. Chlorophyll also needs to be represented, either as a fixed or variable ratio to the biomass, as it is essential for photosynthesis. This can be in addition to the explicit inclusion of bacteria, sediments, the benthic ecosystem, and more. For instance, ERSEM 15.06 (Butenschön et al., 2016), a model with fully flexible stoichiometry, includes up to 59 pelagic and 36 benthic state variables.

An issue with including additional complexity is being able to validate each component. Biogeochemical observations are sparse at the best of times, and become all the more so the more specific the variable. Therefore, additional processes are often only validated in terms of their contribution to quantities such as total chlorophyll concentration. Of studies which have compared models of different complexities, it has generally been found that adding complexity does not necessarily improve model skill (Friedrichs et al., 2007; Kriest et al., 2010; Ward et al., 2013; Kwiatkowski et al., 2014; Xiao and Friedrichs, 2014).

Even the most complex models obviously remain a simplification, and the traditional split between phytoplankton and zooplankton may itself want revisiting. It is increasingly clear that the majority of plankton are not in fact exclusively autotrophs or heterotrophs, but mixotrophs: individual organisms that gain energy by both photosynthesis and the consumption of others, analogous to the Venus flytrap on land. This is not yet commonly represented in models, and may prove to be of great importance (Flynn et al., 2013; Mitra et al., 2014).

In some cases, choice of model complexity will be limited by computational cost. Each extra state variable has to be physically transported at every time step, and once coupled with state-of-the-art physics models, biogeochemical models can become extremely expensive to run. When operational forecasts must be generated within a limited timeframe, or climate simulations run for thousands of years, this often restricts the choice of model complexity (as well as physical resolution) to the simpler end of the spectrum.

In practice, the optimal complexity will depend on the available computing resources and the task at hand: models are tools, and different jobs require different tools. For instance, if the aim is to simulate decadal change in basin-scale primary production and air-sea CO_2 flux, many processes can be safely neglected or parameterised, and a simpler model may be the preferred choice. Whereas simulating small-scale variations in phytoplankton community structure and nutrient ratios will require more variables and processes by necessity. Furthermore, accurate modelling of a coastal ecosystem may require different components than for an open ocean ecosystem.

Biogeochemical Data Assimilation

The theory of data assimilation and its application to physical ocean forecasting systems is introduced elsewhere in this book (Chapter 17 by Hoteit et al., Chapter 18 by Jacobs et al.). The same principles and techniques apply to biogeochemical data assimilation, so are not repeated, but particular considerations for biogeochemical variables are discussed here. A number of reviews and discussions have also been published in recent years (e.g., Gregg et al., 2009; Dowd et al., 2014; Gehlen et al., 2015; Ford and Barciela, 2015).

As with biogeochemical modelling, the lack of first principles equations (such as the Navier-Stokes equations for physics) can potentially lead to a greater flexibility of approaches. In particular, there is scope for data assimilation methods to inform the building of models, and the appropriate variables and processes to include (Ward et al., 2013). In theory, this approach could be developed as a way to construct models based on observations.

More so than in physical data assimilation, the primary application of data assimilation to biogeochemistry has been for parameter estimation [see Schartau et al. (2017) for a recent and comprehensive review]. This aims to adjust model parameters so that a resulting model run better fits a given set of observations, and is usually applied *a priori* as a tuning exercise. As such, while a valuable part of model development, parameter estimation is not routinely applied during the production process of operational forecasts and reanalyses. But there is no reason why state and parameter estimation cannot be combined, so that model parameters are also adjusted during a model run, based on the assimilation of observations. Time-dependent parameters have previously been estimated this way (Losa et al., 2003; Mattern et al., 2012; Doron et al, 2013; Simon et al., 2015), but combined state/parameter estimation has yet to make the transition to operational forecasting and reanalysis systems (Matear and Jones, 2011). The first challenge for most models will be to select appropriate parameters (growth and grazing rates are likely candidates), convert these single numbers to fields that can vary spatially and evolve in time, and define sensible ranges for parameters that are mostly poorly constrained by experimental studies. The bigger challenge is then to vary these realistically, either directly during the data assimilation process, or as a balance relationship, perhaps extending the approach of Hemmings et al. (2008).

As in physical forecasting and reanalysis systems, biogeochemical data assimilation is increasingly used for state estimation, updating the model state variables to produce an analysis, the best estimate of the ocean state at a given time. The most common set of data assimilated is satellite ocean colour (McClain, 2009), as this is the only source of routine global observations of marine biogeochemistry. A full description of the ocean colour processing chain is provided elsewhere in this book (Chapter 9 by Volpe et al.), but the basic procedure is as follows. What ocean colour sensors directly measure is the radiation at different wavelengths leaving the top of the atmosphere. An atmospheric correction model is then used to calculate the water-leaving radiances, converted to remote sensing reflectance, from which optical properties and chlorophyll concentration can be estimated. Algorithms have also been developed to estimate primary production (Carr et al., 2006), and to split chlorophyll up into contributions from different PFTs (Brewin et al., 2011, 2017). The further up the processing chain the more biologically descriptive the product, but the greater the uncertainties.

The majority of ocean colour assimilation studies assimilate total chlorophyll concentration. This has some special considerations compared with physical variables such as sea surface temperature (SST). In particular, chlorophyll distributions are highly non-Gaussian, which poses issues for traditional assimilation methods, which typically assume a Gaussian error distribution. Often a logarithmic (Campbell, 1995) or anamorphic (Brankart et al., 2012) transformation is applied in order to normalise the error distribution, but some studies are exploring assimilation methods such as particle filters (van Leeuwen, 2009), which do not require a Gaussian distribution. Non-Gaussianity is a fundamental issue for data assimilation systems (Bocquet et al., 2010), and assimilation with non-Gaussian observations and models may lead to undesirable effects if not properly treated, potentially including issues with physical data assimilation, as discussed in the fourth section.

Furthermore, chlorophyll products are generated from satellite observations at different wavelengths using an empirical model, and can have relatively large observation errors. Meanwhile, chlorophyll is not the most straightforward variable to accurately model, so model errors are often large, too. Perhaps the most demanding challenge though is in using these observations of near-surface chlorophyll to also improve the sub-surface and non-observed biogeochemical model variables such as nutrients and carbon.

In part, this last point drives the choice of assimilation method. Many centres choose to use a scheme which is inherently multivariate, with the most common being the ensemble Kalman filter (EnKF; Evensen, 2003) and singular evolutive extended Kalman (SEEK) filter (Pham et al., 1998). These generate increments to multiple state variables based on relationships in the background error covariances. This brings obvious advantages in terms of updating the full model state, but relies on accurate evolution of the background error covariances, which is not straightforward. Furthermore, such methods can be computationally expensive, particularly the EnKF which often requires an ensemble of around 100 members. An alternative is to use a univariate assimilation method such as optimal interpolation or 3D-Var which just updates the model chlorophyll, but with the option of using a set of balance relationships to update other state variables. The most sophisticated such scheme in use is the nitrogen balancing scheme of Hemmings et al. (2008). This seeks to determine whether phytoplankton growth or loss errors dominate at each grid point, and partitions increments to the other state variables accordingly, constrained by the principle of conservation of mass.

Because of the reduced uncertainties, there is a great attraction in assimilating optical properties rather than chlorophyll. This is particularly the case in a shelf seas environment with large concentrations of suspended particulate matter (SPM) and coloured dissolved organic matter (CDOM, also known as gelbstoff), where ocean colour chlorophyll algorithms struggle, and where light directly affects more components of the ecosystem than chlorophyll does. Optical properties have been successfully assimilated by Shulman et al. (2013), Ciavatta et al. (2014), Jones et al. (2016), and Gregg and Rousseaux (2017). This has been shown to improve the representation of the ecosystem as a whole, and even of chlorophyll in particular, over assimilating chlorophyll, and is an approach likely to be increasingly explored.

However, it does rely on the biogeochemical model having a sufficiently sophisticated optical model. This remains lacking from most models, and such treatment of optics is of secondary importance for simpler models designed for the open ocean, and when computational restrictions limit model complexity. In these cases, it may be of benefit to move the other way up the ocean colour processing chain, and develop the assimilation of chlorophyll separated by plankton functional type. This has been investigated by Ciavatta et al. (2018) and Skákala et al. (2018), and results suggest that it improves the internal ecosystem dynamics over assimilating total chlorophyll, increasing reanalysis and forecasting skill. Further investigation is required, but it could be that, whilst undoubtedly beneficial, total chlorophyll is often the least appropriate ocean colour product to assimilate into a biogeochemical model.

On the other hand, there is a largely untapped potential to directly assimilate chlorophyll imagery into physical models (Gaultier et al., 2013; Titaud et al., 2010). High-resolution chlorophyll images contain a lot of information about eddies and surface currents, which could be used to update sea surface height fields, complementing existing altimetry products. This could either be done directly, or as part of a fully coupled physical-biogeochemical data assimilation scheme, which would help maintain consistency between physics and biogeochemistry.

With the increasing availability of in situ biogeochemical observations, more studies will incorporate assimilation of data other than ocean colour, which will allow more accurate updating of the multivariate biogeochemical state. However, observational coverage is still likely to remain relatively sparse for years to come, especially in shelf seas, so assimilating these data will provide its own challenges. In situ biogeochemical observations are even sparser than physical ones, and very rarely available in near-real-time, limiting the potential to assimilate them into operational models. Use of in situ biogeochemistry for state estimation has so far been largely limited to research studies related to specific cruises (e.g. Anderson et al., 2000), with observations instead providing valuable independent data sets for validation. An exception is the pCO_2 assimilation scheme of While et al. (2012).

But a revolution (of sorts) is promised over the coming years. The Argo programme has transformed physical oceanography, and is now being extended to include biogeochemistry under the flag of Biogeochemical-Argo (Johnson and Claustre, 2016; Claustre et al., 2010). The prospect of regular global profiles of chlorophyll, nitrate, pH, and oxygen means researchers are gearing up to assimilate these. Further observations will be provided through the increasing deployment of gliders. Advances in biogeochemical in situ observing systems are detailed elsewhere in this book (Chapter 6 by Telszewski et al.). Development of schemes to assimilate Biogeochemical-Argo data is underway at a number of centres. In common with some other areas of data assimilation, there are challenges around how to make the fullest use of sparse observations. More specific to marine biogeochemistry is the challenge of optimally integrating the assimilation of different in situ variables, and ocean colour. This will be a key focus in the development of these schemes.

