Authors' response to an anonymous review (reviewer #2) of "Evaluation of NorESM-OC (versions 1 and 1.2), the ocean carbon-cycle stand-alone configuration of the Norwegian Earth System Model (NorESM1)" by Schwinger et al.

We thank the reviewer for reviewing our manuscript and for his/her constructive and helpful comments. Our detailed response to the points raised is given below (reviewer's comments in italic font, our response in normal font).

The paper falls well within the scope of Geoscientific Model Development. It is overall well structured, although in detail, information can not always be found where one would expect to find it (some parameter values are only reported in the discussion section, instead of the model description part). The English language is generally good. There remain a few minor shortcomings that can nevertheless be easily fixed.

In the revised manuscript all parameter values are defined in the model description part.

The presentation and description of the model is rather complete, the discussion of the results is to some extent merciless: the authors put much emphasis on the biases in their model results. I would be pleased to read a few more sentences about the strengths of their model.

We added a few sentences on this at the beginning of the model description section. "The main benefit of an isopycnal model is the good control on the diapycnal mixing (less numerical diffusion across isopycnic surfaces) that helps to preserve water masses during long model integrations. This is of particular interest for a coupled physical-biogeochemical model, since sharp gradients in tracer concentrations between water masses (e.g. in the thermocline region) can potentially be better represented by the model. The isopycnic framework provides a terrain-following vertical coordinate, and therefore overflow of water masses is modelled without numerical obstacles (although mixing at steep slopes must be parameterised carefully)."

General model presentation: scope

The model is clearly global in scope. Time frames for applications are not clearly stated. The paper presents simulation experiments over a few centuries, following 800 or 1000 year spin-ups. Can it also be reasonably used on longer time-scales (a few thousands to a few tens of thousands of years)?

We added this information in the introduction as follows: "Although NorESM-OC is computationally less expensive than the fully coupled NorESM, computational constraints limit the time-scale, for which the model can be applied, to a few hundred to thousand years on current hardware. We note, that this limitation is mainly due to the physical ocean model and the costly transport of tracers. By using an efficient offline

method (e.g., Khatiwala et al., 2005; Kriest et al., 2012) it would be possible to apply HAMOCC on time-scales at least one order of magnitude longer, as long as all relevant external inputs are provided (e.g., fluxes due to continental weathering). Similar versions of HAMOCC have been applied on time-scales of 50 000 years (e.g., Heinze et al., 2003, who use an offline setup with annually averaged ocean circulation fields)."

Code: availablility and quality

The code is not openly available, but only upon signing licence agreements (at least two, one for NorESM, one for HAMOCC, but perhaps even more – this is not clear). As a reviewer, I was nevertheless granted access to the code. It is written in Fortran (much FORTRAN77-style, with some Fortran 90 elements). A C-style pre-processor is required to compile it. Some files are well commented, others almost devoid of comments. The presence of OpenMP compiler directives suggests that it has been prepared for usage on multi-platform shared memory multiprocessing computing facilities, which is definitely a noteworthy feature of interest to prospective users.

Unfortunately, the manuscript does not provide much information about the code, nor about numerical methods adopted in general (except for time-stepping, which seems to follow a leap-frog scheme in combination with a filter with unspecified characteristics though). These are, however, the informations that are typically expected in a Development and technical paper in Geoscientific Model Development. I would recommend to add a short, descriptive paragraph about such aspects, the more since the code is not publicly available. Finally, it would also be interesting to read about typical execution times of the three versions.

As recommended we added a subsection on technical aspects at the end of the model description section, which reads:

"**2.5 Technical aspects**

The HAMOCC code was originally written in FORTRAN77. It was later re-written to take advantage of FORTRAN90 elements (e.g. ALLOCATE statements) and to conform to FORTRAN90 free source code format. Certain model options, for example the selection of a POC sinking scheme, are implemented via C-preprocessor directives. HAMOCC should compile on any platform that provides a FORTRAN compiler and a C-preprocessor.

