Responses to the referees and changes to the manuscript

We want to thank the two referees for their helpful and constructive reviews, which have greatly improved the manuscript. Below please find our responses to all of your points. The track-changes (latexdiff) version of the manuscript follows at the end of this pdf.

Dear Sakina-Dorothée,

We like to thank your for the effort and the time spent for a careful and positive review. Your comments are much appreciated, as they are constructive and helpful. We listed our responses to all of your points below and hope the manuscript is now satisfactory.

Remarks on the Methods

Equations 1 and 2: *why don't you also test the sensitivity of the model to the parameters of these two equations? Please justify it in the text.*

Reply: These two temperature configurations are empirical functions, directly simulating expected or observed temperature dependencies of N_2 fixation. One of the main goals of this study is to compare a more realisitic temperature dependency (OPEM-H) to the default function used in the UVic (OPEM). Thus, our study compares two different model configurations. The structural difference becomes ultimately fixed by the temperature function employed, which includes the values assigned to the respective parameters. In addition, our focus is on variations in physiological parameters, but the parameters of the two different temperature equations have no clear physiological meaning. They determine (define) the two model configurations exclusively. We explain this now in the text on p. 4, lines 94–103: *"Both of these equations are empirical functions directly simulating expected or observed temperature dependencies of N_2 fixation. We consider Eq. (2) more realistic and hence analyse its effect on model behaviour. However, since the parameters in these two equations have no clearly identifiable physiological meaning, we consider a sensitivity analysis of the parameters in Eqs. (1) and (2) beyond the scope of the present study. Note that some models do not enforce any temperature limitation on nitrogen fixation (e.g., Dunne et al., 2012; Ilyina et al., 2013; Jickells et al., 2017). In the present ocean, waters colder than about 15 °C are generally replete with fixed inorganic nitrogen. For existing parameterisations of N_2 fixation, which are functions of the nitrate deficit with respect to phosphate, there has been little indication of substantial impacts of the formulation of temperature control at low temperatures on the distribution of nitrogen fixation (Somes and Oschlies, 2015; Landolfi et al., 2017). Such differences in formulation may, however, gain importance in environmental conditions different from today's."*.

Table 1: *More clearly indicate in the legend of Table 1 that the identified "best" values for trade-off simulations were the same for the two model configurations OPEM and OPEM-H. It is indicated in the text later, but it has to be clearly mentioned here for the reader.*

Reply: We have added *"Note that the trade-off simulations share the same parameter combination."* to the table caption.

Lines 102–107: *I am not entirely convinced by the arguments given here to justify why the parallel setup is better than systematic calibration approaches. Indeed, one can imagine a systematic calibration where X values are systematically tested for the 13 parameters. In that case, I do not see how this would lead to individual model simulations that would depend on other/previous combinations of parameters, neither how it would prevent re-evaluation with different metrics. However, the first item would be true for a parameterisation based microgenetic algorithm for instance. The authors may need to rephrase this sentence to make it accurate (e.g., by replacing the term "systematic" by an other one).*

Reply: We agree, by "systematic" we originally thought of a typical path-dependent minimization (or maximization) in the parameter-cost function manifold, which is an iterative process. We replaced *"systematic"* with *"iterative"*, now on p. 4, line 113.

Remarks on the Results

Table 2: *Could you add a column with observation values, at least for the depth levels where the data are available (and also add rows of simulated values at these different depth levels), so that the reader can also estimate if the observed concentrations/fluxes fall into the range of simulated values? Otherwise, I have the feeling that this table could be removed. As it is, is in unclear, even from lines 170–175, which main result(s) the reader should keep in mind from this Table.*

Reply: We added one column with observational estimates, based on global averages of either available observations or of data-based (data-driven) model results for those tracers/rates whose measurements are sparse on the global scale. For better comparison we adjusted the depth ranges accordingly. Now Chl concentration is the average of the top 50 m, and NCP is calculated from 0 to 100 m. We removed POC export and show only NCP, since at steady state it is equivalent to POC export flux.

Positive comment: *Figure 1 nicely shows that the model outputs are highly sensitive to ν_{det} and $Q^{N}_{0,\,phy}$ (and to $g_{\max}$ and ϕ_{phy} at the second order)!*

Section 3.3.3 *is quite long, but it presents a very detailed and interesting description and associated comments of the results presented in Figure 1. Keep it as it is.*

Reply: We appreciate your positive comments.

Figure 3: *Justify in the methods the reason(s) of your choice of performing a regional splitting into latitudinal bands. This is missing in the article. Indeed, you mention later line 259: "sensitivities of dissolved N:P ratio to parameters in [...] three geographical settings (low, high latitudes and global)". It has to be mentioned (and justified) earlier.*

Reply: We added the justification in the Methods section, now on p. 5, lines 127–129: *"We also evaluate the sensitivities of surface particulate elemental ratios (C:N, C:P and N:P), as well as nitrate to phosphate ratios for different latitude bands (40°S to 40°N, 60°S to 70°S, and globally). This is because dissolved and particulate elemental ratios in general show very different behaviour between lower and higher latitudes (Martiny et al., 2013a)."*

Line 250: *"where diazotrophs are abundant in high latitudes": yet, this is not visible from your results. If this comes from Pahlow et al. 2019 please indicate it.*

Reply: We have added *"(see Fig. 15 in Pahlow et al., 2020)"* on p. 11, lines 271–272.

Figure 2: *Add a legend for black (OPEM) and grey (OPEM-H) as you did in Figure 1. Same comment for Figure 3.*

Reply: Done

Line 285–288: *It seems to me that these sentences should rather be included in the Methods section, not in the Results section.*

Reply: We moved this part to the Methods section (now on p. 7, lines 185–189).

Figure 3: *It may be nice to highlight the values that differ between OPEM and OPEM- H, for instance with rectangles (around the bars) and or stars (below or above the bars), so that it would clearly strike the eyes that the differences between the two configurations are obtained for the 60°S-70°S latitudinal band for C:N, C:P, and N:P. Also indicate in the legend the choices you made for the "different latitude bands".*

Reply: We follow your suggestion and have added asterisks below or above the bars where sensitivities are very different between OPEM and OPEM-H. We also indicate our choices of our latitude bands in the legend.

Figure 5 *is not easy to read as it is because purple and black symbols look very similar. Smaller symbols may be used to help. Drawing horizontal and vertical lines to better underline the location of the WOA 2013 values (green square) may also be a good idea (although the figure is very well described lines 306–313).*

Reply: We shrank size of the symbols, and changed the color key, which appears to make it easier to read. In additional to drawing lines to indicate the location of the WOA 2013, we added its nitrate and oxygen values in the legend.

Lines 313–314: *"Overall, we stress that the minimum-cost and trade-off solutions appear at the margin of the full spread of the ensembles, which could be interpreted as indicating a model deficiency.": I do not understand what you mean here. For me, it seems that they are in a patch of simulations with symbols in black, indicated log10 of cost values lower than 8, which seems OK. What are you referring to by the term "model deficiency"?*

Reply: Ideally, we would obtain the ensemble solutions evenly spread around the WOA 2013 data-based estimate. That the cost values close to this point are low is a necessary consequence of the definition of the cost function. Our results clearly show, however, that the mean of the ensemble solutions does not correspond to the observational estimate. Given all the non-linearities, this is not really surprising. Referring to this as a "model deficiency" might be an overly critical statement. We removed the term "model deficiency" and rephrased our explanation on p. 17, lines 332–333: *"The ensemble solutions are unevenly spread around the WOA2013 data-based estimates. This highlights that our trade-off solutions could not have been identified had we only considered the ensemble means."*

Line 315: *"Figures 6 and 7 show zonally averaged NO_3^- and O_2 in simulations with low and high NO_3^- and the trade-off simulations": Would it be possible to delineate these simulations in Figure 5? Indeed, it is unclear if the concentrations presented Figures 6 and 7 come from one simulation only, or from several (how many?) simulations. When describing these two figures, also underline the fact that the outputs from OPEM and OPEM-H are very similar here. If this is indeed the trade-off simulation (as indicated in the legend), then the results should be the same and there is no need to show twice the same figures.*

Reply: The panels in Figures 6 and 7 are zonally averaged NO_3^- and O_2 from three simulations that result from the three parameter sets that generated the lowest and highest NO_3^- inventories, and the trade-off simulation in the OPEM configuration. We have revised the sentence to read *"Figures 6 and 7 show zonally averaged NO_3^- and O_2 in simulations with the lowest and highest NO_3^- and the trade-off simulation in the OPEM configuration."* on p. 17, lines 334–335. We indicate the low and high NO_3^- simulations with solid black symbols in Figure 5, while the trade-off simulations are already highlighted. We now underline the fact that the results from OPEM and OPEM-H are very similar and removed the panels for OPEM-H from these figures. We have also added zonal averages from the WOA 2013 for comparison.

Line 332: *"because of intense denitrification in the ODZ" => the last (and first) time that you used the abbreviation ODZ was line 193. As it has not been used since, I recommend giving the full name here again and not just the abbreviation (as you do it later line 402).*

Reply: We have added the full name *"oxygen deficient zones"*, now on p. 17, line 352.

Line 334: *"widespread ODZs, occupying much of the deep water in the northern and equatorial Pacific as well as the Indian Ocean (Figure 6)" => Please indicate these areas clearly on the Figure 6, using arrows for instance.*

Reply: We have added in the supplement a global 2D map (Figure S1) showing oxygen concentrations in the deep water (1240 to 5490 m) for a simulation with very low globally averaged oxygen. This simulation is the same as the low oxygen OPEM simulation shown in Figure 6.

Figure 8: *clearly mention in the legend that the two trade-off simulations for OPEM and OPEM-H are in fact the same, and use only one symbol for this trade-off simulation for figure clarity.*

Reply: The two trade-off do have same parameter combination, but their costs are slightly different, so we need to keep both symbols. We have added *"Note that trade-off simulations share same parameter combination but have slightly different cost values."* in the caption.

Remarks on the Discussion

The section 4.1.1.: (*especially the lines 348–362*) *provides new and very interesting hypotheses on the link between NO3 inventory at global scale and phytoplankton physiology. I appreciate this section.*

Reply: We appreciate this comment.

Line 404: *"ODZ volumes in the trade-off simulations are more than twice that in the WOA 2013 (Figure 10)" => I do not see where it is visible on the Figure 10. I guess it could be inferred from Figure 10C from an expert eye, but I would rather give the precise value in the legend of Figure 10, with the corresponding vertical lines on Figure 10C, if you decide to keep the text as it is. Besides, this is the fist mention of Figure 10, that will be mentioned again line 439. I recommend clearly describing this figure here and later in the discussion, to fully explain and exploit it.*

Reply: We have added ODZ volume for the WOA2013 ($7.45 \times 10^{14} \mathrm{m}^3$) in the caption and also added a description of this figure earlier in this section on p. 23, lines 437–439: *"To evaluate how water-column denitrification affects the cost value of our simulations, we arrange our simulations in the order of their cost values and plot the volume of oxygen deficient zones (ODZs) against cost values for both the OPEM and OPEM-H configurations in Figure 10A to C."*. On line 439 we misplaced an extra (Figure10) here. This sentence is referring to Figure (S1 -S6) (now Figure (S2-S7)) in the previous sentence. We have removed the misplaced (Figure 10) now.

Line 436–437: *"A peculiarity of our cost function is that it complements the data-model misfit, i.e. the residuals of spatial mean log-transformed values, with an additional term that resolves differences in spatial variances" => Yes, indeed! I have particularly appreciated this.*

Reply: We are happy that this is perceived in a positive way.

Line 439: *"The cost function's variance term introduces a strong penalty to approximately 30 % of all ensemble model solutions (Figure 10)." => As mentioned above, Figure 10 lacks a clear description. I do not see what in Figure 10 supports this, but I am sure the authors could give more explanation for helping the reader through this.*

Reply: We apologize, because we misplaced an extra (Figure10) here. This sentence is actually referring to Figure (S1–S6) (now Figure (S2–S7)) in the previous sentence. We have removed the misplaced (Figure 10) now.

Additional remarks

I am wondering why keeping the quarter of the 400 simulations with the highest (worst) cost values in all the analyses, and not keeping only the 200 to 300 best ones?

Reply: This is a fair and meaningful comment, because it reflects an aspect we also discussed internally. We concluded that our analyses should involve a global (total / full) sensitivity analysis (sensitivities to all variations in the full parameter-cost function manifold), rather than local sensitivity analyses in the vicinity around the

trade-off solutions. This way we think we reveal information about the full model behaviour. From our internal discussion we learned that another difficulty would be to justify a threshold limit around the vicinity, which becomes even more subjective in our situation where the minima ("best") do not exactly match the trade-off solutions. There is no justification for why we could use only the best 200, 237, 299, or 300 solutions for our analysis. We have decided instead to add an explanation on p. 5, lines 129–130: *"We keep all 400 simulations because we want to obtain the sensitivity information for the full parameter ranges."*.

Technical corrections

Minor comments and typos:

Lines 44–46: *"Our new ecosystem model [...] offers new features and it improves the representation of some biogeochemical tracers on the global scale (see accompanying study, Pahlow et al., 2019)" => Which biogeochemical tracers? Give examples in brackets.*

Reply: We provide examples here, so now the sentence reads *"Our new ecosystem model [...] offers new features and it improves the representation of some biogeochemical tracers on the global scale (e.g., net community production (NCP) and particulate C:N:P in the surface water, see Part I, Pahlow et al., 2020)."* on p. 2, lines 45–47.

Lines 48–49: *"This model approach yields mass flux estimates with spatial and temporal variations in the elemental C:N:P stoichiometry of both inorganic nutrients and organic matter." => Add at the end of this sentence: "as observed in situ" and give some references to justify (e.g. Martiny, A.C., Vrugt, J.A., Primeau, F.W., Lomas, M.W., 2013. Regional variation in the particulate organic carbon to nitrogen ratio in the surface ocean. Global Biogeochem. Cycles 27, 1–9.)*

Reply: We have added *"as observed in situ"* and cite Martiny et al. (2013) and Loh et al. (2000) now on p. 2, line 51.

Line 79: *"Our setup comprises ensembles of 400 simulations for each of two model configurations. The two model configurations differ in how temperature affects diazotro- phy." => This could be replaced by "Our setup comprises ensembles of 400 simulations for each of the two model configurations that differ in how temperature affects diazotrophy."*

Reply: Done, now the sentence is on p. 3, lines 85–86.

Line 102: *"the parallel setup with different parameter combinations has a some advantages" => Remove "a".*

Reply: Done, now the sentence is on p. 4, line 113.

Line 103: *Replace "Individual" by "individual".*

Reply: Done, now it is on p. 4, line 114.

Legend of Figure 4: *Replace "minmum-cost" by "minimum-cost".*

Reply: Done.

Line 337: *a space is missing after the term "quota".*

Reply: Done, now the space is on p. 18, line 358.

Line 360: *"our simulations: A more intense..." replace "A" by "a".*

Reply: Done, now it is on p. 22, line 383.

Line 378: *You may want to change "do contribute some variations to most of the tracers" by "do contribute to some variations of most of the tracers"*

Reply: Done, now the sentence is on p. 22, lines 405–406.

Line 393: *Figure 5 instead of Fig. 5 (for homogeneity).*

Reply: Done, now it is on p. 23, line 429.

Line 421: *" The mean global estimates ± 1 standard deviation in OPEM and OPEM-H are..."=> You may want to replace "± 1" by "$\pm$".*

Reply: Done, now it is on p. 24, line 466.

Line 496: *"and" instead of "adn"*

Reply: Done, now it is on p. 28, line 542.

We thank the referee for the constructive review. The comments and questions are useful, which helped us to introduce changes and improve our manuscript. We listed our responses to all of your points below and hope the manuscript is now satisfactory.

General comments

This paper aims to optimize and calibrate important parameters used in the lower trophic marine ecosystem component of UVic-ESM and is a companion paper to the model description paper (Pahlow et al., 2019). In this study, authors set up cost functions to minimize the misfit between model outputs and observations for nitrate, phosphate, dissolved oxygen, and surface chlorophyll-a. Of the 13 parameters they have chosen to calibrate, the subsistence N quota of phytoplankton and remineralization rate have the highest sensitivity. Overall, the paper is nicely written and organized. Optimization schemes are well described and the parameters are calibrated rigorously. However, I do have some important points that need to be clarified before I am ready to recommend publications of this paper in the GMD.

1. *What is the "best" model choice? The authors state in line 7 – "For identifying the "best" model we therefore also consider ... water-column denitrification". I was not ultimately clear after reading this paper, what the "best" model choice is. Is it OPEM/OPEM-H with the lowest overall total cost function or "trade-off" model which does not necessarily have the lowest cost function (7th best) but does best at representing N cycle? I may have missed this but if water-column denitrification and N2 fixation are indeed very important, why did you not include these in your cost function?*
   Reply: The "best" model solutions in line 7 refer to the trade-off simulations in each of the OPEM and OPEM-H configurations. We changed "best model solutions" on p. 1, line 8 to "reference parameter sets" to avoid this confusion. We also modified p. 1, lines 6–8 to: *"The simulations closest to the data with respect to our metric exhibit very low rates of global N_2 fixation and denitrification, indicating that in order to achieve rates consistent with independent estimates, additional contraints have to be applied in the calibration process. For identifying the reference parameter sets..."* During our analysis we had considered the implementation of observed N_2 fixation rates to our cost function, as suggested by the referee. But we quickly learned that this is not straightforward, mainly because of the scarcity of observed rates on the global scale, which introduces a large imbalance (between the many terms for each of NO_3^-, O_2, and DIC, and the one term for global N_2 fixation) in our cost function. The spatial and temporal coverage of these data is very different from the monthly resolved tracer concentrations we consider for our cost function. Such an imbalance requires the introduction of some regularization, which would make the cost function less objective than it is now. Instead, we interpreted the identified model solutions in terms of a multi-objective optimisation. In this manner, the consideration of global N_2 fixation rates is treated as a second objective, in addition to our cost function being the first objective. This is the reason why we refer to these model results as *trade-off solutions*. We address the problem now more explicitly on pp. 24–26, lines 474–478: *"Incorporating N_2 fixation as a single global rate estimate into our Likelihood-based cost function as a single additional term would, without some difficult-to-define regularization, become overwhelmed by the many tracer and variance terms defined in Eqs. (6) and (7). Rather, the additional information is treated as a second objective, namely that global N2 fixation should be greater than 60 Tg N yr^{-1} (see above), which is similar to applying a multi-objective approach for model calibration (e.g., Sauerland et al., 2019), where a trade-off between two or more objectives (cost functions) is resolved."*

2. *What is the selling point of this "optimized" flexible C:N:P model? Authors state that most NPZD models do not adequately describe the behavior of plankton physiology such as non-Redfieldian plankton stoichiometry. However, outside the UVic framework, there are quite a few ESMs in the market already with flexible C:N:P including those in CMIP5 (see Bopp et al., 2013) and CMIP6 (see Arora et al., 2019). There are also some studies that utilize Pahlow's phytoplankton model (Kwiatkowski et al., 2018, 2019). My question then is what is the selling point of this model over other existing*

models out there? Is it the computational efficiency and how useful is this model for studying climatic conditions such as the last glacial maximum or future projections (lines 39)? I think some discussions on model comparisons would be useful.

Reply: The combination of optimality-based nitrogen fixation (Pahlow et al., 2013) and optimal current feeding for zooplankton (Pahlow, 2010), together with the flexible C:N:P stoichiometry are the novel features in the OPEM. Kwiatkowski et al. (2018, 2019) adopted a previous optimality-based model for phytoplankton growth (Pahlow et al., 2009) that does not include optimal resource allocation for nitrogen fixation. It also lacks the optimal current feeding model of zooplankton. None of the biogeochemical modules of the ocean models in CMIP5 and CMIP6 resolve dynamics with respect to the optimality conditions applied in OPEM. PELAGOS (Vichi, Pinardi, and Masina, 2007) is the only model application with variable C:N:P in phytoplankton in CMIP5 (Bopp et al., 2013) and CMIP6 (Arora et al., 2019). It does not consider diazotrophy, and other models resolve either variable N:P (TOPAZ2, Dunne et al., 2013) or variable C:P (MARBL (CESM2), Danabasoglu et al., 2020). In addition to the variable C:N:P stoichiometry, the optimality-based formulations of primary producers and zooplankton have a demonstrated ability to describe processes observed in the laboratory as well as in mesocosm studies and hence provide a strong mechanistic foundation for OPEM. We have added a comparative description in the introduction in this ms on p. 2, lines 51–55: *"PELAGOS (Vichi et al., 2007), the only ocean model with variable C:N:P in phytoplankton in CMIP5 (Bopp et al., 2013) and CMIP6 (Arora et al., 2019), has no diazotrophs, others either have only variable N:P (TOPAZ2, Dunne et al., 2013), or variable C:P (MARBL, Danabasoglu et al., 2020). While some of the existing models have a variable C:N:P based on the optimality-based model for phytoplankton growth (Kwiatkowski et al., 2018, 2019), optimality-based N_2 fixation is not included."*, as well as extended that in Part I on p. 3, lines 65–68: *"We view the implementation of OPEM as one step towards the ultimate goal of reconciling plankton-organism behaviour as observed in the laboratory with global marine biogeochemistry. Therefore, the variable stoichiometry of primary producers should be considered but one, albeit central, aspect of the mechanistic foundation of OPEM."* in Pahlow et al. (2020).

3. *How sensitive is "sensitive"? Authors discuss the sensitivity of each parameter in Section 3.1 but one thing I find problematic is that all the graphs in Figure 1 – 3 have different y-scale increments. Since sensitivity is non-dimensional, they should ideally all have the same axis for a fair comparison since authors frequently say things like "Sensitivity of XXX is low" (e.g., line 196) or "No single parameter dominates sensitivity" (line 217). Although such rigorous statistical treatments may not be expected for this kind of modeling work, I want some general clarifications on how authors interpreted whether something is very sensitive or not.*

   Reply: The different scales of the y-axes result from our definition of sensitivity in Eq. (3), in which the tracer difference is divided by the average. Some tracers, e.g., DIC or dissolved Fe, vary much less relative to their average concentration, because they are more strongly determined by boundary conditions (air-sea exchange, atmospheric deposition) or exhibit a huge background concentration (DIC) that is, on the timescales considered, not affected by biotic or physical processes. Naturally, we expect that our measure of sensitivity, although normalised and thus being dimensionless, varies on different scales for different tracers or rates. We see no reason for why this may appear problematic, since our focus here is on contrasting the effects of the different parameters and not of the different tracers. In fact, it is important for us to take advantage of the different y-axis scales. Doing so reflects more clearly how sensitive each individual tracer is to variations in the different parameters. The information of the differences between the tracers' general sensitivities to parameter variations is maintained, but according to our chosen style of presentation the emphasis is on the sensitivity to variations in the individual parameters.