Physical-biogeochemical Coupling

Impact of model physics on biogeochemistry

As biogeochemical state variables are modelled as tracers, model physics is of fundamental importance. Physical and biogeochemical models may be coupled online, in which case they are integrated together, with physical fields used directly by the biogeochemical model at each time step. Alternatively, the biogeochemical model may be forced offline using previously-generated physical model output, usually at a lower temporal resolution. Online coupling is in theory most accurate, but offline coupling allows more flexibility and spreading of computational cost.

Whether the coupling is online or offline, the equations are the same. Currents transport plankton and nutrients, and vertical velocities drive mixing in the water column. The same advection and diffusion schemes as used by the physical model for temperature and salinity are often used for the biogeochemical state variables too, but this need not be the case. A particular consideration is that tracer concentrations should never be negative, meaning a positive definite advection scheme should be used.

Despite the similarity in approach to ocean circulation, the biogeochemical modeller often has different priorities for physical model accuracy than the physical modeller does. In particular, tracers are very sensitive to mixed layer depth (MLD) and vertical velocities, which are notoriously difficult to model with accuracy. The MLD is critical in determining the mixing of nutrients and carbon into the surface ocean, as well as the depth from the surface to which phytoplankton are mixed. Small changes in the concentration of nutrients and plankton in the euphotic zone, where there is enough light available for photosynthesis, can have a large impact on primary production, and the maximum MLD over the integration period is more important than the mean. Furthermore, stratified regimes with a shallow MLD but deeper euphotic zone can lead to a deep chlorophyll maximum developing at the base of the mixed layer, where nutrient concentrations are greater than at the surface, and there remains sufficient light for growth. Similarly, vertical velocities and other contributors to vertical mixing control the amount of nutrients and carbon brought from the deep to the surface ocean. Biogeochemical quantities tend to be much more sensitive to this than physical ones such as SST or sea level.

Spatial scales and model resolution play an important role too. On longer timescales the thermohaline circulation and carbon pumps (Volk and Hoffert, 1985) control the transport of carbon and nutrients in the deep ocean, as well as patterns of ocean productivity. On shorter timescales, the mesoscale and submesoscale are a major control on primary production and air-sea fluxes, and these are not always well represented in models. Upwelling of nutrients along fronts can result in phytoplankton blooms, whilst eddies can either supply nutrients and fuel blooms (Lévy et al., 2001, 2012a), or aid the sequestering of carbon to the deep ocean.

A good discussion of these and other challenges is given by Holt et al. (2014). Given their sensitivity, tracers and biogeochemical models can also act as a valuable diagnostic for issues with modelled ocean physics.

As well as controlling the physical transport of tracers, temperature and salinity have direct effects on biogeochemistry. Many organisms grow at different rates depending on the temperature, as commonly represented by the parameterisation of Eppley (1972). The carbon cycle is particularly sensitive, with pCO_2 being a function of temperature and salinity as well as DIC and alkalinity. Decreasing temperature or increasing salinity increases the solubility of CO_2, with impacts on air-sea CO_2 flux and the solubility pump. Alkalinity is a charge balance and as such is closely tied to seawater constituents and hence salinity (Zeebe and Wolf-Gladrow, 2001).

Sometimes, the impact of physics on biology in the real ocean itself is unclear, an example of which is the mechanism behind triggering the spring bloom. In 1953, Harald Sverdrup published the critical depth hypothesis (Sverdrup, 1953). The critical depth is defined as the depth at which phytoplankton losses integrated over the water column equal phytoplankton growth, so that net primary production is zero. In winter, light levels are low and so the critical depth is shallow, whilst the MLD is deep, mixing phytoplankton away from the surface, prohibiting growth. During spring, the MLD shallows and the critical depth deepens. The critical depth hypothesis states that the spring bloom is initiated at the point where the MLD becomes shallower than the critical depth, because nutrients and phytoplankton are being kept in the euphotic zone and so growth must exceed loss. This hypothesis has been widely accepted and been a key component of biological oceanography for decades.

Recently though, studies have started to question this paradigm (Behrenfeld and Boss, 2014), arguing that the physical events described above are merely correlated with the spring bloom, and are not the cause. Many observational studies, both in situ and based on ocean colour, suggest that net primary production often becomes positive during winter, earlier than would be explained by the critical depth hypothesis. Furthermore, phytoplankton loss rates are not, as assumed by Sverdup (1953), constant. It has also been shown that bloom initiation coincides with the spring switch in net heat flux from out of the sea to downwards into the ocean (Smyth et al., 2014). As a replacement theory, Behrenfeld (2010) proposed the dilution-recoupling hypothesis. This states that in winter, deep mixing causes a decoupling between growth rates and losses from grazing ("dilution phase"), by reducing predator-prey interactions. This allows phytoplankton biomass to increase (bloom) until environmental conditions lead to a recoupling of growth and loss rates.

The debate over the exact mechanisms at play continues, and the study by Kuhn et al. (2015) provides a fine example of the opportunity for biogeochemical models to inform the understanding of the fundamentals of biological oceanography.

Impact of physical data assimilation on biogeochemistry

As discussed in other chapters of this book, assimilation of physical data, such as temperature, salinity, and altimetry, is a fundamental component of operational forecasting and reanalysis systems. It brings major improvements to analysis and forecast skill, and few operational centres would be without it. Given the reliance of biogeochemistry on model physics, physical data assimilation should also improve biogeochemical simulations.

Counter-intuitively though, the opposite is usually found, and physical data assimilation can significantly degrade biogeochemical fields. This is most commonly the case around the equator in global models (While et al., 2010; El Moussaoui et al., 2011; Park et al., 2018), but has also been observed in regional models (Raghukumar et al., 2015).

The reason for this is an impact of physical data assimilation on vertical mixing processes. The assimilation can result in spuriously large and noisy vertical velocities, bringing excessive concentrations of nutrients and carbon to the surface, as shown in Fig. 22.2. In the nutrient-limited tropics this leads to massively increased primary production and chlorophyll concentrations, and anomalous outgassing of CO_2.

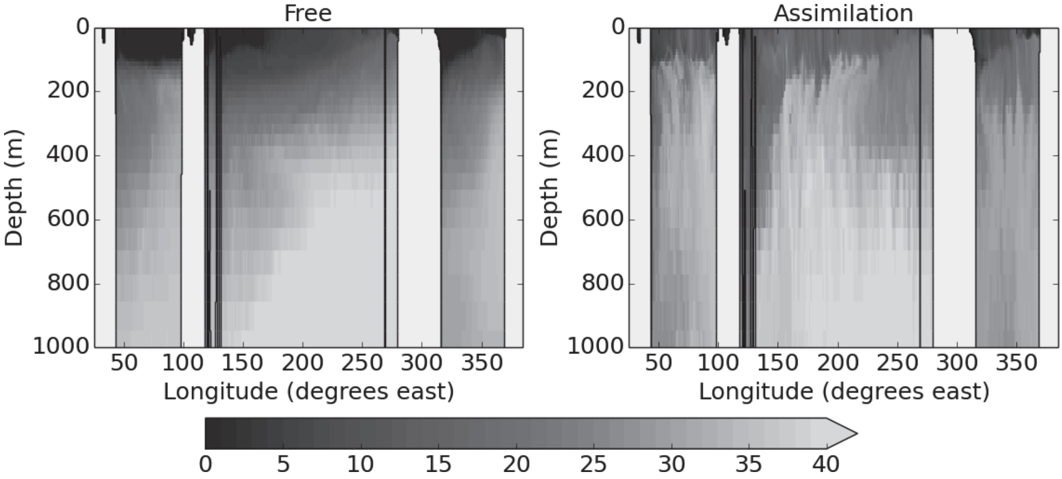


Figure 22.2. Equatorial cross-section of DIN (mmol m^{-3}) from equivalent runs of the NEMO-HadOCC physical-biogeochemical model with no data assimilation (left) and assimilation of physical data (right).

The exact mechanisms remain unclear, and addressing the issue has proven far from trivial. The blame is often laid on vertical velocities, but experiments have shown that filtering the vertical velocities alone does not prevent spurious tracer transport. Other fields relevant to vertical mixing are also affected, including vertical eddy diffusivity and density gradients. The impact on density may be particularly key. Major issues are typically only found when the assimilation updates the 3D temperature and salinity fields, either by directly assimilating temperature and salinity profiles, or as a balance when assimilating sea level anomaly from satellite altimetry. It is desirable for the assimilation to introduce finer-scale eddy structures than in the original model, but doing so without introducing undesirable levels of mixing (to quote Holt et al. (2014): “The last thing many ocean models need is more vertical mixing”) will be a challenge.

Preventing the issue at source is the long-term goal, though a number of methods to alleviate the issue in the short-term have been proposed. These have had limited success however. Not assimilating physics data in the affected regions does generally work (Park et al., 2018), but is not entirely satisfactory. Divergence damping (Dobricic et al., 2007) and incremental pressure correction (Waters et al., 2017) each help reduce vertical velocities, but may need to be used in conjunction with other techniques. The sensitivity of the biogeochemistry to elevated nutrients in nutrient-limited regions means that any truly effective solution needs to be comprehensive, as even a small increase in surface nutrients can significantly increase primary productivity. Biogeochemical model parameters could potentially be tuned to compensate, but probably not within sensible ranges. Furthermore, tuning biogeochemical parameters to compensate for biases in model physics is neither satisfying nor recommended, and may prevent the model from accurately responding to long-term changes. One method that may work in some regions is the nutrient balance proposed by While et al. (2010). This applies direct adjustments to nutrient concentrations based on the physical assimilation increments, to ensure the nutricline and pycnocline remain aligned. This addresses the problem found in some regions of the assimilation deepening the MLD, instantly putting increased nutrients into the mixed layer. However, it does not address the increased vertical mixing seen in other regions, particularly the tropics.

The issue does not hinder physics forecasting in the way it does biogeochemistry for two main reasons. First, accurate simulation of vertical mixing is not as critical to the variables that other users of physics output are primarily interested in, such as SST. Second, assimilating physical data will continually correct these variables, masking any problems. Nonetheless, it will still negatively affect physical as well as biogeochemical forecasts, and a robust solution is likely to lead to improvements to all systems. A long-term approach will include the development of strongly coupled physical-biogeochemical data assimilation schemes.

Impact of biogeochemistry on physical fields

Whilst biogeochemical models cannot be run without physical forcing, it is less common for ocean models to contain any feedback from the biogeochemistry to the physics, particularly on the timescales of operational forecasting and reanalysis. However, a number of such feedbacks exist in the real world, and will increasingly be represented as models improve.

