

## Response to Referee #1:

Thanks very much for your comments, suggestions and recommendation with respect to improve our paper. The response to all your comments are listed below. There was an extensive discussion among the authors regarding how to revise the content, and this paper is subjected to a major revision including an update of all retrievals using new inputs (e.g.,  $S_a$  based on standard deviation of a dedicated WACCM run from 1980 to 2020), re-plot all figures, condense/reorganize the content and focus more on the scientific topics. Thus, the response is delayed, and we are sorry for this.

### General comments :

The authors have used a new FTIR dataset to infer tropospheric ozone seasonal evolution and photochemical production regime at Hefei in China. Comparisons of the new dataset with OMI observations, and the GEOS-Chem and WRF-chem model data have shown good agreements. Back trajectories analyses have been used to attribute the contribution regions, and seasonal variabilities, to the high ozone levels observed at Hefei. The chemical sensitivity to ozone production has been studied at Hefei by using proxies such as CO and HCOH.

Although the authors use a new dataset, the novelty of some results is hard to admit. For instance, it is presented the fact that tropospheric ozone column is higher in spring/summer as a key result, which is a known scientific idea (same for better agreements comparing smoothed profiles relative to unsmoothed profiles). A reorganization of the paper's structure is needed, with less focus on the know results and more thinking about what is the paper contribution to scientific progress. In addition, the goal of the comparisons of the new dataset with independent data (atmospheric models and satellite observations) is unclear, as well as the use of two different model (global and regional). The objectives of the paper should be clarified and listed in a concise way. The number of figures should be reduced to fit the main scientific results. Discussions about results, such as model and observations comparisons, are missing and would improve the scientific impact of the paper.

The thorough section describing the retrievals is well written and I would advise the authors to submit this paper to a more technical journal, such as Atmospheric

Measurement Technique, if not addressing these comments.

**Response:** This paper has been subjected to a major revision based on the comments from three referees. All your comments are appreciated and have been addressed in the revised version. Main changes/improvements are listed as follows:

1) We have updated all retrievals with new  $S_a$  deduced from standard deviation of a dedicated WACCM run from 1980 to 2020, which should be more close to actual natural variation compared to the previous version. This improvement doesn't change the results of this paper.

2) We have reorganized the paper's structure, with less focus on known results and more describing about what is scientifically new. The objectives of the paper are clarified and listed in a concise way. The number of figures is reduced to focus more on the main scientific results. We have condensed quite a lot the descriptions of site/instrument, retrieval, theoretical basis but added many discussions/explanations regarding the observed results and photochemical regime. The figures and descriptions that are useful for understanding this paper but not scientific new are now shifted to the supplement (e.g., previous figures 2 - 5).

3) After an extensive discussion among the authors, we deleted all paragraphs and figures regarding comparisons with the correlative data, i.e., OMI, GEOS-Chem and WRF-Chem data, due to the following reasons:

a) The scientific topic of our manuscript is the investigation of the ozone seasonal evolution, source and photochemical production regime in polluted eastern China. The main interesting message we would like to present is the application of the FTS tools to determine if the tropospheric  $O_3$  is produced by  $NO_x$  or VOC, and give a recommendation about what could be done to mitigate the high  $O_3$  levels. This can not only improve the understanding of regional photochemical  $O_3$  production regime, but also contributes to the evaluation of  $O_3$  pollution controls. In the revised version, we leads straightly to this recommendation. For things which are not important for the main message, especially the deviation or something which probably misleads a potential reader, are removed. Accordingly, we removed the comparison with the models and the satellite.

b) This topic regarding comparisons with the correlative data, i.e., OMI, GEOS-Chem and WRF-Chem data, is interesting, but it cannot be clarified clearly within a few sentences or paragraphs and is basically a separate paper. Considering that this paper is already very long (referee's comments), we keep the intention of investigating the ozone seasonal evolution, source and photochemical production regime and removed all comparison with the correlative data.

4) We have responded to all referees' comments point-by-point and revised the manuscript accordingly.

