Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC2
https://doi.org/10.5194/gmd-2021-327-RC2, 2021

Comment on gmd-2021-327

Anonymous Referee #2

Referee comment on "Ocean biogeochemistry in the Canadian Earth System Model version 5.0.3: CanESM5 and CanESM5-CanOE" by James R. Christian et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-327-RC2, 2021

Review of Christian et al.: "Ocean biogeochemistry in the Canadian Earth System Model version 5.0.3: CanESM5 and CanESM5-CanOE"

Summary:

The manuscript presents a description of the marine biogeochemistry component of the CanESM5-CanOE model, and an evaluation of this, together with its less complex sibling, CanESM5. The evaluation is focused on marine biogeochemistry, specifically the models' oxygen cycles, carbonate chemistry, nutrient cycles, plankton abundances and export production. Evaluation compares both to observational data (where available), and to a selection of peer CMIP6 models (including the multi-model mean of this CMIP6 ensemble). The analysis identifies biases in nutrient distributions, historical trends in export production, and differences in net ocean CO2 uptake.

Assessment:

First of all, I would just want to make clear that I think that this sort of publication is valuable for documenting in detail the performance of major components of ESMs, especially where, as here, the authors make an effort to cover the model end-to-end, as well as compare with peer models. Given the broad span of tracers, processes and geographical / historical patterns, there is no natural end to where such analysis should begin or stop. That said, I have a number of major criticisms of the current draft of this manuscript:

- While the focus is on marine biogeochemistry, it is remiss not to include information about the performance of the physical model that underpins this. The reader has no information on how well this latter model performs in terms of surface properties, mixing and ventilation and interior circulation. This need not be exhaustive, but some material on this (even if only summarising from another evaluation manuscript) seems important. Especially where this has a potential bearing on biogeochemistry performance.
- The evaluation itself seems almost arbitrary in its choice of targets and, in particular, the order in which these are introduced and discussed. For me, oxygen and carbonate chemistry parameters are essentially "downstream" of the main drivers of biogeochemistry in the ocean. Nutrient cycles, productivity and carbon / alkalinity seem much more important to first-order patterns. And this leads to oddities in the manuscript, for instance where patterns of oxygen biases, likely due to "upstream" biases in production, are discussed prior to anything about these likely sources. Also, the focus on export production instead of primary production is rather odd.
- Latterly, the manuscript looks at historical trends in anthro CO2 uptake, export production and ocean anoxia. While these are valuable to look at, the manuscript offers little by way of explanation for them. These aren't easy issues to tackle, and are probably beyond the scope of this manuscript, but it seems remiss not to include analysis. For instance, it might be informative to show whether there are spatial components to these trends, or links to physical phenomena. In the specific case of export production, the authors allude to long-known issues with increasing stratification and decreasing primary production, but show no evidence of either (further, I think the physics might be the same in both models).
- Finally, on the model naming side, while the manuscript eventually settles down, it is confusing at first about the identities and compositions of the CanESM models under consideration. The introduction and, especially, the abstract are messy on this front, and will likely confuse readers. A clear and simple statement

In addition, I have a number of specific comments on details of the manuscript and include these below. I should add that some of these reflect my own style / presentation

preferences, and the authors should not feel obliged to address these if they disagree.

Overall, while I appreciated much of this manuscript, I judge that it requires major revisions before it can be accepted.

Specific comments:

Ln. 19-22: This opening is confusing; which models are being examined here?; if there are more than one, and it sounds like there are, say this in the opening sentence

Ln. 24: a brief mention of the aspects of the model being intercompared would be useful (e.g. nutrients, carbon, productivity, etc.)

Ln. 33: re: export decline - why?; is this explored in the manuscript?

Ln. 33-35: re: plankton - is this worth mentioning in the abstract?; it's without context and isn't clear whether it's something that can be compared to observations

Ln. 35-36: re: phytoplankton - is this worth mentioning in the abstract?; again, there's just no context here - e.g. is this different between the different model versions considered?