Most obviously, as included in Earth system models (ESMs) used for climate projections, is the marine carbon cycle. By altering air-sea CO_2 flux and global ocean carbon uptake, this helps determine atmospheric CO_2 concentrations and therefore global climate. These processes give rise to complex feedback loops linking ocean physics, chemistry and biology (Heinze and Gehlen, 2013; Gehlen et al., 2014; Riebesell et al., 2009).

Another feedback beginning to be included in state-of-the-art ESMs is the cycling of the gas dimethyl sulphide (DMS). Some species of phytoplankton, notably coccolithophores, naturally synthesise dimethyl sulphoniopropionate (DMSP). This breaks down in the water to become DMS, which is outgassed to the atmosphere once it reaches the sea surface. The DMS then reacts with oxygen in the atmosphere to form sulphur dioxide, leading to the creation of sulphate aerosols, which act as cloud condensation nuclei. This process has been widely studied in the context of the CLAW hypothesis (Charlson et al., 1987), which proposes a negative feedback loop in the climate system. It suggests that global temperature rises will lead to increased phytoplankton growth through physiological effects, and so increased DMS production. In turn, this will lead to more clouds being formed, reflecting solar radiation and so cooling the planet. It is unclear how big a role this feedback actually plays, and there is also evidence that global warming will instead decrease phytoplankton growth through increased stratification and restricted nutrient supply, thereby turning it into a positive climate feedback. DMS production may also decrease as a result of ocean acidification (Six et al., 2013; Schwinger et al., 2017). No consensus has yet been reached in the scientific community (Gunson et al., 2006), but recent studies have suggested that the overall impact of DMS cycling on long-term climate change may in fact be minimal (Quinn and Bates, 2011). It is likely to contribute to natural variability though, and its inclusion in operational forecasting models and reanalyses is a challenge for the future.

Organic matter from phytoplankton and bacteria can also form cloud condensation nuclei directly (Wilson et al., 2015). Bubbles form on the surface of the ocean, and organic matter can become incorporated. The bubbles burst, releasing salt and organic matter into the atmosphere, known as sea spray aerosol, acting as cloud condensation nuclei.

Phytoplankton directly impact atmospheric physics further through their contribution to ocean albedo. The presence of chlorophyll colours the surface ocean green, altering the wavelengths of light which are absorbed and reflected by the ocean, and so the amount of shortwave radiation heating the lower atmosphere. Whilst a smaller contribution than factors such as sea state, sea surface chlorophyll is included in the widely-used albedo parameterisation of Jin et al. (2004), and a climatological representation of this is now included in weather forecasts produced at the Met Office. In future, a direct feedback from a marine biogeochemical model may be included in coupled weather and climate models.

A larger and direct biophysical feedback exists within the ocean, operating on similar principles (Morel, 1988). The depth over which shortwave solar radiation is absorbed is controlled by the presence of chlorophyll in the open ocean, as well as SPM and CDOM in coastal waters. When chlorophyll is present, more radiation is absorbed in the surface layers, and so the associated warming is concentrated nearer the surface. This raises SST, cools the sub-surface, and shallows the MLD. It has also been found to amplify seasonal cycles of temperature and sea ice cover (Manizza et al., 2005), and enhance rainfall during the South Asian monsoon (Turner et al., 2012). It is beginning to be included in ocean models, with chlorophyll taken either from a satellite climatology or directly from a biogeochemical model, and is expected to have impacts on all time scales.

A more controversial suggestion of a significant biophysical feedback in the ocean is that creatures such as copepods and jellyfish make a major contribution to physical mixing (Dewar et al., 2006). One study (Katija and Dabiri, 2009) measured the amount of turbulent mixing generated by motile plankton, extrapolated the results to the global ocean, and concluded that their contribution to ocean mixing could be greater even than that provided by wind, which would be a huge omission from current ocean models and generally accepted theory. The suggestion has obviously been treated with caution, but the contribution of different processes to ocean mixing remains uncertain (Dewar et al., 2006), and so the role of zooplankton is not implausible. Even if only partially the case, it is a phenomenon warranting further study and potential inclusion in state-of-the-art ocean models.

Finally, (perhaps), sea ice can sustain large populations of algae, the presence of which can potentially modify the ice albedo and speed up seasonal melting, accelerated further by their impact on light penetration once in the water (Taskjelle et al., 2017). Whilst the overall significance of this is unclear, a number of models of the ice ecosystem have now been developed (Vancoppenolle and Tedesco, 2017).

The impact of biophysical feedbacks on forecasting skill has yet to be investigated, but any process found to make a relevant contribution will progressively be incorporated into operational systems.

Overview of Current Applications

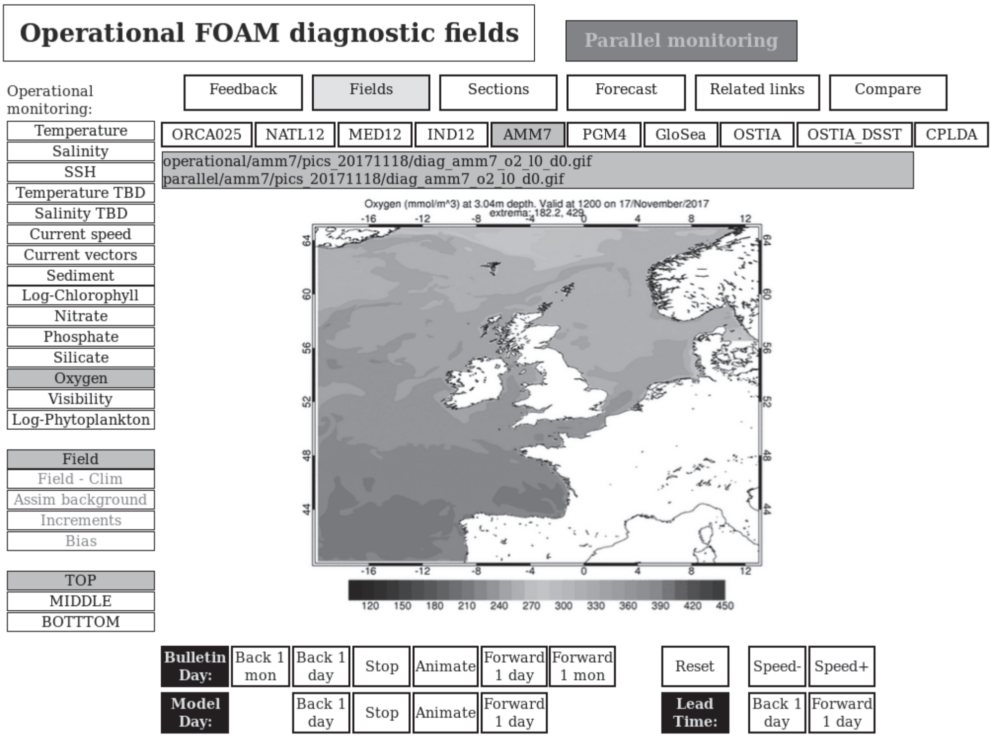
Operational forecasting

Most operational biogeochemical forecasting systems operate at regional scales, as these are the scales where there is most human interest in short-term change of the marine environment. Shelf seas provide 90% of fisheries and almost all leisure use of the sea. Short-range forecasts of algal blooms, water clarity and environmental indicators are of interest to fisheries, regulators, navies and others. At the global scale, forecasts of primary production can be used as inputs to fisheries models (Cheung et al., 2009; Lehodey et al., 2015) or as boundary conditions for regional models. In addition, applications for monitoring the global carbon cycle are emerging (Gehlen et al., 2015).

In Europe, operational forecasts for a number of regions are produced and made publicly available as part of CMEMS (http://marine.copernicus.eu/). These cover the global ocean, Arctic Ocean, Baltic Sea, Black Sea, Mediterranean Sea, Iberian-Biscay-Irish (IBI) Seas, and Northwest European Shelf Seas. Regional forecasts are also produced in other countries around the world, including Australia, Canada, Indonesia, USA, and others. Many of the centres involved in producing operational biogeochemical forecasts collaborate as part of the GOV MEAP-TT. A related, emerging field is the operational forecasting of marine ecology (Payne et al., 2017).

Biogeochemical forecasting systems vary considerably in terms of model complexity, data assimilation, and coupling techniques (Brasseur et al., 2009; Gehlen et al., 2015). This depends on available resources, customer requirements, and how closely tied a biogeochemical system is to a physical counterpart. Systems generally run on either a daily or a weekly basis, either coupled online to an equivalent physics forecasting system, or forced offline by forecast physical fields. Data assimilation, normally of ocean colour, is increasingly employed, but this is not ubiquitous. Forecast accuracy and validation strategies are covered elsewhere in this book (Chapter 19 by Hernandez et al.).

At the Met Office, a daily analysis and six-day forecast is produced for the Northwest European Shelf Seas (NWS), and made available through CMEMS. This is run operationally as part of the FOAM suite (Blockley et al., 2014), which provides global and regional ocean forecasts. The physical model used is NEMO (Madec, 2008), in a series of one-way nested configurations. An analysis and forecast is initially produced for the 1/4° global configuration. This then provides boundary conditions for a 1/12° model of the North Atlantic, which in turn provides the boundary conditions for a 7 km resolution model of the NWS (O'Dea et al., 2012, 2017). Each uses atmospheric forcing from the Met Office numerical weather prediction (NWP) model. The NWS configuration of NEMO is coupled online with the ERSEM biogeochemical model (Edwards et al., 2012), providing coupled physical-biogeochemical forecasts for CMEMS and other customers. Physics data are assimilated using the 3D-Var NEMOVAR scheme (Waters et al., 2015). Assimilation of ocean colour data has recently been developed, and will be implemented operationally in the near future. The operational system is accompanied by a comprehensive monitoring system, as shown in Fig. 22.3.



operational/amm7/pics_20171118/diag_amm7_o2_l0_d0.gif
parallel/amm7/pics_20171118/diag_amm7_o2_l0_d0.gif

Figure 22.3. Monitoring of the Met Office operational forecasting system.

A similar approach is taken by other centres contributing to CMEMS, although forcing strategies and update schedules differ. For instance, the global ocean product from Mercator Ocean is updated weekly, and run using offline physical forcing downscaled from daily physics forecasts. The Mediterranean Sea product run by OGS is updated biweekly, also using offline forcing, delivering seven days of analysis/hindcast and ten days of forecast, and incorporating ocean colour data assimilation (Teruzzi et al., 2014). Despite some differences in methodology, in general a relatively standardised approach is taken to the generation and delivery of products within CMEMS.

