

## Responses to reviewer 1

We thank the reviewer for his/her valuable comments. We answer point to point to the comments in blue:

The authors have undertaken several efforts to address the previous review comments and I think the paper is improved. Below I have provided some specific comments on the revised manuscript as well as a comment on their response, and some general comments.

### COMMENT ON AUTHORS' RESPONSE 1.

The authors were not sure about my comment on figure 6 in my previous review, repeated here: "If you are showing absolute numbers (are you sure you want to do that?), then it would be good to show comparison numbers from (e.g.) a reanalysis product. Climate models are biased for the global mean, so I am sure that they will be so for a smaller region." What I meant was that the physical climate world normally deals in anomalies, or changes in parameters from a given period. One reason to do this is that the models are biased for the absolute value, but they might get a trend that agrees with the observations. The authors are showing absolute values for temperature, humidity and rainfall, whereas differences to (say) the 1850 time slice value would be more standard practice. If the authors want to keep absolute values, my point was that it would be useful to evaluate the models against (say) a reanalysis product. Incidentally, given the biases in both the underlying climate and the chemistry, why report absolute ozone concentrations?

We agree with the reviewer that models are likely biased but the difficulty here to show anomalies is to use the adequate reference. In this paper, we focus on the period between 1990 and 2100 and if we use a reanalysis product of the contemporary period to calculate the anomaly, it will not affect the trend found from 2000 and 2100. In Figure 6, the box-whisker plot of 2000 has to be taken as a reference which gives mean values similar to the contemporary period. The trend has to be understood by taking the 2000 period as the reference. We added this point into Figure 6 caption.

2. For Figure 9 the authors are arguing that they should not apply a field significance test. I would strongly argue that it is best practice to do this. The data are presented as a map, and therefore just relying on local significance is insufficient. As mentioned before, I suggest consulting Wilks (2016, BAMS, doi: 10.1175/BAMS-D-15-00267.1).

We considered the field significance in our indication of significant as suggested by the reviewer. We use a field significance test (Benjamini and Hochberg, 1995; Wilks, 2006) that satisfied the false discovery rate (FDR) criterion with  $\alpha_{FDR} = 0.10$ . The FDR method was performed using p values

from local Student t-test that was computed for each grid points with 95% confidence level. We added this point in the new version of the paper (see P6, L5-L7) and change consequently Figure 9 and adapt the text accordingly.

## GENERAL COMMENTS

1. For consideration by the authors: The authors now cite Schnell et al. (2015) in the context of hourly ozone analysis, but I would still think that this present manuscript would be more widely useful if the hourly ozone results were analyzed. I appreciate that this is a substantial undertaking, and the manuscript could stand as it is with the monthly mean analysis, but I think it would make the manuscript more relevant.

The purpose of our paper is more focused on climate than air quality. For this main reason, we think the time scale of 1 month is sufficient for our analysis. We prefer to leave the manuscript as it is with the monthly mean analysis.

2. I'm sorry if this sounds like nitpicking, but I would re-encourage the authors to revisit the use of paragraphs. There are still paragraphs going over 20-30 lines (e.g., start of Sections 2.2, 3.1, 3.2 etc etc). New point, new paragraph!

Thank you for this remark. We separate the text from sections 2.2-4.5 into different paragraphs following the recommendation of the reviewer and revisit the text of all these sections. The corrections are highlighted in blue in the revised version of the paper.

## SPECIFIC COMMENTS (incl. minor corrections)

**P1, L8 (and elsewhere):** You say the "behavior of the annual cycle" is good, but I think this is (still) ambiguous. I would say something like "the shape correlates but the values are biased".

We replace the sentence "The ensemble mean of ACCMIP models simulates very well the behavior of the annual cycle of surface ozone."

by (P1, L7-L9)

"The shape of the annual cycle of surface ozone simulated by ACCMIP models is similar to the observations one but the model values are biased high."

**P3, L30:** UM-CAM and STOC-HadAM3 produce meteorology without interaction with the concentrations of radiatively active species calculated by the chemistry scheme (NOT "without interaction with climate")

We replace the sentence "which produce their own meteorology without any interaction with climate"

by (P3, L30-L31)

“which produce their own meteorological fields with no interaction with the concentrations of radiatively active species calculated by the chemistry scheme.”