**Related change:** The changes/improvements listed above have been done in the revised paper.

Specific comments:

Concerning the structure of the paper, it needs to be reorganized with a shorter abstract focusing a key results, more detailed introduction about the proxies used to assess the chemical sensitivity to ozone production, more sub-sections and tables summarizing the results of the comparisons, less figures, and appropriate English language.

**Response:** We have reorganized the paper's structure, shortened the abstract to focus on a key results, and included more detailed introduction about the proxies used to assess the chemical sensitivity to ozone production. In addition, more sub-sections and tables are used, and the number of figures are reduced to focus on the main scientific results. The revised paper has been corrected by a copy-editing service to improve the language.

**Related change:** The changes/improvements listed above have been done in the revised paper.

Consistency is also needed across the paper: define once (NDACC has never been defined in the abstract but appears in the keyword section, same for HCOH and VOC in the abstract, ...) and use them along the manuscript (ozone or O<sub>3</sub>?). Change old references with newest and avoid Wikipedia as a reference.

**Response:** All acronyms are now defined when they are first mentioned and also used consistently along the manuscript. Most old references are replaced with the newest

ones and the Wikipedia reference is removed.

**Related change:** All these problem have been addressed in the revised paper.

In the introduction section, more explanations about why using proxies such as CO and HCOH would help the reader understanding how to assess the chemical sensitivity to ozone production.

**Response:** We have added more detailed introduction about the proxies used to assess the chemical sensitivity to ozone production in the introduction section, which would help the reader understanding how to assess the chemical sensitivity to ozone production.

**Related change:** Detailed introduction about the proxies used to assess the chemical sensitivity to ozone production have been included in the revised paper.

In the retrieval strategy section, you mention a meteorological station onsite. Do you correct the NCEP profiles with these data? If yes, it should be clarified.

**Response:** As done at the other FTIR sites of the network, we did not correct the NCEP profiles with these data because this step normally makes the a priori profile (pressure, temperature) inconsistent. The pressure/temperature profiles have to obey some rules and this is fulfilled in the model data. The correction is also not that crucial, because the layers chosen depend only to a small extent on the temperature. When creating HDF files for the NDACC database, people usually have a field for surface temperature. But it is optional.

**Related change:** None

Define the use of the coincidence criteria when comparing to OMI (Why 3 hours and 0.7 ?) and other independent data.

**Response:** After an extensive discussion among the authors, we deleted all paragraphs and figures regarding comparisons with the correlative data, i.e., OMI, GEOS-Chem and WRF-Chem data. Now this problem doesn't exist in the revised version. Please check above clarification (page 2) for the reason.

**Related change:** Please check the revised version for details.

In section 5.1, it is mentioned a trend analysis on a 4-years time series. The word

“trend” needs to be changed.

**Response:** This has been done in the revised version.

Are back trajectories used to investigate the regions of influence of high tropospheric ozone at Hefei? If yes, it needs to be clarified and better structured in a sub-section. The end of section 5.1 needs to be better structured to emphasis on the scientific conclusions.

**Response:** The back trajectories are used to determine the origin of the air masses. This has been clarified and the previous section has been re-structured into two sub-sections.

**Related change:** This has been done in the revised version. Please check section 4 for details.

Concerning the comparisons with models, you may want to clarify their use; it is unclear if it is to assess the new dataset quality or investigate the model performances to reproduce observations. Explain the scientific interest of comparing the FTIR dataset with a global and a regional model. Discussions about results concerning comparisons between model/satellite and FTIR observations are missing and would raise the scientific level of the paper. Why is there a shift in the seasonal maximum in GEOS-Chem? Why do OMI and the FTS exhibit different seasonality? For WRF-Chem it is mentioned that the difference could be attributed to uncertainties in the input files, but what about the meteorological data, and/or the chemistry? This has to be further analyzed and explained.

**Response:** After an extensive discussion among the authors, we deleted all paragraphs and figures regarding comparisons with the correlative data, i.e., OMI, GEOS-Chem and WRF-Chem data. Now all these problems don't exist in the revised version. Please check above clarification (page 2) for the reason.

