Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2018-747-RC1, 2018

Atmospheric Chemistry and Physics Discussions

ACPD

Interactive comment on "Local and remote temperature response of regional SO_2 emissions" *by* Anna Lewinschal et al.

Anonymous Referee #1

Received and published: 9 October 2018

This manuscript presents a very interesting and important set of coupled atmosphere-ocean simulations, in which SO2 emissions have been increased or decreased separately in different geographic regions. Studies which have systematically performed aerosol forcing experiments in different individual regions like this are rare, and so the addition of these simulations with the NorESM model is a very valuable addition to the existing literature, that allows for comparison with other models. This manuscript also goes further by looking at both increasing and decreasing aerosol perturbations, allowing the authors to test the linearity of the response, and showing that there may be considerable non-linearities which is an important caveat when constructing regional metrics.

I recommend the manuscript be accepted provided the following points are addressed:

Printer-friendly version

Discussion paper

Scientific comments:

L26/Introduction: The authors should additionally mention the work of Conley et al. (JGR:A 2018, https://doi.org/10.1002/2017JD027411) and Kasoar et al. (npj Climate and Atmospheric Science 2018, https://doi.org/10.1038/s41612-018-0022-z) which are both highly relevant and could be compared directly with this study. Conley et al. present global temperature changes to removing US SO2 emissions in three different atmosphere-ocean models, while Kasoar et al. show temperature responses due to removing SO2 emissions individually from either North America, Europe, South Asia, or East Asia in the HadGEM3 atmosphere-ocean model, directly analogous to the present study with NorESM.

L182: Are the 'climatology' aerosols diagnosed from the control simulation? Or do they come from somewhere else? Essentially, I want to double-check that a free-running control simulation would by construction have zero RF – which might not be the case if the 'climatology' isn't equal to the online control aerosol distribution.

Figure 3: Though a nice way of presenting this info, I'm not sure this figure is critical given that some of the information can also be discerned from the error bars e.g. on Fig. 5. The current Fig. 3 could probably be moved to the supplement. Instead, what would be much more useful to include here would be global maps of the temperature changes in each of the experiments, perhaps with shading or stippling to indicate significance at each grid point. This would allow the same comparison as the present Figure 3 in terms of seeing how similar and significant the responses are in different regions, but would also allow for the pattern of the temperature responses to be compared against other studies. I find it odd that this paper does not currently include a single plot which just shows the geographic temperature changes from these experiments, which to me is the first thing I would look at.

Table 2: It could be useful to also include the climate sensitivity values (dT/ERF and dT/RF) in this table, as they are used later in the discussion.

Table 2: Please include uncertainty ranges

L206/Figure 3 plus subsequent figures that have error bars: How is significance determined? The paper frequently discusses whether results from different regions are significantly different from each other, but I am unclear how this was tested for. Similarly error bars are sometimes quoted as one standard deviation, but in the absence of a large ensemble of simulations I'm unclear what it's a standard deviation of.

L234: Quoting a correlation coefficient for four data points (three of which are pretty much on top of each other) is arguably misleading – it's bound to be close to 1, but this doesn't necessarily tell you much about the strength of the relationship given that not much variation was sampled

L239-249/Figure 4b: This is a very interesting result which I find it hard to get my head around. The whole reason that ERF is widely used is because it has generally been shown to be a better predictor of dT than the instantaneous RF – at least across different and varied forcings. Here you find the opposite – but moreover finding that emission change is an even better predictor of dT! Given that sulfate aerosol has little atmospheric absorption and affects the surface temperatures pretty much entirely through TOA radiative forcing, I would really like to understand why the ERF correlates worse with dT than the emission change. Do the authors have any ideas, physically, how this comes about in these experiments? E.g. maybe some large land-surface responses in the fixed SST experiments, which mean that a substantial portion of the final temperature response is subsumed in rapid adjustments? (N.B. As noted later in the manuscript though, once you include the 0xEU experiment, then ERF does become a better predictor of dT again).

L262: The SA response can't be weaker in all the latitude bands, or else it would also be weaker globally.

L263-L270/Figure 6: I'm confused by the different indications of significance. E.g in Figure 6d, the 10xSA ERF in the tropics has an error bar which does not cross zero, and yet it is shaded to indicate that it's not significant. Yet in Fig 6b, the 5xNA ERF in the NHml has a huge error bar that spans zero, but is shaded to indicate that it is significant.

L316-318: Consider citing Teng et al. (GRL 2012, doi:10.1029/2012GL051723) which provides a similar example of aerosol forcing over Asia resulting in remote warming over the US, in a different model

L415: Units of climate sensitivity seem to have been inverted here. Check that the number being quoted isn't actually the feedback parameter.

L422: Why is the goal to use model-dependent sensitivities? Surely for integrated assessment modelling, you would like to use a model-independent choice of climate sensitivity? So then, does it matter if you assume a different climate sensitivity and get a different answer (scaled up or scaled down) as a result?

L424: why do you compare with the GISS-E2 transient sensitivity and not equilibrium, given that the NorESM simulations aren't transient?

L427-428: I don't agree with this conclusion. The way I read it, using the Shindell and Faluvegi coefficients has reproduced NorESM well here because GISS-E2 has an ECS (i.e. 2 x CO2) climate sensitivity of 0.6 K/(Wm-2) (Flato et al. IPCC 2013), which happens to be very similar to the sulfate climate sensitivity found for NorESM here. Hansen et al. (JGR 2005, doi:10.1029/2005JD005776) shows that in GISS-E2, the ERF-based sulfate sensitivity is similar to CO2. So this doesn't explain why the authors get such different climate sensitivities for sulfate and 2 x CO2 simulations in NorESM. Maybe due to differences in methodology defining the equilibrium state? Or in calculating ERF (e.g. fixed-SST versus Gregory regression?). At any rate, it would be interesting to understand why NorESM seems to have such a different climate sensitivity for sulfate here compared with the previously published 2 x CO2 values. The message of e.g. the Hansen et al. (2005) paper is that ERF is a more forcing-independent predictor of temperature change, so it's surprising that the global climate sensitivity of NorESM varies so much between forcers. One final point, is that climate sensitivities in general differ hugely between models for the same forcing agent (e.g. the range in Flato et al. is from ∼0.5 to 1.5 K/(Wm-2)). This is presumably the case for sulfate as much as any other forcer. So, the coincidence that GISS-E2 has a similar climate sensitivity to NorESM doesn't really show that there is smaller variation across models in the sulfate climate sensitivity compared with between different forcers in the same model; this seems quite unlikely to be the case across most models in fact.

L486-487: If saturation of aerosol indirect effects is the explanation here, then shouldn't there be a similar difference in the RF/em as there is in the ERF/em? I don't see how the RF/em can be unaffected by CDNC saturation such that it only shows up as a difference in ERF/em. On a related point: In Figure 4, the error bars for the 0xEU response per em or per RF are enormous and span the entire range of the other experiments. Can the authors be confident that the sensitivity to an emissions reduction actually is any different to an emissions increase, given the considerable overlap of the error bars? It might just be that the smaller forcing and smaller response from the 0x experiment has higher uncertainty because the signal is small compared to internal model variability.

Technical/grammar/typographic comments:

L69-70: Confusing wording in this sentence, please re-phrase

L176-177 and Eq3: Inconsistent use of r subscript (emission region or response region?)

L217: add 'typically' or similar caveat

Figure 4: The caption should explain how the quantities are normalised. Currently, have to refer to the main text to find out that everything is normalised to the 5xNA experiment in this plot. L263: Should Figure 6 have been referenced here?

L299: increase -> increases

L495: skills -> skill

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-747, 2018.