The manuscript presents an important piece of work to evaluate tropospheric ozone trends for 30 years or longer from ozonesonde and surface measurement and from two chemical models. The authors have conducted a careful analysis on the trend calculation, compared the consistency and inconsistency with previous studies, uncovered the model disability to reproduce the ozone trends, and discussed the possible reasons. In general this work is well-motivated, the analyses are comprehensive and easy to follow. Using two global models for long-term ozone trend analyses is particularly appreciable. Interestingly at almost the same time another paper comes up in ACPD using the same model (GEOS-Chem) to evaluate and attribute global tropospheric ozone trends (Wang et al., 2022, ACPD, in review), and I am glad to see that the two papers review the same important topic, and consistently point to the positive ozone trends from observations and the challenge for model to reproduce these trends.

I recommend this work to be published in ACP after moderate revisions. Below are some questions and comments that may help the authors to further strengthen their analyses and improve the presentation.

1. **Terminology.** The authors may use term “background ozone” or “baseline ozone” in literature review and analyses (e.g. Line 97, Line 474, Line 202, ...). Please try to clarify the terms when they first appear in the text. The authors may have known that policy-relevant background ozone has a more rigorous definition in US as ozone concentration without national anthropogenic sources. This should not be the same as “rural” defined in TOAR. Is there a clear definition on “baseline ozone”?

2. **Trends from ozonesonde and their representativeness of free tropospheric ozone.** The robustness of trends derived from ozonesonde is an important topic being argued a lot. The authors have tried to make sure that they select ozonesonde sites with
frequent sampling and the trends are consistent with surface sites nearby. I appreciate all the efforts. There are also studies suggesting that more profiles per month are required for robust trend quantification at a single ozondesonde site, and aircraft observations from the IAGOS project may provide better quantification of tropospheric ozone trends (Gaudel et al., 2020; Chang et al. 2020, 2021). I wonder whether the authors can try to reconcile their analyses with these existing literatures. It would be great if the authors can also utilize the aircraft (IAGOS) observations for evaluating the ozonesonde trends but if not some discussions are also acceptable.

3. **GEOS-Chem sensitivity simulation.** I appreciate that the authors applied the UCX scheme in their GEOS-Chem simulation for a better presentation of stratospheric chemistry. I wonder when the authors fixed anthropogenic emissions in their sensitivity simulation, did ozone-depletion-species also be fixed in the model? Or in other words, whether stratospheric influences from GEOS-Chem should be analyzed from the "Meteorology" simulation or in the "Emission" simulation? This should be clarified both in Section 2.2.1 and in Section 5.2.

4. **GMI vs GEOS-Chem simulations.** The authors use two chemical models and find that the GMI model produces larger ozone trends than GC. I am curious that whether the difference in anthropogenic emission inventory (and trends) may contribute. Could the authors show the emission trends used in the two model?

5. Section 3.3. In Figure 5 we see a site in eastern China with negative surface ozone trend that is not consistent with other site in the East Asia. Is this negative trend spread across all seasons? Is there any previous report that explains the negative trend?

6. Section 3.4, Line 489-492: I didn’t get clear information from the analyses at WLG site. The authors argue that “expect similar trends in the Japan FT since increasing ozone over Japan is influenced by transport from China”. But WLG is located at the west of the eastern China (not influenced by polluted outflow there), while the Japanese sites may be affected by the outflow but should mostly be constrained in springtime, why we should expect similar ozone trends at both sites?

7. Line 497: where does “up to 1.7 ppbv per decade” come from?

8. Line 517: What does “primary influences” mean?

9. Section 3.5. I few that the title of “Drivers of observed ozone change” is misleading. I am not sure whether the change of 5th ozone alone can be used to indicate ozone drivers of emission or transport. Why not integrate this part with the model sensitivity studies? In particular, I am not convinced by the statement in Line 549 that “multiple previous analyses suggest that regional ozone increases are best explained by transport”. Transport
from where? Why transport is a more likely reason than NOx reduction to explain 5th ozone increase in US and Europe?

10. Line 566-567: Again, it would be important to compare the emission trends used in GMI and GC.

11. Section 5.1. Line 690-691. Wang et al. (2022, in review) shows that aircraft emissions may be an important source of trend underestimation in chemical models.

12. Section 5.2. I am curious about the stratospheric influence on tropospheric ozone. First, are both models show consistent pattern of ozone trends in the upper troposphere and lower stratosphere? A latitude-altitude plot of modelled ozone trends from the two models would be plausible. Second, I wonder what processes may contribute to increasing STE in the GMI model, is it more likely due to recover of ozone hole or changes in STE dynamics?

Reference:


