Assimilation of chlorophyll data into FOAM-HadOCC, a coupled ocean physical and biological model

Daniel Lea[1], Rosa Barciela[1], Karen Edwards[1], David Ford[1] and Matthew Martin[1]
[1]Met Office, UK, daniel.lea@metoffice.gov.uk

1. Introduction

We implement a chlorophyll assimilation scheme using processed ocean colour data provided by Globcolour. The data is assimilated into a coupled physical and biological model based on the FOAM (Forecast Ocean Assimilation Model) using the NEMO (Nucleus for European Modelling of the Ocean) model (Madec 2008) and HadOCC (Hadley Centre Ocean Carbon Cycle) models. The physical model assimilates various data types including sea surface temperature and in-situ profile data using an optimal interpolation-type scheme. The chlorophyll data is assimilated by calculating 2D chlorophyll increments which are balanced with changes in the nutrients, zooplankton, phytoplankton, detritus, pCO_2 and alkalinity using the Hemmings (2008) scheme.

The results of one year hindcasts of the system with and without chlorophyll assimilation are compared. A more extensive description of the work done is in Edwards et al. 2009. A demonstration pre-operational version of this chlorophyll data assimilation system will also be examined.

2. System Description

The hydrodynamic component of the coupled model used in this study is the FOAM-NEMO system which is described in detail in Storkey *et al.* (submitted). This system has been developed for analysis and forecasting of the properties of the deep ocean. The model includes a sea-ice component, the second version of the Louvain-le-Neuve sea-ice model (LIM2, Timmerman *et al.*, 2005). For this work we use a 1° global grid with 42 vertical levels.

The physical data assimilation scheme is described in detail in Martin *et al.* (2007). The FOAM system is capable of assimilating in-situ and satellite sea-surface temperature (SST) data, satellite altimeter sea-level anomaly (SLA) data, satellite sea-ice concentration data, and temperature and salinity profile data. In this application, we are not assimilating the SLA or sea-ice concentration data. The SLA data are not assimilated because a scheme still needs to be developed to change the biological profiles due to changes in the density profile. The HadOCC model proved to be very sensitive to the assimilation of sea-ice data so this was turned off until further testing can be done.

The ecosystem component, HadOCC, is a relatively simple NPZD (nutrient, phytoplankton, zooplankton, and detritus) model that also includes DIC (dissolved inorganic carbon) and alkalinity (Palmer and Totterdell, 2001). The model uses nitrogen has a currency with a fixed carbon to nitrogen ratio. Figure 1 provides a schematic of the HadOCC model and shows the flow of carbon and nitrogen between the model components along with the processes responsible for the transfer. The HadOCC model is coupled on-line to the FOAM-NEMO system allowing the FOAM-NEMO physical fields to drive the ecosystem model variables

which are then updated at the same temporal and spatial scales. It is only a one-way coupling in that there is no feedback from the biological model to the physical model.

The model is forced at the surface by 6-hourly mean fluxes (wind stress, heat flux, and evaporation minus precipitation and sea ice concentration) from the Met Office's Numerical Weather Prediction (NWP) system.

