Responses to referees for manuscript "*Analysis of present day and future OH and methane lifetime in the ACCMIP simulations*" by Voulgarakis et al. (2012).

Response to Referee #1

We thank the 1st referee for the useful comments. Below we respond to the reviewer's comments and describe the changes that we made, following this review.

Length of manuscript: We follow the advice of the referee and shorten the paper to some extent. We recognize that Fig. 3 and Table 5 from the submitted manuscript were not essential for the discussion, and remove them and the discussion related to them. Furthermore, we have now rephrased, shortened or removed some sentences throughout the text to be more to-the-point and more coherent.

More thorough analysis on factors driving present-day OH diversity, with an emphasis on OH production/loss rates: We believe that we have gone as far as possible in attempting to understand present-day model uncertainty, by performing the regression analysis presented in Sect. 6 (now Sect. 7). An even deeper analysis would require a) more sensitivity simulations, and b) more data from the models on the chemical fluxes. Neither of the above is available. However, we believe that by highlighting where the differences are and by discussing the drivers that may be the most important (now Sect. 7) we provide valuable information that will help the scientific community identify what needs to be addressed in future model intercomparisons.

Models which are outliers: The reviewer mentions the outliers and requests further explanations on why some models do not behave similarly to others. However, this is what Sect. 6 (now Sect. 7) does, and we feel that it goes as far as possible with discussing the model spread and its drivers. Specifically, for future trends, based on our regression analysis, it appears that responses to temperature changes are a key factor driving uncertainty. Two of the models, which ignore stratospheric ozone effects on photolysis, have a stronger response to climate change. This finding provides insight into a process that has not been discussed in a multi-model context in past studies. We also discuss that even though for the present-day, diversity in modelled OH and methane lifetime is more difficult to explain through a regression analysis than for the future, still, the treatment of photolysis and NMVOCs seem to be key.

Comment stating that "…it seems we did not converge since the TAR report, but added complexity in terms of future climate change": We recognize that this is the case if we include the whole range of ACCMIP models, but we do not necessarily feel that this is a negative aspect of the models' evolution. As models get more complicated, they also get more interesting, but they also inevitably include more uncertainties. Exploring these uncertainties is one of the major foci of this study. However, we also now provide alternative multi-model statistics following the exclusion of models that have a clear deficiency that will affect the simulations in a known way. We will come back to this modification further down.

Speculation over possible causes of the 2000-2100 changes in OH: For RCP8.5, the manuscript demonstrates rather conclusively that the main driver of OH and methane lifetime changes is methane itself. For the other RCPs, the trends are rather weak and mixed between models, so that it is difficult to discuss what drives the trends.

Some models not including the effect of stratospheric ozone on photolysis: This is true only for two models; the rest of them include this interaction, although they have not necessarily provided the photolysis rates output. Those that did not include the interactions are models that have extensively been run, at least up to now, for studying composition-climate topics under this configuration. Thus, it is fair to include them as a representation of this category of models. In addition, they provide valuable insight into underlying mechanisms (as the 2nd reviewer also notes), and when compared to other models, they help us understand the importance of stratospheric ozone depletion and recovery on tropospheric composition. As mentioned above, we now also provide alternative multi-model statistics following the exclusion of models that have a clear deficiency that will affect the simulations in a known way. We will come back to this further down.

Sensitivity simulations done by one model: We agree that it would be ideal to have all the sensitivity simulations performed by each individual model. However, the amount of work required for this would be enormous and is impractical in a study spanning so many research groups. We went further than what was requested by the ACCMIP protocol and performed some extra simulations with the GISS-E2-R model. We believe that these sensitivity simulations contribute to what the reviewer stresses in his comments: a more in depth analysis. Note that past multi-model studies (e.g. Koch et al., 2009) have also used a similar approach (using a single model for sensitivity experiments).

