Biogeosciences Discuss., doi:10.5194/bg-2016-304, 2016
Manuscript under review for journal Biogeosciences
Published: 25 July 2016

Observing and modelling phytoplankton community structure in the North Sea: can ERSEM-type models simulate biodiversity?

David A. Ford[1], Johan van der Molen[2], Kieran Hyder[2], John Bacon[2], Rosa Barciela[1], Veronique Creach[2], Robert McEwan[1], Piet Ruardij[3], and Rodney Forster[2,4]

[1]Met Office, FitzRoy Road, Exeter, EX1 3PB, UK
[2]Centre for Environment, Fisheries and Aquaculture Science (Cefas), Pakefield Road, Lowestoft, NR33 0HT, UK
[3]Royal Netherlands Institute for Sea Research (NIOZ), PO Box 59, 1790 AB Den Burg, Texel, The Netherlands
[4]University of Hull, Cottingham Road, Hull, HU6 7RX, UK

Correspondence to: David Ford (david.ford@metoffice.gov.uk)

Abstract. Phytoplankton form the base of the marine food chain, and knowledge of phytoplankton community structure is fundamental when assessing marine biodiversity. Policy makers and other users require information on marine biodiversity and other aspects of the marine environment for the North Sea, a highly productive European shelf sea. This information must come from a combination of observations and models, but currently the coastal ocean is greatly under-sampled for phytoplankton data, and outputs of phytoplankton community structure from models have therefore yet to be properly validated. This study presents a novel set of in situ observations of phytoplankton community structure for the North Sea using accessory pigment analysis. The observations allow a good understanding of the patterns of surface phytoplankton biomass and community structure in the North Sea for the observed months of August 2010 and 2011. Two physical-biogeochemical ocean models, the biogeochemical components of which are different variants of the widely-used European Regional Seas Ecosystem Model (ERSEM), were then validated against these and other observations. Both models were a good match for sea surface temperature observations, and a reasonable match for remotely sensed ocean colour observations. However, the two models displayed very different phytoplankton community structures, with one better matching the in situ observations than the other. Nonetheless, both models shared some similarities with the observations in terms of spatial features and inter-annual variability. A comparison of the formulations and parameterisations of the two models suggests that diversity between the parameter settings of model phytoplankton functional types, along with formulations which promote a greater sensitivity to changes in light and nutrients, is key to capturing the observed biodiversity. These findings will help inform future model development, which should be coupled with detailed validation studies, in order to help facilitate the wider application of marine biogeochemical modelling to user and policy needs.

1 Introduction

Marine biogeochemical model complexity has long been a subject of debate (e.g. Anderson, 2005). Simpler models require fewer, often better understood parameterisations, but omit processes which are known to be important. More complex models explicitly include these processes, but require an increased number of tuneable parameters, the ranges of which are often poorly constrained by observations and/or poorly defined where they aggregate over a range of species. No consensus exists within the community, but recent studies have shown that simple to moderately complex models still do the best job of reproducing basic biogeochemical descriptors such as primary production and carbon fluxes (e.g. Kwiatkowski et al., 2014; Friedrichs et al., 2007; Ward et al., 2013; Xiao and Friedrichs, 2014). Further studies have suggested that "Models of Intermediate Complexity for Ecosystem assessments" (MICE) are the most appropriate for fisheries management (Plagányi et al., 2014).

Ultimately though, models are tools, and the most appropriate tool should be chosen for the task at hand - different scientific, societal and managerial questions will require models of different complexities. For instance, describing a complex coastal environment will likely require the explicit inclusion of processes which are less important when considering, say, global-scale carbon budgets. Furthermore, some users require more detailed information about the marine environment than simple models can provide, necessitating the use of more complex models.

This demand for detailed information applies to the North Sea, with users and policy makers requiring information about topics including eutrophication and nutrient ratios (Skogen et al., 2014), productivity in relation to fisheries (Chassot et al., 2007), harmful and nuisance algal blooms (Blauw et al., 2010; Kurekin et al., 2014), water clarity (Dupont and Aksnes, 2013; Capuzzo et al., 2015), biodiversity (Brandsma et al., 2013), effects of climate change (e.g. Van der Molen et al., 2013; Wakelin et al., 2015a), effects of trawling (Allen and Clarke, 2007; Van der Molen et al., 2013), and impacts of marine renewable energy generation (Van der Molen et al., 2014; Van der Molen et al., 2016a). In particular, indicators of Good Environmental Status (GES) are required in the context of the Marine Strategy Framework Directive (MSFD; Borja et al., 2013). These include descriptors of food-web structure, trophic status, and biodiversity, and elements of these can be assessed with various modelling approaches (Piroddi et al., 2015; Hyder et al., 2015).

Ecosystem models are central to the delivery of marine ecosystem-based management that is specified in existing legislation (e.g. MSFD (EU, 2008), CFP (EU, 2013), WFD (EU, 2000)). These models are important for the design of management measures and to assess the social, economic and environmental performance of management in relation to targets (Defra, 2014; Sutherland et al., 2006). This is done through improving our understanding of the links between pressures (human and environmental) and the response of the system to these pressures. However, ecosystem models are not used frequently in the United Kingdom (UK) and Europe in support of policy and management (Hyder et al., 2015), despite increasing use in the USA and Australia (e.g. Fulton and Link, 2014; Fulton et al., 2007). For models to have a larger impact on policy development and decision-making, modelling approaches need to be more transparent, verifiable, and repeatable than they are at present, as any outputs can be subject to legal challenge (Hyder et al., 2015). Poor uptake of ecosystem models by decision-makers is due to a lack of confidence and understanding of models, relating to use of models, modelling terminology, outputs in the right currency, treatment of uncertainty, rigorous quality standards, and presentation of model products (Hyder et al., 2015). The use of ecosystem models will become increasingly important as the complexity of marine legislation increases (e.g. Boyes and Elliott, 2014). Hence, simple assessment of the skill of models in predicting outcomes (validation - e.g. Mackinson, 2014), model comparisons (e.g. Kwiatkowski et al., 2014), and the clear treatment of the uncertainty associated with predictions (e.g. Thorpe et al., 2015; Gårdmark et al., 2013; Stewart and Martell, 2015; Tebaldi and Knutti, 2007) are needed to increase the confidence in and uptake of models (Hyder et al., 2015).

At the base of the marine food chain are phytoplankton, and phytoplankton community structure is a fundamental consideration in any assessment of marine biodiversity (Garmendia et al., 2013). Changes in community structure can result from large-scale environmental changes such as temperature rises or eutrophication, with different organisms favouring different conditions. Some organisms that favour changed conditions may be harmful to human health (Roselli and Basset, 2015; Bruggeman, 2009). Alternatively, top-down control by benthic or pelagic grazers can change the size-structure of phytoplankton by selective removal of larger species, resulting for instance in an increased proportion of pico-phytoplankton in areas with dense shellfish aquaculture (Smaal et al., 2013). Phytoplankton vary in size by up to nine orders of magnitude for cell volume (Finkel et al., 2010), with variations in community structure reflected in the size and species of their predators, and the number of links in the food chain (Ryther, 1969; Chavez et al., 2011). Larger cells such as diatoms are consumed directly by copepod grazers, giving a higher transfer of energy and ultimately impacting commercial fish stocks (Jennings and Collingridge, 2015). As the physical structure of the North Sea becomes increasingly well understood due to advances in hydrodynamic modelling (Van Leeuwen et al., 2015) and availability of long-term observations (Greenwood et al., 2010; Núñez-Riboni and Akimova, 2015), the potential to predictively model plankton population structure and distribution increases too.

A common way to model plankton community structure is to take a phytoplankton functional type (PFT) approach, such as is done in variants of the European Regional Seas Ecosystem Model (ERSEM; Baretta et al., 1995). This approach groups phytoplankton into a number of PFTs, based on their general function within the ecosystem (Le Quéré et al., 2005). If information on phytoplankton community structure is to be modelled and provided to users, then it must be validated. Commonly though, validation studies go no further than total chlorophyll concentration (e.g. Edwards et al., 2012; De Mora et al., 2013). This is largely because there is a lack of observations that contain more detail about community structure against which to compare. Algorithms for deriving phytoplankton community structure from remotely sensed satellite ocean colour observations, either in the form of PFTs or phytoplankton size classes (PSCs) are being developed (Brewin et al., 2011; Brito et al., 2014), but have not yet reached maturity and are not yet widely available to the general scientific community. Moreover, such remote sensing products require a similar level of validation (Brotas et al., 2013). In situ observations are sparse, particularly in shelf seas, and the measured variables may not be easily matched to model outputs, which do not always aggregate neatly over species or size classes.

This study presents a novel set of in situ observations of phytoplankton community structure in the North Sea using accessory pigment analysis (Sherrard et al., 2006), noting that coastal seas are greatly under-represented in the existing global collection of pigment data (Peloquin et al., 2013). Pigment data were analysed so as to give the relative distribution of different size classes, allowing a robust comparison with outputs from ERSEM-type models. Two variants of ERSEM, run by the Centre for Environment, Fisheries and Aquaculture Science (Cefas) and the Met Office, both public bodies in the UK, were validated against these and other observations. The aims of the study were to determine what these new observations add to current scientific understanding of North Sea biogeochemistry, assess the extent to which the models can reproduce the observations, and discuss the implications for current and future user and policy needs, observing strategies and model development.

2 Observations

The International Bottom Trawl Survey (IBTS) is a multi-national ecological research effort established by the International Council for the Exploration of the Sea (ICES) in the early 1970s. Surveys using fisheries research vessels currently take place in the first and third quarter of the year and cover the entire North Sea, using standardised sampling gears and protocols. With cruise lengths of typically 6-8 weeks, each vessel undertakes a gridded survey of the North Sea, repeated each year, in which stations are sampled for groundfish (the primary target of the survey), but also secondary targets such as benthos, seabed litter, and hydrographic parameters. Individual station sampling is often accompanied by visual seabird and cetacean estimates, underway acoustics, and online monitoring of near-surface water quality using FerryBox-type instruments (Petersen et al., 2008). The IBTS thus fits the needs of a multi-disciplinary survey capable of collecting data on human pressures and ecosystem responses for legislation such as the MSFD (http://www.jpi-oceans.eu/multi-use-infrastructure-monitoring). The open data policy of ICES has resulted in many significant publications in fisheries research (Jennings et al., 2002; Daan et al., 2005) and fisheries policy (Rombouts et al., 2013; Shephard et al., 2015).

Prior to 2010, phytoplankton had not been systematically sampled on the UK IBTS. Advances in the autonomous sampling and detection of particles in the water column (e.g. online flow cytometry (Thyssen et al., 2015)), and also the need for high-quality in situ data for validation of satellite remote sensing data, indicated that the addition of phytoplankton to the survey would be beneficial. Hence, sampling of PFTs using high-pressure liquid chromatography (HPLC) - pigment fingerprinting, and analytical flow cytometry (results reported elsewhere) were initiated on the third quarter IBTS cruise of the RV *Cefas Endeavour* in August-September 2010 and subsequent years.

Seawater samples from depths of 4 m ('surface') were collected using 10 l Niskin bottles when weather conditions permitted, or from the ship's bow-intake flow-through clean seawater supply during adverse weather conditions. A known amount of water, typically 1000 ml, was passed through a 200 µm gauze to remove larger zooplankton and debris, then filtered within 1 h on 47 mm GFF filters, which were folded in half, wrapped in aluminium foil and snap frozen in liquid nitrogen dry shippers. On return to shore, samples were transferred to a -80 °C freezer for a storage period of 1-2 months before shipping of samples on dry ice to an accredited HPLC laboratory (DHI Water Quality Institute; Horsholm, Denmark) for chlorophyll-a (chl-a) quantification and full accessory pigment analysis (Schlüter et al., 2011).

Pigment data from the surface stations were quality data controlled in several steps: first, with an initial comparison of HPLC chl-a against independent measures of chlorophyll fluorescence from the fluorometers on the ship's Ferrybox and CTD system. This step corrected a small number of mis-labelled samples. In a second step, anomalies within a sample were detected using methods described by Aiken et al. (2009), e.g. regression of total accessory pigments against chl-a concentration and search for outliers.

Diagnostic pigment analysis was then used on the quality-controlled data set to relate the composition of specific accessory pigments to the relative contribution of different size classes to the total phytoplankton biomass. The designation of specific accessory pigments to algal taxonomic groups of different size, e.g. fucoxanthin and peridinin for large-cell diatoms and dinoflagellates, has been widely established in the biological oceanographic literature (Uitz et al., 2006, 2008). The equations used to estimate the contribution of pico-phytoplankton (0-2 µm), nano-phytoplankton (2-20 µm) and micro- or net phytoplankton (>20 µm) were later modified by Hirata et al. (2008, 2011) and Brewin et al. (2010). The various methods differ in the degree to which the marker pigments chlorophyll-b (chl-b) and 19-hex-fucoxanthin (19-hex) are attributed to the three size classes. Here, chl-b and 19-hex were assigned equally to the pico-phytoplankton and nano-phytoplankton size classes. Results are expressed as a proportion of the total chl-a concentration for each station.

3 Models

Two different physical-biogeochemical modelling systems were used in this study: GETM-ERSEM-BFM, run by Cefas, and NEMO-ERSEM, run by the Met Office. Each is described in turn below, followed by a discussion of their differences and similarities. Existing configurations of each model were used, with no attempt made to increase their similarity.

3.1 GETM-ERSEM-BFM

GETM (General Estuarine Transport Model) is a public domain, three-dimensional (3D) finite difference hydrodynamical model (Burchard and Bolding, 2002; http://www.getm.eu). It solves the 3D partial differential equations for conservation of mass, momentum, salt and heat. The ERSEM-BFM (European Regional Seas Ecosystem Model - Biogeochemical Flux Model) version used here is a development of the model ERSEM III (see Baretta et al., 1995; Ruardij and Van Raaphorst, 1995; Ruardij et al., 1997; Vichi et al., 2003; Vichi et al., 2004; Ruardij et al., 2005; Vichi et al., 2007; Van Leeuwen et al., 2013; Van der Molen et al., 2013; Van der Molen et al., 2014; Van der Molen et al., 2016a; http://www.nioz.nl/northsea_model), and describes the dynamics of the biogeochemical fluxes within the pelagic and benthic environment. The ERSEM-BFM model simulates the cycles of carbon, nitrogen, phosphorus, silicate and oxygen, and allows for variable internal nutrient ratios inside organisms, based on external availability and physiological status. The model applies a functional group approach and contains six phytoplankton groups, four zooplankton groups, and five benthic groups, the latter comprising four macrofauna and one meiofauna groups. Pelagic and benthic aerobic and anaerobic bacteria are also included. The pelagic module includes a number of processes in addition to those included in the oceanic version presented by Vichi et al. (2007) to make it suitable for temperate shelf seas: (i) a parameterisation for diatoms allowing growth in spring, (ii) enhanced transparent exopolymer particles (TEP) excretion by diatoms under nutrient stress, (iii) the associated formation of macro-aggregates consisting of TEP and diatoms, leading to enhanced sinking rates and a sufficient food supply to the benthic system especially in the deeper offshore areas (Engel, 2000), (iv) a *Phaeocystis* functional group for improved simulation of primary production in coastal areas (Peperzak et al., 1998, Ruardij et al., 2005), (v) a new suspended particulate matter (SPM) resuspension module that responds to combined currents and surface waves, and uses a concentration-dependent settling velocity for improved simulation of the under-water light climate (Van der Molen et al., 2016b), and (vi) resuspension of particulate organic material coupled to the SPM model. The model includes a 3-layer benthic module comprising 53 state variables, which enables it to resolve a high level of detail of benthic processes and benthic-pelagic coupling. New features of the benthic model are: (i) benthic diatoms, and (ii) active oxygen uptake of deposit feeders from the water column.