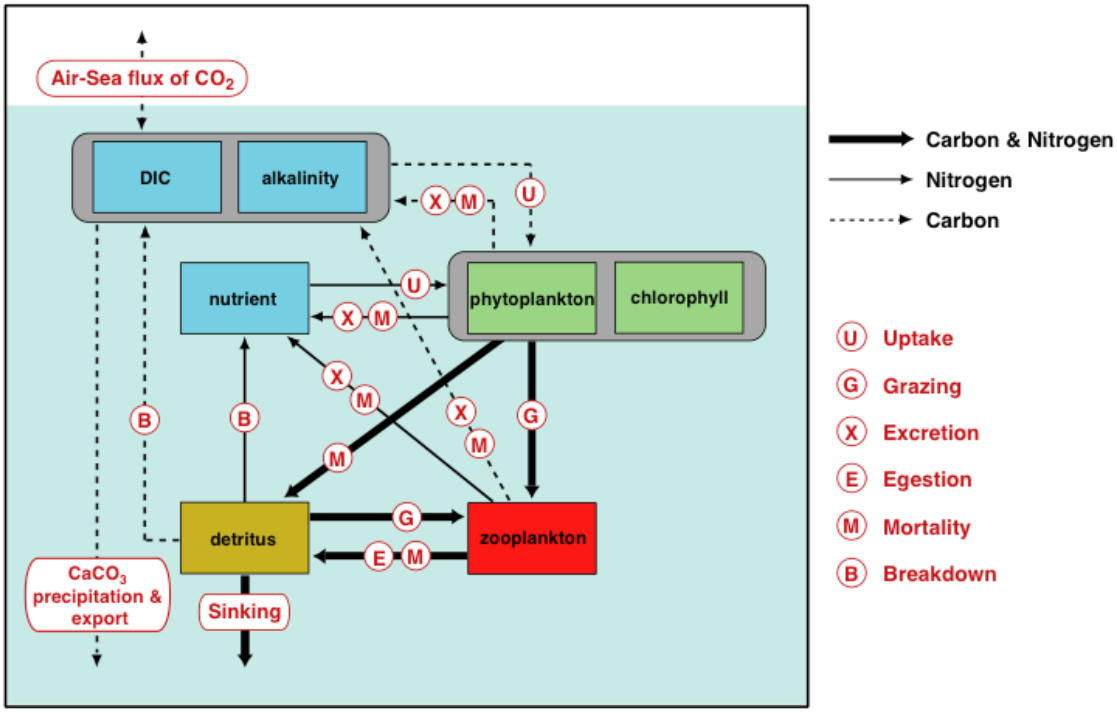

Figure 1: HadOCC schematic.

3. Observations

3.1 Data

The biological data used in the experiments presented here are daily merged L3 (level 3) 4-km resolution gridded surface chlorophyll (CHL1) data (mg m^{-3}). This product is provided by the GlobColour service team, and merges data from the MERIS and MODIS sensors using the GSM model. Detailed information about the product can be found at http://www.globcolour.info and in the GlobColour Product User Guide (Barrot 2007).

4 Assimilation Scheme

The assimilation scheme assumes Gaussian errors. It is well known that phytoplankton biomass is closer to a logarithmic distribution in nature therefore we must convert to log10 chlorophyll, log10(Chl), in the assimilation.

There are three stages to each day's run. Firstly, the model is run with the observation operator applied (stage 1) where model values to observation locations at the observation times. "Feedback files" are then created for each of the assimilated variables (temperature,

salinity and chlorophyll). These contain the observation location, observation time, observation value and model value for each of the observations. These feedback files are passed to both the physical and biological assimilation schemes, and contain all the necessary information these schemes require. Once the observation operator run is complete, the assimilation scheme is run and the increments calculated (stage 2). Finally (stage 3), the model is run again for the same day, with the increments applied using the incremental analysis update (IAU) technique (Bloom *et al.*, 1996).

4.1 Error covariance

The log10(Chl) error covariances were calculated from the output of the control run (with no biological assimilation), using both the observations and model configuration that were to be used in the assimilation runs. Using the innovations (observation minus background values) from the feedback files, the covariances were calculated using the method described by Hollingsworth and Lönnberg (1986) which assumes the observation errors are spatially uncorrelated, attributing any spatial correlation of the observation minus background errors purely to the background errors. The covariance is fitted as a function of the separation distance of observations, and can be extrapolated to zero separation to obtain the background variance for that grid point. The difference between this and the total variance is taken to be the observation variance. The fitting routine also separates the contributions at the mesoscale and synoptic scale. In this case the length scales were 60km and 400km respectively. These are kept constant, but their relative contributions will vary with their variances.

4.2 Details of biological data assimilation scheme

The scheme to assimilate ocean colour data is described in detail in Hemmings *et al.* (2008). This is a nitrogen balancing scheme that begins when a 2-dimensional analysis of log10(Chl) is performed producing a field of surface log10(Chl) increments. These can then be converted into surface phytoplankton increments using the model's nitrogen: chlorophyll (N:Chl) ratio and then into 3D increments for all of the state variables in the ecosystem model. The N:Chl ratio is based on a combination of a fixed C:N ratio and the model (fixed) carbon:chlorophyll (C:Chl) ratio.