Outside of CMEMS, there are currently few examples of short-range operational biogeochemical forecasts being produced, with any operational applications typically more focussed on ecology and fisheries (Payne et al., 2017), and often based on statistical relationships with physical fields and satellite ocean colour data, rather than coupled physical-biogeochemical modelling. Such applications can include forecasting distributions of pathogens such as bacteria of the genus Vibrio (Jacobs et al., 2014), risk of coral bleaching (Liu et al., 2018), and harmful algal blooms (Davidson et al., 2016). In India, potential fishing zone advisories are issued by INCOIS based on satellite SST and ocean colour data (Deshpande et al., 2011), with plans to develop a model forecasting system (Thushara and Vinayachandran, 2016; Vijith et al., 2016). In Australia, an operational physical-biogeochemical forecasting system for the Great Barrier Reef is being assembled as part of the eReefs project (Schiller et al., 2014; Baird et al., 2016; Jones et al., 2016), for applications including sediment transport and water quality. In Indonesia, weekly regional biogeochemical forecasts are produced operationally using the NEMO-PISCES model (Gutknecht et al., 2016), as part of the INDESO project which monitors and forecasts variables from physics to fish.

Aside from short-range forecasts, there is an increasing demand for seasonal and decadal forecasts of marine biogeochemistry (Rousseaux and Gregg, 2017; Siedlecki et al., 2016; Hobday et al., 2018; Séférian et al., 2014). This is of importance for fisheries and aquaculture, as well as informing environmental commitments, for instance relating to the commitment to Good Environmental Status under the European Union's Marine Strategy Framework Directive (MSFD).

Climate studies

For many biogeochemical models, the primary motivation for development is inclusion in ESMs, for use in climate projections. This allows full consideration of the marine carbon cycle, how ocean carbon uptake will vary into the future, and the resulting feedback on global climate (Friedlingstein et al., 2006). Ocean acidification and ecosystem change can also be studied. The regular reports of the IPCC draw on results submitted as part of the international CMIP projects, most recently CMIP5 (Taylor et al., 2012).

As climate models potentially need to be run for thousands of years, computational expense is a critical factor. Biogeochemical models run as part of ESMs will therefore normally be relatively simple with only a small number of PFTs, and most are still only run at 1° or lower resolution, where eddies must be parameterised. Results from multiple ESMs are often combined in a multi-model ensemble for the analysis of climate change and ocean acidification impacts on marine biogeochemistry and ecosystems (Orr et al., 2005; Bopp et al., 2013; Gehlen et al., 2014). This approach gives insights into regional differences in ecosystem stressors that may occur as a result of climate change. Current models suggest a seesaw pattern with increased primary productivity at high latitudes and decreased productivity over the tropics and mid-latitudes. These changes are at first order explained by increased stratification in response to warming. A warmer ocean leads to enhanced stratification and fewer nutrients upwelled to the euphotic zone. At low- to mid-latitudes this increases the nutrient limitation and results in lower primary productivity. At high latitudes, however, enhanced stratification reduces light limitation as phytoplankton remain nearer the surface, which along with warming and an extended growing season, results in increased primary productivity (Bopp et al., 2013). However, this broad picture hides significant regional and inter-model variability (Laufkötter et al., 2015). More complex ecosystem models, with multiple and flexible plankton types, can also give insight into how changes in temperature and nutrient supply may affect ecosystem productivity and the makeup of plankton communities in different regions (Dutkiewicz et al., 2013; Barton et al., 2016).

Regional studies use global climate models as drivers to produce downscaled projections at higher resolution. For example, biogeochemical modelling has been used to investigate the impact of climate change and other anthropogenic stressors on the marine ecosystem of the NWS (Wakelin et al., 2015). Regional model outputs can be combined with fish models to investigate how climate change may affect food supply and promote sustainable fisheries (Barange et al., 2014; Fernandes et al., 2015).

Reanalysis

Reanalyses, in which models and data assimilation are used to recreate the past ocean state, provide a link between operational forecasting and climate studies. They can be used to assess long-term environmental change, such as eutrophication or alterations to plankton community structure and phenology (e.g. bloom timings). They can also be used to monitor the global ocean carbon sink, and study its variability. Reanalyses could also provide validation of climate models, by seeing how well they reproduce the observed ocean when given realistic forcing and initial conditions. The requirements for physical reanalyses, such as consistent and stable inputs, are discussed elsewhere in this book (Chapter 19 by Haines), and these considerations apply equally to biogeochemistry.

The same NEMO-ERSEM configuration used for operational forecasting of the NWS at the Met Office is also used to produce reanalyses of the region covering recent decades, again available through CMEMS. The version currently available (at CMEMS v3) just assimilates SST, but the latest reanalysis, being produced for CMEMS v4, introduces the assimilation of ocean colour data. This will allow greater exploration of long-term variability and change, and can be compared to other regional reanalyses (Ciavatta et al., 2016). Reanalyses are also available for the other regions considered as part of CMEMS.

The Met Office also produces reanalyses of the global ocean (Ford and Barciela, 2017), using NEMO coupled with the HadOCC biogeochemical model (Fig. 22.1). Ocean colour is assimilated using 3D-Var plus the nitrogen balancing scheme of Hemmings et al. (2008), and there is the option to assimilate pCO_2 (While et al., 2012) and physics (Waters et al., 2015) data. A particular motivation for the work is to explore variability in the global ocean carbon cycle. An example is shown in Fig. 22.4, which shows the monthly mean air-sea CO_2 flux in the Tropical Pacific from a NEMO-HadOCC reanalysis, and how this compares to climatology (Takahashi et al., 2009) and variability of the multivariate El Niño Southern Oscillation (ENSO) index (Wolter and Timlin, 1993, 1998). The Tropical Pacific is a major upwelling region, which brings nutrient- and carbon-rich waters to the surface, resulting in strong outgassing of CO_2 to the atmosphere. In El Niño periods, such as 1997/1998 and 2009/2010, upwelling is weaker, reducing CO_2 outgassing. In La Niña periods, such as 1999 and 2007/2008, upwelling is enhanced, increasing CO_2 outgassing. This strong inter-annual variability is captured well by the reanalysis. NEMO-HadOCC reanalyses are also being used to learn about biases in climate model projections submitted to CMIP5.

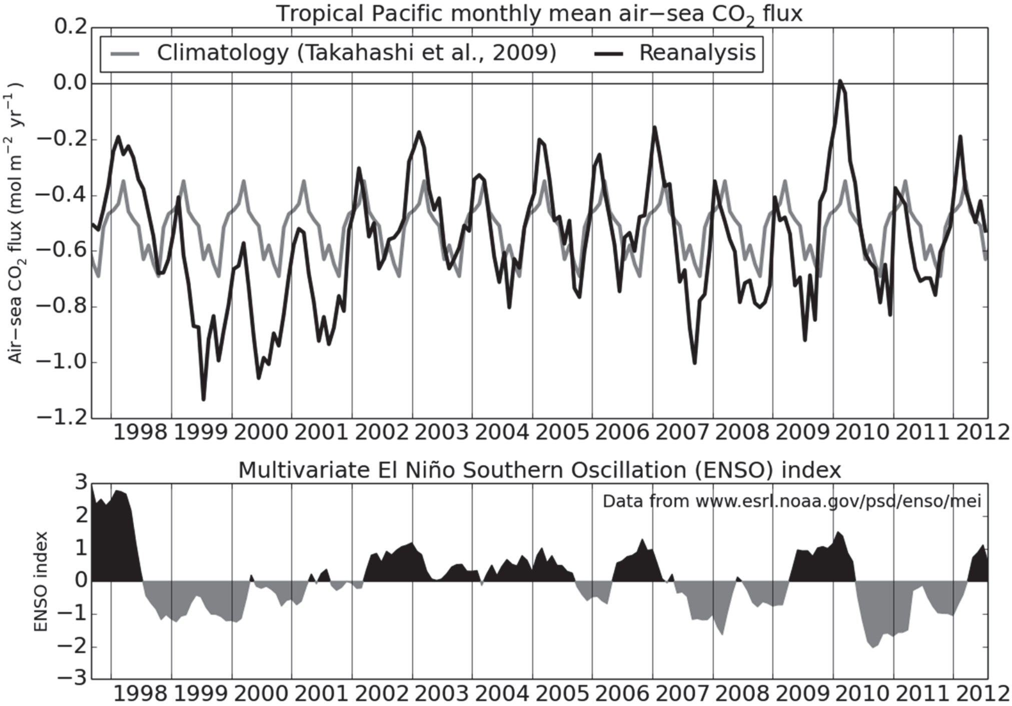


Figure 22.4. Top: Tropical Pacific monthly mean air-sea CO_2 flux (negative means outgassing) from a repeating climatology (Takahashi et al., 2009) and a reanalysis assimilating ocean colour data (Ford and Barciela, 2017). Bottom: multivariate ENSO index (Wolter and Timlin, 1993, 1998).

Towards a Fully Integrated Prediction Framework

As scientific understanding and computing resources continue to advance, so do the modelling systems used to predict the weather, ocean, and climate. As discussed elsewhere in this book (Chapter 16 by Harris), coupled ocean-atmosphere forecasting models are becoming increasingly mature, and starting to be used operationally. On climate time scales, the next generation of ESMs aim to include ever more components of the Earth system. There is also a strong desire to ensure consistency and traceability between regional and global, and weather and climate, simulations. Marine biogeochemistry should be actively included, and there is presently an opportunity to increase the value of biogeochemical predictions to society.

For modelling and forecasting biogeochemistry, there are a number of particular considerations, which are already facing the current generation of systems. Physical ocean forecasting systems are often run at the highest spatial resolution that can be afforded, meaning that coupling a biogeochemical model at the same resolution is not an option. This may necessitate offline coupling using downscaled physical fields, or the online coarsening approach of Lévy et al. (2012b), in which physics and biogeochemistry are coupled online, but the biogeochemical transport is calculated at lower resolution. Alternatively, given the issues around biogeochemical model complexity discussed in the second section, a possible approach might be a "plug-and-play" framework, in which the same model could be run with fewer PFTs, or different parameterisations, depending on the application. This could be achieved using a dedicated coupler such as FABM (Bruggeman and Bolding, 2014). Even if high-resolution physical fields can be used though, exploiting these to benefit biogeochemistry is not necessarily straightforward. The (sub-)mesoscale processes represented, and amount of extra energy present in high-resolution simulations, can potentially have large impacts on biogeochemical fields (Lévy et al., 2001, 2012a), which may not be beneficial if the biogeochemical model parameterisations have been developed and tuned at coarser resolution. Furthermore, currently neglected biogeochemical processes may need to be included.

