

Authors’ response to an anonymous review (reviewer #1) 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).

My only major comment would be that, while intercomparing the various versions of the model, the manuscript does nothing to contextualise the performance of NorESM-OC within the context of other Earth system models. The CMIP5 archive is something of a treasure trove on this point, and most of the evaluations made in the manuscript could be repeated with output from it. However, I would only suggest adding an overview of this (e.g. Taylor diagrams?), and not amending the manuscript throughout the main task here, and I would agree with the authors on this, is to evaluate performance between traceable versions of NorESM-OC.

We see the reviewers point that it would be interesting for the reader to see how NorESM-OC performs relative to similar models. We therefore propose to add a figure showing the carbon uptake of our model compared to the models used in Le Quéré et al. (2015). We believe that this would satisfactorily meet the reviewer’s concern because *i*) estimation of contemporary C-uptake is and has been an application of major importance for the stand-alone configuration of our model, *ii*) C-uptake provides a kind of integrated measure since it depends on several aspects of model performance, *iii*) the models used in Le Quéré et al. (2015) are truly comparable to NorESM-OC also in terms of atmospheric forcing (all these models are forced by reanalysis data). We have added this as Fig. 21 of the revised manuscript.

Pg. 2, ln. 4: “We present simulation results . . .” this sentence could perhaps be a little clearer on how many configurations are examined; the mention of multiple resolutions makes it a little opaque

We rephrased this sentence as follows. “We present simulation results of three different model configurations (two different model versions at different grid resolutions) using two different atmospheric forcing data sets.”

Pg. 3, ln. 12: The authors simulate the ocean component of an ESM under reanalysis forcing; did they consider running under atmospheric output from the ESM instead?

We did not consider this kind of experiments in this study, but it is certainly a very useful application of stand-alone configurations. We added a short sentence on this in the revised manuscript: “Idealised experiments include set-ups where (purposefully manipulated) atmospheric output from fully coupled ESM runs is used to force the stand-alone configuration of the same or another model.”

Pg. 8, ln. 1: “We anticipate . . . of 10,000 years . . .” would I be correct in assuming this “anticipation” is either too computationally expensive to test, or is not actually possible because v1.2 of the model rolls in several changes and not just this one (i.e. it is not possible to separate differences due to the time-stepping change from those from other sources)?

We did tests (1000 years with the low resolution configuration) comparing model runs that differed only in the time-stepping method, and we found that the simulation results were not significantly different. Nevertheless, the global correction that accounts for the non-conservation of tracer mass due to the time-smoothing error in version 1 leads to a small but continuous unphysical re-distribution of tracer mass. By extrapolating these differences we come to the “anticipation” that for time scales of 10,000 years the fully consistent time stepping in version 1.2 would be relevant. The reviewer is correct in the assumption that this would be rather computationally expensive to test by actually running the model for 10,000 years (about a wall-clock year of run time with the low resolution configuration).

Pg. 8, ln. 13: “derived by Wanninkhof (1992)” it might be worth examining Wanninkhof (2014) for updates here; I believe that some of the issues mentioned subsequently are tackled there

Yes, the updated formulation of Wanninkhof (2014) includes a re-fit of the Schmidt-Numbers avoiding the mentioned problems at high SSTs. We decided to adopt the Gröger and Mikolajewicz (2011) formulation to keep our HAMOCC version close to the one used by the Hamburg-group (Ilyina et al., 2013). We mention the updated formulation of Wanninkhof (2014) in the revised manuscript as follows: “We note that the updated air-sea gas exchange formulation provided by Wanninkhof (2014) also includes re-fitted Schmidt numbers, avoiding problems at high SST.”

Pg. 10, ln. 15: A remark about the shape of the nutrient limitation relationship would be useful; also something about the relationship between nutrients (e.g. Leibigs Law or something else?)

We added this information as follows. “The nutrient limitation is expressed through a Monod function $f_{\text{nut}} = X/(K + X)$, where K is the half-saturation constant for nutrient uptake ($K = 4 \times 10^{-8} \text{ kmol P m}^{-3}$), and

$$X = \min \left(\text{PO}_4, \frac{\text{NO}_3}{R_{\text{N:P}}}, \frac{\text{Fe}}{R_{\text{Fe:P}}} \right).$$

$R_{\text{N:P}} = 16$ and $R_{\text{Fe:P}} = 366 \times 10^{-6}$ are the (constant) nitrogen to phosphorus and iron to phosphorus ratios for organic matter used in the model.”

Pg. 10, ln. 23: Presumably this means that opal “production” is associated with export only, and is not representative of all of the opal production by diatoms (some of which is dissolved before it can be exported)?

Yes, this is correct. We clarified this by adding the word “export”. “The distribution of

calcium carbonate and biogenic silica export production depends on the availability of silicic acid ...”

Pg. 11, ln. 4: “we performed a re-tuning” by eye?

We did not apply an objective optimization procedure for the model tuning. Data assimilation schemes are being developed for our biogeochemistry module, but are not yet available. Instead, our model tuning was based on our experience (as to which model parameter could be tuned to achieve a certain effect) and a long series of trial and error simulations to finally arrive at the (mostly) improvements presented in this manuscript. So, yes, the tuning was partly performed “by eye”, partly based on statistical measures (biases, correlations, etc. as presented in the manuscript). We clarified this by adding the following sentence. “The tuning was not based on an objective procedure, but relied on the authors’ experience and a series of “trial and error” simulations.”

Pg. 11, ln. 16: Is there a potential problem here because Si:P ratios in the real world are not constant?; in iron-limited HNLC regions, for instance, Si:P is typically elevated (since diatom cell cycle length is increased affording more time to uptake silicon); the global nature of the “solution” here to DIC/TA biases potentially promises trouble regionally

Indeed, we find that the surface Si distribution is less well reproduced than that of phosphate. In the revised manuscript we added a few sentences on the possible shortcomings in the parameterisation of the Si cycle in connection with elevated Si concentrations in the North Pacific. “We note that the pattern of elevated silicate concentrations in the North Pacific and Arctic is similar to the pattern of PO_4 concentration in the model. While this is in good agreement with observations for phosphate, observed maximum Si concentrations are 50% smaller than modeled maximum concentrations. This might indicate that our ecosystem model is not tuned well enough or that its structure is oversimplified with respect to silica cycling (e.g., fixed Si:P ratio, fixed constant sinking speed).” Although the fixed Si:P is a limitation of the model, we show that surface silica concentrations compare better to observations in model version 1.2, and that DIC/TA biases are significantly reduced by tuning of the fixed Si:P ratio.

Pg. 14, ln. 16: It would perhaps be helpful to illustrate to the reader the depth distribution of sinking particles, as well as remineralisation, produced by these different schemes (e.g. for the same quantity of POC at 100m, whats its fate down a static water column?)