The model setup for the North Sea uses a spherical grid with a spatial resolution of approximately 11 km and 25 layers in the vertical (see http://www.nioz.nl/northsea_model and also Lenhart et al., 2010). The model was forced with tidal boundary conditions from a shelf-scale model, temperature and salinity boundary conditions from a global hindcast (ECMWF-ORAS4; Balmaseda et al., 2013; Mogensen et al., 2012; http://www.ecmwf.int/products/forecasts/d/charts/oras4/reanalysis/), climatological nutrient boundary conditions, observations-based river run-off and riverine nutrient loads (the National River Flow Archive (www.ceh.ac.uk/data/nrfa/index.html) for UK rivers, the Agence de l'eau Loire-Bretagne, Agence de l'eau Seine-Normandie and IFREMER for French rivers, the DONAR database for Netherlands rivers, ARGE Elbe, the Niedersächsisches Landesamt für Ökologie and the Bundesanstalt für Gewässerkunde for German rivers, and the Institute for Marine Research, Bergen, for Norwegian rivers; see also Lenhart et al., 2010), and atmospheric forcing from the ECMWF ERA-40 and operational hindcast (ECMWF, 2006a,b). Details of the model configuration and forcing are given in Table 1.

3.2 NEMO-ERSEM

The hydrodynamic component of the Met Office modelling system is NEMO (Nucleus for European Modelling of the Ocean; Madec, 2008). NEMO is an open source community model originally developed for global ocean modelling (e.g. Storkey et al., 2010), but which has also been recently developed for use in shelf seas (O'Dea et al., 2012). The version used in this study (CO5; O'Dea et al., in prep.) is based on NEMO v3.4, and is a development of that described in O'Dea et al. (2012) and Edwards et al. (2012).

The version of ERSEM used is an alternative development of the original code of Baretta et al. (1995), led by Plymouth Marine Laboratory (PML), and is described in detail by Blackford et al. (2004) and Edwards et al. (2012). The pelagic component includes four phytoplankton and three zooplankton functional groups, and one bacterial group. The benthic component includes aerobic and anaerobic bacteria, suspension feeders, bottom feeders, and the meiobenthos. SPM is simulated as per Sykes and Barciela (2012).

As part of the FOAM (Forecasting Ocean Assimilation Model; Blockley et al., 2014) suite of models, NEMO-ERSEM is run operationally at the Met Office on a daily basis, providing five-day forecasts of physical and biogeochemical variables for the North-West European Shelf Seas. Analyses and forecasts are publicly available through the Copernicus Marine Environment Monitoring Service (CMEMS; http://marine.copernicus.eu), which is the operational service building on the MyOcean project. Physical and biogeochemical reanalysis products (Wakelin et al., 2015b) are also available through CMEMS, and results from the recent NEMO-ERSEM reanalysis were used in this study.

Details of the model configuration are summarised in Table 1. The model was run on the 7 km resolution Atlantic Meridional Margin (AMM7) domain, covering the entire North-West European Shelf Seas, including the North Sea. There are 51 vertical levels, using a hybrid σ-S coordinate system with the stretching function of Siddorn and Furner (2013). This uses terrain-following coordinates whilst ensuring a fixed surface resolution of 1 m. Remotely sensed and in situ observations of sea surface temperature (SST) were assimilated using a 3D-Var implementation of the NEMOVAR data assimilation scheme (Waters et al., 2015; O'Dea et al., 2012). River inputs were taken from the E-HYPE model (Donnelly et al., 2015) for flow values, and from the same climatology as in Edwards et al. (2012) for nutrients and sediments. Lateral boundary conditions for physical variables were taken from a reanalysis of the GloSea5 seasonal forecasting system (MacLachlan et al., 2014) at the Atlantic boundaries, and from the IOW-GETM model (Stips et al., 2004) at the Baltic boundary. For biogeochemistry, lateral boundary conditions for nutrients were taken from the World Ocean Atlas monthly climatology (Garcia et al., 2010) at the Atlantic boundaries, and zero flux boundary conditions were applied at the Baltic boundary. Zero flux boundary conditions were applied for all other biogeochemical variables at all boundaries. Surface forcing was from the ERA-Interim reanalysis (Dee et al., 2011). The NEMO-ERSEM reanalysis covers the period January 1985 to July 2012, but for practical reasons was run in three sections. The final section, which this study uses, started in November 2003, with physics initial conditions taken from the corresponding date of a non-assimilative hindcast of the entire reanalysis period, and biogeochemical initial conditions taken from a winter date of the run of NEMO-ERSEM described in Edwards et al. (2012).

3.3 Comparison of the two models

Even though both models used versions of ERSEM, it is reasonable to expect differences in the results. Such differences are inevitably a result of the accumulation of differences between the models. It should be noted that both models were run as usual, and no attempts were made to increase similarity. To help understand the differences in model behaviour, this section summarises the main differences between the two models. We focus on two types of differences: general level differences (Table 1), and differences in phytoplankton parameters and parameterisations (Table 2). For the sake of readability, and to limit repetition, the following summary is kept at a fairly basic level; for (numerical) detail the reader is referred to the tables.

The two hydrodynamics models were different, and in general used different domains, resolutions, and forcing data. The NEMO-ERSEM model had a larger domain, at higher resolution, and used more advanced atmospheric forcing. Moreover, in NEMO-ERSEM, SST was assimilated, while GETM-ERSEM-BFM had no data assimilation. NEMO-ERSEM's river runoff originated from a model, that of GETM-ERSEM-BFM from observations. The GETM-ERSEM-BFM model used time series of riverine nutrient inputs whereas the NEMO-ERSEM model used a climatology. The SPM model of NEMO-ERSEM contained explicit size fractions and cohesive interactions, but was only forced by flow velocities, while that of the GETM-ERSEM-BFM model was non-cohesive, with implicit size-related behaviour and included resuspension by both currents and waves (Van der Molen et al., 2016b). The models also used different initial conditions and spin-up sequences.

Both ERSEM versions share a common origin, both use the same base nutrients (N, P, Si, C), and are both based on a functional type approach. They share four phytoplankton types, three zooplankton types, and a basic bacteria type. Both have a three-layered benthic module, with similar nutrient regeneration mechanisms.

ERSEM-BFM had a number of additional functional types compared to PML-ERSEM: *Phaeocystis* colonies, benthic diatoms, carnivorous zooplankton, filter feeder larvae, epibenthos, benthic predators, and benthic and pelagic nitrifying bacteria. Furthermore, it used a CO_2 module, whereas in PML-ERSEM this was switched off.

The models used different methods for nutrient affinity, nutrient stress and sinking, and light susceptibility. For nutrient affinity, ERSEM-BFM used 10-100 times higher affinity values for nutrient uptake. There are two ways to measure phytoplankton nutrient uptake in experiments (Veldhuis and Admiraal, 1987): (i) a short-duration experiment in which nutrients are added to nutrient-deprived algal cultures and uptake rates into the internal nutrient buffer are measured; and (ii) an experiment that lasts a full day in which uptake rates needed for daily growth are measured. The parameters for GETM-ERSEM-BFM were based on short-duration experiments, whereas those for NEMO-ERSEM were based on long-duration experiments. The short-duration parameterisation allows for improved incorporation of the dependencies of cell properties such as cell size and buffer capacity. These features were needed to resolve the competition between diatoms and *Phaeocystis* colonies during excessive spring blooms in the Dutch coastal zone, which terminate through phosphate depletion. In ERSEM-BFM, nutrient stress of pelagic diatoms leads to excretion of all (new fixated) organic C that cannot be used for growth as carbohydrates (TEP). At high levels of diatoms, this excretion leads to the simulation of the effect of macro-aggregate formation through binding by these carbohydrates, through increases in the sinking rate of live and dead particulate matter. PML-ERSEM used a more implicit approach to sinking. For light susceptibility, both models used a photosynthesis-irradiance (P-I) curve approach, but PML-ERSEM defined it through the initial slope (alpha), whereas ERSEM-BFM defined it through the light intensity at maximum photosynthetic rate (Ke). For several elements where both models used the same approach, parameter settings were different: maximum productivity, respiration, excretion, minimum quota for P, lysis, and C:Chl ratios. For these, there was typically more differentiation in settings between phytoplankton functional types in ERSEM-BFM than in PML-ERSEM.

3.4 Aggregating model PFTs to match observed PSCs

To allow validation of phytoplankton community structure from the models against the IBTS observations, the four PFTs from NEMO-ERSEM and six PFTs from GETM-ERSEM-BFM must be appropriately aggregated to match the observed PSCs. Diatoms (both models), dinoflagellates (both models) and resuspended benthic diatoms (GETM-ERSEM-BFM only) were considered to be micro-phytoplankton. Flagellates (both models) and *Phaeocystis* colonies (GETM-ERSEM-BFM only) were considered to be nano-phytoplankton. The pico-phytoplankton PFT (both models) was directly mapped to the pico-phytoplankton PSC. For consistency with the IBTS observations, the PFTs and PSCs were expressed as fractions of total chlorophyll concentration, rather than biomass.

4 Results

4.1 Observations

Each year, the IBTS cruise starts in early August in the Southern Bight of the North Sea off the Thames estuary (51.5° N) and proceeds northwards via a series of longitudinal transects, with each transect taking between one to three days, depending upon the width of the North Sea at each point. The final transect between the Shetland Islands and Norway at 61° N was reached by early September for the 2010 and 2011 IBTS cruises. The spatially-averaged annual mean surface temperature for the North Sea was 9.9 °C in 2010 and 10.0 °C in 2011, which were very close to the long-term average of 10.0 °C for the 1985-2014 period. Hence, the years surveyed represent near-average conditions for temperature.

A continuous recording of chlorophyll fluorescence showed good agreement with the quantity of extracted chl-a determined by HPLC ($r^2 = 0.64$ for 2010 and 0.65 for 2011). The number of same-day match-ups between in situ chl-a and satellite-derived

Biogeosciences Discuss., doi:10.5194/bg-2016-304, 2016
Manuscript under review for journal Biogeosciences
Published: 25 July 2016

chl-a was low for both years, but a comparison of eight-day averaged surface chl-a from MERIS with in situ data showed an excellent qualitative agreement for both years (Fig. 1). Satellite coverage was more complete in 2011 than 2010. Time series plots and maps of the two cruises showed a number of regularly-occurring features that can be observed at this time of year (labelled 'A' to 'J' on Fig. 1).

A zone of high chl-a (> 2 mg m^{-3}) was observed with all methods in the coastal waters of Belgium, The Netherlands, Germany and Denmark. This zone extended between points 'A' and 'B' for the map of 2010, and points 'F' and 'H' for 2011. High chlorophyll values (> 2 mg m^{-3}) were observed in the outer Thames estuary and close to the English east coast as far north as the Humber estuary, but the English coastal zone was not as clearly demarked by high chl-a as the continental coast. The continuous recording of the first 7-10 days of the IBTS thus alternated between moderate chl-a (1 to 2 mg m^{-3}) and high chl-a as the vessel covered the southern North Sea between 51.5° N and 55° N. An exceptional bloom event with chl-a of over 6 mg m^{-3} was recorded at location 'G' in 2011, and was clearly visible in MERIS and MODIS images.

The central section of the North Sea between 55° N and 58° N was covered during the second and third weeks of the IBTS. This section showed low chl-a values (< 1 mg m^{-3}) across most of the zone (Fig. 1), particularly in the region north of 'I', 56.5° N to 58.5° N, 0° E to 3° E, which was a large region with values < 0.5 mg m^{-3}. To the east, the Danish coastal waters ('B' and 'H') showed high chl-a. The inshore English coast north of the Humber, and Scottish coastal waters, are low in chl-a compared to those further south. A moderate chl-a bloom was evident in the chlorophyll fluorescence trace, MERIS image and extracted chl-a at position 'C' in 2010, and a high chl-a patch was evident close to the Scottish coast at Aberdeen at position 'D'.

The northern North Sea was sampled in weeks three and four (from 28th to 29th August onwards) and was similar in 2010 and 2011. An arc of high chl-a was detected from north of the Scottish mainland through the Orkneys and Shetlands, e.g. from 'D' to 'E' in 2010, with particularly high values at 'E'. In 2011, high values were observed from the Orkneys through to north of the Shetlands at 'J'. The ferrybox chlorophyll fluorescence recorded a further large bloom on 6th September 2011, but this event was not sampled for pigments.

PFTs were determined on the basis of accessory pigment composition. In general, pigment diversity was lower in coastal waters and in the southern North Sea and reached peak diversity in the stratified central North Sea. Fucoxanthin was the dominant accessory marker pigment in the southern North Sea, and 19-hex was dominant in the northern North Sea. Pico-phytoplankton were represented in this analysis by the marker pigments zeaxanthin, chl-b and 19-hex; these pigments were rare in the southern North Sea below a line from East Anglia to the Wadden Sea, hence pico-phytoplankton contribution was estimated in this region to be less than 10 % of total phytoplankton biomass (Fig. 2). The contribution of pico-phytoplankton increased with increasing latitude so that the area with highest contribution from the smallest PFT was found in both years to be located north of 57° N and east of 0° E. Nano-phytoplankton were represented by the pigments 19-hex, 19-but, alloxanthin and chl-b. The highest percentage contribution of nano-phytoplankton was found in both years to be located in the central and northern North Sea, including the high chl-a regions around the Shetland and Orkney islands. The largest PFT, micro-phytoplankton, were represented by the pigments fucoxanthin and peridinin. The distribution of this group showed highest abundance in the southern North Sea high chl-a regions near the continental coast, but also in location 'G' (2011) and between 'C' and 'D' in 2010.

The combination of continuous underway logging with autonomous instruments, high precision pigment measurements at selected stations, and good satellite earth observation coverage allowed the patterns of surface phytoplankton biomass and PFT distribution in the North Sea to be well understood. Together, this provided a solid observational base with which to test biogeochemical model accuracy.

4.2 Model validation - domain-scale

This section presents validation of physical and biogeochemical model variables against a range of observation-based products, in order to assess the models' skill at broader scales than the IBTS observations measured. SST has been validated against OSTIA (Operational Sea Surface Temperature and Sea Ice Analysis; Donlon et al., 2012), which is an objective analysis product based on remotely sensed and in situ SST observations. Sea surface chlorophyll and SPM have been validated against remotely sensed ocean colour products, developed by Ifremer using the OC5 algorithm (Gohin et al., 2002, 2005, 2008). Due to the limited availability of observations, nutrient concentrations have been validated against the 1° resolution World Ocean Atlas climatologies (Garcia et al., 2010). Detailed validation of phytoplankton community structure against the IBTS observations follows in Sect. 4.3.

Since the focus of this study is on the phytoplankton community structure in August 2010 and 2011, most of the validation presented here is for these two months. For more general model validation the reader is referred to Edwards et al. (2012) for NEMO-ERSEM, and Lenhart et al. (2010), Aldridge et al. (2012), Van Leeuwen et al. (2013), Van der Molen et al. (2013), and Van der Molen et al., (2016a,b) for GETM-ERSEM-BFM in various configurations. However, some statistical assessment has been performed here for SST, chlorophyll and SPM over the period March 2010 to October 2011. Two seasons have been defined for this assessment: the growing season and winter. The growing season is defined as March to October, and is averaged over 2010 and 2011. Winter is defined as November 2010 to February 2011. Statistics have been calculated in observation space by performing a bilinear interpolation of the daily mean model fields to the observation locations. Calculations have been performed for $\log_{10}$(chlorophyll) rather than for chlorophyll in order to provide a more Gaussian distribution (Campbell, 1995).