   We agree with the referee that we need to be more careful with statements like "the sensitivity of X is low", in particular, if we cannot provide a common reference point. It is more appropriate to refer to relative differences in the sensitivities, which we have considered in our corrections. Thus, we now use phrases like

"sensitivity of X is lower than Y". The text of the description of "No single parameter dominates sensitivity" for N_2 fixation has been corrected accordingly. The rephrased description on p. 10, lines 238–241, now reads: *"The simulated global N_2 fixation rate is sensitive to many parameters, apart from A_0 and Q_0^Pdia. Similar relative changes of most parameter values introduce changes to the global N_2 fixation rate that are of similar magnitude. Interestingly, N_2 fixation is sensitive also to zooplankton parameters, indicating that zooplankton grazing on diazotrophs is an important factor controlling not just diazotroph biomass but also N_2 fixation."*

4. *The highest sensitivity of C:N over C:P and N:P? Regarding the sensitivity, I was quite surprised looking at Figure 3 that C:N has much larger sensitivity compared to C:P and N:P. The current understanding in the scientific community is that C:N is more homeostatic compared to C(N):P for autotrophs, heterotrophs, and for detritus (Galbraith and Martiny, 2015; Geider and La Roche, 2002; Martiny et al., 2013; Sterner and Elser, 2002). Looking at the companion paper by Pahlow et al. (2019), steady-state C:N also seems to overestimate observation (Table 3 and Figure 7). I think this is an important point to address given that C:N (and therefore QoN) affects all aspects of the model output and that the whole point of this model is incorporating flexible C:N:P.*

   Reply: The finding that C:N is more homeostatic than C:P or N:P for particulate matter applies to the spatio-temporal variability in the current ocean and thus could be compared to our trade-off (reference) simulations, which is the topic of Part 1 (Pahlow et al., 2020). The present article, however, describes sensitivities of globally-averaged elemental ratios to parameter variations among 400 simulations with different parameter settings. Thus, the sensitivities discussed here have a very different meaning to the spatio-temporal variability of the elemental ratios in the surface ocean. Particulate C:N and C:P are not only directly affected by $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$ and $Q^{\mathrm{P}}_{0,\,\mathrm{phy}}$, but also by the NO_3^- and PO_4^{3-} inventories. The marine NO_3^- inventory varies strongly owing to N_2 fixation and denitrification. In contrast, the PO_4^{3-} inventory is conserved in the UVic model, allowing only shifts in the spatio-temporal distribution of PO_4^{3-}. Hence, the sensitivity of globally averaged particulate C:N across simulations with different parameter sets is greater than that of C:P. We have added a discussion to clarify this issue and avoid a possible misunderstanding of the nature of these variations in elemental ratios in the manuscript on p. 22, lines 385–392: *"The strong impact of $Q^N_{0,\,phy}$ on the NO_3^- inventory and globally averaged phytoplankton C:N causes a higher sensitivity of globally averaged C:N than C:P (Figure 3). A higher $Q^N_{0,\,phy}$ results in a higher NO_3^- inventory and a lower phytoplankton C:N, both tending to lower particulate C:N and vice versa. On the other hand, C:P is not as sensitive because we have a constant PO_4^{3-} inventory in the UVic model. Surface particulate matter C:N is less variable compared to C:P and N:P in field observations along regional gradients (Galbraith and Martiny, 2015; Geider and Roche, 2002; Martiny et al., 2013; Sterner and Elser, 2002), which is an apparent contrast to our results, where the sensitivity of C:N to $Q^N_{0,\,phy}$ is the highest among the particulate elemental ratios. However, our sensitivities are with respect to parameter variations among many simulations, rather than spatial or temporal gradients in the one real ocean."*

   The overestimated steady-state C:N in Part I (Pahlow et al., 2020) results from N_2 fixation in the trade-off simulations being much lower than in the current ocean due to the lack of benthic denitrification. Lower N_2 fixation results in a lower supply of nitrogen and consequently an overall higher particulate C:N at low latitudes. We have added a discussion on p. 23, lines 432–436: *"Also, this means that global N_2 fixation (same as global denitrification in our spun-up steady-state simulations) is underestimated, and since it occurs mostly at 40°S to 40°N (see Fig. 13 in Part I, Pahlow et al., 2020), particulate carbon to nitrogen (C:N) ratios could be overestimated due to a missing input of nitrogen to the surface ocean. This could explain the overestimated surface particulate C:N at low latitudes (see Table 3 and Figure 16 in Part I, Pahlow et al., 2020)."* We have also added a statement about this topic in Part I (Pahlow et al., 2020) on lines 415–418: *"Both the high surface C:N and low P:C in mid-latitude regions might result from the underestimation of N_2 fixation, owing to the lack of benthic denitrification. Enhanced N_2 fixation would add fixed N to the surface ocean, partly releasing phytoplankton from N limitation and intensifying P limitation, and could thus bring C:N and C:P ratios closer to the observations."*

We disagree with the statement that "the whole point of this model is incorporating flexible C:N:P." Flexible C:N:P can be (and has been) implemented in several ways. Although we consider the representation of variable C:N:P in the OPEM very important, our main goal here is, nevertheless, improving the mechanistic foundation of biotic process descriptions in Earth system models.

Specific comments

Equations: *Diazotrophy rate increases indefinitely with temperature with this formulation. But the growth rate of diazotrophs should hit the limit at some optimal value (e,g., 28 degrees Celsius for Trichodesmium; Breitbarth, E., A. Oschlies, and J. LaRoche (2007), Physiological constraints on the global distribution of Trichodesmium-effect of temperature on diazotrophy, Biogeosciences (BG), 4(1), 53–61). What is the justification of this temperature formulation? I feel like Eppley (1972) is not quite up to date.*

Reply: Because there is no equation number indicated, the question is not clear to us. In Equation 1 the rate of N_2 fixation indeed increases indefinitely with temperature. While we agree in principle that Eppley (1972) is not quite up to date, this is exactly the reason for introducing Equation 2, which is based on observations, where maximum diazotrophy rate occurs around 25 °C. However, Eppley (1972) is the temperature function in the original UVic, and we wanted to be clear about which changes in model behaviour are due to the optimality-based, variable-stoichiometry formulations, and which are due to the new temperature function. Thus, we set up two model configurations to identify the influence of the temperature dependence of diazotrophy on model behaviour.

Line 85: *The temperature dependence of nitrogenase activity in the terrestrial system was used. Are there not any data from marine ecosystem literature?*

Reply: No, at least we are not aware of any.

Table 1: *How are the "Range" chosen for these parameters?*

Reply: The parameter ranges are based on literature values. We have revised the description on p. 4, lines 106–108: *"We vary 15 parameters in total, within the variational ranges shown in Table 1, which are based on reference ranges according to literature values."* and added references for the parameters in Table 1.

Table 2: *Maybe it would be nice to have some "target" values for comparison from WOA 2013 or other datasets.*

Reply: We have added the values for the WOA 2013 and other datasets for reference.

Line 202: *What are the sinks for DFe?*

Reply: We have a sink for DFe to the sediment, we added this to the text, now on p. 10, line 223: *...iron has a fixed source from atmospheric deposition and a sink in the sediment,...*

Figure 2: *Phytoplankton (1st column) and diazotrophs (2nd column) have different y-axis range. For a fair comparison, they should have the same y range (at least for the same given row).*

Reply: As explained above in our reply to the third general comment, we have decided to keep the y-axis ranges as they were. To avoid confusion, we added *"Note the different y-axis ranges in the different panels."* to the caption.

Line 246: *"their biomass is higher". What is "biomass"? Is it C quota or C+N+P or Chl? I do not see "biomass" in Figure 2.*

Reply: We use the term biomass to refer to C, so higher biomass means higher POC content. The higher diazotroph biomass (Carbon, vertically-integrated and temporally-averaged biomass, mmol C m^{-2}) can be seen in Fig. 15 in Pahlow et al. (2019) (now Pahlow et al., 2020). We revised the sentence to *"is generally larger because of the growth of diazotrophs at high latitudes (see Fig. 15 in Part I, Pahlow et al., 2020)"* on p. 11, lines 264–265 to indicate this. Note that we changed Pahlow et al. (2019) to Pahlow et al. (2020) in the manuscript.

Line 254–257: *The logical behind explaining C:P pattern is not clear. Why does NO3:PO4 supply stoichiometry only affect low latitudes? Why that fact P-limitation is not present in S. Ocean explain the negative correlation between C:P and* $Q^{N}_{0,\,phy}$*?*

Reply: To clarify the explanation, we revised this part as: *"At low latitudes, the effects of* $Q^{P}_{0,\,phy}$ *are suppressed by variations in phytoplankton C, which is affected by* $Q^{N}_{0,\,phy}$ *and the consequent change in nitrate concentration. Nitrate and phosphate are not limiting in the high-latitude Southern Ocean where, under N- and P-replete conditions, cellular C:P is mainly determined by* $Q^{P}_{0,\,phy}$ *and a higher* $Q^{P}_{0,\,phy}$ *would result in a higher cellular P:C (lower C:P). Therefore, the global C:P of total particulate matter, which is dominated by ordinary phytoplankton, is negatively correlated with* $Q^{P}_{0,\,phy}$*."* on p. 11, lines 277–281.

L282: *The description of "trade-off solutions". I went to Pahlow et al. (2019) but I could not easily locate where the discussion is. Could you direct me specifically to where it is?*

Reply: The two calibrated reference simulations in Pahlow et al., (2020) are the "trade-off solutions" in this manuscript. We changed *"in the companion paper Pahlow et al. (2019)"* to *"in Part I (reference simulations in Pahlow et al., 2020)"* on p. 16, line 306 to avoid such confusion.

Figures 6 and 7: *What does "low nitrate" and "high nitrate" mean? I may have missed it but are they different model configurations or are they taken from different oceanographic regions?*

Reply: The "low nitrate" and "high nitrate" are the simulations with the lowest and highest globally averaged nitrate concentrations in the OPEM configuration. We revised the description in the caption of Figure 6 as: *"Zonally averaged* NO_3^- *in the World Ocean Atlas 2013 (A), the simulations with the lowest and highest* NO_3^- *inventory (B, D), and the trade-off simulation (C) in the OPEM configuration. Globally averaged* NO_3^- *concentrations are shown in each panel. Simulations shown here are marked with solid black and open red triangles in Figure 5. Note that the outputs from OPEM and OPEM-H are very similar and only OPEM results are shown here."* and in the text on p. 17, lines 334–335: *"Figures 6 and 7 show zonally averaged* NO_3^- *and* O_2 *in simulations with the lowest and highest* NO_3^- *and the trade-off simulation in the OPEM configuration."* . We now show the corresponding simulations in Figure 5. We only show simulations from the OPEM now, because distributions in the OPEM-H are very similar to the OPEM.

Also Figures 6 and 7: *It would be nice to have a zonal average from WOA 2013 for comparison.*

Reply: We added zonal averages from the WOA 2013 and removed simulations of the OPEM-H configuration since the distributions are very similar to those of OPEM.

Line 381: *N:P of diazotrophs is critically important for determining the outcome of competition between diazotrophs and non-diazotrophs so it should be discussed in more depths here (e.g., Weber and Deutsch, 2012).*

Reply: While N:P of diazotrophs was proposed to be very important for determining the outcome of competition between diazotrophs and non-diazotrophs, results of our sensitivty analysis of the OPEM do not support this. In Figure 2 we can see while N:P of diazotrophs is most sensitive to diazotroph subsistence P quota ($Q^{\mathrm{P}}_{0,\,\mathrm{dia}}$), dizotrophs biomass (carbon) itself is much less sensitive to $Q^{\mathrm{P}}_{0,\,\mathrm{dia}}$ than to $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$ and to $Q^{\mathrm{N}}_{0,\,\mathrm{dia}}$. In our view,

the competitive abilities for N and P are more important than the N:P ratio for determining the outcome of such competition. We have added a discussion about how N:P affects competition and N_2 fixation of diazotrophs on p. 23, lines 409–417, which now reads: *"Diazotroph subsistence N and P quotas ($Q^N_{0,\,dia}$ and $Q^P_{0,\,dia}$) in general have much less influence on particulate stoichiometry than $Q^N_{0,\,phy}$ and $Q^P_{0,\,phy}$ because diazotrophs are much less abundant than ordinary phytoplankton. However, diazotroph biomass (carbon) itself is more sensitive to $Q^N_{0,\,dia}$ than $Q^N_{0,\,phy}$, which shows that the diazotroph subsistence quotas are still important for both their elemental stoichiometry and ability to compete with ordinary phytoplankton. While elemental stoichiometry has been suggested to be an important factor for determining the outcome of the competition between diazotrophs and non-diazotrophs, and consequently N_2 fixation (Deutsch and Weber, 2012; Weber and Deutsch, 2012), we find that N_2 fixation is no more sensitive to $Q^N_{0,\,dia}$ than to the remineralisation rate (ν_{det}), $Q^N_{0,\,phy}$, or zooplankton grazing parameters ($g_{\max}$, ϕ_{phy}, and ϕ_{dia}). Nevertheless, our analysis agrees with the argument that global N_2 fixation is mainly determined by rates of fixed-N loss (Weber and Deutsch, 2014), which in our model is largely affected by ν_{det} and $Q^N_{0,\,phy}$."*

Line 407: *I think authors should also mention the fact that physical component/ocean circulation is very important for the global distribution of oxygen and nitrate.*

Reply: We have added a statement that the physical component/ocean circulation is very important for the global distribution of oxygen and nitrate. Now on p. 24, lines 450–451: *"While the physical component (ocean circulation) of the UVic model is also very important for the global distribution of oxygen and nitrate, our results suggest that…"*

References

Arora, V. K. et al. (2019). "Carbon-concentration and carbon-climate feedbacks in CMIP6 models, and their comparison to CMIP5 models". In: *Biogeosciences Discussions*, pp. 1–124. DOI: 10.5194/bg-2019-473.

Bopp, L. et al. (2013). "Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models". In: 10.10, pp. 6225–6245. ISSN: 1726-4170. DOI: 10.5194/bg-10-6225-2013.

Danabasoglu, G. et al. (2020). "The Community Earth System Model Version 2 (CESM2)". In: *Journal of Advances in Modeling Earth Systems* 12.2. DOI: 10.1029/2019MS001916.

Deutsch, Curtis and Thomas Weber (2012). "Nutrient Ratios as a Tracer and Driver of Ocean Biogeochemistry". In: *Annual Review of Marine Science* 4.1, pp. 113–141. DOI: 10.1146/annurev-marine-120709-142821.

Dunne, John P. et al. (2013). "GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part II: Carbon System Formulation and Baseline Simulation Characteristics". In: *Journal of Climate* 26.7, pp. 2247–2267. DOI: 10.1175/JCLI-D-12-00150.1.

Galbraith, Eric D. and Adam C. Martiny (2015). "A simple nutrient-dependence mechanism for predicting the stoichiometry of marine ecosystems". In: *Proceedings of the National Academy of Sciences* 112.27, pp. 8199–8204. DOI: 10.1073/pnas.1423917112. eprint: http://www.pnas.org/content/112/27/8199.full.pdf. URL: http://www.pnas.org/content/112/27/8199.abstract.

Geider, Richard and Julie La Roche (2002). "Redfield revisited: variability of C:N:P in marine microalgae and its biochemical basis". In: *European Journal of Phycology* 37.1, pp. 1–17. DOI: 10.1017/S0967026201003456.

Martiny, Adam C. et al. (Apr. 2013). "Strong latitudinal patterns in the elemental ratios of marine plankton and organic matter". In: *Nature Geosci* 6.4, pp. 279–283. URL: http://dx.doi.org/10.1038/ngeo1757.

Pahlow, M. et al. (2020). "Optimality-Based Non-Redfield Plankton-Ecosystem Model (OPEMv1.0) in the UVic-ESCM 2.9. Part I: Implementation and ModelBehaviour". In: *Geoscientific Model Development Discussions*, pp. 1–34. DOI: 10.5194/gmd-2019-323.

Sterner, R. W. and J. J. Elser (2002). *Ecological Stoichiometry: The Biology of Elements from Molecules to the Biosphere*.

Vichi, M., N. Pinardi, and S. Masina (2007). "A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part I: Theory". In: *Journal of Marine Systems* 64.1, pp. 89–109. DOI: https://doi.org/10.1016/j.jmarsys.2006.03.006.

Weber, Thomas and Curtis Deutsch (2012). "Oceanic nitrogen reservoir regulated by plankton diversity and ocean circulation". In: *Nature* 489. URL: http://dx.doi.org/10.1038/nature11357.

— (2014). "Local versus basin-scale limitation of marine nitrogen fixation". In: *Proceedings of the National Academy of Sciences* 111.24, pp. 8741–8746. DOI: 10.1073/pnas.1317193111.

Optimality-Based Non-Redfield Plankton-Ecosystem Model (OPEM v1.0) in the UVic-ESCM 2.9. Part II: Sensitivity Analysis and Model Calibration

Chia-Te Chien, Markus Pahlow, Markus Schartau, and Andreas Oschlies
GEOMAR Helmholtz Centre for Ocean Research Kiel
Correspondence: Chia-Te Chien (cchien@geomar.de)

Abstract.

We analyse 400 perturbed-parameter simulations for two configurations of an optimality-based plankton-ecosystem model (OPEM), implemented in the University of Victoria Earth-System Climate Model (UVic-ESCM), using a Latin-Hypercube sampling method for setting up the parameter ensemble. A likelihood-based metric is introduced for model assessment and selection of the model solutions closest to observed distributions of NO_3^-, PO_4^{3-}, O_2, and surface chlorophyll *a* concentrations. ~~According to our metric the optimal model solutions comprise~~ The simulations closest to the data with respect to our metric exhibit very low rates of global N_2 fixation and denitrification~~. These two rate estimatesturned out to be poorly constrained by the data~~, indicating that in order to achieve rates consistent with independent estimates, additional contraints have to be applied in the calibration process. For identifying the ~~"best" model solutions~~ reference parameter sets we therefore also consider the model's ability to represent current estimates of water-column denitrification. We employ our ensemble of model solutions in a sensitivity analysis to gain insights into the importance and role of individual model parameters as well as correlations between various biogeochemical processes and tracers, such as POC export and the NO_3^- inventory. Global O_2 varies by a factor of two and NO_3^- by more than a factor of six among all simulations. Remineralisation rate is the most important parameter for O_2, which is also affected by the subsistence N quota of ordinary phytoplankton ($Q^{N}_{0,\,phy}$) and zooplankton maximum specific ingestion rate. $Q^{N}_{0,\,phy}$ is revealed as a major determinant of the oceanic NO_3^- pool. This indicates that unraveling the driving forces of variations in phytoplankton physiology and elemental stoichiometry, which are tightly linked via $Q^{N}_{0,\,phy}$, is a prerequisite for understanding the marine nitrogen inventory.

1 Introduction

Earth system climate models (ESCMs) are powerful tools for analysing variations in climate, while resolving interdependencies between changes in the atmosphere, on land, and in the ocean (Flato, 2011; Prinn, 2013). In this regard, the dynamics of marine ecosystems is a critical link. On long timescales it regulates atmospheric CO_2 on the basis of biotic uptake of carbon dioxide (CO_2) over vast oceanic regions and due to the export of photosynthetically fixed carbon into the deep ocean, which affects the Earth's climate (Reid et al., 2009; Sigman and Boyle, 2000). Plankton ecosystem models are widely applied to understand marine biogeochemical cycles, by estimating fluxes of major elements, e.g., nitrogen, phosphorus, and carbon, as well as the

sources and sinks of marine oxygen ~~(Maier-Reimer et al., 1995; Six and Maier-Reimer, 1996; Schmittner et al., 2005; ?; Vallina et al., 201~~ (Maier-Reimer et al., 1995; Six and Maier-Reimer, 1996; Schmittner et al., 2005; Bopp et al., 2013; Vallina et al., 2017; Everett et al., 201 .

The basic structure of most marine ecosystem models has been designed for resolving mass fluxes between nutrients, phytoplankton, zooplankton and detritus, typically referred to as NPZD models. Mathematical formulations that describe growth and fate of marine phytoplankton and zooplankton biomass have been successfully applied over a range of scales, from local 0D-ecosystem models (e.g., Fasham et al., 1990; Edwards, 2001) to global 3D models (Sarmiento et al., 1993; Keller et al., 2012; Nickelsen et al., 2015). However, most of these NPZD models lack a sound mechanistic foundation, preventing them from explicitly accounting for the organisms' regulation of their internal physiological state. For example, N_2 fixation by algae is often diagnosed from the availability of dissolved nutrients, so that it only occurs when the ratio of nitrate-to-phosphate concentrations falls below the Redfield ratio of 16:1 (Deutsch et al., 2007; Ilyina et al., 2013). As these assumptions neglect a number of environmental and ecological controls (e.g., grazing, often also temperature), they do not adequately describe the behaviour of plankton organisms and their sensitivity to changes in their environment. With the introduction of refined mechanistic (physiological) descriptions we here aim at alleviating this deficiency. In this study we introduce a new marine ecosystem model coupled to the University of Victoria Earth System Climate Model (UVic-ESCM, based on the configurations of Keller et al., 2012; Getzlaff and Dietze, 2013; Nickelsen et al., 2015). Doing so we anticipate the model not only to provide improved mass flux estimates, but also to exhibit more realistic sensitivities of these fluxes to varying climate conditions, e.g., in simulations of the last glacial maximum or in future projections.

In order to better represent plankton physiology, the new ecosystem model relies on optimality-based considerations for phytoplankton growth, including N_2 fixation (Pahlow et al., 2013; Pahlow and Oschlies, 2013), as well as zooplankton behaviour (Pahlow and Prowe, 2010). These two optimality-based models have been shown to be superior to traditional model approaches in reproducing phytoplankton and zooplankton growth and grazing under various environmental conditions (e.g., Fernández-Castro et al., 2016). Our new ecosystem model, the optimality-based plankton ecosystem model (OPEM v1.0) coupled to the UVic-ESCM, offers new features and it improves the representation of some biogeochemical ~~tracers~~ properties on the global scale ~~(see accompanying study, ?)~~(e.g., net community production (NCP) and particulate C:N:P in the surface water, see Part I, Pahlow et a . One of the novel features is the representation of variable quotas of carbon (C), nitrogen (N), and phosphorus (P) in ordinary phytoplankton, diazotrophs, and particulate organic matter (detritus) exported to the deep ocean. This model approach yields mass flux estimates with spatial and temporal variations in the elemental C:N:P stoichiometry of both inorganic nutrients and organic matter as observed in situ (Loh and Bauer, 2000; Martiny et al., 2013b). PELAGOS (Vichi et al., 2007), the only ocean model with variable C:N:P in phytoplankton in CMIP5 (Bopp et al., 2013) and CMIP6 (Arora et al., 2019), has no diazotrophs, others either have only variable N:P (TOPAZ2, Dunne et al. (2013)), or variable C:P (MARBL, Danabasoglu et al., 2020). While some of the existing models have a variable C:N:P based on the optimality-based model for phytoplankton growth (Kwiatkowski et al., 2018, 2019), optimality-based N_2 fixation or zooplankton behaviour are not included.

Here we analyse the new model's performance and evaluate model-ensemble results against observations. Since the model is based on ~~plankton~~ plankton-organism physiology, it includes new parameters whose values have not been estimated for

global model applications. Also, we set up two configurations, OPEM and OPEM-H, with different temperature ~~dependences~~ dependencies for diazotrophs to investigate the effects of different empirical temperature functions on distributions of diazotrophs and N_2 fixation. Our analysis relies on ensembles of solutions of the two different model configurations, where every single simulation within each ensemble is subject to a different combination of parameter values. The ensembles allow assessing the sensitivity of biogeochemical tracer distributions and budgets to variations of the model's parameters. We introduce a likelihood-based metric that quantifies the global misfit between model results and observations. Amongst the ensemble simulations we regard those model solutions as the best that yield low misfits according to the metric and are also close to current estimates of water-column denitrification. The specific objectives of the present paper are (1) to identify and compare those model solutions that correspond to the best representation of observed tracer concentrations and (2) to specify the sensitivity of simulations to variations of the model's parameter values. We make inferences about the model's overall behavior, especially focusing on data constraints, limitations and advantages of resolving variable C:N:P stoichiometry for estimations of global net primary production (NPP), net community production (NCP), biogenic C export, and the global O_2, N, and C inventories.

2 Materials and Methods

2.1 The non-Redfield, optimality-based plankton ecosystem model in the UVic-ESCM

The optimality-based plankton ecosystem model (OPEM) has been implemented into the UVic-ESCM (Weaver et al., 2001; Eby et al., 2013), version 2.9, in the configuration of Nickelsen et al. (2015) with the isopycnal diffusivity modifications by Getzlaff and Dietze (2013), vertically increasing sinking velocity of detritus (Kriest, 2017), and several bug-fixes (some of which were already introduced by Kvale et al., 2017). The UVic-ESCM comprises three components including a simple one-layer atmospheric energy-moisture balance model (Weaver et al., 2001), a terrestrial model and a three-dimensional general ocean circulation model. The horizontal resolution of the land and ocean model components is $1.8°$ latitude $\times$ $3.6°$ longitude, and the ocean has 19 vertical levels with a thickness ranging from $50\,\mathrm{m}$ in the surface layer to $590\,\mathrm{m}$ in the deep ocean.

The OPEM and its implementation into the UVic-ESCM ~~,~~ are described in ~~the companion paper (?)~~Part I (Pahlow et al., 2020). Briefly, the major new features of the new model include (1) an optimality-based model of phytoplankton growth and diazotrophy with variable C:N:P stoichiometry (Pahlow et al., 2013), (2) the optimal current-feeding model for zooplankton (Pahlow and Prowe, 2010), and (3) variable stoichiometry in detritus. The focus on physiology in the construction of the OPEM enables us to study how biogeochemical tracer distributions and fluxes respond to different assumptions about plankton physiology.

2.1.1 Simulation setup

Our setup comprises ensembles of 400 simulations for each of two model configurations ~~. The two model configurations~~ that differ in how temperature affects diazotrophy. The original temperature dependence of diazotrophs ($f_{\mathrm{dia}}(T)$) in the UVic-ESCM (and other models, e.g., Aumont et al., 2015), which we also employ for the OPEM configuration, limits both growth

and N_2 fixation of diazotrophs to above 15 °C,

$$f_{\text{dia}}(T)\_\text{OPEM} = \max(1.066^T - 2.6, 0)/2 \quad (1)$$

where T is seawater temperature. In the OPEM-H configuration, the temperature dependence of nitrogenase activity in terrestrial systems by Houlton et al. (2008) is implemented as affecting only N_2 fixation,

$$f_{\text{dia}}(T)\_\text{OPEM-H} = 0.0266 * (1.066^T)^{(4.22-1.3166*ln(1.066^T))}{}^{[4.22-1.3166*\ln(1.066^T)]} \quad (2)$$

while growth and nutrient uptake of diazotrophs follow the same temperature dependence as ordinary phytoplankton ~~(see ?).~~ (see Part I, Pahlow et al., 2020). Both of these equations are empirical functions directly simulating expected or observed temperature dependencies of N_2 fixation. We consider Eq. (2) more realistic and hence analyse its effect on model behaviour. However, since the parameters in these two equations have no clearly identifiable physiological meaning, we consider a sensitivity analysis of the parameters in Eqs. (1) and (2) beyond the scope of the present study. Note that some models do not enforce any temperature limitation on nitrogen fixation ~~(e.g., Dunne et al., 2012; Ilyina et al., 2013; Jickells et al., 2017)~~ (e.g., Dunne et al., 2013; Ilyina et al., 2013; Jickells et al., 2017). In the present ocean, waters colder than about 15 °C are generally replete with fixed inorganic nitrogen. For existing parameterisations of N_2 fixation, which are functions of the nitrate deficit with respect to phosphate, there has been little indication of substantial impacts of the formulation of temperature control at low temperatures on the distribution of nitrogen fixation (Somes and Oschlies, 2015; Landolfi et al., 2017). Such differences in formulation may, however, gain importance in environmental conditions different from today's.

For all simulations we impose preindustrial (A.D. 1850) boundary conditions with a CO_2 concentration of 284 ppm. The models have been integrated over a period of at least 10,000 years, until they reached steady-state.

The 400 parameter combinations are obtained via Latin Hypercube Sampling (LHS) (McKay et al., 1979). We vary 15 parameters in total, within the variational ranges shown in Table 1, which are based on reference ranges according to literature values. In order to reduce the number of possible parameter combinations, we vary nutrient affinities for macronutrient uptake and half-saturation concentration for iron uptake for ordinary phytoplankton and diazotrophs in constant proportions ($A_0 : A_{0,\text{D}} = 4 : 3$, $K_{Fe} : K_{\text{Fe,D}} = 1 : 2$), so that diazotrophs have a lower nutrient affinity (Pahlow et al., 2013) and higher Fe half-saturation concentration (Dutkiewicz et al., 2012; McGillicuddy Jr., 2014; Ward et al., 2013) than ordinary phytoplankton. Since our parameter sets are independent of each other, the simulations can be carried out in parallel. Apart from the computational time, the parallel setup with different parameter combinations has ~~a some advantagescompared to systematic (often iterative)~~ some advantages compared to iterative model calibration approaches, e.g., parameter-optimisation: (i) ~~Individual~~ individual model simulations do not depend on any other (i.e. previous) combinations of parameter values, (ii) the ensemble results can always be re-evaluated with different metrics, perhaps with substantial differences between selected "best" solutions, depending on the error model applied, and (iii) the ensembles provide insight to the sensitivities and thus to uncertainties of particular model results with respect to parameter variations.

2.2 Sensitivity Analysis and Model Calibration

2.2.1 Sensitivity analysis

The sensitivity (Sensitivity_T) of a tracer T to a parameter P is defined here as

$$\text{Sensitivity}_T = \frac{\Delta T}{\Delta P} \times \frac{\overline{P}}{\overline{T}} \tag{3}$$

where the Δ indicates the change and the overbar the ensemble mean of P or T. If $\text{Sensitivity}_T < 0$, the tracer and the parameter vary in opposite directions. We evaluate the sensitivities of globally and annually averaged net primary production (NPP), net community production (NCP), ~~particulate organic carbon (POC) export,~~ nitrogen fixation by diazotrophs (N_2 fixation), and the concentrations of oxygen (O_2), nitrate (NO_3^-), DIC, ~~POC,~~ dissolved and particulate iron (DFe and PFe), Chl, ordinary phytoplankton, diazotrophs, particles (ordinary phytoplankton + diazotrophs + zooplankton + detritus) and their elemental stoichiometry to the parameters listed in Table 1. We also evaluate the sensitivities of surface particulate elemental ratios (C:N, C:P and N:P), as well as nitrate to phosphate ratios for different latitude bands (40°S to 40°N, 60°S to 70°S, and globally). This is because dissolved and particulate elemental ratios in general show very different behaviour between lower and higher latitudes (Martiny et al., 2013a). We keep all 400 simulations because we want to obtain the sensitivity information for the full parameter ranges.

2.2.2 Likelihood-based metric assessing global biogeochemical model results

We consider four different types of observations for quantitatively assessing the model simulations. The first three are the objectively analysed monthly (upper $550\,\text{m}$) and annual (below $550\,\text{m}$) concentrations of nitrate, phosphate, and oxygen of the World Ocean Atlas 2013 (WOA 2013, Garcia et al., 2013a, b). The fourth is the monthly mean chlorophyll concentration derived from remote sensing data (MODIS/Aqua level 3), based on monthly climatologies for 10 years from 2008 to 2017, provided by the ocean biology processing group (Ocean Biology Processing Group, 2014). The satellite-derived chlorophyll (Chl) concentrations are used for data-model comparison only for the UVic model's top layer, i.e. the upper $50\,\text{m}$.

We define our metric in terms of spatial averages of 17 distinct biogeochemical biomes, as derived and described by Fay and McKinley (2014). The individual biomes are regarded as regions of common biogeochemistry and thus account for spatial differences between ocean regions on the largest possible (global) scale. Using 56 biogeochemical provinces, as defined by Longhurst (2007), might have hampered our data-model comparison, because a higher resolution of individual regions can accentuate spatial pattern errors in tracer concentrations, resulting from ~~differences~~ model errors in advection and mixing. In our view the biomes of Fay and McKinley (2014) are coarse enough for avoiding this problem, but still sufficiently informative for identifying representative parameter values.

The underlying error model of the likellihood based metric assumes a Gaussian (normal) distribution, which is well represented by using the first two moments of log-transformed tracer concentrations, in particular for the upper ocean layers (Schartau et al., 2017). For every depth-level of the UVic model ($k \in \{1, 2, 3, \ldots, 19\}$), average ~~log-transformed~~ $\log_{10}$-transformed tracer concentrations (~~ln X~~$\log_{10}$X) of type X are determined as spatial arithmetic means for our 17 biomes (indexed as j in Eq. 4) for the

Table 1. Parameter names, reference and variational ranges, identified "best" values for the trade-off simulations ~~in~~ (OPEM and OPEM-H), units and descriptions. Note that the trade-off simulations share the same parameter combination.

Symbol	~~Range~~ Reference range	Variational range	OPEM/ ~~OPEM-H~~ OPEM-H	Units	Definition
$A_{0,\,\mathrm{phy}}$	70–1000[a]	120–280	229	$\mathrm{m^3\,(mol\,C)^{-1}\,d^{-1}}$	phytoplankton potential nutrient affinity
$Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$	0.038–0.086[a]	0.04–0.06	0.04128	$\mathrm{mol\,(mol\,C)^{-1}}$	phytoplankton subsistence N quota
$Q^{\mathrm{N}}_{0,\,\mathrm{dia}}$	0.13[a]	0.06–0.12	0.067	$\mathrm{mol\,(mol\,C)^{-1}}$	diazotroph subsistence N quota
$Q^{\mathrm{P}}_{0,\,\mathrm{phy}}$	0.0008–0.002[a]	0.0013–0.0023	0.0022	$\mathrm{mol\,(mol\,C)^{-1}}$	phytoplankton subsistence P quota
$Q^{\mathrm{P}}_{0,\,\mathrm{dia}}$	0.0027[a]	0.0025–0.0035	0.00271	$\mathrm{mol\,(mol\,C)^{-1}}$	diazotroph subsistence P quota
$k_{\mathrm{Fe,\,phy}}$	0.035–0.12[c-g]	0.04–0.08	0.066	$\mathrm{\mu mol\,m^{-3}}$	phytoplankton half-saturation constant for Fe
$g_{\max}$	0.49–5[a]	1–2	1.75	$\mathrm{d^{-1}}$	zooplankton maximum specific ingestion rate
ϕ_{phy}	174–765[h]	100–200	118	$\mathrm{m^3\,(mol\,C)^{-1}}$	capture coefficient of phytoplankton
ϕ_{dia}	$1.05 \cdot \phi_{\mathrm{phy}}$[i]	150–250	232	$\mathrm{m^3\,(mol\,C)^{-1}}$	capture coefficient of diazotrophs
ϕ_{det}	ϕ_{phy}[c–f]	20–100	94	$\mathrm{m^3\,(mol\,C)^{-1}}$	capture coefficient of detritus
ϕ_{zoo}	0–3230[h]	100–200	118	$\mathrm{m^3\,(mol\,C)^{-1}}$	capture coefficient of zooplankton
$\lambda_{0,\,\mathrm{phy}} = M_{0,\,\mathrm{dia}}$	0.001–0.015[c-f]	0.01–0.03	0.018	$\mathrm{d^{-1}}$	specific mortality rate
ν_{det}	0.05–0.07[c-g]	0.04–0.09	0.087	$\mathrm{d^{-1}}$	remineralization rate

[a](Pahlow, 2005; Pahlow et al., 2013), [b](Pahlow and Prowe, 2010), [c](Keller et al., 2012), [d](Somes and Oschlies, 2015), [e](Somes et al., 2017) [f](Landolfi et al., 2017), [g](Landolfi et al., 2015), [h](Su et al., 2018), [i](Wang et al., 2019)

observations and model results:

$$\left(\overline{\log_{10}\mathrm{X}}\right)_{jk} = \frac{1}{\mathrm{N}_{jk}} \sum_{n=1}^{\mathrm{N}_{jk}} \left(\ln\log_{10} \left[\frac{\max(\mathrm{X}_{(n)}, \mathrm{X}_{(0)})}{\mathrm{X}_{(0)}} \right] \right), \qquad \mathrm{X} \in \{\mathrm{Chl},\, \mathrm{O_2},\, \mathrm{NO_3}^-,\, \mathrm{PO_4}^{3-}\} \tag{4}$$

where N_{jk} is the number of available data points within biome j in depth level k. Prior to ~~log-transformation~~the $\log_{10}$-transformation, all tracer concentrations have been normalised to lower detection (uncertainty) thresholds ($\mathrm{X}_{(0)}$) respectively. Measured or derived concentrations below these thresholds are treated as noise and therefore remain unresolved. Thus, the ~~log-transformed~~ $\log_{10}$-transformed normalised concentrations are non-negative. The threshold-values are: $\mathrm{Chl}_{(0)} = 0.1\,\mathrm{mg\,m^{-3}}$, $\mathrm{O_{2\,(0)}} = 1\,\mathrm{mmol\,m^{-3}}$, $\mathrm{NO_3}^-{}_{(0)} = 0.05\,\mathrm{mmol\,m^{-3}}$, and $\mathrm{PO_4}^{3-}{}_{(0)} = 0.01\,\mathrm{mmol\,m^{-3}}$.

Our metric is derived from a likelihood, assuming a Gaussian error distribution for the residuals, which describe the discrepancy between mean values derived from observations ($\overline{\ln \mathrm{X}^{(\mathrm{obs})}}$ $\overline{\log_{10}\mathrm{X}^{(\mathrm{obs})}}$) and model simulations ($\overline{\ln \mathrm{X}^{(\mathrm{mod})}}$ $\overline{\log_{10}\mathrm{X}^{(\mathrm{mod})}}$).

Hereafter we refer to this metric as our cost function (J). Our cost function is split up into two major parts:

$$J = \sum_{k=1}^{5} J_k^{(u)} + \sum_{k=6}^{19} J_k^{(l)} \tag{5}$$

$$J_k^{(u)} = \sum_{i=1}^{12} \sum_{j=1}^{17} \left[\mathbf{d}^T \, R^{-1} \, \mathbf{d}\right]_{ijk} + \left(\mathbf{v}^{(\mathrm{obs})} - \mathbf{v}^{(\mathrm{mod})}\right)_{ijk}^{T} V_{ijk}^{-1} \left(\mathbf{v}^{(\mathrm{obs})} - \mathbf{v}^{(\mathrm{mod})}\right)_{ijk} \tag{6}$$

$$J_k^{(l)} = \sum_{j=1}^{17} \left[\mathbf{d}^T \, R^{-1} \, \mathbf{d}\right]_{jk} + \left(\mathbf{v}^{(\mathrm{obs})} - \mathbf{v}^{(\mathrm{mod})}\right)_{jk}^{T} V_{jk}^{-1} \left(\mathbf{v}^{(\mathrm{obs})} - \mathbf{v}^{(\mathrm{mod})}\right)_{jk} \tag{7}$$

where $\mathbf{d}$ is the residual vector (see Eq. (8) below), R the covariance matrix (Eq. 9), $\mathbf{v}^{(\mathrm{obs})}$ and $\mathbf{v}^{(\mathrm{mod})}$ the spatial variance estimates of the log$_{10}$-transformed observed and modelled tracers, and V^{-1} are diagonal matrices with the variances (uncertainties) of $\mathbf{v}^{(\mathrm{obs})}$. The first part ($J_k^{(u)}$) of the cost function resolves seasonal changes between the surface and 550 m depth, corresponding to the upper five depth-levels of the model. The second part ($J_k^{(l)}$) represents the lower depth range below 550 m and does not account for seasonal changes, as only annual mean data are available.

The residual vector ($\mathbf{d}$) (whose components represent the tracer types X) used for J describes the differences between the ~~log-transformed~~ log$_{10}$-transformed observations and their model counterparts:

$$\mathbf{d}_{ijk} = \left(\overline{\log_{10} \mathbf{X}_{ijk}^{(\mathrm{obs})}} - \overline{\log_{10} \mathbf{X}_{ijk}^{(\mathrm{mod})}}\right) \tag{8}$$

where i and j are the month and biome indices, respectively. We recall that $\mathbf{d}$ has four components only for the UVic model's top layer ($k = 1$) where chlorophyll data are regarded as well. For $k > 1$ the residual vector contains three components: O_2, NO_3^-, and PO_4^{3-}. Both parts of the cost function ($J_k^{(u)}$ and $J_k^{(l)}$) in turn contain two terms, one with respect to the residuals, as defined in Eq. (8), and another that accounts for the differences between the spatial variances (vectors $\mathbf{v}_{ijk}^{(\mathrm{obs})}$ and $\mathbf{v}_{ijk}^{(\mathrm{mod})}$) within each biome (and month for $J_k^{(u)}$) at each depth-level. The covariance matrices R_{ijk} account for temporal correlations (C_{jk}) between different variables ($\mathbf{X}^{(\mathrm{obs})}$), that are specified for every biome and depth level separately:

$$R_{ijk} = S_{ijk} \cdot C_{jk} \cdot S_{ijk} \tag{9}$$

where the elements of the diagonal matrices S_{ijk} are the standard errors of the mean ~~log-transformed~~ log$_{10}$-transformed tracer concentrations ($\overline{\ln \mathbf{X}_{ijk}^{(\mathrm{obs})}}$ $\overline{\log_{10} \mathbf{X}_{ijk}^{(\mathrm{obs})}}$) calculated in Eq. (4) for every month i, biome j, and depth level k. For $J_k^{(l)}$ the R_{jk} contain only the squared standard errors of the annual data as diagonal elements ($R_{jk} = S_{jk}^2$).

With the consideration of standard errors instead of standard deviations, we implicitly impose weights to differences in the spatial expansion (i.e. number of data points of the gridded product used) of individual biomes. Overall, the final cost function J resolves spatial differences between regions (biomes) as well as temporal differences for those depth levels where monthly data are available. It is thus a ~~trade-off in combining~~ combination of time-varying and spatial information for the assessment of our biogeochemical model results on a global scale.

In order to estimate uncertainty ranges for selected model results (globally-averaged N_2 fixation, NO_3^-, O_2, DIC concentrations, NPP, NCP), we apply a bootstrap method to obtain an uncertainty quantification for our simulated values based on the 400

available ensemble model simulations. We collect the best solutions (lowest cost function value) of 1000 randomly selected subsets of 100 out of our 400 ensemble members. The mean and 95% confidence interval of these subsets provide an uncertainty range in the vicinity of the value of the full ensemble.

3 Results

Table 2 lists the ranges of selected simulated tracers and processes for the full ensemble of parameter values generated by the Latin Hypercube Sampling for the OPEM and OPEM-H configurations. Our results exhibit wide ranges of tracer concentrations and fluxes in these two configurations. In particular, globally-averaged NO_3^- concentrations range from 10.2 to 66.2 $\mathrm{mmol\,m^{-3}}$ and integrated N_2 fixation from 0 to 518 $\mathrm{Tg\,N\,yr^{-1}}$. Tracers in OPEM and OPEM-H show similar ranges, except for globally averaged NO_3^-, which ranges from 10.2 to 66.2 $\mathrm{mmol\,m^{-3}}$ in OPEM and 13.0 to 55.0 $\mathrm{mmol\,m^{-3}}$ in OPEM-H.

Table 2. Ranges of global averages of major tracer concentrations or fluxes in the OPEM and OPEM-H configurations. Chl ~~concentrations~~ concentration is for the upper 50 m (surface layer of the UVic grid) and NCP is for the upper 100 m. Observations and reference model simulations are ~~depth integrated~~listed in the Reference column.

Tracer	OPEM	OPEM-H	Reference	Units
Oxygen	99.6–219	103–214	176[a]	$\mathrm{mmol\,m^{-3}}$
Nitrate	10.2–66.2	13.0–55.0	31[b]	$\mathrm{mmol\,m^{-3}}$
DIC	2.239–2.439	2.248–2.430	2.317[c]	$\mathrm{mol\,m^{-3}}$
DFe	0.47–0.71	0.47–0.69	0.57[d]	$\mathrm{\mu mol\,m^{-3}}$
PFe	0.44–0.75	0.44–0.70	1.17[d]	$\mathrm{nmol\,m^{-3}}$
Chl	~~37.6–101.2~~ 0.123–0.332	~~38.0–103.5~~ 0.128–0.336	0.309[e]	$\mathrm{mg\,m^{-2}}$ $\mathrm{^{-3}}$
NPP	27.8–88.0	27.2–88.0	52[f]	$\mathrm{Pg\,C\,yr^{-1}}$
NCP	~~0.86–3.01~~ 8.0–16.4	~~0.79–3.20~~ 7.8–16.2	13.5[g]	$\mathrm{Pg\,C\,yr^{-1}}$
N_2 Fixation	0–480	0–518	140[h]	$\mathrm{Tg\,N\,yr^{-1}}$

[a]WOA 2013 (Garcia et al., 2013a)

[b]WOA 2013 (Garcia et al., 2013b)

[c]GLODAPv2 (Olsen et al., 2016)

[d](Nickelsen et al., 2015),

[e]MODIS/Aqua level 3, 2008–2017 (Ocean Biology Processing Group, 2014)

[f](Westberry et al., 2008)

[g](Li and Cassar, 2016)

[h](Luo et al., 2012)

3.1 Sensitivity to Model Parameters

3.1.1 Biogeochemical tracer inventories and governing processes

The sensitivities of globally averaged biogeochemical properties to the variations of each of the 13 parameters in Table 2 are comparable for OPEM and OPEM-H (Figure 1). Global mean oxygen concentration is most sensitive to ν_{det} (remineralization rate). Higher ν_{det} increases oxygen consumption in shallow water, where oxygen resupply from the atmosphere is stronger. Less oxygen is consumed below the surface ocean, hence the total oxygen inventory increases. Maximum ingestion rate ($g_{\max}$) and grazing rate on ordinary phytoplankton (ϕ_{phy}) also correlate positively with oxygen. Higher $g_{\max}$ or ϕ_{phy} means more ordinary phytoplankton is grazed and less particles are formed, which then decreases oxygen consumption through remineralization. Oxygen is less sensitive to ϕ_{dia}, because the biomass of diazotrophs is much smaller than that of ordinary phytoplankton.

A surprising finding is that oxygen is sensitive to, and positively correlated with, the subsistence nitrogen quota of ordinary phytoplankton ($Q^{\text{N}}_{0,\text{ phy}}$). From a classic point of view, oxygen levels in the ocean are dominated by physical supply processes as well as biogeochemical consumption processes such as remineralization (Feely et al., 2004). Nevertheless, in our simulations the sensitivity to $Q^{\text{N}}_{0,\text{ phy}}$ is more than half (58%) of that to ν_{det} in OPEM and 48% in OPEM-H (Figure 1). In our model, $Q^{\text{N}}_{0,\text{ phy}}$ has no effect on the spatial distribution of cellular C:N ratios in phytoplankton, which is determined by ambient light and nutrient conditions. However, $Q^{\text{N}}_{0,\text{ phy}}$ affects the average phytoplankton C:N ratio. The average phytoplankton C:N ratio decreases when $Q^{\text{N}}_{0,\text{ phy}}$ increases, with less carbon being fixed for the same NO_3^- supply. Oxygen consumption (due to remineralization) per mole of nitrogen thus decreases in consequence. $Q^{\text{N}}_{0,\text{ phy}}$ in turn affects NO_3^-: A higher $Q^{\text{N}}_{0,\text{ phy}}$ yields a higher oxygen level and hence less denitrification in oxygen deficient zones (ODZs) and therefore leads to more NO_3^-. In fact, we identify this as a major process that controls the NO_3^- inventory in our simulations (Figure 1). While NO_3^- is also sensitive to other parameters, its sensitivity to $Q^{\text{N}}_{0,\text{ phy}}$ is more than twice that to any other parameter (Figure 1).

The sensitivity of dissolved inorganic carbon (DIC) is generally low, because of the relatively large DIC pool compared to the variations in fluxes among the different parameter sets. Similar to oxygen, DIC is most sensitive to ν_{det}, $Q^{\text{N}}_{0,\text{ phy}}$, $g_{\max}$ and ϕ_{phy}. Faster carbon recycling in the surface layer due to higher ν_{det} generates a higher surface DIC concentration and hence more outgassing, which decreases the DIC inventory. A somewhat lower DIC inventory is also induced by a larger $Q^{\text{N}}_{0,\text{ phy}}$, as less carbon is fixed and exported per unit nitrogen in phytoplankton, and by enhanced zooplankton grazing with larger $g_{\max}$.

Dissolved iron (DFe) is most sensitive to the remineralisation rate (ν_{det}). Unlike NO_3^-, which has dynamic source (N_2 fixation) and sink (denitrification) processes, iron has a fixed source from atmospheric deposition and ~~the~~ a sink in the sediment, and the size of the DFe pool is mainly determined by its internal cycle. A higher remineralisation rate prolongs the residence time and thus increases the DFe pool. The parameter ν_{det} also indirectly affects the internal DFe cycle via its effect on O_2. While the detritus remineralisation rate drops when O_2 falls below $5\,\text{mmol}\,\text{m}^{-3}$ (Nickelsen et al., 2015), scavenging of DFe stops below the same oxygen threshold. Detritus remineralisation rate dominates variations in DFe when globally averaged O_2 is above $135\,\text{mmol}\,\text{m}^{-3}$, in which case DFe is positively correlated with ν_{det} and O_2. When globally averaged O_2 is below $135\,\text{mmol}\,\text{m}^{-3}$, the wide-spread ODZs (below $5\,\text{mmol}\,\text{m}^{-3}$) inhibit the scavenging of DFe and this effect dominates. As a result, DFe becomes anti-correlated with O_2. Particulate iron (PFe) is also positively correlated with ν_{det} when globally

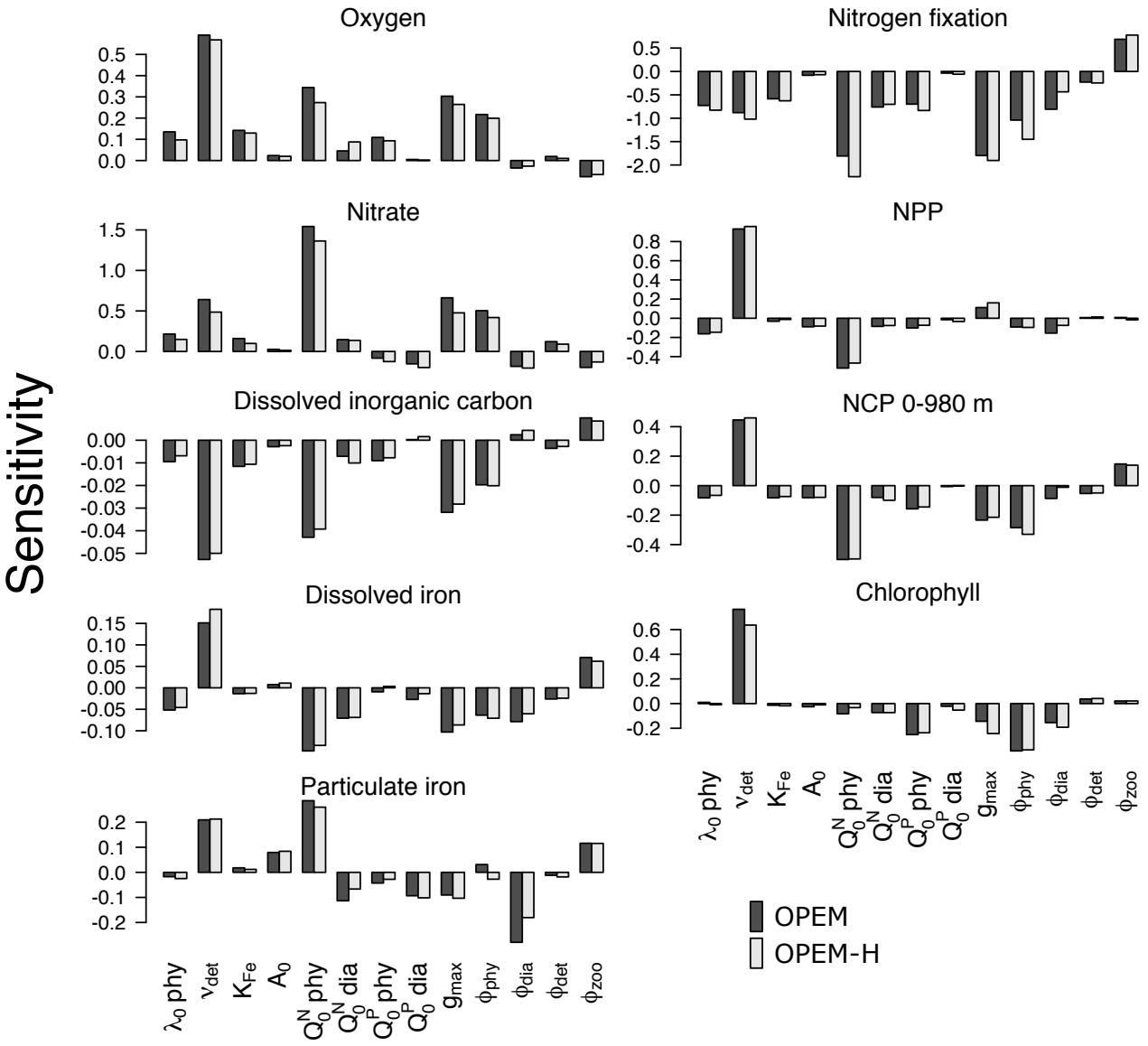

Figure 1. Sensitivities of globally averaged O_2, NO_3^-, dissolved inorganic carbon, dissolved iron, particulate iron, N_2 fixation, net primary production (NPP), Chlorophyll, and net community production (NCP) integrated from 0 to 980 m to individual model parameters, computed according to Eq. (3). Note the different ~~vertical scales~~ y-axis ranges in the different panels.

averaged O_2 is above $135\,\mathrm{mmol\,m^{-3}}$, but below that PFe shows no correlation with ν_{det}. When globally averaged O_2 is below $135\,\mathrm{mmol\,m^{-3}}$, inhibition of scavenging of DFe in ODZs decreases PFe there but a higher DFe increases PFe elsewhere, because PFe is coupled to DFe through scavenging and remineralisation. As mentioned above, $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$ controls the average nitrogen quota in phytoplankton and thus in particles. Since PFe is proportional to the amount of nitrogen in particles, $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$

also affects PFe. This (positive) sensitivity is much stronger than the indirect (negative) effect via DFe leading to opposite sensitivities of DFe and PFe to $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$. Other than ν_{det} and $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$, PFe is also sensitive to ϕ_{dia} because dead diazotrophs enter the particulate pool (detritus) and diazotrophs are very sensitive to ϕ_{dia} (Figure 2).

~~No single parameter dominates the sensitivity of N_2 fixation in the simulations (Figure 1), which resembles the result of Tang et al. (2019) that no single environmental property predicts~~ The simulated global N_2 fixation rate is sensitive to many parameters, apart from $A_{0,\,\mathrm{phy}}$ and $Q^{P}_{0,\,\mathrm{dia}}$. Similar relative changes in most parameters introduce changes to the global N_2 fixation ~~, even with a data-based machine-learning method~~rate that are of similar magnitude. Interestingly, ~~other than ν_{det} and $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$,~~ N_2 fixation is ~~also sensitive~~ sensitive also to zooplankton parameters, indicating that zooplankton grazing on diazotrophs is an important factor controlling not just diazotroph biomass but also N_2 fixation~~. Compared to $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$, g_{max}, ϕ_{zoo} and ν_{det}, N_2 fixation is not very sensitive to the iron half-saturation constant $k_{\mathrm{Fe,\,phy}}$, probably because iron limitation occurs mainly in regions where relatively high nitrate concentrations impede N_2 fixation anyway.~~ .

Of particular interest are the sensitivities of global net primary production (NPP) and net community production (NCP). Particle fluxes in marine biogeochemical models tend to agree most closely with sediment trap data for depths of about $1000\,\mathrm{m}$ or below (Kriest et al., 2012). Therefore, different from Table 2, showing NCP for the upper $100\,\mathrm{m}$ for comparison with observations and other (reference) model simulations, here we integrate NCP from 0 to $980\,\mathrm{m}$ (7^{th} layer of the ocean in the UVic-ESCM), which in steady state is equivalent to POC export flux at $980\,\mathrm{m}$. NPP is sensitive to ν_{det} and $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$. A higher ν_{det} causes faster nutrient recycling in surface waters, which increases NPP and reduces particle export and hence NCP. Increasing $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$ lowers both NPP and NCP and hence also the fixed-carbon inventory. A higher ingestion rate of zooplankton (g_{max}) removes more particles and thus is negatively correlated with NCP. Chl is the principal agent of C fixation in the OPEM and hence Chl has a similar sensitivity pattern as NPP except for g_{max} and ϕ_{phy}.

3.1.2 Ordinary phytoplankton, diazotrophs, particles, export and their elemental stoichiometry

First we discuss the proportions of carbon, nitrogen and phosphorus in ordinary phytoplankton and diazotrophs, since variations in elemental stoichiometry in autotrophs originate in differential uptake of nutrients under different environmental conditions.

Globally averaged C, N, P concentrations and ratios of globally averaged N and P of ordinary phytoplankton and diazotrophs are sensitive to ν_{det}, $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$, ϕ_{phy} and ϕ_{dia} (Figure 2). As expected, C, N and P of ordinary phytoplankton and diazotrophs increase for higher ν_{det}, which generates higher nutrient concentrations in the surface ocean. They are also sensitive to zooplankton grazing, especially to ϕ_{phy} and ϕ_{dia}. $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$ and $Q^{\mathrm{P}}_{0,\,\mathrm{phy}}$ are negatively correlated with ordinary phytoplankton C, indicating that the negative effect of higher subsistence quotas on competitive ability dominates their effect on biomass. A similar behavior is found in diazotrophs except that $Q^{\mathrm{N}}_{0,\,\mathrm{dia}}$ is also negatively correlated with diazotroph N and hence also nitrogen fixation (Figure 1). Although an increase in $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$ makes ordinary phytoplankton less competitive, it also raises the oceanic NO_3^- inventory, which eventually leads to more phytoplankton N (Figure 2) and less nitrogen fixation (Figure 1).

Diazotroph C, N and P are generally more sensitive to parameter variations than phytoplankton, due to the much smaller total biomass of diazotrophs, which is also the reason why diazotrophs are less sensitive in OPEM-H, the model configuration in which their biomass is generally larger ~~(Figure 2)~~because of the growth of diazotrophs at high latitudes (see Fig. 15 in Part I, Pahlow et al., 2

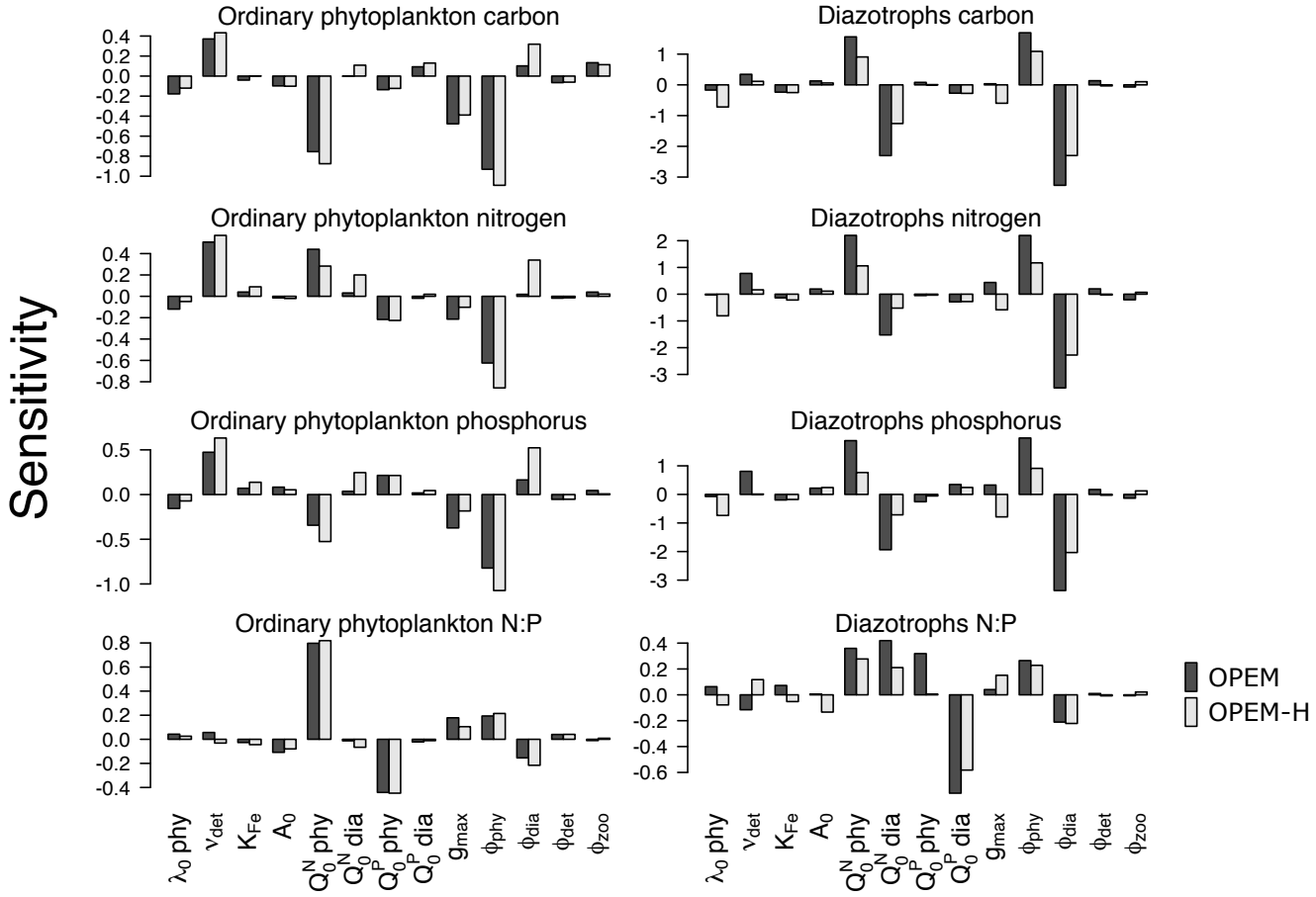

Figure 2. Parameter sensitivities of globally averaged concentrations of ordinary phytoplankton and diazotrophs carbon, nitrogen, phosphorus, and ratios of globally averaged N and P. Black and grey shading denote OPEM and OPEM-H configurations, respectively. Note the different y-axis ranges in the different panels.

. Since ordinary phytoplankton dominates autotrophic biomass, it tends to control nutrient distributions. This explains why ordinary phytoplankton parameters such as $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$ and ϕ_{phy} have strong effects on diazotrophs but not vice versa. The zooplankton grazing preferences ϕ_{phy} and ϕ_{dia} drive the competition between ordinary phytoplankton and diazotrophs and hence have strong and opposing effects on their biomass. Owing to the relatively small total biomass, diazotroph C is more sensitive to changes in ϕ_{phy} and ϕ_{dia} than ordinary phytoplankton C.

Particulate C:N and N:P ratios are most sensitive to $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$ (Figure 3). This sensitivity is related to biomass, as we see from the OPEM-H configuration, where (non-N_2 fixing) diazotrophs are abundant ~~in high latitudes~~ at high latitudes (see Fig. 15 in Part I, Pahlow et a and consequently the sensitivity of high-latitude C:N to $Q^{\mathrm{N}}_{0,\,\mathrm{dia}}$ is high, even higher than to $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$ (Figure 3). We do not find this behavior for high-latitude regions in the OPEM configuration, as well as low-latitude regions, where diazotrophs are not

as abundant. The parameter $Q^{P}_{0,\,phy}$ was expected to be the most important parameter for particulate C:P ratios, just like $Q^{N}_{0,\,phy}$ is for the C:N ratio. However, this is only true for the OPEM at high latitudes.

At low latitudes, ~~and for the global ocean,~~ particulate C:P ratios are most sensitive to $Q^{N}_{0,\,phy}$ (Figure 3). The supply of nitrate and phosphate at different latitudes is the major reason for this pattern. ~~Phosphate is not a limiting nutrient~~ At low latitudes, the effects of $Q^{P}_{0,\,phy}$ are suppressed by variations in phytoplankton C, which is affected by $Q^{N}_{0,\,phy}$ and the consequent change in nitrate concentration. Nitrate and phosphate are not limiting in the high-latitude Southern Ocean where, under N- and P-replete conditions, cellular C:P is mainly determined by $Q^{P}_{0,\,phy}$ and a higher $Q^{P}_{0,\,phy}$ would result in a higher cellular P:C (lower C:P). Therefore, the ~~cellular~~ global C:P ~~ratio of~~ of total particulate matter, which is dominated by ordinary phytoplankton, ~~which dominates total particles,~~ is negatively correlated with $Q^{P}_{0,\,phy}$. ~~The effects of $Q^{P}_{0,\,phy}$ are suppressed by the prevalence of N limitation in low latitude regions, and hence $Q^{N}_{0,\,phy}$ affects particulate C:P variations more than $Q^{P}_{0,\,phy}$.~~

The sensitivities of dissolved N:P ratio to parameters in the three geographical settings (low, high latitudes and global) follow similar patterns. However, we find sensitivities to be generally higher in the low-latitudes, especially to variations of the phytoplankton parameters. Again this is because NO_3^- is often limiting in lower latitudes, particularly in the oligotrophic gyres, where the dissolved nitrogen pool is more sensitive to changes in phytoplankton as well as N_2 fixation. This is also why grazing pressure on diazotrophs (ϕ_{dia}) has a much stronger effect at low than at high latitudes.

