

## Reply to Reviewer's comments

### *General Comments*

*The manuscript presents OH reactivity measurements from two urban sites in China and compared the OH reactivity data to calculated and modelled reactivity determined from the individually measured, co-observed OH sinks. Ozone production efficiency (OPE) is calculated from measured and modelled reactivity and the authors conclude that missing OH reactivity can increase ozone production efficiency at both sites. Understanding total OH reactivity by considering the dominant species contributing to OH reactivity and identifying missing OH reactivity and how this influence ozone production in urban environments is important and a suitable subject for ACP. The conclusion that more aged air-masses have a higher % of missing reactivity is an important finding also. Unfortunately there are several major problems with the manuscript currently which mean that the results and interpretation of the results are over-shadowed: The technical concerns (already thoroughly covered by Reviewer 1) relating to the quality of the OH reactivity data from the CRM instrument under high NO<sub>x</sub> conditions where large corrections have been applied need to be addressed before final publication. Furthermore, a more comprehensive comparison between the observed reactivity and calculated and modelled reactivity should be included and discussed to strengthen the overall conclusions drawn. I struggled to evaluate much of the discussion and conclusions, largely due to the poor English, but also because data discussed in the text did not appear in the referred figures or table: the modelled reactivity is not included in Fig. 11 upper panel and the breakdown of modelled reactivity is duplicated in tables S5 and S6. The figure axes and figure and table captions are inadequate to understand the data presented and there are inconsistencies between the data presented in the figures and discussion provided in the text. I have made a number of recommendations below where further clarification is needed or where the discussion should be improved before final publication can be considered.*

**Response:** Appreciate the general comments on this manuscript. We accept the comments on poor English and modified to this version. Wish this time it's easier to understand.

*Line 30: '... by adding unmeasured oxygenated...' this suggests that the model was constrained to assumed concentrations of OVOCs, but I don't think this was the case so this sentence needs revising*

**Response:** Thanks for asking. We are constraining the box model with several measured carbonyls, such as acetaldehyde, acetone.

*Line 34: change '.. such as aldehydes..' to '.. such as unmeasured aldehydes..'*

**Response:** Appreciate the suggestions and accepted.

*Model description: Line 230: How were the VOC data inputted into the model given the 1 hour time resolution of these measurements and 5 min time resolution of the model?*

Response: Thanks for asking. For VOCs data, we were using the linear interpolation method to achieve the 5 min time resolution and then put the data into model.

*In section 4 the authors consider the contribution unmeasured primary and secondary VOCs may make to missing reactivity. To strengthen this discussion some commentary is needed on the sensitivity of modelled OH reactivity to some of the assumed model parameters: Line 251: Are there local sources, e. g. roads, which mean that unconstrained products are not in steady state? How different is modelled reactivity on day 1 vs day 3 spin up? Line 251-253: How sensitivity is modelled OH reactivity to the treatment of dry deposition in both locations? How was the changing boundary layer height treated in the model? Could this influence the diurnal profile of the modelled OH reactivity?*

Response: Thanks for asking. Firstly, we admit that the model we were using in this manuscript was really a simple one. The close primary emissions, surely would introduce large uncertainty on the model job. However, as we were using this observation-based model, we could not input the emission data. The way we dealt with this problem was trying to do the moving average to avoid the sudden increase or decrease of certain species. For different spin-up time test, we found modelled OH reactivity could be different within 10% between two situations. However, for the literature suggestion and the 24 hour deposition time we chose, we decided to choose 3 day spin-up. For the dry deposition, some OVOCs and secondary species were quite sensitive to the dry deposition choice. From literature (Lu et al, 2012; 2013), we chose 24 hour dry deposition, and the modelled OVOCs presented similar diurnal variations of observations, with a 15%-30% difference depending on species. For the boundary layer height, we determined to set a well-mixed boundary layer height of about 1 km. However, this could be a source of uncertainty due to the diurnal variation of boundary layer height. Moreover, from Lu et al (2013) and Tan et al (2016), we would know that there could be missing chemistry in the nocturnal boundary layers, which would introduce unconstrained species and reactions in model work. However, due to the lack in boundary layer height measurements, we decided to set a constant boundary layer height. But the advice above are all important and appreciate all the questions.