MICOM is parallelised by dividing the global ocean domain horizontally into (logically) rectangular tiles which are processed on one processor core each. Communication between cores is implemented using the Message Passing Interface (MPI) standard. Since HAMOCC is integrated into MICOM as a subroutine call it inherits this parallelism. Note that HAMOCC, in addition, has been parallelised for shared memory systems using the OpenMP standard. This feature is, however, not tested and not supported for the current set-up. For the simulations presented in this manuscript, the ocean component has been run on 190 (Mv1), 309 (Mv1.2), and 155 (Lv1.2) cores on a Cray XE6-200 system, yielding a model throughput of 11 (Mv1), 20 (Mv1.2), and 64 (Lv1.2) simulated model years per (wall-clock) day."

We also better specified the characteristics of the time-smoothing applied with the leap-frog time stepping in MICOM: "To prevent excitation of numerical noise, the MICOM leap frog time-stepping includes a time-smoothing applied to the mid time-level of temperature, salinity and layer thickness fields at each time step, i.e. $x_{\mathrm{mid}} = w_1 x_{\mathrm{mid}} + w_2(x_{\mathrm{old}} + x_{\mathrm{new}})$ with $w_1 = 0.875$ and $w_2 = 0.0625$. Since the physical fields undergo this time smoothing, but the biogeochemical tracers do not, ...".

In the "Code availability" section we clarified the number of licenses that is required: "...availability of the code is subject to signing two license agreements — one for the use of NorESM, and, additionally, for the use of HAMOCC signing of the MPI-ESM license agreement is required. ..."

Specific comments:

page 4, line 19: "[...] NorESM-OC1 corresponds to the fully coupled NorESM1-ME [...]" would better read "[...] NorESM-OC1 corresponds to the version included in the fully coupled NorESM1-ME [...]"

This has been adopted.

Although the paper reads as if NorESM-OC1 was a sub-model isolated from NorESM1, its setup is highly similar to the earlier isopycnic ocean carbon cycle model of Assmann et al. (2009). How does NorESM-OC1 actually differ from the model described by Assmann et al. (2009)?

We added information on this in Sec. 2.3 as follows: "The model described by Assmann et al. (2010) was the starting point for the development of the ocean biogeochemistry component of NorESM1. Main differences between this model and NorESM-OC1 are updates of mixed layer and eddy diffusivity parameterisations, the use of CICE as ice model, and the details of atmospheric forcing, which followed Bentsen and Drange (2000), wheras NorESM-OC uses the approach of Large and Yeager (2004). The version of HAMOCC employed by Assmann et al. (2010) is very similar to the one in NorESM-OC1 with the notable difference of an update of the carbon chemistry scheme (see below)."

page 7, lines 14–17: I presume that the preformed tracer values are set to the mixed-layer concentrations each time that a water parcel leaves the surface mixed-layer. Please rephrase for clarity.

Yes, the preformed tracers are set to the tracer concentration in the mixed layer at each time step. We clarified this by adding "at each time step".

page 8, line 26 – page 9, line 7: the detailed approximation adopted for TA is given, but it would also be interesting to know how the speciation of the acid-base systems is calculated, or, equivalently, how pH is calculated from the TA expression.

We added this information as follows: "We use an iterative carbonate alkalinity correction method (Follows et al., 2006; Munhoven, 2013) to calculate the oceanic partial pressure of CO_2 (pCO_2) prognostically as a function of temperature, salinity, dissolved inorganic carbon (DIC), and TA. Using an initial guess for the hydrogen ion concentration $[H^+]_0$, all non-carbonate contributions to TA are calculated according to Dickson et al. (2007) and added to or subtracted from (1) in order to obtain an initial guess for carbonate alkalinity CA_0 ($CA = [HCO_3^-] + 2[CO_3^{2-}]$). Given CA and DIC and the equilibrium constants of the carbonate system K_1 and K_2, the corresponding hydrogen ion concentration can be calculated:

$$[H^+] = \frac{K_1}{2\,\mathrm{CA}} \left(\mathrm{DIC} - \mathrm{CA} + \sqrt{(\mathrm{DIC} - \mathrm{CA})^2 + 4\,\mathrm{CA}\,K_2/K_1(2\,\mathrm{DIC} - \mathrm{CA})} \right). \quad (2)$$