The increments to phytoplankton are used to determine surface increments for the other nitrogen tracers in the model (nutrients, zooplankton, and detritus). Importantly, the balancing factors are not independent of the background state and attempt to conserve nitrogen subject to its availability at each model grid point. Surface increments are then assigned to each level above the mixed layer depth. For levels below the mixed layer depth, compound increments are applied. These compound increments are equal to the sum of any primary and secondary increments where the primary increments are based on the surface increments. The secondary increments are nutrient profile correction increments which are based on the expectation that nutrient concentrations increase monotonically with depth throughout most of the ocean. Secondary increments are required only where the surface-layer and primary increments would create an unrealistic sub-surface nutrient minimum. Finally, the increments for the carbon tracers in the model (DIC and alkalinity) are based on constant C:N ratios within the model.

5 Model Runs/Experiments

The coupled hydrodynamic-ecosystem model was spun up for the full year of 2007. Initial conditions for the physics came from the operational FOAM-UM system on 31 December 2006. The initial conditions for the ecosystem variables were from an earlier FOAM-UM-HadOCC system representing northern hemisphere winter-time conditions. Physical data assimilation was turned on after 3 months and a further 9-months of model integration were included in the spin up period.

To test the impact of assimilating GlobColour ocean colour data on the ecosystem model diagnostics. We compare a "control" run, without biological data assimilation, and a biological assimilating run for calendar year 2008. The "L3 assim" run includes the daily assimilation of the GlobColour L3 ocean colour data as well as physical data. Both integrations in this experiment used the same spun-up initial condition, surface forcing from the NWP model and the same observations in the physical data assimilation scheme.

6 Assessment

To determine the impact of the assimilation of ocean colour data on the ecosystem model, we compare the control and L3 assim model runs with each other and with available data. Time series plots of root mean square (RMS) and mean error statistics of observations minus model background are shown below.

A comparison of the errors (RMS and mean) in the control and chlorophyll assimilating run for the global ocean is provided in Figure 2a and shows that the assimilation reduces both the RMS and mean errors and provides an improvement in the chlorophyll fields throughout the year. Figures 2b and 2c present a similar comparison for the North Atlantic and Arctic Oceans only. In both the global (Figure 2a) and North Atlantic (Figure 2b) averages, it is apparent that the assimilation is performing as expected and reducing both the RMS and mean errors. The main source of error remaining in the global statistics is in the Arctic region as seen in Figure 2c where there is very little difference between the control and L3 assim runs. Little or no data are being assimilated in the Arctic resulting in high errors during the bloom period. In Figure 2b, we can clearly see a seasonal cycle of over-prediction of chlorophyll in the spring and under-prediction in the autumn/winter in the control run.

