Marine ecosystem models for earth systems applications: The MarQUEST experience

J. Icarus Allen a,⁎, James Aiken a, Thomas R. Anderson b, Erik Buitenhuisc, Sarah Cornelled, Richard J. Geider e, Keith Haines f, Takaufumi Hirata a, Jason Holt g, Corinne Le Quéré c, h, Nicholas Hardman-Mountford a, Oliver N. Ross e, Bablu Sinhab, James While f

a Plymouth Marine Laboratory, Prospect Place, West Hoe, Plymouth. PL1 3DH, UK
b National Oceanography Centre European Way Southampton SO14 3ZH, UK
c University of East Anglia, School of Environmental Sciences, Norwich NR4 7TJ, UK
d QUEST, Department of Earth Sciences, University of Bristol, BS8 1Bj, UK
e Department of Biological Sciences University of Essex Wivenhoe Park, Colchester Essex CO4 3SQ, UK
f ESSC, Harry Pitt Building, Earley Gate, Whiteknights, Reading, RG6 6AL, UK
g Proudman Oceanographic Laboratory, Brownlow Street Liverpool L3 5DA, UK
h British Antarctic Survey, High Cross, Madingley Road, Cambridge, CB3 0ET, UK

Available online 4 January 2010

Abstract

The MarQUEST (Marine Biogeochemistry and Ecosystem Modelling Initiative in QUEST) project was established to develop improved descriptions of marine biogeochemistry, suited for the next generation of Earth system models. We review progress in these areas providing insight on the advances that have been made as well as identifying remaining key outstanding gaps for the development of the marine component of next generation Earth system models. The following issues are discussed and where appropriate results are presented; the choice of model structure, scaling processes from physiology to functional types, the ecosystem model sensitivity to changes in the physical environment, the role of the coastal ocean and new methods for the evaluation and comparison of ecosystem and biogeochemistry models. We make recommendations as to where future investment in marine ecosystem modelling should be focused, highlighting a generic software framework for model development, improved hydrodynamic models, and better parameterisation of new and existing models, reanalysis tools and ensemble simulations. The final challenge is to ensure that experimental/observational scientists are stakeholders in the models and vice versa.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

We live in a rapidly changing environment. One of the great challenges of our age is to understand and predict the consequences of changes in, inter alia, climate, biogeochemical cycles and human resource use, and mitigate the impacts. Marine science is a vital component in responding to this challenge. Over recent decades, climate models have increased the sophistication of their representations of physical ocean processes (circulation and heat transport), yet an understanding of the contemporary world requires more than this. The oceans cover 71% of the Earth surface area, and contain 95% of its volatiles and clearly play a major role in the biogeochemistry of the Earth. If we are interested in how important constituents of the atmosphere such as carbon dioxide, oxygen, nitrogen and sulphur gases interact with the oceans, it is necessary, even at first order, to understand the marine ecosystem. Ecosystems, by their very nature, are dynamic, and change in space, time and composition over a range of scales. Large-scale marine ecosystem models are tools with which we can potentially incorporate this range of variability and the processes underlying it, into a management approach.

Some of the major science challenges for marine ecosystem models are:

- to capture fundamental Earth system processes that regulate the elemental cycles of Carbon, Nitrogen, Phosphorous, Silicon, Sulphur, Iodine, Iron, etc., and the biogeochemical feedback processes between the land, ocean and atmosphere and their implications for climate;
- to assess regional ecosystem response to climate and its implications for sustainable oceans;
- to forecast the state of the marine ecosystem (near-real time to seasonal forecast);
- to identify and attribute the relative influences of multiple drivers on the ecosystem including natural fluctuations and anthropogenic perturbations.

The standard paradigm which underpins biological models (including marine ones) is the ecosystem concept; i.e., the ecosystem as a system consisting of interacting biotic and abiotic components,
resulting in a stable entity in which the exchange of materials between the living and non-living parts follows circular paths (Odum, 1953). A consequence of this is a systems analysis approach to ecology, whereby an ecosystem is viewed as being analogous to a machine. This approach offers a pragmatic method of improving our understanding of the complexity of natural systems (O’Neill, 2001). The great advantage of this concept is that the myriad complex interactions at a species level can be cut through by focusing on a small subset of average or integrated properties of all the populations within the area of study. A further advantage is that it allows the identification of emergent properties such as energy flow and nutrient cycling, and the investigation of the functional stability of this abstract structure. The weakness of the systems approach lies in its limited ability to explain the relative stability of ecological systems in a changing environment. Additionally there is the issue of complexity: abstraction means that important processes can be missed. With its focus being on a self-regulating system, and consequently on local and short term stability (i.e., recovery from disturbance) it is ill-equipped to explore flexibility in the sense of maintaining variability in space and time as conditions change (O’Neill, 2001). The ecosystem approach provides the conceptual framework underpinning the current generation of coupled biophysical models, with their limited ability to respond to environmental change. This limited ability may of course be due to missing or mis-parameterised processes.

It is increasingly evident that a deeper understanding of the marine biogeochemical system is required (Doney, 1999). One approach to this has been to incorporate more explicit ecosystem descriptions into models. Simple NPZD schemes (incorporating one Nutrient term, one primary Producer, one consumer (Zooplankton), and one Detritus) employed since the late eighties (e.g., Fasham et al., 1990) may often capture bulk properties and the essential dynamics of events such as the North Atlantic bloom. This description can be elaborated somewhat (3N–2P–2Z–2D models, for instance, Aumont et al., 2003) and may begin to capture certain key feedbacks in much of the world’s ocean. However, in order to describe the multidimensional behaviour of ecosystems and their interaction with many interlinked biogeochemical cycles, the degree of elaboration may have to increase substantially (e.g., Hood et al. 2006).

Land models began by classifying vegetation into biomes, with defined vegetation characteristics. The desire to capture biogeochemical processes better, especially the carbon cycle, resulted in efforts to develop physiologically-based representations of plant functional types (e.g., Sitch et al., 2003) that have measurable characteristics (e.g., for carbon allocation within the plant structure). By reducing the complexity of species diversity in terms of ecological function, and representing the landscape as patches of plant functional types rather than biomes, climate and ecosystem models could then be linked together.

In recent years we have begun to apply a similar approach to marine ecosystems and organisms, pioneered by the European Regional Sea Ecosystem Model (ERSEM; Baretta et al., 1995). However, the situation for marine ecosystems is in many respect quite different from that on the land: abiotic controls on ecosystem distribution are profoundly different, and a unique property of marine ecosystems is that they exist in dynamic environments, being influenced by ocean physics, currents and mixing on (macro-) scales ranging from tens of metres in estuaries and coastal zones to thousands of kilometres in ocean gyres. This does not preclude classification of the marine environment by biome (e.g. Longhurst et al., 1995, Hardman-Mountford et al., 2008), however, these differences raise fundamental questions about how best to incorporate our physiological understanding into ecosystem models. Furthermore, marine plankton functional types (PFTs) are currently defined largely by the geochemical transformations made by the group of organisms, for instance calcifiers, silicifiers, DMS producers, pico-heterotrophs, etc. In the ocean, widely different classes of organism often fill each of these descriptions, and it is not yet clear whether a useful classification can be made without going to considerably greater detail. A fundamental problem therefore, is to find the appropriate level of complexity that will enable ecosystem models to have the most skill in predicting biogeochemical fluxes. We must bear in mind that the level of complexity also depends on how well we can parameterise things; the quest for greater detail has to be tempered by our ignorance of the ecology and physiology of the organisms in question.

