Dear Editor Dr. Leiming Zhang,

Thank you very much for handling our manuscript. Please find below our itemized responses to the reviewers' comments and a marked-up manuscript. We have addressed all the comments raised by the reviewer and incorporated them in the revised manuscript.

Thank you for your consideration.

Sincerely,

Haolin Wang et al.

Reviewer #1

Comment [1-1]: This study applied an extensive dataset, including aircraft, ozonesonde and global CTMs to study the trends of global ozone changes in the troposphere. Furthermore, the authors performed sensitivity simulations from the CTM to analyze the attribution of global ozone burden changes. They found a consistent ozone burden changes from both the aircraft and ozonesonde observations from 1995 to 2017 which was also captured by the CTM with latest emission inventory. The study also compared the trends simulated by the ensemble CMIP6 models, and concluded that the higher ozone trends from the CMIP6 was potential caused by the overestimation of anthropogenic emissions used in the earlier versions of CEDS emission inventory. The manuscript is generally well-written and the presentation quality is very good as well. I have two concerns about the methods in which I want the authors to address before this manuscript accepted for publication.

Response [1-1]: We thank the reviewer for the positive comments on our study. Below, we provide a point-by-point response to the reviewer's comments and summarize the changes that have been made in the revised manuscript.

Comment [1-2]: As comparison with the multi CMIP6 models, we can see the GEOS-Chem model used here has a really coarse resolution. I wonder how this issue will affect the trend analysis especially for regions in China and India which are experiencing significant changes both in climate and emissions at urban cores.

Response [1-2]: We thank the reviewer for bringing this up. While we can principally apply the GEOS-Chem model at a finer global 2°×2.5° resolution, we choose the 4°× 5° resolution to reduce simulation time. Our simulation includes a 10-year spin-up run and four sets of 23-year run (1995-2017). The simulations were conducted consecutively without any break to make sure the modelled ozone trends are consistent. It takes four hours with 48 CPUs for a one-month GEOS-Chem standard simulation at 4°× 5° (http://wiki.seas.harvard.edu/geos-chem/index.php/GEOS-Chem_13.3.0#1-month_benchmarks), so a 33-year simulation costs 66 natural days to finish. Running the model at 2° × 2.5° would cost at least eight times as much computational time and resources, which is rather inapplicable.

We agree with the reviewer that coarse-grid simulations (4°× 5°) may limit the ability of the model to capture finer-scale ozone trends, in particular at near surface where ozone and its precursor has a short lifetime. Artificial mixing of surface ozone precursors in coarse model grids may lead to higher-than-actual ozone production efficiency and therefore positive ozone biases (Wild and Prather, 2006; Yu et al., 2016; Young et al, 2018), especially for grids covering urban cores in China and India where significant changes in emissions are occurring. In light of this we do not use surface observations for model evaluation, and have limited the discussion of surface ozone trends. The limitation of model resolution, however,

should be alleviated for ozone in the free troposphere, where ozone has longer chemical lifetime so that we expect ozone there is better mixed than at near surface (Petetin et al., 2016). Nevertheless, increasing model spatial and temporal resolution is still preferable in future modeling studies of long-term ozone trends.

We have clarified in the text: "We run the GEOS-Chem model at a horizontal resolution of 4° (latitude) × 5° (longitude), with 72 vertical layers extending from surface to 0.01 hPa. One-month model simulation at this resolution costs 4 hours with 48 CPUs (http://wiki.seas.harvard.edu/geos-chem/index.php/GEOS-Chem_13.3.0#1-month_benchmarks). Yielding 33-year (including 10-year spin-up simulation) global simulation of ozone trends thus require computation time of more than 60 natural days. As such we do not use a finer resolution of 2°× 2.5° that would otherwise cost at least eight times as much computational time and resources as in this study. This relatively coarse resolution of 4°× 5°may limit the ability of the model to capture finer-scale ozone trends, in particular at near surface where ozone and its precursor has a short lifetime. Artificial mixing of surface ozone precursors in coarse model grids may lead to higher-than-actual ozone production efficiency and therefore positive ozone biases which may further influence trend analyses (Wild and Prather, 2006; Yu et al., 2016; Young et al, 2018; Yin et al., 2021). The limitation of model resolution, however, should be alleviated for ozone in the free troposphere, where ozone has longer chemical lifetime and should be better mixed than at near surface (Petetin et al., 2016). In light of this we do not use surface ozone observations for model evaluation, and mainly focus the trend analyses on above 950 hPa."

Reference:

Petetin, H., Thouret, V., Athier, G., Blot, R., Boulanger, D., Cousin, J. M., Gaudel, A., Nédélec, P., and Cooper, O.: Diurnal cycle of ozone throughout the troposphere over Frankfurt as measured by MOZAIC-IAGOS commercial aircraft, Elem. Sci. Anth., 4, 10.12952/journal.elementa.000129, 2016.

Wild, O. and Prather, M. J.: Global tropospheric ozone modeling: Quantifying errors due to grid resolution, J. Geophys. Res., 111, D11305, https://doi.org/10.1029/2005jd006605, 2006.