3.2 Cost function values of the ensemble simulations

3.2.1 Constraining global rate estimates and inventories

The cost function (introduced in Section 2.2.2) was devised for identifying the best solutions among the ensemble runs. For the model's upper layers ($0 - 550\,\mathrm{m}$) observational monthly mean concentrations of nitrate and phosphate enter the cost function, thereby reflecting regional and seasonal variations in the N:P uptake ratio of ordinary phytoplankton and diazotrophs. Variations in nitrate and phosphate availability affect the growth of diazotrophs and thus determine global N_2 fixation in both OPEM and OPEM-H. In our UVic configurations, water column denitrification is the only fixed-N loss term. Therefore, the simulated N_2 fixation is expected to match water column denitrification under a steady-state nitrogen cycle. Nevertheless, the simulation with the lowest cost yields a global N_2-fixation rate estimate of $38.8\,\mathrm{Tg\,N\,year^{-1}}$ (Figure 4A), much lower than recent estimates of water column denitrification (55.8 - $72.9\,\mathrm{Tg\,N\,year^{-1}}$; Somes et al., 2017; Wang et al., 2019).

The cost function penalises solutions that yield N_2 fixation rates greater than $90\,\mathrm{Tg\,N\,year^{-1}}$, but shows no clear relation to N_2 fixation at lower rates (Figure 4A). For example, among the simulations with the 5 lowest cost function values in the OPEM configuration, the global ocean N_2 fixation rate varies between 8 and $40\,\mathrm{Tg\,N\,year^{-1}}$. These model solutions also differ with respect to their O_2 inventories. The tendency of the cost function to favor very low global N_2 fixation is caused by a compensatory effect, whereby improving NO_3^- deteriorates O_2 and vice versa ~~(see also ? and the Discussion section below)~~(see also Part I, Pahlow et al., 2020, and the Discussion section below). Thus, instead of selecting the reference parameter sets based only on the cost function, we also take the ability to yield reasonable N_2 fixation rates into account, whereby we ~~deem rates of about as reasonable~~ignore simulations with rates below $60\,\mathrm{Tg\,N\,year^{-1}}$, since this ~~matches current~~

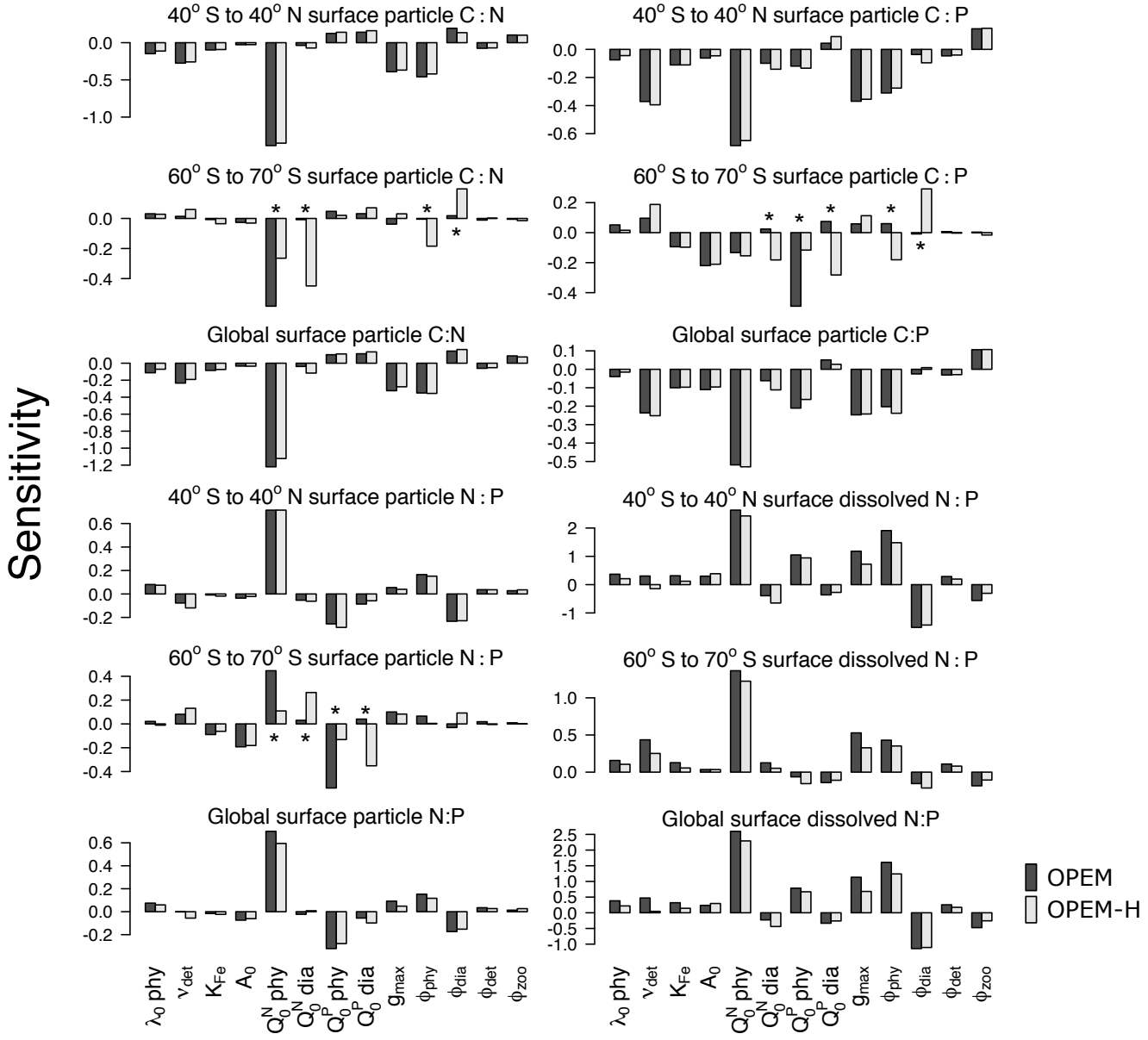

Figure 3. Parameter sensitivities of averaged surface (0–130 m) particulate elemental C:N, C:P, and N:P ratios for different latitude bands (40°S to 40°N, 60°S to 70°S, and the global ocean). Asterisks indicate sensitivities that are very different between OPEM and OPEM-H. Note the different y-axis ranges in the different panels.

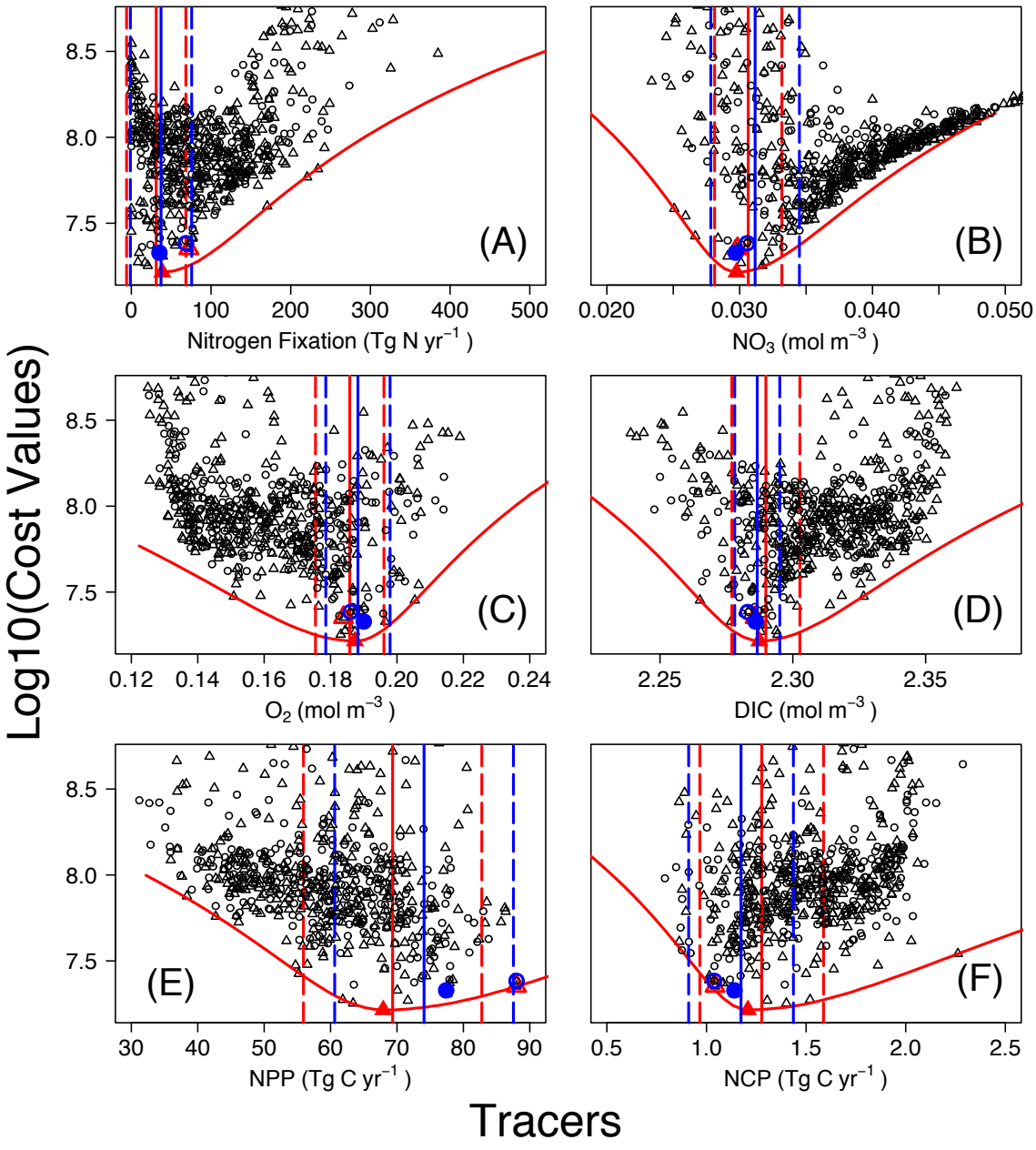

Figure 4. Costs vs. tracer concentrations and fluxes for annual N_2 fixation (A), globally averaged NO_3^- (B), O_2 (C) and dissolved inorganic carbon (DIC) (D) concentrations, as well as annual net primary production (NPP) (E) and net community production (NCP, here integrated over the depth range 0 to 980 m) (F). Red and blue symbols and lines are for OPEM (triangles) and OPEM-H (circles), respectively. Solid and open symbols represent ~~minimum-cost~~ minimum-cost and trade-off simulations, respectively. Vertical solid and dashed lines represent mean and 95% confidence interval of best solutions of 1000 randomly selected subsets of 100 ensemble members. Red parabolas fit the lowest costs at different rates or tracer concentrations.

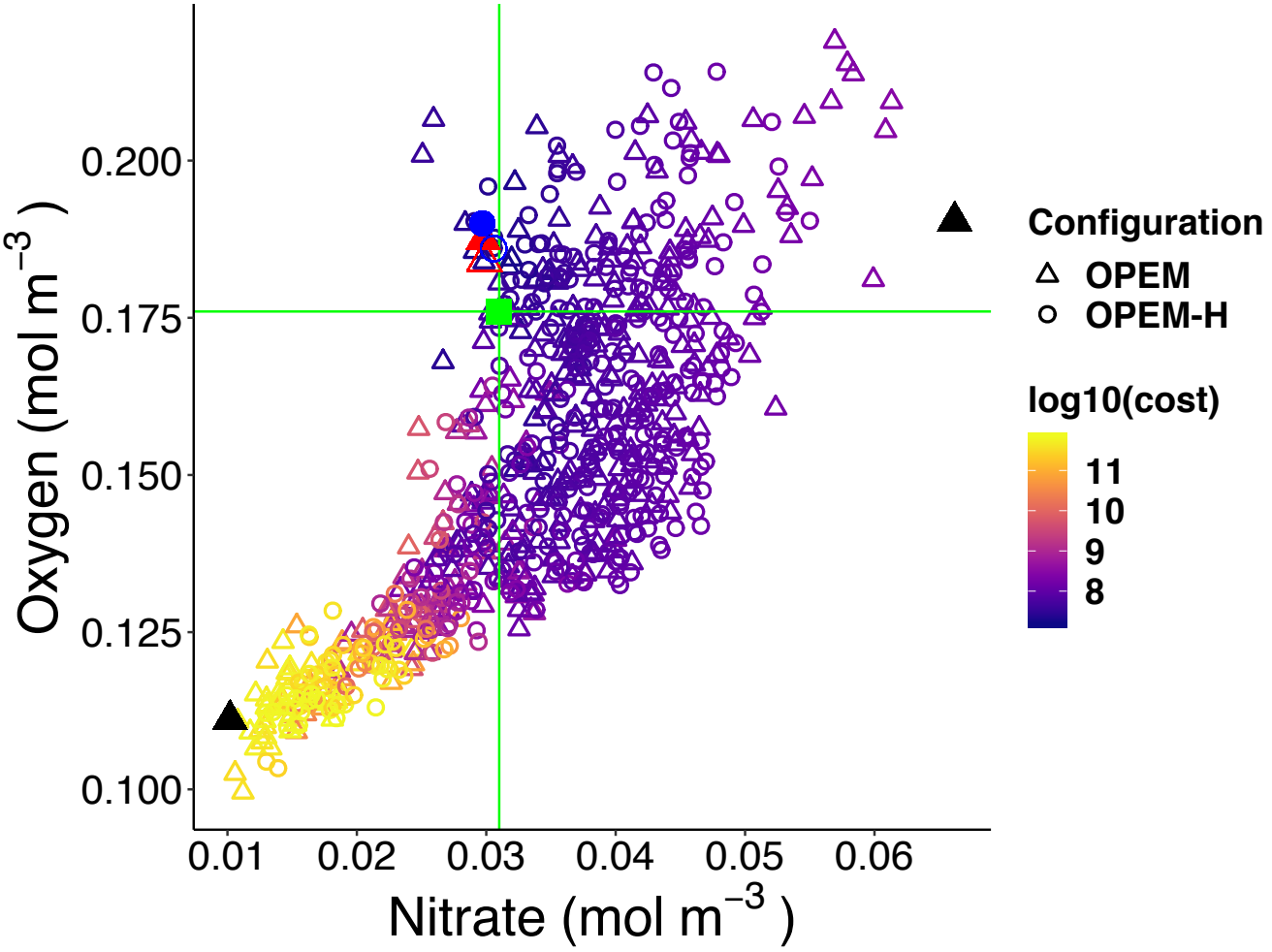

Figure 5. Globally averaged oxygen vs. nitrate in OPEM and OPEM-H. Color represents cost value. Solid red triangle and blue circle annotate the simulations with minimum cost in OPEM and OPEM-H, respectively, and open red triangle and blue circle are the trade-off simulations. The green square ~~indicates~~, horizontal and vertical lines indicate mean oxygen and nitrate concentrations of $0.176\,\mathrm{mol\,m^{-3}}$ and $0.031\,\mathrm{mol\,m^{-3}}$, respectively, in the WOA ~~2013 value~~2013. Solid black triangles highlight the lowest and highest NO_3^- simulations used in Figure 6 and 7.

is the lower boundary of current data-based estimates of water-column denitrification ~~(Somes et al., 2017; Wang et al., 2019)~~ (DeVries et al., 2012). As these solutions represent a somewhat subjective trade-off between low cost and reasonable N_2 fixation, we refer to them as trade-off solutions and details of their behaviour are shown and discussed ~~in the companion paper ?~~in Part I (reference simulations in Pahlow et al., 2020). For OPEM the trade-off solution corresponds to the seventh-lowest cost function value, and the fourth-lowest for OPEM-H.

~~To understand the uncertainty range of our model results, we apply a bootstrap method to obtain an uncertainty quantification for our N_2 fixation rate estimates, based on the available ensemble model runs. We collect the best solutions (lowest cost function value) of 1000 randomly selected subsets of 100 ensemble members. Mean and 95% confidence interval of these estimates provide an uncertainty range in the vicinity of the N_2 fixation rate estimate of the full ensemble. Globally averaged N_2 fixation rates of our trade-off solutions of OPEM and OPEM-H are just outside and within this uncertainty range, respectively (Figure 4A).~~ In the following we will describe the lowest-cost solutions together with the trade-off solutions, as well as respective uncertainty ranges obtained from the bootstrap method described in the Materials and Methods section. The width of the uncertainty ranges (95% confidence intervals) in Figure 4 indicates the metric's ability to constrain the inventory or rate under consideration.

Globally averaged N_2 fixation rates of our trade-off solutions of OPEM and OPEM-H are just outside and within this uncertainty range, respectively (Figure 4A). The global NO_3^- inventory turns out to be remarkably well constrained (Figure 4B). The mean global estimates are $30.6\,\mathrm{mmol\,N\,m^{-3}}$ and $31.4\,\mathrm{mmol\,N\,m^{-3}}$ for OPEM and OPEM-H, respectively. Ensemble solutions that deviate from these estimates have high costs and therefore the uncertainty ranges remain narrow. The trade-off and minimum-cost solutions are hardly distinguishable. The uncertainty of the simulated global O_2 is comparable to that of the NO_3^- inventory. Global mean O_2 concentrations of OPEM and OPEM-H are $186\,\mathrm{mmol\,O_2\,m^{-3}}$ and $187\,\mathrm{mmol\,O_2\,m^{-3}}$. Our metric effectively constrains global DIC estimates, $2.290\,\mathrm{mol\,C\,m^{-3}}$ for OPEM and $2.287\,\mathrm{mol\,C\,m^{-3}}$ for OPEM-H (Figure 4D), although DIC data have not been explicitly considered in the cost function.

While the trade-off solutions exhibit NO_3^-, O_2 and DIC inventories well within their respective uncertainty ranges, we find somewhat larger deviations for the predicted global mean net primary production (NPP, Figure 4E). For OPEM and OPEM-H the trade-off solutions produce a, respectively, 30 % and 14 % higher NPP than the minimum-cost solutions. The net community production (NCP, here integrated over the depth range 0 to 980 m) estimates in Figure 4F are better constrained than NPP for both configurations. The trade-off solution of OPEM corresponds to a global NCP of $1.043\,\mathrm{Tg\,C\,year^{-1}}$, which is close to the trade-off estimate of OPEM-H, where NCP $= 1.039\,\mathrm{Tg\,C\,year^{-1}}$.

Figure 5 shows globally averaged concentrations of O_2 versus NO_3^- of all ensemble members. The spread of the ensembles differs between the two tracers (by a factor of two for O_2 and by a factor of six for NO_3^-). Most solutions overestimate the global average NO_3^- concentration obtained from the WOA 2013 (Garcia et al., 2013a, b) and underestimate O_2. Solutions where both tracers strongly underestimate the WOA 2013 data are penalised by the cost function (Figure 5). The minimum-cost and trade-off solutions of OPEM and OPEM-H are close to the WOA 2013 estimates. The respective optimal solutions have slightly higher global mean O_2 concentrations than the WOA 2013 and are in good agreement with respect to NO_3^-. In spite of larger costs, the trade-off solutions of both OPEM and OPEM-H are closer to the WOA 2013 estimate than the minimum-cost

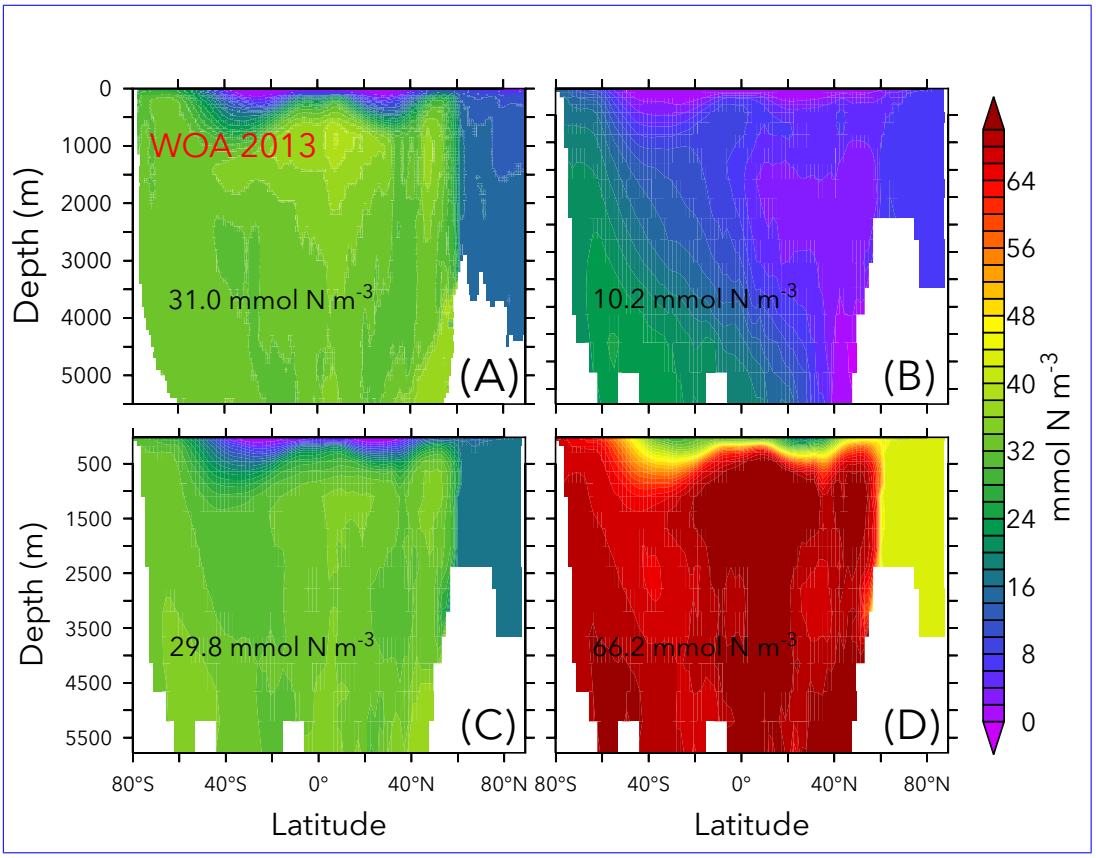

Figure 6. Zonally averaged NO_3^- in ~~low-~~ the World Ocean Atlas 2013 (A), the simulations with the lowest and ~~high-~~highest NO_3^- inventory (B, D), and the trade-off ~~simulations for OPEM~~ simulation (~~upper row~~C) ~~and OPEM-H (lower row)~~in the OPEM configuration. Globally averaged NO_3^- concentrations are shown in each panel. Simulations shown here are marked with solid black and open red triangles in Figure 5. Note that the outputs from OPEM and OPEM-H are very similar and only OPEM results are shown here.

solutions (Figure 5). ~~Overall, we stress that the minimum-cost and trade-off solutions appear at the margin of the full spread of the ensembles, which could be interpreted as indicating a model deficiency~~The ensemble solutions are unevenly spread around the WOA 2013 data-based estimates. This highlights that our trade-off solutions could not have been identified had we only considered the ensemble means.

