Interactive comment on “The Impact of Future Emission Policies on Tropospheric Ozone using a Parameterised Approach” by Steven Turnock et al.

Anonymous Referee #1

Received and published: 7 March 2018

Overall, I think this will be an important and useful paper detailing the impact of a great many future emission scenarios on surface ozone and radiative forcing. In general, the paper is well written with lots of detailed tables and clear figures. I would recommend publication following the changes detailed below.

Major Comments

1. Overall, I found the explanation for the parameterization rather confusing and not straightforward. Please think about how to make the explanation more precise. Some details... P4, L5: “the scale factor, f, is replaced by g”. This is a rather confusing way to put it. Why don’t you write out the full parameterization from the beginning by defining the various terms in equation (2) dependent on the constituent and not explain the parameterization by first defining f, then replacing f with g? Once could easily expand equation (2) to include the definitions for the various constituents. In addition, equation (2) is written as one might write out a computer code, but does not make sense from a mathematical viewpoint. Where does the factor (2f-g) come from. The factor g is evidently different for both CH4 and NOx from that given in equation 2? (equation 3). This should be discussed at the beginning and not mid-way through the paper. Where exactly is the ozone adjustment factor used (page 10)? Section 2.1 is titled “Original Ozone Parameterization”, but as far as I understand it is also the parameterization used in the present paper.

2. I think the paper could do a better job of emphasizing which results should be believed and which should be treated with skepticism. There are a number of emission scenarios where the emission change is over 50% (either with a positive or a negative change). To what extent should the results from these scenarios be believed? Results that should be treated with caution could be clearly indicated in the tables. The radiative forcing calculation does not seem particularly accurate. The authors claim that the parameterization reproduces the ACCMIP changes fairly well (p13, l23), but if I understand correctly the parameterized radiative forcing should be 20 to 30 MW m-2 larger than the ACCMIP results (as it does not account for the climate feedbacks which represent a large part of the ACCMIP signal). Thus, it looks like in most cases the future mean radiative forcing is dramatically underestimated (although, perhaps with the extremely large error bars in ACCMIP it is difficult to really say anything meaningful). Unless I missed it, the authors compared the change in the ozone burden between the parameterization and HadGEM2-ES but not the overall radiative forcing. Without some more evidence, and in a context that does not assume changes in climate, it is somewhat difficult to see what the parameterized radiative forcing calculation adds. I am willing to be convinced otherwise, but do need some convincing. The paper often makes somewhat vague statements about the comparison of the parameterization to explicit results (e.g., it states that the parameterization is valid, or compares well...). It would be nice to see in the conclusion a somewhat more explicit discussion of when and under what scenarios the parameterization should be believed: e.g., should it be...
believed under scenarios with large changes (e.g., +/- 50%), over regions where decreasing NOx increases ozone, over southeast Asia even with small emission changes, for the radiative forcing in 2100 given the strong climate influence etc? In other words, the certainty bounds should be discussed and quantified in more detail with an overall summary given in the conclusions.

Minor Comments:
Page 1:
L26 "are valid". This is really rather strong language as the accuracy of the parameterization differs depending on the region. It would be better to quantify this a bit more, saying instead something like "are reasonably accurate for most regions" or "are within the model spread for most regions". Once you say they are valid, it is difficult to quantify how valid are they?

L24. The neglect of climate change is mentioned in regards to radiative forcing but not to changes in surface ozone concentration. It would be important to emphasize that changes in climate and associated changes in climate dependent precursor emissions are neglected at the outset (e.g., in L24).

L32 "across different regions". This is a bit confusing. It might be better to say: "will regionally increase by 1 to 8 ppbv".

L33 I wander if it would be clearer to say "change in radiative forcing... from 2010 to 2050"?

Page 4
L31: which regions are not sources?
Page 5:
L3. "Models covered". Rather awkward English usage. Models don’t cover...

C3

L15-16. It would be worthwhile to add a line that the paper discusses in detail how the results TF-HTAP2 are incorporated below.

L25. It is not clear how the parameterization has been improved.

L24-34. In Table 1 the emission differences are averaged for both the MACCity and EDGAR inventories? It is not clear why the authors did not average the emission differences as used in the TF-HTAP2 models versus those in the TF-HTAP1 models. It is unclear how or why internal consistency (whatever that means) should be relevant here. The actual difference in emissions would seem to be more relevant in comparing the change in O3.

L35 - how many models contributed to the parameterized ozone response? -In general it is not really clear what was done here. Were the emission differences in table 1 used to compute the parameterized change in ozone in table 2? Was the parameterized response from each model computed separately to give the standard deviation in the parameterized response?

L37-38. "similar to the TF-HTAP2 multi-model mean values". This seems a little misleading as the parameterized responses are also similar to the TF-HTAP1 values. It would be more insightful to quantify the extent to which the parameterization quantifies the changes between TF-HTAP1 and TF-HTAP2.

Page 6
L5. "adjusted". I assume the authors explain how the emissions are adjusted below.

L11-21. Both the source and receptor regions are changed between HTAP1 and HTAP2. It is unclear from the description here how you discretize the response in the HTAP1 models into both smaller source regions and smaller receptor regions. The discussion in 2.3.2 seems to only concern the source region adjustment. Sections 2.3.2 and 2.3.3 should be clarified as to the exact procedure used.

Page 7
This seems to be largely a repeat of what is said above.

"significant improvements". In what way? Are more models are used, or are the source-receptor regions are better defined, or do the authors feel the parameterization itself has been improved in some fundamental way? The improved parameterization is again mentioned on page 9, line 19 (and probably elsewhere). Please be explicit on how exactly the parameterization is improved.

"is working well". This is a little hard to tell from the figures. It would be valuable to show the percentage error as a function of month for the two responses.

Page 8.

The results over South Asia are really quite strange as the parameterization based on HadGEM2 fits the multi-model parameterization. The authors seem to be arguing on page 8 and 9 that this is a difficult region to simulate and that perhaps it is not surprising that the parameterization based on HadGEM with the large titration might not be able to simulate this region accurately. However, this seems to be only half of the story. Why does the parameterization based specifically on HadGEM match the multi-model parameterization? And does the multi-model parameterization capture the multi-model response in this region?

Page 14

"compare well with ACCMIP multi-model means for intermediate emission scenarios..." What about RCP4.5? Isn’t this a intermediate emission scenario? The comparison from RCP4.5 does not look that good.