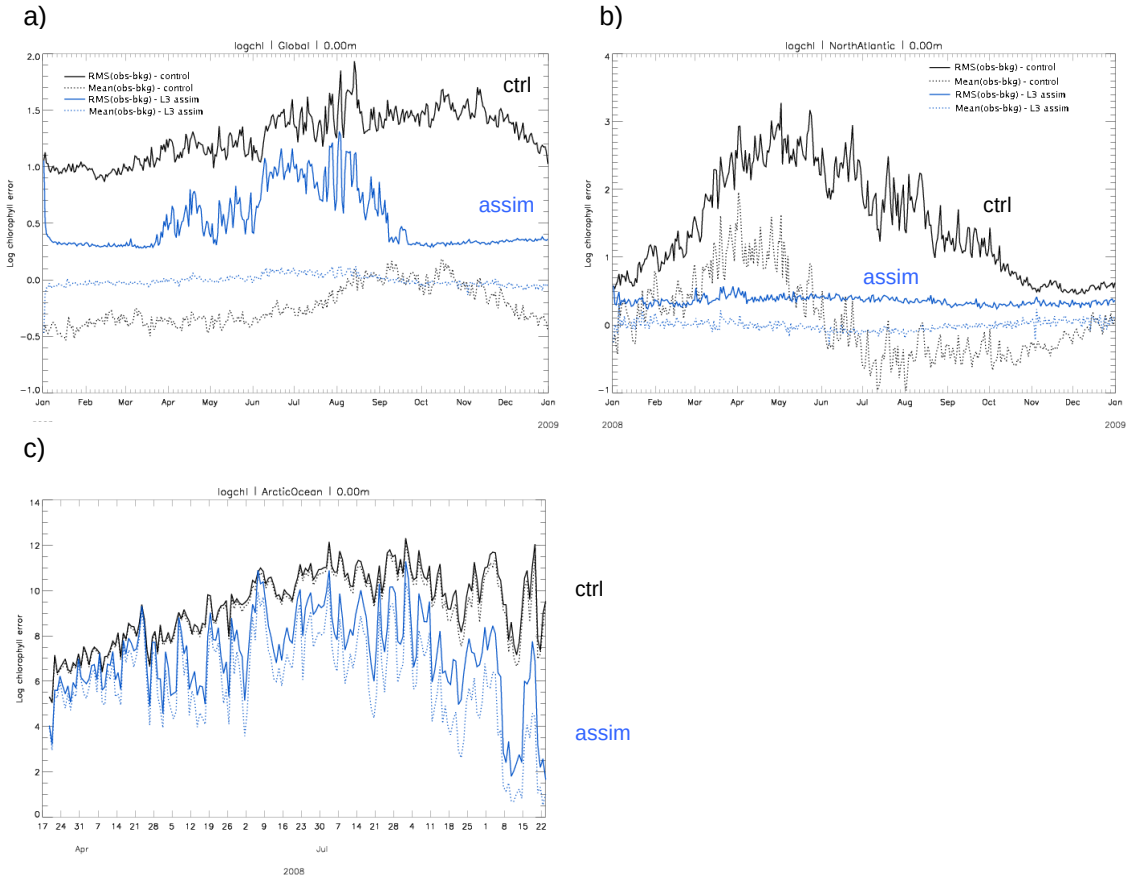

Figure 2. Comparison of the global average RMS (solid) and mean errors (dotted) between the daily observed and modelled chlorophyll for 2008. Note that the chlorophyll is on a log scale. Black = control; blue/grey = L3 assim; RMS error = solid; mean error = dotted. a) Global, b) North Atlantic, c) Arctic Ocean (for April to August).

As part of this project, a monthly surface chlorophyll climatology has been generated using the GlobColour L3 binned data from1998 to 2007. These data provide a more independent source of validation for the model results. Similar to the analysis presented above, In figure 3 the climatology and model chlorophyll fields are compared for the North Atlantic. Again, the chlorophyll assimilating run has lower RMS and mean errors when compared to the control run.

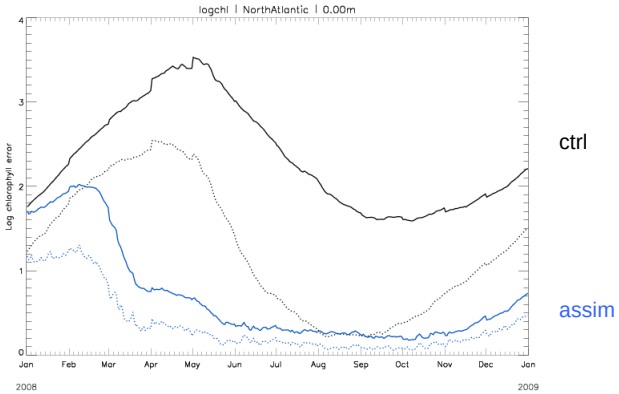

Figure 3: Comparison of RMS and Mean errors between the monthly climatology generated using L3 GlobColour data and daily model chlorophyll for the North Atlantic in model integrations with and without the biological data assimilation. Control = black; L3 assim = blue; RMS error = solid; mean error = dotted.

The spatial differences in the areas of productivity are also quite important. Figure 4a shows the annual average log chlorophyll from the climatology. In general, chlorophyll concentration in the sub-tropical gyres is quite low with higher values at higher latitudes and along the equator in the Pacific. The pattern of chlorophyll in the control run (Figure 4b) shows a very different pattern with high values throughout most of the sub-tropical gyres and again in the Antarctic. It has relatively low values in the sub-polar gyres. Comparing this with the annual average chlorophyll in the L3 assim runs (Figure 4c), it is apparent that the assimilation scheme has greatly changed both the distribution and overall abundance of phytoplankton with the L3 assim run showing the same general pattern as in the climatology. The main area of difference between the climatology (Figure 4a) and the L3 assim run (Figure 4c) is in the Arctic Ocean, where, as described above, we are not assimilating many of the observations. Other differences occur in the Antarctic and upwelling regions which may be due to limited observations in these regions

.

a)

b)

c)