A numerical model is a temporally dynamic cartoon of a system; ideally it should capture the essential bulk properties and feedbacks of the system we are interested in preferably at the level of the data available to support it. Marine ecosystem models should also produce emergent properties (i.e. outcomes which are not a direct extrapolation of the choices made in the model design).

The Marine Biogeochemistry and Ecosystem Modelling Initiative in QUEST (MarQUEST) project is a program of activities to develop improved descriptions of marine biogeochemistry, suited for the next generation of Earth system models. Specifically, it has aimed to take account of the global role of marine ecosystems in biogeochemical cycles, especially the processes determining plankton functional-type composition and its influence on elemental cycling, and the interactions of these processes with the physics and chemistry of the ocean. Particular emphasis has been given to

- choosing appropriate model structures based on the representation of plankton functional types for marine ecosystems;
- proposing improved ways to scale processes from physiology to functional types;
- quantifying the importance of the representation of ocean physics for the representation of marine ecosystems and biogeochemical cycles;
- developing techniques to incorporate coastal-ocean processes into global models and;
- developing new methods to evaluate and compare ecosystem and biogeochemistry models.

Here we present progress on each of these topics, and discuss the experience and prospects.

2. Key issues

2.1. Progress in the representation of plankton functional types for marine ecosystems

A major challenge is the specification and construction of a PFT model suitable for resolving the interactions of multiple biogeochemical cycles in an Earth system model. PlankTOM (Plankton Types Ocean Model) is a global model that incorporates the C, N, Si, P and Fe cycles and phytoplankton chlorophyll. The model is being constructed following the ideas proposed in Le Quéré et al. (2005), and is under constant development, with the identification, prioritisation and parameterisation of the PFTs being informed through a series of international discussions, the Dynamic Green Ocean workshops (http://www.lgmacweb.env.uea.ac.uk/green_ocean). PlankTOM’s ten PFTs and their biogeochemical roles are summarised in Table 1. Le Quéré et al. (2005) suggest the following guidelines for defining a PFT;

1. The PFT should have an explicit biogeochemical role,
2. The PFT should be defined by a distinct set of physiological, environmental, or nutrient requirements controlling its biomass and productivity,
3. The behaviour of the PFT should lead to competitive interaction with other PFTs, for instance, through selective depletion of nutrients or grazing and
4. The PFT should be of quantitative importance in at least some region of the ocean.

Le Quéré et al. (2005) further argue that there are distinct advantages to building biogeochemical models that use PFTs based on
are insufﬁcient observations to constrain many of the rates. The lack of observational data leads to difficulties in aggregating diversity within functional groups into meaningful state variables with meaningful parameters. In recent years, the trend of publishing observational data as electronic supplements to papers and the rise in submission of these data to data centres for archiving has improved the accessibility of observations for use by modellers. In turn, this has stimulated progress in deﬁning the traits (equations, parameters and concentrations) of PFTs, including the statistical signiﬁcance with which parameters can be distinguished. Thus taxonomic groups with distinct physiological traits can potentially be identiﬁed, meeting one of the criteria listed above that Le Quéré et al. (2005) proposed for deﬁning unique PFTs. When the other criteria are also met, further PFTs can be distinguished. Coupled with an exponential growth in computer power, this allows for an increase in the number of trophic levels that can be integrated in one model and may lead to a better representation of the ecological diversity within trophic levels.

2.2. Scaling processes: physiology to functional types

For decades, models of terrestrial plant photosynthesis have been ﬁrmly based on biochemical mechanisms as exempliﬁed by the Farquhar et al. (1980) model of C3 photosynthesis. In contrast, until recently most models of phytoplankton growth have been unabashedly empirical. The commonly used approach in most NPZ and PFT models employs the Monod equation to describe growth rate as a function of the concentration of a limiting nutrient for nutrient-limiting growth, in conjunction with a light-response curve for nutrient-replete growth (Fasham et al., 1990). The desire to capture important aspects of phytoplankton physiology in phytoplankton growth models prompted Geider et al. (1996, 1998)) and Flynn et al., 1997, Flynn (2001), amongst others, to develop dynamic models of growth and acclimation, focusing on photosynthesis, photo-acclimation and nutrient uptake/assimilation. Although some of these models have been described as “mechanistic”, the mechanisms are caricatures of algal physiology and biochemistry. However, even these caricatures are already too computationally expensive to be used in global-scale models of ocean biogeochemistry embedded within 3-D ocean general circulation models. In global-scale models, the trade-off between biological realism and computational tractability often falls heavily in favour of tractability. Now, however, this should be redressed, given the substantial and continuing increases in computational power. A more important constraint on incorporating biological realism is the difﬁculty in providing generic parameterisations as the complexity of the biological model increases. This is important in the context of Levins (1966) who argued that you can only have two out of three of complexity, generality (e.g. size of domain) and accuracy.

A major step forward in ocean biogeochemical models was subdividing generic phytoplankton used in early NPZD models into a small number of phytoplankton functional types (see Section 2.1). Use of PFTs has allowed the metabolic complexity of pelagic ecosystem models to more fully reﬂect a range of ecosystem processes such as N2 ﬁxation and calcification, in addition to primary production. Within the phytoplankton PFTs, further progress has been made by allowing combinations of variable chl:C:N:P:Fe (e.g. Baretta-Bekker et al., 1998; Moore et al. 2002; Blackford et al. 2004, Buitenhuis and Le Quéré, unpublished). However this has come at the cost of increasing the number of biomass variables within each phytoplankton functional type and the cost of placing greater demands on the calibration and validation of models.

It can be argued that the next step that needs to be taken is to increase the sophistication of the description of phytoplankton physiology by moving beyond empirical Monod and Droop-type models (Hood et al. 2007). However, this presents an even greater challenge than the introduction of variable chl:C:N:P:Fe stoichiometry. An important issue, which has not yet received enough attention, is how much physiological detail is needed to model the changing role of the phytoplankton as the ocean environment changes. As described below, we have made steps towards addressing this issue in MarQUEST.

---

<table>
<thead>
<tr>
<th>PFT</th>
<th>“Biogeochemical function”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteria</td>
<td>Remineralise dissolved and particulate organic matter</td>
</tr>
<tr>
<td>Pico-heterotrophs</td>
<td></td>
</tr>
<tr>
<td>Pico-autotrophs</td>
<td>Contribute to primary production, but not to export of carbon</td>
</tr>
<tr>
<td>Phytoplankton</td>
<td>Control total amount of reactive N</td>
</tr>
<tr>
<td>N-fixers</td>
<td></td>
</tr>
<tr>
<td>Calciﬁers</td>
<td>Produce more than half the marine carbonate ∆n, sensitive to pH</td>
</tr>
<tr>
<td>DMS producers</td>
<td>Inﬂuence atmospheric sulphur cycle</td>
</tr>
<tr>
<td>Mixed</td>
<td>Background biomass of phytoplankton</td>
</tr>
<tr>
<td>Silicifers</td>
<td>Contribute to export of carbon to deep ocean</td>
</tr>
<tr>
<td>Zooplankton</td>
<td>Graze on small phytoplankton, control blooms</td>
</tr>
<tr>
<td>Proto</td>
<td>Graze on all sizes of plankton, produce fast-sinking faecal pellets which export carbon</td>
</tr>
<tr>
<td>Meso</td>
<td></td>
</tr>
<tr>
<td>Macro</td>
<td>Graze on all sizes of phytoplankton, produce fast-sinking faecal pellets</td>
</tr>
</tbody>
</table>

---

Taxonomical groups that are already used by biological oceanographers in ﬁeld and laboratory observations:

• Observations of rates and concentrations can be directly used in model equations, parameterisation and evaluation.