Taylor plots (Taylor, 2001) of SST, $\log_{10}$(chlorophyll) and SPM are shown in Fig. 3. SST is a good match for the observations in both the growing season and in winter, although lower correlations are found for both models in August 2010 and 2011 than for the whole seasons. Slightly better statistics are obtained for NEMO-ERSEM than for GETM-ERSEM-BFM, reflecting the assimilation of SST data into NEMO-ERSEM. The statistics for $\log_{10}$(chlorophyll) differ more between models and between seasons, and the models are a less good match for the observations than with SST, as is common in physical-biogeochemical models. With SPM, the two models show large differences in variability. GETM-ERSEM-BFM has a much higher standard deviation than the observations in both seasons, especially the growing season, whilst the standard deviation of NEMO-ERSEM is too low all year round.

Maps of mean SST for August 2010 and 2011, the months during which most of the IBTS observations were collected, are plotted in Fig. 4, from GETM-ERSEM-BFM, NEMO-ERSEM, and OSTIA. There is a great deal of similarity between NEMO-ERSEM and OSTIA, which is unsurprising since NEMO-ERSEM assimilates SST data, but both NEMO-ERSEM and GETM-ERSEM-BFM are able to simulate the spatial features seen in OSTIA, as well as the inter-annual variability between 2010 and 2011. Boundary effects can be seen in GETM-ERSEM-BFM, which has a smaller domain.

Similar maps for sea surface chlorophyll concentration are plotted in Fig. 5, from the two models and the OC5 products. Daily ocean colour coverage is incomplete due to cloud cover, so the observations plotted here are simply a composite of all observations available during the month, rather than a true monthly mean. To ensure a fair comparison, the daily mean model fields were bilinearly interpolated to observation locations, and equivalent composites plotted rather than the true model mean. Van der Molen et al. (2016b) presented a comparison of sub-sampled model results, accounting for cloud cover, of SPM with the true model mean, which suggested noticeable differences in winter, but only small differences in summer. The match between the models and the observations is less good for chlorophyll than for SST, but both models were still able to capture the main observed features throughout the domain. In the central and northern North Sea, which has the lowest chlorophyll concentrations, values were generally under-estimated by GETM-ERSEM-BFM and over-estimated by NEMO-ERSEM. High coastal chlorophyll values were better simulated by GETM-ERSEM-BFM, whilst the Norwegian Trench is better represented by NEMO-ERSEM. GETM-ERSEM-BFM has more spatial variability than NEMO-ERSEM, despite having a lower model resolution. As with SST, there is notable inter-annual variability in the observations, with higher chlorophyll concentrations in 2011 than 2010. Both models captured this variability, although it is less evident in GETM-ERSEM-BFM, and over-pronounced in NEMO-ERSEM.

The models are similarly compared to the OC5 SPM products in Fig. 6. NEMO-ERSEM and GETM-ERSEM-BFM both under-estimate SPM in the central and northern North Sea, with NEMO-ERSEM also under-representing the plume of SPM off south-east England. Overall, the two models give very different results for SPM, and the reasons for and potential consequences of this are discussed in Sect. 5.

Maps of mean surface nitrate, phosphate and silicate for each model for August 2010 are shown in Fig. 7, alongside the corresponding World Ocean Atlas climatology fields. Only 2010 is plotted because very similar patterns are seen in the models for both years, and the climatologies do not include inter-annual variability. It should also be noted that the climatologies are of relatively coarse 1° resolution, so provide only a basic representation of the North Sea, but are the only source of data with full spatial coverage available for such a comparison. For nitrate, the main limiting nutrient in the North Sea, GETM-ERSEM-BFM shows high coastal concentrations, and very low concentrations elsewhere. NEMO-ERSEM has a similar pattern, but with a much less extreme range of values. Both are in broad agreement with the climatology. Likewise for phosphate and silicate, NEMO-ERSEM and GETM-ERSEM-BFM show differing distributions to each other, but nonetheless are both in broad agreement with the magnitudes and spatial patterns of the climatologies.

4.3 Model validation against PFT observations

This section presents validation of the models against the chlorophyll and PFT observations collected on the IBTS cruises. Hereafter, Micro is used to refer to the fraction of total chlorophyll represented by the micro-phytoplankton size class, and similarly Nano and Pico. The most northerly of the IBTS observations were located outside the GETM-ERSEM-BFM model domain; these have been excluded from the assessment to ensure the same observations were used to validate both models.

Maps of mean surface model PFTs for August 2010, as fractions of total chlorophyll, are plotted in Fig. 8. These show very different distributions for NEMO-ERSEM and GETM-ERSEM-BFM, much more so than the difference in total chlorophyll might suggest. In particular, NEMO-ERSEM shows dominance by pico-phytoplankton in the southern North Sea, similar fractions of pico-phytoplankton and flagellates in the rest of the domain, and generally low concentrations of diatoms and dinoflagellates. In contrast, GETM-ERSEM-BFM shows dominance by diatoms in coastal regions, and by pico-phytoplankton in the centre of the domain, with generally lower fractions of the remaining PFTs. The two PFTs unique to GETM-ERSEM-BFM, resuspended benthic diatoms and *Phaeocystis* colonies, only show notable concentrations in certain coastal areas. The reasons for the differences between the two models are discussed in Sect. 5, but they are likely to have arisen from the differences in parameterisations detailed in Sect. 3.3.

The bias, root mean square error (RMSE) and correlation of modelled versus observed total chlorophyll concentration are shown in Table 3. GETM-ERSEM-BFM has a slight bias towards too high chlorophyll, whilst the bias for NEMO-ERSEM is near-zero. GETM-ERSEM-BFM chlorophyll values are typically higher than those from NEMO-ERSEM, but also show a greater range. These features are consistent between the two years. However, whilst GETM-ERSEM-BFM has a similar correlation value for both years, the correlation for NEMO-ERSEM is much higher in 2010 than 2011. It should be noted though that these statistics are based on a relatively small number of points, so any conclusions drawn from this comparison are not guaranteed to be robust, particularly given the domain-scale spatial variability (see Sect. 4.2 and Fig. 5).

A comparison of phytoplankton community structure in the models and IBTS observations has been made by aggregating the model PFTs into the three observed PSCs, as described in Sect. 3.4. The distribution of relative PSC fractions with total chlorophyll is plotted for each data set in Fig. 9. Consistent with results from previous studies (e.g. Devred et al., 2011), as observed chlorophyll increases, Micro tends to increase, and Nano and Pico decrease. This pattern is also seen to some extent in GETM-ERSEM-BFM, but less so in NEMO-ERSEM (and only in 2010), although NEMO-ERSEM has a smaller range of chlorophyll concentrations. In the IBTS data there is a clear overall dominance of Micro. This is well reproduced by GETM-ERSEM-BFM, but the opposite is found in NEMO-ERSEM. The exception to this is a group of observations at low chlorophyll concentrations, most notably in 2011, in which Micro is least abundant, better matching typical NEMO-ERSEM results. These observations were all taken in the central North Sea, and this behaviour is discussed further in Sect. 5.

Three variables which always sum to one can be displayed in a single space, barycentric coordinates, using a ternary plot (e.g. Jupp et al., 2012). Phytoplankton community structure is plotted this way in Fig. 10. The observations form a distinct line in this space, from the centre of the plot to the corner representing dominance by Micro. At lower chlorophyll concentrations (not shown in Fig. 10, but consistent with Fig. 9) there are roughly equal fractions of Micro, Nano and Pico. As chlorophyll concentration increases, Nano and Pico decrease in roughly equal amounts, with Micro increasing accordingly. GETM-ERSEM-BFM displays a similar pattern, with values in the same area of the plot as the observations, although with a less distinct relationship. NEMO-ERSEM values show a very different distribution however. There is some overlap with the observations in 2011, but otherwise a much less Micro-dominated regime is evident.

The ternary plot can also be used as a colour key to produce a map of phytoplankton community structure, as in Fig. 11. This plots the August mean community structure for each model and each year, overlaid with the IBTS point observations in circles. Plotting the community structure in such a fashion demonstrates that whilst GETM-ERSEM-BFM and NEMO-ERSEM give very different results in terms of the magnitudes of the PSC fractions, there are nonetheless some broadly consistent features in terms of spatial patterns, which are also evident to some extent in the observations. For instance, both models (although NEMO-ERSEM less so in 2010) show a distinct split in community structure between the southern and northern North Sea, and around bathymetric features such as Dogger Bank and coastlines. Such a split can be clearly seen in the 2011 observations, which show very little variation throughout the central North Sea, but is less clear in the 2010 observations. A difference between the years in the community structure in the central North Sea is also seen in NEMO-ERSEM, and to a lesser extent GETM-ERSEM-BFM, although the direction of change in the models is from Pico-dominated to Micro-dominated, the opposite of that in the observations. Although in most cases the community structure in the models does not match that of the observations, GETM-ERSEM-BFM is a very good match for the observations in the southern North Sea, an area particularly dominated by diatoms in the model. Silicate in this region is near-depleted in GETM-ERSEM-BFM (see Fig. 7), but abundant in NEMO-ERSEM. In GETM-ERSEM-BFM, distinct blue patches can be seen off East Anglia, south Dorset and the German Bight, which are mostly areas where *Phaeocystis* colonies dominate in the model.

5 Summary and discussion

This study has presented a new set of in situ phytoplankton pigment observations for the North Sea, processed to give information on phytoplankton community structure. Two physical-biogeochemical models, the biogeochemical components of which are different variants of ERSEM, were then validated against these and other observations. Both models were a good match for SST observations, and a reasonable match for chlorophyll observations, but gave contrasting results for SPM. Furthermore, the two models displayed very different phytoplankton community structures. GETM-ERSEM-BFM was able to reproduce many of the features of the observations, particularly in the southern North Sea, whereas NEMO-ERSEM was a poor match for the observations, except at the lowest chlorophyll concentrations. Nonetheless, both models shared some similarities with each other and the observations in terms of spatial features and inter-annual variability.

The distribution of total phytoplankton biomass across the North Sea during summer of both years showed a high degree of consistency between three different observational methods: satellite remote sensing, high-frequency continuous measurement of chlorophyll fluorescence, and chl-a quantification at discrete stations. A similar set of spatial features can be observed in 2010 and 2011, which can be explained by the underlying hydrodynamics (Van Leeuwen et al., 2015). The central, strongly-stratified region of the North Sea has very low nutrient concentrations and correspondingly low chl-a. In areas where vertical mixing, riverine input or horizontal advection bring nutrients into the upper water column, phytoplankton biomass is elevated. As well as observing the total quantity of phytoplankton, deriving the composition of size- and functional types is important for a better understanding of ecosystem function and energy flows to higher trophic levels (Chavez et al., 2011). Accessory pigments have been widely used in biological oceanography to investigate community composition, but caution must be applied when interpreting results, and support from other methods should be used where possible (Schlüter et al., 2014). The original equations used by Hirata et al. (2008, 2011) to convert pigments to pico-, nano-, and micro-phytoplankton size classes underestimated the fraction of pico-phytoplankton compared to flow cytometric observations, and were modified by increasing the contribution of chl-b and 19-hex to this class. Results in both years showed a consistent pattern of decreasing micro-phytoplankton abundance with distance from the coast, and with increasing latitude, and this is supported by previous pigment-based studies. Work in the German Bight has also shown a change from a coastal, diatom-dominated community to a more diverse, small-celled community further offshore (Brandsma et al., 2013; Wollschläger et al., 2015).

Both NEMO-ERSEM and GETM-ERSEM-BFM have shown the ability to reproduce the physics and broad-scale biogeochemistry of the North Sea. However, results are more varied when considering specific aspects such as the phytoplankton community structure in August. In some ways this is to be expected, as this study has used existing model versions which have not been previously validated against or tuned to such observations. Furthermore, August is a challenging month to model in the North Sea, as evidenced by the reduced SST skill for this month compared with the seasonal average. Nonetheless, this kind of specific information is in increasing demand, and confidence must be had in the results if they are to be provided to users.

NEMO-ERSEM and GETM-ERSEM-BFM gave very different representations of SPM concentrations, which impacts ecosystem functioning through light limitation. NEMO-ERSEM uses the two-size class SPM model described by Sykes and Barciela (2012). This was implemented by Sykes and Barciela (2012) in the POLCOMS (Proudman Oceanographic Laboratory Coastal Ocean Modelling System; Holt and James, 2001; Holt et al., 2005) physical framework, in which it gave skilful results compared with observations. However, the model has not yet received the same degree of tuning and development since being implemented in NEMO-ERSEM, which may explain the consistent under-estimation of SPM found in this study. Furthermore, the high vertical resolution of NEMO-ERSEM means that the settling velocities must sometimes be artificially limited when used by the SPM model, in order to avoid breaking the Courant-Friedrichs-Lewy (CFL) condition (Courant et al., 1928), thus reducing resuspension. Changes to the settling parameters would be expected to lead to improvements. GETM-ERSEM-BFM uses an alternative SPM model (Van der Molen et al., 2016b), which only has one size class but includes resuspension by waves as well as currents, and which was developed within the GETM-ERSEM-BFM framework. This generally matches spatial distributions of SPM better, but often has concentrations which are extremely high or low compared with satellite data, leading to a degradation in some error statistics.

The starkest contrast between the model results presented in this study is in the simulated phytoplankton community structures, which differ far more than might be expected given the corresponding total chlorophyll concentrations. GETM-ERSEM-BFM gave a wider range of combinations of biomass in the three size classes resolved by the observations than NEMO-ERSEM did. This was also reflected in more spatial variability and stronger spatial gradients, which resulted in a better match of the coastal to offshore change in phytoplankton community structure evident in the observations. The limited biomass in the two additional PFTs in GETM-ERSEM-BFM (benthic diatoms and *Phaeocystis* colonies, Fig. 8) suggests that these were not the primary cause of this difference in response. It cannot be ruled out that the different physical models, including the different SPM methods, and differences in the model configurations, play a role (Table 1). However, it is most likely that this difference is caused in part by the greater difference in parameter settings in GETM-ERSEM-BFM between the four PFTs that the two models have in common, in combination with additional sensitivity introduced in GETM-ERSEM-BFM by the different methods of calculating nutrient affinity, nutrient stress, sinking, lysis and light susceptibility (Table 2). The more uniform parameter settings of NEMO-ERSEM promote a more uniform response of the PFTs, as is evident in the results.

GETM-ERSEM-BFM provided a better match for the IBTS observations of phytoplankton community structure than NEMO-ERSEM did. The exception to this was in the low-chlorophyll waters of the central North Sea, the region of the domain with the weakest currents and largest residence times. Here the community structure of the observations more closely resembled that typical of NEMO-ERSEM. There are indications that recent versions of NEMO-ERSEM, applied to the global ocean rather than to the North-West European Shelf Seas, perform better at reproducing observed community structures (De Mora et al., 2016). Together with the results presented here, this suggests that NEMO-ERSEM may be more representative of an open ocean environment, whereas the settings in GETM-ERSEM-BFM are better suited to the complex coastal environment of the North Sea.

Whilst there were large contrasts in the corresponding ratios of PSCs between all three data sets, there was more agreement between the data sets about the spatial patterns of community structure. For instance, each had contrasting structures between the southern and central North Sea, and in coastal areas. Furthermore, inter-annual variability in the central North Sea was clearly evident in the observations, and also each of the models. This can be compared to differences in SST between the two years and the two models, suggesting a set of physical drivers which the models were able to capture. This implies that even if models are currently unable to accurately represent the exact community structures, they can still be used to assess the distribution of different habitats, and when and where these may change.

