Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-88-AC1, 2016

Geoscientific Model Development Discussions

Interactive comment on "The Radiative Forcing Model Intercomparison Project (RFMIP): Experimental Protocol for CMIP6" *by* Robert Pincus et al.

Robert Pincus et al.

robert.pincus@colorado.edu

Received and published: 16 June 2016

Dear Keith -

Thanks very much for these helpful comments. Our responses are below.

2-(6-7): I agree with the "normally interpreted" here, but there have been important indications that diversity in the forcing is also responsible . . .

Yes, and indeed the point of RFMIP is that some of the diversity in response is likely to due diversity in forcing. In revisions will make this idea more explicit and cite the Chung and Soden paper in support of the point.

2-13: "Observational estimates of the radiative forcing . . . " I didn't quite get this sentence. The reference to Skeie et al. seems strange . . .

It's true that estimates of radiative forcing as in the Skeie paper necessarily involve rather a lot of modeling, both with respect to radiative transfer and to composition, so that calling that work an "observational estimate" is overstating the case. We'll choose a more nuanced way to make this point in the revisions.

2-10: IRP – presumably this means "instantaneous radiative perturbation" but this isn't spelled out, nor is IRP used in some of the following paragraphs (it is given in full). I'm also not quite sure why IRP is "more precise" than IRF, but perhaps I miss something.

We introduced the IRP nomenclature late in the writing of the manuscript. We are trying find language that clearly distinguishes between theoretical changes to fluxes, i.e. IRF or IRP, that are never actually realized, and the actual change in energy balance to which the climate system responds, i.e. ERF including adjustments. But we clearly didn't make this argument convincingly enough (reviewer 2, Johannes Quaas, is similarly confused) so we'll revisit this carefully in the revisions.

2-22 and 2(32-33): The masking effect of clouds is also normally included in the IRP (and stratosphere-adjusted forcing) – I wasn't sure why it was just mentioned in the context of ERF. The difference with ERF is that the clouds can respond to the forcing.

Good point – we were trying to emphasize that even clear-sky instantaneous flux changes might differ among models for the same change in composition, but we'll mention cloud masking at 2-22 as well.

6-5 "somewhat surprising" and 6-10 "might be expected to have some error" seem contradictory. . . . Perhaps this shows that codes should be tested in implementations as close to those used in the ESM as possible, and this is a point made by the Chung and Soden paper referred to above.

We are trying to distinguish between true uncertainty, which is very small, and the approximation/parameterization errors we aim to assess.

We debated asking people to test the radiation models as you describe, basically as they are implemented in the host model. The drawback is that then each model will work on its own vertical grid, so different models won't be solving precisely the same problem.

6-11 "current spectroscopic knowledge" – I am not sure what the evidence is that radiation parameterisation error is due, to a significant degree, to spectroscopic knowledge. For example, Kratz (10.1016/j.jqsrt.2007.10.010 – see especially his Table 6) shows rather small impacts of changing HITRAN database for any post-1990 data base . . . and that is also my overall experience.

It's true that spectroscopic knowledge in the longwave, as reflected in the HITRAN databases, has not changed much in the last 25 years, at least for "normal" atmospheric conditions. But the number of known absorption lines in the shortwave continues to increase, as the HIRTRAN compilers themselves note (10.1016/j.jqsrt.2013.07.002), and it's in the shortwave where parameterizations vary most widely, especially for absorption, a point made by Pincus et al., 2015 and evident in the studies of hydrologic sensitivity (De Angelis et al, 2015, and Fildier and Collins, 2015). We'll emphasize the shortwave with a phrase in the revisions.

6(17-21) I slightly lost the plot here – firstly I am not sure what evidence there is that cloud optical properties is "likely to have a larger impact" is (especially in the context of global mean forcing), and if it is the case, it seems to undermine the reason for focusing on clear skies. Our point is that differences in the cloud distributions among models – where and when the clouds are – will almost certainly have a larger impact than model errors in treating those clouds in the radiation scheme. It's the errors we're

6(25) "obscured" – I think the intermodel spread is pretty evident in Collins et al. (2006) where the tropopause standard deviation is almost as big as the mean; I think what has happened in more recent years is a realisation that this really matters for hydrologic(al)

sensitivity.

The narrow range of conditions have obscured important parameterization errors in **sensitivity**, like the sensitivity of solar absorption to water vapor. One can perhaps see this point in retrospect in comparisons like Collins et al. 2006 but its hard to distinguish, say, bias from incorrect sensitivity with a single atmosphere. We'll express this point more clearly in the revisions.

7(2) I had a few comments on this table. One is that specification of "HFC", rather than a specific HFC, seems too vague to me . . . Second the +4K experiment is well motivated, but not really elaborated on – I guess the focus is on the temperature dependence of transmittance, but this experiment might also have some Planck function dependencies . . . note a typo for the ozone experiment, where it says O2 not O3.

The atmospheric conditions specify the HFCs precisely. Our current draft specifies CFC11, CFC12, HCFC22 although this list is under discussion.

The +4K experiment is indeed aimed at the temperature dependence of transmittance. It's true that we won't be able to tease apart the Planck function dependence although we're not aware of any evidence that this is a source of error.

We'll add more detail to this table and the surrounding discussion. Thanks for catching the typo as well.

7-23: Not entirely sure why LBLRTM is singled out here . . .

