

## Response to Referee#3

We would like to thank Referee#3 for their extremely thoughtful and useful review. In this document, the referee's comments are repeated; our responses are in red.

### Comments from Referee #3

This is a useful and well written study of tropospheric ozone changes and associated radiative forcings resulting from an analysis using ACCMIP global model simulations for past present and future conditions. The study updates previous studies on the same theme- used in previous IPCC assessments. Obviously this study intends to inform the IPCC AR5 process.

I have little concern on the scientific credibility of the methods used in this paper, I have however some minor suggestions regarding sharpening some of the conclusions, and especially on the resulting uncertainties of tropospheric ozone radiative forcing.

Specifically in the abstract and conclusion I would like to see clearer statements on the uncertainties of our ability to calculate RF from ozone and its precursors, based on quantified and unquantified uncertainties. The paper is mentioning a +/- 1 sigma uncertainty of 30 %; this however implies that we have a sufficient confidence in the pre-industrial values of ozone. Unfortunately, current models systematically seem to overestimate the (few) pre-industrial measurements available. If the measurements are true- this would probably imply some missing process in the models that would lead to a greater RF and hence the uncertainty would at least double. I would like to see this issue somehow taken into account.

We have added discussion and caveats associated with pre-industrial ozone levels (also see response to Referee #1). We now present a comparison of modelled pre-industrial ozone with the available measurements. This does show that the models overestimate these measurements. However, these measurements are highly uncertain, so it is unclear whether this disagreement is a real problem or not. Pre-industrial methane levels are reasonably well-constrained, but concentrations and distributions of other ozone precursors are not, so it is difficult to ascertain if pre-industrial concentrations and distributions of NO<sub>x</sub>, CO and NMVOCs (and the NMVOC speciation: in the ACCMIP simulations, the NMVOC split is kept the same at all times) are accurately modelled. These distributions are mainly controlled by their pre-industrial emissions (e.g., Mickley et al., 2001). Our study has not investigated how uncertainties in these pre-industrial emissions translate into pre-industrial ozone levels, and hence pre-industrial to present day ozone RF, but it is clearly a major source of the overall uncertainty, and we acknowledge that. A further study would be required to quantify this component of the uncertainty in the ACCMIP simulations.

Compared to earlier works, does this mean a major improvement or is the answer essentially unchanged given the uncertainties?

The value and uncertainty for the tropospheric ozone radiative forcing (RF) is essentially unchanged. However, we believe this study is an important step forward, as the tropospheric O<sub>3</sub> RF now has a more robust scientific basis. That is because this study uses multiple new models, with new emission's estimates, and new RF calculations from updated radiative transfer models, all in a consistent inter-comparison framework. Significant uncertainties remain, but we think this study represents progress.

Finally with regard to RFs used in the prescribed RF experiments for IPCC (AMIP5), is there anything to say on the accuracy of these RFs with regard to the parts coming from ozone and CH<sub>4</sub>?

All the ACCMIP models that simulated ozone and also performed CMIP5 runs used their own ozone in those CMIP5 runs. Eight of those models are listed in Shindell et al (2012) (B: CICERO-OsloCTM2, F: GFDL-AM3, G: GISS-E2-R, H: GISS-E2-R-TOMAS, I: HadGEM2, K: LMDzORINCA, L: MIROC-CHEM, and N: NCAR-CAM3.5). These eight models are a fairly representative subset of all the ACCMIP models in terms of their O3 RF value (see Table 3): four are higher than the mean value (B, F, L, and N), by up to 49 mW m<sup>-2</sup>; four are lower than the mean value (G, H, I and K), by up to -71 mW m<sup>-2</sup>.

For methane, all CMIP5 runs (except LMDzORINCA) used observed methane in the past; LMDzORINCA used methane emissions. For the future only the GISS and LMDzORINCA models used their own simulated methane, based on RCP emissions. All other models used prescribed CH<sub>4</sub> concentrations from RCP (Meinshausen et al., 2011).

