Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2019-772-RC3, 2019

Atmospheric Chemistry and Physics Discussions

Interactive comment on "A Machine Learning Examination of Hydroxyl Radical Differences Among Model Simulations for CCMI-1" *by* Julie M. Nicely et al.

Leif Denby (Referee)

l.c.denby@leeds.ac.uk

Received and published: 10 November 2019

I am only commenting on the machine learning aspect of the submitted manuscript. Apologies for overlooking for not providing more general feedback.

1. In section 3.2 I would rephrase the sentence containing "mimic the tropospheric chemistry" to include "predict the instantaneous OH" concentration. As is written now it might give the impression that the time evolution is predicted by the neural networks as the research presented is about reactions.

2. I find the sentences "Briefly, one NN is trained for one model, for one simulation

month at a time." and "To reduce computational demands, we establish NNs for four months, one for each season..." a little contradictory. Is training done on one month or on four months of input? How is it possible to do both? It might be that the reader should simply study the referenced paper, but I find this a little unclear.

3. It would be nice to a brief comment on why models were trained for each month separately. Was this done because the temporal variability couldn't be captured by a single model? Does the skill of each model vary through the month? I assume that at the ends of the month (where there is transition between which model is used) there might be a reduction in skill. But maybe the predictions match seamlessly when switching between models.

4. The "Inter-model comparison" is nice. With the restriction on the numerical range of the values which are substituted I feel that feature importance could similarly be inferred by simply shuffling (across time) all values for a specific feature, similarly to how it's done for random forests. Is there a reason why this wasn't attempted here? Isn't there a concern that using the presented method that one might infer low feature importance for fields that simply vary little between models?

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-772, 2019.