*Results: Line 317 and 319: Are these the peak OH concentrations at both sites? Given the photochemical age of the air masses wouldn't a mean OH concentration be most appropriate for this calculations? The authors should discuss briefly the sensitivity of the photochemical age to [OH] used.*

Response: Thank you very much for this suggestion. These are both the peak OH concentrations in diurnal variations. Yes, this is one mistake and for the mean OH

concentrations should be the most appropriate. Also in the latest manuscript, we include a short sentence concerning the influence from OH concentration on photochemical age calculation. Appreciate your help.

*Line 348: Please provide details on the data used to generate the pie-charts – is this the campaign average picture? How does this change in Heshan during the pollution episodes? It would be informative to include a time series of calculated OH reactivity, modelled OH reactivity and measured OH for the whole of the two campaign periods somewhere in the manuscript.*

Response: Thanks for asking. It is the campaign average. However, it only included the data when there were OH reactivity measurement results. The data used to generate the pie-charts were as follows. However, we thought with this pie-charts, we did not need to present the data as well. For the second question, we could include the figure the referee asked for Heshan campaign. However, in Beijing observation, due to the discontinuity for OH reactivity and VOCs measurements through these three weeks, the modeled reactivity would need many times of interpolation and would thus introduce a great uncertainty here. So we decide to focus on certain processes rather than the campaign average.

*L459: ‘..more significant role..’ give % contributions.*

Response: Thanks for asking. In latest version of manuscript, we rephrase this sentence as below: The OVOCs had also significant contribution, and measured OVOCs had a sharing of 10% in total reactivity in Beijing while 7% in Heshan. We think the comparison was influenced by many factors, so we give up this direct comparison. Appreciate for this question.

*L456-457: What was the level of the NO correction applied to the measurement data during morning rush-hour?*

Response: Thanks for asking. This is one important question. For some periods in morning rush hour, when NO mixing ratio over 20 ppbV was observed, the NO correction could be over 40 s<sup>-1</sup> for measured reactivity. While the absolute reactivity was about the same level. However, this correction was checked for different species and verified. So we think this results were valid.

*Discussion: L363: what is meant by relative reactivity?*

Response: The relative reactivity means here the ratio between VOCs reactivity and NMHCs mixing ratios. It's not a strict definition here. However, we did not bring out this phrase at first in the latest version of manuscript. Appreciate your question.

*Line 366: ‘..not very high..’ apart from Paris, Heshan VOC reactivity is highest. This*

*section needs to be revised to accurately reflect the data in Fig. 10.*

Response: Thanks for asking. We re-phrase this paragraph as below: The measured VOCs reactivity (obtained by subtracting inorganic reactivity from total OH reactivity), 11.2s<sup>-1</sup> in Beijing and 18.3s<sup>-1</sup> in Heshan (Fig 10), was actually not at high end comparing with the levels from literatures. Tokyo presented a similar level of VOCs reactivity (Yoshino et al., 2006) and Paris had an even higher level of VOCs reactivity which was obtained in wintertime (Dolgorouky et al., 2012). The measured NMHCs levels (obtained by adding all hydrocarbon mixing ratios together) were also not very high, with Beijing 2013 being around 20 ppbV and Heshan 2014 higher than 35 ppbV. The relative VOCs reactivity, defined by the ratio of the VOCs reactivity to the measured NMHCs levels, the values for both Beijing and Heshan were very high.