I therefore recommend publication of this paper when addressing the minor revision outlined above and below in the detailed comments.

Detailed comments:

p. 26049- line 5 mention that prescribed boundaries were used, most importantly common anthropogenic emission inventories- which are better constrained for the present day then for the past. The RCPs also give a limited view on the envelope of future air pollution emissions.

We have added.

p. 26049- line 20 I would for the sake of the abstract just talk about preindustrial numbers- explain elsewhere that this was derived by adding 0.04 Wm<sup>-2</sup> to 1850 numbers. It is a bit confusing right now.

We have adjusted the Abstract.

p. 26049 line 20 An important piece of information that should be in the abstract is the relative contribution of methane relative to other air pollutants- which will be very different in the scenarios. I propose to make two sentences -separating 2030 and 2100, and include in brackets the amount attributable to O3 from CH4.

We have separated out the 2030 and 2100 numbers. We cannot separate out the effects of CH4 in the future, as the necessary experiments have not been performed. We would only be able to make a qualitative statement with respect to the impact of future CH4 levels on O3, e.g., based on the past attribution experiments. We consider this is not appropriate for inclusion in the Abstract.

p. 26050: The forthcoming IPCC-AR5 WG1 report contains an updated evaluation of tropospheric ozone trends during the last decades- updated from the earlier HTAP (2010) report- based on work by Parrish and Cooper.

We have updated the Introduction.

p. 26051 l. 24 methodology used in/recommendations issued by IPCC: can you be more specific about the IPCC reference what is being assumed. Refer here to the later section which explains this in more detail.

We have added.

p. 26052l. 17 O. Wild did an attribution of O<sub>3</sub> to NO<sub>x</sub>, CO, VOC on the one hand and CH<sub>4</sub> on the other hand. Surprisingly that work showed a large uncertainty associated with O<sub>3</sub> from CH<sub>4</sub>.

We have added the reference to Wild et al (2012), and discuss the uncertainty associated with O<sub>3</sub> derived from CH<sub>4</sub>.

p. 26502 this is appr. 50 % higher; is somewhere in this paper discussed what this could mean for the IPCC scenario work?

(p26052) The RF values in the current paper supersede those in Cionni et al (2011). All of the O<sub>3</sub> RF calculations, in both papers, were offline, and should have no direct impact on IPCC scenario work. (We are unsure exactly what the referee means by IPCC scenario work.)

p. 26503 Model specific results in a supplement is necessary. Nevertheless, it would sometimes be good to discuss more specific model results, if they are outliers and determining the signal. Are these outliers for good reasons- or do the outliers point to model bugs- and should not be included in the paper.

(p26053) Some discussion of model outliers has been included in the results section.

p. 26504 Here or earlier an exact definition of RF is needed (perhaps again following IPCC: "The definition of RF from the TAR and earlier IPCC assessment reports is retained. Ramaswamy et al. (2001) define it as 'the change in net (down minus up) irradiance (solar plus longwave; in W m<sup>-2</sup>) at the tropopause after allowing for stratospheric temperatures to readjust to radiative equilibrium, but with surface and tropospheric temperatures and state held fixed at the unperturbed values"

(p26054) We have added the exact reference to the IPCC definition of RF as suggested.

Probably described also later, but it may be confusing to realize that each model's meteo is changing- while probably for the calculations a fixed p meteo was used.

Yes, each model has a different climate (meteorology), whereas the RF calculations use a single fixed climate/meteorology. This has been clarified.

p. 26504 l. 22: do you analyse what is the impact of this on ozone and resulting RF? Some of the attribution experiments may help.

(p26054) We did not isolate the impact of different models using different natural emissions, nor the effect that some models allow natural emissions to vary with climate. Changes in natural emissions may be responsible for a component of the pre-industrial (PI) to present-day (PD) RF. However, full analysis of this is beyond the scope of this paper.

p. 26505: Model K: how different was methane from observed values = and will the impact of this be evaluated?