**P4, L10:** "models, however" -> "models; however" or "models. However" (Bad style to use "however" to start a new clause in the middle of the sentence)

Done (P4, L12)

**P5, L29:** I'm not sure about the "IQR for outliers". A sample outside of the central 50% is not generally considered an "outlier". There are several definitions of course, but perhaps you could use something like Tukey's Fences (see here for an introduction: <https://en.wikipedia.org/wiki/Outlier>)

In fact we used the method of Tukey-s Fences with a coefficient of 1.5. Outliers are defined as any values <25th percentile value-1.5 IQR (interquartile range) or values >75th percentile value + 1.5 IQR. This is clarified now in the revised version (see P6, L2-L3).

**P7, L34:** "Moreover" -> "However" or "But" (since introducing a counterpoint)

Done (P8, L18)

**P13, L30:** Be clear upfront that this is the chemical budget, excluding horizontal and vertical transport.

We replace the sentence “we focus on the evolution of the ozone budget along the 21st century”

by (P15, L15)

“we focus on the evolution of four ozone budget terms (excluding horizontal and vertical transport) along the 21st century”

**P14, L7-8:** I must have missed this the first time, but there are no strong reasons for basically defining an emergent constraint on the chemical budget changes ("models that are closest to the observations are the ones with increasing chemical terms". You might consider which models include climate-dependent biogenic emissions, which will have a strong impact on future emissions.

We removed this sentence that does not bring any added value to our paper.

**Figure 5:** Good that there is a color bar, but needs more description of how it was applied. Perhaps put the data in quintiles?

We added in the caption (Fig. 5) more description of how color bar was applied: “The colors associated with each metric value were determined as follows: the values of each metric have been rescaled between 0 and 1 corresponding to the model that is close to and far from the observations, respectively. The interval [0;1] has been subdivided into 6 equal intervals,

each representing a different color. The value of each metric is given by the color of the interval to which the rescaled value belongs.”

**Figure 6** (see also comment above): rainfall is normally reported as mm/day (convertible from what the models output)

As suggested by the reviewer, we changed the unit of the precipitation. It is now in mm/day in Fig. 6.

## Responses to reviewer 2

We thank the reviewer for his/her valuable comments. We answer point to point to the comments in blue:

The authors generally address my comments. In particular, the authors improve the readability of the text; the English language of the text; and the conclusions. However, as the other referee notes, there is still room for improvement. I encourage the authors to address further these three points. The authors should also address the specific comments below.

### Specific comments

**P. 3 L. 13:** I suggest: “...to contribute to the Intergovernmental...”.

Done (P3, L13)

**P. 8 L. 24:** models -> model.

Done (P9, L10)

**P. 9 L. 9:** Style point: I suggest you do not start a sentence with an acronym.

Done (P9, L30)

**L. 20:** Perhaps provide more details as to the effects mentioned and their impact on model results.

To clarify this point, we added the sentence (P10, L12) “More specifically, the increased humidity causes an ozone destruction which leads to a decrease in surface ozone.”

**P. 12 L. 22:** enhance -> enhances.

Done (P13, L34)

**L. 24:** I suggest you remove “indeed”. Omit needless words.

Done

**P. 13 L. 2:** Could authors provide more details explaining this behaviour?

We provided details explaining this behavior in the conclusion of the same section (section 4.4) “For the RCP8.5, the future climate change associated with a net increase in CH<sub>4</sub> concentration offsets the benefit of the emission reductions. In particular, for 2030 and 2100, the surface ozone concentration remains constant even if the NO<sub>x</sub> emissions are decreasing.”

**L. 21:** Remind the reader in what ways is RCP8.5 atypical.

We replace the sentence

“The RCP8.5 is atypical and different from the other scenarios. The surface

ozone over the MB remains constant over the period 2000-2100 with a strong increase in temperature, specific humidity and CH<sub>4</sub> concentration, unlike the global tropospheric ozone, which should increase by 18% in 2100 (Young et al., 2013).”

by (P15, L3-L5):

“For the RCP8.5 scenario, the surface ozone over the MB remains constant over the period 2000-2100 with a strong increase in temperature, specific humidity and CH<sub>4</sub> concentration, unlike the global tropospheric ozone, which should increase by 18% in 2100 (Young et al., 2013).”

**P. 15 L. 7:** For style, I suggest that you use “first” as you use “second” later. Done (P17, L7)

**L. 20:** In what way is this change “non-significant”?

To clarify this point we change the sentence: “non-significant change” by “statistically non-significant changes” (P17, L20-L21).

**P. 33 Fig. 12:** Please indicate in the caption the period over which you calculate the future relative change.

In the caption of Fig 12, we added the sentence: “The future relative change was calculated over the periods 2027-2040 and 2085-2110 (see Table. 3).”