• Observational oceanographers have often already processed and synthesised observational data sets. Amongst other beneﬁts, this reduces the risks of misinterpretation by modellers not specialised in handling these groups’ data.

• Many system level properties have already been deﬁned in the literature.

Examples of such system level properties include global rates that are based on datasets independent from the ones used to construct the model, zooplankton PHYTOPLANKTON ratios as an indicator of trophic interactions, and inter-annual variability of chlorophyll a (e.g. Prentice et al., 2004, Buitenhuis et al., 2006).

However, the normal tension between simpliﬁcation and realism in representing a very complex system is underlain in our marine models, zooplankton PHYTOPLANKTON ratios as an indicator of trophic interactions, and inter-annual variability of chlorophyll a (e.g. Prentice et al., 2004, Buitenhuis et al., 2006) with a resolving power that is limited in various ways. For example, there is a need for information on the small-scale variability of plankton communities, which is not readily accessible in standard observing systems. There is also a need to better understand the processes that govern the distribution and abundance of plankton species and groups.

Table 1

PLANKTOM10. The 10 plankton functional types indentiﬁed as being biogeochemically signiﬁcant for marine biogeochemical cycling in an earth system model by Le Quéré et al. (2005).
by using cell-based models to examine tradeoffs between the costs and benefits of different photo-acclimation strategies.

One of the goals of MarQUEST has been developing a physiologically-based description of phytoplankton responses to changing light levels, respiration and nutrient limitation in order to deliver a better understanding of the effects of vertical mixing on phytoplankton growth dynamics. Two issues that we have addressed in tackling this issue are (1) establishing the errors inherent in treating discrete biological entities (e.g., phytoplankton cells) as continuous variables and (2) providing a physiological model that allows costs and benefits of physiological tradeoffs to be examined. To address this issue, the dynamic model of C, N and Chl dynamics described by Geider et al. (1998) was modified to use the cell as the basic structural unit of the phytoplankton community in order to separate energy reserve and functional carbon pools, and to describe how respiration during prolonged dark periods (>24 h) leads to debilitation of the photosynthetic apparatus and eventually to cell death (Ross and Geider, 2009). To date, simulations have been made with nutrient-replete conditions and (2) providing a physiological model that allows costs and benefits of physiological tradeoffs to be examined. To address this issue, the dynamic model of C, N and Chl dynamics described by Geider et al. (1998) was modified to use the cell as the basic structural unit of the phytoplankton community in order to separate energy reserve and functional carbon pools, and to describe how respiration during prolonged dark periods (>24 h) leads to debilitation of the photosynthetic apparatus and eventually to cell death (Ross and Geider, 2009). To date, simulations have been made with nutrient-replete conditions and (2) providing a physiological model that allows costs and benefits of physiological tradeoffs to be examined. To address this issue, the dynamic model of C, N and Chl dynamics described by Geider et al. (1998) was modified to use the cell as the basic structural unit of the phytoplankton community in order to separate energy reserve and functional carbon pools, and to describe how respiration during prolonged dark periods (>24 h) leads to debilitation of the photosynthetic apparatus and eventually to cell death (Ross and Geider, 2009). To date, simulations have been made with nutrient-replete conditions and (2) providing a physiological model that allows costs and benefits of physiological tradeoffs to be examined. To address this issue, the dynamic model of C, N and Chl dynamics described by Geider et al. (1998) was modified to use the cell as the basic structural unit of the phytoplankton community in order to separate energy reserve and functional carbon pools, and to describe how respiration during prolonged dark periods (>24 h) leads to debilitation of the photosynthetic apparatus and eventually to cell death (Ross and Geider, 2009). To date, simulations have been made with nutrient-replete conditions

**Fig. 1.** Comparing the distribution of CHL:C ratios in the SML in Lagrangian and Eulerian-type simulations. In (a) and (c) the turbulent SML (diffusivity of 5 \( \times 10^{-2} \) m\(^2\) s\(^{-1}\)) is 50 m deep and overlays a non-turbulent bottom layer (10\(^{-3}\) m\(^2\) s\(^{-1}\)) with a diffusivity profile that is continuous and differentiable throughout. In (b) and (d) the entire 200 m water column is turbulent. The maximum surface irradiance at noon was 650 W m\(^{-2}\), of which 45% was assumed to be photosynthetically active, with a semi-sinusoidal distribution over a 12 h light period. The absorption coefficient in the water column was 0.09 m\(^{-1}\).
2.3. The importance of high quality physics: how do we get it right?

The behaviour of ocean ecosystems is critically dependent upon the environment in which they exist. This environment depends upon the physical behaviour of the ocean or, in numerical simulations, of the behaviour of the physical ocean model used. In particular, phytoplankton growth rates depend on the nutrient supply through physical mixing, while the amount of light available to phytoplankton is dependent upon the Mixed Layer Depth (MLD) and the inherent optical properties of the water (related to concentrations of coloured dissolved organic matter (CDOM), suspended particulate material (SPM) and phytoplankton pigment concentrations). As such, it is important that the most accurate description possible of the true ocean state is used in conjunction with biogeochemical models. A major goal of MarQUEST is to assess the influence of the quality of physical simulations of the performance of ecosystem models.

2.3.1. Sensitivity of PFT models to differences in hydrodynamic forcing

The sensitivity of the present generation of global ecosystem models to the representation of the physical environment was investigated in an experiment reported by Sinha et al. (in press). The PlankTOMS (Le Quéré et al., 2005) biogeochemical model was embedded, using identical equations and parameter values for the modelled ecosystem, in two different physical ocean general circulation models of comparable resolution, namely NEMO (Nucleus for European Modelling of the Ocean, Madec, 2008) and OCCAM (Ocean Circulation and Climate Advanced Modelling Project, Marsh et al., 2005). Initial conditions, light forcing of biological production and nutrient inputs (atmospheric iron) were arranged to be the same in two multi-year simulations. Further, both physical models were forced at the surface with slightly differing versions of the National Centres for Environmental Prediction (NCEP) reanalysis of historical atmospheric data, the chief difference being that OCCAM was run with six-hourly surface meteorological fields and included a diurnal cycle in incident shortwave radiation whereas NEMO was run with daily forcing and did not include a diurnal cycle. Importantly, in both simulations the same surface photosynthetically active radiation (PAR) was applied to the biology. Results showed significant sensitivity in the predicted distributions of PFTs to the different formulations of the physics employed (Fig. 2), both globally and regionally. There was clear evidence that differences in the amount of upwelling and convective mixing, and in the depth and seasonality of upper mixed layers predicted by the Ocean General Circulation Models (OGCMs), caused switches in the community structure of the simulated ecosystem (e.g., from one dominated by mixed phytoplankton and microzooplankton to one dominated by diatoms and mesozooplankton). Small differences in physics (i.e., within the range typically expected from running different ocean models) were thus able to exert a disproportionate effect on the ecosystem through biological feedback mechanisms, including bottom-up control by nutrient supply, top-down control by zooplankton and changes in the relative competitiveness of the various plankton functional types. On the other hand global bulk properties did not change substantially (e.g. primary production of 45.9 Gt yr$^{-1}$ in NEMO versus 47.1 Gt yr$^{-1}$ in OCCAM and export (at 150 m) of 8.9 Gt yr$^{-1}$ in NEMO versus 9.6 Gt yr$^{-1}$ in OCCAM), indicating that these properties are likely set by external influences (light, nutrient availability) and are less dependent on the type of community structure that develops.