If we understand the reviewer correctly, he/she asks for the general properties of the different schemes. However, in the case of KR02, the sinking speed depends on the available mass and numbers of particles as well as the strength of aggregation, which all is variable in space and time. Kriest and Oschlies (2008) present a general analysis and comparison of several schemes (among them KR02, STD, and Martin curve, which is mostly equivalent to WLIN). In the revised manuscript we add a panel to Fig. 23 showing the normalised POC flux with depth for the simulations presented in the manuscript. We also add a short sentence referring the reader to Kriest and Oschlies (2008). “For a more

detailed discussion of the KR02 scheme in comparison to the assumption of constant or linearly increasing sinking speed, we refer the reader to Kriest and Oschlies (2008).”

Pg. 16, ln. 10: “... a restoring time scale of 365 days (version 1) and 350 days (version 1.2) ...” any reason why this is 15 days shorter?; presumably a parameterisation oversight?

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.”

Pg. 17, ln. 13: “with only small trends of 0.00007, 0.021, and 0.048 Pg C yr⁻¹ century⁻¹ for Mv1, Mv1.2, and 15 Lv1.2, respectively” any explanation for why the longer duration simulations with v1.2 have markedly higher CO₂ trends than the shorter duration v1 model?; one might instinctively expect smaller values for longer simulations; perhaps plot up the net CO₂ flux with time?

The explanation for this is a smaller decadal to centennial internal variability in the Mv1 compared to Mv1.2 and Lv1.2, which we attribute to the details of the salinity relaxation. First, we apply a reduced salinity relaxation south of 40 degree in Mv1.2 and Lv1.2, and second, the balancing of the salinity restoring flux (introduced in model version 1.2 leads to a generally more saline surface ocean in Mv1.2 and Lv1.2, i.e. a less stable total water column. This can be clearly seen in Fig. 1 of this response, where we plot the last 200 years of the spin-up runs for the three configurations. Nevertheless, a trend of the order of 0.05 Pg C yr⁻¹ 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 (also addressing some questions raised by reviewer #2).

“The relatively large uptake of carbon in Mv1 at the end of the spin-up is due to a lower CaCO₃ 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.”

*Pg. 18, ln. 8: “Since the unbalanced salinity relaxation flux removes salt ...” - this sentence reads as if it is saying that the reduction in S in Mv1 is due to the relaxation flux being applied across *all* model configurations; I think this sentence and the preceding one should be combined into: “While all three configurations include salinity relaxation, this is not balanced in the case of Mv1, with the result that average salinity falls by 0.2 psu during the course of the integration.”*

Thank you for this suggestion, this has been adopted.

Pg. 18, ln. 14-16: Worth reporting these sorts of numbers in a table?

In the revised manuscript we added two tables. The first one reporting the total numbers for PP and export productions (Table 4), and a table summarising C-uptake and storage for the three model versions/configurations (Table 5).

Pg. 19, ln. 12: “... a long transient increase in strength for about 300 years ...” Any chance of including a plot of the AMOC strengths of the models from their spin-up phases?; not least to give some idea of interannual variability in the absence of interannual forcing variability (if any)

Pg. 20, ln. 15: Since the plot makes a point of examining the time-series of AMOC, would it be possible to present the RAPID estimates on the same plot?

We followed these two suggestions by (i) extending the plot of AMOC strength to include also the part of the historical time period that was forced with CORE normal year forcing (1762–1947), and (ii) by including the RAPID estimate for 2004–2014 (McCarthy et al., 2015) in the plot.

Pg. 20, ln. 19: “... the climatology of de Boyer Montégut et al. (2004)” This is calculated how?

Pg. 20, ln. 22: On a related point, are model MLDs comparable to the climatology?; e.g. could MLD be calculated for the model in the same way as its done for the observation-based climatology?; if this is already the case, please make this clear

De Boyer Montégut et al. (2004) provide three climatologies using three different thresh-

old criteria, a temperature criterion (0.2 degree), a density criterion (0.03 kg m^{-3}), and a variable density threshold criterion (corresponding to a 0.2 degree decrease in temperature at local temperature and salinity conditions). Our model uses a bulk mixed layer formulation, that is, the mixed layer extent is calculated based on energy considerations, and the mixed layer is not vertically resolved (except for an extra top model layer of 5 to 10 m thickness). Due to the lack of resolution it is not possible to use threshold criteria applied to model output to obtain an equivalent definition of MLD as used in observation based estimates. The best we can do is to compare to a climatology that uses a density criterion, since this is conceptually the most similar definition. In the original manuscript we compared our model results to the mean and the range of the three de Boyer Montégut et al. (2004) climatologies which is (as mentioned also by the reviewer in a comment further down) not the best way to represent uncertainty. For the revised manuscript we propose to compare modeled MLD to the density criterion MLD only. We revised Fig. 4 and the manuscript text accordingly.

“Seasonal cycles of modeled average bulk mixed layer depth (MLD) compared to an observation based MLD climatology (de Boyer Montégut et al., 2004) are shown for several regions in Fig. 4. The climatology uses a threshold criterion for density, that is, MLD is defined by the depth where density has increased by 0.03 kg m^{-3} relative to its near surface value. We note that the depth of the bulk mixed layer in our model is calculated based on energy gain and dissipation in the surface ocean, and that modelled and observed quantities are therefore not directly comparable. A MLD climatology based on a density criterion is nevertheless suitable for comparison with our model, since the criterion measures stratification directly.”

Pg. 21, ln. 12: “clearly indicates a too deep mixing” Grammar

We rephrased this sentence as follows. “...indicates a deeper than observed mixing compared to CFC-11 profiles from the GLODAP data base.”

Pg. 21, ln. 13-16: Does this have anything to do with sea-ice?; some models can exhibit large polynyas in the SO, with the result that mixing, and ventilation, can be extreme

We do not find large polynyas in the model runs presented here. We clarify this by adding the following sentence to the revised manuscript. “These high concentrations are mainly caused by large fluxes occurring during Antarctic winter north of the ice edge.”

Pg. 22, ln. 5-7: “We note that the Eppley-VGPM algorithm produces global PP estimates at about the mean value ...” This might well be true, but in my experience the spatial patterns of different estimates are wildly different, making the choice of such an “intermediate” product less clear

We deleted this sentence from the revised manuscript. As suggest by the reviewer in a comment further down, we have revised our comparison of PP with observations. We now use the average of three different products, namely VGPM, Eppley-VGPM, and CbPM (Behrenfeld and Falkowski, 1997; Westberry et al., 2008) provided by the Ocean Productivity website <http://www.science.oregonstate.edu/ocean.productivity/>. We plot

the mean and range of these estimates in Figs. 6, 7, and 8. Section 3.2 and the Appendix have been revised accordingly.

