Review of "Climate responses to anthropogenic emissions of short-lived climate pollutants" by L. Baker et al.

This paper discusses the results from a set of idealized simulations performed by 4 different chemistry-climate models. The analysis focuses on a selected set of climate diagnostics, such as surface temperature, precipitations and run-off. While the overall set of results is interesting, there are several issues that the authors should address before the paper is suitable for publication

1) Normalized analysis has been shown to provide much more insight into the comparability of forcing-response relationships (Shindell and Faluvegi being a clear example). While it is clear that it is useful to provide the un-normalized numbers, the paper would greatly benefit from adding a description of the normalized results. For example, if ones takes the numbers from HadGEM/ECHAM/NorESM for BC, we see that the response in delta(T) is almost the same as the scale OC response. As it should. A table documenting the radiative forcing associated with each perturbation run should be included.
2) The figures only show stippling where models agree on the sign. That is a pretty low bar to pass (and I guess they still don't pass it). I would however provide estimates of the statistical significance based on the interannual variability. Similarly, zonal mean figures (5-7) are shown even for areas where models do not agree. What is the meaning of those figures in that case!?
3) The control experiment is much too short for the analysis that is being performed here, where the goal is to identify the response to a forcing much smaller than 2xCO2. As one can see for ECHAM, the global surface temperature is still trending at the end of the fifty years. Knowing that, it is necessary to show and discuss the trends in the climate state for the control experiment continued over the 50 years for which the perturbation is calculated. It would not be surprising if part of the "signal" was actually present in the control experiment as well. An approach might be to take into account the model drift over the 50 years.
4) There is no description of how the control experiments are performed. What is the level of GHGs? Where does the ocean initial state come from?

Minor comments

1) Author list: It is CICERO, not CISERO
2) Page 3825, line 5: There are much better and recent references to the impact of ozone on health and agriculture than HTAP. For example Tai et al., Nature Climate Change, 2014.
3) Page 3825, line 19-21: it might be useful to use the AR5 nomenclature (ACI-ARI)
4) Page 3826, line 23: there is a wealth of recent papers highlighting those connections. Please provide a better list of references.
5) Page 3831: it is not really clear what the added value of including CAM4 is. This version does not have indirect effects. Why bother? Why only BC?

6) Page 3833, line 5: ozone would be affected though, because of the methane impact.
7) Page 3833: this whole section need to include a documentation of the aerosol budget. Also, it needs to include a discussion of the differences in precipitation between the models. Finally, the aerosols should be compared to existing papers such as Samset et al., Observational evidence for overestimation of modeled black carbon radiative forcing. *Atmos. Chem. Phys.*, 14, 12465-12477, doi:10.5194/acp-14-12465-2014, 2014.
8) Page 3834, line 22: why are HadGEM and ECHAM similar when their SO4 burdens are so different?
9) Page 3836, line 2: include references discussing the shift in ITCZ
10) Page 3836, line 23-25: how do we know that this is the causal link?
11) Page 3836, line 26: the discussion of run-off might be much more useful if it is recast in terms of river flows, maybe for the largest rivers.
12) Page 3838, lines 19-21: how do the authors know that they are not simply looking at noise?
13) Page 3842, lines 5-9: it is not that clear that the ITCZ shift is related to climate sensitivity. Please substantiate!
14) Page 3842, lines 21-25: as the papers by Sand et al have indicated (among others) the location of the BC forcing is quite important.
15) Page 3843, lines 25-26: I would expect that concentrations are available from those simulations. Therefore a statement other than "probably have" should be made.