Met Office/NOC Joint White Paper

NEMO Perspectives

1. Summary

We suggest two NEMO development strands to maintain a competitive European ocean modelling effort through to the mid-2020s and beyond. The first is the incremental evolution of the NEMO model, which could take us through to around 2020, and the second is a next generation NEMO modelling initiative that will allow full exploitation of emerging supercomputer architectures and benefit from a fully multi-scale approach.

2. Introduction

This White Paper draws on work carried out in the NERC funded *Next Generation Ocean Dynamical Core Roadmap Project* (PIs Holt and New) that will be reporting shortly. The project investigates how meeting the ocean model development drivers in a changing computational landscape might best be realised over the coming decades. Workshops with computational scientists, ocean model developers and ocean model users were held to discuss the requirements for future ocean models and a potential pathway to deliver on those requirements. The analysis suggests that the current NEMO code base has a finite competitive life time, and over the course of this review period (2015-2025) would be expected to 'lose its edge' and become increasing inefficient and problematic to use, largely due to expected changes in computer architectures. The key question is then whether this existing code base can be re-factored for these future computer architectures or whether a new approach is needed.

3. Drivers for Change

Given the premise (discussed below) that to remain usable ocean models will have to go through significant redesign, it seems pertinent also to re-examine the overarching scientific drivers for developing physical ocean models. To achieve a representation of the ocean that is as accurate as possible when compared to observations and correctly represents the underlying mechanisms and processes, we need:

1. To improve the representation of diapycnal mixing, by reducing numerical mixing and more accurately modelling physical mixing.
2. To accurately model the mesoscale and effectively parameterise the submesoscale in open ocean simulations.
3. To improve the representation of coastal and shelf seas, and of key small-scale topographic features (e.g.narrow sills) in basin and global scale models, through better process representation, resolution and parameterization, involving multi-scale modeling approaches.

Alongside these, properties for effective ocean model design need to be considered to ensure that future codes will deliver on both the computational and scientific drivers. A future ocean model should:

- Have good discrete dissipation properties: i.e. minimal numerical diffusion;
- Give conservative solutions for mass, tracer, energy, momentum, PV and perhaps Enstrophy
- Have good discrete wave dispersion properties: e.g.. numerical modes are controlled
- Be computationally efficient (for large and small models)
- Fit coastline at least as well as quadrilaterals, ideally as well as triangles
- Have flexible vertical coordinates/methods (including ALE)
- Have multi-scale capability and geometric flexibility
- Be accurate in 'realistic' and idealised test cases
- Be portable, adaptable and usable as a community model
- Be supplied with a set of tools that allows efficient handling of the increasingly large amounts of model output, including facilities for runtime calculation of derived statistics such as time and spatial means.

The present NEMO model based upon the OPA code, as a C-grid finite difference model using quadrilateral elements, does well against many of these criteria. However, in common with all major ocean models the code design does not lend itself well to future computing architectures, being based largely on a coarse-grained parallelism.

4. The computational landscape in the UK over the next 5-20 years

Central to the UK NEMO partners' view on the future perspectives of NEMO is the high performance computing environment on which NOC and Met Office implementations will be run. Extrapolating the trend from known architectures to the future gives peak performance of ~100Pflop/s by 2019 and ~5000Pflop/s by 2023. This closely follows Moore's Law / TOP500 trends, and predicts the UK maintains a capacity about a factor of ten lower than the US at any one time (or lags by 3-4 years). There are of course many unknowns in this prediction including the UK Governments continued commitment to HPC, and the share of the resource that the UK marine community may receive, but there is presently no indication that these will falter.

The current expectation is that this increase in computer power will be achieved through a substantial increase in core count per physical chip and memory amount/band width will not keep up with this increase; hence memory per core is expected to steadily decrease. For example, memory per core has decreased from 3GB to 1GB over the life of the UK's main academic HPC facility HECTOR. Of course there may be radical changes in computer architectures that overcome the present power barrier, but we cannot predict these. Hence, it is prudent to assume the current trends will persist. The present approach of one MPI task per core will soon become impractical as there will not be the memory available to support an optimal fraction of the model grid on each core on a chip[1]. Hence, alternative approaches will be needed and mixed MPI, openMP is a very attractive option. This, however, comes at the cost of increasingly complex and difficult-to-maintain code that may put unreasonable expectations on the physical oceanographer (who is not a software engineer) in developing the code. When this point will occur with NEMO is difficult to predict as this model is primarily limited by memory bandwidth rather than amount. However, we can expect difficulties over the course of the next UK HPC refresh (in the next few years), and certainly by the following one (i.e. by 2019); this is likely to manifest itself as a gradual decline in efficiency rather than a sudden shutdown. The implications of doing nothing will be that to run a given NEMO configuration, we will have to increasingly under-utilise the available cores to gain memory bandwidth. This means NEMO will become increasingly expensive to run over time in terms of wallclock hours per gridpoint per core occupied. If the cost (€) per physical chip (with some size of memory) remains constant, then we are under utilising a potentially growing capability and so losing competitiveness. If the cost increases (as is likely) then we go backwards[2].