Pg. 22, ln. 12-14: “These large discrepancies are reduced in model version 1.2 . . .” Is it possible (e.g. via run models that are not shown here) to be sure that the improvements stem from the BGC changes as opposed to the physics changes?

Yes, we have done such tests, which show that the changes in the BGC module are indeed the cause for the (by far largest portion) of the differences in PP between model version 1 and 1.2. (see Fig. 2 of this response). We add this information to the revised manuscript as follows. “Sensitivity tests with the same physical model but both versions of the biogeochemistry module (not shown) indicate that changes in the physical fields between model versions do not contribute significantly to this result.”

Pg. 23, ln. 5-6: “This missing PP on the shelves . . .” You could make this clearer by calculating PP in open ocean areas only for VGPM and the model runs

Thank you for this suggestion, this has been done. We have excluded data from shelf regions in our analysis in the revised manuscript throughout Section 3.2 and in Figures 6, 7, and 8.

Pg. 23, ln. 18-26: Is it possible to determine a map of nutrient limitations from the model?; it might help diagnose another reason for differences between them

This is possible, see Fig. 3 of this response for the configuration Lv1.2. Broadly in agreement with observations nitrate is the limiting nutrient at low latitudes. Iron is the limiting nutrient in the southernmost Southern ocean, in the northernmost North Pacific and parts of the Arctic Ocean in the model. Our model has no iron limitation in the equatorial Pacific and the North Atlantic. Compared to results from nutrient fertilisation experiments compiled by Moore et al. (2013) the iron limited regions in our model appear to be either too small (Southern Ocean, North Pacific) or absent (equatorial Pacific, North Atlantic). Moreover, even if we remove iron limitation completely, model results do not change significantly, that is, in those regions that are iron limited, the limitation is not very strong. Therefore, we do not find that differences in the seasonal cycle of PP and PP patterns can be explained by differences in nutrient limitation. The major differences between the model versions stem from the re-tuning of the ecosystem parameterisation as pointed out in our response above (see also Fig. 2 of this response).

Pg. 24, section 3.3: BGC tracers are the core of the model, while production is just one process within the model; Id suggest swapping the sections around and making this 3.2; production could come just before export - which is arguably more natural anyway

We agree with the reviewer that the order “tracer-production-export” would be more natural from a conceptual viewpoint. However, we describe differences between model versions and configurations here, and the differences in PP is at the origin of much of the differences in the tracer distribution between Mv1, Mv1.2 and Lv1.2. We refer back to the results of section 3.2 frequently in section 3.3, and we think that the manuscript

is much easier to understand if organised in this order. We therefore would prefer to leave the order of sections unchanged.

Pg. 24, section 3.3: Why no chlorophyll?; is this because the fixed chl:C ratio here causes problems?

We do not analyse chlorophyll here because it is not of major importance in our NPZD-ecosystem model. The chlorophyll concentration (calculated by the fixed C:Chl ratio mentioned by the reviewer) is only used in the model to calculate light absorption through phytoplankton. A more realistic variable C:Chl ratio could possibly improve the simulation of light availability, but there is a couple of improvements we think will have a larger impact (e.g. improved vertical resolution in the mixed layer which is currently rather low).

Pg. 25, ln. 16: “Moreover, nitrogen fixation, . . .” – A map of this perhaps?

Pg. 25, ln. 17: “. . . which occurs in the surface ocean as soon as $[\text{NO}_3] < R_{\text{N:P}}[\text{PO}_4]$, . . .” Does this mean that all N_2 -fixation occurs in the right place?; cf. ostensible temperature limits, etc.

We added a map (Fig. 13 of the revised manuscript) showing the difference $[\text{PO}_4] - [\text{NO}_3]/16$ and rewrote this paragraph as follows. “The modelled distributions of nitrate and phosphate are similar in terms of biases and general spatial structure since the model uses a fixed stoichiometric ratio ($R_{\text{N:P}} = 16$) for the composition of organic matter. The difference $[\text{PO}_4] - [\text{NO}_3]/R_{\text{N:P}}$ (Fig. 13) is positive everywhere in broad agreement with observations, indicating that nitrate is depleted relative to phosphate with respect to the canonical N:P ratio of 16. Large values of this difference are found in the tropical Pacific, in the model as well as in observations, but more pronounced in the model. This pattern is due to the tropical Pacific oxygen minimum zone (OMZ) where NO_3 is consumed during denitrification to remineralise organic matter and release PO_4 . The oxygen minimum zones are excessively large in our model, particularly in the tropical Pacific (see Sec. 3.3.2). These results show, that the simple parameterisation of nitrogen fixation, which occurs in the surface ocean as soon as $[\text{NO}_3] < R_{\text{N:P}}[\text{PO}_4]$ in our model, is active over the whole surface ocean, which is probably unrealistic. This simple parameterisation should be viewed more as a means to keep the model ocean close to the assumed stoichiometric N:P ratio than a realistic parameterisation of nitrogen fixation. At depth, we find major deviations from the similarity of nitrate and phosphate distributions in the OMZ of the tropical Pacific. Here, our model shows a local minimum of nitrate (due to the too strong denitrification caused by too low oxygen values) instead of a local maximum as observed and as it is found for phosphate.”

Pg. 26, ln. 2-6: What do the distributions of biogenic opal and CaCO_3 export look like?; and how do they compare to observationally-derived estimates (e.g. in total)?

We have added maps of POC export production as well as maps of CaCO_3 to organic carbon and opal to organic carbon in exported matter (Fig. 9 of the revised manuscript).

We also added a new table summarising the key numbers (PP, exports) for all model versions. Section 3.2 has been expanded by a new subsection which reads as follows.

“3.2.1 Export production of POC, CaCO₃, and opal

Figure 9 shows the mean export production of POC, as well as the CaCO₃ and opal to organic carbon ratios in exported matter. The spatial pattern of POC export closely resembles the pattern of PP, since the fraction of PP that is exported as POC (about 15–25 %, not shown) shows only small spatial variations. Total annual carbon export (averaged over 2003 to 2012, Table 4) is 8.8 (Mv1), 7.1 (Mv1.2), and 6.1 Pg C yr⁻¹ (Lv1.2).

Since the partitioning between opal and CaCO₃ export production is parameterised dependent on available silicate in our model, CaCO₃ export dominates over opal export only in regions depleted of surface silicate, which are the subtropical gyres in the Atlantic and the south Pacific. Due to the increased opal to phosphorus uptake ratio $R_{\text{Si:P}}$ in model version 1.2 (clearly visible in Fig. 9 g to i) the CaCO₃ production is maintained or even slightly expanded into the western North Pacific despite the much lower PP (and surface silicate consumption) at high latitudes in this model version. We note that the simple parameterisation of opal and CaCO₃ export production is qualitatively supported by opal to particulate inorganic carbon (PIC) ratios derived from sediment traps. For example, Honjo et al. (2008, see their Fig. 7) show that high opal/PIC ratios are constrained to ocean regions with high surface silica concentrations, a pattern that is qualitatively reproduced by our model.

