We thank all the reviewers for helpful comments that will improve our manuscript. Our responses are given in red.

Anonymous Referee #4

Received and published: 11 April 2017

General comments

The paper is interesting and in general well-written. As it is clearly described in the paper it builds on the work by Collins et al (2013), make use of a methodology developed in Shindell & Faluvegi (2009) and with data largely from Bellouin et al (2016). The study makes contributions regarding the estimation of Absolute Regional Temperature Potentials for NH3, the effect of aerosols on the ARTP for O3 precursors, extended analysis of the warming effect of BC on snow and ice and perhaps most importantly they analyze summer and winter specific metrics.

My key comment relates to the division of metrics into four latitude bands. I understand that the paper here builds on the framework by Shindell & Faluvegi (2009), but why is it 4 latitude bands, and not 6, 8 or any other number of latitude bands that is a relevant separation for the ARTPs? This needs to be discussed and problematized. More importantly, why is not a separation between temperatures impacts on land surfaces and ocean surfaces used? The land-ocean separation may be critical for regionalized metrics due to the significant land-ocean warming contrast (Joshi et al, 2008; Boer, 2011) and the very different climate impacts on land and ocean areas. I do not think the authors need to change their calculations, but a discussion regarding the relevance of the approach they take, what important aspects they miss with the regionalization they use and how the regionalization can be developed further is needed in the paper. Finally, overall I think the paper is a valuable contribution to the scientific literature and deserves to be published after the general comments above and the specific comments below have been taken into account.

The first order answer on why we use four latitude bands is that we base our calculations on the literature. The reviewer is right that it would be preferable to do our calculations on a more detailed level, but such a framework does not exist at the time. In response to reviewer 1, we have added one paragraph in Section 3.4 about what research we would like. Research on that would be highly welcome, but out of the scope for our study. Some of the reason why Shindell & Faluvegi (2009) did four latitude bands is the relative mixing time in the meridional direction versus the zonal direction. Shindell et al. (2010b) find that responses to inhomogeneous forcing extends roughly 3500 km or 30° in the meridional direction versus more than 10 000 km in the zonal direction. We have added to Section 2.2:

"Our study separates between four latitude response bands, in line with the typical width of response bands to inhomogeneous forcing found by Shindell et al. (2010), while more detailed modelling will be possible with a finer-masked RCS matrix available."

We have added a paragraph in the uncertainty section (3.4) on the land-ocean issue. This could potentially be done in future research, but should probably consider differences between the different species.

"The temperature response will vary by location, such as land surface versus ocean surface. These differences are not accounted for in our study, but the increased efficacy in the RCS matrix towards the

NH can be partly attributed to larger land area fraction in the NH (Shindell et al., 2015). The temperature increase is in general larger over land than ocean (Boer, 2011) driven by several local feedbacks (Joshi et al., 2008). We do not have data to break down this effect for our emission regions, but results in Shindell (2012) indicate that the land response may be 20 % larger than the average."

We have also added a section on what research is most needed to reduce uncertainty, as a response to reviewer 1.

Shindell, D., Schulz, M., Ming, Y., Takemura, T., Faluvegi, G., and Ramaswamy, V.: Spatial scales of climate response to inhomogeneous radiative forcing, Journal of Geophysical Research: Atmospheres, 115, D19110, 10.1029/2010JD014108, 2010.

Specific and technical comments

1. Page 1 line 31-32. The authors write "CH4 is often also included because its lifetime of around 10 years is shorter than or comparable to climate response timescales." I am not sure if I have seen this argument before. Please, justify with a reference why this is the reason why CH4 is included among the short lived climate forcers.

In the present literature, CH4 is sometimes included as an SLCF and sometimes as a well-mixed gas. Climate & Clean Air Coalition considers CH4 to be a SLCF. We are following IPCC AR5, Myhre et al. (2013). They write: "These compounds do not accumulate in the atmosphere at decadal to centennial time scales, and so their effect on climate is predominantly in the near term following their emission." In mitigation strategies, CH4 can be treated differently as CH4 has clearly a shorter lived impact on the climate than CO2. We have changed the sentence to:

"CH4 is included in the definition because its lifetime of around 10 years is shorter than timescales for stabilizing the climate (Aamaas et al., 2016).

2. Page 4 line 131-132. The authors write "We assume that the time evoluation of temperature in each response band follows the global mean temperature". Is this a valid assumption? For example, Cherubini et al (2016) do a related, but brief analysis using MAGICC, where they estimate regional metrics based on emissions that take place in either NH-ocean, NH-land, SH-ocean and SH-land, and where the NH and SH temperature response is analysed. If one contrast the assumption that the "time evaluation of temperature in each response band follows the global mean temperature" with results presented in Cherubini et al (2016) figure 3 the assumptions appears to be rather crude, especially on short time scales. The authors need in greater length justify this assumption.

Emission metrics are a simple tool, and we want to keep it simple. MAGICC can therefore not be introduced into the methodology. We have followed what is common practice in the literature, such as by Collins et al. (2013). The short answer why we assume the same temporal response pattern is that simple and well-tested parameterizations of such temperature ratios for different latitude bands or for land vs. ocean do not exist. While we agree this is a simplification, based on Figure 3 in Cherubini et al. (2016), this simplification only gives large uncertainties in the first 5-10 years after emission. Our main case is ARTP(20), hence, this simplification plays a minor role. We have added this after the sentence:

"Cherubini et al. (2016) show that this simplification is problematic for the first 5-10 years after emissions, but leads to less uncertainty after 20 years, which is our focus."

