Response to Referee #3:

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.

1 Overall remarks

The paper reports on about three years of tropospheric ozone and formaldehyde measurements from a new FTIR instrument in Heifei, China. The data are compared to a number of correlative data, including tropospheric NO2 from the OMI satellite instrument, and results from chemical transport models.

The authors give a very long and detailed description of their instrument and retrieval technique. They then analyse their observations using the correlative data mentioned above. Overall, their results, such as annual cycle, correlations, and trajectory analyses are plausible. However, the authors tend to discount differences and poor correlations, and to ignore the very coarse altitude resolution of their tropospheric ozone data, which average over a very wide altitude range, and have relatively little sensitivity to the planetary boundary layer, where a substantial part of the smog related ozone photo-chemistry takes place. Largely I concur with the comments by the other two reviewers. The paper does not present major new insights. However, it is important to report on new instruments and on tropospheric chemistry findings in China. Therefore, and also considering that this is a special issue for the last Quadrennial Ozone Symposium, I recommend publication after a few major deficits have been addressed.

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 NOx 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.

Response to "the authors tend to discount differences and poor correlations, and to ignore the very coarse altitude resolution of their tropospheric ozone data, which average over a very wide altitude range, and have relatively little sensitivity to the planetary boundary layer, where a substantial part of the smog related ozone photo-chemistry takes place."

Briefly: Many scientists have proved that column technique (OMI, GOME, or airborne results) can be used to investigate PO_3 sensitivity (Martin et al. 2004a; Duncan et al. 2010; Choi et al., 2012; Witte et al., 2011; Jin and Holloway, 2015; Mahajan et al., 2015; Schroeder et al., 2017; Jin et al., 2017). The NO_2 used in this study is the same as most previous studies, the sensitivity/resolution of FTS O_3 is close to that of OMI (Liu et al., 2010), the FTS HCHO is verified to be robust in troposphere in view of future satellite validation (Vigouroux et al., 2018). Thus, column technique used in this study is reasonable. We do acknowledge the paper by Schroeder et.al. (2017) which was published during the preparation of the manuscript. Schroeder et.al. (2017) question the usability of the column technique to infer PO3 sensitivity. However, this manuscript does take into account much of the criticism mentioned by Schroeder et.el (2017): we calculated the transition thresholds with the measurements in Hefei rather than straightly applied the thresholds estimated by either previous studies. The FTIR measurements have a much smaller footprint than the satellite measurements. Also we concentrate on measurements recorded during midday, when the mixing layer has largely been dissolved. And furthermore, the measurements are more sensitive to the lower parts of the troposphere, which can be inferred from the normalized AVK's. This reason is simply, that the AVK's show the sensitivity to the column, but the column per altitude decreases with altitude.

In detail: Over polluted areas, both HCHO and tropospheric NO_2 have vertical

distributions that are heavily weighted toward the lower troposphere, indicating that tropospheric column measurements of these gases are fairly representative of near surface conditions. Many studies have taken advantage of these favorable vertical distributions to investigate surface emissions of NOx and VOCs from space (Boersma et al., 2009; Martin et al., 2004a; Millet et al., 2008; Streets et al., 2013). Martin et al. (2004a) and Duncan et al. (2010) used satellite measurements of column $HCHO/NO_2$ ratio to explore near-surface O_3 sensitivities from space and disclosed that this diagnosis of O_3 production rate (PO_3) is consistent with previous finding of surface photochemistry. Witte et al. (2011) used a similar technique to estimate changes in PO_3 to the strict emission control measures (ECMs) during Beijing Summer Olympic Games period in 2008. Recent papers have applied the findings of Duncan et al. (2010) to observe O_3 sensitivity in other parts of the world (Choi et al., 2012; Witte et al., 2011; Jin and Holloway, 2015; Mahajan et al., 2015; Schroeder et al., 2017; Jin et al., 2017).

Related change: Several references where the column technique (OMI, GOME, or airborne results) is used to investigate PO_3 sensitivity have been included in the revised version. Please check the introduction part for details.

2 Suggested Changes

The description of the FTIR technique and FTIR profile retrieval in lines 148 to 248, as well as the averaging kernel smoothing used for comparison (lines 249 to 279) is pretty much standard. This could all be omitted, or moved to an appendix. A short paragraph and a few references in the main text are enough.

Response: In the revised version, the previous lines 148 to 248 have been condensed dramatically. The previous figures 2 - 5 have been shifted to the supplement. Please check section 3 for details. We still keep some of them because this paper is the first time to present O_3, HCHO, and CO time series at Hefei site. We think a brief introduction regarding site/retrieval setting/error analysis is useful. The previous lines 249 to 279 are all removed when condensing this paper.

lines 339 to 342, lines 367 to 370: I think these simple attributions to "model input files" are not valid. The wide averaging kernels and low sensitivity of the FTIR

tropospheric ozone columns to boundary layer ozone, as well as the limited horizontal resolution of the model data could play a very large role here. Please reword or omit these parts.

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.

Appendix A: Basically this is textbook / Rogers (2000), right? So this could/ should be omitted.

Response: Appendix A is a textbook stuff but useful for understanding this paper. It has been shifted to the supplement in the revised paper.

Fig. 6: I am not sure how meaningful this comparison of ozone profiles is. Both have very poor altitude resolution, and profile shape is determined to a very large degree by a priori assumptions. Comparison with a real tropospheric ozone profile from ozone-sondes or lidar would be much more meaningful. Maybe drop this Figure and its discussion? Similar considerations apply to Figs. 8 and 10.

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.

In most respects, I concur with the detailed recommendations by the other two reviewers. However, after shortening, and addressing the major comments, I think this manuscript is publishable in ACP.

Response: We have reorganized the paper's structure, with less focus on known results and more describing about what is the paper contribution to scientific progress. We have responded to all referees' comments point-by-point and revised the content accordingly. Thanks very much for your recommendation.