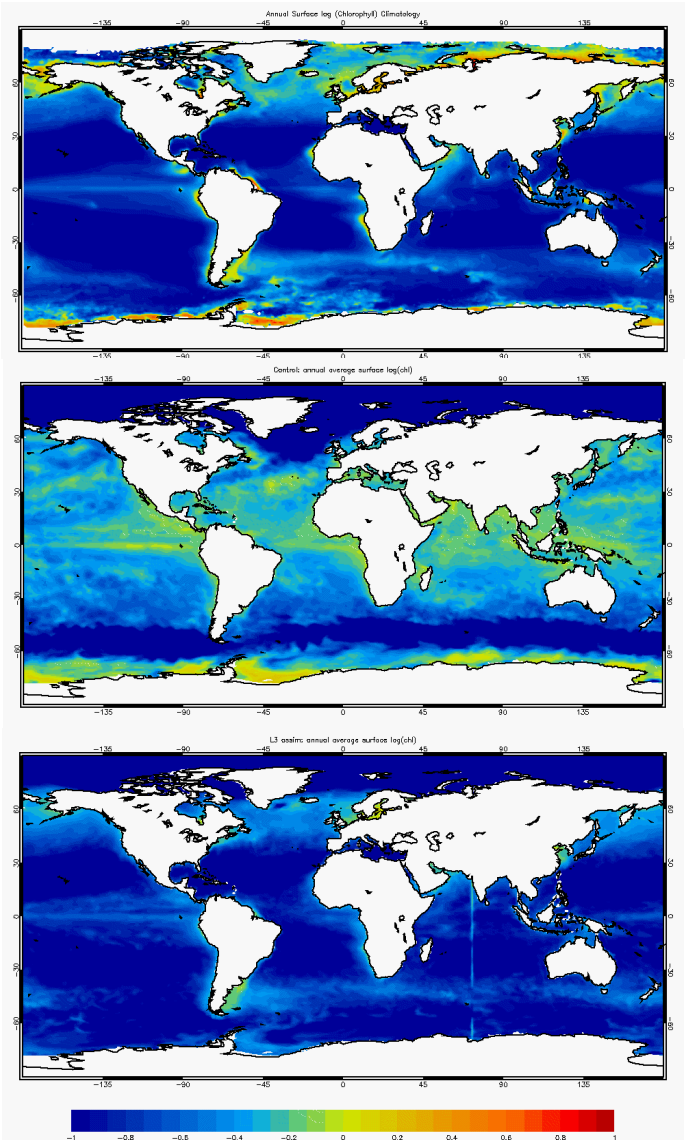

Figure 4. a) The annual average chlorophyll climatology, generated with L3 GlobColour data, on a log scale from -1 to +1. b). climatology from the control run. c). L3 assim run. The vertical line apparent at ~60E is due to a missing overlap region in the plotting not an error in the model.

The model results of unassimilated variables can also be assessed. However suitable observations are sparse. One source is the World Ocean Atlas 2005 (WOA05) which provides a global climatology of objectively analyzed fields on a 1-degree latitude-longitude grid (Garcia *et al.*, 2006 for nutrient analysis). Comparing the model results to the climatology it is apparent that neither the control nor the chlorophyll assim run matches particularly well and both have too much nitrate in relation to climatology.

7 Daily run suite

A daily near-real-time system has been running assimilating the GlobColour L3 chlorophyll since 14th June. This suite:

- downloads the GlobColour data for the previous day;
- performs a 1-day analysis using all available physical and ocean colour data; and
- provides a 5-day forecast (since 29th June only).

The system runs at 14:40 GMT, as the GlobColour data are available from 14:00 GMT, each day on a pseudo-operational basis, mimicking the daily operational FOAM suite. The latest observation file, which is valid for the previous day, is downloaded from the GlobColour FTP server, and quality controlled by the Met Office's observation processing system using the method of Ingleby and Huddleston (2007). Physical observations and surface fluxes, also valid for the previous day, are copies from the FOAM operational suite. However it is worth noting that due to delays in receipt of some types of data, there are less physical observations available each day than for the hindcasts. As with the hindcasts, the model is run for one day with the observation operator applied, then the biological and physical data assimilation is performed, and finally the model is run again for the same day with the increments applied using the IAU as explained previously. This produces the analysis for the day the observations are valid. The model is then run forward for a further five days to produce a five-day forecast, with daily output produced.