Also, it would be good to reduce the amount of models for some analysis: In some parts, we now provide alternative statistics following the exclusion of HadGEM2 and UM-CAM, since they have a well-known and important deficiency (non-interactive photolysis). We also exclude CICERO-OsloCTM2 for the alternative statistics of 2000-2100 changes, since this model did not include any changes in the meteorology. However, we do not exclude these models completely, as they are valuable for understanding model behaviour (as the 2nd reviewer also stresses) and the importance of specific processes, and also provide a sense of the robustness of the intercomparison statistics. The parts of the paper that have been modified are (using new figure/table order): Tables 1, 2, 4, S1, S2, S3; Figures 1, 2, 9, 10; parts of the text related to these tables/figures and the abstract.

Table 5: We agree with the reviewer that the interpretation of Table 5 is difficult, and we now remove it to reduce the length and improve the clarity of the paper. This also follows comment 34 of the 2nd reviewer.

Moving parts of Sect. 6 (now Sect. 7) to an earlier part of the paper: We see that there would be some logic from making this re-organisation, but we really would like to keep this discussion section separate. Since we believe it is one of the most valuable and novel parts of this manuscript, we would prefer it if this discussion is not scattered in different parts of the text. We now rename this section to "Discussing diversity in model results" to make it even clearer that this discussion section should stand on its own.

Comment on OH and NMVOCs in MOCAGE: In order to state something like this with certainty, we would need more chemical reaction fluxes from the models. However, as stated earlier, such datasets are not available. Thus, we can only speculate at this stage, but we feel that this discussion is valuable as it underlines points on which future model intercomparisons should follow (specifically for NMVOCs we stress that again in the conclusions section).

GISS-E2-R (and also LMDzORINCA) using free-running methane: It is true that the GISS-E2-R model has been run with a set-up that differs with the rest of the models, when it comes to methane emissions. However, the modellers that performed the GISS-E2-R simulations aimed for experimenting with a set-up that is unconventional, but may be more standard in future intercomparisons, as more models include interactive emissions. This set-up has provided valuable insight into the effects of methane feedbacks on its own lifetime, and how important it is including such feedbacks in future greenhouse gas scenarios. Repeating the simulations with different methane settings would be extremely time-consuming, since coupled ocean-atmosphere models (such as the version of GISS-E2-R used) take a long time to perform multi-decadal simulations.

For more information on interactive methane in GISS-E2-R and in LMDzORINCA we have already referenced Shindell et al. (2012b), and we now also reference Shindell et al. (2004) and Szopa et al. (2012) in the paragraph describing the models' methane treatment (Sect. 2.1). Burdens will be similar but not identical as a result of methane feedbacks on its own lifetime. We have added this caveat to section 2.1

The reviewer mentions the differences between NMVOC emissions in models in some parts of his/her comments: The differences arise from the fact that some models include interactive isoprene emissions while others do not, that in the rest of the cases modellers were free to choose their own natural emissions as no standard inventory was suggested, and also that some models emit CO as a proxy for VOCs rather than directly emitting VOCs. These features are mentioned/explained in several parts of the text. We agree that this is not ideal but this is not an issue only with ACCMIP, but with all composition-climate model inetercomparisons that have been undertaken so far (Stevenson et al., 2006; Shindell et al., 2006b; Fiore et al., 2009). Differences also arise from different NMVOC chemical mechanisms employed in the models (Lamarque et al., 2012a). For diversity in present-day natural emissions see Naik et al. (2012).

We add the following sentences in the Conclusions to stress this point:

"NMVOC emissions in ACCMIP were highly variable in the different models resulting from diversity in chemical mechanisms and the biogenic source implemented in the models. Future model intercomparisons would benefit from the availability of detailed NMVOC diagnostics, such as emissions from specific sources, and OH diagnostics such as production and loss fluxes for a better understanding of the model-to-model diversity in OH."

Minor Comments

Page 22948, line 9: Changed.

Page 22948, line 20: Changed.

Page 22948, line 24: Rephrased.

Page 22949, line 1: Done.