The hydrogen ion concentration $[H^+]_1$ thus obtained is used to re-iterate these calculations until convergence is reached. Here, we use $[H^+]$ from the previous time step as an initial guess for the calculation of CA, and stop iterations once the relative change $([H^+]_{i+1} - [H^+]_i)/[H^+]_{i+1})$ becomes smaller than $\epsilon = 5 \times 10^{-5}$. Only at the first time step of integration we use $[H^+]_0 = 10^{-8}\ \mathrm{mol\,kg^{-1}}$. As pointed out by Follows et al. (2006), the changes in $[H^+]$ from one time step to the next are small, and one or a few iterations of this procedure usually suffice."

page 12, lines 6–8: I may possibly misinterpret this, but I am wondering how the loss of material (and alkalinity) through the sediment is compensated for, if the model does not take into account the influx of carbon, alkalinity and nutrients? Does it operate in some kind of closed-loop configuration (mass an alkalinity that are lost via the deep- sea sediment are re-injected at the top)? Please clarify how the global inventories of the model ocean are conserved.

There is no compensation for the mass and alkalinity lost to the deep-sea sediment in the model versions presented here, which limits the time-scale of applicability. We clarified this by adding: "The material lost to the sediments is therefore not replaced by some mechanism in the model configurations presented here. The model drift caused by these losses to the sediment is small on the time-scales considered in this manuscript, particularly at the surface. Nevertheless, this simplification limits the applicability of NorESM-OC1 and 1.2 to time-scales of the order of 1000 years. Input of DIC and nutrients in particulate and dissolved forms through rivers is currently implemented into NorESM."

page 13, lines 12–13: while I find it rather obvious how M can be a prognostic variable, I do not immediately see how NOS can be handled as such. Please give a few details about how NOS is predicted.

We rewrote this paragraph to accommodate some more details on how NOS is treated as follows. "NOS is treated like other particulate tracers in the model, i.e. NOS is advected and diffused by ocean circulation and treated in HAMOCC's sinking scheme (see below) using the average sinking speed for particle numbers as given in Kriest (2002).

Additionally, aggregation of particles decreases NOS, while photosynthesis, egestion of fecal pellets, and zooplankton mortality increase NOS. The size distribution is affected by these process as follows. Sinking removes preferentially the large particles and leaves behind the smaller ones, thereby steepening the slope ε of the size spectrum in surface layers. Aggregation "flattens" the slope of the size spectrum, because it reduces the number of particles, but not mass. These processes are parameterised as in scenario "pSAM" of Kriest (2002),... We assume that all other biogeochemical processes do not impact ε (e.g., photosynthesis increases M and NOS proportionally, such that the slope of the size distribution is not affected), the exception being zooplankton mortality, which flattens the size spectrum through the addition of (large) zooplankton carcasses."

page 15, beginning of section 2.4.1: I recommend to introduce this section by one or two general sentences that summarize which data are required to force the model, before coming to the specific versions of datasets used.

We followed this recommendation by adding an introductory sentence to this section: "The atmospheric forcing required to run NorESM-OC comprises temperature, specific humidity and wind at 10 m reference height as well as sea level pressure, precipitation and incident short wave radiation fields."

page 16, line 8: Only the fate of the Antarctic freshwater influx is detailed. How is the rest of the freshwater treated?

We clarified this by adding: "All runoff fluxes are mapped to the ocean grid and smeared out over ocean grid cells within 300 km of each discharge point to account for unresolved mixing processes."

page 16, line 8: 365 days for v1 and 350 days for v1.2: why this difference?

This was also asked by the first reviewer, and we repeat our response to this question here: There was a typo in our original manuscript. The relaxation time scale stated for model version 1.2 should read 300 instead of 350 days. The somewhat shorter relaxation time scale applied in version 1.2 was introduced to approximately counterbalance the weakening of the relaxation flux due to the balancing option. In this context we also found an error in the following sentence describing the balancing option. Correcting these errors and adding the information asked by the reviewer, this paragraph now reads: "In order to stabilise the model solution, we apply salinity relaxation towards observed surface salinity with a restoring time scale of 365 days (version 1) and 300 days (version 1.2) for a 50 m thick surface layer. The restoring is applied as a salt flux which is also present below sea ice. In model version 1.2 balancing of the global salinity relaxation flux was added as an option, which allows to keep the global mean salinity constant over long integration times. That is, positive (negative) relaxation fluxes (where "positive" means a salt flux into the ocean) are decreased (increased) by a multiplicative factor if the global total of the relaxation flux is positive (negative). The somewhat shorter relaxation time scale applied in version 1.2 was chosen to approximately counteract the weakening of the relaxation flux due to this balancing procedure."

page 17, line 8: is atmospheric pCO2 prescribed or is it prognostic?