Figures 6 and 7 show zonally averaged NO_3^- and O_2 in simulations with ~~low and high~~ the lowest and highest NO_3^- and the trade-off ~~simulations~~simulation in the OPEM configuration. The high-NO_3^- ~~simulations have~~ simulation has similar NO_3^- and O_2 patterns to the trade-off ~~simulations~~simulation, despite the very different mean NO_3^- and O_2 concentrations. The patterns are different in the low-NO_3^- ~~simulations~~ simulation because of stronger deoxygenation and denitrification, which occur mostly in North Pacific deep water. The greater similarity of global mean O_2 than NO_3^- reflects the influence of atmospheric O_2 but also indicates that NO_3^- is more sensitive to changes in the physiology of the diazotrophs.

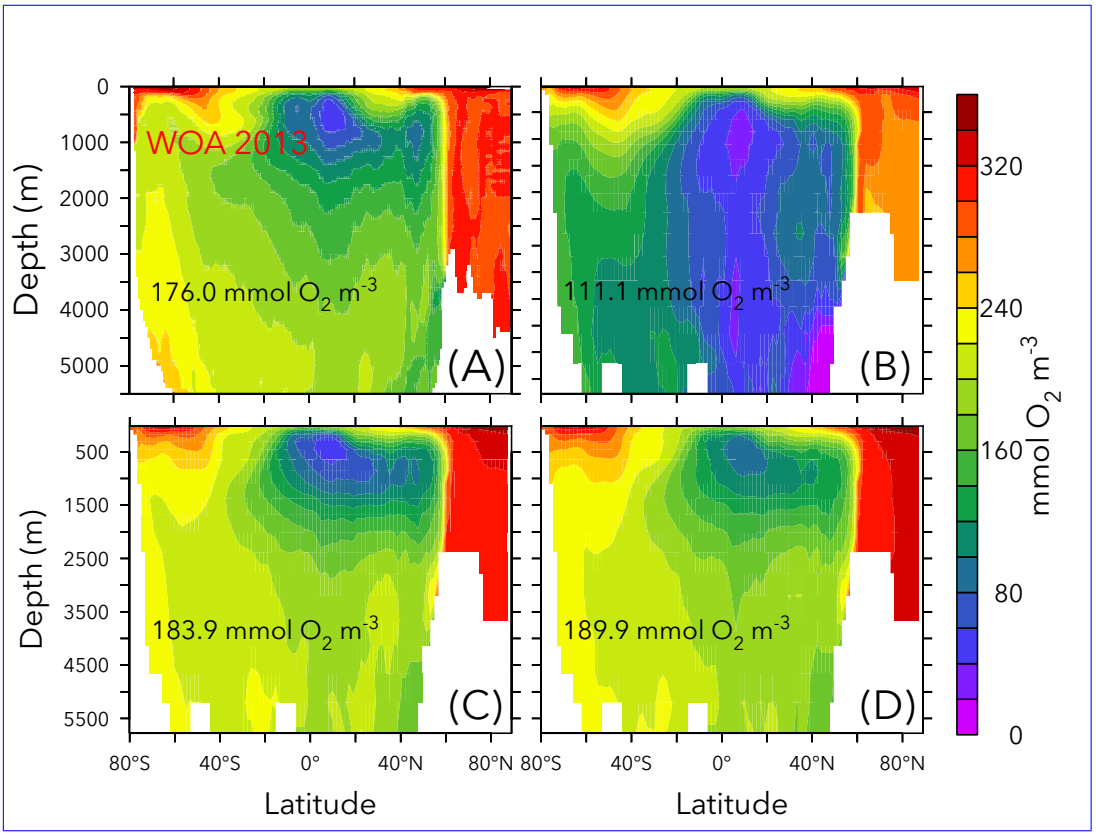

Figure 7. Same simulations as in Figure 6 but showing the results for O_2.

3.2.2 How well can model parameters be constrained?

Cost is conspicuously correlated only with ν_{det}, $Q^{\text{N}}_{0,\,\text{phy}}$, and ϕ_{dia} (Figure 8). O_2 and NO_3^- are sensitive to ν_{det} and $Q^{\text{N}}_{0,\,\text{phy}}$ but not to ϕ_{dia} (Figure 1), which indicates that ϕ_{dia} becomes more important at lower-cost simulations. The minimum-cost and trade-off simulations in OPEM and OPEM-H are usually closer to each other when parameters show strong correlations with costs (Figure 8).

Figure 9 shows how different biomes contribute to the misfit and variance parts of the total cost. For simulations with high cost function values ($J > 10^{10}$), we find the variance term to be dominant in the deep ocean (below 550 m). Among the 17 biomes this is well expressed in NP.SPSS (North Pacific subpolar seasonally stratified), NP.STSS (North Pacific subtropical seasonally stratified), NP.STPS (North Pacific subtropical permanently stratified), Pac.EQU.E (Eastern Pacific equatorial), Pac.EQU.W (Western Pacific equatorial), and IND.STPS (Indian Ocean subtropical permanently stratified) biomes, overwhelming contributions from all other parts of the cost function and all other biomes for the 100 simulations with the highest total costs. These high-cost simulations tend to have low NO_3^- and O_2 concentrations (Figure 5). Low NO_3^- concentrations are coupled to low O_2 because of intense denitrification in the ~~ODZs~~oxygen deficient zones (ODZs). Accordingly,

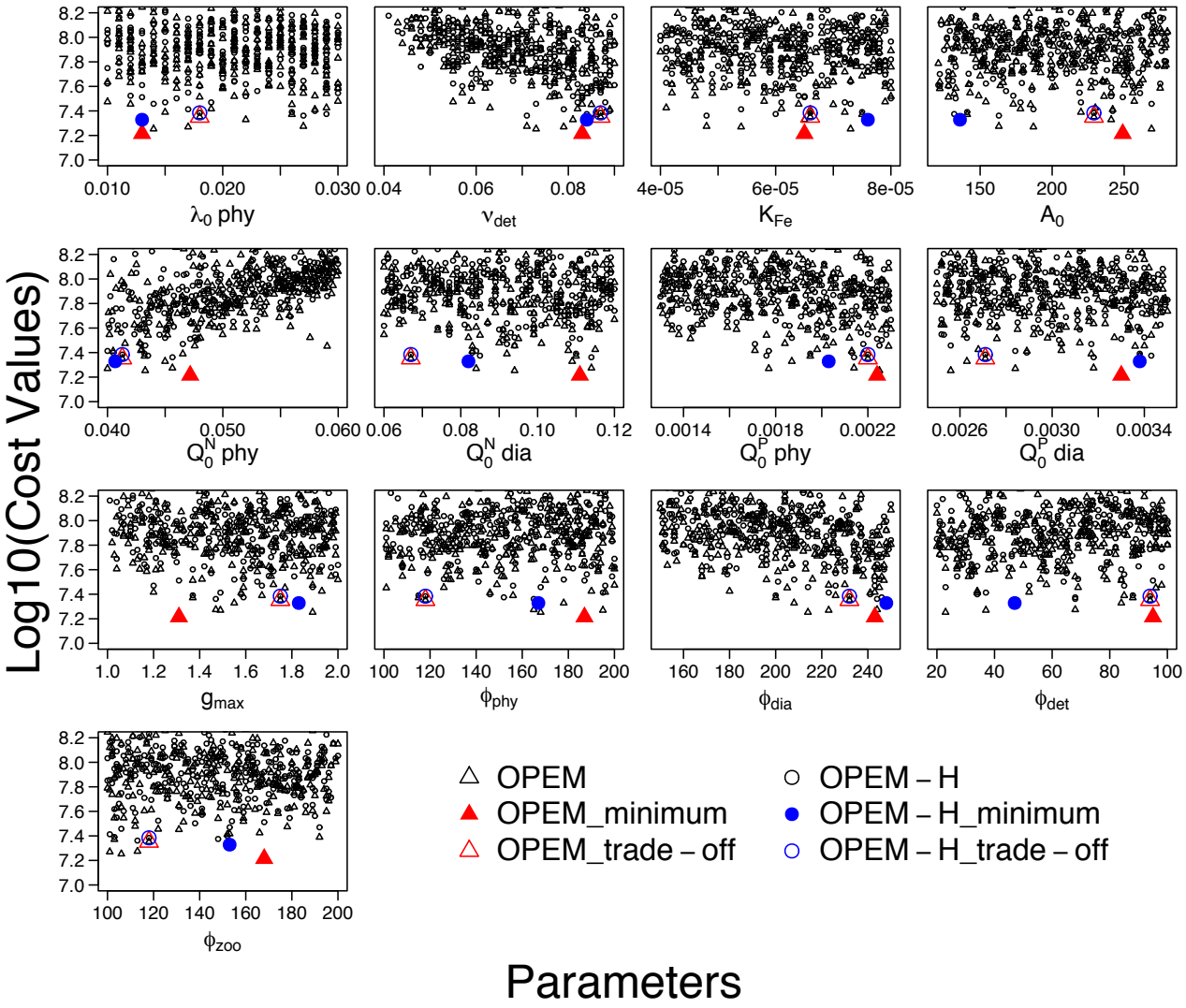

Figure 8. Lower parts (cost $< 10^{8.2}$) of cost-value distributions for the parameter ranges in Table 1. Solid red triangles and blue circles represent the minimum-cost simulations in OPEM and OPEM-H, respectively, and open red triangles and blue circles are the trade-off simulations. Note that the trade-off simulations share the same parameter combination but have slightly different cost-values.

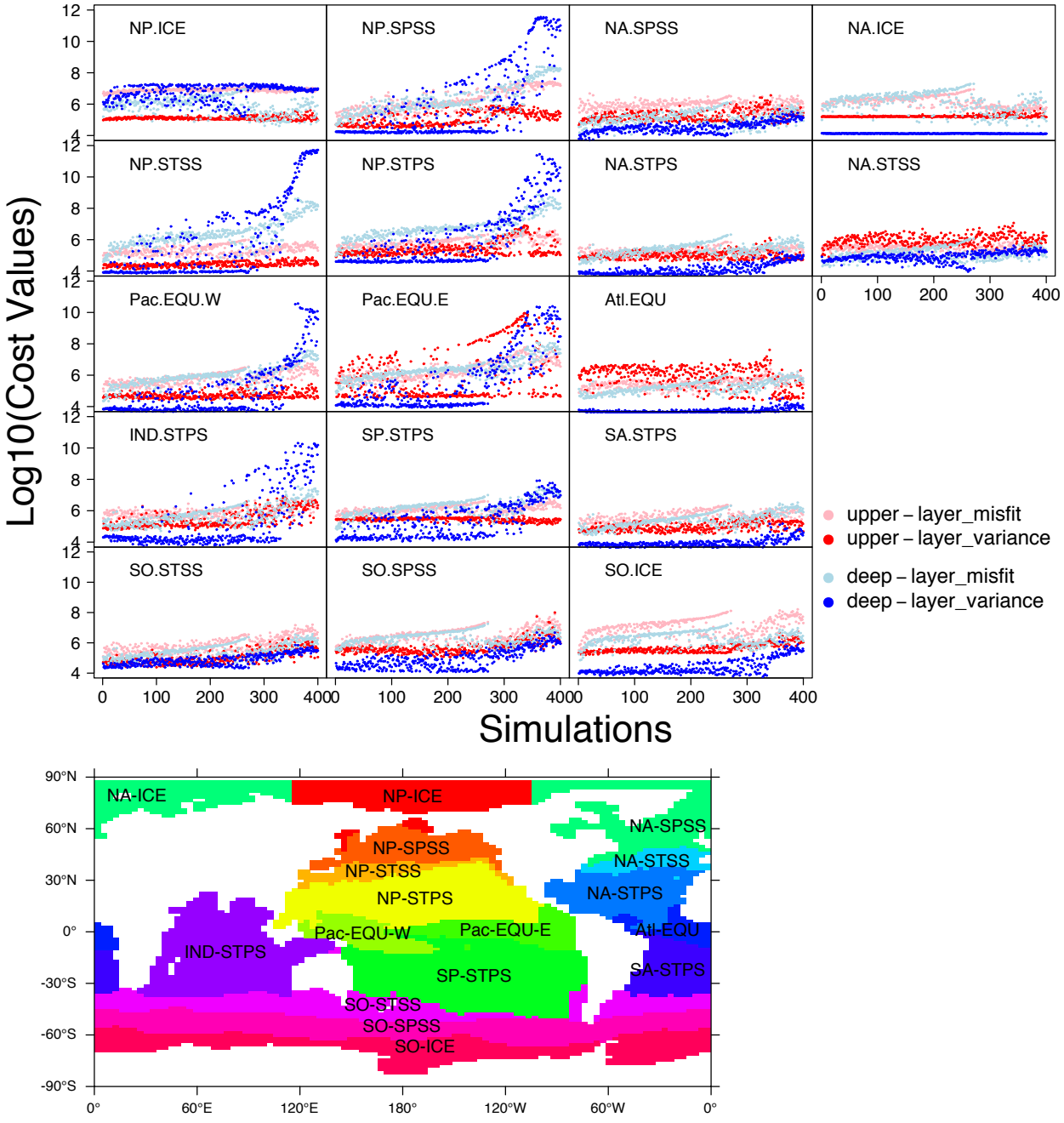

Figure 9. Top panels: Cost-value distributions in the 17 biomes in OPEM. The order of the simulations is based on the total cost from low to high in OPEM. Upper-layer and deep-layer in the legend represent upper ($0 - 550\,\mathrm{m}$) and lower (below $550\,\mathrm{m}$) components of the cost function (Eq. 5). Misfit and variance are calculated by the first and second parts of the cost function components (Eqs. 6 and 7), respectively. Bottom: Map of biome locations.

simulations with very low NO_3^- inventories suffer from widespread ODZs, occupying much of the deep water in the northern and equatorial Pacific as well as the Indian Ocean (Figure ~~6~~S1). This is the main reason for the high variance in the deep water of these biomes (Figure 9).

4 Discussion

4.1 Parameter sensitivities

4.1.1 Remineralisation rate ν_{det} and phytoplankton subsistence nitrogen quota $Q^N_{0,\,phy}$

Remineralisation rate (ν_{det}) and phytoplankton subsistence nitrogen quota ($Q^N_{0,\,phy}$) are the two parameters with the strongest correlations for most tracers as well as particulate elemental stoichiometry. The importance of ν_{det} was expected, because it is an important driver of nutrient recycling in the surface ocean (Thomas, 2002; Anderson and Sarmiento, 1994; Eppley and Peterson, 1979), which strongly affects NPP, NCP, Chl, DIC, DFe and N_2 fixation (Kriest et al., 2012). ν_{det} also determines the rate of O_2 consumption, hence also the NO_3^- level, due to denitrification in ODZs (Cavan et al., 2017). The strong influence of $Q^N_{0,\,phy}$, however, was unexpected. The subsistence quota was first introduced by Droop (1968) in phytoplankton growth models. While it has been applied in Earth System Models (Kwiatkowski et al., 2018; Wang et al., 2019), a sensitivity analysis similar to the present study has not been done before. A higher $Q^N_{0,\,phy}$ implies that more nitrogen is required for phytoplankton growth, but it also can be interpreted as a lessening of carbon fixation for a given nitrogen supply. Our results demonstrate a strong effect of $Q^N_{0,\,phy}$ on NPP, Chl, POC export (NCP, here integrated over the depth range 0 to 980 m) and consequently oxygen consumption and denitrification.

These results also put forward a new point of view on the relation between NO_3^- inventory and carbon export. In classic biogeochemistry, a larger NO_3^- inventory in the ocean stimulates primary production and POC export. This feedback is intuitive and easy to understand, as for a given C:N in phytoplankton, carbon is proportional to the nitrogen pool. This feedback is well recognized and has been widely applied in marine sciences, especially since it forms the foundation of one of the hypotheses explaining the lower atmospheric pCO_2 during the last glacial maximum (LGM) (McElroy, 1983; Falkowski, 1997). However, our analysis of the model ensemble with different parameter combinations suggests another, very different point of view. NO_3^- concentration is positively correlated with $Q^N_{0,\,phy}$, but negatively with NPP and POC export (NCP, Figure 1), which means that an increased NO_3^- inventory can be related to a lower POC export if caused by a change in $Q^N_{0,\,phy}$. The dynamic C:N ratio in our model explains part of this negative correlation. When the NO_3^- inventory increases due to an increase in $Q^N_{0,\,phy}$, the nitrogen demand in phytoplankton also increases, which yields a lower C:N ratio in phytoplankton, and hence changes in carbon fixation due to increases in NO_3^- inventory remain relatively small. The increase in $Q^N_{0,\,phy}$ increases nitrogen in phytoplankton structure and decreases the C:N ratio in phytoplankton as well as detritus. The two effects together both lower POC production and raise the NO_3^- inventory. Changes in ν_{det} also contribute to the negative correlation between NO_3^- and POC export (NCP) in our simulations: ~~A~~ a more intense remineralisation in the surface ocean reduces POC export, and thus decreases oxygen consumption and denitrification, resulting in a larger nitrate inventory.

The strong impact of $Q^{N}_{0,\,phy}$ on the NO_3^- inventory and globally averaged phytoplankton C:N causes a higher sensitivity of globally averaged C:N than C:P (Figure 3). A higher $Q^{N}_{0,\,phy}$ results in a higher NO_3^- inventory and a lower phytoplankton C:N, both tending to lower particulate C:N and vice versa. On the other hand, C:P is not as sensitive because we have a constant PO_4^{3-} inventory in the UVic model. Surface particulate matter C:N is less variable compared to C:P and N:P in field observations along regional gradients (Galbraith and Martiny, 2015; Geider and Roche, 2002; Martiny et al., 2013a; Sterner and Elser, 200 , which is an apparent contrast to our results, where the sensitivity of C:N to $Q^{N}_{0,\,phy}$ is the highest among the particulate elemental ratios. However, our sensitivities are with respect to parameter variations among many simulations, rather than spatial or temporal gradients in the one real ocean.

4.1.2 Zooplankton parameters

While in many global biogeochemical models zooplankton is described by non-mechanistic formulations, such as Holling-type functions (Holling and Buckingham, 1976), in this study we apply a more realistic zooplankton model (Pahlow and Prowe, 2010). Among the five zooplankton parameters, the maximum specific ingestion rate ($g_{\max}$) and the capture coefficients of phytoplankton (ϕ_{phy}) and diazotrophs (ϕ_{dia}) are the most important, whereas the preference for detritus (ϕ_{det}) is generally less important. Grazing on zooplankton itself (ϕ_{zoo}) counters the effect of $g_{\max}$ because it lowers zooplankton biomass and thus total ingestion. These parameters together dominate controls on N_2 fixation and Chl (Figure 1), and C, N and P of ordinary phytoplankton and diazotrophs (Figure 2). It is interesting that zooplankton parameters also exert some control on particulate N:P as well as the dissolved nutrient pools (Figure 3). This can be understood via their controls on N_2 fixation and the ensuing changes in N:P in the dissolved and particulate pools.

4.1.3 Other parameters and the OPEM-H configuration

Other parameters in the sensitivity analysis appear less important for the tracer distributions, but this does not necessarily mean that they are negligible. Specific mortality rate ($\lambda_{0,\,phy}$) and the phytoplankton half-saturation constant for Fe ($k_{\text{Fe},\,phy}$) ~~do contribute some variations to~~ do contribute to some variations of most of the tracers (Figure 1), and particulate C:P is somewhat sensitive to potential nutrient affinity (A_0). Phytoplankton subsistence P quota ($Q^{P}_{0,\,phy}$) affects major tracers much less than phytoplankton subsistence N quota ($Q^{N}_{0,\,phy}$), but it is still important for particulate C:P and particulate N:P ratios, particularly at high latitudes and globally (Figure 3). Diazotroph subsistence N ~~quota (~~and P quotas ($Q^{N}_{0,\,dia}$ ~~) and diazotroph subsistence P quota (~~and $Q^{P}_{0,\,dia}$) in general have much less influence on particulate stoichiometry than $Q^{N}_{0,\,phy}$ and $Q^{P}_{0,\,phy}$ because diazotrophs are much less abundant than ordinary phytoplankton. ~~Nevertheless, they~~ However, diazotroph biomass (carbon) itself is more sensitive to $Q^{N}_{0,\,dia}$ than $Q^{N}_{0,\,phy}$, which shows that the diazotroph subsistence quotas are still important for ~~N_2 fixation and the elemental stoichiometry of diazotrophs~~both their elemental stoichiometry and ability to compete with ordinary phytoplankton. While elemental stoichiometry has been suggested to be an important factor for determining the outcome of the competition between diazotrophs and non-diazotrophs, and consequently N_2 fixation (Deutsch and Weber, 2012; Weber and Deutsch, 2012) , we find that N_2 fixation is no more sensitive to $Q^{N}_{0,\,dia}$ than to the remineralisation rate (ν_{det}), $Q^{N}_{0,\,phy}$, or zooplankton grazing

parameters ($g_{\max}$, ϕ_{phy}, and ϕ_{dia}). Nevertheless, our analysis agrees with the argument that global N_2 fixation is mainly determined by rates of fixed-N loss (Weber and Deutsch, 2014), which in our model is largely affected by ν_{det} and $Q^{\text{N}}_{0,\text{phy}}$.

In general, tracer sensitivities to parameters in OPEM-H configuration are similar to those in OPEM. O_2 and NO_3^- levels are slightly less sensitive to the remineralisation rate, $Q^{\text{N}}_{0,\text{phy}}$, and $g_{\max}$ in OPEM-H because this configuration allows (facultative) diazotroph to grow in high-latitude cold waters, hence the overall biomass of diazotrophs is greater ~~(?)~~(Part I, Pahlow et al., 2020) . This is also the reason why $Q^{\text{N}}_{0,\text{dia}}$ and $Q^{\text{P}}_{0,\text{dia}}$ exert a stronger effect on surface-particle elemental stoichiometry at high latitudes in OPEM-H (Figure 3).