The implications of this study are profound. On the one hand the great importance of employing realistic physics was demonstrated, requiring technical advances (e.g. increasing resolution, improvements to forcing datasets and parameterization of physical processes such as mixing) and dialogue between ecosystem modellers, ocean modellers and observationalists to ensure that the most appropriate ocean models are used for biogeochemical studies. On the other hand the peculiar sensitivity of the biological equations themselves was evident, exacerbating the requirement for realistic physics. The latter problem, which Anderson (2005) coined “all in the interactions”, presents a worrisome concern for modellers, demanding a fresh look at some of the fundamental assumptions currently built in to our ecosystem models in context of the general field of complexity theory. At higher levels of model complexity, associated processes and feedbacks appear to be remarkably sensitive to the physical environment and parameter choices, although bulk properties (associated with lower complexity)

![Fig. 2](image.png) (a) Dominant phytoplankton functional type (green = mixed phytoplankton, yellow = diatoms, brown = coccolithophores) for March–May and September–November, 1994 for the NEMO simulation (b) same as (a) for the OCCAM simulation.
such as total chlorophyll and primary production may be relatively less
affected. We need to invest more effort in understanding the origin of
this sensitivity and where appropriate in developing alternative and/or
complementary approaches to the representation of complexity (e.g.
Bruggeman and Kooijman 2007).

Our experiments also suggest that relatively small changes in
large-scale ocean physics could lead to structural changes in marine
ecosystems, though the consequences on the carbon fluxes are not
immediately evident. More targeted modelling work is required to
elucidate this problem.

2.3.2. Reanalysis simulations

All physical ocean models have some biases with respect to
observations, but data assimilation offers a way to reduce these biases
and improve a model’s representation of true ocean physics. To test
the effects of physical data assimilation on biogeochemical models,
experiments have been conducted with the NEMO-ORCA2 2° ocean
model with an embedded ecosystem. The assimilation method
described in Haines et al. (2006), Smith and Haines (2009) was
used to assimilate temperature and salinity profile data obtained from
the ENSEMBLES database (Ingleby and Huddleston 2007). Experi-
ments were run for the period 1990–2007, with the first 10 years
being treated as a spin-up period. Model forcing was taken from the
DRAKKAR Forcing Set 3 (Molines et al., 2006); operational analyses
from the European Centre for Medium-Range Weather Forecasts were
used to extend the forcing set up until 2007. The embedded biological
model used for the experiments was the plankTOM5 plankton
functional-type model (Le Quéré et al., 2005).

Two experiments were undertaken: a control, which included no
assimilation, and an assimilating run. Results from these experiments
showed that biases in the physical fields were reduced by the
assimilation. When averaged over the top 270 m of the ocean, RMS
differences between the model temperature and the WOA05
(Locarnini et al., 2006) climatology were 0.7 °C for the assimilating
run, compared to 1.2 °C for the control, with Smith and Haines (2009)
demonstrating that some of this residual is due to the climatology not
being fully representative of ocean physics in the early 21st century.
Similarly, RMS biases in the depth of the SML, when compared to the
climatology of de Boyer Montégut et al. (2004), improved from 20.8 m
in the control to 19.9 m in the assimilating run. Unfortunately, relative
to the control, this first assimilating run significantly overestimated
both the primary production and the surface chlorophyll distribution
despite the reduced biases in the physical fields. The principal cause of
the excess primary production was an increase in surface phosphate
availability. This excess phosphate was due to data assimilation
changing the nutrient properties of water masses and acting to increase
vertical and horizontal nutrient mixing and flux within the model,
particularly within the subtropical gyres. The majority of the excess
phosphate was brought up from depth via vertical advection within the
western boundary currents, and was then laterally mixed around the
gyres. However, assimilation also generates spurious mixing fluxes of
material, and away from the gyres this can also be a source of bias.

To tackle the problem of excess nutrient availability a new assimila-
tion experiment, the nutrient increment experiment, was conducted. In
this experiment nutrients (phosphate and silicate) were adjusted through
each data assimilation cycle, so as to maintain the nutrient-density N(p)
relationship in each water column. This is similar to the method employed
in Haines et al. (2006) and Smith and Haines (2009), where a salinity–
temperature relationship is maintained as temperature data are
assimilated. Results from the nutrient increment experiment are shown in
Fig. 3, alongside equivalent results from the control experiment.

Fig. 3a–d shows that the nutrient increment run has greatly
improved phosphate distributions throughout the tropics, and at higher
latitudes. The subtropical gyres remain a problem as in this region there
is still a superabundance of phosphate relative to climatology.
Nevertheless the nutrient increment is an important step forward. For
dexample, using data assimilation it should be possible to ensure that
different physical models, such as those described in Section 2.3.1, can
be bias corrected to give similar upper ocean physical fields, thus
ensuring more consistent ecosystem behaviour.

However in our nutrient incrementing model run chlorophyll is still
systematically biased above SeaWiFS data (shown in Fig. 3e,f), even
outside the subtropical gyres. This occurs despite the more realistic
nutrient and physical environments, suggesting that the biological
model is now producing too much plankton for the available nutrient.
This could be either due to an overestimation of phytoplankton carbon,
or phytoplankton chl:nutrient ratio. Phytoplankton carbon assimilation
was taken from observations and have not been tuned (Buitenhuis et al.
2006), Buitenhuis, Rivkin, Sailey, Le Quere submitted to L&O). The C:
nutrient ratio in this model is fixed (and thus potentially under-
estimated in these nutrient limited regions), and the chl:C ratio is also
underestimated in the subtropical gyres (data not shown). Such results
imply that collaborative interaction between biogeochemical and
physical modellers will be required to make further progress.

2.4. Representing the coastal ocean in Earth system models

Accurately simulating the coastal zones, the continental shelves
and the ocean margins represents a particular challenge for Earth
system models, owing to the complex biophysical interactions in
these regions that occur on scale orders of magnitude smaller than in
the open ocean. That this class of geographical regime deserves special
attention is apparent in three important respects: their role in global
biogeochemical cycles, their role in global ocean dynamics, and their
socio-economic significance.

