Response to reviews of "Evaluation of tropospheric ozone and ozone precursors in simulations from the HTAPII and CCMI model intercomparisons - a focus on the Indian Subcontinent".

We would like to thank the editor and the two reviewers for their time in handling and reviewing our manuscript. We reply below to each of the reviewers comments in turn.

Reviewer 2:

Hakim et al., present results for intercomparison of HTAPII and CCMI models on the evaluation of ozone levels over India sing different observations, where surface observations are scarce. It is an interesting study and identifies some key challenges of climate models in reproducing observed ozone levels over india. I have a number of questions and comments in addition to reviewer #1 before the manuscript can be accepted for publication in ACP.

We thank anonymous reviewer 2 for their time and insightful comments to help improve our paper. We reply to these comments below.

1) Are there no rural sites to be used in model evaluation?

We have used a systematic set of surface observations from the MAPAN network. Whilst there are a limited number of rural sites across the domain (see comment from reviewer 1) these are generally located in regions of very clean air (i.e. in the Himalaya region) or are not available to us. We are not sure if the reviewer feels that rural sites would improve the evaluation, we infer they do, but we think that this is key and the next step forward to understanding how to improve model representation of this area. See response to reviewer 1 above and the modified manuscript for further details.

2) Figure 1 caption: Please make clear that these are total (anthropogenic + natural emissions).

We have modified the Figure caption accordingly.

3) Line colors in Figure 2 are difficult to be attributed to the individual models, please consider changing color scale.

We thank the reviewer for pointing this out and have changed the line colours in Figure 2 for consistency with the other figures.

4) Figure 3. Change "Mean" to "MMM" in the figure to be consistent with the text

We have changed this to be consistent.

5) In Table 1 caption, also refer to Fig. 4 for monitoring site locations.

We have modified Table 1 to make the links clearer.

6) Are there any filtering for missing data in the calculation of monthly mean observations?

Inevitably there are always gaps in any data record. In our analysis we considered missing values to be "NA" values and these were omitted while taking the mean. The data gaps are pretty small and we have calculated these here for the reviewer. The percentage of days missing per station across the year analysed were:
Delhi – 5.4% (20 days)
Patiala – 8.4% (31 days)
Udaipur – 0% (0 days)

Jabalpur – 1.9% (7 days)
Pune – 3.8% (14 days)
Guwahati – 1.9% (7 days)
Chennai - 0% (0 days)

7) What is special about Chennai that leads to poor temporal model evaluation?

The model the grid box(es) that we used to evaluate the performance at Chennai were heavily influenced by having large ocean fractions. This was the only station we looked at with the models that was heavily influenced by the coast and coupled to that Chennai is affected by both summer and winter monsoons. This combination is likely responsible for the poor model temporal evolution and we have modified the text to make the case for further work looking at coastal sites in the region. The text now reads "Observations at Chennai peak in April and October, i.e. during pre and post summer monsoon season. Models show poor correlation with the seasonal cycle of ozone at Chennai. To some extent this might be affected by the model's ability to simulate summer monsoon (from the south-west) and winter monsoons (from the north-east) that affect Chennai. It would be worth comparing model simulations with ozone observational data at Mumbai on the west coast of India, which receives rainfall only during the summer to understand the role of the monsoon near these coastal sites and we suggest further analysis assesses the performance of the models at the coastal impacted locations specifically."

8) Among the monitoring sites, Delhi seems to have much higher NOx and CO values compared to the other sites. Values reaching to almost 200 ppb for NOx and 5 ppm for CO do suggest that this station is not a typical sub-urban station. Can the authors comment on this?

Please see our reply to Reviewer 1 Comment 2.

9) It is interesting that for CO, the two stations with poorest correlations have the lowest biases judged from Fig. 9. Can the authors comment on this?

We agree with the reviewer that it is interesting but also note that there is no reason why correlation should be related to bias. It is interesting and we have to consider whether or not it is related to the parameterised processes like chemistry (which arguably are more likely to be related to the correlation coefficients) or inputs (i.e.

emissions). We can't say at this stage which is cause and which is effect but we agree with the reviewer and note the interesting feature in the main paper. We have added the following text to the paper to make the point "Interestingly, the model simulations at Jabalpur and Hyderabad show lowest correlations with the observations in spite of having the lowest biases. This could point towards some important processes which the models are struggling to simulate but further work would be needed to clarify this."

10) What is the difference between AATOC and MTOC?
AATOC is annual average of total ozone column and MTOC is the spatial mean of AATOC over the region considered in this study. We have modified the text to make this clearer "In order to evaluate the model simulations and observations we first compare the mean total ozone column (MTOC), defined as the spatial mean of AATOC over the study domain."

11) What do the PC1 and PC2 components refer to in these analyses? Is PC1 the monsoon system?
The reviewer is correct in suggesting that the timing of PC1 reflects the monsoon system. We have made changes in the paper to make this clearer "Hence maximum variance in tropospheric ozone is explained by the monsoon over South Asia (i.e. PC1 reflects the monsoon)." However, PC2 is more complex and at present we don't have a strong physical argument for what it represents. More work is required to tackle this which is beyond the scope of the present manuscript.

12) Can the reason why EOF do not provide clear understanding in comparison to other studies be the temporal resolution of o3? Can i.e. looking at daily or sub daily resolutions give more answers (like in i.e. Solazzo et al., 2017)
We very much like the Solazzo et al (2017) study and would like to follow that approach up in further studies looking at the ground based observations if we are able to get sub hourly data (at present hourly data is the highest time frequency). The satellite data we analysed to generate the EOFs are monthly means. The orbiting of the satellite means that at present daily data is the highest time frequency available and we could look at daily slices but there are issues with sampling, due to clouds etc, which mean that the analysis is much more robust by looking at monthly mean data. However, if and when higher time resolution satellite data becomes available it would be incredibly useful to perform new analyses with these data to better understand the temporal dependence of O_3 both in models and reality.

13) Adding MMM in Figure 13. could be helpful to interpret results
We have modified the Figure 13 accordingly.