Several studies have revealed that N_2 fixation occurs at high latitude regions (Sipler et al., 2017; Harding et al., 2018; Shiozaki et al., 2018; Mulholland et al., 2019), which supports a wider temperature range of N_2 fixation, similar to what we have in OPEM-H. In the trade-off simulation for OPEM-H we do find some N_2 fixation in the eastern North Pacific and the Arctic Ocean ~~(?)~~(Part I, Pahlow et al., 2020). The different temperature function for diazotrophy is also the reason for the differences in the sensitivities of particulate C:N:P to diazotroph subsistence quotas in high-latitude regions (Figure 3).

4.2 Model limitations

The strong correlation between O_2 and NO_3^- (~~Fig.~~Figure 5) indicates that O_2 and denitrification are tightly coupled. Lack of benthic denitrification leaves water column denitrification as the only loss of NO_3^- and O_2 becomes the primary factor controlling the NO_3^- inventory. This ~~also~~ implies that sensitivities of NO_3^- to the model-parameters could be different when benthic denitrification is incorporated in our model. Also, this means that global N_2 fixation (same as global denitrification in our spun-up steady-state simulations) is underestimated, and since it occurs mostly at $40°$S to $40°$N (see Fig. 13 in Part I, Pahlow et al., 2020) , particulate carbon to nitrogen (C:N) ratios could be overestimated due to a missing input of nitrogen to the surface ocean. This could explain the overestimated surface particulate C:N at low latitudes (see Table 3 and Figure 16 in Part I, Pahlow et al., 2020) .

To evaluate how water-column denitrification affects our cost function, we arrange our simulations in the order of their cost values and plot the volume of oxygen deficient zones (ODZs) against cost for both the OPEM and OPEM-H configurations in Figure 10A to C. Several of our simulations~~have~~ , mostly among those with the 200 lowest cost values (Figure 10A), have a relatively small misfit in O_2 and NO_3^- compared to the WOA 2013, and ~~have~~ high N_2 fixation rates, comparable to those estimated in previous model ~~simulations (e.g., Somes et al., 2017; Wang et al., 2019). On the one hand~~ studies (e.g., Somes et al., 2017; Wang et al., 2019). For these simulations, low O_2 is connected with high rates of water-column denitrification in the eastern equatorial Pacific Ocean (Pac.EQU.E), causing a depression of NO_3^- concentration and a rather high variance in NO_3^- concentration, both of which conflict with the observations. Hence cost in this biome is very high, especially in the upper $550\,\text{m}$ (Figure 9), where denitrification is strongest. On the other hand, although the volume of ~~oxygen deficient zones (ODZs)~~ ODZs in the minium-cost simulations in OPEM and OPEM-H is greater than in the WOA 2013 (Figure 10C), they yield rather low N_2 fixation rates (38.8 and $35.1\,\text{Tg}\,\text{N}\,\text{year}^{-1}$ for OPEM and OPEM-H, respectively). ODZ volumes in the trade-off simulations are more than twice that in the WOA 2013 (Figure 10) and yield global N_2 fixation rates close to current estimates of water-column denitrification (about $70\,\text{Tg}\,\text{N}\,\text{year}^{-1}$, Somes et al., 2017; Wang et al., 2019). The mismatch

between ODZ volume and N_2 fixation rate indicates that a refined description of water-column denitrification setting may be needed (Sauerland et al., 2019). ~~Clearly,~~ While the physical component (ocean circulation) of the UVic model is also very important for the global distribution of oxygen and nitrate, our results suggest that, clearly, only by considering all major nitrogen sources and sinks, such as atmospheric deposition and benthic denitrification, a better representation of ~~N2~~ N_2 fixation and the global marine nitrogen cycle can be achieved.

4.3 Likelihood-based metric

4.3.1 Applicability of the cost function and usefulness of introducing variance information

The cost function introduced above is a metric that quantifies the discrepancy between objectively analyzed observational data and simulation results. Our cost function proves useful for exploring the 400 ensemble model solutions and identifies model solutions that reproduce deep ocean gradients in the NO_3^-:PO_4^{3-} ratio better than a classic fixed-stoichiometry model ~~(?)~~ (Part I, Pahlow et al., 2020). In addition, the optimal model solutions yield improved NCP rate estimates integrated over the top 100m (Part I, Pahlow et al., 2020). In particular, the trade-off solutions of OPEM and OPEM-H can resolve observed latitudinal patterns in dissolved and particulate C:N:P within the upper productive ocean layers ~~(0–, ?)~~(0–130 m, see Part I, Pahlow et al., 2020). The consideration of monthly mean O_2, NO_3^-, PO_4^{3-} data for the upper $550\,\mathrm{m}$ and surface Chl remote sensing data introduces important constraints on the representation of the relation between light and nutrient limitation, thereby also specifying the degrees of N and P limitation.

Even within the 5% of the simulations with the lowest costs, the estimates of global N_2 fixation rate vary considerably. The mean global estimates ~~±1~~ ± standard deviation in OPEM and OPEM-H are $(37 \pm 26)\,\mathrm{Tg\,N\,yr^{-1}}$ and $(51 \pm 29)\,\mathrm{Tg\,N\,yr^{-1}}$, respectively. We initially expected that the NO_3^- and PO_4^{3-} data in the cost function would effectually constrain N_2 fixation. This is clearly not the case and additional information has to be considered. One explanation may be that considerable N_2 fixation can occur during short periods and may also be confined to regions smaller than the biomes. Regional differences with respect to N_2 fixation remain unresolved if only biome-specific monthly mean NO_3^- and PO_4^{3-} data are considered for the upper layers in the cost function.

Also, the minimum-cost solution yields very low global N_2 fixation rates. Thus, for the identification of the trade-off solutions we had to consider prior information about global water column denitrification, whose rate is balanced by N_2 fixation according to our models. Incorporating N_2 fixation as a single global rate estimate into our Likelihood-based cost function as a single additional term would, without some difficult-to-define regulatization, become overwhelmed by the many tracer and variance terms defined in Eqs. (6) and (7). Rather, the additional information is treated as a second objective, namely that global N_2 fixation should be greater than $60\,\mathrm{Tg\,N\,yr^{-1}}$ (see above), which is similar to applying a multi-objective approach for model calibration (e.g., Sauerland et al., 2019), where a trade-off between two or more objectives (cost functions) is resolved. A refined cost function may incorporate monthly mean N:P ratios or N* values based on WOA 2013 data (e.g., for the upper $130\,\mathrm{m}$) for clustered sub-regions of some biomes. Such addition to the cost function would require some careful preprocessing, e.g., cluster analysis of the spatial N:P or N* patterns, but may suffice to constrain simulated N_2 fixation rates.

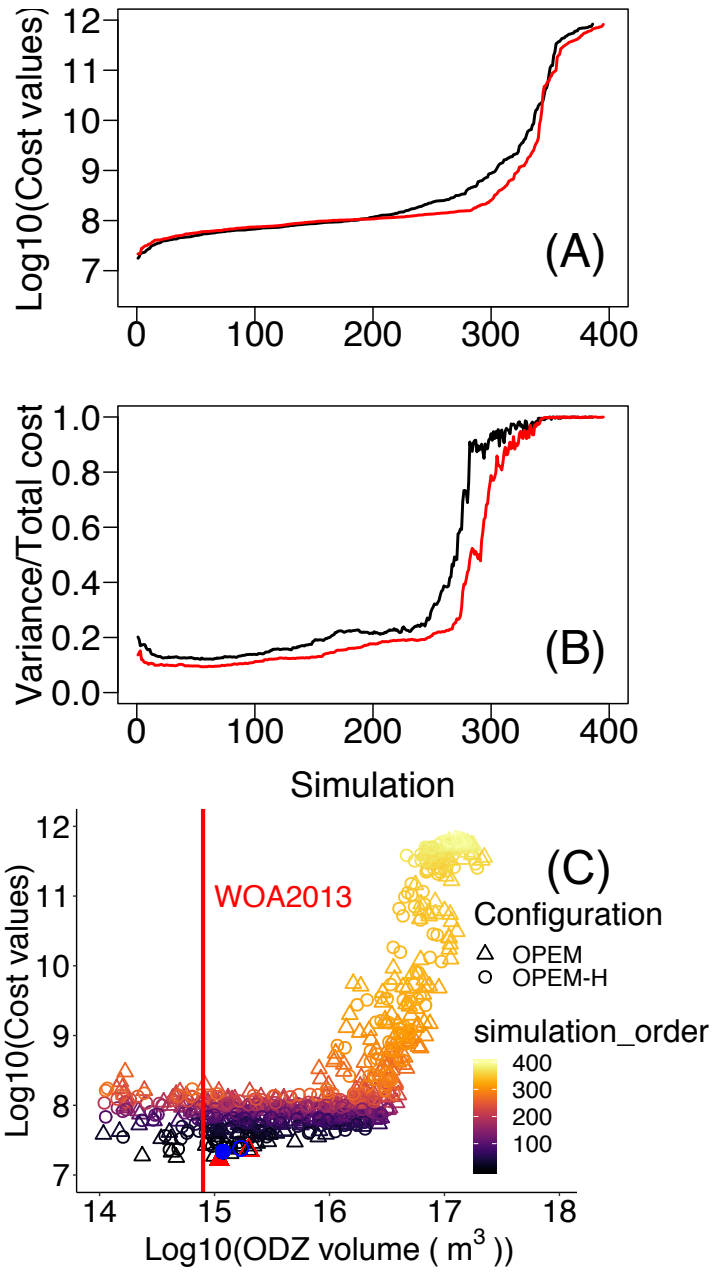

Figure 10. Cost values across all parameter sensitivity simulations ordered from low to high for the two model configurations. Cost values in both misfit and variance (A) and the contributions of variance (B). Black and red lines are for OPEM and OPEM-H, respectively. Total cost versus volume of ODZ (oxygen deficient zone $< 5\,\mathrm{mmol\,O\,m^{-3}}$) in the simulations (C), color represents the simulation order as shown in (A) and (B). The red vertical line indicates ODZ volume in the WOA 2013 ($7.45 \times 10^{14}\mathrm{m}^3$), ~~Solid~~ the solid red triangle and blue circle ~~annotate~~ represent the simulations with ~~minmum~~ minimum cost in OPEM and OPEM-H, respectively, and open red triangle and blue circle are the trade-off simulations.

A peculiarity of our cost function is that it complements the data-model misfit, i.e. the residuals of spatial mean ~~log-transformed~~ $\log_{10}$-transformed values, with an additional term that resolves differences in spatial variances. How the neglect of this term affects the global mean tracer concentrations and flux estimates is depicted in Figures (~~S1~~ S2 – ~~S6~~S7) in the supplemental material. The cost function's variance term introduces a strong penalty to approximately $30\,\%$ of all ensemble model solutions~~(Figure 10)~~. The highest cost-function values ($J > 10^9$) are associated with discrepancies in spatial variances that exceed the misfits in the ~~log-transformed~~ $\log_{10}$-transformed tracer concentrations. For large parts of the ensemble solutions the variance term contributes between 15 and $20\,\%$ to the total costs. Interestingly, for those model solutions that yield low cost function values ($J < 4 \times 10^7$) the relative contribution rises again when the misfit in the ~~log-transformed~~ $\log_{10}$-transformed tracer concentrations gradually decreases (Figure 10B).

4.3.2 ~~Contribution~~ Contributions of biomes

The 17 biomes derived by Fay and McKinley (2014) represent a scale similar to that addressed in global efforts to establish surface-ocean air-sea carbon-flux estimates (Wanninkhof et al., 2013; Rödenbeck et al., 2015). Accordingly, our cost function can be easily extended by incorporating air-sea CO_2 flux estimates in the future. Further improvements may be possible by introducing sub-regions in some biomes, e.g., for constraining N_2 fixation rate estimates, as discussed above.

For low cost function values the contribution of the variance term is generally small in most biomes for the deep layers (Figure 9), where variances of the ~~log-transformed~~ $\log_{10}$-transformed tracer concentrations compare very well between the simulations and the WOA 2013. For high costs this term can become dominant, e.g., for some biomes in the North Pacific as well as the Indian Ocean. A remarkable exception is the North Pacific Arctic biome (NP-ICE), where the deep layer's variance term remains dominant for most of the ensemble solutions. This is somewhat different in the Arctic biome of the North Atlantic (NA-ICE) and the Southern Ocean (SO-ICE), where the variance term remains low throughout almost the entire ensemble. For SO-ICE the cost function is mainly affected by the misfit in ~~log-transformed~~ $\log_{10}$-transformed tracer concentrations. The misfit is associated mainly with discrepancies between observed and simulated NO_3^- within the SO-ICE biome. Interestingly, these misfits in both upper and deeper layers drop again after around the 280$^{\text{th}}$ simulation. Simulations with high NO_3^- do not result in total cost values as high as in simulations with very low NO_3^- (Figure 5), but they have larger misfits for NO_3^- in SO-ICE. A similar behaviour can be seen in the other Southern Ocean biome (SO-SPSS) as well as in NA-ICE.

The upper layer's variance term contributes strongly for low costs in North Atlantic biomes. This is particularly striking for the Equatorial Atlantic biome (Atl-EQU). The main reason is water column denitrification that results in a high variance in NO_3^-. Likewise the Eastern Equatorial Pacific biome (Pac-EQU-E) reveals major model limitations in the upper layers. Overall, the unfolding of biome-specific contributions to the cost function clearly points to those regions where improving model performance appears most worthwhile. Our present cost function may then be reapplied to quantify and highlight specific model improvements.

5 Conclusions

We demonstrate sensitivities of various tracers and processes to parameters in two configurations of a new optimality-based plankton-ecosystem model (OPEM) in the UVic-ESCM. While OPEM-H predicts a wider geographical range for N_2 fixation ~~(?)~~ (Part I, Pahlow et al., 2020) and shows some differences in the sensitivities of diazotroph C, N and P to parameters when compared to OPEM, the tracer sensitivity to model parameters is very similar in both configurations. The trade-off simulations in the OPEM and OPEM-H happen to have the same parameter set. Among our model simulations, varying model parameters within reasonable ranges results in variations in O_2 by a factor of two and in NO_3^- concentration by a factor of six. The sensitivity analysis provides important information regarding the new models' behaviour. The O_2 inventory is mainly influenced by the remineralisation rate (ν_{det}) as well as phytoplankton subsistence nitrogen quota ($Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$) and zooplankton maximum specific ingestion rate (g_{max}). Changes in $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$ strongly impact the NO_3^- inventory, as well as the elemental stoichiometry of ordinary phytoplankton, diazotrophs and detritus. $Q^{\mathrm{N}}_{0,\,\mathrm{phy}}$ also affects N_2 fixation, Chl, DIC and iron levels. Furthermore, our sensitivity analysis resolves correlations between various biogeochemical tracers. For example, POC export is negatively correlated with the NO_3^- inventory. We would like to point out that these changes in model behaviour are solely caused by variations in parameters. Thus, the correlations between tracers and rates might not stand when tracer variations are caused by other factors. For example, an increase in the NO_3^- inventory due to anthropogenic emissions may be accompanied by an increase in POC export (Fernández-Castro et al., 2016). Also, although we did evaluate sensitivities of particulate elemental stoichiometry at different latitudes, most tracer sensitivities and correlations should be considered valid only for global but not regional scales.

We introduce a new likelihood-based metric for model calibration. The metric appears capable of constraining globally averaged O_2, NO_3^- and DIC concentrations as well as NCP. In particular, the minimum-cost and trade-off model solutions resolve observed latitudinal patterns in particulate C:N:P within the surface layers ($0 - 130\,\mathrm{m}$). However, the metric does not effectually constrain the models' global N_2 fixation rate estimates. ~~Incorporating additional terms such as monthly mean N^* in the surface layer into the cost function might provide an additional constraint on simulated N_2 fixation rates.~~ Individual contributions of the biomes to the cost function provide details of how tracer distributions in each biome respond differently under different ecosystem settings. The consideration of spatio-temporal variations in the stoichiometry of NO_3^-, ~~PO_4^-~~ PO_4^{3-}, and O_2 in our metric favours model solutions with low N_2 fixation rates that are solely balanced by low rates of water column denitrification. From our findings we conclude that an explicit consideration of benthic denitrification and atmospheric deposition seem critical for improving the representation of the complete global nitrogen cycle in our model.

Code availability. The University of Victoria Earth System Climate Model version 2.9 (Original Model) is available at http://www.climate.uvic.ca/model/. The OPEM v1.0 code is available at http://dx.doi.org/10.3289/SW_1_2020. The instructions needed to reproduce the model results described in this article are in the supplemental material.

Author contributions. Chia-Te Chien and Markus Pahlow performed the ensemble solutions and selected the reference simulations. Markus Schartau set up the likelihood-based metric. All authors contributed to the manuscript text.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Chia-Te Chien, Markus Pahlow and Markus Schartau were supported by the BMBF-funded project PalMod. Markus Pahlow was supported by Deutsche Forschungsgemeinschaft (DFG) by the SFB754 (Sonderforschungsbereich 754 "Climate-Biogeochemistry Interactions in the Tropical Ocean", www.sfb754.de) and as part of the Priority Programme 1704 (DynaTrait).

References

Anderson, L. A. and Sarmiento, J. L.: Redfield ratios of remineralization determined by nutrient data analysis, Global Biogeochemical Cycles, 8, 65–80, https://doi.org/10.1029/93GB03318, 1994.

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., Schwinger, J., Bopp, L., Boucher, O., Cadule, P., Chamberlain, M. A., Christian, J. R., Delire, C., Fisher, R. A., Hajima, T., Ilyina, T., Joetzjer, E., Kawamiya, M., Koven, C., Krasting, J., Law, R. M., Lawrence, D. M., Lenton, A., Lindsay, K., Pongratz, J., Raddatz, T., Séférian, R., Tachiiri, K., Tjiputra, J. F., Wiltshire, A., Wu, T., and Ziehn, T.: Carbon-concentration and carbon-climate feedbacks in CMIP6 models, and their comparison to CMIP5 models, Biogeosciences Discussions, pp. 1–124, https://doi.org/10.5194/bg-2019-473, 2019.

Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., and Gehlen, M.: PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies, Geoscientific Model Development, 8, 2465–2513, https://doi.org/10.5194/gmd-8-2465-2015, 2015.

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Seferian, R., Tjiputra, J., and Vichi, M.: Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models, Biogeosciences, 10, 6225–6245, https://doi.org/10.5194/bg-10-6225-2013, 2013.

Cavan, E. L., Trimmer, M., Shelley, F., and Sanders, R.: Remineralization of particulate organic carbon in an ocean oxygen minimum zone, Nature Communications, 8, 14 847 EP –, https://doi.org/10.1038/ncomms14847, 2017.

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS001916, 2020.

Deutsch, C. and Weber, T.: Nutrient Ratios as a Tracer and Driver of Ocean Biogeochemistry, Annual Review of Marine Science, 4, 113–141, https://doi.org/10.1146/annurev-marine-120709-142821, 2012.

Deutsch, C., Sarmiento, J. L., Sigman, D. M., Gruber, N., and Dunne, J. P.: Spatial coupling of nitrogen inputs and losses in the ocean, Nature, 445, 163–167, http://dx.doi.org/10.1038/nature05392, 2007.

DeVries, T., Deutsch, C., Primeau, F., Chang, B., and Devol, A.: Global rates of water-column denitrification derived from nitrogen gas measurements, Nature Geoscience, 5, 2012.

Droop, M. R.: Vitamin B12 and Marine Ecology. IV. The Kinetics of Uptake, Growth and Inhibition in Monochrysis Lutheri, 48, 689–733, https://doi.org/10.1017/S0025315400019238, 1968.

Dunne, J. P., John, J. G., Shevliakova, E., Stouffer, R. J., Krasting, J. P., Malyshev, S. L., Milly, P. C. D., Sentman, L. T., Adcroft, A. J., Cooke, W., Dunne, K. A., Griffies, S. M., Hallberg, R. W., Harrison, M. J., Levy, H., Wittenberg, A. T., Phillips, P. J., and Zadeh, N.: GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part II: Carbon System Formulation and Baseline Simulation Characteristics, Journal of Climate, 26, 2247–2267, https://doi.org/10.1175/JCLI-D-12-00150.1, 2012.

Dunne, J. P., John, J. G., Shevliakova, E., Stouffer, R. J., Krasting, J. P., Malyshev, S. L., Milly, P. C. D., Sentman, L. T., Adcroft, A. J., Cooke, W., Dunne, K. A., Griffies, S. M., Hallberg, R. W., Harrison, M. J., Levy, H., Wittenberg, A. T., Phillips, P. J., and Zadeh, N.: GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part II: Carbon System Formulation and Baseline Simulation Characteristics, Journal of Climate, 26, 2247–2267, https://doi.org/10.1175/JCLI-D-12-00150.1, 2013.

Dutkiewicz, S., Ward, B. A., Monteiro, F., and Follows, M. J.: Interconnection of nitrogen fixers and iron in the Pacific Ocean: Theory and numerical simulations, Global Biogeochemical Cycles, 26, https://doi.org/10.1029/2011GB004039, 2012.

Eby, M., Weaver, A. J., Alexander, K., Zickfeld, K., Abe-Ouchi, A., Cimatoribus, A. A., Crespin, E., Drijfhout, S. S., Edwards, N. R., Eliseev, A. V., Feulner, G., Fichefet, T., Forest, C. E., Goosse, H., Holden, P. B., Joos, F., Kawamiya, M., Kicklighter, D., Kienert, H., Matsumoto, K., Mokhov, I. I., Monier, E., Olsen, S. M., Pedersen, J. O. P., Perrette, M., Philippon-Berthier, G., Ridgwell, A., Schlosser, A., von Deimling, T. S., Shaffer, G., Smith, R. S., Spahni, R., Sokolov, A. P., Steinacher, M., Tachiiri, K., Tokos, K., Yoshimori, M., Zeng, N., and Zhao, F.: Historical and idealized climate model experiments: an intercomparison of Earth system models of intermediate complexity, Clim. Past, 9, 1111–1140, https://doi.org/10.5194/cp-9-1111-2013, 2013.