5. Target configurations

It is helpful to consider a range of possible model configurations that would meet the user requirements over the period and when these become practical as 'routine' physical models (i.e. ocean components of IPCC input and operational models) or routine for ensemble runs and with embedded ecosystems. Costs can be estimated on the basis of number of grid cells, timestep and a penalty for using unstructured meshes (taken to be 5). It is clear that treatment of time stepping is the major consideration as resolutions are refined, particularly in the near coastal zone. Beyond 1/12° global and 1/60° shelf sea resolutions, unstructured grid approaches become increasingly attractive. This is particularly the case given that these estimates suggest it would be three computer refreshes before a 1/36° global model would be considered routine; a very challenging proposition for NEMO without a complete rewrite. We therefore also consider the possibility of a multi-scale approach to allow resolution to be targeted at key regions, e.g. small-scale topographic constraints and regions of submesoscale-mean flow interaction. Hence, by the time the current generation of advanced models have become routine, computer architectures will have developed to the extent that the next generation of advanced model will require a radical change before they too can become routine.

---

[1] For example for the current HadGEM3 atmosphere-ocean physics model at N96 atmosphere (120km) and ORCA025 running on the Met Office IBM Power7 system, and load balanced between atmosphere, ocean and sea ice, about half of the available memory is used for the ocean nodes. Assuming we wish to increase number of gridpoints linearly with number of PEs, a further reduction of a factor of 2 in the memory per PE would lead to memory becoming the limiting factor in ocean parallelisation. At this point the number of PEs allocated to the ocean would have to increase, leading to load balancing and scalability becoming limitations on overall coupled model performance.

[2] Say currently a chip has 32 cores and 32GB of memory and in 2019 it has 1000 cores and 32GB. If we can still only run of 32 of these and the 2019 chip costs the same as the current one we have gained no benefit from the extra cores. But if it costs more e.g. due to manufacturing costs then we will achieve worse performance per € in 2019 than in 2013.

Grid	Size (k cells)	Relative Cost (time step)	Relative Cost (no time step)	Timescale for routine use
Global Scale				
1/12	8150	27.0	9.0	2016
1/36	73350	973.7	108.2	2023
1/12+1/36 (multi-scale)	8700	647	72	2021
Shelf scale (NWS)				
1/60	1776	15.7	1.3	2015
1/120	7104	125.7	5.2	2023
Shelf to Coastal				
~1km - 200m (US)	5030	1963.8	18.5	2019

Table 1: A representative range of grids expected to be similar to those in widespread use for UK ocean modelling over the coming decades. Costs are relative to ORCA025, with assumptions that the timestep either does not change, or changes linearly with the grid length. Timescales are for first active operational use (see use cases below). The global multiscale option is based on refining the ORCA083 grid to resolve topography (down to 3.1km), and includes a factor 5 penalty for an unstructured grid approach. See section 9 for further discussion.

6. Gung Ho

Within the UK, and elsewhere, there are efforts to design and build new geophysical fluid dynamics codes that are designed from the outset to be computationally efficient on the next generation of computers, whilst remaining relatively simple to use and develop. The Gung Ho project has been developing such a framework for the next generation atmospheric model to be used by the UK academic and operational NWP/climate prediction communities. To date the project has gone a significant way towards investigating options for the numerics (having investigated grid designs and numerical schemes suitable for atmospheric modelling) and the computational environment within which that should sit. Presently the solution strategy is likely to be based on some variant of the finite element method. Although the Gung Ho work is focused upon the needs of the atmospheric modelling community, and therefore the applicability to ocean modelling needs to be investigated, there is potential for components of the Gung Ho modelling solution to be used as the core of a new ocean model. The Gung Ho project proposes to develop code in separate components (Figure 1): a computational science or ‘driver’ component, which is likely to be derived from freely available generic software tools, and the model components which solve the model equations. The philosophy behind this structure is that the scientific equations (which are coded as algorithms in the Algorithmic layer and computed using reusable local kernel operators in the Kernel layer) are expected to be coded primarily by ocean/atmospheric scientists, and are kept separate from the computational infrastructure which is the domain of the computational scientists. The blue vertical line in the schematic (Figure 1) shows the separation between the Parallel System (PSy), where parallelization, communications and computational tasks are coded, and the Algorithmic (Alg) and kernel components. The algorithmic layer is primarily where the model specific equations will exist, although it is expected that some aspects of the kernels may also have to be updated with changes to the scientific equations.