Total modelled opal export (between 95 and 120.5 Tmol Si yr⁻¹, Table 4) is within the uncertainty range of the estimate of 105 ± 17 Tmol Si yr⁻¹ given by Tréguer and De La Rocha (2013). The ratio of CaCO₃ export to organic carbon export of 7.1 to 7.9 % is within the range estimated by Sarmiento et al. (2002, 6 ± 3 %).”

Pg. 27, ln. 15: While iron isnt quite at the stage of having a global observational climatology, Geotraces has some fields that might help; and even in the absence of an observational comparison, it might be helpful to compare the models to elucidate differences; irons reach is longer than simply total iron concentration

The iron cycle in our model is rather simplistic. The spatial pattern of iron in the surface mainly reflects the aeolian input and some upwelling of iron in the Southern Ocean. The maximum iron concentration is determined by relaxation towards a value of $0.6 \mu\text{mol m}^{-3}$ (when modelled iron is larger than this value) to mimic the process of complexation with ligands. Therefore, below 700 m, the iron concentration is constant at $0.6 \mu\text{mol m}^{-3}$ everywhere. As mentioned above, the iron limitation in our model is not very strong, and the parameterisation of the iron cycling is definitely something to be improved in the future. For the current model version presented here, we do not think that it makes much sense to discuss the iron cycle in detail.

Rather, we propose to amend the model description as follows. “A fraction of the dust deposition (1 %) is assumed to be iron, and part of it is immediately dissolved and available for biological production. To mimic the process of complexation with ligands,

iron concentration is relaxed towards a value of $0.6 \mu\text{mol m}^{-3}$ when modelled iron is larger than this value. We note that this parameterisation of the iron cycle is rather simplistic. The spatial pattern of iron in the surface mainly reflects the aeolian input and upwelling of iron in the Southern Ocean. At depth, iron concentration is determined by accumulation of remineralised iron and the assumed complexation. Therefore, the iron concentration approaches a constant value of $0.6 \mu\text{mol m}^{-3}$ at depth larger than approximately 700 m. We find that the resulting iron limitation in our model is rather weak, and given these limitations we will not focus on the iron cycle in this manuscript.” In the conclusions, we also mention the iron cycle (together with nitrogen fixation) as parameterisations to be improved in the future. “There are several directions for future model development that can be identified from our results. The parameterisations of iron cycling and nitrogen fixation are simplistic and should be improved for future model versions, particularly since the observational basis for both processes has improved considerably in recent years. ”

Pg. 28, ln. 22: “. . . results in increased CaCO₃ production and a considerable reduction of the alkalinity and DIC biases . . . ” Can this be squared with any observational evidence?; it can certainly be squared with model evidence (cf. Kwiatkowski et al., 2014; here, a number of models have low CaCO₃ production in the tropics and excessive alkalinity and DIC)

The simple parameterisation used in HAMOCC to set opal and CaCO₃ content of exported material is qualitatively supported by opal to PIC ratios derived from sediment traps (e.g. Honjo et al., 2008, Fig. 7). This information has been added to the new section on export production (Sec. 3.2.1 of the revised manuscript, see above).

We have no direct observational support for improved CaCO₃ export in model version 1.2, but rather use the indirect evidence of improved surface alkalinity. We propose to amend the manuscript as follows.

“Since modelled alkalinity in low latitudes is quite sensitive to CaCO₃ export (see also Kwiatkowski et al., 2014), we take the reduction of alkalinity biases as indirect evidence for an improved distribution of CaCO₃ export in Mv1.2 and Lv1.2.”

Pg. 30, ln. 25: “. . . the GLODAP data base” The Khatiwala et al. (2013) estimate of anthropogenic CO₂ is probably a better estimate

Yes, we agree. We added a comparison of the Khatiwala et al. (2013) estimate with our model results as follows. “Compared to a recent synthesis of anthropogenic carbon storage estimates (Khatiwala et al., 2013, 155 ± 31 Pg C for the year 2010) modelled DIC_{ant} storage of 186 (Mv1), 175 (Mv1.2), and 159 (Lv1.2) Pg C is higher, but still within the uncertainty range of the synthesis estimate.”

Pg. 31, ln. 16: As mentioned previously, having a figure that illustrates how each of these schemes remineralises organic matter down the water column would also be helpful (or a plot of how the OM is attenuated)

In the revised manuscript we added a new panel to Fig. 23 showing the average POC fluxes (normalised to the POC flux at 100 m depth). “The global average sinking speed profiles for the four experiments and the resulting POC fluxes (normalised to the flux at 100 m depth) are shown in Fig. 23a and b, respectively.”

Pg. 31, ln. 28: “At the end of the spin-up runs ...” How do these fit with the long spin-ups already done?; also, are these long enough to approach equilibrium?; or is the assumption that they are long enough for only transient drift to remain?

These are not additional spin-up runs. Each model configuration is spun-up for 1000 years. For the comparison of the different POC sinking schemes we use data from the years 1001–1010. The model is generally not in full equilibrium after 1000 years but there is only a small drift remaining. We clarify this by rewording the sentence as follows. “At the end of the spin-up runs (years 1001–1010) we find As mentioned above, the model is generally not in full equilibrium after 1000 years, but the remaining drift is small.”

Pg. 33-34: I dont know the answer myself, so its perhaps cheeky to ask, but could the authors comment on whether direct POC flux measurements or indirect AOU (or other tracer) measurements better constrain export and remineralisation; there may be no good answer at the moment, so the authors use of both is probably best

We added a short discussion on the advantages/disadvantages of using direct/indirect methods to estimate POC-fluxes as follows. “We have based the evaluation of the different POC sinking schemes on indirect ($\text{PO}_4^{\text{remin}}$) and direct (sediment trap) measurements. While the indirect method has inherent inaccuracies related to the calculation of AOU and the assumed stoichiometry of remineralisation, the direct measurement of POC fluxes by sediment traps also comes with large systematic uncertainties (see Honjo et al., 2008, and references therein). It is difficult to decide whether one of the two methods provides a more reliable evaluation and we therefore use both approaches here.”

Pg. 35, ln. 16: A more general comment some studies (e.g. Kwon et al., 2009; Kriest, Oschlies & Khatiwala, 2012) examine the tuning of such models of export, whereas the manuscript uses them as is; while the authors do mention alternative sinking velocities for STD-fast at one point, they could help here by drawing further attention to this and / or commenting on the tuning of such models (e.g. if they have any unreported experience on the success or otherwise of this)

We do not have any further experience with tuning the POC sinking scheme for our model. The computational constraints do not allow for much more than the runs presented in this study. We added a comment on this as follows. “We finally note that a comprehensive sensitivity analysis or a rigorous tuning of the different POC sinking schemes would require accelerated off-line integration techniques as applied by e.g. Kwon et al. (2009) or Kriest et al. (2012), which we to date have not available for our model.”