8 Discussion and future developments

Year-long integrations of the FOAM-NEMO HadOCC system have been completed on a 1° grid with and without the assimilation of the GlobColour chlorophyll data. Finally, a near real time run is being performed daily in a quasi-operational status.

The assimilation of merged L3 GlobColour chlorophyll data into the coupled physical-biogeochemical modelling system has been shown to significantly improve the chlorophyll (and therefore phytoplankton) fields when compared to the control run which included physical data assimilation only. Assimilation of chlorophyll data also significantly reduced the model errors seasonality. However, it also resulted in a small increase in the amount of nitrate in the sub-tropical gyres at the same time, which is related to the way the assimilation scheme propagated the increments to the unobserved biological variables in the model.

This study also provided useful insight into the coupled ocean model's performance and highlighted areas where further work is required. Future steps include:

- The control run in these experiments included physical data assimilation. The impact of that assimilation on the biological tracers should be assessed. There is some evidence to suggest that the physical data assimilation is partly responsible for the unrealistically high

nutrient field in the sub-tropical gyres. If this is the case, schemes for better constraining the model nitrate need to be investigated.

- In this initial assessment, we used a fixed C:Chl ratio in HadOCC. The impact of using a more realistic and variable C:Chl ratio should be assessed.
- As mentioned above, for the integration period, calendar year 2008, very little *in situ* biogeochemical data is available for comparison with the model tracer variables. Effort to gather this data, especially for comparison to sub-surface values, and integration of this data into the assessment process needs to be continued.
- Other model improvements are currently being considered. Since HadOCC is largely used for climate-scale integrations, the model has been written to use daily average solar radiation as a forcing. However, since we are using 6-hourly fluxes, we could interpolate the instantaneous values provided by the NWP model and include a daily cycle within HadOCC.

References

Barrott G, 2007: ESA DUE GlobColour - Global Ocean Colour for Carbon Cycle Research - Product User Guide, Ref GC-UM-ACR-PUG-01, Version 1.2. http://www.globcolour.info/products_description.html

Bloom SC, Takacs LL, Da Silva AM, and Ledvina D. 1996: Data assimilation using incremental analysis updates. *Mon. Weather. Rev.* **124**: 1256-1271.

Garcia HE, Locarnini RA, Boyer TP, and Antonov JI. 2006: World Ocean Atlas 2005, Volume 4: Nutrients (phosphate, nitrate, silicate). S. Levitus, Ed., *NOAA Atlas NESDIS 64, U.S. Government Printing Office*, Washington, D.C., 396 pp.

Hemmings JCP, Barciela RM and Bell MJ. 2008: Ocean color data assimilation with material conservation for improving model estimates of air-sea CO_2 flux. *J. Mar. Res.* **66**: 87-126.

Hollingsworth A and Lönnberg P. 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus* **38A**, 111-136.

Ingleby B and Huddleston M. 2007: Quality control of ocean temperature and salinity profiles – historical and real-time data, *J. Mar. Sys.* **65**, pp 158-175

Edwards KP, Barciela RM, Ford D, Lea D and Martin M. 2009: Assimilation of GlobColour Products into Operational Models, GlobColour Final Report. Reference: ARG/001-001/0812_CCN_METOFFICE.

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

Martin MJ, Hines A and Bell MJ. 2007: Data assimilation in the FOAM operational short-range ocean forecasting system: A description of the scheme and its impact. Q. *J. R. Meteorol. Soc.* **133**: 981-994.

Palmer JR and Totterdell IJ. 2001: Production and export in a global ocean ecosystem model. *Deep-Sea Res.* I, **48**, 1169-1198.

Stark JD, Donlon CJ, Martin MJ and McCulloch ME. 2007: OSTIA : An operational, high resolution, real time, global sea surface temperature analysis system. Oceans '07 IEEE Aberdeen, conference proceedings. Marine challenges: coastline to deep sea. Aberdeen, Scotland. *IEEE*. 21 June 2007.

Timmermann R, Goosse H, Madec G, Fichefet T, Ethe C, and Duliere V. 2005: On the representation of high latitude processes in the ORCA-LIM global coupled sea ice-ocean model. *Ocean Modelling* **8**: 175-201.