Yin, H., Lu, X., Sun, Y., Li, K., Gao, M., Zheng, B., and Liu, C.: Unprecedented decline in summertime surface ozone over eastern China in 2020 comparably attributable to anthropogenic emission reductions and meteorology, Environ. Res. Lett., 16, 124069, 10.1088/1748-9326/ac3e22, 2021.

Yu, K., Jacob, D. J., Fisher, J. A., Kim, P. S., Marais, E. A., Miller, C. C., Travis, K. R., Zhu, L., Yantosca, R. M., Sulprizio, M. P., Cohen, R. C., Dibb, J. E., Fried, A., Mikoviny, T., Ryerson, T. B., Wennberg, P. O., and Wisthaler, A.: Sensitivity to grid resolution in the ability of a chemical transport model to simulate observed oxidant chemistry under high-isoprene conditions, Atmos. Chem. Phys., 16, 4369–4378, https://doi.org/10.5194/acp-16-4369-2016, 2016.

Young, P. J., Naik, V., Fiore, A. M., Gaudel, A., Guo, J., Lin, M. Y., Neu, J. L., Parrish, D. D., Rieder, H. E., Schnell, J. L., Tilmes, S., Wild, O., Zhang, L., Ziemke, J. R., Brandt, J., Delcloo, A., Doherty, R. M., Geels, C., Hegglin, M. I., Hu, L., Im, U., Kumar, R., Luhar, A., Murray, L., Plummer, D., Rodriguez, J., Saiz-Lopez, A., Schultz, M. G., Woodhouse, M. T., and Zeng, G.: Tropospheric Ozone Assessment Report: Assessment of global-scale model performance for global and regional ozone distributions, variability, and trends, Elem. Sci. Anth., 6, p. 10, https://doi.org/10.1525/elementa.265, 2018.

Comment [1-3]: When doing the attributing analysis, taking aircraft for example, the authors fixed all the other emissions at 1995 level, and then varying the aircraft emissions. Since the world is experiencing significant emission changes for both developed regions (emission decreasing) and developing regions (emission

increasing), so how this method will affect the contribution from the aircraft for the ozone production without considering the realistic emissions in specific years?

Response [1-3]: Thanks for your question. For estimating ozone trend from aircraft emissions, we conduct the simulation FixAircraft with aircraft emissions fixed at 1995 levels while other emissions are varied from 1995 to 2017. We then estimate the aircraft emission contributed ozone trend as the difference in trends between the BASE and FixAircraft simulations. In this method, we do have considered the aircraft ozone production with the realistic emissions from other sources in specific years. We have clarified in the text "In the third simulation (FixAircraft), we fix global aircraft emissions at 1995 levels, and use the difference in ozone trend between BASE and FixAircraft to estimate the contribution of aircraft emissions alone to ozone trends.".

Comment [1-4]: L55-56: rephrase this sentence.

Response [1-4]: We have modified the text as "Tropospheric ozone is produced chemically from anthropogenic and natural precursors, it is also transported from the stratosphere, and is removed by chemical loss and dry deposition."

Comment [1-5]: L57: I feel change to "The ozone lifetime ranges/spans from xxx to ..." reads better. Just a suggestion.

Response [1-5]: We have modified the text as "The ozone lifetime spans from hours in the polluted boundary layer to a few weeks in the free troposphere, sufficiently short that ozone distributions and trends are highly variable."

Comment [1-6]: L70-79: I would encourage the authors to summary the findings from the IPCC report, instead of citing the whole paragraph for their own paper.

Response [1-6]: Thanks for your suggestion. We have carefully evaluated and discussed this issue. Our co-author, Dr. Owen Cooper, was a contributing author to the tropospheric ozone assessment by IPCC. He and his colleagues were instructed to produce a very short statement that concisely and accurately summarized global tropospheric ozone trends. This was a difficult task given the regional variability of ozone trends around the world, and every word in the IPCC statement was scrutinized for accuracy. Given that IPCC has already produced an accurate and concise summary statement on global ozone trends, it's not possible for us to summarize the statement further as important information will be lost; in addition, we prefer not to change any of the words as they were carefully chosen for accuracy. For these reasons we suggest it's best to provide a direct quote from IPCC.

Comment [1-7]: L202: The units for the trend should be "Tg yr-1" or "Tg decade-1".

Response [1-7]: Corrected.

Comment [1-8]: L214: change to "(Zheng et al., 2018)"

Response [1-8]: Corrected.

Comment [1-9]: Fig. 1: Explain the color indications for low panel

Response [1-9]: Thanks for pointing this out. We have clarified in the text: "The lower panel shows the location of selected ozonesonde sites in 1995–2017 used in this study, grouped by six latitude bands with an interval of 30° as denoted by different colors."

Comment [1-10]: Fig. 2: Explain the meanings in the bracket in left panels. The unit for the right panel is not accurate. Also reading the right panels for the trends of NOx CO and NMVOC, the authors can use the unit of "Tg/decade".

Response [1-10]: Thanks for pointing this out. We have clarified in the figure, and added the following text in the caption of Fig.2 "The total global anthropogenic emission trends with *p*-value are shown in left panels."