acp reviewer 2

Kevin Smalley

March 2017

We thank the second reviewer for their helpful remarks.

Smalley et al. analyse CCM model predictions of stratospheric water changes over the 21st century. A multivariate linear regression is applied to the models' stratospheric water entry mixing ratios ("$[H_2O]_{entry}$"), with the explanatory variables being a "tropospheric temperature index", a "Brewer Dobson strength" index, and a QBO index; this analysis follows the method of Dessler et al. (2013). Overall, the analysis is straight- forward, and the results are clearly described. I do not comment on the aspects of the statistical analysis brought up by the other reviewer.

However, this reviewer cannot quite see that "Our approach provides more insight into model processes than simply comparing $[H_2O]_{entry}$ or TTL temperatures." (Page 7/Line 19).

We strongly disagree with this comment. Comparing water vapor and TTL temperatures tells you nothing about the contribution of individual processes that are responsible for $[H_2O]_{entry}$ variability. Or analysis breaks down variability in $[H_2O]_{entry}$ by process. That being said, we modified the text on page 8, lines 8-10 to try to make this clearer.

Rather, the paper is somewhat superficial (it certainly does not help that (Page 2/Line 13): "Finally, a warmer troposphere tends to increase $[H_2O]entry$, although whether this is through influence on TTL temperatures or some other mechanism such as convective ice lofting, is not clear."), and results are few. It would be great if the authors would work out the connection between tropopause temperatures and $[H_2O]_{entry}$ in the models, and the connection between "tropospheric temperature" and tropopause temperature.

We disagree that the paper is superficial. We view this as an important new technique to diagnose processes in CCMs, which can reveal problems in the CCMs not apparent by just looking at $[H_2O]entry$ and TTL temperatures. That said, we have added more text (lines 7-13 on page 2) that discusses the connection between tropospheric temperature, TTL temperature, and $[H_2O]_{entry}$.

The QBO results would also deserve some further analysis - for the 21st century analysis, annual mean data is analysed. This evidently removes much of the variance associated with the QBO, and it appears that the lack of influence of the QBO (as e.g. shown in Figure 2) is due to a lack of a trend in the QBO index. This evidently begs the question why the model does not have a QBO trend when it has been argued that the tropospheric expansion associated with global warming would have an impact on the lower stratospheric QBO - and as such would be reflected in the QBO index. While this may not have an impact on $[H_2O]entry$ (because the QBO influence at the rising tropopause level main remain constant over time), it would be useful to have some more information why the QBO index (as e.g. shown in Figure 2) does not show a trend.

We first note that this comment shows the usefulness of our analysis (contradicting the reviewer's earlier comment): just comparing $[H_2O]entry$ and TTL temperatures would not reveal this problem with the QBO. That said, we disagree with the overall comment. Our paper is designed to understand how these processes (BDC, QBO, ΔT) affect $[H_2O]entry$, not why the processes evolve as they do. Understanding why the BDC, QBO, etc. evolve as they do over the 21st century is far beyond the scope of this paper. Our paper is nonetheless an extremely useful result — by identifying this issue, our paper will spur additional research into why the QBO is not impact water vapor in the way suggested by the models.

Two additional minor comments:

Please provide a reference for the statement that "Virtually all climate models ..." (page 2/Line 14)

This sentence has been removed from the current manuscript

and some more information about the differences in results for models that participated in CCMI-I and CCMVal-2 would be useful.

After lengthy consideration, we've decided that there's no easy way to summarize the differences in the models in these two groups because there are no systematic differences. Trying to summarize the differences in the text therefore was unwieldy and created difficult-to-read, boring text. If people are interested in this, they can determine it using the Tables in our paper.