The separation of the computational science layer from the natural science layers is the crucial point: it isolates the natural scientist as far as possible from the complexities needed to ensure efficient scalability. The expectation is that this computational science layer will provide a set of tools and approaches, many of which will be easy to adapt for ocean as well as atmospheric modelling, while a number of the kernel operators will also be applicable in an ocean model. The degree of flexibility available for the design of the ocean model has yet to be established (i.e. how far it would be able to deviate from the atmospheric approach). The framework is likely to be able to accommodate differing element shapes and types, and vertical grid approaches, but differing solution strategies (e.g. choosing a Finite Volume rather than a Finite Element approach) are likely to be more problematic and require more ocean specific development effort, or indeed may not be practical.

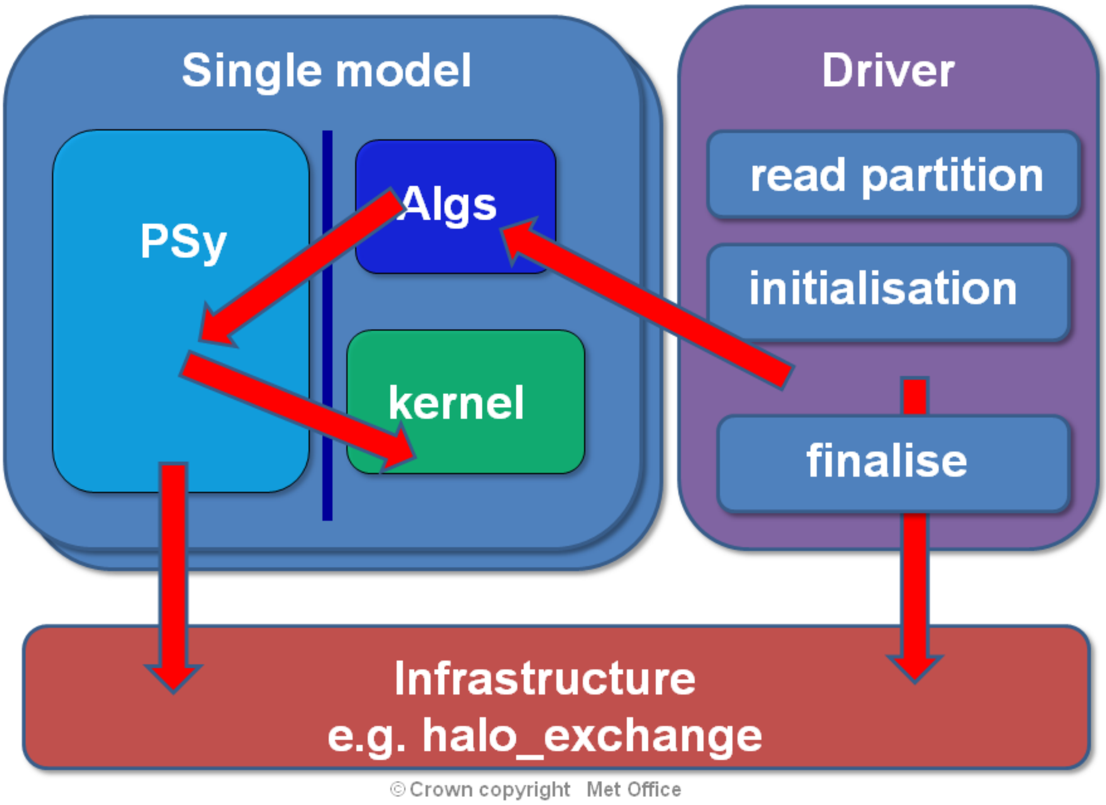

Figure 1: A schematic of the coding framework for Gung Ho.

7. Proposed way forward

The Ocean Roadmap report (soon to be published) describes two complementary ways forward for ocean modelling in the UK. The time over which the current NEMO code becomes increasingly inefficient on evolving computer architectures depends on how much effort is spent on incrementally developing this code base to meet these challenges. However, there will come a point where the recoding becomes so intensive and extensive that the activity really constitutes a complete rewrite of the code. Moreover, incrementally optimising the code, without radically restructuring it (for example to separate science and computer-science layers) will seriously hamper the ability of the community to maintain the code, its usefulness and its 'developability'. The serious challenge lies in meeting the optimisation requirements and maintaining a code base usable by its current broad community. We do not believe this is achievable with the current NEMO code after about 2019. There are essentially two options on the table to address this:

1. A science neutral rewrite of the NEMO code, eg using a mixed MPI/ openMP coding approach to address the issue of computational efficiency. In addition, developments to the physics included in NEMO would be needed, as outlined in the "Target Configurations" section below.