(p26055) This model's simulations are analysed in more detail in Szopa et al. (2013), which we now refer to. The evolution of CH<sub>4</sub> concentrations is quite close to observed.

p. 26506 I think this 'climate' change effect is perhaps one of the newer aspects of these simulations. However, I didn't find it in the abstract-even if the conclusion is that it is not a major factor.

(p26056) We have added a sentence to the Abstract.

p. 26056 l 23: Here it said that meteo fields were constant? Meteo from one single model, or from each individual model. It is explained on p. 26057 that indeed one model was used: is there any sense of what uncertainty could arise from this (meteo not necessarily consistent with the individual models). As mentioned before, I think the concept should be introduced earlier.

This uncertainty is not fully explored, although by using three different offline RF schemes with different underlying meteorologies, it is partially. We think it is a relatively minor source of uncertainty, but this should be checked in future (one of the co-authors, J-F Lamarque, finds only small effects in a paper submitted elsewhere). Unfortunately it is outside the scope of this study.

p. 26058 A recent analysis of tropospheric ozone forcing using HTAP models and GFDL code was recently described in Fry et al (JGR, 2012). It would be good to have some qualitative evaluation of differences with that work as well.

A comparison with results from Fry et al (2012) has been included.

p. 26059 l. 7 To me the two different tropopause definitions give pretty similar results. There is a quite extensive discussion on the use of various other assumption, but I think the section should end with a bottom-line point of how the authors summarize the various effect (and not postpone to a later section).

We have added.

p. 26059 l. 15 This points to the fact that it is not so clear what has been used for pre-industrial stratospheric ozone (perhaps I missed it earlier).

We have attempted to clarify what changes in stratospheric O3 were used in these models with rather different behaviour over high southern latitudes. The information is given in Section 2.1. About half of the models (C, D, E, F, G, H, L, M, and N) employ detailed stratospheric chemistry schemes, and simulate the 1850s by driving these schemes with prescribed concentrations of ODS (ozone depleting substances) and other species. The models that most strongly show decreases in Southern Hemisphere (SH) ozone over 1850s-2000s (G and M) are in this group. From the individual model plots (Figures S1, S2, S3) in the Supplementary Material, it can be seen that these models have relatively high values for SH O3 in the 1850s, and this certainly contributes to the SH O3 decreases in these models.

p. 26060 clouds). We only use a single representation of cloud distributions (from the 64-level HadAM3 model) in the E-S calculations; cloud fields from individual models were not used: this information should be more upfront and the reason for doing so explained. Earlier you say that the uncertainty in clouds is being explored but I didn't find it-only that you did calculations with and without clouds and varying tropopause conditions, but that is not quite the same.

We have clarified this. We compare the effect of clouds/no clouds in the Edwards-Slingo and Oslo RF schemes (Table 4). These RF calculations used two different cloud distributions/properties. Uncertainties related to the representation of clouds were not explored beyond that. Exploring the impacts on O3 RF of using a variety of different cloud fields would be interesting, but is beyond the scope of this study.

The next few comments refer to the section on attribution of the O3 RF to different emissions. This section contained errors and has been updated and rewritten, see Author Comment #1.

p. 26062 l. 14 for spin-up times of 6 months or more, it probably was. Is that true for this analysis?

Yes – spin-ups were all >6 months. To clarify, it is all the species related to CH4 that will take time to come in to equilibrium in these runs, as the CH4 concentrations are held fixed.

p. 26062 l. 21: why do you consider use of single factor better than from the individual models. Models are known to have fairly different feedback factors. Would that induce an additional uncertainty?

See Author Comment #1: we now calculate 'f' values for each model and use these in the CH4 equilibrium calculations.

p. 26063 :However, the sum of the indirect effects on methane must sum to zero, but actually sum to -98mWm<sup>-2</sup>: I didn't get this statement why should this sum be zero? Can you elaborate?