A further challenge is understanding the limits of biogeochemical predictability, and properly characterising and presenting uncertainty information. Errors can arise and propagate from a variety of sources, both physical and biogeochemical, and may either add or cancel. The ability to quantify uncertainty through comparison with observations is limited due to the sparsity of data, the high level of uncertainty present in both satellite and in situ observations themselves, and the fact that observed and modelled quantities are often not a like-for-like match. Furthermore, the spatial and temporal scales at which models have skill must be considered. Models may be able to capture large-scale patterns but not small-scale features (Saux Picart et al., 2012), and understanding and quantifying these limits of predictability is important for communicating with, and gaining the confidence of, end users (Hyder et al., 2015).

When taken to its logical conclusion, the current direction of travel in weather and climate modelling leads to centres developing fully integrated prediction frameworks, allowing two-way coupling between all relevant components of the Earth system, which can be used for modelling and assimilation on all time and space scales. These would have the flexibility to allow components and processes to be switched on and off, parameterisations to be changed, and resolutions to vary. In the Met Office, for example, the aim is to maintain a seamless approach to atmosphere and ocean modelling across time scales (Martin et al., 2010). The UM (Cullen, 1993) atmosphere and NEMO (Madec, 2008) ocean models are used for regional and global, weather and climate modelling, with the aim of traceability between applications. These are increasingly used for coupled modelling and data assimilation (Lea et al., 2015). The United Kingdom Environmental Prediction (UKEP) project (Lewis et al., 2018) aims to extend this by also including the land surface, waves, and marine biogeochemistry, for regional predictions covering the UK. The components and processes included will develop further as these projects mature.

Such an approach has various potential benefits. It would allow the different biophysical feedbacks detailed in the fourth section to be included, benefitting physics simulations and in turn allowing the biogeochemistry to benefit from improved physics. Physical and biogeochemical forecasts would be fully consistent, providing clarity to end users. Conclusions and issues from simulations on one time scale could directly inform others, for instance using reanalyses to learn about climate models. This potentially makes the model much more robust, particularly when predicting future change, having been tested and validated in a wider range of situations. It would also make it quicker and easier to extend to new regions and novel applications, as developments made for one purpose would be compatible elsewhere. This is further the case due to the shared development effort of a large community working on the same system.

A fully integrated system would present a number of challenges though. Different applications may desire specialist approaches, and compromises may have to be made that otherwise would not. For marine biogeochemistry in particular, where there remains considerable model uncertainty, there is much to be said for maintaining a diverse set of different models, and this should be balanced with the advantages that shared developments towards a community model would bring. Furthermore, with every extra component and process added, particularly with two-way feedbacks, it becomes increasingly difficult to diagnose issues, and validate the impact of any changes. It is likely that many developments would improve some components and degrade others, and marine biogeochemical forecasts will always be of lower priority than weather forecasts when deciding what changes are acceptable. The computational expense of a fully integrated high-resolution system would also be rather large.

As well as the potential for better including marine biogeochemistry in physical systems, improved coupling with the wider marine ecosystem needs to be considered. An integrated approach to this is already taken by end-to-end ecosystem models (Fulton, 2010), which aim to include marine physics, biogeochemistry, fish, marine mammals, seabirds, and humans in a single model. Opinions differ as to whether it is better to use a single end-to-end model or a series of more specialist coupled models. Again, there is a benefit to maintaining a diversity of approaches, particularly whilst the subject is still maturing, and coupled physical-biogeochemical models can feed into these developments. This can help address the challenges of increasing the use of biogeochemical and ecosystem models by policy makers, as discussed by Hyder et al. (2015).

Summary

Marine biogeochemistry is increasingly being included in operational forecasting and reanalysis systems, using progressively more sophisticated models and data assimilation schemes. These systems are less mature than their physical counterparts, but are starting to provide valuable information to a range of users. Future development of biogeochemical systems promises many exciting challenges, through more accurate modelling, exploiting new data sources, increased integration with physical ocean and atmosphere forecasting systems, and better communication of products to end users (Hyder et al., 2015). This will help address important societal needs, such as the impact of toxic algae on human health, the maintenance of diverse and sustainable ecosystems, and understanding the role of the ocean in a changing climate.

Acknowledgements

The authors would like to thank Rabitah Daud, Dianne Deauna, Matías Dinapoli, Jenny Machinford, and an anonymous reviewer for their comments on the draft manuscript, which helped to improve the quality of the final version.

References

Anderson, L. A., A. R. Robinson, and C. J. Lozano, 2000: Physical and biological modeling in the Gulf Stream region: I. Data assimilation methodology. *Deep Sea Research Part I: Oceanographic Research Papers*, **47**, 1787-1827, https://doi.org/10.1016/S0967-0637(00)00019-4.

Anderson, T. R., 2005: Plankton functional type modelling: running before we can walk? *Journal of Plankton Research*, **27**, 1073–1081, https://doi.org/10.1093/plankt/fbi076.

Aumont, O., C. Ethé, A. Tagliabue, L. Bopp, and M. Gehlen, 2015: PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies. *Geoscientific Model Development*, **8**, 2465-2513, https://doi.org/10.5194/gmd-8-2465-2015.

Baird, M. E., and Coauthors, 2016: Remote-sensing reflectance and true colour produced by a coupled hydrodynamic, optical, sediment, biogeochemical model of the Great Barrier Reef, Australia: comparison with satellite data. *Environmental Modelling & Software*, **78**, 79-96, https://doi.org/10.1016/j.envsoft.2015.11.025.

Barange, M., and Coauthors, 2014: Impacts of climate change on marine ecosystem production in societies dependent on fisheries. *Nature Climate Change*, **4**, 211-216, https://doi.org/10.1038/nclimate2119.

Baretta, J. W., W. Ebenhöh, and P. Ruardij, 1995: The European regional seas ecosystem model, a complex marine ecosystem model. *Netherlands Journal of Sea Research*, **33**, 233-246, https://doi.org/10.1016/0077-7579(95)90047-0.

Barton, A. D., A. J. Irwin, Z. V. Finkel, and C. A. Stock, 2016: Anthropogenic climate change drives shift and shuffle in North Atlantic phytoplankton communities. *Proceedings of the National Academy of Sciences*, **113**, 2964-2969, https://doi.org/10.1073/pnas.1519080113.

Behrenfeld, M. J., 2010: Abandoning Sverdrup's Critical Depth Hypothesis on phytoplankton blooms. *Ecology*, **91**, https://doi.org/10.1890/09-1207.1.

Behrenfeld, M. J., and E. S. Boss, 2014: Resurrecting the ecological underpinnings of ocean plankton blooms. *Annual Review of Marine Science*, **6**, 167-194, https://doi.org/10.1146/annurev-marine-052913-021325.

Blockley, E. W., and Coauthors, 2014: Recent development of the Met Office operational ocean forecasting system: an overview and assessment of the new Global FOAM forecasts. *Geoscientific Model Development*, **7**, 2613-2638, https://doi.org/10.5194/gmd-7-2613-2014.

Bocquet, M., C. A. Pires, and L. Wu, 2010: Beyond Gaussian statistical modeling in geophysical data assimilation. *Monthly Weather Review,* **138**, 8, 2997-3023, https://doi.org/10.1175/2010MWR3164.1.

Bopp, L., and Coauthors, 2013: Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models. *Biogeosciences*, **10**, 6225-6245, https://doi.org/10.5194/bg-10-6225-2013.

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

Brasseur, P., and Coauthors, 2009: Integrating biogeochemistry and ecology into ocean data assimilation systems. *Oceanography*, **22**, 206-215.

Brewin, R. J. W., and Coauthors, 2011: An intercomparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing. *Remote Sensing of Environment*, **115**, 325-339, https://doi.org/10.1016/j.rse.2010.09.004.

Brewin, R. J., and Coauthors, 2017: Uncertainty in Ocean-Color Estimates of Chlorophyll for Phytoplankton Groups. *Frontiers in Marine Science*, **4**, 104, https://doi.org/10.3389/fmars.2017.00104.

Bruggeman, J., and K. Bolding, 2014: A general framework for aquatic biogeochemical models. *Environmental Modelling & Software*, **61**, 249-265, https://doi.org/10.1016/j.envsoft.2014.04.002.

Butenschön, M., and Coauthors, 2016: ERSEM 15.06: a generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels. *Geoscientific Model Development*, **9**, 1293-1339, https://doi.org/10.5194/gmd-9-1293-2016.

Campbell, J. W., 1995: The lognormal distribution as a model for bio-optical variability in the sea. *Journal of Geophysical Research: Oceans*, **100**, 13237–13254, https://doi.org/10.1029/95JC00458.

Carr, M. E., and Coauthors, 2006: A comparison of global estimates of marine primary production from ocean color. *Deep Sea Research Part II: Topical Studies in Oceanography*, **53**, 741-770, https://doi.org/10.1016/j.dsr2.2006.01.028.

Charlson, R. J., J. E. Lovelock, M. O. Andreaei, and S. G. Warren, 1987: Oceanic phytoplankton, atmospheric sulphur, cloud. *Nature*, **326**, 655-661.

Cheung, W. W. L., V. W. Y. Lam, J. L. Sarmiento, K. Kearney, R. Watson, and D. Pauly, 2009: Projecting global marine biodiversity impacts under climate change scenarios. *Fish and Fisheries*, **10**, 235–251, https://doi.org/10.1111/j.1467-2979.2008.00315.x.

Ciavatta, S., R. Torres, V. Martinez-Vicente, T. Smyth, G. Dall'Olmo, L. Polimene, and J. I. Allen, 2014: Assimilation of remotely-sensed optical properties to improve marine biogeochemistry modelling. *Progress in Oceanography*, **127**, 74-95, https://doi.org/10.1016/j.pocean.2014.06.002.

Ciavatta, S., S. Kay, S. Saux-Picart, M. Butenschön, and J. I. Allen, 2016: Decadal reanalysis of biogeochemical indicators and fluxes in the North West European shelf-sea ecosystem. *Journal of Geophysical Research: Oceans*, **121**, 1824–1845, https://doi.org/10.1002/2015JC011496.

Ciavatta, S., R. J. W. Brewin, J. Skákala, L. Polimene, L. de Mora, Y. Artioli, and J. I. Allen, 2018: Assimilation of Ocean-Color Plankton Functional Types to Improve Marine Ecosystem Simulations." *Journal of Geophysical Research: Oceans,* **123**, no. 2, 834-854, https://doi.org/10.1002/2017JC013490.

Claustre, H., and Coauthors, 2010: Bio-optical profiling floats as new observational tools for biogeochemical and ecosystem studies. *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2)*, Venice, Italy, 21-25 September 2009, Hall, J., D. E. Harrison, and D. Stammer, Ed., ESA Publication WPP-306, https://doi.org/10.5270/OceanObs09.cwp.17.