Edwards, A. M.: Adding Detritus to a Nutrient–Phytoplankton–Zooplankton Model:A Dynamical-Systems Approach, Journal of Plankton Research, 23, 389–413, https://doi.org/10.1093/plankt/23.4.389, 2001.

Eppley, R. W. and Peterson, B. J.: Particulate organic matter flux and planktonic new production in the deep ocean, Nature, 282, 677–680, https://doi.org/10.1038/282677a0, 1979.

Everett, J. D., Baird, M. E., Buchanan, P., Bulman, C., Davies, C., Downie, R., Griffiths, C., Heneghan, R., Kloser, R. J., Laiolo, L., Lara-Lopez, A., Lozano-Montes, H., Matear, R. J., McEnnulty, F., Robson, B., Rochester, W., Skerratt, J., Smith, J. A., Strzelecki, J., Suthers, I. M., Swadling, K. M., van Ruth, P., and Richardson, A. J.: Modeling What We Sample and Sampling What We Model: Challenges for Zooplankton Model Assessment, Frontiers in Marine Science, 4, 77, https://doi.org/10.3389/fmars.2017.00077, 2017.

Falkowski, P. G.: Evolution of the nitrogen cycle and its influence on the biological sequestration of CO2 in the ocean, Nature, 387, 272–275, https://doi.org/10.1038/387272a0, 1997.

Fasham, M. J. R., Ducklow, H. W., and McKelvie, S. M.: A nitrogen-based model of plankton dynamics in the oceanic mixed layer, Journal of Marine Research, 48, 591–639, https://www.ingentaconnect.com/content/jmr/jmr/1990/00000048/00000003/art00006, 1990.

Fay, A. R. and McKinley, G. A.: Global open-ocean biomes: mean and temporal variability, Earth Syst. Sci. Data, 6, 273–284, https://doi.org/10.5194/essd-6-273-2014, 2014.

Feely, R. A., Sabine, C. L., Schlitzer, R., Bullister, J. L., Mecking, S., and Greeley, D.: Oxygen Utilization and Organic Carbon Remineralization in the Upper Water Column of the Pacific Ocean, Journal of Oceanography, 60, 45–52, https://doi.org/10.1023/B:JOCE.0000038317.01279.aa, 2004.

Fernández-Castro, B., Pahlow, M., Mouriño-Carballido, B., Marañón, E., and Oschlies, A.: Optimality-based *Trichodesmium* Diazotrophy in the North Atlantic Subtropical Gyre, J. Plankton Res., 38, 946–963, https://doi.org/10.1093/plankt/fbw047, 2016.

Flato, G. M.: Earth system models: an overview, Wiley Interdisciplinary Reviews: Climate Change, 2, 783–800, https://doi.org/10.1002/wcc.148, 2011.

Galbraith, E. D. and Martiny, A. C.: A simple nutrient-dependence mechanism for predicting the stoichiometry of marine ecosystems, Proceedings of the National Academy of Sciences, 112, 8199–8204, https://doi.org/10.1073/pnas.1423917112, http://www.pnas.org/content/112/27/8199.abstract, 2015.

Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Mishonov, A. V., Baranova, O. K., Zweng, M. M., Reagan, J. R., and Johnson, D. R.: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation, in: World Ocean Atlas 2013, edited by Levitus, S., vol. 3, NOAA Atlas NESDIS 75, https://doi.org/http://www.nodc.noaa.gov/OC5/indprod.html, 2013a.

Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Mishonov, A. V., Baranova, O. K., Zweng, M. M., Reagan, J. R., and Johnson, D. R.: Dissolved Inorganic Nutrients (phosphate, nitrate, silicate), in: World Ocean Atlas 2013, edited by Levitus, S., vol. 4, NOAA Atlas NESDIS 76, https://doi.org/http://www.nodc.noaa.gov/OC5/indprod.html, 2013b.

Geider, R. and Roche, J. L.: Redfield revisited: variability of C:N:P in marine microalgae and its biochemical basis, European Journal of Phycology, 37, 1–17, https://doi.org/10.1017/S0967026201003456, 2002.

Getzlaff, J. and Dietze, H.: Effects of increased isopycnal diffusivity mimicking the unresolved equatorial intermediate current system in an earth system climate model, Geophys. Res. Lett., 40, 2166–2170, https://doi.org/10.1002/grl.50419, 2013.

Harding, K., Turk-Kubo, K. A., Sipler, R. E., Mills, M. M., Bronk, D. A., and Zehr, J. P.: Symbiotic unicellular cyanobacteria fix nitrogen in the Arctic Ocean, Proceedings of the National Academy of Sciences, 115, 13 371, https://doi.org/10.1073/pnas.1813658115, 2018.

Holling, C. S. and Buckingham, S.: A behavioral model of predator-prey functional responses, Behavioral Science, 21, 183–195, https://doi.org/10.1002/bs.3830210305, 1976.

Houlton, B. Z., Wang, Y.-P., Vitousek, P. M., and Field, C. B.: A unifying framework for dinitrogen fixation in the terrestrial biosphere, Nature, 454, 327–330, https://doi.org/10.1038/nature07028, 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, Journal of Advances in Modeling Earth Systems, 5, 287–315, https://doi.org/10.1029/2012MS000178, 2013.

Jickells, T. D., Buitenhuis, E., Altieri, K., Baker, A. R., Capone, D., Duce, R. A., Dentener, F., Fennel, K., Kanakidou, M., LaRoche, J., Lee, K., Liss, P., Middelburg, J. J., Moore, J. K., Okin, G., Oschlies, A., Sarin, M., Seitzinger, S., Sharples, J., Singh, A., Suntharalingam, P., Uematsu, M., and Zamora, L. M.: A reevaluation of the magnitude and impacts of anthropogenic atmospheric nitrogen inputs on the ocean, Global Biogeochemical Cycles, 31, 289–305, https://doi.org/10.1002/2016GB005586, 2017.

Keller, D. P., Oschlies, A., and Eby, M.: A new marine ecosystem model for the University of Victoria Earth System Climate Model, Geoscientific Model Development, 5, 1195–1220, 2012.

Kriest, I.: Calibration of a simple and a complex model of global marine biogeochemistry, Biogeosciences, 14, 4965–4984, https://doi.org/10.5194/bg-14-4965-2017, 2017.

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

Kvale, K. F., Khatiwala, S., Dietze, H., Kriest, I., and Oschlies, A.: Evaluation of the Transport Matrix Method for simulation of ocean biogeochemical tracers, Geosci. Model Dev., 10, 2425–2445, https://doi.org/10.5194/gmd-10-2425-2017, 2017.

Kwiatkowski, L., Aumont, O., Bopp, L., and Ciais, P.: The Impact of Variable Phytoplankton Stoichiometry on Projections of Primary Production, Food Quality, and Carbon Uptake in the Global Ocean, Global Biogeochemical Cycles, 32, 516–528, https://doi.org/10.1002/2017gb005799, 2018.

Kwiatkowski, L., Aumont, O., and Bopp, L.: Consistent trophic amplification of marine biomass declines under climate change, Global Change Biology, 25, 218–229, https://doi.org/10.1111/gcb.14468, 2019.

Landolfi, A., Koeve, W., Dietze, H., Kähler, P., and Oschlies, A.: A new perspective on environmental controls of marine nitrogen fixation, Geophysical Research Letters, 42, 4482–4489, https://doi.org/10.1002/2015GL063756, 2015.

Landolfi, A., Somes, C. J., Koeve, W., Zamora, L. M., and Oschlies, A.: Oceanic nitrogen cycling and N2O flux perturbations in the Anthropocene, Global Biogeochemical Cycles, 31, 1236–1255, https://doi.org/10.1002/2017GB005633, 2017.

Li, Z. and Cassar, N.: Satellite estimates of net community production based on O2/Ar observations and comparison to other estimates, Global Biogeochemical Cycles, 30, 735–752, https://doi.org/10.1002/2015GB005314, 2016.

Loh, A. N. and Bauer, J. E.: Distribution, partitioning and fluxes of dissolved and particulate organic C, N and P in the eastern North Pacific and Southern Oceans, Deep Sea Research Part I: Oceanographic Research Papers, 47, 2287 – 2316, https://doi.org/https://doi.org/10.1016/S0967-0637(00)00027-3, 2000.

Longhurst, A. R.: Ecological Geography of the Sea, 2nd ed., Academic, Burlington, Vt., 2007.

Luo, Y.-W., Doney, S. C., Anderson, L. A., Benavides, M., Berman-Frank, I., Bode, A., Bonnet, S., Boström, K. H., Böttjer, D., Capone, D. G., Carpenter, E. J., Chen, Y. L., Church, M. J., Dore, J. E., Falcón, L. I., Fernández, A., Foster, R. A., Furuya, K., Gómez, F., Gundersen, K., Hynes, A. M., Karl, D. M., Kitajima, S., Langlois, R. J., LaRoche, J., Letelier, R. M., Marañón, E., McGillicuddy Jr., D. J., Moisander, P. H., Moore, C. M., Mouriño Carballido, B., Mulholland, M. R., Needoba, J. A., Orcutt, K. M., Poulton, A. J., Rahav, E., Raimbault, P., Rees, A. P., Riemann, L., Shiozaki, T., Subramaniam, A., Tyrrell, T., Turk-Kubo, K. A., Varela, M., Villareal, T. A., Webb, E. A., White, A. E., Wu, J., and Zehr, J. P.: Database of diazotrophs in global ocean: abundance, biomass and nitrogen fixation rates, Earth System Science Data, 4, 47–73, https://doi.org/10.5194/essd-4-47-2012, 2012.

Maier-Reimer, E., Mikolajewicz, U., and Winguth, A.: Interactions between ocean circulation and the biological pumps in the global warming, Max-Planck-Institut für Meteorologie, Report 171, 1995.

Martiny, A. C., Pham, C. T. A., Primeau, F. W., Vrugt, J. A., Moore, J. K., Levin, S. A., and Lomas, M. W.: Strong latitudinal patterns in the elemental ratios of marine plankton and organic matter, Nature Geosci, 6, 279–283, http://dx.doi.org/10.1038/ngeo1757, 2013a.

Martiny, A. C., Vrugt, J. A., Primeau, F. W., and Lomas, M. W.: Regional variation in the particulate organic carbon to nitrogen ratio in the surface ocean, Global Biogeochemical Cycles, 27, 723–731, https://doi.org/10.1002/gbc.20061, 2013b.

McElroy, M. B.: Marine biological controls on atmospheric CO2 and climate, Nature, 302, 328–329, https://doi.org/10.1038/302328a0, 1983.

McGillicuddy Jr., D. J.: Do Trichodesmium spp. populations in the North Atlantic export most of the nitrogen they fix?, Global Biogeochemical Cycles, 28, 103–114, https://doi.org/10.1002/2013GB004652, 2014.

McKay, M. D., Beckman, R. J., and Conover, W. J.: A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, Technometrics, 21, 239–245, http://www.jstor.org/stable/1268522, 1979.

Mulholland, M. R., Bernhardt, P. W., Widner, B. N., Selden, C. R., Chappell, P. D., Clayton, S., Mannino, A., and Hyde, K.: High Rates of N2 Fixation in Temperate, Western North Atlantic Coastal Waters Expand the Realm of Marine Diazotrophy, Global Biogeochemical Cycles, 0, https://doi.org/10.1029/2018GB006130, 2019.

Nickelsen, L., Keller, D. P., and Oschlies, A.: A dynamic marine iron cycle module coupled to the University of Victoria Earth System Model: the Kiel Marine Biogeochemical Model 2 for UVic 2.9, Geosci. Model Dev., 8, 1357–1381, https://doi.org/10.5194/gmd-8-1357-2015, 2015.

Ocean Biology Processing Group: Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Ocean Color Data, NASA OB.DAAC, NASA Goddard Space Flight Center, Greenbelt, MD, USA, https://doi.org/10.5067/ORBVIEW-2/SEAWIFS_OC.2014.0, accessed 2018/07/04, 2014.

Olsen, A., Key, R. M., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Pérez, F. F., and Suzuki, T.: The Global Ocean Data Analysis Project version 2 (GLODAPv2) – an internally consistent data product for the world ocean, Earth System Science Data, 8, 297–323, https://doi.org/10.5194/essd-8-297-2016, 2016.

Pahlow, M.: Linking chlorophyll-nutrient dynamics to the Redfield N:C ratio with a model of optimal phytoplankton growth, Mar. Ecol. Prog. Ser., 287, 33–43, https://doi.org/10.3354/meps287033, 2005.

Pahlow, M. and Oschlies, A.: Optimal allocation backs Droop's cell-quota model, Marine Ecology Progress Series, 473, 1–5, https://www.int-res.com/abstracts/meps/v473/p1-5/, 2013.

Pahlow, M. and Prowe, A.: Model of optimal current feeding in zooplankton, Marine Ecology Progress Series, 403, 129–144, http://www.int-res.com/abstracts/meps/v403/p129-144/, 2010.

Pahlow, M., Dietze, H., and Oschlies, A.: Optimality-based model of phytoplankton growth and diazotrophy, Marine Ecology Progress Series, 489, 1–16, https://www.int-res.com/abstracts/meps/v489/p1-16/, 2013.

Pahlow, M., Chien, C.-T., Arteaga, L. A., and Oschlies, A.: Optimality-Based Non-Redfield Plankton-Ecosystem Model (OPEMv1.0) in the UVic-ESCM 2.9. Part I: Implementation and ModelBehaviour, Geoscientific Model Development Discussions, pp. 1–34, https://doi.org/10.5194/gmd-2019-323, 2020.

Prinn, R. G.: Development and application of earth system models, Proceedings of the National Academy of Sciences, 110, 3673, https://doi.org/10.1073/pnas.1107470109, 2013.

Reid, P. C., Fischer, A. C., Lewis-Brown, E., Meredith, M. P., Sparrow, M., Andersson, A. J., Antia, A., Bates, N. R., Bathmann, U., Beaugrand, G., Brix, H., Dye, S., Edwards, M., Furevik, T., Gangstø, R., Hátún, H., Hopcroft, R. R., Kendall, M., Kasten, S., Keeling, R., Le Quéré, C., Mackenzie, F. T., Malin, G., Mauritzen, C., Ólafsson, J., Paull, C., Rignot, E., Shimada, K., Vogt, M., Wallace, C., Wang, Z., and Washington, R.: Chapter 1 Impacts of the Oceans on Climate Change, vol. 56, pp. 1–150, Academic Press, https://doi.org/10.1016/S0065-2881(09)56001-4, 2009.

Rödenbeck, C., Bakker, D. C. E., Gruber, N., Iida, Y., Jacobson, A. R., Jones, S., Landschützer, P., Metzl, N., Nakaoka, S., Olsen, A., Park, G.-H., Peylin, P., Rodgers, K. B., Sasse, T. P., Schuster, U., Shutler, J. D., Valsala, V., Wanninkhof, R., and Zeng, J.: Data-based estimates of the ocean carbon sink variability — first results of the Surface Ocean pCO_2 Mapping intercomparison (SOCOM), Biogeosciences, 12, 7251–7278, https://doi.org/10.5194/bg-12-7251-2015, 2015.

Sarmiento, J. L., Slater, R. D., Fasham, M. J. R., Ducklow, H. W., Toggweiler, J. R., and Evans, G. T.: A seasonal three-dimensional ecosystem model of nitrogen cycling in the North Atlantic Euphotic Zone, Global Biogeochemical Cycles, 7, 417–450, https://doi.org/10.1029/93GB00375, 1993.

Sauerland, V., Kriest, I., Oschlies, A., and Srivastav, A.: Multiobjective Calibration of a Global Biogeochemical Ocean Model Against Nutrients, Oxygen, and Oxygen Minimum Zones, J. Adv. Model. Earth Syst., 11, 1285–1308, https://doi.org/10.1029/2018ms001510, 2019.

Schartau, M., Wallhead, P., Hemmings, J., Löptien, U., Kriest, I., Krishna, S., Ward, B. A., Slawig, T., and Oschlies, A.: Reviews and syntheses: parameter identification in marine planktonic ecosystem modelling, Biogeosciences, 14, 1647–1701, https://doi.org/10.5194/bg-14-1647-2017, 2017.

Schmittner, A., Oschlies, A., Giraud, X., Eby, M., and Simmons, H. L.: A global model of the marine ecosystem for long-term simulations: Sensitivity to ocean mixing, buoyancy forcing, particle sinking, and dissolved organic matter cycling, Global Biogeochemical Cycles, 19, n/a–n/a, https://doi.org/10.1029/2004GB002283, gB3004, 2005.

Shiozaki, T., Fujiwara, A., Ijichi, M., Harada, N., Nishino, S., Nishi, S., Nagata, T., and Hamasaki, K.: Diazotroph community structure and the role of nitrogen fixation in the nitrogen cycle in the Chukchi Sea (western Arctic Ocean), Limnology and Oceanography, 63, 2191–2205, https://doi.org/10.1002/lno.10933, 2018.

Sigman, D. M. and Boyle, E. A.: Glacial/interglacial variations in atmospheric carbon dioxide, Nature, 407, 859–869, https://doi.org/10.1038/35038000, 2000.

Sipler, R. E., Gong, D., Baer, S. E., Sanderson, M. P., Roberts, Q. N., Mulholland, M. R., and Bronk, D. A.: Preliminary estimates of the contribution of Arctic nitrogen fixation to the global nitrogen budget, Limnology and Oceanography Letters, 2, 159–166, https://doi.org/10.1002/lol2.10046, 2017.

Six, K. D. and Maier-Reimer, E.: Effects of plankton dynamics on seasonal carbon fluxes in an ocean general circulation model, Global Biogeochemical Cycles, 10, 559–583, https://doi.org/10.1029/96GB02561, 1996.

Somes, C. J. and Oschlies, A.: On the influence of "non-Redfield"dissolved organic nutrient dynamics on the spatial distribution of N2 fixation and the size of the marine fixed nitrogen inventory, Global Biogeochemical Cycles, 29, 973–993, https://doi.org/10.1002/2014GB005050, 2015.

Somes, C. J., Schmittner, A., Muglia, J., and Oschlies, A.: A Three-Dimensional Model of the Marine Nitrogen Cycle during the Last Glacial Maximum Constrained by Sedimentary Isotopes, Frontiers in Marine Science, 4, 108, https://www.frontiersin.org/article/10.3389/fmars.2017.00108, 2017.

Sterner, R. W. and Elser, J. J.: Ecological Stoichiometry: The Biology of Elements from Molecules to the Biosphere, 2002.

Su, B., Pahlow, M., and Prowe, A. E. F.: The role of microzooplankton trophic interactions in modelling a suite of mesocosm ecosystems, Ecological Modelling, 368, 169 – 179, https://doi.org/https://doi.org/10.1016/j.ecolmodel.2017.11.013, 2018.

Tang, W., Li, Z., and Cassar, N.: Machine Learning Estimates of Global Marine Nitrogen Fixation, Journal of Geophysical Research: Biogeosciences, 124, 717–730, https://doi.org/10.1029/2018JG004828, 2019.

Thomas, H.: Remineralization ratios of carbon, nutrients, and oxygen in the North Atlantic Ocean: A field databased assessment, Global Biogeochemical Cycles, 16, 24–1–24–12, https://doi.org/10.1029/2001GB001452, 2002.

Vallina, S. M., Cermeno, P., Dutkiewicz, S., Loreau, M., and Montoya, J. M.: Phytoplankton functional diversity increases ecosystem productivity and stability, Ecological Modelling, 361, 184–196, https://doi.org/10.1016/j.ecolmodel.2017.06.020, 2017.

Vichi, M., Pinardi, N., and Masina, S.: A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part I: Theory, Journal of Marine Systems, 64, 89 – 109, https://doi.org/https://doi.org/10.1016/j.jmarsys.2006.03.006, 2007.

Wang, W.-L., Moore, J. K., Martiny, A. C., and Primeau, F. W.: Convergent estimates of marine nitrogen fixation, Nature, 566, 205–211, https://doi.org/10.1038/s41586-019-0911-2, 2019.

Wanninkhof, R., Park, G. H., Takahashi, T., Sweeney, C., Feely, R., Nojiri, Y., Gruber, N., Doney, S. C., McKinley, G. A., Lenton, A., Le Quéré, C., Heinze, C., Schwinger, J., Graven, H., and Khatiwala, S.: Global ocean carbon uptake: magnitude, variability and trends, Biogeosciences, 10, 1983–2000, https://doi.org/10.5194/bg-10-1983-2013, 2013.

Ward, B. A., Dutkiewicz, S., Moore, C. M., and Follows, M. J.: Iron, phosphorus, and nitrogen supply ratios define the biogeography of nitrogen fixation, Limnology and Oceanography, 58, 2059–2075, https://doi.org/10.4319/lo.2013.58.6.2059, 2013.

Ward, B. A., Wilson, J. D., Death, R. M., Monteiro, F. M., Yool, A., and Ridgwell, A.: EcoGEnIE 1.0: plankton ecology in the cGEnIE Earth system model, Geoscientific Model Development, 11, 4241–4267, https://doi.org/10.5194/gmd-11-4241-2018, 2018.

Weaver, A. J., Eby, M., Wiebe, E. C., Bitz, C. M., Duffy, P. B., Ewen, T. L., Fanning, A. F., Holland, M. M., MacFadyen, A., Matthews, H. D., Meissner, K. J., Saenko, O., Schmittner, A., Wang, H., and Yoshimori, M.: The UVic earth system climate model: Model description, climatology, and applications to past, present and future climates, Atmosphere-Ocean, 39, 361–428, https://doi.org/10.1080/07055900.2001.9649686, 2001.

Weber, T. and Deutsch, C.: Oceanic nitrogen reservoir regulated by plankton diversity and ocean circulation, Nature, 489, http://dx.doi.org/10.1038/nature11357, 2012.

Weber, T. and Deutsch, C.: Local versus basin-scale limitation of marine nitrogen fixation, Proceedings of the National Academy of Sciences, 111, 8741–8746, https://doi.org/10.1073/pnas.1317193111, 2014.

Westberry, T., Behrenfeld, M. J., Siegel, D. A., and Boss, E.: Carbon-based primary productivity modeling with vertically resolved photoacclimation, Global Biogeochemical Cycles, 22, https://doi.org/10.1029/2007GB003078, 2008.