Careful thought needs to be given therefore to what products and information can be offered to users, which address user and policy requirements with a sufficient level of skill. There are many challenges related to the uptake and use of models (Hyder et al., 2015), and studies have looked at the potential for use of models for MSFD (Piroddi et al., 2015), and the issues of uncertainty in climate projections (Payne et al., 2015). The main drivers for decision makers relate to sustainable exploitation of resources whilst protecting the environment, so it is important to consider how best biogeochemical models can be used to support and underpin the evidence upon which decisions are based (Hyder et al., 2015). Biogeochemical models can provide insight into many of the key challenges for sustainable use and protection of our seas including eutrophication (Lenhart et al., 2010), harmful algal blooms (Shutler et al., 2015), non-native species, aquaculture (Aldridge et al., 2012), climate change and ocean acidification (Van Leeuwen et al., 2016), and marine spatial planning (Van der Molen et al., 2014). It is not possible to identify here exactly how biogeochemical models may be used in the future to support decision making, but it is clear that requirements will include insights into the attribution of changes in ecosystems to underlying drivers of environmental change, improving the efficiency of monitoring programmes, assessment of indicators, and economic impacts of management options

(Hyder et al., 2015). Decision making timescales are often at odds with model development, so it is important to be able to adapt existing models to address decision makers' needs at short notice. To achieve this, close communication is needed between modellers and decision makers, so that the potential of biogeochemical models to support decision making is not lost in translation (Hyder et al., 2015).

Continual model developments will be required to address these and other user needs. A comprehensive review of the challenges faced is given by Holt et al. (2014), but specific issues have emerged from this current study which should inform future model development. This will be particularly relevant in the context of the UK Shelf Seas Biogeochemistry (SSB) research programme (http://www.uk-ssb.org), in which Cefas and the Met Office are both participants. One of the aims of the SSB programme is to create a common version of ERSEM to be used by the UK research community, by combining features of the two versions of ERSEM used in this study. An initial combined version is described by Butenschön et al. (2015), and this will be further developed within the SSB programme. It is clear from this study that the details of the model components and parameterisations can lead to very different results, and validation against a range of data, using a range of methods, is vital throughout the model development cycle.

A further development will be the assimilation of biogeochemical data. Ocean colour data assimilation is being increasingly utilised by the reanalysis and forecasting community (Gehlen et al., 2015), and has already been successfully demonstrated for ERSEM (Ciavatta et al., 2011, 2014). A suitable ocean colour assimilation scheme for operational purposes is being developed as a collaboration between the Met Office and PML, to be implemented in the SSB ERSEM version and run operationally as part of CMEMS. This will also give the opportunity to take advantage of the advent of remote sensing PFT/PSC products, incorporating such data into the assimilation and routine validation.

Information on the marine environment can come from three sources: in situ observations, remote sensing data, and models. These three sources are inter-linked and all are vital - sufficient scientific understanding of the North Sea and other environments cannot be obtained if any of these three sources are removed. Models provide complete 3D spatial and temporal coverage, can be used to simulate a range of hypotheses, and are relatively inexpensive. However, as demonstrated in this study, observations are necessary for the validation and development of models, and model data cannot be relied upon in isolation. Remote sensing data provide considerably greater observational coverage than in situ measurements, but this coverage is still limited to the sea surface and cloud-free conditions, and empirical algorithms based on in situ data are used in the construction of remote sensing products. These satellite data must be comprehensively ground-truthed against in situ observations for confidence to be had in them. Continuing in situ observations are therefore required to under-pin model and remote sensing data, as well as to provide unique insights into the marine environment. In turn, modelling studies can be used to help inform sampling strategies of future observing programmes, to help provide value for money without sacrificing accurate scientific understanding.

Finally, it remains to answer the question posed in the title of this paper: "Can ERSEM-type models simulate biodiversity?" The evidence from this study suggests that ERSEM-type models definitely have the potential to accurately simulate biodiversity, but certain model formulations and parameterisations are required to do so, and current ERSEM versions do not reliably do so at this stage. Appropriate model development, informed by detailed validation studies, will be a major but achievable challenge, and will help facilitate the wider application of marine biogeochemical modelling to wide-ranging user and policy needs.

Acknowledgements. Sampling of the North Sea was funded by the European Union project PROTOOL (EU FP7, Grant No. 226880). The authors wish to thank Brian Harley and Sophie McCully and the officers and crew of RV *Cefas Endeavour* for their assistance during the IBTS surveys. David Ford and Rosa Barciela were funded through the Met Office Innovation Fund. David Ford, Rosa Barciela and Robert McEwan also received funding from the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 283367 (MyOcean2), and from the Copernicus Marine Environment Monitoring Service. This paper is a contribution to the NERC-Defra funded Marine Ecosystems Research Programme (NERC award NE/L002981/1). Johan van der Molen and Kieran Hyder were funded through Cefas Seedcorn project DP235.

References

Aiken, J., Pradhan, Y., Barlow, R., Lavender, S., Poulton, A., Holligan, P., and Hardman-Mountford, N.: Phytoplankton pigments and functional types in the Atlantic Ocean: a decadal assessment, Deep-Sea Res. Pt II, 56(15), 899-917, 2009.

Aldridge, J., van der Molen, J., and Forster, R.: Wider ecological implications of Macroalgae culitvation, The Crown Estate, 95 pp, ISBN: 978-1-906410-38-4, 2012.

Allen, J. I. and Clarke, K. R.: Effects of demersal trawling on ecosystem functioning in the North Sea: a modelling study, Mar. Ecol.-Prog. Ser., 336, 63-75, 2007.

Anderson, T. R.: Plankton functional type modelling: running before we can walk?, J. Plankton Res., 27(11), 1073-1081, 2005.

Balmaseda M. A., Mogensen, K., and Weaver, A.: Evaluation of the ECMWF Ocean Reanalysis ORAS4, Q. J. Roy. Meteor. Soc., 139(674), 1132-1161, 2013.

Baretta, J. W., Ebenhöh, W., and Ruardij, P.: The European Regional Seas Ecosystem Model, a complex marine ecosystem model, Neth. J. Sea Res., 33(3), 233-246, 1995.

Blackford, J. C., Allen, J. I., and Gilbert, F. J.: Ecosystem dynamics at six contrasting sites: a generic modelling study, J. Marine Syst., 52(1), 191-215, 2004.

Blauw, A. N., Los, F. J., Huisman, J.,, and Peperzak, L.: Nuisance foam events and *Phaeocystis* globosa blooms in Dutch coastal waters analyzed with fuzzy logic, J. Marine Syst., 83, 115-126, 2010.

Blockley, E. W., Martin, M. J., McLaren, A. J., Ryan, A. G., Waters, J., Lea, D. J., Mirouze, I., Peterson, K. A., Sellar, A. and Storkey, D.: Recent development of the Met Office operational ocean forecasting system: an overview and assessment of the new Global FOAM forecasts, Geoscientific Model Development, 7(6), 2613-2638, 2014.

Borja, A., Elliott, M., Andersen, J. H., Cardoso, A. C., Carstensen, J., Ferreira, J. G., Heiskanen, A. S., Marques, J. C., Neto, J. M., Teixeira, H., and Uusitalo, L.: Good Environmental Status of marine ecosystems: What is it and how do we know when we have attained it?, Mar. Pollut. Bull., 76(1), 16-27, 2013.

Boyes, S. J. and Elliott, M.: Marine legislation - The ultimate "horrendogram": International law, European directives and national implementation, Mar. Pollut. Bull., 86(1), 39-47, 2014.

Brandsma, J., Martínez, J.M., Slagter, H.A., Evans, C., and Brussaard, C.P.: Microbial biogeography of the North Sea during summer, Biogeochemistry, 113, 119-136, 2013.

Brewin, R. J., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R. M. and Hardman-Mountford, N. J.: A three-component model of phytoplankton size class for the Atlantic Ocean, Ecol. Model., 221(11), 1472-1483, 2010.

Brewin, R. J., Hardman-Mountford, N. J., Lavender, S. J., Raitsos, D. E., Hirata, T., Uitz, J., Devred, E., Bricaud, A., Ciotti, A., and Gentili, B.: An intercomparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing, Remote Sens. Environ., 115(2), 325-339, 2011.

Brito, A. C., Sá, C., Brotas, V., Brewin, R. J., Silva, T., Vitorino, J., Platt, T., and Sathyendranath, S.: Effect of phytoplankton size classes on bio-optical properties of phytoplankton in the Western Iberian coast: Application of models, Remote Sens. Environ., 156, 537-550, 2014.

Brotas, V., Brewin, R. J., Sá, C., Brito, A. C., Silva, A., Mendes, C. R., Diniz, T., Kaufmann, M., Tarran, G., Groom, S. B., and Platt, T.: Deriving phytoplankton size classes from satellite data: Validation along a trophic gradient in the eastern Atlantic Ocean, Remote Sens. Environ., 134, 66-77, 2013.

Bruggeman, J.: Succession in plankton communities: A trait-based perspective, Ph.D. thesis, Vrije Universiteit Amsterdam, 2009.

Burchard, H. and Bolding, K.: GETM: A General Estuarine Transport Model; Scientific Documentation, Tech. Rep. EUR 20253 EN, European Commission, 2002.

Butenschön, M., Clark, J., Aldridge, J. N., Allen, J. I., Artioli, Y., Blackford, J., Bruggeman, J., Cazenave, P., Ciavatta, S., Kay, S., Lessin, G., van Leeuwen, S., van der Molen, J., de Mora, L., Polimene, L., Sailley, S., Stephens, N., and Torres, R.: ERSEM 15.06: a generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels, Geoscientific Model Development, 8, 7063-7187, 2015.

Campbell, J. W.: The lognormal distribution as a model for bio-optical variability in the sea, J. Geophys. Res.-Oceans, 100(C7), 13237-13254, 1995.

Capuzzo, E., Stephens, D., Silva, T., Barry, J., and Forster, R. M.: Decrease in water clarity of the southern and central North Sea during the 20th century, Global Change Biol., 21(6), 2206-2214,, 2015.

Chassot, E., Mélin, F., Le Pape, O., and Gascuel, D.: Bottom-up control regulates fisheries production at the scale of eco-regions in European seas, Mar. Ecol.-Prog. Ser., 343, 45-55, 2007.

Chavez, F. P., Messié, M., and Pennington, J. T.: Marine Primary Production in Relation to Climate Variability and Change, Annu. Rev. Mar. Sci., 3, 227-260, 2011.

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

Ciavatta, S., Torres, R., Martinez-Vicente, V., Smyth, T., Dall'Olmo, G., Polimene, L., and Allen, J. I.: Assimilation of remotely-sensed optical properties to improve marine biogeochemistry modelling, Prog. Oceanogr., 127, 74-95, 2014.

Courant, R., Friedrichs, K., and Lewy, H.: Über die partiellen Differenzengleichungen der mathematischen Physik, Math. Ann., 100(1), 32-74, 1928.

Daan, N., Gislason, H., Pope, J. G., and Rice, J. C.: Changes in the North Sea fish community: evidence of indirect effects of fishing?, ICES J. Mar. Sci., 62(2), 177-188, 2005.

De Mora, L., Butenschön, M., and Allen, J. I.: How should sparse marine in situ measurements be compared to a continuous model: an example, Geoscientific Model Development, 6(2), 533-548, 2013.

De Mora, L., Butenschön, M., and Allen, J. I.: The assessment of a global marine ecosystem model on the basis of emergent properties and ecosystem function: a case study with ERSEM, Geoscientific Model Development, 9(1), 59-76, 2016.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553-597, 2011.

Defra: Making the most of our evidence?: a strategy for Defra and its network, Department for Environment, Food and Rural Affairs, 2014.

Devred, E., Sathyendranath, S., Stuart, V., and Platt, T.: A three component classification of phytoplankton absorption spectra: Application to ocean-color data, Remote Sens. Environ., 115(9), 2255-2266, 2011.

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., and Wimmer, W: The operational sea surface temperature and sea ice analysis (OSTIA) system, Remote Sens. Environ., 116, 140-158, 2012.

Donnelly, C., Andersson, J. C., and Arheimer, B.: Using flow signatures and catchment similarities to evaluate the E-HYPE multi-basin model across Europe, Hydrolog. Sci. J., 61(2), 255-273, 2015.

Dupont, N. and Aksnes, D. L.: Centennial changes in water clarity of the Baltic Sea and the North Sea, Estuar. Coast. Shelf S., 131, 282-289, 2013.

ECMWF: European Centre for Medium-Range Weather Forecasts (ECMWF) 40-year Re-Analysis (ERA-40) model data, NCAS British Atmospheric Data Centre, http://catalogue.ceda.ac.uk/uuid/775634f7e339b5262067e28a5d7b679d, 2006a.

ECMWF: Assimilated Data from the European Centre for Medium-Range Weather Forecasts (ECMWF) operational analysis program, NCAS British Atmospheric Data Centre, http://catalogue.ceda.ac.uk/uuidf6ce34fc/c462480467660a36d9b10a71, 2006b.

Edwards, K. P., Barciela, R., and Butenschön, M.: Validation of the NEMO-ERSEM operational ecosystem model for the North West European Continental Shelf, Ocean Sci., 8, 983-1000, 2012.

Engel, A.: The role of transparent exopolymer particles (TEP) in the increase in apparent particle stickiness (α) during the decline of a diatom bloom, J. Plankton Res., 22, 485-497, 2000.

EU: Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy, European Parliament and Council of the European Union, Off. J. Eur. Communities, 327, 1-72, 2000.

EU: Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive), European Parliament and Council of the European Union, Off. J. Eur. Union, 164, 19-40, 2008.

EU: Regulation (EU) No 1380/2013 of the European Parliament and of the Council of 11 December 2013 on the Common Fisheries Policy, amending council Regulations (EC) No 1954/2003 and (EC) No 1224/2009 and repealing Council regulations (EC) No 2371/2002 and (EC) No 639/2004 and Council Decision 2004/585/EC, Off. J. Eur. Union, 354, 22-61, 2013.

Finkel, Z. V., Beardall, J., Flynn, K. J., Quigg, A., Rees, T. A. V., and Raven, J.A.: Phytoplankton in a changing world: cell size and elemental stoichiometry, J. Plankton Res., 32(1), 119-137, 2010.

Friedrichs, M. A., Dusenberry, J. A., Anderson, L. A., Armstrong, R. A., Chai, F., Christian, J. R., Doney, S. C., Dunne, J., Fujii, M., Hood, R., and McGillicuddy, D. J.: Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups, J. Geophys. Res.-Oceans, 112(C8), 2007.

Fulton, E. A. and Link, J. S.: Modeling Approaches for Marine Ecosystem-based Management, in: The Sea, Volume 16: Marine Ecosystem-Based Management, editors: Fogarty, M. J. and McCarthy, J. J., Harvard University Press, 2014.

Fulton, E. A., Link, J. S., Kaplan, I. C., SavinaRolland, M., Johnson, P., Ainsworth, C., Horne, P., Gorton, R., Gamble, R. J., Smith, A. D., and Smith, D. C.: Lessons in modelling and management of marine ecosystems: the Atlantis experience, Fish Fish., 12(2), 171-188, 2007.

Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Zweng, M. M., Baranova, O. K., and Johnson, D. R.: World Ocean Atlas 2009, Volume 4: Nutrients (phosphate, nitrate, silicate), edited by Levitus, S., NOAA Atlas NESDIS 71, U.S. Government Printing Office, Washington, D.C., 398 pp., 2010.