**Related change:** Please check the revised version for details.

Technical corrections :

- line 20 and 22 : define acronym CO, HCOH, and NO<sub>2</sub>

**Response:** We have defined these gases in the revised version.

- line 26 : by “the” FTS

- line27 : “occur” with no s

**Response:** This sentence has been removed when condensing the paper.

- line 32 and 33 : choose the precision, one or two decimal?

**Response:** Both are two decimal in the revised version.

- line 34 : by “atmospheric models” GEOS-Chem and WRF-Chem

**Response:** This sentence has been removed when condensing the paper.

- line 41-43 : rephrase the sentence

**Response:** We have rephrase it as “ Compared with SON/DJF season, the observed tropospheric O<sub>3</sub> levels in MAM/JJA are mainly influenced by transport of air masses from densely populated and industrialized areas while the broad and high O<sub>3</sub> level and variability in MAM/JJA is determined by the photochemical O<sub>3</sub> production.” Please check abstract for details.

- line 45-50 : state that HCHO is a VOC and define VOC

**Response:** We state that HCHO is a VOC and define VOC in the revised version. Please check the second sentence in the abstract for details.

- key words : NDACC never defined in the abstract

**Response:** As far as we know, the key words part is not a mandatory part of ACP, and thus we have removed the key words part in the revised version. The definition for NDACC has been done in the main text (introduction).

- line 55 : add a reference

**Response:** This has been done in the revised version.

- line 56 : add a reference

**Response:** This has been done in the revised version.

- line 53-71 : references are old

**Response:** Some old references have been replaced by the references published recently.

- line 75-77 : why so many references? Are they all relevant? You may use the most relevant one.

**Response:** This paragraph focuses on descriptions of the NDACC network. In the revised version, I removed the whole paragraph since it doesn't have much

contributions to the main point of this paper. According, all references (if not referred in elsewhere) are also removed.

- line 78-79 : define all chemical species.

- line 84 : state the accuracy or use a reference

**Response:** The whole paragraph has been removed, see above.

- line 87 : avoid Wikipedia as a scientific reference

**Response:** This reference has been replaced by two scientific papers.

- line 88 : first time ozone is written O<sub>3</sub>. Be consistent across the manuscript

**Response:** In the revised version, all “ozone” are replaced by “O<sub>3</sub>” . Now it is consistent across the paper.

- line 91 : “PM<sub>2.5</sub>”

- line 93 : “the” FTS

- line 96 : “Most NDACC sites”

**Response:** These have been done in the revised version.

- line 99 : Is the Hefei site a NDACC site? It is not clear here

**Response:** Hefei has ran both NDACC and TCCON conventions for more than 4 years, but is still a candidate site rather than an official one because of certain data publicity policy by Chinese government, and not because of the data quality. We are in progress to become an official TCCON site and we believe it will be also possible to be an official NDACC in near future.

**Related change:** Most site/instrument descriptions are removed and two reference are cited here.

- line 103 : add a reference for OMI

**Response:** A reference has been included in the revised version.

- line 105 : “the” site description

- line 107 : clarify the sentence “ozone related gases”

- line 114 : reference to Figure 1b

- line 114-115 : rephrase

- line 117 : clarify why it is an important region

- line 118 : add a reference

- line 119-120 : rephrase
- line 123-125 : rephrase
- line 125 : change “demonstrated” to “showed”
- line 126 : “typical observation day in August”

**Response:** All above related sentence has been removed when condensing the paper. Most site/instrument descriptions can be found in our previous paper (Yuan et al.,2017; Wei et al., 2017).

- line 121 : change “the same as” to “similar to”
- line 139 : define MIR
- line 143 : “for O<sub>3</sub> measurements”

**Response:** These have been done in the revised version.

- line 144 : are you certain filters are used to avoid detector non-linearity? What about signal to noise ratios?

**Response:** Filters are used for both, avoid detector non-linearity and improve the signal to noise ratios. However, this sentence has been removed when condensing the paper.

- line 148 : clarify “ozone related gases”

**Response:** This has been changed to “FTS retrievals of O<sub>3</sub>, CO and HCHO” in the revised version.

- line 151 : how much is an adequate accuracy?
- line 152-155 : is this sentence misplaced ? If not explain why it “confirms”
- line 155 : delete “so”
- line 157 : delete “overpass”

**Response:** To avoid misunderstanding, this paragraph has been removed when condensing the paper. Accuracy estimation can be found in section 3.3. The whole section 3 is used to confirm tropospheric O<sub>3</sub>, CO and HCHO are robust in Hefei.

