Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2020-9-AC2, 2020

Interactive comment on "One size fits all? – Calibrating an ocean biogeochemistry model for different circulations" *by* Iris Kriest et al.

Iris Kriest et al.

ikriest@geomar.de

Received and published: 6 April 2020

We thank referee 2 very much for his/her encouraging and helpful detailed comments! Please find below our reply and answer to the referee's suggestions and comments (in blue).

Specific comments: Line 28: the use of two maxima here is a little confusing; can this be clarified at all?; a range of 180% might even be explicable by such mismatch of OMZ definitions, but I doubt that's what's happened here

Indeed, for some OMZ criteria the range reported by Bopp et al. (2013) is even larger than the 180% stated here, and very large (more than 20 times the observed volume) for a rather low OMZ criterion of 5 mmol O2/m3. We will replace this sentence

by: *"Bopp et al. (2013) report a range of variation that is several times the observed volume, depending on the criterion (maximum oxygen concentration) used for OMZ definition."*

Line 38: "by hand" is, in part, related to keeping the number of tuning simulations down; were a model to be inexpensive to run, automatic tuning sampling a large number of parameter sets would likely be preferred

We will add: *"To keep the number of computationally expensive, global simulations low, usually (...)"*

Line 59: "seems to be" -> "has proven to be"

Will be changed.

Lines 60-69: a very nice framing of the problem; thanks

Line 81: it seems remiss not to include even a sentence or two explaining what the TMM is (or what it comprises); it would spare your readers to present something here

We will add a few sentences on this: *"All model simulations and optimisations apply the Transport Matrix Method (Khatiwala, 2007, 2018), as an efficient "offline" method for ocean passive tracer transport. The TMM represents advection and mixing in the form of transport matrices that have been calculated from an ocean circulation model simulation prior to the biogeochemical simulations performed here. For our model simulations we apply monthly mean transport matrices (TMs), as well as monthly wind speed, temperature and salinity for air-sea gas exchange."*

Line 87: move this domain information into a table instead?

We would prefer to refrain from adding more information to table 1 (which gives the spatial resolution), and also rather not add another table to the main paper. However, we now cross-reference between this section and Table 1.

Line 104: interesting numbers are reported for this domain, but not the others; it would be good to have all, or at least some other common information across the TMs

Because the original circulation models differ among each other (in their configuration), and because the offline circulation (in form of TMs) might differ from the basic circulation model we'd prefer not to extend too much on the original circulation model properties, but rather focus on the properties we can derive from the offline circulation. However, we suggest to present the physical diagnostics discussed in subsections 2.2 and 2.3 in an additional table in the supplement. (See also below.)

Lines 109-110: what does "accurately represent" here mean?; is it in reference to the circulation strengths you mention, or something else?

What we wanted to say in in fact, that these modified circulations have not been compared or evaluated agains the previously published circulations. We suggegstto replace this by *"We note that none of the circulation configurations, aside from UHigh, have been previously evaluated against the most commonly used UVic ESCM FCT configuration (e.g., Weaver et al., 2001; Schmittner et al., 2005; Somes et al., 2013)."*

Line 123: could a calculation of ventilation timescale (mean, max, and between basins) help separate out the differences between the TMs?

This is what we tried using the ideal age tracer. See below, we suggest to add the values of ideal age of different water masses, as well as globally to Table S1 in the supplement.

Line 129: "up to more than 800 years"; any idea why?; that does sound surprising at face value

One possible explanation of the strong increase in ideal age could be a lack of bottom water formation and ventilation in this model, which affects, for example CFC-11 (Dutay et al., 2002). In fact, the mixing criterion applied in our paper (>400m) would be too shallow to mix young surface and old deep waters, anyway. We therefore propose to skip *", despite its large area of deep mixing"*, and not give an explanation in this

subsection. Instead, we propose to mention it below, at the beginning of subsection 3.2 (see below, our reply to Line 169).

Line 130: ah-ha - ages mentioned here, but perhaps these could be added to the table I mentioned before for clarity

See above; we would rather not "burden" the table with this information, but have added a table on outcrop area, area of MLD and ideal age in the different water masses and globally to the supplement.

Line 132: "tracer observations"?; do you mean radiocarbon?

Yes, but also CFCs, temperature, salinity, phosphate, and oxygen. We will change this sentence to: *"constrained with transient (radiocarbon, CFCs) and hydrographic (temperature, salinity, nutrients, oxygen) tracer observations (Khatiwala et al., 2012)."*

Lines 137-138: this section feels like it could do with a sentence explaining how this information will be used later; however, it's certainly very helpful to elucidate how models might be good / bad

See our reply below, we have rewritten the first paragraph of this section and added a sentence on this.