2. Moving to a new modelling approach, as described below.

Option (2) could be conducted either using the Gung Ho approach (scoped below) or an alternative coming out of the NEMO consortium. The GungHo approach is proposed below because of its potential to leverage value from this major existing effort.

It is important to make the distinction between the current NEMO *code*, which would be replaced or significantly updated in either option (1) or option (2), and the NEMO *consortium*, which is expected to remain in place irrespective of decisions on how the NEMO code develops.

The need for a new approach for ocean modelling is clear if we are to both maintain competitiveness on future computer architectures, and exploit the potential of multi-scale modelling. **To this end we propose aligning a NEMO ocean model development effort with the Gung Ho project, and propose the resulting model forms the new dynamical core for use by the NEMO community**. An ocean model core (referred to as NEMO-G below) based on Gung Ho computational science layers (and where practical dynamical layers as well) presents an opportunity to move the NEMO consortium to a second phase, building on the strengths of the consortium, and allowing it to focus on key ocean modelling issues while taking advantage of more generic developments in geophysical model software engineering.

8. Development of the NEMO model

The timescales required for the various stages outlined below are informed by assumptions on the changing high performance computing environment and estimates of the key model types/resolutions that will need to be computationally efficient on that architecture.

8.1 OPA code base

The OPA code base will need to be kept competitive until a replacement option is available. Although a complete code rewrite is not considered an effective approach, there will need to be effort put in to optimisation of NEMO using the OPA dynamical core to maintain the code's utility prior to its replacement, and to high priority scientific developments for immediate requirements.

If a proof of concept for the NEMO-G approach for NEMO is approved by the NEMO consortium as the preferred development pathway (see below), then the OPA dynamical core and the NEMO science modules associated with it would cease to be developed, and the NEMO systems team effort would gradually move to developing the NEMO-G based code.

8.2 NEMO-G code base

Over the next few years (2013-2015?) a parallel activity would need to be spun up to examine the feasibility of the NEMO-G approach, by building prototypes and idealised models. This activity would include contributions from the current NEMO consortium but also draw in external expertise as appropriate. It would need to define decision points/gateways to move onto greater commitment to the NEMO-G approach. At some point if the gateways are passed (to be approved by the NEMO Steering Committee), the system team would transition to a focus on NEMO-G development. It would probably be necessary to expand the systems team to include expertise in new approaches. This could occur by expanding the NEMO consortium to include interested groups in Europe with a proven track record in the chosen modelling approaches.

The target of an operational NEMO-G model by the early 2020s is considered challenging but achievable if there is a strong commitment to the approach across the NEMO consortium.

If the NEMO-G framework based upon the Gung Ho code and methodologies is proven not to be effective then the NEMO community will need to discuss fallback options. Other computational modelling codes may be available to replace the Gung Ho framework, although there are no obvious candidates at the time of writing. If none become available, then a science neutral rewrite of NEMO/OPA will be required.

9. UK Applications of the NEMO model

This section addresses a UK forward look for specific applications of the NEMO model, and the requirements for appropriate developments of NEMO.

9.1 Global short-range and seasonal forecasting: NEMO based component of a coupled forecast system

Short-range and seasonal forecasting systems are expected to remain using the $1/4^{o}$ resolution NEMO configuration for the coming few years for both ocean-wave-ice-atmosphere coupled deterministic and ensemble prediction using NEMO-CICE-WWIII-UM.

The major science challenges for these systems are presently in developing coupled initialisation and the representation of air-sea exchange for the benefit of NWP timescales. Representing the mesoscale, sub-mesoscale and vertical mixing processes is therefore the primary focus for scientific development of the ocean components of short-range forecasting. The lack of explicit representation of the mesoscale is a limitation both for the evolution of the instantaneous fields and for the mean state. It is expected that as the computational resource and scientific robustness of eddy resolving models allow, then these systems will become higher resolution. By 2016 a deterministic $1/12^{o}$ coupled system should be envisaged. It is likely that this will become part of an ensemble system over the following years.

A $1/36^{o}$ is unlikely to be used in the near future as the problems of computational cost and of manipulating the data are prohibitive. However, within ten years one could expect these problems to begin to be resolved, and some short-range forecasting applications may be able to take advantage of these sorts of resolution. Multi scale unstructured grid approaches may be expected to be coming into effect around this stage, and the use of NEMO based upon the OPA code base is likely to limit the use of this model. The priority of moving to 1/36º would need to be tensioned against the value of investing the same computational resource in other model developments, e.g improved atmospheric resolution; at present little evidence is available to inform such a decision.

