Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-150-RC1, 2018

Geoscientific Model Development Discussions

Interactive comment on "Limitations of the 1 % experiment as the benchmark idealized experiment for carbon cycle intercomparison in C^4MIP" *by* Andrew H. MacDougall

C.D. Jones (Referee)

chris.d.jones@metoffice.gov.uk

Received and published: 29 July 2018

Review of "Limitations of the 1% experiment as the benchmark idealised experiment for carbon cycle inter comparison in C4MIP", by Andrew MacDougall.

This is a well written, clear description of a proposed alternative experiment to the now-standard 1% experiments often used to quantify and compare carbon cycle feedbacks in coupled climate carbon cycle models (so-called C4MIP experiments).

I found this a useful and thoughtful paper which makes some very salient comments about existing experimental design and offers some insights into the limitations of the

standard experiments compared to new "logistic" CO2 pathways. The paper show cases the new pathways using the UVIC EMIC.

In general, both personally and as a co-chair of C4MIP, I find this level of analysis and engagement very pleasing to see, and it will certainly help drive the further evolution of C4MIP in the future (I'm not yet ready to think about CMIP7 though!). C4MIP is explicitly aimed at ESMs, although we welcome EMIC participation. But perhaps for a next generation we should more explicitly engage with EMICs and provide additional simulations which EMICs can lead on to supplement joint ESM/EMIC runs. In fact it was a requirement of CMIP6 that no MIPs added new experiments which had not been tried by at least some models. They (very reasonably) wanted to avoid too many brand new experiments being suggested and possibly wasting time of model groups. So it is really positive to see suggestions like this also being tested with a model.

I list below some comments which I hope will be useful both for the improvement of this manuscript and also in general as part of the evolving discussion. There are some areas of literature which can be helpful, and there are some issues which are relevant to ESMs more than EMICs (mainly around computational expense). But overall I very much like this paper and would recommend publication with only minor amendments.

My main question really is not just the choice of scenario - what do you recommend about an analysis technique. You do not mention performing coupled/uncoupled simulations with the logistic pathway - so how would you look at climate-carbon and CO2-carbon feedbacks? Would you still want to do COU, BGC and RAD versions of the logistic pathway? (which would increase computational cost of course). How do these metrics (beta/gamma) evolve in time?

or are you suggesting keeping the 1% run for the feedback separation and using the logistic run to look more at emissions/TCRE/AF?

It would be good to be clear on the intended USE as well as scenario that you are suggesting.

Otherwise, I list some comments below which I hope you find useful. It would be great to involve you in future discussions around C4MIP analysis and experimental design.

Chris Jones

1. In several places, including the abstract and conclusions the paper mixes up features of the models/results with features of the experiment itself. For example you say sink-to-source transition is "absent from the 1% experiment". I think you should be a bit stricter in which phrasing you use - the sink-to-source transition is neither present nor absent in the 1% experiment - but it will depend on the results. It may or may not occur depending on the model. You might be able to say it is more likely in one set of model runs than another, but it is not a "feature of the experiment".

2. The paper gives a nice overview of the history of the 1% simulation. There has, though, been more discussion around the choice of this for C4MIP than acknowledged here (it's not true to say, "a clear rationale for... 1% experiment... is absent". The best paper on this is Gregory et al (2009, J.Climate). They look in some detail at the Friedlingstein 2006 paper and discuss some of the limitations you mention. They conclude that the 1% should be used and cumulative airborne fraction is a good measure. This is closely related to subsequent papers which derived TCRE or similar metrics relating cumulative emissions to warming levels. Gregory et al also perform and acknowledge differences between scenarios due to rate of change - beta and gamma feedback metrics are seen to vary in 0.5%, 1% and 2% rates of rise.

3. I like how you show various outputs change in time during the various simulations (airborne fraction etc). Can you also derive and show TCRE? You may find that this is actually better behaved in terms of being more constant in time and between scenarios. Which is a nice feature of it in fact.

