Reply to reviewer 2. M. Prather

Maarten Krol at al.

February 5, 2018

First of all we would like the thank the reviewer for his insightful comments, and starting this discussion. The reviewer makes one big point, and several other points that we address below.

- *The protocol has a fundamental difficulty: for the Age of Air (AoA) tracers with fixed boundary conditions are used, which cannot be compared to real tracers like SF_6, ^{222}Rn, and e90.*: In order to derive transport times from transport model simulations, there are several options. In all cases a time scale has to be extracted from simulations, as time is not a transportable quantity. One option is to have tracers with different lifetimes (e.g. 222Rn, e90, ..) and fixed emissions. Alternatively, tracers with 'known' emissions and a source-sink imbalance can be used to investigate the age of air masses (e.g. SF6, CO2). In all cases, time-scales of mixing have to be derived from known decay rates, or from atmospheric accumulation rates. Some tracers have the clear advantage that atmospheric observations are available (SF6, CO2, 222Rn). This protocol is experimenting with boundary conditions with a prescribed growth rate, with the growth rate serving as the 'clock'. This offers advantages, because the boundary conditions can be applied in a flexible way (NH, SH, land, ocean, ...). As the reviewer rightfully remarks, we also face challenges. However, according to us there is not a fundamental difficulty with the protocol, other than the issues we discuss in the manuscript. Yes, we can have a rectifier/diode effect, but studying this in different models may help diagnose model–model differences. For instance, our surface boundary conditions are applied only in the lowest 100 m. This introduces a sensitivity for boundary layer mixing. Models with efficient BL mixing (like TM5) consequently find 'young air' in the upper boundary layer, while other models find 'older air'. With a flux boundary condition, these differences would also appear. Therefore, we do not really get the point made by the reviewer: "*since the emissions cannot be stopped and must get out of the boundary layer daily*". The issue is that some models tend to mix the boundary layer slower than other models, and thus simulate different gradients of emitted species. Most clearly, this is observed for the SF6 and 222Rn comparisons, both having flux boundary conditions, but models have a different rate at which the emissions are distributed vertically. To summarise:

With fixed emissions, an identical amount of tracer is mixed differently over the domain, depending on the mixing characteristics of the model. In the AoA protocol, the amount of tracer in the model intentionally depends on mixing. After translation to time, these differences can then be interpreted as different atmospheric mixing time scales. The challenges that we face are: (1) indeed, there are no known tracers with a fixed concentration boundary. However, the protocol is designed to simulate tracers from which transport timescales can be derived, and this advantage is explored here. (2) depending on the forcing volume and location, large gradients in the artificial AoA tracers may occur. Coping with large gradients is a known challenge for advection schemes. We discuss this issue in the manuscript in Section 3.6: Mapping the tropopause. Here we compare the tropopause derived from tracer 'e90' to the tropopause derived from AoA tracer 'surface'. In general, reasonable agreement is found, but differences are larger for some models. One tentative result is that models with less gradient conserving advection schemes (e.g. ACTM, Lin and Rood) show larger differences than the more gradient conserving models (like TM5, Russel and Lerner slopes scheme). In the revised manuscript we will discuss the issues with the protocol more comprihensively.

- *Missing references*: Thanks for pointing these out. We will refer to them in the revised manuscript.
- *Are the monthly mean mixing ratios 3D? Are the hourly station data 1D (profiles) or just layer 1?*: Yes, the monthly mean output is 3D. Modellers were asked to sample the model at 247 stations at hourly time intervals. At 119 station locations, profiles of concentrations and meteorological data were also requested. We will add some more information in the revised manuscript.
- *P8 L14 ? do you mean "adapted" instead of adopted?*: we did intend to use adapted here. Will be corrected.
- *P11 L29 ? is "IH" defined earlier? might be better just spelling it out everywhere, after all you did not define VT for vertical transport?*: IH is used 29 times in the manuscript, while vertical transport is used only 8 times. IH is defined on page 2.
- *P12 L13 ? please give real numbers - -how many years older?*: This is quantified in figure 9. We will add the range 0.10 – 0.17 years.
- *S to N transport patterns*: This is an interesting issue. We speculate on page 22 the the asymmetry between Figures 3 and 4 is caused by a 'seasonal rectifier' effect, meaning that over NH landmasses the forcing is strongly coupled to the season. Due to wintertime stabilisation and strong mixing in the summer, the AoA metric shows a larger seasonal variation in the NH compared to the ocean-dominated SH. Indeed, we found much larger seasonal cycles in the time series that make up Figure 3,

compared to the time-series on which Figure 4 is based. How this relates to transport patters that transport the SH air masses to the NH (i.e. the coupling to monsoon circulations), remains to be explored and data to do so are available. We would like to restrict the current manuscript to a description of the first results.