 # Global Ocean :

The resolution of the ocean for short-range ocean and seasonal forecasting is presently in the “grey” area of eddy permitting, and there is considerable evidence to suggest moving to an eddy resolving ocean will be of benefit to this use case. Initialisation of the mesoscale, and subsequent retention of this scale within short range model forecasts (for applications such as ship-routing) is presently limited by the resolution of both the ocean models and observations. Resolving geographic detail and eddies are important in the transport of heat horizontally (e.g. Gulf Stream separation) and in restratification and thus play a role in determining the mean state over seasonal timescales. It is not clear to what extent explicit representation of eddy processes in the ORCA025 means parameterised representation is not required, and thus it is expected that the role of eddies is not optimally included at this resolution. Coupled feedbacks may amplify this short-coming in the seasonal forecasting systems. It is thus expected that the short-range and seasonal forecasting systems would benefit from resolutions up to the $1/36^{o}$.

The NEMO or a replacement ocean component will need to develop improved methods for resolving air-sea exchange and the vertical mixing/restratification processes that affect the near-surface tracer and current fields. It is expected that over the coming years improved vertical mixing schemes will become available and their use will be a priority for short-range and seasonal forecasting systems.

 # Coupler:

The coupler used in present operational systems is the OASIS3 system. This is being transitioned to the OASIS-MCT which will be the coupler used for the foreseeable future. There is an ongoing discussion on the use of a dedicated coupler versus direct interfacing of models. The former gives flexibility, but adds complexity and potentially some computational overhead; the latter is potentially more efficient but becomes more challenging as the number of components increases, especially if those components are developed by multiple groups. Given that the systems developed for the forecasting systems will share components and development with the Earth System Modelling systems (which will have more components and a stronger driver for interoperability of different component models) the preference is to use a dedicated coupler, despite the fact that for the forecasting systems a direct interfacing system may be more efficient. The efficient design of such flexible coupling systems requires international discussion and cooperation over the coming years; it is not considered further here but may influence the design of the next generation ocean model.

 # Atmosphere model:

The atmosphere model used primarily within the UK will be the Met Office Unified Model until around 2020, when the transition to a Gung Ho based model will begin. It is expected that by the time this happens all global short-range and seasonal forecasting applications will be fully coupled to the atmosphere. The present atmosphere resolution for the seasonal forecasting system is the N216 (~60km) UM model, but experiments for short-range forecasting with resolutions of up to N768 (~16km) are planned. The expectation is that within the medium-term coupled systems will be trialled with N1024 (~10km) atmospheres.

For coupled systems the expectation is that the ocean and atmosphere should be of similar resolutions this would suggest that ORCA12-N1024 would be the obvious optimal resolution using an ORCA12 ocean, although assessment of the impact of resolution and the relative gains from increased atmosphere versus ocean resolution will be looked at and will determine the phasing of the respective resolution increases.

Global convective scale models (~1-2km) of the atmosphere are not expected to be coupled to the ocean for operational forecasting over the next decade or so. However, regional convective scale coupled systems are already being implemented for research purposes in regional models, and are expected to become operational within the next few years. It is therefore likely that in the next decade or so global, coupled convective-scale systems will be being developed for research purposes.

 # Global Biogeochemical model:

Biogeochemistry is not yet used in operational short-range forecasting or seasonal forecasting systems. Pre-operational systems have been developed for short-range forecasting, and there is reason to believe that the lack of resolution of eddies in the physics may be one of the limiting factors in the skill of biogeochemical modelling systems.

The major requirement for a biogeochemical component to the short-range forecasting/seasonal forecasting systems is as part of a reanalysis system and any future forecasting system will need to have the ability to include a relatively complex pelagic ecosystem model, preferably with the option to run in-line at a reduced resolution spatially and temporally, and to feed back through light penetration to the NEMO model.

 # Sea-ice model:

The sea-ice model presently used is the CICE model developed at Los Alamos, and the forecasting systems will evolve following the strategies developed at Los Alamos. With the move to increasingly high resolution models assumptions made by the sea ice models on the rheology start to break down so the ice model used will be expected to develop appropriately. Small scale interactions will become increasingly important at higher resolutions and it is expected that coupling with wave effects will become included over the time period being discussed.

CICE is presently being developed for use within MPAS, an unstructured grid model of both the ocean and atmosphere, and there is already a basic EVP code running on an unstructured grid.

Developments to the physics parameterisations can be expected, with updates to the snow model being expected soon. There will be a range of rheologies developed. Icebergs and biogeochemistry will also be included within the CICE model. The CICE code will be optimised over the coming years for use with GPU technologies.

 # Wave model:

There is a project on-going within the Met Office to couple WaveWatch III to NEMO-CICE and the UM for short range forecasting. WWIII is increasingly a framework for wave modelling rather than a model in itself, and is being adapted to for use at increasingly high resolutions.