Ln. 39: if you're going to put the specific numbers in the abstract, you should also include some reference to the observational estimate (or range)

Ln. 39: re: anthro CO_2 uptake differences - why?; is this explored in the manuscript?

Ln. 41: some of this material seems more appropriate for the methods section than the introduction

Ln. 41: more generally, the manuscript makes some specific choices on the model components and processes analysed, but does not articulate the science cases for these; the introduction is the place for doing this

Ln. 48: just out of interest, why not v3.6_stable?; that's a common CMIP6 configuration; (although having a different version isn't necessarily an issue - it's more variety in the CMIP6 ensemble)

Ln. 50: this is PISCES-v2?

Ln. 66: how does tracer number relate to compute cost?

Ln. 66: the relationship between CMOC and the models here is unclear; how does this relate to CanESM1, 2 and 5?; this should be unambiguous

Ln. 69: re: CFCs - if the physics is the same between the models, you should only need to run the CFCs and SF6 in a single simulation anyway; they are non-interactive with the BGC

Ln. 71: re: “prohibitively expensive” - this isn't clear as many of these tracers are single tracers; or are you thinking that you might want to duplicate all of your BGC model tracers to have parallel reservoirs of natural C and C14?

Ln. 74: does CMOC not have a version number?; seems odd as the version here is different (via O2) from earlier versions; not necessarily meaningfully (since O2 is a "downstream" tracer), but this does suggest at least a code difference

Ln. 74: there would be much less risk of confusion here if the different versions of CMOC here were identified with version numbers

Ln. 80-81: this specification of Fe limitation as being calculated from surface nitrate could do with a bit of justification; presumably this has been given before in a previous outline of the model; please include this here

Ln. 95: re: 100% burial - is this true regardless of the saturation state of the seafloor?; i.e. above / below CCD

Ln. 99: maybe referring to CanESM5-CMOC might be better than CanESM5 on its own

Ln. 105: I would have expected some sort of summary about the physical ocean (and sea-ice) model used here; including resolution and major options selected; especially as the common NEMO configuration v3.6_stable has not been used here

Ln. 117: modified from what?; do you mean from previous versions of this model, or from elements of PISCES?; the latter is implied, but it would be useful to have a "... modified from corresponding PISCES components to varying degrees."

Ln. 118-119: and now we have another name for the model, NEMO-CMOC; some standardisation of naming would be helpful

Ln. 121: does the model use the preferred carbonate chemistry of Orr et al. (2017)?; MOCSY

Ln. 126: this is ambiguous; previously it's implied that all calcite reaching the seafloor is buried, whereas this line implies otherwise

Ln. 131: re: phytoplankton functional types - ... which could be named here

Ln. 130-137: this summary paragraph of the model focuses on the phytoplankton, and doesn't mention other components, e.g. zooplankton; that tends to imply they're the same as before; anyway, this would be a good place to introduce other elements

Ln. 142: are we getting an explanation of what these size categories are meant to represent?; e.g. prokaryotic vs. eukaryotic

Figure 1: as already noted, I'd suggest properly introducing all of the elements of this model in section 2

Figure 1: for ease of comparison, and especially as it is not described here, it might be an idea to include the corresponding schematic of CMOC;

Table 1: might it be worth noting where parameters have a corresponding parameter in CMOC, and whether the values are the same?; some model processes certainly overlap

Table 1: re: parameter k_Ca - does $CaCO_3$ dissolve above the CCD?

Table 1: re: parameter K_NH4ox - might it be better to describe K_NH4ox as a maximum nitrification rate, which is then diminished by an irradiance function with a half-saturation, K_E

Table 1: re: parameter K_NO3 - presumably N_2-fixation occurs in ignorance of PO_4 availability?

Ln. 160: explain what's going on here with NH_4 and NO_3; does this functional form have a source describing it?

Ln. 169: is E irradiance?

Ln. 175: re: C_XS - this term needs expansion (or a reference to a later equation number if it appears below); I'm uncertain what you mean by this, or what ecological process it's meant to represent; equation 16a seems to be the right one

Ln. 180: "excessively low" in a model stability sense?; or is there an actual observed threshold here?