It is well established that shelf seas are regions of exceptionally high
biological productivity. This arises from the re-supply of nutrients to
otherwise depleted surface waters by various processes, including
heterotrophic nutrient recycling (by zooplankton and bacteria in pelagic
and benthic ecosystems), coastal upwelling, cross-frontal transport, and
riverine inputs. There is still significant doubt as to the accuracy of
satellite estimates of primary production (Carr et al., 2006) in the near-
coastal region. They are invariably contaminated by non-biotic influence
on the optical properties (Sathyendranath et al., 1989; Smyth et al.,
2006). However there is good qualitative agreement with in situ
observations showing that shelf seas are regions of high primary pro-
duction (e.g., Joint and Pomeroy, 1993), and that near-coastal regions
can be areas of exceptionally high production (e.g., Cadee and Hegeman,
2002). Hence, while accounting for only 8% of the surface area of the
global ocean, estimates suggest the coastal ocean contributes 30% of the
oceanic primary production (Longhurst et al., 1995; Walsh et al., 1988;
Walsh et al., 1991), 80% of the organic matter burial, 90% of the sedi-
mentary mineralization and 50% of the deposition of calcium carbonate
(Gattuso et al., 1998; Mantoura et al., 1991; Pernetta and Milliman,
1995; Walsh et al., 1988). With regard to the global carbon cycle, there
is still no clear consensus as to whether the coastal regions are a net source
or sink of CO2 to the atmosphere, nor to how much carbon is exported to
the deep ocean (Borges et al., 2005). This uncertainty arises in part
because of the complex biophysical interactions in shelf seas. Primary
production draws down atmospheric CO2, but unlike in the open ocean,
where particulate carbon can sink directly to the deep ocean and be
removed from contact with the atmosphere, the shelf-sea carbon pump
(Tsunogai et al., 1999) often requires lateral transport to remove carbon
from the ventilated surface waters. This opens up the possibility for re-
mineralization and hence there is no straightforward relationship
between primary production and carbon sequestration in shelf seas.
Shelf seas also impact on global ocean circulation through the
frictional effects on ocean currents at the shelf-slope regions and the
formation of dense water in high latitude shelf seas, which make a
major contribution to the meridional overturning circulation.
The properties and processes important in the coastal ocean are determined by two important characteristics of shelf seas: their depths are shallow and they are close to land. The former leads to long-waves (e.g. tides and coastal trapped waves) growing in amplitude, high levels of friction/mixing, and turbulent boundary layers often being of similar thickness to the water depth. The latter leads to complex flow regimes around the coastline, restrictions between shelf-sea basins, estuaries and sea lochs/fjords, and to terrestrial inputs. These inputs take the form of freshwater, nutrients, optically active constituents from rivers, sediments from cliff erosion, and aeolian deposits of nutrients. Because riverine nutrient inputs are often accompanied by SPM and CDOM, coastal regions are characteristically light-limited, where the water's optical properties rather than nutrient availability determine the phytoplankton growth (differing significantly from typical open ocean conditions). A unifying characteristic of all these processes is that their horizontal scales are small in the global context. For example, the wavelength and adjustment length (Rossby Radii) of long barotropic and baroclinic waves both tend to scale by the square root of the water depth $h^{0.5}$ so decrease by an order of magnitude as the water depth reduces from 4000 m to 40 m. Hence the barotropic Rossby radius at

Fig. 3. Biological outputs of assimilation experiments. Global distribution of phosphate concentration from the control run (a), and the nutrient increment run (b). The residual: WOA05 phosphate minus the control run (c) and the nutrient increment run (d). The chlorophyll residual from the control run (e) and the nutrient increment (f) run minus SeaWiFS chlorophyll (http://oceancolor.gsfc.nasa.gov/SeaWiFS/). Note that the large negative values are due to SeaWiFS overestimating chlorophyll in case 2 waters. In the case of phosphate and its residuals, values given are the average between the surface and 270 m in depth.
mid-latitudes is \(\sim 200\) km and the baroclinic Rossby radius is typically 2–20 km. The dynamic scales in turn determine the scales of the distribution of material transported from the land–sea interface, such as in river plumes/coastal currents (1–300 km globally; Warrick and Fong, 2004), and exchange processes at fronts.

From the scale of the ocean models used in the IPCC 4th assessment report, it is immediately apparent that the current generation of coupled climate models is a long way from being able to represent shelf-sea processes. The challenge is how to introduce aspects of coastal-ocean physical processes, for which we have high-resolution knowledge, into a coarse resolution OGCM, without substantially impacting the required computational resource, i.e. maintaining the grid and time step of the OGCM. This necessitates a highly parameterized approach.

2.4.1. The vertical resolution in shelf seas

Many OGCMs with geopotential coordinates lose resolution in coastal regions. Using variable vertical coordinates (e.g., \(Z^*\)) allows these levels to be used to enhance coastal-ocean resolution without impacting on the run time. Care is needed, however, not to introduce errors in the dynamics associated with steeply sloping coordinates. Results from the ORCA-1 application of NEMO in the North Sea (Fig. 4) demonstrate the inadequacy of this global model in shelf seas. Only nine levels remain at this water depth (73 m), which when combined with the lack of tidal mixing to sharpen the thermocline from below, lead to a very diffuse thermocline compared with the sharp thermoclines observed in shelf seas (Sharples et al., 2001). Given the strong relation between phytoplankton growth and vertical temperature structure in the region (Sharples, 1999) it is evident that an ecosystem model run coupled to this physics model would be severely compromised further emphasising the difficulties associated with utilising biogeochemical models in OGCMs identified in Section 2.3.1 above.

2.4.2. Tidal mixing

On centennial time scales, tides are approximately invariant, changing predominately with bathymetric evolution at the coastline. Hence, to a good approximation, high-resolution tidal models (e.g., Lyard et al., 2006) can be used to drive tidal mixing models in shelf seas. The latter are well established vertical 1D models driven by the sea surface slope (e.g., GOTM). These can be run at each coastal-sea grid point to provide a better estimate of mixing and stratification. Fig. 5 shows an example where a 1D water column mixing application of POLCOMS

Fig. 4. Time-series of temperature in the central North Sea seasonally stratified region, from a) the ORCA-1 application of NEMO.

Fig. 5. Time-series of temperature in the central North Sea seasonally stratified region. The full 3D MRCS POLCOMS model (Holt and Umlauf, 2008) and a 1D model forced by tidally oscillating sea surface slope.
forced by tidally oscillating sea surface slope is compared with a fully three-dimensional POLCOMS application. While some of the details differ, the overall structure is quantitatively very similar. The ORCA-1 model captures the seasonal stratification but does not reproduce the sharp thermocline observed in shelf seas (e.g. Sharples et al 2001, Fig. 4). Moderately improved vertical resolution (18 levels as opposed to 9) and mixing from the benthic boundary layer leads to improvement in POLCOMS, which is straightforwardly realisable by 1D vertically resolved processes.

2.4.3. River plumes, coastal currents and coastal polynias

Phenomena arising from small-scale dynamical processes represent a particular challenge. However, parameterisations of river plume spreading are available (e.g. Yankovsky and Chapman, 1997). By implementing these on a grid-cell by grid-cell basis based on cyclonic distance to a river mouth, a simple parameterisation of a coastal current can be established. This allows us to represent the important correlation between rivers as a source of nutrients and a source of optically active material (CDOM and SPM). Polynias are pools of open water within areas of sea-ice. Like river plumes and coastal currents, these features can be represented by simple parameterisations, which can then be used to represent the effects of ice-free regions on coastal production.

Addressing this category of phenomena requires the introduction of an additional degree of dimensionality into the ocean model, characterising qualitatively different types of coastal water in ways that are analogous to tiling in land models. In these two cases, there might be two classes in each: shelf sea water, and within-river plume water and within-polynia water, respectively. The parameterisation of the processes determines the fraction in each and the forcing (light/riverine nutrient availability) is applied separately to each fraction. Without a substantially more complex approach, the grid cell has to be treated as homogeneous for the purposes of horizontal transport.