- *The "oldest" tropospheric air noted here was also part of the Prather e90 paper, at least compare*: We will include a comparison in the revised manuscript.
- *Easter Island*: Typo will be corrected.
- *BL differences*: We are pretty sure that BL parameterisation differences are the cause. Instead of adding further analyses, we propose to remove "likely".
- *P15 L7 ? In addition to Diallo et al 2012, check the Prather 2008 PNAS paper that demonstrated the convergence of two schemes with increased vertical resolution.*: Thanks for pointing out this reference. We will include it in the revised manuscript.
- *P16 L7 ? Remember that the SF6 is flux-driven and may show different features than the AoA tracers.*: We will highlight this issue more strongly in the revised manuscript, also in response to the first point raised.
- *P18 L1 ? SF6 accumulation in the BL*: We have been careful here with the wording. We mention SF_6 accumulation over the source regions over land when refer to the top panel of Figure 8, in which a clear enhancement over the source regions is seen in the simulated SF6 fields.
- P19 ? figures 9 and 10 : With Figure 9 we try to quantify the interesting asymmetry between the tracers 'NHsurface' and 'SHsurface'. Close to the equator the results become very sensitive to the way the model is sampled at the surface. This was illustrated with Figure 1. When forcing a tracer at the surface, e.g. in the SH, only the lowest 100 m of the grid-cel is forced (Table 1). After forcing, advective processes may transport AoA tracers into the box before sampling is performed. Close to the equator, large gradients in the mixing ratio of these NHsurface and SHsurface AoA tracers are expected, leading to results that depend on the sampling strategy of the models. Concerning the TOMCAT results: Indeed TOMCAT simulates the correct latitudinal gradient over the clean pacific (Figure 8, middle panel). However, the point we try to make is that this is a combined effect of slow vertical mixing (Figure 8, upper panel) and slow IH transport (quantified through the composite AoA in Figure 10). In contrast, NIES is also not "ball-park" in Figure 10. This is attributed to a combination of limited vertical mixing and fast IH transport. In that respect, the comparison to HIPPO data presented in the Appendix is interesting. Shifts to modeled mixing ratios needed to match the HIPPO data are largest for NIES and TOMCAT, but opposite in sign.

- *Please spell out IH and BL etc in the conclusions*: Will be done.
- *The Conclusions are a bit too speculative*: We agree on the fact that more work is needed to separate the effects of advection, convection, mass-fixers etc. We will modify the conclusions accordingly. Concerning the rectifier effect, we will try to explain this better in the revised manuscript. The bottom line is that when mixing is fast in the NH, the surface is in relatively fast contact with the free troposphere. Since inter-hemispheric transport proceeds most efficiently in the free troposphere (as seen in Figures 3 and 4), this implies transport of 'young' air to the SH. In winter, the situation reverses. Averaged over 11 years this leads to an average AoA that is younger than a situation with small seasonal variation in vertical mixing (i.e. the SHsurface tracer). To illustrate this, we build a simple three box model, with the lowest box representing the 'forced' surface layer. This box is in contact with box two, representing the boundary layer. Box two, in turn, is connected to box three, representing the free atmosphere. For this simple example, boxes are taken of equal size. The mixing time between the surface and boundary layer box is taken as 1 day. The mixing time between the boundary layer box and the free atmospheric box is taken as one week. To represent the NH, these times are modulated with a seasonal variation of 50%. To represent the SH, a smaller modulation of 5% is chosen. We now conducted the AoA experiments and calculated (in arbitrary units) the AoA in the NH and SH simulations in the three boxes (see Figure 1, upper panel). Logically, seasonal variations are larger in the NH case. If we now compare the running averages in the lower panel, we notice a small rectifier effect: The mean of the upper atmosphere SH boxes is systematically older than the corresponding NH boxes. This rectifier effect is caused by a covariation of the mixing and forcing. For the revised manuscript, we plan to include a worked out example with more realistic values as an Appendix.

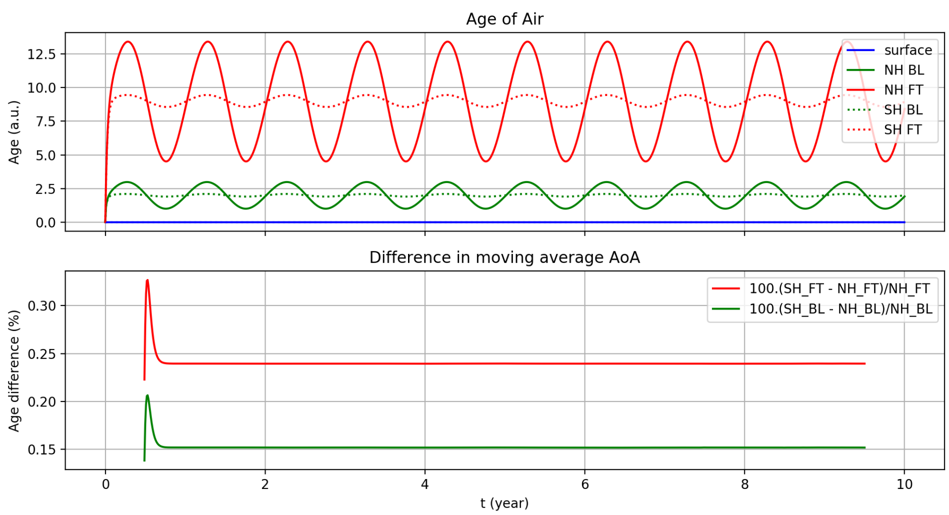

Figure 1: Results of a simple three-box AoA experiment in which the mixing between the boxes is modulated with a seasonal cycle. See main text for a description of the experiment.