It is prescribed. We clarified this by rephrasing "... using the CORE normal year forcing and a prescribed atmospheric CO_2 concentration of 278 ppm...".

page 17, lines 13–14: the carbon flux of 0.26 GtC/yr from Mv1 is about 50 times larger than the fluxes for the other two and not truly negligible (it is of about the same order of magnitude as the atmospheric CO_2 consumption rate by continental weathering). Is this correct, and if so, what is the reason? Furthermore, the drifts of Mv1.2 and L1.2 are of the order of 5–10%, which is far from negligible over 100 yr. This may point out significant deviation from the sought near-equilibrium state.

The carbon-flux in Mv1 is due to the too low $CaCO_3$ production, and the time-scale at which the model would equilibrate with regard to this process (under prescribed atmospheric CO_2 particularly) is probably of the order of 100000 years. The negligible carbon-uptake in model version 1.2 is a result of the re-tuning of $CaCO_3$ production. If we had implemented riverine inputs to the ocean we would expect to have an outgassing of a few tenths of Pg carbon per year if $CaCO_3$ production was reasonably tuned. The larger drift in Mv1.2 and Lv1.2 is caused by a larger decadal to centenial scale internal variability in these configuration (see response to the first reviewer, Fig. 1). Nevertheless, a trend of the order of 0.05 Pg C century^{-1} is still rather small compared to an anthropogenic change in oceanic C inventory of the order of 100 Pg C over less than a century. Further, our ocean sink estimates are calculated relative to a control run, i.e. any offset and trend in the model is taken into account. We added the following text to the revised manuscript in order to better clarify this.

"The relatively large uptake of carbon in Mv1 at the end of the spin-up is due to a lower $CaCO_3$ to POC production ratio found in this configuration. A full equilibration of the model with respect to this process would require a considerably longer spin-up. The re-tuning of the ecosystem parameterisation, which increases $CaCO_3$ production in version 1.2, leads to carbon fluxes much closer to zero at the end of the spin-up period. The larger trends in Mv1.2 and Lv1.2 compared to Mv1 despite longer spin-up time is due to larger decadal to centennial scale internal variability in these configurations (i.e., the systematic long term drift is much smaller). We attribute this to the details of the salinity relaxation (balancing of the restoring flux, weaker relaxation south of 40° S). We finally note that even these larger trends are tiny compared to the changes in ocean carbon uptake due to anthropogenic carbon emissions, and that we calculate our estimates of anthropogenic carbon uptake relative to a control run to account for offsets and trends due to a not fully equilibrated model."

page 19, lines 19–23: "[...] if this trend would be removed.": is it possible to remove it? Where does it actually come from?

We agree that this formulation was a bit unclear. The reason for this trend is discussed further down in the text. We rephrased this sentence as follows. "Compared to the other model simulations the annual and decadal scale variability of AMOC strength appears

to be similar, but superimposed there is a negative trend of 8 Sv over the simulation period (see below)."

Further down we add a sentence to make it clearer that the discussed peculiarity of the salinity relaxation scheme when the balancing of the relaxation flux is activated leads to this negative trend: "This effect is particularly pronounced in the Lv1.2 configuration leading to the negative AMOC trend of 8 Sv described above."

page 27, line 26: I find these biases rather strong (50% or so). Comments?