2.5. New methods to evaluate and compare ecosystem models and biogeochemistry models

Generally when assessing skill we are asking: How well does a model represent truth over a specified range of conditions? However, because truth cannot be accurately measured, we use observations as a surrogate and ask instead: How well does a model fit the data? Both our model predictions and the observations reside in a halo of uncertainty; the true state of the system is assumed to be unknown, but lying within the observational uncertainty (Fig. 6a). A model starts to have skill when the observational and predictive uncertainty halos

---

**Fig. 6.** Schematic diagram of the relationships between, model prediction (P), observations (O) and the true state of the system (T). Both P and O are assumed to have a halo of uncertainty. Panel a) shows the case for a model with no skill and b) shows the case for the ideal model. Reproduced from Stow et al., 2009.
2.5.1. Metrics

An assessment of the confidence that can be placed on model results (known as model validation) must take into account the complex combination of model and observational uncertainties. Model errors derive from inaccuracies in the model structure itself, process descriptions, parameterisation, initialisation and forcing functions. Errors in observations arise from basic measurement error, inadequate sampling of a process (i.e. aliasing of small-scale signals to large scales due to under-sampling) or lack of replication in highly heterogeneous systems and issues of methodology. A crucial issue is balancing precision (how well the model fits each data point) with trend (i.e., how well it reproduces the observed seasonal cycles). For example, even when the trend is well reproduced small differences in the timing of an event can lead to large errors in precision. The choice of error statistic is crucial and a comprehensive validation process must consider several.

Our general approach to model-data comparison is to first extract the observations from the model (i.e., reconstructing the data set using the model). This model data set can subsequently be aggregated in space and/or time depending on the timescale of comparison required. For example, in short term forecasting (e.g., use in operational models) then the focus is on precision, which requires the rather unforgiving direct like-with-like comparison of model and data in space and time; this is appropriate when seeking to reproduce the short term, high frequency variability of the system. On the other hand, if the focus is on seasonal or longer timescales, then the model and data need to be matched to the timescales in question while the requirement for instantaneous agreement can be relaxed.

A variety of univariate and multivariate methods can be used to assess model skill (e.g., Stow et al 2009; Allen and Sommerfield 2009, Doney et al., 2009; Jolliff et al., 2009), the choice of which is dependent on the questions being asked and the data available for confronting the model.

2.5.2. Data

Data is essential to model skill assessment. The process not only requires data to be available (and ideally at an appropriate distribution in space and time), but also an understanding of what the data tells us about the system and what it does not. There are a number of issues to consider including:

What exactly has been measured and how does it relate to what has been modelled?

In the simplest cases, such as temperature or nitrate, the model variable maps exactly onto what has been measured. However biological variables are frequently much more complicated. The first issue is that of aggregation. With reference to bulk biomass functional-type models, most modelled planktonic variables describe the bulk properties of a size class and/or a particular biogeochemical function, which is quantified as a concentration of a chemical component (e.g., nitrogen or carbon). This has to be compared with observed data, which are often semi-qualitative, e.g. species type and abundance.

What are the assumptions and uncertainties of the observations?

This includes measurement errors, replicates, and the uncertainties in any assumptions made. This is particularly pertinent where the measurement itself is the result of a complex retrieval process (e.g. satellite chlorophyll).

How representative is the observation at the time and space scales of the model?

This is a crucial question particularly when considering biological variables which in reality often display considerable temporal and sub model grid scale variability and is important when considering short term forecast skill for biological variables.

2.5.3. Distributions of phytoplankton functional types

There is a severe shortage of data for the distributions of phytoplankton functional types at the global scale. Satellite earth observation provides extensive spatial and temporal coverage of the surface of the ocean, primarily for temperature, sea surface height and ocean colour. A number of new products including inherent optical properties (Garver and Siegel, 1997; Lee et al., 2002; Smyth et al., 2006), dominant phytoplankton type (e.g. Aiken et al., 2007; Alvain et al., 2005; Hirata et al., 2008a; Sathyendranath et al., 2004; Uitz et al. 2006) and phytoplankton type-specific primary production estimates (Uitz et al., 2008, Hirata et al., in press) are currently being developed and evaluated. Advancing the understanding and interpretation of satellite signals has enormous potential to relax the difficulties in the validation of models.

Of particular relevance is the recent development of a semi-analytic ocean colour inversion technique (Smyth et al., 2006) which is less dependent on empiricism than conventional satellite chlorophyll a biomass derivations. It derives optical properties such as backscattering by particles, absorption by phytoplankton and coloured dissolved organic matter (CDOM) at several wavelengths. These optical properties are more directly attributable to the in-water constituents (Zaneveld et al. 2006) than is the raw ocean light field (water-leaving radiances), as used, for example, by Alvain et al. (2005). Phytoplankton absorption, scattering and particle size distribution have been used to estimate the dominant PFT contributing to the remotely-sensed signal (Hirata et al., 2008a; Hirata et al., 2008b), the underlying rationale being that these optical properties represent phytoplankton pigment concentration, which is in turn related to the size of the cell and the taxonomic group, of which PFTs constitute a proxy designation (for a fuller discussion of PFT remote sensing see Nair et al. 2008).

Although these estimates are based on measurements of ocean colour and not a direct measurement of in situ abundance, two different approaches: phytoplankton absorption and particle scattering (albeit both derived from the same satellite data set) show a reasonably consistent result (Fig. 7a,b). Validation of the satellite-derived PFT estimates against in situ measurements has demonstrated a reasonably successful classification of PFTs: 26 satellite data points were matched up with in situ data on PFTs determined from High Performance Liquid Chromatography (HPLC)-derived pigment measurements taken during the Atlantic Meridional Transect 07 in 1998, using Diagnostic Pigment Analysis. The satellite and in situ PFTs coincided at 19 stations (Hirata et al., 2008a). This agreement allows validated satellite PFTs to be used as an alternative to “real” measurements, even though they are a product derived from satellite measurement of ocean colour. Without satellite data, validation of PFTs derived from biogeochemical models would be even less practicable and far more arduous, especially at the global scale. There is of course uncertainty associated with satellite-derived PFT estimates, as the principle of the derivation is to utilise coherent relationships between in situ PFT presence, diagnostic pigment distribution and optical properties. Such relationships may vary between different ecosystem types (or bio-optical ocean provinces) requiring that regular calibration of algorithms is carried out to ensure the robustness of the results.

3. Future directions

In order to develop ecosystem models which can enhance our knowledge of marine ecosystems well enough to equip us to address major societal issues (i.e., have predictive capability), we need to ask
the question where should future effort and investment be focused? Our experience has flagged several areas of equal and interlinked importance: model complexity, parameterisation of physiology and foodwebs, ocean and coastal physics, hydrodynamic model resolution, model diversity and experimental design (including reanalysis and ensembles).