The effect of waves on the surface exchange, and its potential impact on the near-surface atmospheric forecasts particularly, lead us to expect the inclusion of waves within the coupled forecasting system over the coming years. The wave model is closely coupled to the atmosphere model and so its resolution is expected to be the same as the atmosphere model used.

 # Ice sheets:

There is an ice cavity model being developed at the British Antarctic Survey that is expected to become incorporated within the ORCA025 model with updated bathymetry and grid over the coming 18 months to 2 years. The inclusion of ice sheets requires the adaptation of present global grids; this work is planned within the next year.

 # Interface with drifting ice:

Lagrangian drift of icebergs for shiprouting is expected to be of increasing interest with the opening of new shipping routes with receding ice sheets. An ability to couple to a lagrangian model of icebergs, with parameterised melting and drift and initialised from observations is of obvious interest.

 # Data assimilation and ensembles:

Data assimilation and ensembles provide much of the skill in short-range and seasonal prediction. Within the UK the present systems are transitioning from the OCNASM data assimilation to a 3DVAR approach using NEMOVAR. This is expected to move to a 4DVAR approach in the coming decade(s), although the cost implications are significant. At present the seasonal-forecasting system uses a lagged ensemble to reduce the cost of providing probabilistic information, but in time this may move towards using some form of ensemble data assimilation approach. Consideration is being given to the use of a hybrid 4DVAR/ensemble approach within the Met Office Numerical Weather Prediction community, and there is the potential for this to be used as part of the ocean data assimilation. The evolving computational constraints will significantly affect the feasibility of the data assimilation algorithms and are likely to affect both the design and implementation of the data assimilation code.

An area of active research is in the coupled data assimilation problem. The present plan is to use a weakly coupled data assimilation approach, whereby the ocean and atmosphere data assimilation schemes perform independently although receiving information from the evolving atmosphere/wave/ocean/ice models respectively. A fully coupled data assimilation is not a tractable problem at present, although would be desirable and it is expected that research in coupled initialisation will continue to look at this problem over the coming years and decades.

9.2 Climate: NEMO based component of an Earth System Model

In the UK, Earth System and Climate modelling has so far largely been undertaken with the ORCA1 configuration of NEMO. However, NERC and the Met Office, under a new strategic partnership (the Joint Weather and Climate Research Programme) are now building a combined UK Earth System Modelling (ESM) group that will make decisions on the next version of the UK ESM. There is an expectation that in the near-future the ESM will transition to using the ORCA025 model presently used in short-range and seasonal forecasting at the Met Office. It is unlikely that ESMs will, on the timescales discussed here, be cost-effective at the ORCA12 or higher resolutions. It is therefore likely that for the next decade long coupled simulations will be predominantly done using the ORCA025 model with varying atmosphere resolutions and Earth System components included. However, the need to maintain efficient use of ORCA1 (or even ORCA2) will persist for (eg.) large ensembles, complex ecosystem simulations (e.g. more complex biogeochemical models as the global models begin to resolve shelf sea scales, and evolutionary/ stochastic approaches) and millennial scale simulations.

The most important issues which NEMO faces in terms of its suitability for earth system/ climate modeling are those concerning ocean mixing, which affects water mass properties (and hence density gradients and circulation patterns) on decadal to centennial timescales.

First, there is the general need to reduce spurious numerical mixing to the lowest possible levels. Such mixing can dominate the explicitly applied mixing in large regions of some ocean simulations (for example in the Southern Ocean and regions of steep isopycnal gradient), making the model simulations unrealistic. Spurious mixing is also generated if the schemes applied to dissipate the turbulent enstrophy cascade operate on level, rather than ispoycnic surfaces. NEMO is already developing a new z-tilde vertical coordinate which is isopycnal to high frequency motions, and this will help reduce the spurious mixing from, for example, inertial and internal gravity waves. However, it is recommended that NEMO should move as quickly as possible to a fully generalized vertical coordinate using the ALE (Arbitrary Lagrangian Eulerian) framework. This will give NEMO additional functionality in reducing spurious numerical mixing, and will enable it to keep up with planned developments for other leading models such as MOM. In addition, efforts should also be targeted at trials of new advection schemes (such as the Prather and PPM schemes), to further reduce such mixing.

Secondly, NEMO should continue to strive to incorporate realistic representations of physical processes which actually do mix the water column in nature. This would include for example, mixing by internal waves radiated from deep topography and localized shelf sources, inertial waves, Langmuir turbulence, shear spiking at the base of the mixed layer, and restratification by sub-mesoscale processes (building on the existing Fox-Kemper scheme).