Cullen, M. J. P., 1993: The unified forecast/climate model. *Meteorological Magazine*, **122**, 81-94.

Davidson, K., D. M. Anderson, M. Mateus, B. Reguera, J. Silke, M. Sourisseau, and J. Maguire, 2016: Forecasting the risk of harmful algal blooms. *Harmful Algae*, **53**, 1-7, https://doi.org/10.1016/j.hal.2015.11.005.

Deshpande, S. P., K. V. Radhakrishnan, and U. G. Bhat, 2011: Direct and indirect validation of potential fishing zone advisory off the coast of Uttara Kannada, Karnataka. *Journal of the Indian Society of Remote Sensing,* **39 (4)**, 547-554, https://doi.org/10.1007/s12524-011-0104-4.

Dewar, W. K., R. J. Bingham, R. L. Iverson, D. P. Nowacek, L. C. St Laurent, and P. H. Wiebe, 2006: Does the marine biosphere mix the ocean? *Journal of Marine Research*, **64**, 541-561, https://doi.org/10.1357/002224006778715720.

Diaz, R. J., and R. Rosenberg, 2008: Spreading dead zones and consequences for marine ecosystems. *Science*, **321**, 926-929, https://doi.org/10.1126/science.1156401.

Dobricic, S., N. Pinardi, M. Adani, M. Tonani, C. Fratianni, A. Bonazzi, and V. Fernandez, 2007: Daily oceanographic analyses by Mediterranean Forecasting System at the basin scale. *Ocean Science*, **3**, 149-157, https://doi.org/10.5194/os-3-149-2007.

Doron M., Brasseur P., Brankart J.-M., Losa S. and Melet A., 2013: Stochastic estimation of biogeochemical parameters from Globcolour ocean colour satellite data in a North Atlantic 3D coupled physical-biogeochemical model, *Journal of Marine Systems.,* **117–118**, 81–95, https://doi.org/10.1016/j.jmarsys.2013.02.007.

Dowd, M., E. Jones, and J. Parslow, 2014: A statistical overview and perspectives on data assimilation for marine biogeochemical models. *Environmetrics*, **25**, 203–213, https://doi.org/10.1002/env.2264.

Dutkiewicz, S., J. R. Scott, and M. J. Follows, 2013: Winners and losers: Ecological and biogeochemical changes in a warming ocean. *Global Biogeochemical Cycles*, **27**, 463–477, https://doi.org/10.1002/gbc.20042.

Edwards, A. M., 2001: Adding Detritus to a Nutrient–Phytoplankton–Zooplankton Model: A Dynamical-Systems Approach. *Journal of Plankton Research*, **23**, 389–413, https://doi.org/10.1093/plankt/23.4.389.

Edwards, K. P., R. Barciela, and M. Butenschön, 2012: Validation of the NEMO-ERSEM operational ecosystem model for the North West European Continental Shelf. *Ocean Science*, **8**, 983-1000, https://doi.org/10.5194/os-8-983-2012.

El Moussaoui, A., C. Perruche, E. Greiner, C. Ethé, and M. Gehlen, 2011: Integration of biogeochemistry into Mercator Ocean systems. *Mercator Océan Newsletter*, **40**, 3-14.

Eppley, R. W., 1972: Temperature and phytoplankton growth in the sea. *Fishery Bulletin*, **70**, 1063-1085.

Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, **53**, 343-367, https://doi.org/10.1007/s10236-003-0036-9.

Fasham, M. J. R., H. W. Ducklow, and S. M. McKelvie, 1990: A nitrogen based model of plankton dynamics in the oceanic mixed layer. *Journal of Marine Research*, **48**, 591–639.

Fernandes, J. A., S. Kay, M. A. R. Hossain, M. Ahmed, W. W. L. Cheung, A. N. Lazar, and M. Barange, 2015: Projecting marine fish production and catch potential in Bangladesh in the 21st century under long-term environmental change and management scenarios. *ICES Journal of Marine Science*, **73**, 1357–1369, https://doi.org/10.1093/icesjms/fsv217.

Finkel, Z. V., J. Beardall, K. J. Flynn, A. Quigg, T. A. V. Rees, and J. A. Raven, 2010: Phytoplankton in a changing world: cell size and elemental stoichiometry. *Journal of Plankton Research*, **32**, 119–137, https://doi.org/10.1093/plankt/fbp098.

Flynn, K. J., 2005: Castles built on sand: dysfunctionality in plankton models and the inadequacy of dialogue between biologists and modellers. *Journal of Plankton Research*, **27**, 1205–1210, https://doi.org/10.1093/plankt/fbi099.

Flynn, K. J., D. K. Stoecker, A. Mitra, J. A. Raven, P. M. Glibert, P. J. Hansen, E. Granéli, J. M. Burkholder, 2013: Misuse of the phytoplankton–zooplankton dichotomy: the need to assign organisms as mixotrophs within plankton functional types. *Journal of Plankton Research*, **35**, 3–11, https://doi.org/10.1093/plankt/fbs062.

Follows, M. J., S. Dutkiewicz, S. Grant, and S. W. Chisholm, 2007: Emergent biogeography of microbial communities in a model ocean. *Science*, **315**, 1843-1846, https://doi.org/10.1126/science.1138544.

Ford, D., and R. Barciela, 2015: Marine biogeochemical data assimilation - literature review and scoping report. Met Office Forecasting Research Technical Report 609, 46 pp, https://www.metoffice.gov.uk/learning/library/publications/science/weather-science-technical-reports.

Ford, D., and R. Barciela, 2017: Global marine biogeochemical reanalyses assimilating two different sets of merged ocean colour products. *Remote Sensing of Environment*, **203**, 40-54, https://doi.org/10.1016/j.rse.2017.03.040.

Fox-Kemper, B., 2018: Chapter 2: Notions for the motions of the oceans. In "New Frontiers in Operational Oceanography", E. Chassignet, A. Pascual, J. Tintoré, and J. Verron, Eds., GODAE OceanView.

Friedlingstein, and Coauthors, 2006: Climate-carbon cycle feedback analysis: results from the C^4MIP model intercomparison. *Journal of Climate*, **19**, 3337-3353, https://doi.org/10.1175/JCLI3800.1.

Friedrichs, M. A. M., and Coauthors, 2007: Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups. *Journal of Geophysical Research: Oceans*, **112**, https://doi.org/10.1029/2006JC003852.

Fulton, E. A., 2010: Approaches to end-to-end ecosystem models. *Journal of Marine Systems*, **81**, 171-183, https://doi.org/10.1016/j.jmarsys.2009.12.012.

Gaultier, L., J. Verron, J. M. Brankart, O. Titaud, and P. Brasseur, 2013: On the inversion of submesoscale tracer fields to estimate the surface ocean circulation. *Journal of Marine Systems*, **126**, 33-42, https://doi.org/10.1016/j.jmarsys.2012.02.014.

Gehlen, M., and Coauthors, 2014: Projected pH reductions by 2100 might put deep North Atlantic biodiversity at risk. *Biogeosciences*, **11**, 6955-6967, https://doi.org/10.5194/bg-11-6955-2014.

Gehlen, M., and Coauthors, 2015: Building the capacity for forecasting marine biogeochemistry and ecosystems: recent advances and future developments. *Journal of Operational Oceanography*, **8**, sup1, s168-s187, https://doi.org/10.1080/1755876X.2015.1022350.

Gregg, W. W., M. A. M. Friedrichs, A. R. Robinson, K. A. Rose, R. Schlitzer, K. R. Thompson, and S. C. Doney, 2009: Skill assessment in ocean biological data assimilation. *Journal of Marine Systems*, **76**, 16-33, https://doi.org/10.1016/j.jmarsys.2008.05.006.

Gregg, W. W., and C. S. Rousseaux, 2017: Simulating PACE Global Ocean Radiances. *Frontiers in Marine Science*, **4**, 60, https://doi.org/10.3389/fmars.2017.00060.

Gunson, J. R., S. A. Spall, T. R. Anderson, A. Jones, I. J. Totterdell, and M. J. Woodage, 2006: Climate sensitivity to ocean dimethylsulphide emissions. *Geophysical Research Letters*, **33**, L07701, https://doi.org/10.1029/2005GL024982.

Gutknecht, E., G. Reffray, M. Gehlen, I. Triyulianti, D. Berlianty, and P. Gaspar, 2016: Evaluation of an operational ocean model configuration at 1/12° spatial resolution for the Indonesian seas (NEMO2. 3/INDO12)–Part 2: Biogeochemistry. *Geoscientific Model Development,* **9 (4)**, 1523-1543, https://doi.org/10.5194/gmd-9-1523-2016.

Haines, K., 2018: Chapter 19: Ocean reanalysis. In "New Frontiers in Operational Oceanography", E. Chassignet, A. Pascual, J. Tintoré, and J. Verron, Eds., GODAE OceanView.

Harris, C., 2018: Chapter 16: Coupled atmosphere-ocean modelling. In "New Frontiers in Operational Oceanography", E. Chassignet, A. Pascual, J. Tintoré, and J. Verron, Eds., GODAE OceanView.

Heinze, C., and M. Gehlen, 2013: Chapter 26: Modelling ocean biogeochemical processes and resulting tracer distributions. *Ocean Circulation and Climate*, G. Siedler, S. Griffies, J. Gould, and J. Church, Ed., Academic Press, 667-694.

Hemmings, J. C. P, R. M. Barciela, and M. J. Bell, 2008: Ocean color data assimilation with material conservation for improving model estimates of air-sea CO_2 flux. *Journal of Marine Research*, **66**, 87-126, https://doi.org/10.1357/002224008784815739.

Hernandez, F., and Coauthors, 2018: Chapter 29: Measuring performances, skill and accuracy in operational oceanography: New challenges and approaches. In "New Frontiers in Operational Oceanography", E. Chassignet, A. Pascual, J. Tintoré, and J. Verron, Eds., GODAE OceanView.

Hobday, A. J., C. M. Spillman, P. Eveson, J. Hartog, X. Zhang, and S. Brodie, 2018: A framework for combining seasonal forecasts and climate projections to aid risk management for fisheries and aquaculture. *Frontiers in Marine Science*, **5**, 137, https://doi.org/10.3389/fmars.2018.00137.

Holt, J., and Coauthors, 2014: Challenges in integrative approaches to modelling the marine ecosystems of the North Atlantic: Physics to fish and coasts to ocean. *Progress in Oceanography*, **129**, 285-313, https://doi.org/10.1016/j.pocean.2014.04.024.