Page 22950, line 19: Changed.

Page 22955, line 5: Temperature change is in percent, as all the other variables on the tables.

Page 22955, line 26: The caption of Table 1 has now been changed.

Page 22956, line 19: Status of the reference has changed.

Page 22958, line 9: This figure has now been removed in response to comment 17 from the second reviewer.

Page 22958, line 21: Done.

Page 22960, line 4: Changed "reducing" to "increasing".

Page 22964, line 7: Rephrased to "There is a strong relationship between the CH_4+OH reaction rate constant (k) and temperature (see Appendix A), which implies …" to make more accurate.

Page 22965, line 4: We have now rephrased to "…due to CO_2-induced cooling of the stratosphere and enhanced circulation leading to a faster recovery of stratospheric ozone in this scenario" to be more precise". Also, we note that in the tropical lower stratosphere ozone is actually expected to decrease (e.g. Hegglin and Shepherd, 2009; Eyring et al., 2010, Lamarque 2010b), and so the column changes are fairly small. In higher latitudes changes are expected to mainly be positive (same references).

Response to Anonymous Referee #2

We thank the 2[nd] referee for the constructive and very useful comments. Below we respond to the reviewer's comments and describe the changes we made, following the reviewer's suggestions.

Major Comments

1. We agree with the reviewer on the structure of the paper. We follow the suggestion and move the discussion on regional changes after the analysis of the OH/methane lifetime drivers (after Sect. 5). This has led to the number of sections increasing by 1. We have made corresponding changes in section/figure numbering in the text.

2. We agree with the reviewer that some of the comments in the text are speculative. However, this is mostly in cases where we did not have further insight (e.g. due to the lack of necessary model diagnostics or suitable sensitivity experiments) so as to expand and provide more concrete evidence. Thus, in some of the cases we have not been able to avoid speculation in the revised version. However, there are statements that we now strengthen using further evidence, while when there are statements that are still speculative, we make sure that we change the phrasing (e.g. use "we speculate", "most likely" etc) to imply that what we state is not fully conclusive. Several of these improvements are listed below, in the responses to the "Minor Comments".

3. We have now made several changes in the text to tighten it up. This involves cases where repetition existed and where the phrasing could be made shorter. We also removed Fig. 3 and Table 5 of the original paper. Please see also response to the 1[st] referee.

4. Following the reviewer's comment, we now treat the models that clearly have well known and important deficiencies as special cases. This involves a) CICERO-OsloCTM2 for the future, which did not include changes in meteorology, b) UM-CAM and HadGEM2, which did not include interactive photolysis. Rather than excluding these models, we now report two different multi-model means with and without these models. More specifically, the parts of the manuscript that have been altered are the following (using the new figure/table numbering): Tables 1, 2, 4, S1, S2, S3; Figures 1, 2, 9, 10; parts of the text related to these tables/figures and the abstract.

Note that we have not excluded the GISS-E2-R model from the future analysis as its handling of methane, though different, is not known to be less good than in the rest of the models (as is the case with e.g. the photolysis handling in UM-CAM and HadGEM2). Indeed, we believe it helps with understanding the uncertainties associated with using fixed methane concentration and hence not simulating the effect of methane lifetime on itself.

Minor comments

1. We rephrased to “while water vapour abundances are largely determined by temperature changes”.

2. The table has now been cleaned-up, as the reviewer suggested. We state what methane emissions GISS-E2-R and LMDzORINCA use. Also, we now provide the global 2000 lightning emissions for all models. However, where modellers provided more than one reference for lightning emissions, we still keep it this way, as elements from all references may have been used.

3. We believe that the treatment of methane is outlined early enough in Sect. 2.1, which describes the different models used in the analysis. Before, we referred to Fig. 3 (was Fig. 7) a bit too late in the text, so now we also do that in Sect. 4.2, which starts with the sentence: “The future evolution of methane burden is shown in Fig. 3, while 2000-2100 changes can be seen in Tables 2, S1, S2 and S3.”