It is true that the oxygen biases in the model are rather strong. As described in the text, this is due to model deficiencies in the Southern Ocean and in the oxygen minimum zones in combination with the relatively long spin-up time. Part of the problem is also the too low C transport into the deep ocean when the standard sinking is used. In fact, a reduction of the global average O_2 bias below 3000 m of 40% is attained with the KR02 and WLIN sinking schemes, as mentioned in Sec. 3.5. Finally, the situation is probably not improved by the fact that the model is isopycnic, since the strong gradients between different water masses with opposite biases are not alleviated by numerical diffusion. We rewrite this paragraph to add these points: "We note that part of the deep ocean oxygen bias is connected to low POC transport into the deep ocean when the standard sinking scheme is used. The large positive bias below 3000 m depth is reduced by 40% in Lv1.2 when using alternative sinking schemes. This is further discussed in Sect. 3.5."

And further down: "We finally speculate that the strong O_2 biases of opposite sign in adjacent water masses are probably more pronounced in our isopycnic model than they would be in an z-coordinate model, since the strong O_2 gradients between these water masses are not alleviated by numerical diffusion."

page 32, lines 8, 25 and 27: please provide references for the cited numbers (they can be found in the figure captions, but it would be good to have them in the text as well).

This has been done, each cited number is provided with the respective citation in the revised manuscript.

page 36, lines 23–29: It is fairly possible that the new sinking parametrisations are simply more efficient in counterbalancing the effects of the (too?) strong Southern Ocean ventilation. Using a model biogeochemical process to reduce biases arising from shortcomings in the model physics can hardly be considered an improvement.

We have multiple lines of evidence for "improvement" based on direct and indirect observations. The direct indication of "improvement" comes from observation based estimates of POC fluxes (Figs. 23 and 26 in the revised manuscript). For example, we show that modeled POC fluxes and sediment trap data compare much better with the KR02 and WLIN sinking schemes. This is independent of any process in the physical model. Even if one has limited confidence in sediment trap derived fluxes because of methodological difficulties, Figs. 22a and b (original manuscript) clearly demonstrate that the STD-slow scheme has a much too low POC flux through 2000 m.

Oxygen biases in the SO and the deep water masses are rather strong as mentioned by the reviewer above, and our results indicate that there are two factors that contribute. First, the too strong ventilation which is also indicated by CFC fields as well as cold and fresh biases in the SO. Second, if the standard sinking scheme is used, a contribution of too little remineralisation in the deep ocean. In Lv1.2 the global average bias below 3000 m is reduced from 0.1 to 0.06 mol m^3 with the KR02 and WLIN schemes. Reductions in O_2 biases are also seen in regions that are not ventilated by the SO, and where we think that the physical model performs reasonably well (e.g. North Atlantic Deep Waters). Given the evidence from direct POC flux observations, we think that it is justified to consider the reduction of O_2 biases an improvement.

page 37, line 11: "[...] a long standing problem [...]": please provide a reference to support for this statement.

This has been done. We cite the original work of Najjar et al. (1992) and the more recent work of Dietze and Loeptien (2013).

Technical comments

Throughout the paper: please consistently use either "parameterisation/parameterise" or "parametrisation/parametrise" as a spelling

All occurrences of "parametrisation/parametrise" have been replaced by "parameterisation/parameterise" in the revised manuscript.

page 2, line 14: "[...] scheme prescribing a linear increase of sinking speed [...]" would better read "[...] scheme that uses a linear increase of the sinking speed [...]"

This has been adopted.

page 3, lines 15–24: please indicate which versions of CESM, CAM-Oslo, MICOM and HAMOCC are used, respectively.

This has been done except for MICOM, since there are no official version numbers attached to the MICOM developed in Bergen.

page 3, line 28: "persue" should read "pursue"

This has been corrected.

page 4, lines 23–24: It is at this point not entirely clear what "on a numerically more efficient grid in 1° and 2° resolution" means. Based upon what we read later on, this should actually read "on a numerically more efficient grid at either 1° or and 2° resolution". Also: are these actual or nominal resolutions? If they are nominal, what are the extremes?

We clarified this by rephrasing "... is configured on a numerically more efficient grid at either 1° or 2° nominal resolution...".