Hoteit, I., 2018: Chapter 17: Data assimilation in oceanography: Status and new directions. In "New Frontiers in Operational Oceanography", E. Chassignet, A. Pascual, J. Tintoré, and J. Verron, Eds., GODAE OceanView.

Hyder, K., and Coauthors, 2015: Making modelling count - increasing the contribution of shelf-seas community and ecosystem models to policy development and management. *Marine Policy*, **61**, 291-302, https://doi.org/10.1016/j.marpol.2015.07.015.

Jacobs, G., and Coauthors, 2018: Chapter 18: Operational ocean data assimilation. In "New Frontiers in Operational Oceanography", E. Chassignet, A. Pascual, J. Tintoré, and J. Verron, Eds., GODAE OceanView.

Jacobs, J. M., M. Rhodes, C. W. Brown, R. R. Hood, A. Leight, W. Long, and R. Wood, 2014: Modeling and forecasting the distribution of Vibrio vulnificus in Chesapeake Bay. *Journal of Applied Microbiology*, **117 (5)**, 1312-1327, https://doi.org/10.1111/jam.12624.

Jin, Z., T. P. Charlock, W. L. Smith Jr., and K. Rutledge, 2004: A parameterization of ocean surface albedo. *Geophysical Research Letters*, **31**, L22301, doi:10.1029/2004GL021180.

Johnson, K., and H. Claustre, 2016: The scientific rationale, design, and implementation plan for a Biogeochemical-Argo float array. *Biogeochemical-Argo Planning Group*, 58, https://doi.org/10.13155/46601.

Jones, E. M., and Coauthors, 2016: Use of remote-sensing reflectance to constrain a data assimilating marine biogeochemical model of the Great Barrier Reef. *Biogeosciences*, **13**, 6441-6469, https://doi.org/10.5194/bg-13-6441-2016.

Katija, K., and J. O. Dabiri, 2009: A viscosity-enhanced mechanism for biogenic ocean mixing. *Nature*, **460**, 624-626, https://doi.org/10.1038/nature08207.

Kriest, I., S. Khatiwala, and A. Oschlies, 2010: Towards an assessment of simple global marine biogeochemical models of different complexity. *Progress in Oceanography*, **86**, 337-360, https://doi.org/10.1016/j.pocean.2010.05.002.

Kuhn, A. M., K. Fennel, and J. P. Mattern, 2015: Model investigations of the North Atlantic spring bloom initiation. *Progress in Oceanography*, **138**, 176-193, https://doi.org/10.1016/j.pocean.2015.07.004.

Kwiatkowski, L., and Coauthors, 2014: iMarNet: an ocean biogeochemistry model intercomparison project within a common physical ocean modelling framework. *Biogeosciences*, **11**, 7291-7304, https://doi.org/10.5194/bg-11-7291-2014.

Laufkötter, C., and Coauthors, 2015: Drivers and uncertainties of future global marine primary production in marine ecosystem models. *Biogeosciences*, **12**, 6955-6984, https://doi.org/10.5194/bg-12-6955-2015.

Le Quéré, C., and Coauthors, 2005: Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology*, **11**, 2016–2040, https://doi.org/10.1111/j.1365-2486.2005.1004.x.

Lea, D. J., I. Mirouze, M. J. Martin, R. R. King, A. Hines, D. Walters, and M. Thurlow, 2015: Assessing a new coupled data assimilation system based on the Met Office coupled atmosphere–land–ocean–sea ice model. *Monthly Weather Review*, **143**, 4678-4694, https://doi.org/10.1175/MWR-D-15-0174.1.

Lehodey, P., I. Senina, S. Nicol, and J. Hampton, 2015: Modelling the impact of climate change on South Pacific albacore tuna. *Deep Sea Research Part II: Topical Studies in Oceanography*, **113**, 246-259, https://doi.org/10.1016/j.dsr2.2014.10.028.

Lévy, M., P. Klein, and A. M. Treguier, 2001: Impact of sub-mesoscale physics on production and subduction of phytoplankton in an oligotrophic regime. *Journal of Marine Research*, **59**, 535-565, https://doi.org/10.1357/002224001762842181.

Lévy, M., D. Iovino, L. Resplandy, P. Klein, G. Madec, A. M. Tréguier, S. Masson, and K. Takahashi, 2012a: Large-scale impacts of submesoscale dynamics on phytoplankton: Local and remote effects. *Ocean Modelling*, **43**, 77-93, https://doi.org/10.1016/j.ocemod.2011.12.003.

Lévy, M., L. Resplandy, P. Klein, X. Capet, D. Iovino, and C. Éthé, 2012b: Grid degradation of submesoscale resolving ocean models: Benefits for offline passive tracer transport. *Ocean Modelling*, **48**, 1-9, https://doi.org/10.1016/j.ocemod.2012.02.004.

Lewis, H.W., and Coauthors, 2018: The UKC2 regional coupled environmental prediction system. *Geoscientific Model Development*, **11**, 1–42, https://doi.org/10.5194/gmd-11-1-2018.

Liu, G., and Coauthors, 2018: Predicting Heat Stress to Inform Reef Management: NOAA Coral Reef Watch's 4-Month Coral Bleaching Outlook. *Frontiers in Marine Science*, **5**, 57, https://doi.org/10.3389/fmars.2018.00057.

Losa, S. N., G. A. Kivman, J. Schröter, and M. Wenzel, 2003: Sequential weak constraint parameter estimation in an ecosystem model. *Journal of Marine Systems*, **43**, 31-49, https://doi.org/10.1016/j.jmarsys.2003.06.001.

Madec, G., 2008: NEMO ocean engine. Note du Pôle de modélisation, Institut Pierre-Simon Laplace (IPSL), France, No 27, ISSN No 1288-1619.

Manizza, M., C. Le Quéré, A. J. Watson, and E. T. Buitenhuis, 2005: Bio-optical feedbacks among phytoplankton, upper ocean physics and sea-ice in a global model. *Geophysical Research Letters*, **32**, L05603, https://doi.org/10.1029/2004GL020778.

Martin, G. M., S. F. Milton, C. A. Senior, M. E. Brooks, S. Ineson, T. Reichler, and J. Kim, 2010: Analysis and reduction of systematic errors through a seamless approach to modeling weather and climate. *Journal of Climate*, **23**, 5933-5957, https://doi.org/10.1175/2010JCLI3541.1.

Matear, R. J., and E. Jones, 2011: Marine biogeochemical modelling and data assimilation. *Operational Oceanography in the 21st Century*, A. Schiller, and G. B. Brassington, Ed., Springer, 295-317.

Mattern, J. P., K. Fennel, and M. Dowd, 2012: Estimating time-dependent parameters for a biological ocean model using an emulator approach. *Journal of Marine Systems*, **96**, 32-47, https://doi.org/10.1016/j.jmarsys.2012.01.015.

McClain, C. R., 2009: A decade of satellite ocean color observations. *Annual Review of Marine Science*, **1**, 19-42, https://doi.org/10.1146/annurev.marine.010908.163650.

Mitra, A., and Coauthors, 2014: The role of mixotrophic protists in the biological carbon pump. *Biogeosciences*, **11**, 995-1005, https://doi.org/10.5194/bg-11-995-2014.

Morel, A., 1988: Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters). *Journal of Geophysical Research: Oceans*, **93**, 10749–10768, https://doi.org/10.1029/JC093iC09p10749.

O'Dea, E. J., and Coauthors, 2012: An operational ocean forecast system incorporating NEMO and SST data assimilation for the tidally driven European North-West shelf. *Journal of Operational Oceanography*, **5**, 3-17, https://doi.org/10.1080/1755876X.2012.11020128.

O'Dea, E., and Coauthors, 2017: The CO5 configuration of the 7 km Atlantic Margin Model: large-scale biases and sensitivity to forcing, physics options and vertical resolution. *Geoscientific Model Development*, **10**, 2947-2969, https://doi.org/10.5194/gmd-10-2947-2017.

Orr, J. C., and Coauthors, 2005: Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature*, **437**, 681-686, https://doi.org/10.1038/nature04095.

Orr, J. C., J.-M. Epitalon, and J.-P. Gattuso, 2015: Comparison of ten packages that compute ocean carbonate chemistry. *Biogeosciences*, **12**, 1483-1510, https://doi.org/10.5194/bg-12-1483-2015.

Palmer, J. R., and I. J. Totterdell, 2001: Production and export in a global ocean ecosystem model. *Deep Sea Research Part I: Oceanographic Research Papers*, **48**, 1169-1198, https://doi.org/10.1016/S0967-0637(00)00080-7.

Park, J.-Y., C. A. Stock, X. Yang, J. P. Dunne, A. Rosati, J. John, and S. Zhang, 2018: Modeling global ocean biogeochemistry with physical data assimilation: A pragmatic solution to the equatorial instability. *Journal of Advances in Modeling Earth Systems*, **10 (3)**, 891-906, https://doi.org/10.1002/2017MS001223.

Payne, M. R., and Coauthors, 2017: Lessons from the First Generation of Marine Ecological Forecast Products. *Frontiers in Marine Science*, **4**, 289, https://doi.org/10.3389/fmars.2017.00289.

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

Quinn, P. K., and T. S. Bates, 2011: The case against climate regulation via oceanic phytoplankton sulphur emissions. *Nature*, **480**, 51-56, https://doi.org/10.1038/nature10580.

Raghukumar, K., C. A. Edwards, N. L. Goebel, G. Broquet, M. Veneziani, A. M. Moore, and J. P. Zehr, 2015: Impact of assimilating physical oceanographic data on modeled ecosystem dynamics in the California Current System. *Progress in Oceanography*, **138**, 546-558, https://doi.org/10.1016/j.pocean.2015.01.004.

Redfield, A. C., 1934: On the proportions of organic derivatives in sea water and their relation to the composition of plankton. *James Johnson Memorial Volume*, R. J. Daniel, Ed., University Press of Liverpool, 177-192.

Riebesell, U., A. Körtzinger, and A. Oschlies, 2009: Sensitivities of marine carbon fluxes to ocean change. *Proceedings of the National Academy of Sciences*, **106**, 20602-20609, https://doi.org/10.1073/pnas.0813291106.

Rousseaux, C.S., and W. W. Gregg, 2017: Forecasting Ocean Chlorophyll in the Equatorial Pacific. *Frontiers in Marine Science*, **4**, 236, https://doi.org/10.3389/fmars.2017.00236.

Sarmiento, J.L., and N. Gruber, 2006: *Ocean Biogeochemical Dynamics*. Princeton University Press, 528 pp.