In most models, methane concentrations are specified at the lower boundary, but then methane is allowed to undergo loss in the rest of the atmosphere. We now specify this in a parenthesis in Sect. 2.1. How this “fixing” of methane affects OH, methane lifetime and methane itself cannot be fully discussed without sensitivity simulations with the same model using fixed and non-fixed surface concentrations. However, in Sect. 4.2 we already discuss this based on the different behaviour of the GISS-E2-R model. We have added the “(in which surface methane emissions rather than concentrations are prescribed)” part to make this clearer.

4. This is because, as we state in the caption, “Multi-model mean values (black dots connected with solid line) were only plotted for timeslices with data from at least 7 models”.

5. We mean either linked by using the same atmospheric component of a CMIP5 model, or by using SSTs/SI from a CMIP5 model. We have now rephrased to make clearer. Also we made the connections clearer in the parentheses. For the historical simulations, which are not the main focus of this analysis, we refer to the Naik et al. (2012) paper.

6. We changed to “time periods”.

7. The reviewer is right: anthropogenic and biomass burning NMVOC emissions are not identical, as models use a range of different oxidation mechanisms, with different lumpings applied. We add a sentence at this point to clarify this. We also add a sentence to Sect. 7 to state that “Generally, the reasons why NMVOC emissions are so diverse are that a) modellers were free to choose their own biogenic sources, and b) a wide range of NMVOC oxidation mechanisms were used in the models.”

A detailed discussion of the role of biogenic emissions is not possible, as these emissions have only been provided individually by a few models.

8. This is true only for global total emissions. There will still be interannual variability in the spatial distribution. We now clarify this in the sentence.

9. We now rephrase to "…led to those models being outliers in terms of global total lightning NO_x emissions (1.3 $Tg(N)yr^{-1}$ and 9.7 $Tg(N)yr^{-1}$, respectively)."

10. P22954, L26: We added "that are mainly" in the sentence to be more accurate. We also changed "present-day-to-future" to "present-day and future". "(no interannual variations)" was added to clarify that SSTs are annually repeating.

11. Several parts of the paper have now been modified as mentioned in response to Major Comment 4. Here, we also add the sentence "The range does not become…" to strengthen the point that our range is not smaller than in ACCENT.

12. Added "models".

13. Differences between the observational estimate and the model results is unlikely due to differences in the tropopause used, as Holmes et al. (2013) found that methane oxidation between 200hPa and the tropopause was less than 1.5% of total tropospheric oxidation by OH. We now clarify this in the text. Note that the brief discussion of the comparison to this observational estimate has now been moved to the first paragraph of this section.

14. Changed "most" to "7", and added "(-4.5%)".

15. Rephrased to "reveals the largest negative change in global methane lifetime levels relative to 2000".

16. Added "most likely" in two sentences of this section to clarify that these statements are not fully conclusive.

17. We have now removed Fig. 3 altogether (and the paragraph linked to it) as we agree with the referee that it is not essential for the discussion.

18. We agree that there certainly are some imperfections in the modelling set-up and how consistent the individual model experiments were. As discussed above, we have already provided alternative statistics excluding models that have clear deficiencies, such as lack of interactive photolysis and lack of meteorological changes. We feel that applying more alternative statistics would complicate the content of the discussion too much, and distract the reader from the main messages.

19. This paragraph has now been removed.

20. Added "most likely".

21. We have now altered or removed several sentences in these two paragraphs to make them more tight and coherent.

22. Changed to “most likely occurring”.

23. To avoid misinterpretation, we replace the sentence starting “This trend could…” with the following:

“The response of the convective updraft mass flux to climate change will likely vary with height and depend on the particular convective parameterization, so trends in lightning NO_x derived from convective mass flux fields will vary with the details of the implementation.”

