

Reply to reviewer #1

Reviewer comments in *italic* and author response starting with [CW]

Overall, the authors have done a good job responding to both reviews. The re-ordering of sections and improvements in the figures helps a great deal.

[CW] Thank you for reviewing the revised paper and for these additional helpful comments to improve the paper.

However, I still have a couple of suggestions for the paper.

Abstract: The Abstract could be improved by including more conclusions from the paper, and more context for this particular work. The first 3 sentences seem quite general and could perhaps be condensed to one introductory sentence.

[CW] We appreciate the suggestions to add more to the abstract, and we do, following your additional comments below. The first 3 sentences have been reduced to this one sentence: "While carbon dioxide is the main cause for global warming, modelling short-lived climate forcers (SLCFs) such as methane, ozone, and particles in the Arctic allows us to simulate near-term climate and health impacts for a sensitive, pristine region that is warming at three times the global rate."

It would be good to include highlights of the conclusions that you discuss in section 5, such as what the models do well, the fact that the multi-model mean is appropriate for the radiation and health impacts determined in the AMAP report.

[CW] We have expanded/added these sentences in the abstract, which include more of our important conclusions from section 5:

"The multi-model mean was able to represent the general features of SLCFs in the Arctic, and had the best overall performance. For the SLCFs with greatest radiative impact (CH₄, O₃, BC, and SO₄), the mmm was within $\pm 25\%$ of the measurements across the northern hemisphere. Therefore, we recommend a multi-model ensemble be used for simulating climate and health impacts of SLCFs. Of the SLCFs in our study, model biases were smallest for CH₄ and greatest for OA. For most SLCFs, model biases skewed from positive to negative with increasing latitude."

It would also be useful to mention which evaluation metrics were found to be most important and that should be used in future comparisons.

[CW] We have added these sentences in the abstract:

"The annual means, seasonal cycles, and 3-D distributions of SLCFs were evaluated using several metrics, such as absolute and percent model biases and correlation coefficients."

"... As model development proceeds in these areas, we highly recommend that the vertical and 3-D distribution of SLCFs be evaluated, as that information is critical to improving the uncertain processes in models."

MOPITT, Fig. 12: I am still surprised that the coverage of MOPITT data is so sparse, and urge the authors to review their processing of the MOPITT retrievals, and their treatment of missing values. Is no data shown for a given grid box if any month of the 2 years has missing data? Perhaps an average of JJA would be appropriate to show here.

[CW] Thank you for highlighting this issue again. We did have an error related to masking null values in the MOPITT retrievals, which caused no data to be shown for a given grid box if any month of the 2 years had missing data. This has now been corrected and we have also changed the figure to show the June-July-August (JJA) mean instead, which for MOPITT has much better coverage in the Arctic region, and the models are actually biased high. Even at mid-latitudes, in the summertime the models do not underestimate CO like they do in the winter and spring (which had dominated the annual mean), and this is evident in our new plot. Therefore, we have updated the text discussing this figure, and we have put the springtime mean in the supplemental material.