Ln. 204-205: how equations 13a and 13b fit into equation 11b is unclear, especially as equation 11b refers only to G_L, which seems to be calculated in equation 12b; the latter point also occurs for small zooplankton

Ln. 208-216: might the clarity of this section be improved by the addition of a diagram that quantitatively illustrates the scale of excess C, N (and possibly Fe) over a span of intake C:N?

Table 2: this kind-of answers my point before about a diagram, although a diagram might still be better (if more difficult to create)

Ln. 241: has the impact of forcing a common zooplankton C:N on detritus compared to dynamic C:N in phytoplankton been explored at all?; does this mean that the majority source of both detritus classes is zooplankton?; it seems odd to make a fuss about C:N in phytoplankton only to entirely overlook the C:N of the more heterogeneously-sourced detritus component

Ln. 263: ah-ha; E = irradiance

Ln. 265: does the absence of PO4 in the model cause any problems for this N2-fixation scheme?; low NO3 is often associated with low PO4

Ln. 284-285: does this scheme produce large-scale spatial patterns in calcite production that match the general high-equatorial, low-polar pattern?

Ln. 290: again, it's implied earlier that 100% of calcite is buried, but this suggests otherwise

Ln. 291-292: is this localised spatially?; i.e. loss at the seafloor is added at the surface immediately above

Ln. 296: some expansion here on the precise links between processes would be helpful; e.g. NO3 vs. NH4

Ln. 317-318: this is a little paradoxical; the closer a seafloor tile is to sources of O2 (surface productivity and the atmosphere), the less oxygenated the sediments; this presumably reflects the supply of organic matter to the seafloor and the resulting oxygen demand; if this is the logic, make this clear

Ln. 328: this also implies that particles can scavenge iron continuously without saturation; I don't imagine this is a problem, but it might make the model's behaviour in areas dominated by slow or fast sinking detritus interestingly different

Ln. 332-340: sensible; I like this

Ln. 342-352: this could be clearer and sourced to relevant work on the topic; Wolf-Gladrow et al. (2007) (which you cite earlier) suggest +1 ALK for N2-fixation to NH4+, -2 ALK for NH4+ to NO3-, and +1 ALK for denitrification of NO3-; here, assuming N2-fixation goes to NO3-, this implies -1 ALK for N2-fixation and +1 ALK for denitrification; anyway, the text here is ambiguous, and should be straightened out and sourced

Ln. 354: links for the data?; and access dates; some of these products are revised periodically

Ln. 360: this is GLODAPv2

Ln. 360: which offline carbonate chemistry calculations are needed?; GLODAPv2 includes pH

Ln. 363: rephrase to "... were used for the absent tracers, phosphate ..."

Ln. 382: why 2x2 degree?; the 33 levels is more understandable

Ln. 382: how regridded?; linear, nearest neighbour, etc.?

Ln. 383: technically, GLODAP follows WOA (which did this vertical grid first)

Ln. 386: I think ignoring variability across the CanESM5 ensemble is not an unreasonable

assumption, but it might be useful to support it for this particular model with some evidence; e.g. a plot of some key property (e.g. NPP, CO_2 flux, SST, etc.) across the ensemble for, say, the decades of interest here; this could be put in supplementary if it breaks the flow

Ln. 394: it would be helpful to name (and source to descriptions / evaluations) the CMIP6 models used in this analysis within this section

Ln. 396: while it may have been done elsewhere, some sort of outline of the performance of the physics model seems necessary to me; even if it's cursory and largely points to this other work; if there is no other work, some expansion would be useful; things like surface physics (incl. mixing), sea-ice, major circulation (AMOC, Drake), MOC would be of interest

Ln. 398: please be clear why you're starting with oxygen; in most models it's largely slaved to other more dynamic model processes and tracers (which are, in turn, strongly influenced by physics processes); if there's a good reason why you're looking at it first, make it clear

Ln. 400: why these depths?; including a more abyssal depth might hint at circulation issues

Ln. 401: for a number of reasons, I would not expect to MEM to be a good comparison; do you know how it compares to observations relative to the performance of its component models?