See Author Comment #1 for a full explanation. The statement in the text was rather too sweeping and has been updated (although it was qualified in the following sentence). The sum of the indirect effects should sum to zero (in the absence of non-linear interactions between terms), as the net effect of NO<sub>x</sub>, CO, NMVOC and CH<sub>4</sub> emissions on CH<sub>4</sub> concentrations (after adjustment of CH<sub>4</sub> to equilibrium in each case) must result in the observed change in CH<sub>4</sub> concentration.

p. 26064: l. 4 Why is this? Because the direct effects of NO<sub>x</sub> on O<sub>3</sub> and OH are in absolute sense much higher? Where is the threshold for this method to be accurate?

See Author Comment #1 – we no longer make this argument in the revised text.

p. 26064:l. 5 is this referring to the -98 mW number?

See Author Comment #1 – we no longer make this argument in the revised text.

p. 26064: l. 12 these numbers can probably be compared to those at p. 26063 l. 21? On the global average the corrections seem to be relevant but not determinant .... I found the discussion above rather hard to follow- please try to further improve the description.

See Author Comment #1 – we no longer make this argument in the revised text. We hope that the revised description is easier to follow.

p. 26065 l. 15 This section would need a quantitative statement on the resulting RF or mention that a robust number could not be determined.

We have clarified (quantified) the impacts of climate change on O3 RF.

p. 26068 Model P: any explanation why results are so different?

Model P (UM-CAM) has a rather long CH<sub>4</sub> lifetime. We think this is related to this model's representation of photolysis, which is probably less accurate than other models.

p. p. 26068 I.14 Since Shindell is part of this study, and contributed to this paper, it seems not impossible to find out what is the difference. Given the importance of that earlier paper, I would expect a somewhat more quantitative attempt to explain this. Was there possibly a bug in the earlier study? In support of the current paper, it was indeed argued in a comment to the Shindell paper that additional studies should confirm the Nature paper. <http://www.nature.com/news/2009/091029/full/news.2009.1049.html> If indeed the current study downplays the importance of CH<sub>4</sub> in this conclusion would merit a more prominent place in the conclusions.

With our new analysis (see Author Comment #1), we make a more detailed comparison with Shindell et al (2005, 2009). Our results (the contributions towards the O<sub>3</sub> and CH<sub>4</sub> RFs from individual emissions) differ somewhat from Shindell et al's, and we make clear what those differences are.

p. 26069 On the role of climate change I would say the abstract should mention that no robust conclusion could be drawn?

We have adjusted the Abstract.

p. 26070 As introduced earlier- some discussion on what we don't know (e.g. the per-industrial to industrial change) and how that would change the conclusion would be helpful.

We have added some discussion to the Conclusions.

### References (in addition to those in the ACPD paper)

Fry, M. M., et al (2012) The influence of ozone precursor emissions from four world regions on tropospheric composition and radiative climate forcing, *J. Geophys. Res.* 117, D07306, doi:10.1029/2011JD017134

Mickley, L.J., D.J. Jacob, and D. Rind, 2001: Uncertainty in preindustrial abundance of tropospheric ozone: implications for radiative forcing calculations. *J. Geophys. Res.*, 106, 3389–3399.

Szopa, S. et al. (2013) Aerosol and ozone changes as forcing for climate evolution between 1850 and 2100, *Climate Dynamics*, in press

Wild, O., Fiore, A. M., Shindell, D. T., Doherty, R. M., Collins, W. J., Dentener, F. J., Schultz, M. G., Gong, S., MacKenzie, I. A., Zeng, G., Hess, P., Duncan, B. N., Bergmann, D. J., Szopa, S., Jonson, J. E., Keating, T. J., and Zuber, A.: Modelling future changes in surface ozone: a parameterized approach, *Atmos. Chem. Phys.*, 12, 2037-2054, doi:10.5194/acp-12-2037-2012, 2012.