Gårdmark, A., Lindegren, M., Neuenfeldt, S., Blenckner, T., Heikinheimo, O., Müller-Karulis, B., Niiranen, S., Tomczak, M. T., Aro, E., Wikström, A., and Möllmann, C.: Biological ensemble modeling to evaluate potential futures of living marine resources, Ecol. Appl., 23(4), 742-754, 2013.

Garmendia, M., Borja, A., Franco, J., and Revilla, M.: Phytoplankton composition indicators for the assessment of eutrophication in marine waters: Present state and challenges within the European directives, Mar. Pollut. Bull., 66, 7-16, 2013.

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

Gohin, F., Druon, J. N., and Lampert, L.: A five channel chlorophyll concentration algorithm applied to SeaWiFS data processed by SeaDAS in coastal waters, Int. J. Remote S., 23(8), 1639-1661, 2002.

Gohin, F., Loyer, S., Lunven, M., Labry, C., Froidefond, J. M., Delmas, D., Huret, M., and Herbland, A.: Satellite-derived parameters for biological modelling in coastal waters: Illustration over the eastern continental shelf of the Bay of Biscay, Remote Sens. Environ., 95(1), 29-46, 2005.

Gohin, F., Saulquin, B., Oger-Jeanneret, H., Lozac'h, L., Lampert, L., Lefebvre, A., Riou, P., and Bruchon, F.: Towards a better assessment of the ecological status of coastal waters using satellite-derived chlorophyll-a concentrations, Remote Sens. Environ., 112(8), 3329-3340, 2008.

Greenwood, N., Parker, E. R., Fernand, L., Sivyer, D. B., Weston, K., Painting, S. J., Kröger, S., Forster, R. M., Lees, H. E., Mills, D. K., and Laane, R. W. P. M.: Detection of low bottom water oxygen concentrations in the North Sea; implications for monitoring and assessment of ecosystem health, Biogeosciences, 7(4), 1357-1373, 2010.

Hirata, T., Aiken, J., Hardman-Mountford, N., Smyth, T. J., and Barlow, R.G.: An absorption model to determine phytoplankton size classes from satellite ocean colour, Remote Sens. Environ., 112(6), 3153-3159, 2008.

Hirata, T., Hardman-Mountford, N. J., Brewin, R. J. W., Aiken, J., Barlow, R., Suzuki, K., Isada, T., Howell, E., Hashioka, T., Noguchi-Aita, M. and Yamanaka, Y.: Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types, Biogeosciences, 8(2), 311-327, 2011.

Holt, J. T. and James, I. D.: An s coordinate density evolving model of the northwest European continental shelf: 1. Model description and density structure, J. Geophys. Res.-Oceans, 106(C7), 14015-14034, 2001.

Holt, J. T., Allen, J. I., Proctor, R., and Gilbert, F.: Error quantification of a high-resolution coupled hydrodynamic-ecosystem coastal-ocean model: Part 1. Model overview and assessment of the hydrodynamics, J. Marine Syst., 57(1), 167-188, 2005.

Holt, J., Allen, J. I., Anderson, T. R., Brewin, R., Butenschön, M., Harle, J., Huse, G., Lehodey, P., Lindemann, C., Memery, L. and Salihoglu, B.: Challenges in integrative approaches to modelling the marine ecosystems of the North Atlantic: Physics to fish and coasts to ocean, Prog. Oceanogr., 129, 285-313, 2014.

Hyder, K., Rossberg, A. G., Allen, J. I., Austen, M. C., Barciela, R. M., Bannister, H. J., Blackwell, P. G., Blanchard, J. L., Burrows, M. T., Defriez, E., and Dorrington, T.: Making modelling count-increasing the contribution of shelf-seas community and ecosystem models to policy development and management, Mar. Policy, 61, 291-302, 2015.

Jennings, S. and Collingridge, K.: Predicting consumer biomass, size-structure, production, catch potential, responses to fishing and associated uncertainties in the world's marine ecosystems., PloS one, 10(7), e0133794, 2015.

Jennings, S., Greenstreet, S., Hill, L., Piet, G., Pinnegar, J., and Warr, K.J.: Long-term trends in the trophic structure of the North Sea fish community: evidence from stable-isotope analysis, size-spectra and community metrics, Mar. Biol., 141(6), 1085-1097, 2002.

Jupp, T. E., Lowe, R., Coelho, C. A., and Stephenson, D. B.: On the visualization, verification and recalibration of ternary probabilistic forecasts, Philos. T. Roy. Soc. A, 370(1962), 1100-1120, 2012.

Kurekin, A.A., Miller, P.I., and Van der Woerd, H.J.: Satellite discrimination of Karenia mikimotoi and *Phaeocystis* harmful algal blooms in European coastal waters: Merged classification of ocean colour data, Harmful Algae, 31, 163-176, 2014.

Kwiatkowski, L., Yool, A., Allen, J. I., Anderson, T. R., Barciela, R., Buitenhuis, E. T., Butenschön, M., Enright, C., Halloran, P. R., Le Quéré, C., and De Mora, L.: iMarNet: an ocean biogeochemistry model intercomparison project within a common physical ocean modelling framework, Biogeosciences, 11(24), 7291-7304, 2014.

Le Quéré, C., Harrison, S. P., Prentice, I. C., Buitenhuis, E. T., Aumont, O., Bopp, L., Claustre, H., Cotrim Da Cunha, L., Geider, R., Giraud, X., and Klaas, C.: Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models, Global Change Biol., 11(11), 2016-2040, 2005.

Lenhart, H. J., Mills, D. K., Baretta-Bekker, H., Van Leeuwen, S. M., Van der Molen, J., Baretta, J. W., Blaas, M., Desmit, X., Kühn, W., Lacroix, G., Los, H. J., Ménesguen, A., Neves, R., Proctor, R., Ruardij, P., Skogen, M. D., Vanhoutte-Grunier, A., Villars, M. T., and Wakelin, S. L.: Predicting the consequences of nutrient reduction on the eutrophication status of the North Sea, J. Marine Syst., 81, 148-170, 2010.

Mackinson, S.: Combined analyses reveal environmentally driven changes in the North Sea ecosystem and raise questions regarding what makes an ecosystem model's performance credible? , Can. J. Fish. Aquat. Sci., 71(1), 31-46, 2014.

MacLachlan, C., Arribas, A., Peterson, D., Maidens, A., Fereday,D., Scaife, A. A., Gordon, M., Vellinga, M., Williams, A., Comer, R. E., Camp, J., and Xavier, P.: Global Seasonal Forecast System 5 (GloSea5): a high resolution seasonal forecast system, Q. J. Roy. Meteor. Soc., 141(689), 1072-1084, 2014.

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

Mogensen, K., Balmaseda, M., and Weaver, A.: The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System4, ECMWF Technical Memorandum 668, 59 pp., 2012.

Núñez-Riboni, I. and Akimova, A.: Monthly maps of optimally interpolated in situ hydrography in the North Sea from 1948 to 2013, J. Marine Syst., 151, 15-34, 2015.

O'Dea, E. J., Arnold, A .K., Edwards, K. P., Furner, R., Hyder, P., Martin, M. J., Siddorn, J. R., Storkey, D., While, J., Holt, J. T. and Liu, H.: An operational ocean forecast system incorporating NEMO and SST data assimilation for the tidally driven European North-West shelf, Journal of Operational Oceanography, 5(1), 3-17, 2012.

O'Dea, E. J., et al.: in preparation.

Payne, M. R., Barange, M., Cheung, W. W., MacKenzie, B. R., Batchelder, H. P., Cormon, X., Eddy, T. D., Fernandes, J. A., Hollowed, A. B., Jones, M. C., and Link, J. S.: Uncertainties in projecting climate-change impacts in marine ecosystems, ICES J. Mar. Sci., fsv231, 2015.

Peloquin, J., Swan, C., Gruber, N., Vogt, M., Claustre, H., Ras, J., Uitz, J., Barlow, R., Behrenfeld, M., Bidigare, R. and Dierssen, H.: The MAREDAT global database of high performance liquid chromatography marine pigment measurements, Earth System Science Data, 5(1), 109-123, 2013.

Peperzak, L., Colijn, F., Gieskes, W. W. C., and Peeters, J. C. H.: Development of the diatom-*Phaeocystis* spring bloom in the Dutch coastal zone of the North Sea: the silicon depletion versus the daily irradiance threshold hypothesis, J. Plankton Res., 20, 517-537, 1998.

Petersen, W., Wehde, H., Krasemann, H., Colijn, F., and Schroeder, F.: FerryBox and MERIS - Assessment of coastal and shelf sea ecosystems by combining in situ and remotely sensed data, Estuar. Coast. Shelf. S., 77(2), 296-307, 2008.

Piroddi, C., Teixeira, H., Lynam, C. P., Smith, C., Alvarez, M. C., Mazik, K., Andonegi, E., Churilova, T., Tedesco, L., Chifflet, M., and Chust, G.: Using ecological models to assess ecosystem status in support of the European Marine Strategy Framework Directive, Ecol. Indic., 58, 175-191, 2015.

Plagányi, É. E., Punt, A. E., Hillary, R., Morello, E. B., Thébaud, O., Hutton, T., Pillans, R. D., Thorson, J. T., Fulton, E. A., Smith, A. D. and Smith, F.: Multispecies fisheries management and conservation: tactical applications using models of intermediate complexity, Fish Fish., 15(1), 1-22, 2014.

Rombouts, I., Beaugrand, G., Fizzala, X., Gaill, F., Greenstreet, S. P. R., Lamare, S., Le Loc'h, F., McQuatters-Gollop, A., Mialet, B., Niquil, N., and Percelay, J.: Food web indicators under the Marine Strategy Framework Directive: From complexity to simplicity?, Ecol. Indic., 29, 246-254, 2013.

Roselli, L. and Basset, A.: Decoding size distribution patterns in marine and transitional water phytoplankton: from community to species level, PloS one, 10(5), e0127193, 2015.

Ruardij, P. and Van Raaphorst, W.: Benthic nutrient regeneration in the ERSEM-BFM ecosystem model of the North Sea, Neth. J. Sea Res., 33, 453-483, 1995.

Ruardij, P., van Haren, H., and Ridderinkhof, H.: The impact of thermal stratification on phytoplankton and nutrient dynamics in shelf seas, J. Sea Res., 38, 311-331, 1997.

Ruardij, P., Veldhuis, M. J. W., and Brussaard, C. P. D.: Modeling the bloom dynamics of the polymorphic phytoplankter *Phaeocystis* globosa: impact of grazers and viruses, Harmful Algae, 4, 941-963, 2005.

Ryther, J. H.: Photosynthesis and fish production in the sea. The production of organic matter and its conversion to higher forms of life vary throughout the world ocean, Science, 166, 72-76, 1969.

Schlüter, L., Henriksen, P., Nielsen, T. G., and Jakobsen, H. H.: Phytoplankton composition and biomass across the southern Indian Ocean, Deep-Sea Res. Pt I, 58(5), 546-556, 2011.

Schlüter, L., Møhlenberg, F., and Kaas, H.: Temporal and spatial variability of phytoplankton monitored by a combination of monitoring buoys, pigment analysis and fast screening microscopy in the Fehmarn Belt Estuary, Environ. Monit. Assess., 186(8), 5167-5184, 2014.

Shephard, S., Greenstreet, S. P., Piet, G. J., Rindorf, A. and Dickey-Collas, M.: Surveillance indicators and their use in implementation of the Marine Strategy Framework Directive, ICES J. Mar. Sci., fsv131, 2015.

Sherrard, N. J., Nimmo, M., and Llewellyn, C. A.: Combining HPLC pigment markers and ecological similarity indices to assess phytoplankton community structure: an environmental tool for eutrophication?, Sci. Total Environ., 361(1), 97-110, 2006.

Shutler, J. D., Warren, M. A., Miller, P. I., Barciela, R., Mahdon, R., Land, P. E., Edwards, K., Wither, A., Jonas, P., Murdoch, N., and Roast, S. D.: Operational monitoring and forecasting of bathing water quality through exploiting satellite Earth observation and models: The AlgaRisk demonstration service, Computers & Geosciences, 77, 87-96, 2015.

Siddorn, J. R., and Furner, R.: An analytical stretching function that combines the best attributes of geopotential and terrain-following vertical coordinates, Ocean Model., 66, 1-13, 2013.

Skogen, M. D., Eilola, K., Hansen, J. L. S., Meier, H. E. M., Molchanov, M. S., Ryabchenko, V. A.: Eutrophication status of the North Sea, Skagerrak, Kattegat and the Baltic Sea in present and future climates: A model study, J. Marine Syst., 132, 174-184, 2014.

Smaal, A. C., Schellekens, T., van Stralen, M. R., and Kromkamp, J. C.: Decrease of the carrying capacity of the Oosterschelde estuary (SW Delta, NL) for bivalve filter feeders due to overgrazing?, Aquaculture, 404, 28-34, 2013.

Stewart, I. J. and Martell, S. J.: Reconciling stock assessment paradigms to better inform fisheries management, ICES J. Mar. Sci., fsv061, 2015.

Stips, A., Bolding, K., Pohlmann, T., and Burchard, H.: Simulating the temporal and spatial dynamics of the North Sea using the new model GETM (General Estuarine Transport Model), Ocean Dynam., 54(2), 266-283, 2004.

Storkey, D., Blockley, E. W., Furner, R., Guiavarc'h, C., Lea, D., Martin, M. J., Barciela, R. M., Hines, A., Hyder, P. and Siddorn, J. R.: Forecasting the ocean state using NEMO: The new FOAM system, Journal of Operational Oceanography, 3(1), 3-15, 2010.

Sutherland, W. J., Armstrong-Brown, S., Armsworth, P. R., Tom, B., Brickland, J., Campbell, C. D., Chamberlain, D. E., Cooke, A. I., Dulvy, N. K., Dusic, N. R., and Fitton, M.: The identification of 100 ecological questions of high policy relevance in the UK, J. Appl. Ecol., 43(4), 617-627, 2006.

Sykes, P. A. and Barciela, R. M.: Assessment and development of a sediment model within an operational system, J. Geophys. Res.-Oceans, 117(C4), 2012.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res.-Atmos., 106(D7), 7183-7192, 2001.

Tebaldi, C. and Knutti, R.: The use of the multi-model ensemble in probabilistic climate projections, Philos. T. Roy. Soc. A, 365(1857), 2053-2075, 2007.

Thorpe, R. B., Le Quesne, W. J., Luxford, F., Collie, J. S., and Jennings, S.: Evaluation and management implications of uncertainty in amultispecies size-structured model of population and community responses to fishing, Journal, 6(1), 49-58, 2015.

Thyssen, M., Alvain, S., Lefèbvre, A., Dessailly, D., Rijkeboer, M., Guiselin, N., Creach, V., and Artigas, L. F.: High-resolution analysis of a North Sea phytoplankton community structure based on in situ flow cytometry observations and potential implication for remote sensing, Biogeosciences, 12(13), 4051-4066, 2015.

Uitz, J., Claustre, H., Morel, A., and Hooker, S. B.: Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll, J. Geophys. Res.-Oceans, 111(C8), 2006.

Uitz, J., Huot, Y., Bruyant, F., Babin, M., and Claustre, H.: Relating phytoplankton photophysiological properties to community structure on large scales, Limnol. Oceanogr., 53(2), 614-630, 2008.

Van der Molen, J., Aldridge, J. N., Coughlan, C., Parker, E. R., Stephens, D., and Ruardij, P.: Modelling marine ecosystem response to climate change and trawling in the North Sea, Biogeochemistry, 113, 213-236, 2013.