Finally, NEMO should develop a better representation of the mixing of deep dense outflows such as those flowing southwards between Greenland and Scotland. This falls between the two areas outlined above. These waters typically mix too quickly (and spuriously) in NEMO, and in the case of the outflows between Greenland and Scotland, give rise to a North Atlantic Deep Water mass (critical for the overturning of the Atlantic circulation) that is too light and too high in the water column. This could be addressed by either the addition of a stream-tube or other bottom boundary layer scheme to NEMO, by a straightforward "plumbing" fix (to force the overflows down to their correct depths), or by using sigma-(terrain-following) coordinates. This latter method has been shown to help with the downslope flows of such water masses, but the resultant state may be (to some extent at least) an artifact of the depth to which the sigma-layers are taken. This aspect would also need a better understanding of the real-world processes which actually do mix these overflows, which in turn would require interaction with observationalists.

To some extent NEMO is already moving in these directions. We therefore support and encourage its enhanced continuation along these lines, in particular with the development of an ALE coordinate system.

 # Coupling to other model components

The generic approach in the UK is to develop coupled systems which are appropriate for all timescales. The coupling of NEMO in the Earth System context to other component models is therefore largely covered in the preceeding section (eg. for the atmosphere, sea-ice, coupler and data assimilation sections). In addition, it is worth commenting that since the ESMs are envisaged to be run with ¼° resolution oceans for centennial timescales, the choice of complexity for the marine biogeochemistry modules will need careful consideration. One possibility to assist with the implementation of more complex biogeochemical models, however, would be to use the NEMO facility for down-grading the resolution of the physical model for input into the biogeochemical model. In addition, the Earth System models will include atmospheric chemistry and aerosols, and models of ice sheets and ice cavities, which will interface with the NEMO ocean model.

9.3 Shelf seas and coastal forecasting, climate impacts and environmental assessment: NEMO based component of a forecasting system

Shelf sea modelling is characterised by the need for many domains, which often need to change depending on the region and/or topic of interest. Rapid and straightforward reconfiguration capability is therefore important. The area of much of the interest for NOC and Met Office is the North-West European continental shelf, which is a tidally dominated, wide-and-shallow shelf at mid-latitudes, with strong density differences at the interface between riverine and marine waters. To resolve the mesoscale (fronts, filaments, jets, eddies) in these regions grid resolutions of the order 1 km are required. A major limitation for these shelf models is in resolution of the coastline and bathymetry (e.g. the shelf slope region).

The numerics for this region need to deal with these O(1) changes in water depth, large tidal amplitudes, and strong density gradients. Bottom friction, and its relationship with the surface wave regime is of increasing importance as water depths get shallower and the bathymetry becomes increasingly dominant. The bathymetry is not static; however, predictive capability in morphological evolution has so far been elusive.

Distinct from open-shelf, coastal is loosely defined here as regions close to coastlines where resolutions of the order 100's of metres are required. Coastal applications often require engineering type solutions, with the influence of human activity often being a considerable factor (for example through dredging, shipping, ports and harbours activities and river flow management). Additionally the morphodynamics become important, with shifting bathymetry with storms or dredging activity. However, off-shore renewable energy installations, fishing, carbon capture and storage projects, and pollution dispersal extends direct human influence far off-shore. This along with the need for more extensive environmental assessment and prediction driven by policy objectives of (e.g.) the Marine Strategy Framework Directive and Marine Conservations Zones, suggests the traditional distinction between shelf and coastal models should be blurred with a unified approach.

The present workhorse model for shelf seas forecasting, climate impacts work and NERC Research Programmes is at 7km, but work is underway (e.g. in the NERC FASTNet Project) to develop a 1/60^{o} (~2km) model in the region, which should be usable for these applications within the next couple of years. This will improve the ability to resolve many important processes on the shelf, although for downscaling to coastal applications and for improved representation of shelf slope processes further refinement to ~1km would be desirable The present generation of 7km models are likely to be in use for a number of years for ensemble forecasting, climate impacts and increasingly complex ecosystem models.

With respect to the incremental development of NEMO, the lack of fine resolution shelf sea (1-2km) and near coastal (50-200m) capability and application is an on-going deficiency. This needs to be addressed for many scientific and stakeholder applications particularly when long integrations and/or ecosystem models are needed, making the use of unstructured mesh models less practical. In the NERC community POLCOMS and FVCOM are favoured for the near coastal activities, and it would make sense to encourage a transition of the POLCOMS work to NEMO (as capabilities are developed), to complement on-going FVCOM work.