24. The only type of methane emissions in the GISS-E2-R model that is distinct from the AR5 recommendations is from wetlands. Shindell et al. (2012b), for RCP8.5, show that there is an increase in wetland emissions, but far smaller than the increase in anthropogenic emissions that follow AR5. Thus, the fact that the GISS-E2-R model shows a stronger trend than other models in this scenario cannot be due to its different handling of wetland emissions. We refer to the Shindell et al. (2012b) article, which clarifies this.

25. The statement has been removed.

26. The clause has been removed.

27. As suggested, we have included mean and standard deviation for RCP2.6 and RCP8.5. We avoid doing this for all scenarios, since the differences are small in RCP4.5/6.0 and we do not want to overwhelm the text with numbers.

28. We have added “possibly” to make it clearer that we have not demonstrated this, since we do not have the available model outputs due to substantial differences in the ozone burden term reporting that precluded consistent results for STE as the closure term of the ozone budget. Increased STE is also discussed in Young et al. (2012), which we now cite at this point in the text.

29. Added the word “partly” to indicate that there may be other reasons why GISS-E2-R responds more strongly.

30. For the first point that is mentioned (methane being the major driver in RCP8.5), the quantitative evidence is given in the next subsection, via a sensitivity experiment. This is clearly stated in the last sentence of Sect. 5.1. For the other cases (non-RCP8.5 scenarios and NOx/CO/NMVOC emissions), we believe that we would not get meaningful results from a regression analysis, as the trends are so weak and/or mixed.

Also, since both reviewers mentioned that the paper is already too long or that it needs tightening, we would prefer to avoid adding more figures that are not essential. Ideally, we would have sensitivity experiments perturbing one of the above emissions at a time, but there is no availability at the moment for future ACCMIP scenarios.

31. We agree that the table is useful and add it.

32. We have rephrased to "We conclude it is…" as the reviewer suggests.

33. We have now defined the response time in the text.

34. We agree with the referee that there are too many insignificant numbers on the table, so we now remove the table, as it is not essential for our discussion. We focus on the related figures, which show the cases where we find a significant relationship. See response to comment 36 for comments on the new version of the figures.

35. We did not conduct formal tests of the differences in significance, and have rephrased that sentence as follows: "There was a strong association with $J(O^1D)$ (p=0.03) and a moderate association with NMVOC (p=0.07), though the latter is based on results from fewer models".

36. Now in figures 9 and 10 (replacements of previous figures 10 and 11) we show the results both using all the available models, and following exclusion of models with well-known deficiencies (as done in other parts of the analysis, for this revision). This makes the relationships somewhat less significant for the present-day. For the future, the relationships with climate variables become even stronger, while stratospheric ozone column also emerges as a possible driving factor. We modify the text (including the abstract) to reflect these changes as well.

37. This part of the text no longer exists. This section has been modified following the exclusion of some models from the regression analysis (see related comments to both reviewers).

38. We agree with the point and now remove water vapour from the statistical analysis.

<u>References not in manuscript:</u>

Hegglin, M. I., and Shepherd, T. G.: Large climate-induced changes in ultraviolet index and stratosphere-to-troposphere ozone flux, Nature Geosci., 2, 687-691, 2009.

Holmes, C. D. et al.: Future methane, hydroxyl, and their uncertainties: key climate and emission parameters for future predictions, Atmos. Chem. Phys., 13, 285-302, doi:10.5194/acp-13-285-2013, 2013.

Koch, D., et al.: Evaluation of black carbon estimations in global aerosol models, Atmos. Chem. Phys., 9, 9001–9026, doi:10.5194/acp-10-7017-2010, 2009.

Lamarque, J.-F.: Impact of Changes in Climate and Halocarbons on Recent Lower Stratosphere Ozone and Temperature Trends, J. Climate, 23, 2599-2611, doi:10.1175/2010JCLI3179.1, 2010b.

Shindell, D.T. et al.: Impacts of climate change on methane emissions from wetlands, Geophys. Res. Lett., 31, L21202, doi 10.1029/2004GL021009, 2004.