Van der Molen, J., Smith, H. C. M., Lepper, P., Limpenny, S., and Rees, J.: Predicting the large-scale consequences of offshore wind turbine array development on a North Sea ecosystem, Cont. Shelf Res., 85, 60-72, 2014.

Van der Molen, J., Ruardij, P., and Greenwood, N.: Potential environmental impact of tidal energy extraction in the Pentland Firth at large spatial scales: results of a biogeochemical model, Biogeosciences, 13(8), 2593-2609, 2016a.

Van der Molen, J., Ruardij, P., and Greenwood, N.: A 3D SPM model for biogeochemical modelling, with application to the northwest European continental shelf, Submitted to J. Sea Res., 2016b.

Van Leeuwen, S. M., Van der Molen, J., Ruardij, P., Fernand, L., and Jickells, T.: Modelling the contribution of deep chlorophyll maxima to annual primary production in the North Sea, Biogeochemistry, 113(1), 137-152, 2013.

Van Leeuwen, S., Tett, P., Mills, D., and Van der Molen, J.: Stratified and non-stratified areas in the North Sea: Long-term variability and biological and policy implications, J. Geophys. Res.-Oceans, 120(7), 4670-4686, 2015.

Van Leeuwen, S. M., Le Quesne, W. F., and Parker, E.R.: Potential future fisheries yields in shelf waters: a model study of the effects of climate change and ocean acidification, Biogeosciences, 13(2), 441-454, 2016.

Veldhuis, M. J. W. and Admiraal, W.: Influence of phosphate depletion on the growth and colony formation of Phaeocystis pouchetii, Mar. Biol., 95, 47-54, 1987.

Vichi, M., Oddo, P., Zavatarelli, M., Coluccelli, A., Coppini, G., Celio, M., Umani, S. F., and Pinardi, N.: Calibration and validation of a one-dimensional complex marine biogeochemical flux model in different areas of the northern Adriatic Sea, Ann. Geophys., 21(1), 413-436, 2003.

Vichi, M., Ruardij, P., and Baretta, J.W.: Link or sink: a modelling interpretation of the open Baltic biogeochemistry, Biogeosciences, 1, 79-100, 2004.

Vichi, M., Pinardi, N., and Masina, S.: A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part I: Theory, J. Marine Syst., 64, 89-109, 2007.

Wakelin, S. L., Artioli, Y., Butenschön, M., Allen, J. I., and Holt, J. T.: Modelling the combined impacts of climate change and direct anthropogenic drivers on the ecosystem of the northwest European continental shelf, J. Marine Syst., 152, 51-63, 2015a.

Wakelin, S., While, J., King, R., O'Dea, E., Holt, J., Furner, R., Siddorn, J., Martin, M., McEwan, R., Blockley, E., and Tinker, J.: Quality Information Document: North West European Shelf Reanalysis - NORTHWESTSHELF_REANALYSIS_PHYS_004_009 and NORTHWESTSHELF_REANALYSIS_BIO_004_011, EU Copernicus Marine Environment Monitoring Service, http://marine.copernicus.eu/documents/QUID/CMEMS-NWS-QUID-004-009-011.pdf, 2015b.

Ward, B. A., Schartau, M., Oschlies, A., Martin, A. P., Follows, M. J., and Anderson, T. R.: When is a biogeochemical model too complex? Objective model reduction and selection for North Atlantic time-series sites, Prog. Oceanogr., 116, 49-65, 2013.

Waters, J., Lea, D. J., Martin, M. J., Mirouze, I., Weaver, A., and While, J.: Implementing a variational data assimilation system in an operational 1/4 degree global ocean model, Q. J. Roy. Meteor. Soc., 141(687), 333-349, 2015.

Wollschläger, J., Wiltshire, K. H., Petersen, W., and Metfies, K.: Analysis of phytoplankton distribution and community structure in the German Bight with respect to the different size classes, J. Sea Res., 99, 83-96, 2015.

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

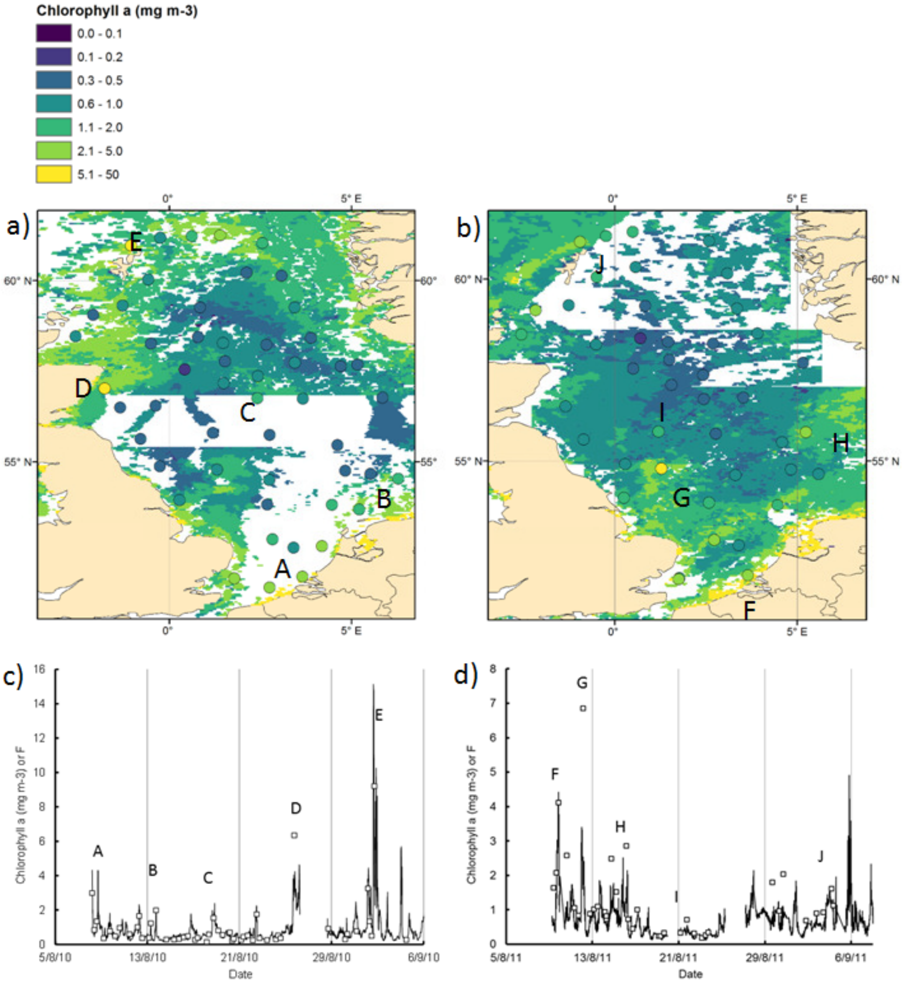


Figure 1. Maps of satellite-derived surface phytoplankton chlorophyll distribution during the summer IBTS cruises of a) 2010 and b) 2011, overlaid with the in situ observations in circles. White areas are where no satellite data were available. Time series of phytoplankton chlorophyll along the IBTS cruise track in c) 2010 and d) 2011 as assessed by continuous measurements of chlorophyll fluorescence (solid black line), and sampling of surface water for quantification of chl-a (open squares). Specific events along the tracks are referenced with a letter.

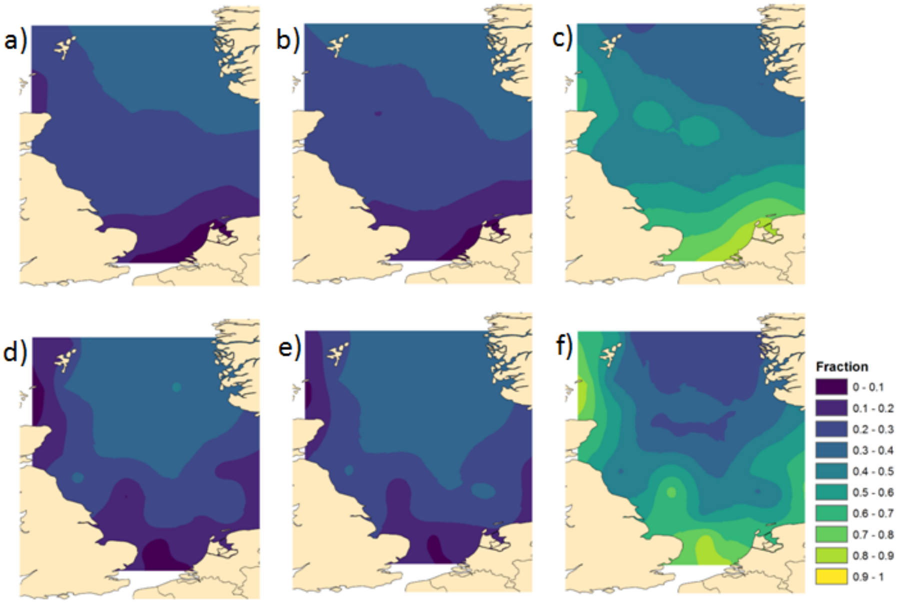


Figure 2. Maps of percentage surface PFT distribution during the summer IBTS cruises of 2010 (upper maps, a-c) and 2011 (lower maps, d-f) for pico-phytoplankton (a, d), nano-phytoplankton (b, e) and micro-phytoplankton (c, f).

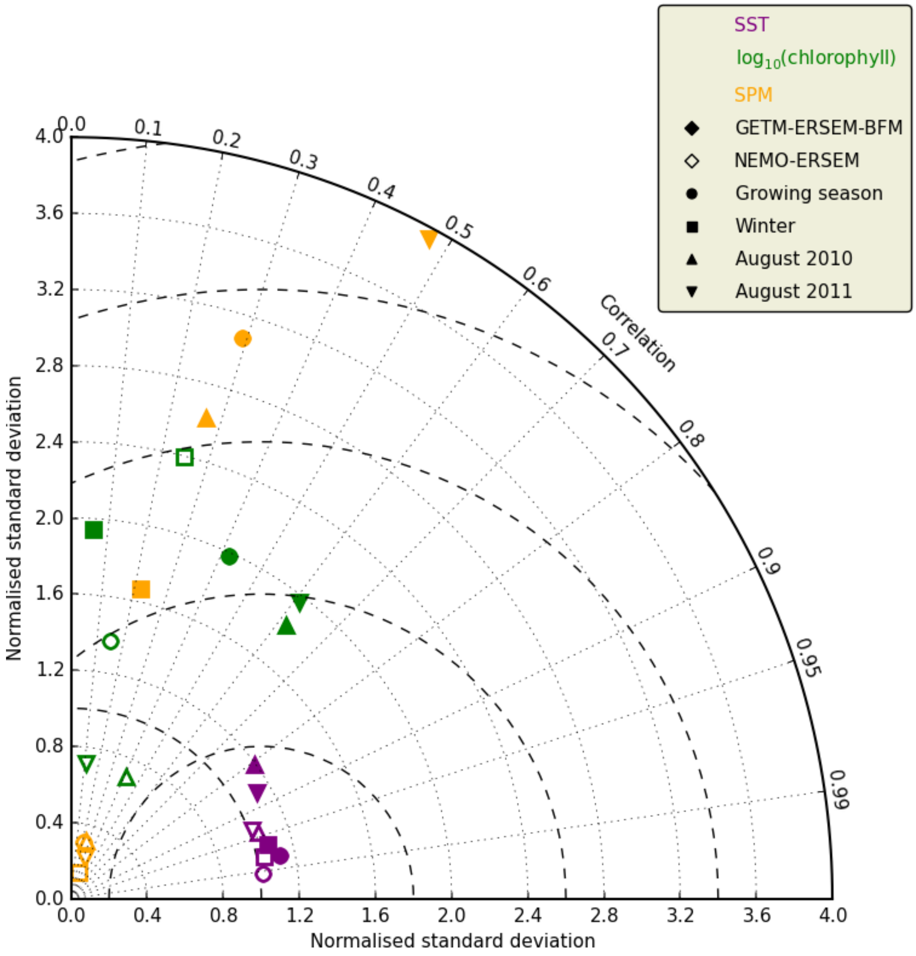


Figure 3. Taylor plot of SST (purple), $\log_{10}$(chlorophyll) (green), and SPM (orange) for the growing season (circles), winter (squares), August 2010 (upwards triangles), and August 2011 (downwards triangles). Filled symbols are GETM-ERSEM-BFM, unfilled symbols are NEMO-ERSEM.

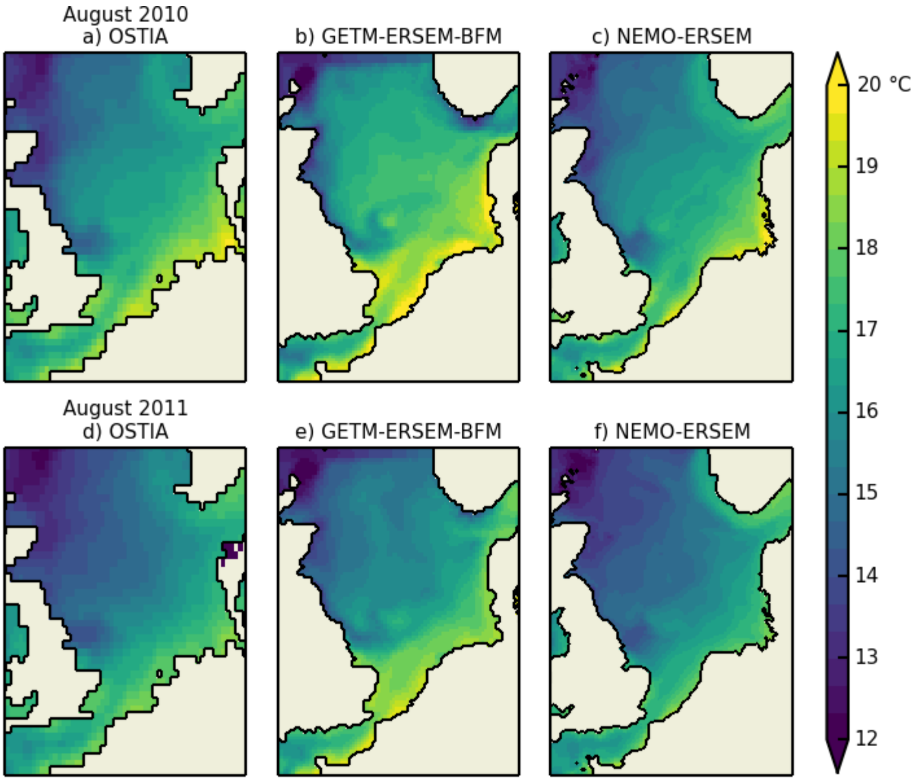


Figure 4. Monthly mean SST for August 2010 (a-c) and August 2011 (d-f): observational data (a, d), GETM-ERSEM-BFM (b, e), and NEMO-ERSEM (c, f).