3.1. How complex should a complex ecosystem model be?

This question remains the holy grail of marine ecosystem modelling, and although our recent experience has provided valuable insight, the answer remains a matter of opinion. As stated previously, a model is a cartoon of a system which hopefully describes the major modes of variability of the simulated system. A good model should describe the available data, represent the system’s dynamics and produce emergent properties (Noble 2003). Too much complexity leads to uncertainty and problems in interpretation of the model dynamics, while too little means that the models cannot reproduce realistic behaviour (Fulton et al., 2003). Assuming that accuracy of prediction is required, there tends to be a trade-off between complexity and generality in models (Levins, 1966). It is perhaps surprising that the most complex ecosystem models, incorporating multiple PFTs, are often being incorporated into global GCMs without thorough testing at the regional scale first (Anderson, 2005). Further, the type of ecosystem model employed should depend on the user objectives. Complex models can be superb as heuristic tools for gaining understanding by exploring and testing hypotheses. The under-determination problem must not however be underestimated. Models may in some cases be “unreasonably effective” at describing phenomena (it is all too easy to fit to data), such that if accurate prediction is the goal then “a veil of measured prudence must descend upon us” (Anderson, 2010).

Defining marine PFTs in terms of their biogeochemical role (calcifiers, DMS producers, and so on), downplays the intricacy of the interactions amongst the different classes of organism that often perform these biogeochemical functions. Very few systematic studies of the effects of complexity have been made. Friedrichs et al. (2007) assessed 12 lower trophic level models of varying complexity objectively in two distinct regions (equatorial Pacific and Arabian Sea) using variational adjoint assimilation of chlorophyll a, nitrate, export, and primary productivity. The same cost function metric was used to assess model skill. They demonstrated that when a single pelagic regime is considered, the simplest models fit the data as well as those with multiple phytoplankton functional groups. The under-determination problem remains ever present. Ward et al. (2010), for example, carried out one-dimensional modelling of the Arabian Sea and were unable to uniquely constrain even ten parameters. Cross-validation experiments revealed that as long as only a few key biogeochemical parameters were optimized, the models with greater phytoplankton complexity were generally more portable (i.e., better able to perform well in diverse regions and physical settings, (Friedrichs et al., 2007). Friedrichs et al. (2007) also found that models with multiple zooplankton compartments did not necessarily outperform models with single zooplankton compartments, even when zooplankton biomass data are assimilated.

Fig. 7. Dominant phytoplankton size classes derived from SeaWIFS for 1994 using inherent optical property approaches: a) absorption approach and b) scattering approach. Blue, green and red colours show the dominance of pico-plankton, nano-plankton and micro-plankton, respectively.
They also demonstrated that even when different models produced model-data misfits, they did so via different pathways. While the more complex models often seem to be more portable in terms of goodness of fit to data, it remains unclear whether such models are mechanistically more realistic.

While there is a great deal of ecosystem model diversity, it generally seems to be more by accident than design, and there is a danger that scientists all too often seek to develop models that simply reinforce theories that have become part of common working knowledge (Anderson, 2010). It is not yet known where the biology should converge. The majority of ocean biogeochemistry models are now based on PFTs, but with different complexity and representations. Since there is no consensus on how many PFTs are needed to represent which key processes, flexibility in approach is needed in order to select appropriate levels of complexity, depending on the question, geographical area, or research agenda. Our experience suggests the construction of model frameworks in which models of different complexity can be compared is to be encouraged.

Our results suggest that models based on PFTs are more responsive than NPZD model to the simulated environment in which they are placed and that they produce a more extensive range of emergent properties. However the question of how real these responses are remains and further data is required to resolve this issue. Robustness in parameterisation is essential if reliable prediction of, for example, climate change scenarios in earth system models is the goal. A follow-up question is how to define a PFT, by biogeochemical role or by ecological function, and how do we resolve the tensions when the two don’t neatly map onto each other. In practice, our choices tend to be (or at least should be) limited by our ability to parameterise the function in question.

3.1.1. Resolving foodweb structure and interactions

Ecosystem dynamics are profoundly influenced by the complex web of trophic interactions, which links all species together. Hannah et al. (in press) point out the importance of food webs in marine modelling. The structure of the foodweb and the strength of its interactions are critical to the stability and persistence of ecosystems (Dunne, 2006; McCann, 2000) and it must be remembered that the structure and dynamics are subject to feedbacks on each other. These aspects should be included in marine ecosystem models as they are crucial; however care is needed when using foodwebs to ensure that they are being usefully applied.

The fundamental problem is that the marine foodweb is poorly defined. Studies are often limited to a low species numbers coupled with a tendency to focus on higher trophic levels (Link, 2002). High levels of aggregation and the limited spatio-temporal extent of such studies compound these limitations. An understanding is required of how the network structure of the foodweb influences what happens when species are added or removed (May, 2006); generally speaking removing the most highly connected species has more impact than random removal. A further issue related to interaction strengths is capturing plasticity in the realised diets of organisms.

As a consequence, we lack a detailed understanding of the structure of some (if not many) parts of the marine ecosystem and we have even less understanding of their interactions, all of which are serious hurdles for modellers (Anderson, 2005). Ultimately, this lack of knowledge is the limiting factor in model construction. The first challenge is to get a better description of the structure and dynamics of the whole foodweb including plankton. Clearly, including all interactions at a species level is intractable so the second challenge is to devise how to aggregate the foodweb in such a way that it captures the important ecosystem dynamics and feedbacks. Such abstraction and simplification lie at the heart of the modelling endeavour—developing a conceptual model of the system of interest is where intuition and genius, as well as knowledge and experience, come to the fore (Anderson, 2010). Some guidelines exist (e.g., Fulton et al., 2003) but this is a work in progress and the subject of ongoing debate (e.g. Anderson, 2005).

3.2. Future challenges and directions for scaling from physiology to functional types

The sequencing of complete genomes of representative phytoplankton taxa (Misumi et al., 2008) has paved the way for application of transcriptomic, proteomic and metabolomic approaches to phytoplankton, providing unprecedented amounts of information on which to base truly mechanistic models of phytoplankton physiology. In the coming decade we will be required to distill from this superabundance of new data the principles for developing better models of phytoplankton growth that can be employed within ecological and biogeochemical contexts. The availability of these new data streams raises new challenges in bottom-up modelling from phytoplankton physiology to functional group representations. Given the constraint that ocean biogeochemical models are unlikely to become much more complex than the Monod and/or Droop equations in the near future, the problem for a bottom-up approach is deciding what to exclude rather than what to include. As an alternative to mechanistic models based on descriptions of metabolic and energetic pathways, it has been argued that physiology can be modelled as a trade-off in resource allocation amongst different metabolic functions with the goal of maximizing growth rate (Shuter 1979; Klausmeier et al. 2004; Pahlow 2005; Armstrong 2006). These formulations are based on a conceptual model that evolution results in organisms that are optimally adapted to their environment. It is known that, because of variability in the environment, there is not a single optimum state that every organism evolves towards (Armstrong 2006). However, application of optimality criteria has been shown to facilitate translating qualitative physiological information into quantitative models with realistic results (Shuter 1979; Klausmeier et al. 2004; Pahlow 2005; Armstrong 2006).