Saux Picart, S., M. Butenschön, and J. D. Shutler, 2012: Wavelet-based spatial comparison technique for analysing and evaluating two-dimensional geophysical model fields. *Geoscientific Model Development*, **5 (1)**, 223-230, https://doi.org/10.5194/gmd-5-223-2012.

Schartau, M., and Coauthors, 2017: Reviews and syntheses: parameter identification in marine planktonic ecosystem modelling. *Biogeosciences*, **14**, 1647-1701, https://doi.org/10.5194/bg-14-1647-2017.

Schiller, A., M. Herzfeld, R. Brinkman, and G. Stuart, 2014: Monitoring, predicting, and managing one of the seven natural wonders of the world. *Bulletin of the American Meteorological Society*, **95 (1)**, 23-30, https://doi.org/10.1175/BAMS-D-12-00202.1.

Schwinger, J., J. Tjiputra, N. Goris, K. D. Six, A. Kirkevåg, Ø. Seland, C. Heinze, and T. Ilyina, 2017: Amplification of global warming through pH dependence of DMS production simulated with a fully coupled Earth system model. *Biogeosciences*, **14**, 3633-3648, https://doi.org/10.5194/bg-14-3633-2017.

Séférian, R., L. Bopp, M. Gehlen, D. Swingedouw, J. Mignot, E. Guilyardi, and J. Servonnat, 2014: Multi-year prediction of Tropical Pacific Marine Productivity. *Proceedings of the National Academy of Sciences*, **111 (32)**, 11646–11651, https://doi.org/10.1073/pnas.1315855111.

Shulman, I., S. Frolov, S. Anderson, B. Penta, R. Gould, P. Sakalaukus, and S. Ladner, 2013: Impact of bio-optical data assimilation on short-term coupled physical, bio-optical model predictions. *Journal of Geophysical Research: Oceans*, **118**, 2215–2230, https://doi.org/10.1002/jgrc.20177.

Siedlecki, S. A., and Coauthors, 2016: Experiments with Seasonal Forecasts of ocean conditions for the Northern region of the California Current upwelling system. *Scientific Reports*, **6**, 27203, https://doi.org/10.1038/srep27203.

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

Six, K. D., S. Kloster, T. Ilyina, S. D. Archer, K. Zhang, and E. Maier-Reimer, 2013: Global warming amplified by reduced sulphur fluxes as a result of ocean acidification. *Nature Climate Change*, **3**, 975-978, https://doi.org/10.1038/nclimate1981.

Skákala, J., D. A. Ford, R. J. Brewin, R. McEwan, S. Kay, B. H. Taylor, L. de Mora, and S. Ciavatta, 2018: The assimilation of phytoplankton functional types for operational forecasting in the North-West European Shelf. Journal of Geophysical Research: Oceans, https://doi.org/10.1029/2018JC014153.

Smyth TJ, I. Allen, A. Atkinson, J. T. Bruun, R. A. Harmer, R. D. Pingree, C. E. Widdicombe, and P. J. Somerfield, 2014: Ocean Net Heat Flux Influences Seasonal to Interannual Patterns of Plankton Abundance. *PLoS ONE*, **9**, e98709, https://doi.org/10.1371/journal.pone.0098709.

Sverdrup, H. U., 1953: On Conditions for the Vernal Blooming of Phytoplankton. *ICES Journal of Marine Science*, **18**, 287–295, https://doi.org/10.1093/icesjms/18.3.287.

Takahashi, T., and Coauthors, 2009: Climatological mean and decadal change in surface ocean pCO_2, and net sea–air CO_2 flux over the global oceans. *Deep Sea Research Part II: Topical Studies in Oceanography*, **56**, 554-577, https://doi.org/10.1016/j.dsr2.2008.12.009.

Taskjelle, T., M. A. Granskog, A. K. Pavlov, S. R. Hudson, and B. Hamre, 2017: Effects of an Arctic under-ice bloom on solar radiant heating of the water column. *Journal of Geophysical Research: Oceans*, **122**, 126–138, https://doi.org/10.1002/2016JC012187.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, **93**, 485-498, https://doi.org/10.1175/BAMS-D-11-00094.1.

Telszewski, M., A. Palacz, and A. Fischer, 2018: Chapter 6: Biogeochemical In-situ Observations – Motivation, Status, and New Frontiers. In "New Frontiers in Operational Oceanography", E. Chassignet, A. Pascual, J. Tintoré, and J. Verron, Eds., GODAE OceanView.

Teruzzi, A., S. Dobricic, C. Solidoro, and G. Cossarini, 2014: A 3-D variational assimilation scheme in coupled transport-biogeochemical models: Forecast of Mediterranean biogeochemical properties. *Journal of Geophysical Research: Oceans,* **119 (1)**, 200-217, https://doi.org/10.1002/2013JC009277.

Thushara, V., and P. N. Vinayachandran, 2016: Formation of summer phytoplankton bloom in the northwestern Bay of Bengal in a coupled physical-ecosystem model. *Journal of Geophysical Research: Oceans,* **121 (12)**, 8535-8550, https://doi.org/10.1002/2016JC011987.

Titaud, O., A. Vidard, I. Souopgui, and F. X. Le Dimet, 2010: Assimilation of image sequences in numerical models. *Tellus A*, **62**, 30–47, https://doi.org/10.1111/j.1600-0870.2009.00416.x.

Turner, A. G., M. Joshi, E. S. Robertson, and S. J. Woolnough, 2012: The effect of Arabian Sea optical properties on SST biases and the South Asian summer monsoon in a coupled GCM. *Climate Dynamics*, **39**, 811-826, https://doi.org/10.1007/s00382-011-1254-3.

Van Leeuwen, P. J., 2009: Particle filtering in geophysical systems. *Monthly Weather Review*, **137**, 4089-4114, https://doi.org/10.1175/2009MWR2835.1.

Vancoppenolle, M., and L. Tedesco, 2015: Numerical models of sea ice biogeochemistry. *Sea Ice*, D. N. Thomas, Ed., John Wiley & Sons, Ltd, https://doi.org/10.1002/9781118778371.ch20.

Vichi, M., N. Pinardi, and S. Masina, 2007: A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part I: Theory. *Journal of Marine Systems*, **64**, 89-109, https://doi.org/10.1016/j.jmarsys.2006.03.006.

Vijith, V., P. N. Vinayachandran, V. Thushara, P. Amol, D. Shankar, and A. C. Anil, 2016: Consequences of inhibition of mixed-layer deepening by the West India Coastal Current for winter phytoplankton bloom in the northeastern Arabian Sea. *Journal of Geophysical Research: Oceans*, **121**, 6583–6603, doi:10.1002/2016JC012004.

Volk, T. and M. I. Hoffert, 1985: Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO2 changes. *The carbon cycle and atmospheric CO2: Natural variations Archean to present*, E.T. Sundquist, and W. S. Broecker, Ed., American Geophysical Union, 99–110.

Volpe, G., B. Buongiorno Nardelli, S. Colella, A. Pisano, and R. Santoleri, 2018: Chapter 9: An operational interpolated ocean colour product in the Mediterranean Sea. In "New Frontiers in Operational Oceanography", E. Chassignet, A. Pascual, J. Tintoré, and J. Verron, Eds., GODAE OceanView.

Wakelin, S. L., Y. Artioli, M. Butenschön, J. I. Allen, and J. T. Holt, 2015: Modelling the combined impacts of climate change and direct anthropogenic drivers on the ecosystem of the northwest European continental shelf. *Journal of Marine Systems*, **152**, 51-63, https://doi.org/10.1016/j.jmarsys.2015.07.006.

Ward, B. A., M. Schartau, A. Oschlies, A. P. Martin, M. J. Follows, and T. R. Anderson, 2013: When is a biogeochemical model too complex? Objective model reduction and selection for North Atlantic time-series sites. *Progress in Oceanography*, **116**, 49-65, https://doi.org/10.1016/j.pocean.2013.06.002.

Waters, J., D. J. Lea, M. J. Martin, I. Mirouze, A. Weaver, and J. While, 2015: Implementing a variational data assimilation system in an operational 1/4 degree global ocean model. *Quarterly Journal of the Royal Meteorological Society*, **141**, 333–349. https://doi.org/10.1002/qj.2388.

Waters, J., M. J. Bell, M. J. Martin, and D. J. Lea, 2017: Reducing ocean model imbalances in the equatorial region caused by data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **143**, 195–208, https://doi.org/10.1002/qj.2912.

While, J., K. Haines, and G. Smith, 2010: A nutrient increment method for reducing bias in global biogeochemical models. *Journal of Geophysical Research: Oceans*, **115**, C10036, https://doi.org/10.1029/2010JC006142.

While, J., I. Totterdell, and M. Martin, 2012: Assimilation of pCO_2 data into a global coupled physical-biogeochemical ocean model. *Journal of Geophysical Research: Oceans*, **117**, C03037, https://doi.org/10.1029/2010JC006815.

Wilson, T.W., and Coauthors, 2015: A marine biogenic source of atmospheric ice-nucleating particles. *Nature*, **525**, 234-238, https://doi.org/10.1038/nature14986.

Wolter, K., and M.S. Timlin, 1993: Monitoring ENSO in COADS with a seasonally adjusted principal component index. Proc. of the 17th Climate Diagnostics Workshop, Norman, OK, NOAA/NMC/CAC, NSSL, Oklahoma Clim. Survey, CIMMS and the School of Meteor., Univ. of Oklahoma, 52-57, https://www.esrl.noaa.gov/psd/enso/mei/WT1.pdf.

Wolter, K., and M. S. Timlin, 1998: Measuring the strength of ENSO events - how does 1997/98 rank? *Weather*, **53**, 315-324, https://doi.org/10.1002/j.1477-8696.1998.tb06408.x.

Xiao, Y., and M. A. M. Friedrichs, 2014: Using biogeochemical data assimilation to assess the relative skill of multiple ecosystem models in the Mid-Atlantic Bight: effects of increasing the complexity of the planktonic food web. *Biogeosciences*, **11**, 3015-3030, https://doi.org/10.5194/bg-11-3015-2014.

Yool, A., E. E. Popova, and T. R. Anderson, 2013: MEDUSA-2.0: an intermediate complexity biogeochemical model of the marine carbon cycle for climate change and ocean acidification studies. *Geoscientific Model Development*, **6**, 1767-1811, https://doi.org/10.5194/gmd-6-1767-2013.

Zeebe, R.E., and D. Wolf-Gladrow, 2001: *CO_2 in seawater: equilibrium, kinetics, isotopes*. Elsevier, 360 pp.