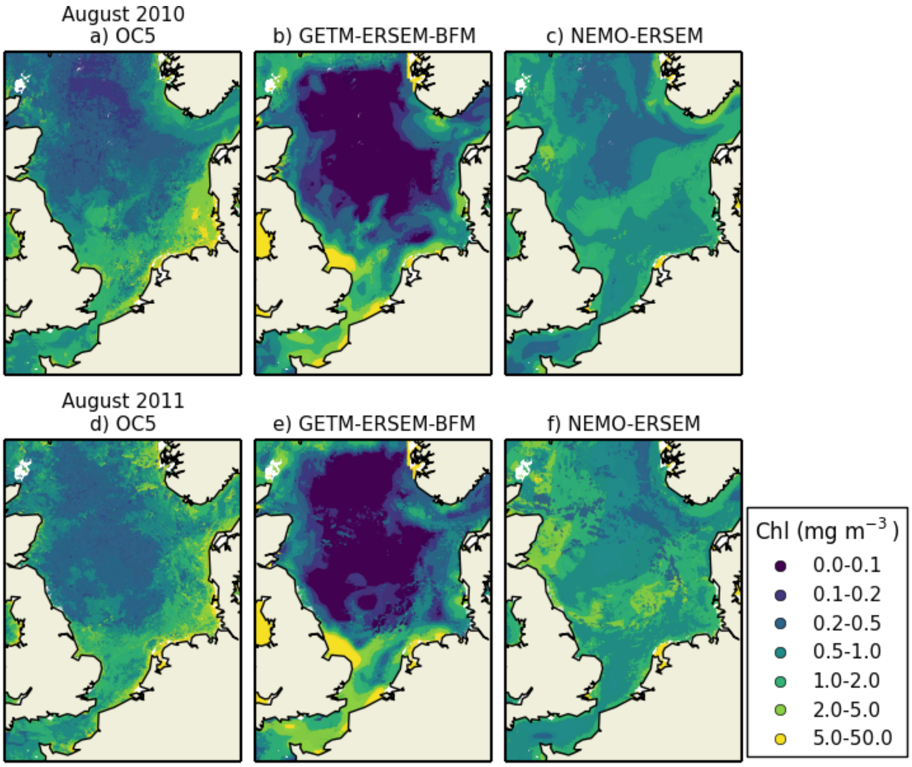


Figure 5. Composites of sea surface chlorophyll at ocean colour observation points for August 2010 (a-c) and August 2011 (d-f): satellite observations (a, d), GETM-ERSEM-BFM (b, e), and NEMO-ERSEM (c, f).

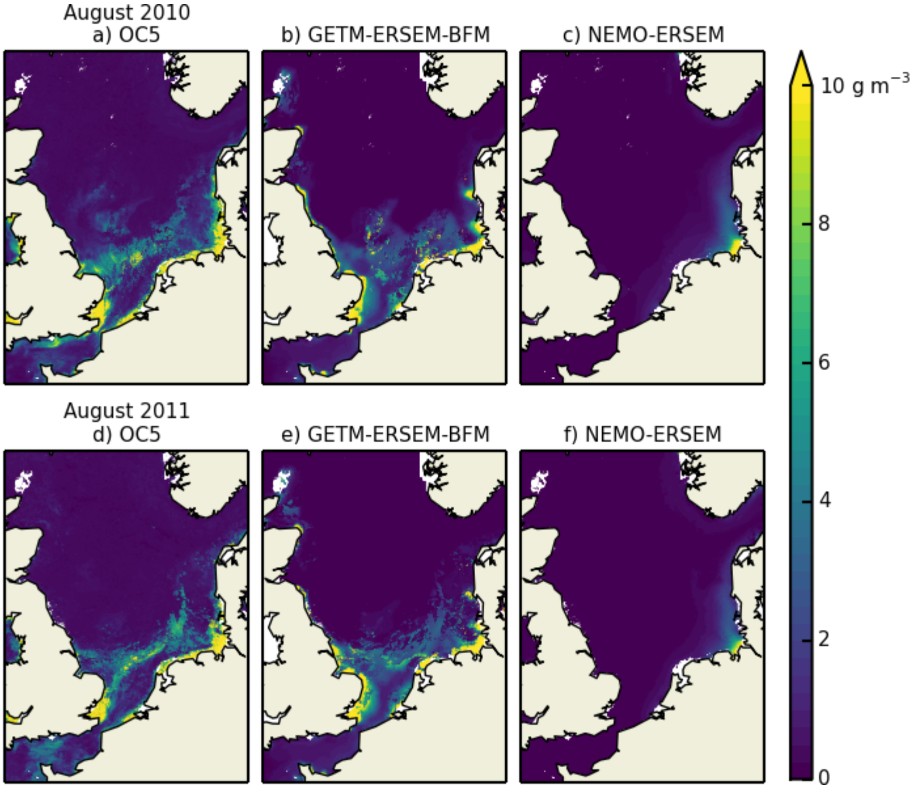


Figure 6. Composites of sea surface SPM at ocean colour observation points for August 2010 (a-c) and August 2011 (d-f): satellite observations (a, d), GETM-ERSEM-BFM (b, e), and NEMO-ERSEM (c, f).

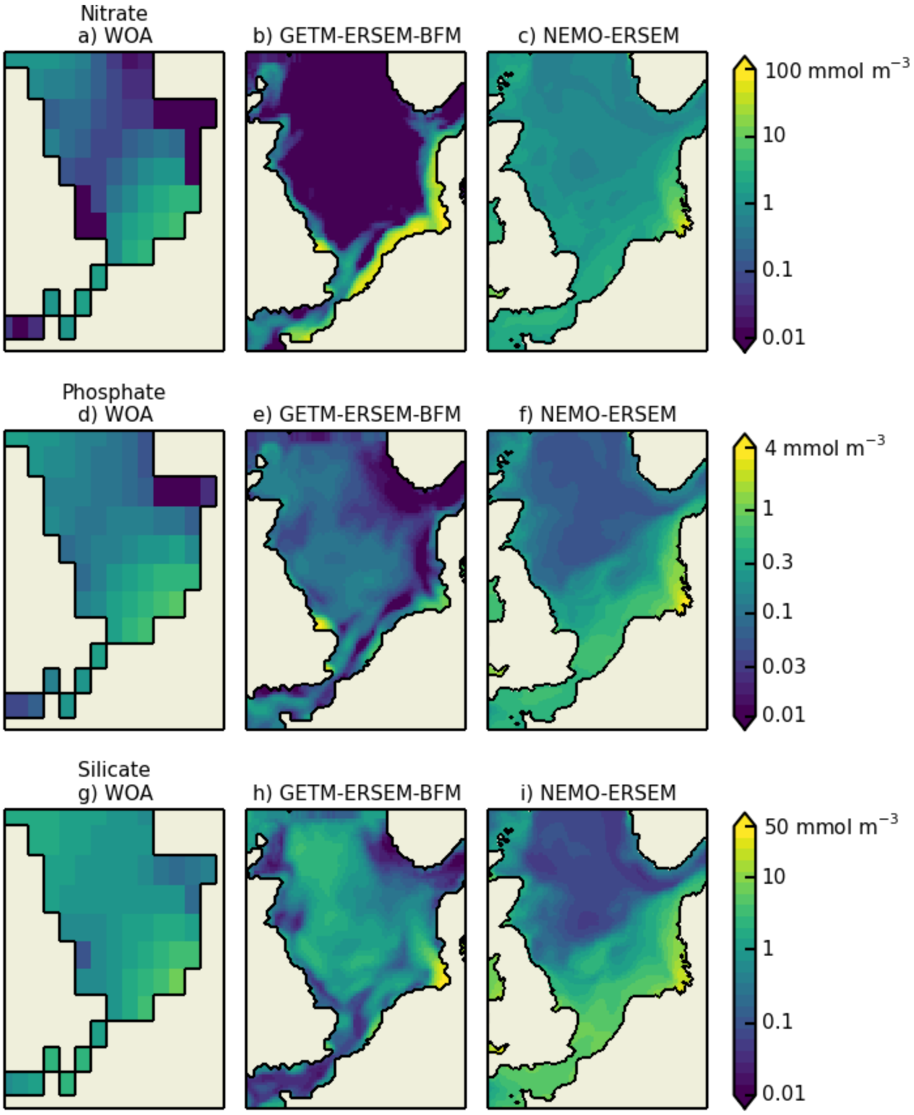


Figure 7. Monthly mean sea surface nitrate (a-c), phosphate (d-f), and silicate (g-i) for August 2010: World Ocean Atlas climatology (a, d, g), GETM-ESREM-BFM (b, e, h), and NEMO-ERSEM (c, f, i).

Biogeosciences Discuss., doi:10.5194/bg-2016-304, 2016
Manuscript under review for journal Biogeosciences
Published: 25 July 2016

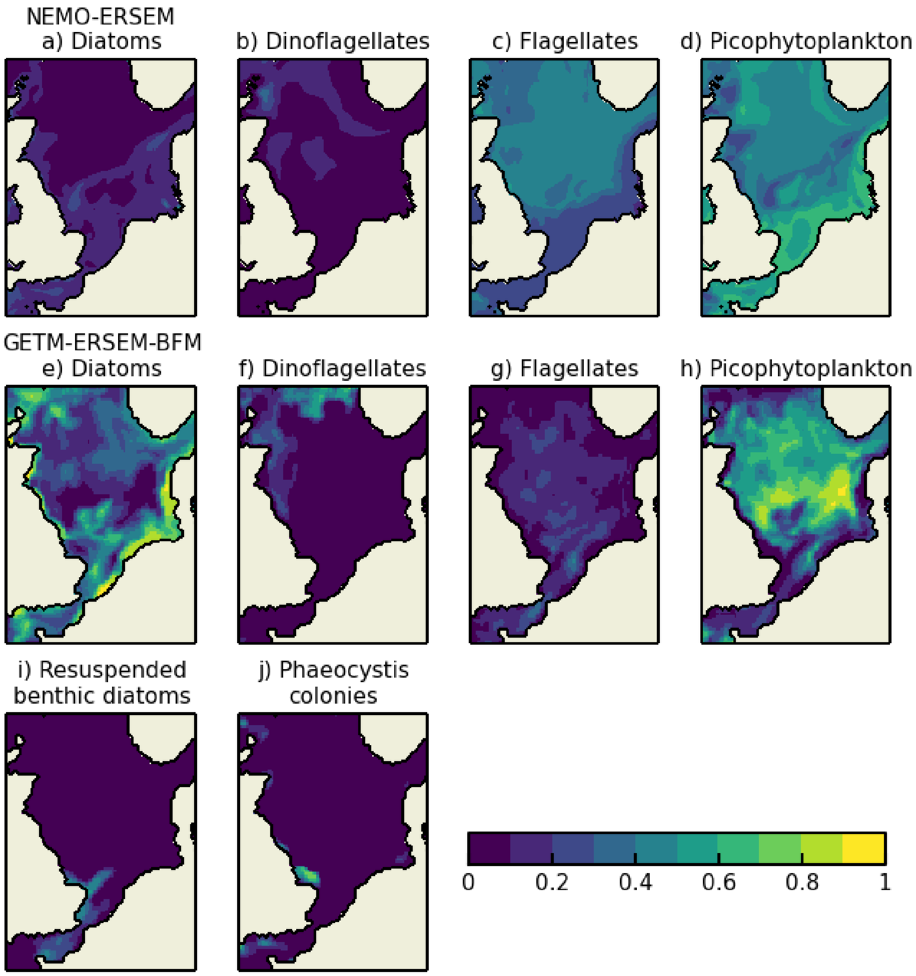


Figure 8. Monthly mean sea surface PFT fractions for August 2010 from NEMO-ERSEM (a-d) and GETM-ERSEM-BFM (e-j).

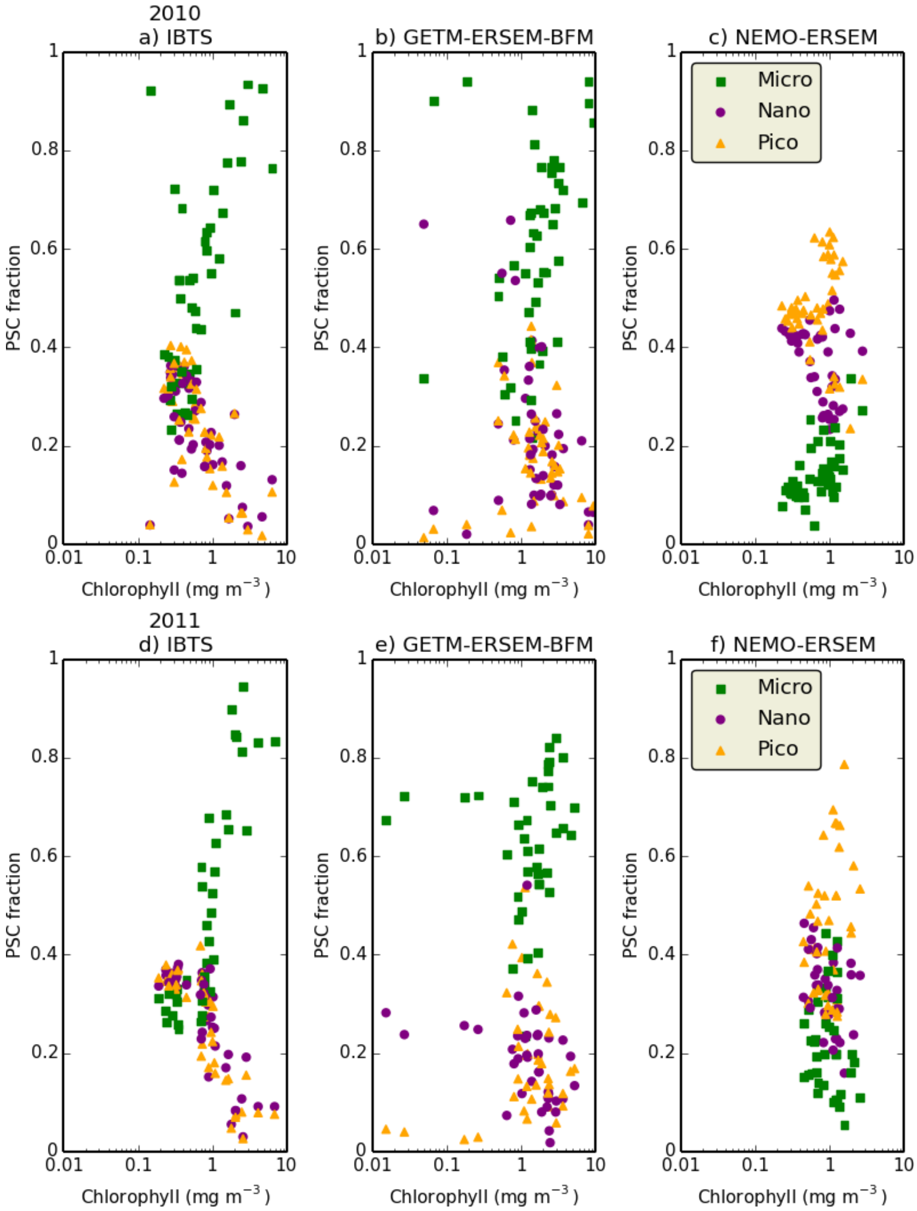
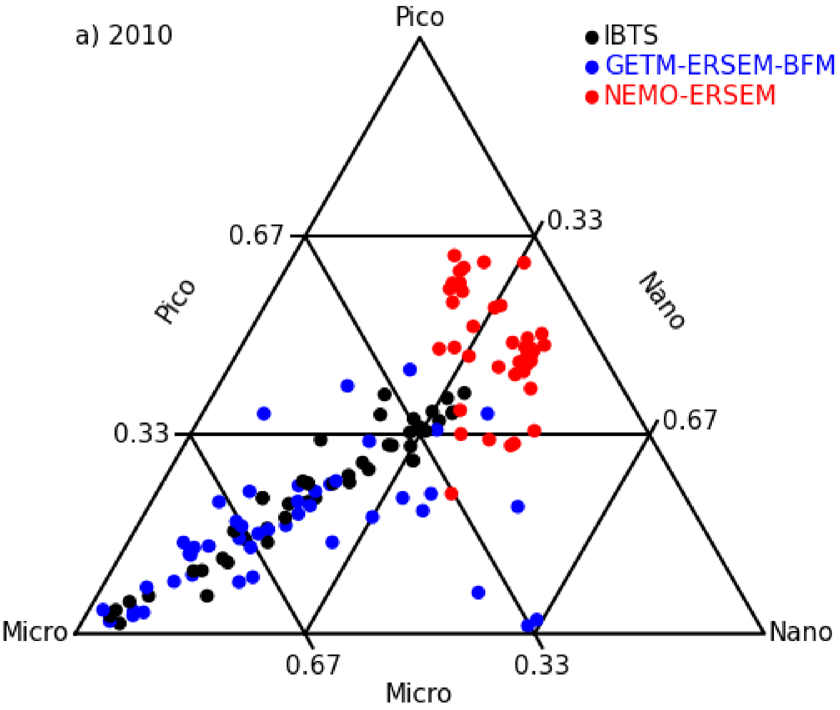


Figure 9. PSC distribution as a function of chlorophyll concentration.