Figure 2: with fewer colours in this scale, it would be easier to discern differences between the models

Ln. 422: you're inferring these as "circulation features" but haven't reported on your model's circulation at all (e.g. MOC)

Ln. 428: "the ensemble mean" = "MEM"

Ln. 433: it's old-fashioned of me, but would a profile of O2 further assist here?; possibly not given its spatial heterogeneity, but a series of vertical Taylor diagram slices is a little hard to take in!

Ln. 444-450: how does this relate to any hard-wired limits in models?; the model I use, for instance, is prevented from consuming oxygen below a limit

Ln. 461: re: “much deeper” - why?; and does this relate to the abyssal issue I raised re: oxygen?; i.e. this is an interesting depth

Figure 5: might a table be better for this information?; maybe combined with other measures of model performance?

Figure 6: too many colours here makes it more difficult to discern differences between the panels

Figure 6: might the depth at which omega aragonite hits some threshold (value 1 would be most obvious) be better?

Ln. 471-472: again, remineralisation is mentioned in the context of biases before anything about production and export is introduced; omega is a downstream variable, so the ordering of the analysis here is perplexing

Figure 7: how reliable are the observations here?; GLODAP is much less data-rich than WOA

Figure 7: geographical plots of surface omega, seafloor omega, and the depth at which omega hits some threshold would seem more valuable to me; and easier to compare between models - these are very similar looking plots whose differences are not easy to discriminate

Figure 7: do you need both aragonite and calcite?; the model seems to use calcite only

Figure 8: actually, I take my earlier comment back, the MEM is pretty much always better than the individual models

Ln. 496: as the N and Fe cycles regulate productivity and therefore ocean interior remineralisation and DIC/ALK, it would perhaps make more sense to discuss these ahead of the more downstream oxygen and carbonate chemistry properties

Ln. 502: HadGEM2-ES's marine BGC included a prognostic Fe cycle; see the full description of Totterdell (GMD, 2019)

Figure 11a-11b: it seems overkill to have both 11a and 11b in the manuscript; I'd suggest deleting 11b

Figure 12: is this scale running across three orders of magnitude?; so is it 1 nmol/m3 to 1000 nmol/m3?; if so, the labelling of this log scale differs from that of the nitrate plots above

Figure 14: why not geographical plots of DIN?

Figure 18: worth plotting some regressions on here?; the data density means that the shape of the curves might be easier to discern then; also, why does the plot's chlorophyll appear "capped" at 1 mg / mg?

Figure 18: re: 17 mg / m3 - this seems a bad idea; why do this?; it looks like you're trying to maximise the appearance of fit

Figure 19a: why crop the scale?; it's not helpful with bar charts

Ln. 666-673: a plot that might be helpful here is the geographical map of cumulative CO2 uptake; for instance, to identify whether the uptake pattern is the same but the magnitude different, or that there are actual differences in the spatial pattern of uptake

Ln. 666-673: another plot which might be useful here is the geographical inventory of anthro CO2 in the models (the CanESM5 ones); again to identify whether there are patterns in the differences between the models

Ln. 681: is this decline in response to stratification happening here?; I thought the models were physically identical?

Ln. 684: geographical plots of export production in the models, and how it changes between, say, 1980 and 2014 would be helpful here

Ln. 684: more generally, it seems strange to include these trends in export production without (a) talking about primary production, and (b) trying to dissect what the source(s) of the trends are

Figure 21: why focus on export ahead of production?; and would it be more interesting to consider export / production over this time period?

Ln. 711-713: broken word here; "depend ... ent"

Ln. 768: are there runtime figures on how more costly it is?; you might expect cost to scale with complexity; e.g. Kwiatkowski et al. (2014) found ~linear relationships with tracer count

Ln. 798: my personal preference is to conclude a paper with a set of bulletpoint conclusions of the main findings

Ln. 807-809: maybe include the source ID for the model on the ESGF system together with the variant labels for the specific ensemble members included in the analysis