LBLRTM is singled out because this is the model with which we, meaning the RFMIP participants, will make reference radiative transfer calculations. Several other groups have agreed to do so informally but we can't commit on their behalf. We will clarify this in the revision.

7-30 Perhaps the most major of my comments. I naturally assumed that the shortwave GHG forcing would be included for the greenhouse gases, but I realised at this point that while the insolation conditions are specified for aerosols, they are not specified for *greenhouse gases . . . it would be advantageous to have a proper day/global average even if it is for a single day, as in the aerosol case. Could the authors at least clarify the situation regarding the GHG SW forcing?*

As Johannes Quaas (reviewer 2) also noted, this table is a little sparse by way of explanation. Solar zenith angle is indeed specified in the experimental protocol, and varies among profiles so that one can compute the diurnal average.

8-3 "parameterization error increases model diversity" – isn't parameterization error a (and possibly the major) CAUSE of model diversity (unless all models have the same error!) As above, RFMIP-IRF will assess the error in clear-sky and aerosol radiative transfer parameterizations. We can call this "error" because the correct answer is known for radiation. The same is not true for clouds or convection or a host of other parameterizations for which there are reference calculations analogous to line-by-line calculations.

8-9 This feels like sloganeering, and I haven't really got a clue what "the 20th century belonged to sulphate" means (there are many contenders to the accolade of "the 20th century belonged to ..."! I might vote for the Beatles), especially given the claim that a strong negative aerosol forcing is implausible and the statement that greenhouse gas forcing dominates at 2-23. Incidentally, is the reference to Carslaw et al., in the following sentence, the right one? There seems little or nothing on sulphate global optical depths in the cited paper.

8-13 "Starting in the mid-1970s" – maybe better "Since the mid-1970's"? I presume the stated changes (five-fold, factor of two etc) refer to the present day compared to mid-1970s?

We have rewritten this paragraph to make it less playful. The reference to Carslaw et al. was to their 'Extended Data Table 2' which gives SO2 emissions, so the original statement assumed (consistent with the literature) that the the conversion of SO2 to sulfate and the lifetime of SO2 did not change dramatically. To avoid confusion we

have made these statements more precise.

"In the 20th Century sulfate is thought to have contributed substantially to the net radiative forcing, although how is disputed (Stevens 2015). What is to disputed is that precursor SO_2 emissions increased greatly, and that these emissions were concentrated over a relatively small portion of the planet. Consistent with other studies, Carslaw et al. (2013) estimate that SO_2 emissions, to which the dominant component of the aerosol contribution to ERF are attributed, increased three-fold through the first hundred years of industrialization. Smith et al., (2011) pinpoint these changes to changes in the North Atlantic sector – a region covering about a tenth of Earth's surface. Beginning in the 1970s air quality controls began to reduce emissions in Western Europe and North America. Present Western European emissions are now estimated to be a fifth, and North American Emissions a half, of what they were in the early 1970s. As emissions over the Atlantic sector declined, emissions over South and East Asia increased so that globally anthropogenic SO2 emissions remained roughly constant. The short life-time of sulfate implies that the regional concentration of emissions would lead to commensurately large regional forcing. So to the extent that sulfate forcing is important globally, regional signals should be readily identifiable, and may help bound the overall radiative forcing attributable to anthropogenic SO_2 emissions."

8-29 "less negative" – isn't this "more negative"?

Yes thank you.

9-3 typo "response"

Thanks for catching the typo, but in reviewing this sentence we also found it unwieldy. We have revised it to read:

By more tightly-constraining the pattern of the aerosol effective radiative forcing across models it should be easier to identify a clear response of the climate system to the imposed aerosol perturbations. To the extent that clear responses can be identified, they may be combined with formal methods of detection and attribution (e.g., Stott et al., 2010) to also estimate the magnitude of the forcing. 9-18 This section was written in a different way to the equivalent in Section 3. I could not see a reference to Table 5, and the section is a bit cluttered by the acronyms of the experiments, which are a bit opaque (to me). Perhaps this section could be restructured (not a big job) to do as was done in Section 3, and leave the acronyms to the table, but ensure the underlying motivation of experiments is spelled out.

9-31 "warming" – sorry if I am getting confused, but as I understand, RFMIP only provides evidence of "forcing" rather than "warming". Perhaps it is the coupling with DAMIP that is being referred to here?

Here we did mean warming, as we will have the temperature record from the historical simulations and we will compare these directly. To make this more clear we have modified the sentence to say: "The historical simulations based on MACv2-SP will be analyzed, also in cooperation with DAMIP, to test the hypothesis by Stevens (2015) that the observed northern-hemispheric warming is inconsistent with an aerosol radiative forcing more negative than about -1W/m$^{2\prime\prime}$

10-2 "warming hole" –same comment as above. In fact the whole of this sentence, seems to go a bit beyond what RFMIP results could achieve (although they give important clues), so I wondered if this was really referring to DAMIP.

Yes, we understand how this might be confusing. We have revised the sentence to make clear that simply the identification of a pattern of warming that is consistent with that which is often attributed to aerosol forcing would provide a constraint on that forcing. The revised sentence reads:"For example the pattern, or lack thereof, of the response across the multi-model ensemble may be helpful to advancing our understanding of the extent to which aerosol forcing underlies the warming hole in the east-central United States . . . "

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-88, 2016.