Pg. 36, ln. 10: So, paradoxically, excessively high and excessively low O2?

Yes, exactly. We made this a bit clearer by rephrasing this sentence as follows. “In strong contrast to these positive O₂ biases, the model develops too large oxygen minimum zones along with a too strong accumulation of remineralised phosphate in the tropical oceans, particularly in the Pacific.”

Pg. 37, ln. 18: “Part of this problem is the distribution of primary production which is too high in a narrow band along the upwelling ...” Is this in any way related to the model being an isopycnal model?

We do not have any evidence that this could be related to fact that the model is isopycnic. There is a CORE intercomparison paper on Pacific circulation in preparation (Mats Bentsen, personal communication), and our model (as well as the other isopycnal models) do perform as good (or bad) as other models in the equatorial Pacific. Since this work is at an early stage however, we do not have a citable reference for this statement.

Pg. 38, ln. 1: “In the Southern Ocean ...” Since this paragraph deals with the model as it is, rather than - per the preceding paragraph - the model as it might become, it should precede the paragraph on shelf improvements

This paragraph deals with the problems that need to be addressed in the Southern Ocean to improve our model, and as such it also deals with the model as it might become. Since the paragraph concludes with some rather general statements, we would prefer to leave it at this position.

Pg. 38, ln. 9: It’s a wholly personal preference, but I think papers finish better with a short, bulletpointed list of the main points / findings

Since this comment is flagged as a personal preference, we hope that it is allowed to respectfully disagree here. We have tried to write a concise ‘summary and conclusions’ section, which is clearly structured as we believe. We do not immediately see the point in repeating our main findings as a bullet-point list, and would therefore prefer to leave the end of the manuscript unchanged.

Pg. 38, ln. 9: A general criticism I’d make of the manuscript’s validation of the model is that the performance of the model is not properly put within the context of similar models; the CMIP5 archive, for instance, offers a range of similar resolution models that could profitably provide such context; most of which arent isopycnal models

Please see our response the general comment at the start of this document.

Table 1: “ kmol m^{-3} ” At the risk of both being a pedant and missing the wood for the trees, presumably kmol/m^3 is being used here because it is equivalent to molar units (i.e. mol/l); if so, why not just use mol/l ?

For some reason kmol m^{-3} is the “traditional” HAMOCC-unit. It is used in the model descriptions Maier-Reimer et al. (2005) and Ilyina et al. (2013), and therefore we think it might be reasonable to stick to this unit in this manuscript, too.

Table 1: “Laughing gas” While - appropriately enough - I laughed when I saw this, I dont think it can be called this in the final manuscript; nitrous oxide, perhaps?

This has been corrected.

Table 2: “Fraction of grazing egested $1 - \varepsilon_{zoo}$ ” This is a little bit confusing; is the symbol really “ $1 - \varepsilon_{zoo}$ ”?; why not “ ε_{zoo} ” and give it a value of 0.2 or 0.1?

Again, this is a traditional HAMOCC formulation. It has been used by Six and Maier-Reimer (1996), who implemented the ecosystem parameterisation and later on by Maier-Reimer et al. (2005) and Ilyina et al. (2013). We therefore propose to stick to this way of expressing this parameter.

Figure 1: Sometimes putting the y-axis on a logarithmic scale is helpful for showing whats happening near-surface

We changed the vertical scale of the zonal mean plots throughout the manuscript to enlarge the upper 1000 m. In the revised plots the 0–1000 m range occupies about 1/3 of the vertical scale. This solution avoids potential problems of a log scale while still emphasizing the fields at the surface.

Figure 1: Since the colour map includes white, these panels could do with having the seafloor drawn on; that would help separate places that have zero difference from those that are rock; also, this would help clarify the bathymetry differences between different model grids

The background colour of the zonal mean plots has been changed to grey throughout the manuscript, such that sea-floor can be distinguished from regions with zero differences (white colour).

Figure 3: What happens with AMOC during the long spin-ups?

Figure 3: “ 10^9 kg s^{-1} ” Convert to Sverdrups?; or is this awkward for an isopycnal model?

Figure 3: Observational data from RAPID appropriate for comparison?

For the revised manuscript, we updated Figure 3. We converted units to Sverdrup and an observational estimate from RAPID has been included. Also, we have extended the plot to include the years 1762–1947, which are forced by the CORE normal year forcing. We mention in the main text that during the spin-up “the AMOC shows a long transient increase in strength for about 300 years before stabilising”.

Figure 4: “The range given for the observation based estimates is solely due to different criteria used to define MLD and not due to other uncertainties” Per a previous comment, how does the MLD method used for the models compare to that of the observations?; also, given how variable different MLD methods can be from one another, reporting uncertainty in this way here seems potentially risky

Please see our detailed response to the previous comment above. We have revised the

figure and compare our model to one climatology only. This climatology is based on a density criterion, which is conceptually most similar to the MLDs calculated by our model. The sentence about the range of observation based estimates has been deleted from the revised manuscript.

Figure 5: Again, a log-scale y-axis might help here; most of the structure here is in the upper water column

Figure 5: Rotate panel e so that its y-axis is aligned with the x-axes on the panels to the left?; i.e. 90S to the left, 90N to the right

The vertical scale of this figure has been changed as explained above. Panel e has been rotated as requested.

Figure 6: Its a weakness on my part, but I prefer my plots to omit unnecessary grid lines (and have coastlines if possible)

Figure 6: Rotation would make panel e easier to understand (though I appreciate it would not then be aligned as in Figure 5)

We added coast lines to all 2d-plots throughout the revised manuscript. We also rotated panel e as requested. Since we do find the grid lines very helpful to read and interpret this figure, we propose to leave them in the revised plots.

Figure 6: Rather than only use VGPM, you could average VGPM with other estimates; its a poor way of simplifying the diversity in observational estimates of PP, but it can be useful given their spread, and its not without precedent

As explained above, we consider three estimates obtained by different algorithms in the revised manuscript. We plot the mean of the three estimates in panel d), and the mean and range of zonal means in panel e).

Figure 7: Why 40 in one hemisphere and 60 in the other?

Expressed in PgC yr^{-1} , the production north of 60N is tiny, and the difference between model versions and observations is small. To some degree this is also true in the southern hemisphere, but there is a notable difference between model version 1 and 1.2, so we split the southern hemisphere into one additional subregions. For the northern hemisphere there is no added value in doing so, therefore we propose to leave Fig. 7 unchanged in this respect.

Figure 7: Again, why just use VGPM?