Line 169: per my earlier remark, how's Drake Passage transport in the models?; it has a relationship with the SO properties mentioned

The transport through the Drake passage might, for example influence the properties and formation of SAMW in the Atlantic section (e.g., Sallee et al., 2010, 2013). However, as noted above we would prefer not to discuss details of the underlying circulation models too much in this paper. We propose to rewrite the beginning of subsection 2.3 as: *"The underlying circulation models, from which the TMs and forcing were extracted, differ in many aspects, such as parameterisation of mixing, forcing, sea-ice, etc., all of which can affect their dynamic behaviour and the quantities and diagnostics described above. For example, in the Southern Ocean the eastward transport of waters through* *the Drake passage can affect the properties and formation of SAMW (e.g., Sallee et al., 2010, 2013), while parameterisation of sea ice in the models might affect the formation and ventilation of AABW (Dutay et al., 2002), with consequences for water mass age. It is beyond the scope of this paper to compare and discuss the details of the circulation models. Instead, to examine the potential impact of their charactistic features on optimal biogeochemical model parameters (section 3.2), we will focus on three diagnostics that can be easily derived from most circulation models (see Table S1 for simulated and observed values)."*

Line 204: how is near steady state defined? (presumably in Kriest et al., 2017 ...)

We have examined the transient behaviour of MOPS coupled to ECCO TMs over a simulation time of 9000 years in Kriest and Oschlies (2015). In that paper we focused on global average oxygen and nitrate (or fixed nitrogen), as these two properties are subject to processes affected by a variety of time scales (large scale circulation; deep remineralisation; air-sea gas exchange or fixed nitrogen gain and loss in different ocean basins). In the analysis we found that after significant variations within the first 3000 years (even with inflection points) the inventories begin to stabilise. We propose to add a note on this, with reference to that paper (in particular: Figure 2): *"Following the simulation of these 10 model setups over 3000 years to near-steady state, after which, for example, global oxygen and nitrate inventory change only by a small amount (see Figure 2 by Kriest and Oschlies, 2015), (...)"*

Line 210: "many local optima"; good - this is a perennial problem with BGC models

Yes, unfortunately.

Line 214: as these properties are tightly constrained in the real ocean (e.g. via the N:P ratio), is there an advantage to using all of them?; i.e. could N, O2 or P, O2 be sufficient?; something like carbon - which has a more plastic relationship with nutrients - might arguably be good too

In the model all three tracers - phosphate, nitrate, oxygen - are subject to different biogeochemical and physical processes, and this is the reason why we are using all of them in the combined misfit function: phosphate, as a conserved property, is not affected by a global bias (and therefore has the least impact on the misfit function; see Kriest et al., 2017). The oxygen inventory can adapt to the combined physical (air-sea gas exchange; subduction and ventilation) and biogeochemical (particle sinking and remineralisation) processes. So can nitrate, as its distribution in the ocean is affected by production, export, sinking and remineralisation. In addition, it is also affected by denitrification and nitrogen fixation, which happen in different areas, connected through large scale circulation as well as (local or regional) processes (e.g., upwelling). Therefore, we think that especially the latter two tracers (nitrate and oxygen) are necessary for the optimisation a model that simulates all three tracers independently. We agree, that at a later point it will be important to include other types of observations, such as DIC and alkalinity; but this would require a model that also parameterises the (effects of) calcification, calcite sinking and dissolution. Finally, given the so far rather unconstrained biological model components (phyto- and zooplankton, detritus, DOP), we think that it might be even more important to include corresponding observations (even if they are sparse in space and time) into the optimisation.

Line 239: you could add the total number of simulations here to indicate the total computational load

We will add a line on this to table 1.

Line 245: earlier, in relation to burial and riverine input, you suggest that the inventory is not actually fixed, and can drift by a few percent; which is right?

Actually, we made a mistake here, and should have skipped "inventory" when referring to all nutrients at the end of section 2.4. Thank you for pointing this out. The form of resupply of buried P (and N) via either river runoff or surface supply alters the spatial distribution of nutrients and oxygen, as well as the inventory of nitrate and oxygen, but the phosphorus inventory is fixed: global average phosphate remains almost the same (within 2e-4) between different model setups mentioned. (The very small change in phosphate inventory is likely due to faster equilibration when buried phosphate is distributed everywhere at the surface, and/or a slightly different contribution of organic P to total P, because of "fertilisation" of the surface.) Effects on nitrate and oxygen are larger, because these inventories can adapt dynamically to a different nutrient supply and turnover. This is currently subject to investigation. We will change the text accordingly: *"We note that in uncalibrated models (e.g., Kriest and Oschlies, 2015) this creates differences of a few percent in the global inventory of nitrate and oxygen. Differences in the regional distribution of nutrients and oxygen are largely comparable in magnitude to those caused by the numerical sinking scheme of detritus (Kriest and Oschlies, 2011). The effect of this process is subject to further research."*

Line 284: why especially oxygen?