*Figure 10a: Why not change the x axis to calculated NMHC reactivity (s<sup>-1</sup>) ? This would then help to demonstrate the cause for this trend, i. e. a) that the type of measured NMHCs in Beijing are indeed more reactive with respect to OH than at other sites or b) missing reactivity is more significant in Beijing and Heshan vs other sites. The discussion provided in 4.1 should be revised once the figure is changed.*

Response: Thanks for asking. Your suggestion is quite good to clarify the importance of the missing reactivity. However, if we are using the calculated NMHCs reactivity as x axis, we would not know that the first part you're telling, that the difference in the known compositions. Because we would only know that NMHCs reactivity in Beijing and Heshan were quite high. However, what we would like to express as well is while the NMHCs mixing ratios were not very high in Beijing and Heshan, the NMHCs reactivity could be high in both places. However, in our figure, there remains the questions that how could we presented both problems in one figure.

*Figure 10b: The y axis label is missing. Also why is Paris not included in this plot?*

Response: Thanks for the suggestion. Accepted and revised.

*Line 400: It is unclear whether the NOAA 2005 dataset is from Beijing. Even it is, it doesn't seem reasonable to simply compare missing reactivity from 2013 with branched alkenes data from 2005. Could a common species, measured both in 2005 and 2013, which is strongly correlated to the branch alkene data be used to scale the 2013 branched alkene data? Why are there only 6pts in figure 11, lower panel?*

Response: Thanks for asking. Yes for sure, the NOAA 2005 dataset was from Beijing. For the calculation in this part, we were firstly trying to track the correlation between the branched-alkenes and missing reactivity. So we compare the results in both diurnal variations. However, as we said in the paper. Even the mixing ratios in 2005 was not enough to explain all the missing reactivity, not even to say the decreasing trends for VOCs species in Beijing since 2005. The reason we only got 6 points were mainly due

to we would compare the data between morning rush hours in the diurnal variations. Appreciate for the questions to clarify this part.

*Figure 11 upper panel: modelled reactivity needs to be added to this plot*

Response: Accepted and modified as suggested. Thanks for the advice.

*Lines 424-426: Key to ozone control strategies, the authors should discuss the primary species from which the modelled species derive.*

Response: Thanks for asking. Sorry for this paper, the research focuses mostly on whether the primary or secondary species contributed more for the missing reactivity. However, we also have the data to answer your questions. These species were formed mostly as the oxidation products of alkenes and aromatics. However, your suggestion is very important and key to our next step – ozone control strategies. The vehicular emissions and solvent use related to alkenes and aromatics in urban areas should be controlled strictly for ozone reduction.

*Table S5: Species names should be provided in full – what is ‘DCB’?*

Response: Thanks for advice. This is also the advice from other reviewers. We will give a full explanation of major species in the tables.

*Lines 444-445: the authors should also compare the calculated and modelled reactivity from 2006 and 2014 too, so the 50% higher measured reactivity in 2014 can be evaluated fully.*

Response: Thanks for the advice. However, VOCs measurements from 2006 was limited to offline canister samples and the species were fewer than 2014. We’re afraid this compare may be not comprehensive.

*Lines 452,453: ‘PAMS 56 hydrocarbons’ and ‘TO-15 OVOCs’ needs defining*

Response: Accepted. We made an introduction to these two sets of standard gases, but due to the limitation of length, we would not supply the details of these species here.

*Lines 465-466; in section 4.3 the authors report that the OH reactivity modelled in Beijing agreed with the measured reactivity in the daytime (lines 423-424), but on lines 465-466 report difference between measured and modelled reactivity in Beijing which changed OPE by 27%. These two statements are inconsistent with each other and as the modelled reactivity is missing from Figure 11 it is unclear which is correct.*

Response: Thanks for asking. This is an interesting question. We’ve got similar questions at first. However, after we double-checked, the results remained the same

(though after we take organic nitrate into consideration, the difference was 21%). I think the difference could result from reasons below: 1) though many of the daytime reactivity were the same for the episodes, still some points the measured and modelled reactivity were different, not even to say the significant difference in rush hours; 2) the difference in species composition in scenario 1 and 2. Different species would introduce different levels of ozone production efficiency. However, this remains a question we need to dig in. Appreciate the question.