- line 161 : add a reference to SFIT4

**Response:** This has been done in the revised version.

- line 195 : use mathematical equations

**Response:** We have used mathematical equations in the revised version and shifted it

to supplement.

- line 230 : “Figures 4 and 5”

**Response:** This has been done in the revised version. The two figures have been shifted to the supplement, now it is Figures S4 and S5.

- line 252 : explain why two sets of models

- line 282 : how much is 0.7 in kilometers at Hefei?

- line 307 : add the GEOS-Chem general reference

- line 320 : what is the nearest grid in kilometer?

- line 345 : add a general reference for WRF-Chem

- line 347 : “Liu et al. (2016)”

- line 348 : “20 x 20 km”

- line 382-383 : add a reference

- line 395 : delete “global”

**Response:** After an extensive discussion among the authors, we deleted all paragraphs and figures regarding comparisons with the correlative data, i.e., OMI, GEOS-Chem and WRF-Chem data. Now all these problems don't exist in the revised version. Please check above clarification (page 2) for the reason.

**Related change:** Please check the revised version for details.

- line 430 : “air pollution”

- line 430-432 : rephrase the sentence

**Response:** This sentence has been replaced by many detailed explanations in the revised version. Please check section 4.2 for details.

- line 435 : state the percentage

**Response:** We have stated the percentage in the revised version. Please check section 4.2 for details.

- line 439 : “considering the fact”

**Response:** This sentence has been replaced by many detailed explanations in the revised version. Please check section 4.2 for details.

- line 441 : “air pollution”

**Response:** This has been done in the revised version. Please check section 4.2 for

details.

- line 454 : “Obvious”. Why it is obvious?

**Response:** After an extensive discussion among the authors, we deleted all paragraphs and figures regarding comparisons with the correlative data, i.e., OMI, GEOS-Chem and WRF-Chem data. Now this problem doesn't exist in the revised version. Please check above clarification (page 2) for the reason.

- line 465 : “not an emission pollutant” is not clear, rephrase

- line 466-467 : explain why the fact that it is complicated means that it shows regional representativeness?

**Response:** This sentence has been removed when condensing the paper.

- line 479 : “as a result”, explain further the link between the two sentences

**Response:** Many explanations have been included in the revised version. Please check section 5.1 for details.

- line 485 : stay at present

**Response:** This has been done in the revised version. Please check section 5.2 for details.

- line 497 : how much are a good and an adequate correlation?

**Response:** In previous version, we regard it as good correlation if the correlation is higher than 0.6, and regard it as moderate correlation if the correlation lies in between 0.4 and 0.6. However, in the revised version, we only present the numbers and don't use the description such as “good” or “moderate” or “poor” to avoid controversy.

- line 502 : “has taken”

**Response:** This sentence has been changed to “Sillman (1995a) and Tonnesen and Dennis (2000) found that in situ measurements of the HCHO/NO<sub>2</sub> ratio could be used to diagnose local photochemical regimes.” and shifted to introduction part.

- line 505-513 : this could go to the introduction section

**Response:** We have shifted these sentence to introduction section.

- line 525 : change “obtainment”

**Response:** Has been changed to “the measurement tool for HCHO in this study was not the same as that of...”

- line 554 : change “validate” since OMI, GEOS-Chem, and WRF-Chem, to my knowledge, have already been validated

**Response:** After an extensive discussion among the authors, we deleted all paragraphs and figures regarding comparisons with the correlative data, i.e., OMI, GEOS-Chem and WRF-Chem data. Now this problem doesn't exist in the revised version. Please check above clarification (page 2) for the reason.

- Figure 1 a : change to see star colors

- Figure 1 b : instead of showing SZA for 1 day, you may want to show the mean SZA for all the data involved in this study

- Figure 1 legend : is it relevant to point out the wetlands? Are the red hexagons SZA or azimuth angle?