Biogeosciences Discuss., doi:10.5194/bg-2016-304, 2016
Manuscript under review for journal Biogeosciences
Published: 25 July 2016



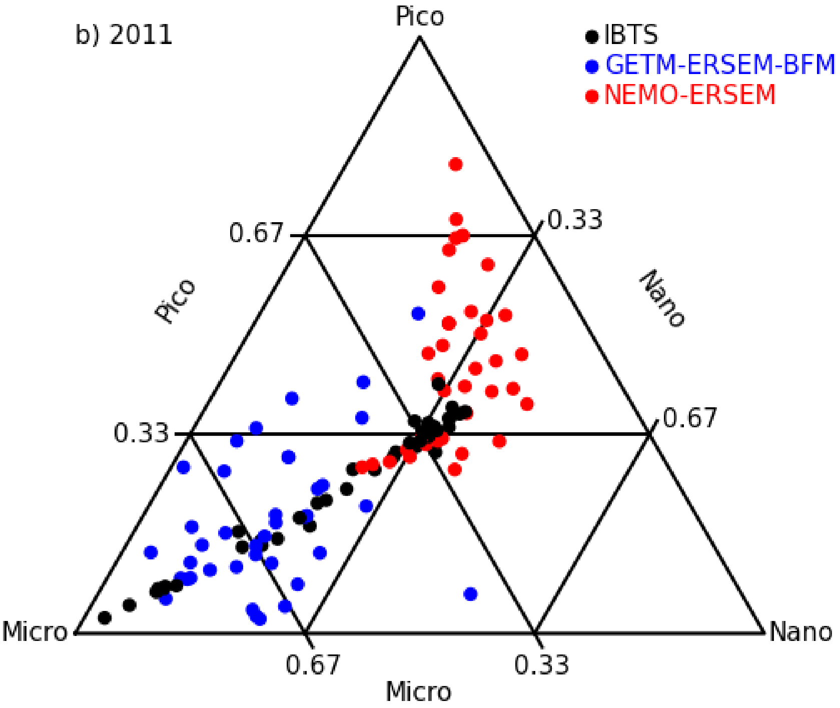
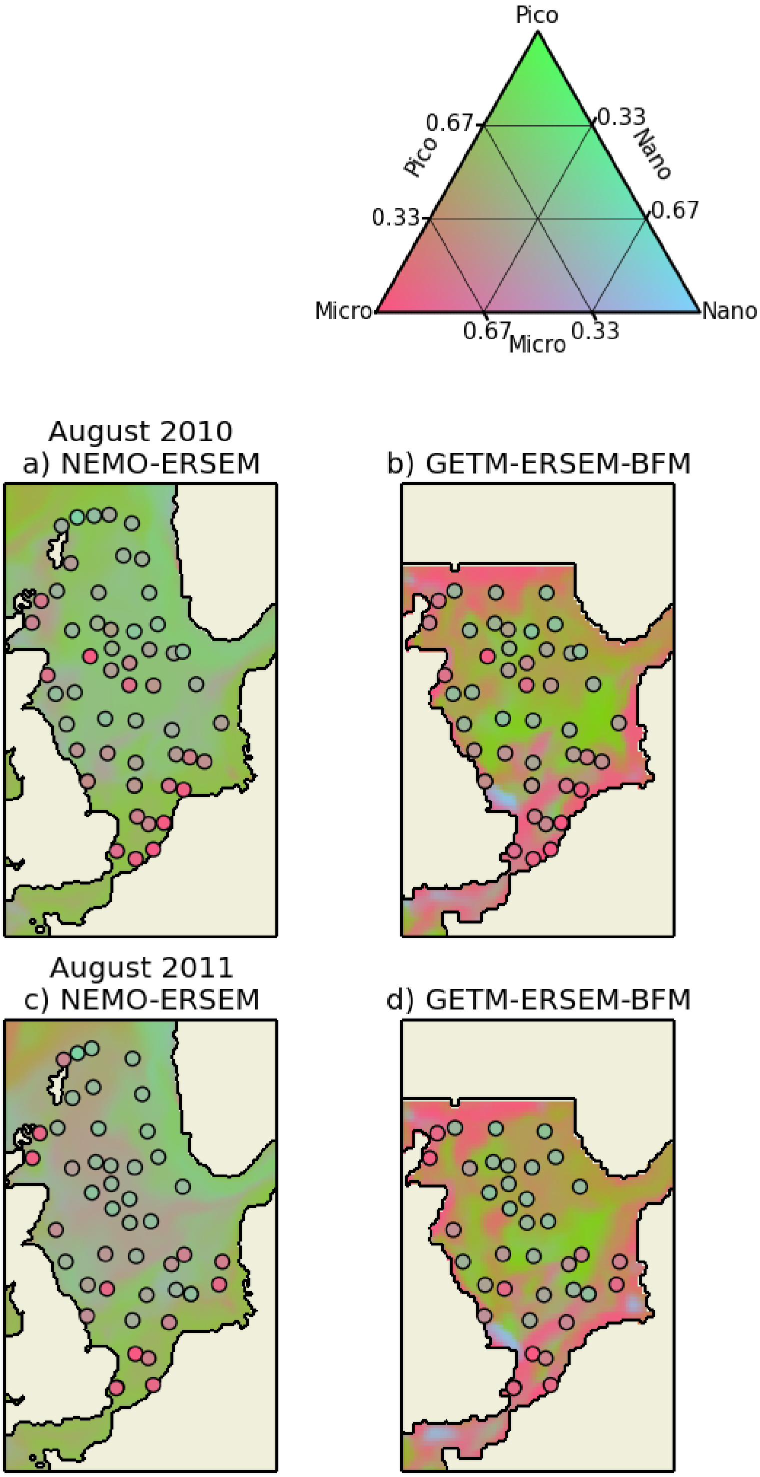


Figure 10. Ternary plots of phytoplankton community structure.



August 2010
a) NEMO-ERSEM

b) GETM-ERSEM-BFM

August 2011
c) NEMO-ERSEM

d) GETM-ERSEM-BFM

Figure 11. Maps of modelled phytoplankton community structure for August 2010 (a, b) and August 2011 (c, d) from NEMO-ERSEM (a, c) and GETM-ERSEM-BFM (b, d). The IBTS observations, sampled in August and early September of each year, are overlaid in circles.

Biogeosciences Discuss., doi:10.5194/bg-2016-304, 2016
Manuscript under review for journal Biogeosciences
Published: 25 July 2016

Table 1. General-level model characteristics.

	NEMO-ERSEM	GETM-ERSEM-BFM
Hydrodynamics	NEMO	GETM
Biogeochemistry	PML-ERSEM	ERSEM-BFM
Domain	North-west European Shelf	North Sea
Horizontal resolution	~7 km	~10 km
Vertical resolution	50 levels, terrain-following with constant 1m top box	25 levels, terrain following General Vertical Coordinates
Tidal boundary	Elevation and currents from Met Office global model, Flather radiation condition	Elevations and currents from shelf model, Flather radiation condition
Temperature and salinity boundary	Met Office global model (GloSea5 reanalysis)	ECMWF global model
Nutrients boundary	World Ocean Atlas climatology	World Ocean Atlas climatology
Meteorological forcing	ECMWF ERA-Interim reanalysis: surface temperature, 2m wind, air pressure, heat fluxes, precipitation. 3-hourly	ECMWF ERA-40 and operational hindcast: surface temperature, 10m wind, air pressure, humidity, cloud cover. 6-hourly
Atmospheric nutrient deposition	Not included	Not included
River inputs	Freshwater flow: E-Hype; nutrients: climatology; sediments: daily climatology of satellite SPM at river points	Cefas data base, interpolated daily values of runoff and nutrients based on various observational sources
SPM concentrations	Modelled, two size classes, full transport, resuspension, aggregation and disaggregation	Modelled, 1 size class with concentration-dependent settling, full transport, resuspension by waves and currents
Nutrients	N, P, Si, C, O (Fe available but not used)	N, P, Si, C, O reduction equivalents
Pelagic autotrophic types	Diatoms, flagellates, dinoflagellates, picophytoplankton	Diatoms, flagellates, dinoflagellates, picophytoplankton, *Phaeocystis* colonies, resuspended benthic diatoms, pelagic nitrifiers
Zooplankton functional types	Mesozooplankton, microzooplankton, heterotrophic nanoflagellates	Filterfeeder larvae, mesozooplankton, omnivorous mesozooplankton, microzooplankton, heterotrophic nanoflagellates
Pelagic bacteria	Pelagic bacteria	Pelagic bacteria
Pelagic detritus	Labile dissolved organic matter, semi-labile dissolved organic matter, small particulate organic matter, medium particulate organic matter, large particulate organic matter	Labile organic carbon, TEP, particulate organic carbon (POC). Degradability of POC depends on nutrient:C quota. Vertical exchange of POC coupled to SPM transport
Type of benthic model	3-layer model: oxic layer, denitrification layer, anoxic layer	3-layer model: oxic layer, denitrification layer, anoxic layer
Seabed characterisation	Distribution of the two modelled SPM size classes, dependent on model dynamics	Porosity interpolated from North Sea Benthos Survey grain size data
Benthic autotrophic types	Not included	Benthic diatoms, benthic nitrifying bacteria, nitrifying archaea
Benthic macrofauna	Deposit feeders, suspension feeders, meiobenthos	Epibenthos, Deposit Feeders, Filter Feeders, Meiobenthos, Benthic predators
Benthic bacteria	Aerobic bacteria, anaerobic bacteria	Aerobic bacteria, anaerobic bacteria
Benthic detritus	Dissolved organic matter, particulate organic matter, buried organic matter	Labile organic carbon, particulate organic carbon
CO_2 method	Available but not used	Benthic and pelagic CO_2, pH, alkalinity
Pelagic nutrient regeneration	Nitrification depends on dynamics of nitrifying bacteria	Nitrification depends on dynamics of nitrifying archaea and bacteria
Benthic nutrient regeneration	Modelling of fluxes based on estimation of nutrient gradients on basis of processes and concentrations in the 3 benthic layers	Modelling of fluxes based on estimation of nutrient gradients on basis of processes and concentrations in the 3 benthic layers for phosphate, ammonium, nitrate, reduction equivalents, silicate, DIC, alkalinity. Dynamic determination of nitrification rate from benthic nitrifier model
Spinup period	Previous hindcast of Edwards et al. (2012), run for 2007 from previously spun-up fields	1991-2001
Production run	2003-2012 (also run for 1983-1989 and 1989-2003, but sections not continuous)	2002-2011
Data assimilation	Satellite and in situ SST; 3D-Var	Not included

Table 2. Parameters of the four coinciding phytoplankton functional types related to inter-species competition.

	NEMO-ERSEM					GETM-ERSEM-BFM				
	Method/ parameter	P1	P2	P3	P4	Method/ parameter	P1	P2	P3	P4
Doubling temperature	Q_{10}	2	2	2	2	Q_{10}	2	2	2	2
Maximum productivity at 10 °C	sum	2.5	2.7	3.3	1.5	sum	3.0	4.875	5.6	1.75
Respiration rate at 10 °C	srs	0.05	0.05	0.05	0.05	srs	0.125	0.1	0.1	0.125
Fraction of PP excreted as PLOC/PDET	pu_ae	0.05	0.2	0.2	0.05	pu_ae	0.05	0.1	0.1	0.05
Activity respiration rate	pu_ra	0.1	0.25	0.25	0.25	pu_ra	0.1	0.1	0.2	0.1
Half-value of SiO_4-limitation	chP1sX	0.3	-	-	-	P_chPs	0.3	-	-	-
Miniumum quota N	qnlP?cX	0.00687	0.00687	0.00687	0.00687	p_qnlc	0.00687	0.00687	0.00687	0.00687
Minimum quota P	qplP?cX	0.0004288	0.0004288	0.0004288	0.0004288	p_qplc	0.0003931	0.0003931	0.0003931	0.0003931
Minimum quota Si	-	-	-	-	-	-	0.09	-	-	-
Multiplication factor for critical N:C ratio	xpcP?nX	1	1	1	1	Not included	-	-	-	-
Multiplication factor for critical P:C ratio	xpcP?pX	1	1	1	1	Not included	-	-	-	-
	xqnP?X	2	2	2	2	p_xqn	2	2	2	2
	xqpP?X	2	2	2	2	p_xqp	2	2	2	2
	-	-	-	-	-	p_xqs	1.5	-	-	-
Affinity for N(3) [nitrate]	quP?n3X	0.0025	0.0025	0.0025	0.0025	p_qun	0.15	0.215	1.29	0.0084
Affinity for N(4) [ammonium]	quP?n4X	0.01	0.01	0.02	0.01	Grouped with N(3)	-	-	-	-
Affinity for P	qurP?pX	0.0025	0.0025	0.0025	0.0025	p_qup	0.15	0.215	1.29	0.0084
Affinity for Si	qsP1cx	0.03	-	-	-	p_qus	0.1	0	0	0
Nutrient stress threshold for sinking	esNIP?X	0.7	0.75	0.75	0.75	Different method: based on TEP production	-	-	-	-
Sinking by formation of macro-aggregates	-	-	-	-	-	-	Threshold process: in presence of sufficient TEP and diatoms	Sticking to macro-aggregates	Sticking to macro-aggregates	Sticking to macro-aggregates
Lysis rate	sdoP?X	0.05	0.05	0.05	0.05 + 0.2	Max. by nutrient stress, p_sdmo	No lysis	0.025	0.15	0.001 + 0.0
Stress excretion	-	-	-	-	-	-	Extraction of TEP (carbohydrates)	-	-	-
Light susceptibility	P-I curve, alpha	2.98	2.98	2.98	2.98	P-I curve, Ke	35	70	124	116
Maximum Chl to C ratio	phimP?X	0.035	0.035	0.035	0.035	p_qchlc	0.02	0.035	0.035	0.035
Minimum Chl to C ratio	phiP?HX	0.025	0.025	0.025	0.025	p_qlPlc	0.0025	0.0035	0.0025	0.00225

Table 3. $\log_{10}$(chlorophyll) statistics against IBTS observations.

Year	Model	Bias	RMSE	Correlation	No. observations
2010	GETM-ERSEM-BFM	0.369	0.597	0.334	46
	NEMO-ERSEM	0.017	0.339	0.446	
2011	GETM-ERSEM-BFM	0.178	0.543	0.343	39
	NEMO-ERSEM	0.067	0.389	0.157	
2010+2011	GETM-ERSEM-BFM	0.282	0.573	0.320	85
	NEMO-ERSEM	0.040	0.363	0.358	