As shown by Kriest et al. (2017), the misfit function is dominated by oxygen, because this tracer can include a global bias. The Pacific, with its very large volume, and old waters (which memorise the errors in biogeochemistry), is especially influential. We will add a reference to this paper and a comment: *"(...) and the comparatively large impact of oxygen on the misfit function (see Kriest et al., 2017)."*

Line 290: so that readers (like this one) don't have to scramble back a few pages, perhaps reiterate the default values once when you mention changes here

Kriest and Oschlies (2015) tested several values for the nitrate and oxygen affinity, so we find it difficult to refer to a default value here, but we suggest to mention the default values for max. rate of nitrogen fixation and the nitrate threshold of that paper, as well as that for b: *" (...) increasing b from a "default" value of 1.1 (Kriest and Oschlies, 2015) to 1.39. It further results in a high nitrate threshold for the onset of denitrification ($DIN_{\text{min}} = 15.8$ mmol m^{-3}, compared to the default value of 4 mmol m^{-3} applied by Kriest and Oschlies, 2015), a low affinity of denitrification to nitrate ($K_{\text{DIN}} =$*

32 mmol m^{-3}), and a low maximum nitrogen fixation rate ($\mu_{\text{NFix}} = 1.19\,\mu$mol m^{-3} d^{-1}; Table 1), which is only about half of the default value applied by Kriest and Oschlies (2015)."

Line 362: it's not *completely* independent when the model is optimised to oxygen concentration; although, I appreciate it's not a target

We agree; we will add *"largely"* to this.

Line 428: should be "m3" not "m-3"; also, might want to contextualise with a percent of mean

Of course, thank you. We will add *"i.e., more than 100 the observed volume."* to put this into context.

Lines 639-640: "on the other hand . . ." is a confusing point; what do you mean?; it seems to suggest that a "benefit of parameter optimisation" is "helping to search for the best parameter set"; that sounds not particularly profound; something instead about "necessary level of model complexity"?

We agree, and suggest to rewrite this as: *"optimisation allows for a "fair" comparison of models of different complexity (after each model has been tuned to match some desired quantity best); it can therefore also help to decide about the necessary level of model complexity."* We will skip the part about model development.

Line 647: "through low ideal age"?

Changed.

Table 1: mean ocean ventilation age from these different circulations might be an interesting metric; or some other relevant integral metric of circulation

See above - we will add a table that provides the ideal ages in different water masses of all 5 circulations. Global average ideal age varies between 583 (UHigh) to 652 (MIT28), reflecting mainly the CDW, and will also given in supplementary table S1. However, as we did not find a relation between any optimal parameter and the CDW age (Table 2), we prefer to not discuss this any further.

Table 3: export production - I like this illustrating of gaps in previous work (which has been more opportunistic than this study)

Figures: a general point I'd make is that red/blue colour bars are usually for situations where a property (e.g. a delta) has a definite central point worth marking (e.g. zero); here they're used broadly, potentially skewing the reader's perspective

We agree that red/blue might suggest a delta scale - on the other hand to our knowledge such a colour scale works well for visually impaired. We have now chosen a colour scale that does not bear such a strong resemblance to a delta scale, but should work for visually impaired as well (applies to Figures 1,2,3,5,S1,S2,S5) (see ferret.pmel.noaa.gov/static/FAQ/graphics/color_friendly_palettes.html)

Figure 1: you may have tried already, but might delta plots be better? (i.e. show the observations as field values but models as differences from this)

We tried, but to our opinion delta plot did not help interpretation.

Figure 2: as the observations are missing ice, I'd be inclined to skip it in the models as well

We will skip this.

Figure 3: see general and Figure 1 comments

See above, we changed to different colour scale.

Figure 4: panel D is rather complicated and ugly

We have skipped this panel, and added it separately as an additional figure in the supplement (see also comment by referee 1).

Figure 5: change to a much uglier palette; same red / blue issue

See above, we changed to different colour scale.

Figure 6: I really like this relating of parameter value to circulation property

Thank you. We have now added a symbol legend, based on a comment by reviewer 1.

Figures 8 and 9: I found these rather difficult to interpret, although I have no suggestions on how to change them

We are aware of the fact that these are difficult to interpret, and tried many different approaches, but also have not found any better was to illustrate this.

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2020-9, 2020.