See our answer above. We include the mean and range of three different observational estimates in the revised figure.

Figure 8: In panel c, including the Indian Ocean is complicated by the presence of the monsoon

In the revised manuscript we change Figure 8c to show the tropical Indian Ocean, and

revised the manuscript text as follows. “In the tropics outside the Indian Ocean there is no significant annual cycle in the model as well as in observations. The seasonal variation in PP caused by the monsoon in the Indian Ocean is well captured by the model although there appears to be a small phase shift of about one month relative to the MODIS based PP estimates (Fig. 8c).”

Figure 9: Rotate panel e again please

This has been done.

Figure 10: Fewer colours in the colour maps here (especially for the delta plots) might make it easier to discern patterns in match-mismatch (e.g. the reds are quite homogeneous)

In the revised version of the manuscript we reduced the number of colours in the colourmap to 16 (or 32 for difference plots) throughout the manuscript.

Figure 12: These panels hint at some odd ventilation feature that elevates N. Pac. silicic acid (which then bleeds into the Arctic); does the same appear in CFC-11?

CFC-11 does not show a suspicious ventilation feature here. Phosphate shows similar elevated surface values in the North Pacific and Arctic, but for phosphate this is also seen in the observations. We conclude that this might be a problem of our ecosystem parameterisation (fixed stoichiometric ratios, fixed constant sinking speed and/or remineralisation rate of opal, to name a few possible sources). To discuss this, we amended the manuscript text as follows. “We note that the pattern of elevated Si values in the North Pacific and Arctic is similar to the pattern of PO_4 concentration in the model. While this is in good agreement with observations for phosphate, observed maximum Si concentrations are 50% smaller than modelled maximum concentrations. This might indicate that our ecosystem model is not tuned well enough or that its structure is oversimplified with respect to silica cycling (e.g., fixed Si:P ratio, fixed constant sinking speed).”

Figure 13: Add a key if possible; also, different symbols might help with the plot (especially for colour blind readers)

This has been done.

Figure 13: Amend to Panel (d) shows results for the 500m depth level with error prone grid points located in the tropics (between 20 and 20S) omitted from the analysis?

This has been done.

Figure 14: Why not go entirely east-west in the Pacific here?; since oxygen is particularly low in the East Pacific, this could be important

We followed this suggestion and plotted a zonal mean O_2 concentration averaged over the entire eastern Pacific. This does not change the general features of the model-observation

comparison (it makes our model look a bit worse though).

Figure 16: Nicely improved!

Thank you.

Figure 17: Uncertainty from Takahashi (and / or other $p\text{CO}_2$ products)?

We included $p\text{CO}_2$ from the Takahashi et al. (2009) data product in the plot.

Figure 18: There's quite a bit of a dip in the observationally-derived fluxes; is its origin explained in the main text?

Yes, this is explained in the main text. The dip in the data was worth a Science publication. "From the early 1990s until 2001 the Landschützer et al. (2015) data shows a marked decline followed by a steep increase of carbon uptake, which they attribute to a saturation followed by a reinvigoration of the Southern Ocean carbon sink."

Figure 19: averaged over the years 1990 to 1998 Why not a full decade?; being unreasonably suspicious, such odd ranges always raise my eyebrows

Actually, the time averaging is unnecessary here anyway. Anthropogenic carbon as a time integrated quantity does not need any averaging to arrive at a climatology. The nominal year for the GLODAP anthropogenic carbon product is 1994. We therefore plotted the corresponding model values for 1994 in the revised plot.

Figure 20: The mismatched sizes of bars shown in panel b seem a bad idea; log scale again?

We added a second y-axis on the right hand side of the revised plot which has an appropriate range for the fluxes at 2000 m and at the bottom.

Figure 21: Why zero in the Arctic of panels a, c, e and g?; does the main text say?

Figure 21: A global integrated profile plot of a, c, e and g might be helpful

We included a figure with global mean profiles of remineralised PO_4 , and added an explanation for the zero remineralised PO_4 in the Arctic as follows. "The absence of significant amounts of remineralised phosphate in the Arctic basin in our model (Fig. 23) is due to a combination of low PP and too strong ventilation (positive O_2 biases, compare Fig. 16)."

Figure 22: Bigger dots?; also, a key would be nice; perhaps a little map that shows the regions in the appropriate colour?

It is difficult to use much bigger dots, since then they would overlap too much. We increased the size of dots a bit for the revised figure. A map showing the colors of each region has been added.

Figures

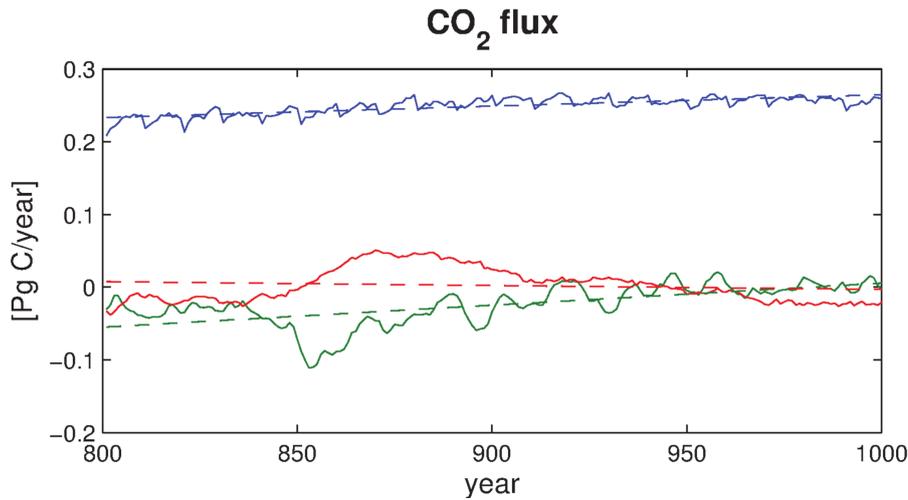


Figure 1: Air-sea CO_2 flux during the last 200 years of the spin up runs for Mv1 (blue line) Mv1.2 (green line), and Lv1.2 (red line). The dotted lines give a linear fit through the respective time series. Note that for Mv1 the spin-up was only 800 years, and the actual years displayed are 600–800.