**Response:** In order to present the objectives of this paper in a concise way, the content has been shortened quite a lot. We removed this figure in the revised version. Detailed site/instrument descriptions can be found in our previous paper (Yuan et al., 2017; Wei et al., 2017).

- Figure 2 : cut altitude at 60 or 80 km

**Response:** This has been done in the revised version and already shifted to supplement. Please check the caption of figure S2 for details.

- Figure 3 : arrange the figure so that the text is readable and is not crossed by the lines. Figure 3 partial column averaging kernel of HCOH: explain what are the influences on retrieved column of a partial Avk of 12.

**Response:** This has been done in the revised version and is shifted to supplement. For partial column averaging kernel of HCOH, we find a bug in our previous plotting script. In the revised version, we fixed this bug and now this problem doesn't exist. This bug has no influence on retrieval but on for PAVK plotting. Thus, every deduction is the same. Please check figure S3 for details.

- Figure 4 and 5 : cut at 60 or 80 km and combine them in one Figure

**Response:** Both have been cut at 60 km, but we did not combine them in one figure because there are so much error components, and the combination is a big mess. We have shifted them to the supplement, please check figures S4 and S5 for details.

- Figure 6, 8, and 10 : insert the number of points included in the comparison and insert the standard deviation of the mean

- Figure 7 and 9 a : insert error bars

- Figure 7 b : shift text

- Figure 11 b : why showing both biased and unbiased data?

**Response:** After an extensive discussion among the authors, we deleted all paragraphs and figures regarding comparisons with the correlative data, i.e., OMI, GEOS-Chem and WRF-Chem data. Now this problem doesn't exist in the revised version. Please check above clarification (page 2) for the reason.

- Figure 12 b : reduce y-axis scale

**Response:** This has been done in the revised version. Please check figure 1b for details.

- Figure 15 : maybe plot all the measurements involved instead of daily means?

**Response:** Now all measurements were included in the revised version. Please check figure 3 for details.

- Figure 16 : reduce the size of the dots. Do you consider error bars to fit the data?

**Response:** We have reduced the size of the dots and grouped them into different seasons in the revised version. The error bars were not included in the fit because the meteorological station data do not have uncertainties. We get the accuracy of each element from the user manual. Please check figure 4 for details.

- Figure 18 : (a) all 3 panels should fit in one page. (b) Do you account for uncertainties when fitting the data? (c)  $R = 919$  with 8 points, are you certain it is a robust statistic?

**Response:**

R(a): In the revise version, the 3 panels have been fitted in one page.

R(b): We account for both slope and correlation. Briefly, we iteratively altered the column HCHO/NO<sub>2</sub> ratio threshold and judged whether the sensitivities of tropospheric O<sub>3</sub> to HCHO or NO<sub>2</sub> changed abruptly. For example, in order to estimate the VOC-limited threshold, we first fitted tropospheric O<sub>3</sub> to HCHO that lies within column HCHO/NO<sub>2</sub> ratios < 2 (an empirical start point) to obtain the corresponding

correlation/slope, and then we decreased the threshold by 0.1 (an empirical step size) and repeated the fit, i.e., only fitted the data pairs with column HCHO/NO<sub>2</sub> ratios < 1.9. This has been repeated. Finally, we sorted out the transition ratio which shows an abrupt change in correlation/slope, and regarded this as the VOC-limited threshold. Similarly, the NO<sub>x</sub>-limited threshold was determined by iteratively increasing the column HCHO/NO<sub>2</sub> ratio threshold till the sensitivity of tropospheric O<sub>3</sub> to NO<sub>2</sub> changed abruptly.

R(c): The previous figure (R = 919 with 8 points) is only used to demonstrate that PO<sub>3</sub> is more sensitive to VOC within VOC-limited region. Actually, the transition occurs close to about 0.6. At the transition ratio, there are much more points than 8. In the revised version, a detailed description for obtaining the transition threshold is presented, this kind of subfigures (only used for examples) are all removed.

**Related change:** Please check figure 6 for details

- Table 1 : enlarge the first column to adjust the word “regularization”. O3668 with exponent and index

**Response:** This has been done in the revised version. Please check table 1 for details.