In the section on uncertainty, we have added several paragraphs on how the temperature response may vary, such as due to the land-ocean contrast, and what new research we would like the most.

3. Equation 2. In equation 2 the indices r,m,s, for the ARTP is dropped. Why? Please, be consistent throughout the equations or explain carefully difference between similar variables used in different equations.

All the indices have now been included.

4. Equation 3 and line 14-143. The authors write: "the general expression for the ARTP can be simplified to". Even though the mathematics behind this approximation is quite simple it should be shown and/or explained in a footnote, in the supplementary material or a reference to a paper where this is done should be included.

We add a reference to Appendix 2 in Fuglestvedt et al. (2010).

5. Page 5 line 152. The authors mention that the average adjustment time of CH_4 is 9.7 years in the three models used. This is relatively short compared to the IPCC AR5 assumption (12.4 years). Can the authors explain why a relatively short atmospheric adjustment time is find in the models used in the paper?

This is a good review question. Our manuscript is a follow up of the RF dataset in Bellouin et al. (2016) (see their Table 7). We have added a reference to this table in our manuscript. The adjustment time is calculated from $\tau tot*f$. As the adjustment time of CH_4 is discussed in Bellouin et al. (2016), we would like to keep the discussion limited in our paper. They found a large variability in the adjustment time between models, which is to be expected and within the model diversity seen in past studies. The low average adjustment time may be due to the selection of models, particularly the inclusion of HadGEM3 with a short adjustment time. We have added this sentence:

"If we use the adjustment time of 12.4 yr from Myhre et al. (2013), the ARTP values would be larger."

6. Page 5 lines 164-165. The authors write "RCS matrices only exist for annual emissions, we assume we can apply the same set of matrices for 165 emissions during NH summer and winter." Please justify this assumption.

This is a good comment, which at present cannot be quantified, as there are no climate model simulations available that has simulated RCS coefficients for seasonal emissions. However, we still believe that there is value added through this approach. The standard annual mean ARTPs quantify the relation between a unit pulse emission and an annual mean temperature response. Applied to a specific mitigation measure (e.g. improvement in wood burning stoves used for heating to reduce BC emissions) would give a seasonal cycle in the amount of mitigation (to a varying degree depending on source). In this case, both the emission → RF and RF → response (RCS) are implicitly assumed to follow the annual mean. In our approach, we resolve the emission → RF part on a seasonal basis, but we have to keep the assumption about the RF → response part. The simple answer why we used RCS for annual emissions is that there is no alternative in the literature. New research on this is highly welcome. This issue can open for a big discussion, which we hope future research will take on. We already state that there is no alternative, but will add these sentences:

"This assumption is a simplification, but is done implicitly when the annual mean RCS are applied to seasonal varying sources, e.g., wood burning heating stoves. We believe that calculating explicitly the RF from each season improve the overall ARTP values."

7. Page 8 lines 277-279. The authors write "For all the species, the response bands with the largest ARTP values are for the responses in the NH mid-latitudes (60% of the cases) and Arctic and the band with the least response the SH mid-high latitudes (see all panels in Fig. 1). This skewness is partly due to the emissions occurring mainly in the NH, but the same pattern is seen for CH4 (Figure 1(O)), for which the emission location is less important." The argument "This skewness is partly due to the emissions occurring mainly in the NH" is confusing. I first thought that the authors were referring to actual real world emissions, but that is totally irrelevant since you study equally sized emission pulses from different regions. Please write clearly what you mean with "This skewness is partly due to the emissions occurring mainly in the NH". There is also a similar argument in line 282 where the authors write "most emissions occurring in NH". Please clarify!

The RF we have used from Bellouin et al. (2016) are based on real-world emissions. The RFs have been normalized per unit emissions, so the reviewer is correct that we in some sense are comparing equal emission pulses. But for basically all the emission regions in this study, most of the emissions of that unit of emissions occur in the NH. The point is that the RF tend to be largest near the emission source, and those emission sources are in the NH in our study. We have reformulated to:

"). This skewness towards the NH is partly due to the emissions occurring in the NH for Europe and East Asia, as well as mainly for the global emissions…"

We have also clarified in line 282 by stating that the emissions occur in the NH "for the emission regions" we looked at.

8. Figure 4 and page 13 line 445-447. The authors write "The relative differences are generally larger for the aerosols than the ozone precursors, as seen in Fig. 4, where only the emissions regions and seasons with a relative difference larger than 20% are presented." Why is only cases where the relative difference between ARTP and AGTP are larger than 20% shown? Wouldn't it be equally relevant to see the cases where the difference is small?

For presentation purposes, we select a few cases. The total number is 70. We think the most interesting is where we find the largest differences and we therefore went for those with larger differences than 20%. Some more information is given in Section 7 in Supporting Information.

References

Joshi, M.M., Gregory, J.M., Webb, M.J. et al., (2008) Mechanisms for the land/sea warming contrast exhibited by simulations of climate change. Clim Dyn 30: 455. doi:10.1007/s00382-007-0306-1

Boer, G.J. (2011) The ratio of land to ocean temperature change under global warming. Clim Dyn 37: 2253. doi:10.1007/s00382-011-1112-3

Cherubini F., J. Fuglestvedt, T. Gasser, A. Reisinger, O. Cavalett, M.A.J. Huijbregts, D.J.A. Johansson, S.V. Jørgensen, M. Raugei, G. Schivley, A. Hammer Strømman, K. Tanaka, A. Levasseur (2016) Bridging the gap between impact assessment methods and climate science. Environmental Science & Policy 64: 129-140

For other references please see the paper by Aamaas et al, 2017