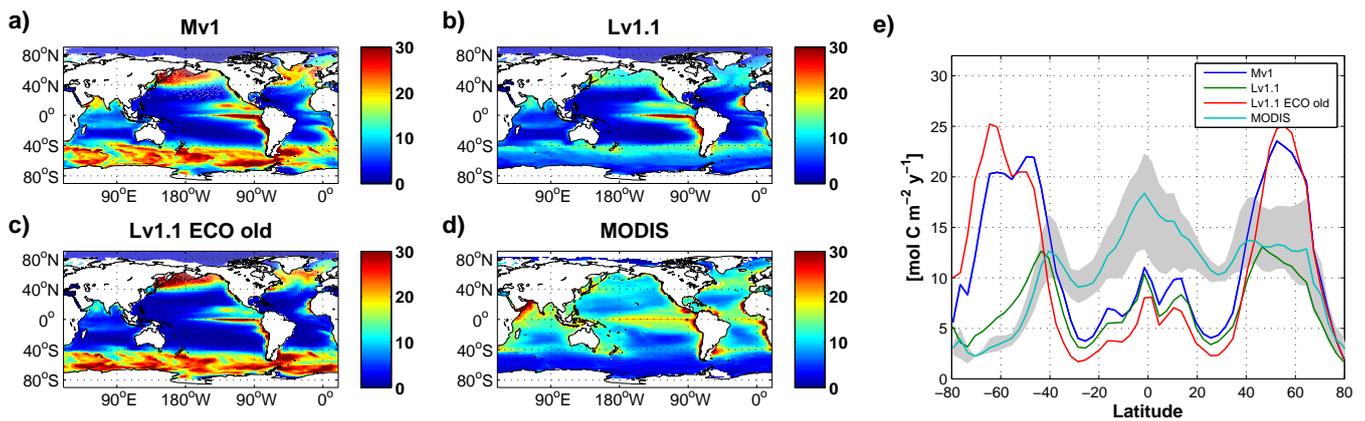


Figure 2: Vertically integrated primary production ($\text{mol C m}^{-2} \text{yr}^{-1}$) averaged over the years 2003–2012 for the model configurations (a) Mv1, (b) Lv1.1 (an intermediate model version very similar to Lv1.2), (c) Lv1.1 but with the ecosystem and light penetration formulation of Mv1, and (d) the mean of three satellite based climatologies (derived from MODIS retrievals). Panel (e) displays the zonal means of each field presented in panels (a–d). The grey shaded area represents the range of the zonal means of the three satellite based climatologies.

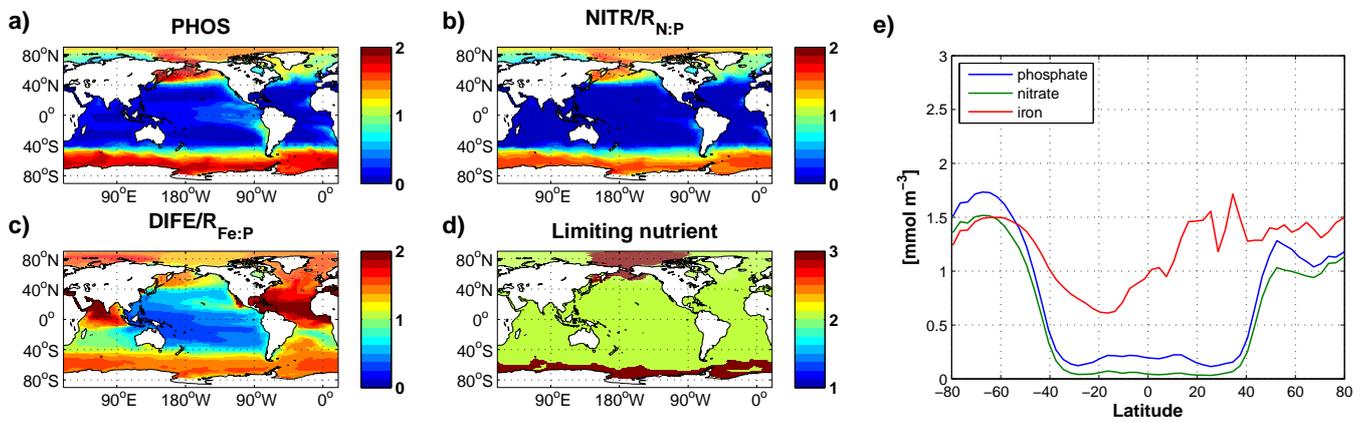


Figure 3: (a) Surface phosphate concentration (mmol m^{-3}), (b) surface nitrate concentration divided by $R_{N:P} = 16$, (c) surface iron concentration divided by $R_{Fe:P} = 366 \times 10^{-6}$, and (d) limiting nutrient (1=phosphate, 2=nitrate, 3=iron) for the model configuration Mv1. Panel (e) displays the zonal means of each field presented in panels (a–c).

References

- Aumont, O. and Bopp, L.: Globalizing results from ocean in situ iron fertilization studies, *Global Biogeochem. Cyc.*, 20, doi:10.1029/2005GB002591, 2006.
- Behrenfeld, M. and Falkowski, P.: Photosynthetic rates derived from satellite-based chlorophyll concentration, *Limnol. Oceanogr.*, 42, 1–20, 1997.
- Buitenhuis, E. T., Rivkin, R. B., Saille, S. F., and Le Quéré, C.: Biogeochemical fluxes through microzooplankton, *Global Biogeochem. Cyc.*, 24, Gb4015, doi:10.1029/2009gb003601, 2010.
- de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D.: Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology, *J. Geophys. Res.* 109, C12003, doi:10.1029/2004JC002378, 2004.
- Doney, S. C., Lima, I., Feely, R. A., Glover, D. M., Lindsay, K., Mahowald, N., Moore, J. K., and Wanninkhof, R.: Mechanisms governing interannual variability in upper-ocean inorganic carbon system and air-sea CO₂ fluxes: Physical climate and atmospheric dust, *Deep-Sea Res. Pt. II*, 56, 640–655, 2009.
- Gröger, M. and Mikolajewicz, U.: Note on the CO₂ air-sea gas exchange at high temperatures, *Ocean Model.*, 39, 284–290, 2011.
- Hauck, J., Völker, C., Wang, T., Hoppema, M., Losch, M., and Wolf-Gladrow, D. A.: Seasonally different carbon flux changes in the Southern Ocean in response to the southern annular mode, *Global Biogeochem. Cyc.*, 27, 1236–1245, 2013.
- Honjo, S., Manganini, S. J., Krishfield, R. A., and Francois, R.: Particulate organic carbon fluxes to the ocean interior and factors controlling the biological pump: A synthesis of global sediment trap programs since 1983, *Prog. Oceanogr.*, 76, 217–285, doi:10.1016/j.pocean.2007.11.003, 2008.
- Ilyina, T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H., and Núñez Riboni, I.: Global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth System Model in different CMIP5 experimental realizations, *J. Adv. Model. Earth Syst.*, 5, 287–315, doi:10.1029/2012MS000178, 2013.
- Khatiwala, S., Tanhua, T., Fletcher, S. M., Gerber, M., Doney, S. C., Graven, H. D., Gruber, N., McKinley, G. A., Murata, A., Ríos, A. F., and Sabine, C. L.: Global ocean storage of anthropogenic carbon, *Biogeosciences*, 10, 2169–2191, doi:10.5194/bg-10-2169-2013, 2013.
- Kriest, I. and Oschlies, A.: On the treatment of particulate organic matter sinking in large-scale models of marine biogeochemical cycles, *Biogeosciences*, 5, 55–72, 2008.

