Comments: A Machine Learning Examination of Hydroxyl Radical Differences Among Model Simulations for CCMI-1 by Nicely et al., 2019

Karl M. Seltzer, Prasad Kasibhatla

Nicholas School of the Environment, Duke University, Durham, NC, USA

Email: karl.seltzer@duke.edu, psk9@duke.edu

General Comments

The manuscript "A Machine Learning Examination of Hydroxyl Radical Differences Among Model Simulations for CCMI-1" by Nicely et al. discusses a topic that is of high interest to the Atmospheric Chemistry and Physics community. Possibly the most perplexing issue in atmospheric chemistry is the unexpected stabilization of global methane concentrations from ~2000-2006. This study attempts to unravel the individual CTM drivers of the hydroxyl radical in a suite of simulations, thus illuminating the changes, and reason for said changes, in the primary termination pathway for methane, as simulated by each CTM.

While this work is important, we do have concerns about how some of the results are presented and methods are employed in this analysis, both of which constitute major comments. We will describe both in more detail below, followed by some minor comments.

Major Comments

1. In Figures 7-10, results from the CH_4 signal, as it relates to changes in tropospheric OH, are presented. While the text does explicitly state that "CH_4" is a normalized value based on the maximum tropospheric value, we believe the presentation of the results in Figures 7-10 and much of the language used throughout the manuscript can lead to substantial confusion on the part of the reader. The reader might reasonably interpret the results as an estimate of the sensitivity of τ_{CH4xOH} to changes in CH_4 abundance (i.e. the CH_4 feedback factor). One example: the inclusion of CH_4 in Figure 10 makes a comparison of the "CH_4" value reported in this study (i.e. NOT the CH_4 feedback factor) with the calculated CH_4 feedback factor from Nicely et al., 2018.

   Based on our interpretation of the methods employed here, the authors did not analyze the CH_4 feedback factor. Since it seems the better characterization is that the global *distributions* of CH_4 concentrations were analyzed, we think the authors need to re-write any discussions related to CH_4 results throughout the manuscript to make this distinction abundantly more clear, and should possibly remove the characterization of "CH_4" in Figures 7-10. Similarly, it is not clear why CH_4 concentrations were normalized. Presumably, the same analysis using non-normalized values of CH_4 would be able to capture the CH_4 feedback?

2. The sensitivity of τ_{CH4xOH} to changes in CH_4 abundance reported by CTM studies are reasonably consistent and range from -0.25 to -0.35 (Prather et al., 2001; Fiore et al., 2009; Holmes et al., 2013, Holmes 2018). That is, the tropospheric OH abundance declines by 0.25%-0.35% for every 1% increase in CH_4 abundance (Prather et al., 2001). The IPCC AR5 reported that global CH_4 abundance grew by ~13% from 1980 to 2010 (Ciais et al., 2013).

Assuming the models used here respond in a similar manner to other published CTM studies, the CH_4 feedback should have yielded a ~3.3%-4.6% decrease in tropospheric OH between 1980-2010 (or equivalently, 1.1%-1.5% per decade). That driver should theoretically be captured in the net results presented in Figure 6.

As noted on Line 457, the mean downward trend in τ_{CH4} of Figure 6 is 1.8% per decade. Therefore, the residual (i.e. all of the other factors outside of the CH_4 feedback) should be ~(-1.8% - 1.3%) → -3.2% per decade (note: 1.3% is the average of 1.1% and 1.5%). This is much larger than the ~residual of -1.9% reported on Line 457 (~residual because it does not include the CH_4 feedback factor). Therefore, since the τ_{CH4} budget does not appear to be closed when adding up all of the variables (including the CH_4 feedback), this suggests that the methods used here have difficulty in deriving the contributions of individual drivers. If so, that would be a fundamental issue with the methods used to derive Figures 7-10. Here are some ways we believe the authors can build confidence in the methods used here:

a. A quick first step would be to add up all of the components for each model in Figure 7 and plot their change, side-by-side, to the values presented in Figure 6 (normalized to 2000 values for consistency). Do the trends match? If yes, since the NN method does not account for the CH_4 feedback and CTMs are known to have a robust and consistent CH_4 feedback, why do they nonetheless match? If no, can the missing CH_4 feedback explain the difference?
b. A lengthier, but maybe necessary test: Experiment with one of the CTMs. For example, re-run GMI with the year 2000 repeating for all variables, except CO. This might only be necessary for a few select years, such as 1985 and 1998. Do these results match the dark blue line in Figure 7e? One or two examples of these types of validation steps would really increase our confidence in the driver analysis.
c. When attributing specific, individual drivers to trends, Random Forests are considered better machine learning tools (Grange et al., 2018). It is likely easy to swap out the NN code in your analysis with a random forest. Experiment with one of the models. For example, run the random forest algorithm for GMI's 2000 results and repeat the process for Figure 7. How different are the results?

Minor Comments

- Figure 3 compares the tropospheric OH columns from WACCM and the ANN-WACCM predicted tropospheric OH columns. As noted on Line 174, the training methods in this analysis were the same as those carried out in Nicely et al., 2017, which stated that the training/validation/testing datasets comprised 80/10/10% of all data. Therefore, it seems that 80% of the data that was used to construct the middle panel of Figure 3 was data that the ANN has seen before (i.e. from the NN training). Shouldn't this part of the evaluation be restricted to only the testing dataset?
- In the paragraphs spanning Lines 423-448, there is a discussion about "spurious results". Are these results "spurious" just because they look out of place in Fig. 7, or are there some other quantifiable ways that might justify the label "spurious"?

- Figure 9b: Don't CTMs have difficulty in capturing observation-derived estimates of IAV (Holmes et al., 2013)? That should be noted.
- Lines 482-498 should likely be removed. The comparison of the CH_4 results here and the CH_4 results in Nicely et al., 2018 are not an 'apples-to-apples' comparison, as noted by the authors in the sentence starting with "On one hand…" from Line 485.

References

Ciais, P., C. Sabine, G. Bala, et al., 2013: Carbon and Other Biogeochemical Cycles. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. 2013.

Fiore, A. M., Dentener, F. J., Wild, O., et al.,: Multimodel estimates of intercontinental source-receptor relationships for ozone pollution, J. Geophys. Res., 114(D04301), doi:10.1029/2008JD010816, 2009.

Grange, S. K., Carslaw, D. C., Lewis, A. C., et al.,: Random forest meteorological normalisation models for Swiss PM10 trend analysis, Atmos. Chem. Phys., 18(9), 6223–6239, doi:10.5194/acp-18-6223-2018, 2018.

Holmes, C. D.: Methane Feedback on Atmospheric Chemistry: Methods, Models, and Mechanisms, J. Adv. Model. Earth Syst., 10(4), 1087–1099, doi:10.1002/2017MS001196, 2018.

Holmes, C. D., Prather, M. J., Søvde, O. A. and Myhre, G.: Future methane, hydroxyl, and their uncertainties: Key climate and emission parameters for future predictions, Atmos. Chem. Phys., 13(1), 285–302, doi:10.5194/acp-13-285-2013, 2013.

Nicely, J. M., Salawitch, R. J., Canty, T., et al.,: Quantifying the causes of differences in tropospheric OH within global models, J. Geophys. Res., 122(3), 1983–2007, doi:10.1002/2016JD026239, 2017.

Nicely, J. M., Canty, T. P., Manyin, M., et al.,: Changes in Global Tropospheric OH Expected as a Result of Climate Change Over the Last Several Decades, J. Geophys. Res. Atmos., 123(18), 10,774-10,795, doi:10.1029/2018JD028388, 2018.

Prather, M., Ehhalt, D., Dentener, F., et al.,: Atmospheric Chemistry and Greenhouse Gases, in Climate Change 2001: The Scientific Basis. Third Assessment Report of the Intergovernmental Panel on Climate Change, 2001.