Atmos. Chem. Phys. Discuss., 9, C6732–C6735, 2009
www.atmos-chem-phys-discuss.net/9/C6732/2009/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Objective assessment of ozone in chemistry-climate model simulations" *by* A. Yu. Karpechko et al.

A. Yu. Karpechko et al.

a.karpechko@uea.ac.uk

Received and published: 5 November 2009

We thank Dr. Thomas Reichler for his positive review of our manuscript. Below are our responses:

19354/14: Could you explain better what you mean with first round of CCMVal? Not everybody is familiar with this project.

We amend the text with the following sentences:

The goal of CCMVal is "to improve understanding of CCMs and their underlying GCMs (General Circulation Models) through process-oriented evaluation" (CCMVal: http://www.pa.op.dlr.de/CCMVal/, access 29 October 2009). The first round was accomplished in support of WMO ozone assessment 2006 (WMO, 2007).

19357/8: Maybe it should be made clear that only zonal mean ozone fields are used.

Done

19358/15-16: This is unclear. What is meant with ": : : combined with the sampling errors by root mean squares"?

We expand the discussion and amend the text as follows:

Measurement errors for individual data values are only available for the Hassler data set. We calculate measurement errors for our diagnostics using the law of combination of errors:

σiml'^2=$\sum$[(∂fl/∂yimn)^2*σimn^2] (X)

where yimn and σimn are individual observations and errors at month m and grid point i and fl is either the function for the mean in the case of the climatology, or the function for the slope parameter in the case of the trend. These measurement errors are combined with the sampling errors (i.e. with the standard error of the mean or with the standard error of the slope parameter) at each month and grid point by root mean squares and the resulting σ are used for all the three profile data sets.

19362/15: When I first read this I was confused. Maybe you should better explain what exactly is done, i.e., that different model rankings are investigated.

We presume the confusion is caused by the text saying we correlate model errors whereas we indeed correlate the model ranks calculated in the same diagnostic but with respect to different observation data sets. The text is rephrased as follows:

"To investigate whether the choice of reference data set has a large impact on the model ranking we calculate Spearman's rank correlation coefficient between model ranks calculated in the same diagnostic but with respect to different observation data sets."

19363/8: Again, it seems this paragraph could be improved by explaining better the

ACPD

9, C6732–C6735, 2009

methodology.

Similarly to the previous comment we correct the text by saying that the model ranks are correlated, not the model errors as is stated in the original version.

19364/21: I am having a hard time to see from Figure 11b the zero grades in the total ozone climatology. Can this be shown clearer?

We will improve the figure quality to make sure the details are clearly seen.

19367/14-17: I do not understand the logic here. The ranking of individual models is sensitive to the chosen metric. But you found quite some good correspondence between your results and Eyring et al. (2006, 2008). So, maybe the simplified ozone diagnostics is not a good metric. But this does not necessarily mean that the multi-model mean is the best estimate of future ozone.

We agree with the comment of Dr. Reichler and thank him for pointing this out. Our results do suggest that the annual mean diagnostics are not a good metric for evaluation of model performance in ozone simulations. The annual mean diagnostics omit the seasonal cycle and the differences between model ranks obtained with and without consideration of the seasonal cycle found here suggest that a good simulation of annual mean values does not guarantee a good simulation of the seasonal cycle, which is quite important in the case of stratospheric ozone. However we have performed other sensitivity tests in which we slightly modified the original metric that includes the seasonal cycle, as briefly mentioned in the last paragraph of Section 4. These sensitivity tests show that the individual models may change their rank depending on the metric chosen. In particular, two models show the lowest combined error among the individual models (i.e. get the highest rank among the individual models) depending on the metric chosen. On the other hand the model ensemble mean always performs better than the individual models in these tests, irrespective of which metric is used. So, assuming that correct simulation of past ozone gives more confidence in future projections, we conclude that the model ensemble mean is the optimal estimate of future ozone. We

ACPD

9, C6732–C6735, 2009

are planning to show the uncertainties estimated using these sensitivity tests in Figure 11a in the revised version.

Typos:

19354/24: vs. 1.0 –> V1.0

Corrected

19364/23: in a contrast –> in contrast

Corrected

Interactive comment on Atmos. Chem. Phys. Discuss., 9, 19351, 2009.