We propose to add more detailed information on the grid spacing in Sec. 2.4 as follows:

"We discuss three different model grid configurations, one for NorESM-OC1, which runs on a displaced pole grid with 1.125° nominal resolution and with grid singularities over Antarctica and Greenland. NorESM-OC1.2 has been set up on a numerically more efficient tripolar grid in 1 and 2° nominal resolution. The tripolar grid has its singularities at the South Pole, in Canada, and in Siberia. The nominal resolutions given here indicate the zonal resolution of the grid while the latitudinal resolution is finer and variable. For the displaced pole grid of NorESM-OC1, the latitudinal grid spacing is 0.27° at the equator gradually increasing to 0.54° at high southern latitudes. The tripolar grid of NorESM-OC1.2 is optimised for isotropy of the grid at high latitudes, and the latitudinal spacing varies from 0.25° (0.5°) at the equator to 0.17° (0.35°) at high southern latitudes for the 1° (2°) nominal resolution. Note that over the northern hemisphere both grid types are distorted to accommodate the displaced pole over Greenland or the dual pole structure over Canada and Siberia."

page 7, sect. 2.3: which version of HAMOCC is finally used? 5 or 5.1? Please check and make sure the text is consistent.

We clarified in the text that HAMOCC5.1 is used.

page 7, line 26: "seriously" would better read "significantly"

page 7, line 27: "[...] the tracer transport fully consistent [...]" should read "[...] the tracer transport scheme to make it fully consistent [...]"

Thank you, these suggestions has been followed.

References

Assmann, K. M., Bentsen, M., Segschneider, J., and Heinze, C.: An isopycnic ocean carbon cycle model, Geosci. Model Dev., 3, 143–167, doi:10.5194/gmd-3-143-2010, 2010.

Bentsen, M. and Drange, H.: Parameterizing surface fluxes in ocean models using the NCEP/NCAR reanalysis data, RegClim, Regional Climate Development Under Global Warming, General Technical Report no. 4, pp. 149–157, 2000.

Dickson, A., Sabine, C., and Christian, J.: Guide to best practices for ocean CO_2 measurements, PICES special publication 3, North Pacific Marine Science Organization (PICES), Sidney, British Columbia, Canada, 2007.

Dietze, H. and Loeptien, U.: Revisiting "nutrient trapping" in global coupled biogeochemical ocean circulation models, Glob. Biogeochem. Cyc., 27, 265–284, doi: 10.1002/gbc.20029, 2013.

Follows, M. J., Ito, T., and Dutkiewicz, S.: On the solution of the carbonate chemistry system in ocean biogeochemistry models, Ocean Modelling, 12, 290–301, 2006.

Heinze, C., Hupe, A., Maier-Reimer, E., Dittert, N., , and Ragueneau, O.: Sensitivity of the marine biospheric Si cycle for biogeochemical parameter variations, Glob. Biogeochem. Cyc., 17, 1086, doi:10.1029/2002GB001943, 2003.

Khatiwala, S., Visbeck, M., and Cane, M. A.: Accelerated simulation of passive tracers in ocean circulation models, Ocean Modelling, 9, 51–69, 2005.

Kriest, I.: Different parameterizations of marine snow in a 1-D model and their influence on representation of marine snow, nitrogen budget and sedimentation, Deep-Sea Res. I, 49, 2133–2162, 2002.

Kriest, I., Oschlies, A., and Khatiwala, S.: Sensitivity analysis of simple global marine biogeochemical models, Global Biogeochem. Cycles, 26, GB2029, doi:10.1029/2011GB004072, 2012.

Large, W. and Yeager, S.: Diurnal to decadal global forcing for ocean and sea-ice models: The data sets and flux climatologies, Tech. Note NCAR/TN-460+STR, National Center of Atmospheric Research, Boulder, Colorado, 2004.

Munhoven, G.: Mathematics of the total alkalinity–pH equation – pathway to robust and universal solution algorithms: the SolveSAPHE package v1.0.1, Geosci. Model Dev., 6, 1367–1388, doi:10.5194/gmd-6-1367-2013, 2013.

Najjar, R. G., Sarmiento, J. L., and Toggweiler, J. R.: Downward transport and fate of organic matter in the ocean: Simulations with a general circulation model, Glob. Biogeochem. Cyc., 6, 45–76, 1992.