Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2020-86-RC1, 2020
© Author(s) 2020. This work is distributed under
the Creative Commons Attribution 4.0 License.

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Projecting ozone hole recovery using an ensemble of chemistry-climate models weighted by model performance and independence" *by* Matt Amos et al.

Anonymous Referee #1

Received and published: 10 March 2020

Review of 'Projecting ozone hole recovery using an ensemble of chemistry-climate models weighted by model performance and independence' by M. Amos et al.

The study by M. Amos and colleagues looks into the historical and future evolution of the Antarctic ozone hole using a weighting technique which accounts for model independence and performance. It is based on an ensemble of Chemisty-Climate Models, which are combined into a weighted multi-model mean by applying a weighting method established in the literature. The authors find that ozone returns to 1980 levels about 3 years later in the weighted mean compared to the unweighted one.

The manuscript is very well written and concise which makes it easy to follow the au-

Printer-friendly version

Discussion paper

thors arguments. In addition, the authors openly address several potential shortcomings of their approach and included an extensive analysis of the performance based on cross-validation, which I find to be an excellent example of best practice. While the results the authors find are not ground-braking, the work, in my opinion, still presents a valuable contribution to the field in particular in highlighting the importance of accounting for model inter-dependence when calculating uncertainty intervals based on MME spread. I found some unclear statements which should be addressed as noted in my minor comments below otherwise I recommend this contribution for publication in ACP.

Minor comments

L1: I don't think that there is a single, agreed-on 'current method' for averaging model ensembles in climate science.

L70: This is a minor point but I'm just pointing out that the REA does not down-weight models if they are more similar (dependent). It rather does the opposite and gives models which are closer to the multi-model mean additional weight as they are considered to be more reliable.

L90: There are several more recent papers addressing and further developing this weighting method: http://doi.wiley.com/10.1029/2017JD027992 https://doi.org/10.1088/1748-9326/ab492f

Equation 1: The sum in the denominator should run over j!=i I assume?

Section 2.1/Figure 1: I fail to understand how the sigma values where chosen from the figure and the text in this section. Can the authors elaborate on that please?

L231: delete 'trend'? Figure 2 shows a time series of TCO not a trend right?

L252: Figure 2 → Figure 3

l268: Maybe give the weight here relative to equal weighting similar to the lowest weight later in the paragraph. This seems to be the more meaningful metric as the absolute

CC BY

weight depends on the ensemble size.

Figure 4: This might just be an ambiguity in the language, so let me see if I understand correctly: Given a perfect model X, there is an unweighted projection Y and a weighted projection Y' correct? So what is shown here is the difference abs(X-Y)-abs(X-Y')? If that is they case maybe rather call it the improvement of the (weighted) average than the average improvement? (except the last bar). I was a bit confused about the "Average improvement" which seems to indicate that this is showing some kind of average over different improvements (if my interpretation is correct).

L296: 'alike any two models are' I don't think that this should be the aim or at least it should be worded more carefully. Two models can be 'alike' for different reasons, one of them being that they simulate the same system and are both 'good' at it so that the converge towards the truth. Or in other words: models should not be punished for arriving at the same answer independently. The assumption behind independence weighting schemes is that an inference on the model dependence can be made based on their output. If, e.g., two models share biases in several metrics this could be a sign that they also share one or several components.

L355: Have the authors tried to create the weights based on the refC2 simulations directly? Is it possible to receive sensible weights from that and how do they compare to the weights retrieved from the hindcasts?

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-86, 2020.

Printer-friendly version

Discussion paper

CC BY