There are activities to develop coastal prediction systems, and at the Met Office a convective scale (1.5km) atmosphere model coupled to a NEMO is presently being trialled to investigate the impact of coupling on atmospheric forecasts. This fits with an increasing drive towards integrated coastal zone modelling with an expectation that over the coming years a UK wide system will be developed that will allow the forecasting of flooding events. It is envisaged that this system will combine fluvial, pluvial and marine (surge, tide and wave) components and potentially inundation modelling to give a flood risk forecast. This will require the development of a system that downscales the shelf-wide tides, surges and waves up estuaries to scales of the order of hundreds of metres. Such a system will need evolving coastlines and, potentially, bathymetry. The ability to fit complex coastlines will be critical.

The physics of the ocean model will need to be appropriate for an environment with strong density gradients and a smaller aspect ratio. An unstructured grid shelf to coast model would be the ideal option for this.

An unstructured grid approach has traditionally been favoured in coastal applications as it allows flexibility in defining the coastline and following the bathymetry, and allows the model to increase in resolution close to the coast whilst maintaining coarser resolution in deeper waters. With current configurations/capabilities, the coastal model would nest within the shelf-models and at its outer edges will have resolutions of the same order as these (~1km). Near the coastline itself the (triangular) unstructured grid would be matched to the coastline with resolutions of the order of 100m. An alternative approach would be a seamless shelf to coast model. This is currently achievable with a model such as FVCOM, but the numerics of that model, particularly at the ocean-shelf interface are open to question.

 # Ocean :

A shelf-coastal model will need to be developed that incorporates the specific physical phenomena present in these regions.

A wetting/drying capability is under development, but this will require extensive testing and will need to adapt to any changes in time stepping approaches.

Crucial for this type of work is an easily relocatable capability, as often new domains are required for each application/project. This needs to include:

- Scripts for domain generation and boundary condition extraction using standard sources of forcing information (e.g. the MyOcean configurations)
- Shelf/coastal standard test cases

Other important aspects of physics are

- non-diffusive advection schemes (noting that physical horizontal eddy diffusivity can be very small in this context)
- Appropriate momentum advection and coastal boundary conditions to ensure accurate flow around headlands
- Appropriate scale dependent sub gridscale parameterisations
- Coupling to surface wave models – for this we need a generic 2-way interface between NEMO and wave modesl (similar to TOP)
- New open boundary approaches, particularly passive baroclinic boundary conditions

NEMO is a hydrostatic model which is a limitation for these regions, where hydraulic jumps and bores (e.g. Bristol Channel) and internal solitons (e.g. in Sea Lochs) can be significant physical features. Although the aspect ratios of these environments are closer to unity than we traditionally consider appropriate for NEMO it is anticipated that scale separation of the horizontal and vertical processes is still appropriate and a full CFD approach will not be needed.

 # Coupler:

A coastal application would be expected to be coupled with a number of other models, certainly including hydrological, inundation, atmosphere and wave models and potentially biogeochemistry and morphodynamic models. As with the global models above, the expectation is that OASIS will be the coupling tool used into the foreseeable future.

Atmosphere model:

The local meteorology is complex, with local effects due to the land-sea interface being important; for example land and sea-breezes having significant local effects on circulation and mixing. Therefore a fully coupled system would be the expectation. An atmosphere model at around the same resolution as the ocean will be used in either a fully coupled or one way coupled context.

Morphodynamic model:

The interaction of currents and waves with the seabed and their role in eroding, transporting and depositing material through the system will be important in determining the quality of coastal simulations. It is expected a morphodynamic model will in time become part of a coastal forecasting system.

Ecosystem and Biogeochemical model:

These are needed to meet many of the user applications for shelf and near coastal modelling. Simulating the ecosystem in these regions requires more complex solutions owing to non-Redfield stoichiometry and often new components of ecosystem need to be modelled specifically (e.g. the phytoplankton species associated with Harmful Algal Blooms). ERSEM is the shelf sea model of choice in the UK, being used in NOC and Met Office (e.g. in NERC-Defra shelf seas biogeochemistry research programme). ERSEM is currently deployed in shelf sea applications of NEMO and coastal applications of POLCOMS and FVCOM. It is envisaged that ERSEM would be deployed rapidly in a coastal application of NEMO.

Wave model:

As for the global model, coastal forecasting systems will use the WWIII system, which will be adapted to incorporate the appropriate physics for the coastal domain. WWIII has an unstructured grid capability, and will be able to be run at the resolutions and in the environments required for coastal modelling.

The interaction of the waves, ocean currents and the frictional effects of the seabed mean that the ocean and wave models at these scales are strongly interdependent. A fully coupled wave-ocean system is therefore essential to properly forecast coastal regions.

Data assimilation and ensembles:

Data assimilation and ensemble prediction will be a central part of any coastal forecasting system. It is likely that the data assimilation approach will follow closely the strategy described above in the global forecasting use case.