4. Top of page 5 lists a nice sequence of phases (accelerating/decelerating emissions etc). I agree it is good to make sure these are assessed. In fact RCP2.6 makes a nice example of this succession and my 2016 ERL paper

(http://iopscience.iop.org/article/10.1088/1748-9326/11/9/095012). In there we show explicitly a sequence of how human and nature sinks/sources gradually transition from positives to negatives and the interesting dynamics of the earth system. To some extent therefore this scenario can achieve (but not in a clean idealised way) they same sequence that you get via your logistic pathway.

5. I don't disagree with your choice of a pathway - it would indeed be useful. There are also many other possible choices which would be useful. Various ones were discussed during our selection of the latest generation of C4MIP experiments, and include:

- 4xCO2 run, BGC mode, extended beyond 150 years - this gives a large signal to noise and the step change helps avoid conflating various timescales of response

- ZEC - as you suggest a sudden stop in emissions and let the model run free - ideally from a "policy relevant" level of CO2 (such as 2xCO2, rather than 4xCO2)

- CO2 pulse (as per Joos et al 2013, ACP)

- 1% ramp-down

- other (faster/slower than 1%) idealised % runs

there were also desires to run other scenarios as well as the idealised cases (e.g. an emissions-driven RCP2.6). We also tried to align with other MIPs - such as LUMIP.

In conclusion therefore - in order to not end up with way too many model years required from model groups, we selected a small and succinct set. It is highly likely, as you suggest, that this is not perfect and there will be value in other simulations too. For CMIP7 we can certainly open this discussion again and evolve our thinking once more.

On reflection I feel the ZEC run in particular would be very valuable.

And in fact the 1% ramp-down has now entered into CMIP6 via CDR-MIP. CDRMIP is explicitly focussed on negative emissions, as the name suggests. Please can you mention this and the negative-pulse experiment discussed in Keller et al (2018, GMD)

So in summary - the main concern over your suggestions is simply computational expense. Your logistic experiment is many hundreds of years - I can value in this, but it needs to be accessible by ESM groups. If we were to require BGC coupled version too then this doubles.

6. Your point about needing to explore low stabilisation and/or peak-and-decline scenarios is well made, and I fully agree. In fact I'd like to point you to my recent phd thesis available here: https://ore.exeter.ac.uk/repository/handle/10871/27943 - this has (I hope!) some useful background on the feedback framework (section 3.1) including discussion of Gregory et al 2009 - then I make some very similar points to you in section 4.3

minor points:

1. Intro. don't confuse CMIP and IPCC - they have very different remits (even if in reality there is overlap of who takes part). CMIP is the modelling community. They design and run the simulations. IPCC assembles experts to assess the literature - these often draw on, but are not limited to, CMIP simulations. IPCC itself neither does, nor recommends science - it does not choose which scenarios for example CMIP should run.

2. 8xCO_2 might be interesting, but (hopefully!) is not policy relevant. I think this would stretch any linearity of the system and not be useful for policy targets. I would expect most ESM groups therefore not to do this one, although EMIC groups, less limited by CPU, may well do.

3. I'm not sure of the value of plotting the compatible fossil fuel emissions for either the 1% or logistic scenarios. To me this is not a relevant quantity. I think experiments should EITHER be "realistic" - i.e. follow a plausible scenario to try to derive useful information about how the real world may unfold, OR be "idealised" - i.e. stripped down or simplified in some way to aide understanding of the system. Both have great value, but shouldn't be mixed. The fossil fuel emissions that would be required to follow a scenario are only really an interesting quantity in the first case. In the second case I don't

think they have either scientific nor policy interest. So I would stick to showing more process-based quantities, such as the land/ocean components, the airborne fraction etc. But not the fossil emissions.

4. in figure 6 - as well as a split into soil/veg carbon. Have you also looked at regional splits? e.g. tropics vs high-latitudes? I could imagine these behave differently and might be interesting to see them separated.

5. on p.7 you say that the increasing ocean-fraction has never been pointed out. While this is true of AR5 (perhaps an omission there), the analysis in Jones et al (2013, J. Climate) does cover this - see the bottom right of our figure 7.

6. a couple of other papers you might want to see: Randerson et al 2015, GCB on the long timescales and how the ocean becomes more important; Schwinger et al 2018, GRL, on ocean carbon reversibility.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-150, 2018.