3.3. Model sensitivity to hydrodynamic and climatic forcing. Getting the physics right; processes, data assimilation, ensembles

The sensitivity of PFT models to the hydrodynamic environment is a double edged sword. On one hand the sensitivity of the model plankton distributions to changes in their environment gives hope that they may improve the representation of the known sensitivity to climate variability (Prentice et al., 2004). On the other hand, just how real these changes are and how much is induced by errors in the model physics in which they are embedded remains an unknown quantity. There is the perennial problem that biological model parameters may be, usually unwittingly, tuned to compensate for inadequate physics (Popova et al., 2006), a problem exacerbated in complex models by the under-determination of parameters by data. There are four main approaches that can be used to address these issues: improving the resolution and process descriptions of the hydrodynamic models; using data assimilation and inversions to constrain simulations with data; running ensemble simulations; and increasing the amount of data that is available for model representations. The first three approaches are computationally expensive, while the fourth one requires substantial researcher efforts.

3.3.1. Hydrodynamic model resolution

There are several ways to address the issue of model resolution and improve understanding of the mesoscale physics and the role of shelf seas in global-scale processes. The traditional response is to refine the horizontal and vertical scales of the physical models to the scale of interest. The problem with this lies in the computational cost associated with it, especially when coupling to complex PFT models which have large numbers of state variables that need to be advected and diffused. As a rule of thumb, once the number of biological
variables transported exceeds the number of physical variables then biological transport dominates the costs. Currently global-scale coupled hydrodynamic ecosystem models are rarely run at horizontal scales of less than 1°.

The most practical option to address coastal zone modelling given the current modelling technology and computer resources is the grid nesting approach. Nesting is a standard practice in downscaling from an ocean basin scale domain to a particular coastal region of interest. The Global Coastal Ocean Modelling System (GCOMS; Holt et al., 2008) provides a flexible framework to set up any number of regional models of the continental shelf over the globe taking lateral boundary conditions from a global OGCM to drive the POLCOMS-ERSEM modelling system (e.g. Allen et al., 2007). The framework enables multiple regional model configurations to be generated from user defined domain boundaries.

Perhaps the most versatile, efficient and elegant solution is to use an unstructured grid ocean model (e.g. Pain et al., 2005). Such models would not immediately address the issue of computational load, but rather distribute the effort more efficiently. The unstructured grid model of St Georges Bank by Chen et al. (2008) using FVCOM demonstrates the efficacy of this approach. If well-resolved coastal seas require a 10-fold reduction in grid-spacing and they occupy 8% of the global ocean area, they would require between 8 and 80 times the computational resource compared to modelling the rest of the world’s oceans. As noted above, simply refining the resolution of the OGCM is an option that would give some modest benefits with current computational resources. It should be borne in mind that the computational capability available to UK research has increased by a factor of ∼150 over the past 8 years (based on national facilities and not accounting for the growth in prevalence of local cluster facilities). There has not been a proportionate increase in either the resolution of the ocean component of earth system models or in the complexity of ecosystem models. This inertia lies in the difficulties in utilising this raw computer power (e.g. developing models that scale well on many thousands of processors) and hence requires close engagement with computer scientists to address this issue.

3.3.2. Reanalysis simulations

Data assimilation is a technique used in meteorology and ocean forecasting to produce the best simulation that the combination of model and available data allow. In particular, these kinds of simulation allow us to make the best possible model quantification of key biogeochemical processes (e.g., the carbon cycle over the periods which have suitable data). Reanalysis simulations can provide robust initial conditions for future climate scenarios.

The application of data assimilation has demonstrated the value of constraining the physical environment; improved descriptions of physical mixing lead to better estimates of the carbon cycle at a global scale. The next stage is to develop complementary assimilation techniques to constrain the biological models. Assimilation techniques for biological models already exist (e.g., Hemmings et al, 2008, Torres et al., 2006, Allen et al., 2002), however, no dominant systematic methodology has emerged and the application of such techniques at a global scale is in its infancy (Gregg 2008).

There are three main limitations to this approach: it is computationally expensive; the availability of suitable data sets at a global scale is currently limited to ocean colour, and, as data assimilation is limited to hindcasts, we cannot make simulations which project future ecosystem states.

3.3.3. Ensemble simulations

Here we refer to the process of running multiple models (‘ensembles’) to acquire a range of future predictions and hence an idea of the range of future ecosystem responses. Ideally the ensembles would be derived from multiple combinations of different biological and physical models which are then assessed to decide which outcomes are more likely (probable) than the others. This is important for example when assessing the sensitivity of PFT models to changes in hydrodynamic environments. Robust validation of hindcast simulations is required to underpin this activity and where necessary to help weight the ensembles. Crucial to this approach is model diversity.

3.3.4. Enlarging datasets for model parameterisation and evaluation

Widely used syntheses of physical and chemical ocean observations (e.g. World Ocean Atlas and Global Ocean Data Analysis Project) have led to an increased appreciation of the uses for data synthesis to define global biogeochemical rates (e.g. Behrenfeld and Falkowski 1997, Calbet and Landry 2004) and concentrations (e.g. Buitenhuis et al. 2006). These data syntheses are useful ways both to look back and make new use of historical observations, and to look forward to identify gaps in our knowledge and suggest the most effective strategies for filling them.

4. Where do we focus future investment?

Delivery of the next generation of ecosystem models requires a range of developments including a generic software framework for model development, improved hydrodynamic models, and better parameterisation of new and existing models, reanalysis tools and ensemble simulations. Simply adding ever more complexity to models may not be productive in terms of generating accurate predictions, given that there is an optimum that depends on the nature of the real-world system, feedbacks of interest and the knowledge base and data available to construct models and provide means of skill assessment. As Anderson (2010) proposes, modellers very probably need instead to get clever and adopt various new approaches that are coming to the fore such as targeting complexity in key species and trophic levels, adaptive parameterisations and the representation of physiological tradeoffs, providing the potential to simulate emergent community structure.

The idea of a model development framework is to take a single physical core and develop biological system modules that can be activated depending on the complexity required and the scientific question being explored. This would allow the same model to work at timescales from days to millennia, and at resolutions required for operational forecasting and for Earth system modelling, whilst maintaining the ability to trace between the hierarchies of complexity. This may be the best bet for getting an understanding of the role of model complexity. The major benefits of such an approach are ease of comparison of models, version control of the model code, reduction of effort in data assimilations and facilitation of the generation of ensembles. Alongside these, information is required to better parameterise model process descriptions, which requires focused resource. Ultimately it is up to the community to define which processes should be prioritised.

Improvement of hydrodynamic process models and increased model resolution is also required. Much of the required work is ongoing and will be delivered by the physical oceanographic community. Perhaps the most important issue is one of liaison with this community to ensure that the next generation of models are fit for purpose (Najjar et al. 2007) and to avoid duplication of effort. More efficient reanalysis schemes are required to make the best use of existing model systems and data particularly to constrain simulations of the recent past. Data is the lifeblood of environmental modelling, without which models can neither be parameterised, or have their structure validated nor their outputs verified. Underpinning all of the previous tasks is the requirement for systematic collection and storage of data sets and this requires communication between data providers and data users.

Finally the cultural disconnection between modellers, laboratory experimentalists and observationalists still remains a limiting factor in the further development of modelling approaches in spite of many...
attempts to address these over the last two decades (Flynn, 2005). In particular non-modellers commonly lack a sense of ownership of numerical models even though their ideas and data have contributed to the models’ development. The final challenge is to ensure that experimental/observational scientists are stakeholders in the models and vice versa.

Acknowledgements

MarQUEST was funded by the Natural Environment Research Council (NERC) as a part of the Quantifying and Understanding the Earth System (QUEST) program.

References