- Kriest, I., Oschlies, A., and Khatiwala, S.: Sensitivity analysis of simple global marine biogeochemical models, *Global Biogeochem. Cyc.*, 26, GB2029, doi:10.1029/2011GB004072, 2012.
- Kwiatkowski, L., Yool, A., Allen, J. I., Anderson, T. R., Barciela, R., Buitenhuis, E. T., Butenschön, M., Enright, C., Halloran, P. R., Le Quéré, C., de Mora, L., Racault, M.-F., Sinha, B., Totterdell, I. J., and Cox, P. M.: iMarNet: an ocean biogeochemistry model intercomparison project within a common physical ocean modelling framework, *Biogeosciences*, 11, 7291–7304, doi:10.5194/bg-11-7291-2014, 2014.
- Kwon, E. Y., Primeau, F., and Sarmiento, J. L.: The impact of remineralization depth on the air-sea carbon balance, *Nat. Geosci.*, 2, 630–635, doi:10.1038/NGEO612, 2009.
- Landschützer, P., Gruber, N., and Bakker, D. C. E.: A 30 years observation-based global monthly gridded sea surface $p\text{CO}_2$ product from 1982 through 2011, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee, doi:10.3334/CDIAC/OTG.SPCO2_1982_2011_ETH_SOM-FFN, URL http://cdiac.ornl.gov/ftp/oceans/SPCO2_1982_2011_ETH_SOM_FFN, 2015.
- Le Quéré, C., Moriarty, R., Andrew, R. M., Peters, G. P., Ciais, P., Friedlingstein, P., Jones, S. D., Sitch, S., Tans, P., Arneeth, A., Boden, T. A., Bopp, L., Bozec, Y., Canadell, J. G., Chini, L. P., Chevallier, F., Cosca, C. E., Harris, I., Hoppema, M., Houghton, R. A., House, J. I., Jain, A. K., Johannessen, T., Kato, E., Keeling, R. F., Kitidis, V., Klein Goldewijk, K., Koven, C., Landa, C. S., Landschützer, P., Lenton, A., Lima, I. D., Marland, G., Mathis, J. T., Metzl, N., Nojiri, Y., Olsen, A., Ono, T., Peng, S., Peters, W., Pfeil, B., Poulter, B., Raupach, M. R., Regnier, P., Rödenbeck, C., Saito, S., Salisbury, J. E., Schuster, U., Schwinger, J., Séférian, R., Segschneider, J., Steinhoff, T., Stocker, B. D., Sutton, A. J., Takahashi, T., Tilbrook, B., van der Werf, G. R., Viovy, N., Wang, Y.-P., Wanninkhof, R., Wiltshire, A., and Zeng, N.: Global carbon budget 2014, *Earth Syst. Sci. Data*, 7, 47–85, doi:10.5194/essd-7-47-2015, 2015.
- Maier-Reimer, E., Kriest, I., Segschneider, J., and Wetzol, P.: The Hamburg Oceanic Carbon Cycle Circulation model HAMOCC5.1 — Technical Description Release 1.1, Reports on Earth System Science 14, Max Planck Institute for Meteorology, Hamburg, Germany, 2005.
- McCarthy, G., Smeed, D., Johns, W., Frajka-Williams, E., Moat, B., Rayner, D., Baringer, M., Meinen, C., Collins, J., and Bryden, H.: Measuring the Atlantic Meridional Overturning Circulation at 26°N, *Prog. Oceanogr.*, 130, 91–111, doi:10.1016/j.pocean.2014.10.006, 2015.
- Moore, C. M., Mills, M. M., Arrigo, K. R., Berman-Frank, I., Bopp, L., Boyd, P. W., E. D. Galbraith, R. J. G., Guieu, C., Jaccard, S. L., Jickells, T. D., La Roche, J., Lenton, T. M., Mahowald, N. M., Marañón, E., Marinov, I., Moore, J. K., Nakatsuka,

- T., Oschlies, A., Saito, M. A., Thingstad, T. F., Tsuda, A., and Ulloa, O.: Processes and patterns of oceanic nutrient limitation, *Nat. Geosci.*, 6, 701–710, doi:10.1038/NNGEO1765, 2013.
- Oke, P. R., Griffin, D. A., Schiller, A., Matear, R. J., Fiedler, R., Mansbridge, J., Lenton, A., Cahill, M., Chamberlain, M. A., and Ridgway, K.: Evaluation of a near-global eddy-resolving ocean model, *Geosci. Model Dev.*, 6, 591–615, doi:10.5194/gmd-6-591-2013, 2013.
- Sarmiento, J. L., Dunne, J., Gnanadesikan, A., Key, R. M., Matsumoto, K., and Slater, R.: A new estimate of the CaCO₃ to organic carbon export ratio, *Global Biogeochem. Cyc.*, 16, 1107, doi:10.1029/2002GB001919, 2002.
- Séférian, R., L., B., Gehlen, M., Orr, J., Ethé, C., Cadule, P. and Aumont, O., Salas y Mélia, D., Voldoire, A., and Madec, G.: Skill assessment of three earth system models with common marine biogeochemistry, *Clim. Dynam.*, 40, 2549–2573, 2013.
- Six, K. D. and Maier-Reimer, E.: Effects of plankton dynamics on seasonal carbon fluxes in an ocean general circulation model, *Global Biogeochem. Cyc.*, 10, 559–583, 1996.
- Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Bakker, D. C. E., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Koertzing, A., Steinhoff, T., Hoppema, M., Olafsson, J., Arnarson, T. S., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Wong, C. S., Delille, B., Bates, N. R., and de Baar, H. J. W.: Climatological mean and decadal change in surface ocean pCO₂, and net sea-air CO₂ flux over the global oceans, *Deep-Sea Res. II*, 56, 554–577, doi:10.1016/j.dsr2.2008.12.009, 2009.
- Tréguer, P. J. and De La Rocha, C. L.: The World Ocean Silica Cycle, *Annu. Rev. Mar. Sci.*, 5, 477–501, doi:10.1146/annurev-marine-121211-172346, 2013.
- Wanninkhof, R.: Relationship between wind speed and gas exchange over the ocean revisited, *Limnol. Oceanogr. Meth.*, 12, 351–362, 2014.
- Westberry, T., Behrenfeld, M. J., Siegel, D. A., and Boss, E.: Carbon-based primary productivity modeling with vertically resolved photoacclimation, *Global Biogeochem. Cyc.*, 22, doi:10.1029/2007GB003078, 2008.