

Some comments on A. Gettelman et al. "Coupled chemistry climate effects from 2050 projected aviation emissions"

(This is not meant as a full scale review but rather a collocation of ideas that occurred to me when reading through the paper. Nevertheless, I am confident that the comments express well-founded criticism.)

The present paper fails to discuss evidence from previous work to an extent that makes it difficult to understand what is actually added here to current knowledge on the subject. Besides other papers I am particularly referring to Huszar et al. (2013), a basic study of future aviation impacts that the authors appear to have overlooked.

- 1) The results given in section 3.1 all in all look scientifically and statistically plausible, yet they have apparently been presented before (Brasseur et al., 2015; Chen and Gettelman, 2016). However, I notice an inconsistency with respect to the contrail cirrus RF estimate for the 2006 baseline scenario between Table 2 (17 mW/m^2) and Figure 1A (12 mW/m^2). This is rather relevant, as the puzzling evidence that contrail cirrus RF increase more strongly over the years than fuel consumption would vanish, if the Table 2 value were taken as the starting value.

Recently, Forster et al. (2016) came up with a study indicating that radiative flux differences derived from free-running fixed SST simulations (I guess that's what "RESTOM" indicates in Figure 1) should amount to at least 100 mW/m^2 in order to reach sufficient statistical significance levels. In case of nudged simulations (resembling the specified dynamics simulations in the present papers) the threshold value may reduce to 10% (Forster et al., 2016, p. 13), which appears to be consistent with the error bars in Figure 1. However, the error bars are clearly overlapping between the different scenarios at all time slices, indicating that the scenarios are statistically indistinguishable.

- 2) I find the ozone pattern difference presentations from Fig. 2 a,b, Fig. 3 a,b, Fig. 5, Fig. 6, rather pointless. While they suggest large areas of statistical significance (for CESM almost over the whole troposphere), this remains unconvincing as contour lines are largely missing (those that *are* shown are mainly referring non-significant structures). Figure 8 of Huszar et al. (2013) offers a more satisfactory description, clearly indicating that patchy stratospheric response patterns are insignificant, despite showing higher concentration difference values compared to the troposphere. It may, hence, be worthwhile to display relative differences, as in many earlier papers (e.g., Grewe et al., 1999) dealing with free running chemistry climate model simulation results.
- 3) The severe problems to assess the (statistical and physical) significance of temperature response patterns simulated from aviation effects have been reported before (e.g., Rap et al., 2010, Fig. 1a; Huszar et al., 2013, Fig. 10, Fig. 12). A point-by-point hypothesis test suggesting statistical significance in strongly confined regions may well turn out to be unfounded, if spatial correlation is accounted for. (Chaotic negative and positive side-by-side differences, as obvious in Fig. 2 c,d, Fig. 3 c,d, are always raising suspicions in this respect. I notice coherent regions of significance only in Figs. 3d, 4d, and 6f shown here.) Significant temperature response is more easily established for global means (Huszar et al., 2013, Fig. 9), but these are bypassed in the present paper. Sometimes, more sophisticated (multivariate) statistical tests have proved helpful to establish pattern significance (e.g., Sausen et al., 1998).

- 4) In section 4.1.2 much text is devoted to allegedly large effects of aviation black carbon emissions without showing any results. To me this is absolutely unconvincing as to underpin what is suggested by Figure 1b.
- 5) Given the general lack of statistical significant simulation results, I find large parts of the concluding section to be insufficiently covered by the results. In my opinion, the simulation strategy followed in this paper is only of very limited value for establishing reliable evidence on the relative importance of individual components in forcing a net aviation climate impact. Even in Huszar et al. (2013) statistical noise has made interpretation of their results problematic and I fail to notice any progress on this in the present paper.

Adding on my main comments, I find the present paper to be written in a rather confusing manner. For example, the description of the simulations is scattered over three different sections (2.1, 2.2, 2.4) and it is not sufficiently recalled in the results section, which of the simulations are actually discussed. A special subsection (3.4.3) is dedicated to alternative fuel effects, yet those are partly addressed already in subsection 3.4.1.

References:

- Forster, P.M., et al., 2016: Recommendations for diagnosing effective radiative forcing from climate models for CMIP6, *J. Geophys. Res.* 121, 12460-12475.
- Grewe, V., et al., 1999: Impact of future subsonic aircraft NO_x emissions on the atmospheric composition, *Geophys. Res. Lett.* 26, 47-50.
- Huszar, et al., 2013: Modeling the present and future impact of aviation on climate: an AOGCM approach with online coupled chemistry, *Atmos. Chem. Phys.* 13, 10027-10048.
- Rap, A., et al., 2010: Estimating the climate impact of linear contrails using the UK Met Office climate model, *Geophys. Res. Lett.* 37, L20703.
- Sausen, R., et al., 1998: Climate impact of aircraft induced ozone change, *Geophys. Res. Lett.* 24, 1203-1206.