Reviewer #1

Comment [1-1]: This study explores how changing anthropogenic emissions have impacts the O3 sensitivity to temperature from 1990 to 2021 in the United States. The observations demonstrate a diminishing sensitivity, and the authors use GEOS-Chem model simulations to show this is due to decreasing anthropogenic NOx emissions, and that both the direct O3-T mechanisms and indirect (non-T meteorological) factors contribute to this. The study is a very nice example of using a model to interpret an observed result. The study is thorough and well-explained and was a pleasure to read. There are three topics that should be better addressed in the text prior to publication:

Response [1-1]: We thank the reviewer for the positive and valuable comments. All of them have been implemented in the revised manuscript. Please see our itemized responses below.

Comment [1-2]: A central result of this study is that the model can only reproduce less than half (42%) of the observed trend in the O_3-T relationship. The paper attributes the source of the trend in the model, but does not sufficiently discuss the possible reasons and implications of the model missing half of the trend. The authors suggest that biases in the MERRA-2 temperature dataset from 1995-1999 contributes to this, but they show that when excluding this part of the record, the model can still only capture 56% of the trend. On lines 304 & 309 they suggest that this is due to biases in the SWUS region, but neither the region nor the degree of bias seen in Figure 5a seem sufficient to explain the majority of the missing trend. If it is indeed due to the SWUS, the paper should (1) discuss why the model cannot capture this region and (2) show that the model can capture the trend without the SWUS and (3) proceed only with non-SWUS results. If the SWUS can only explain a small part of this model bias, then the paper should discuss what other factors could contribute to the model bias, how this could be further explored in future studies and/or how the model could be improved. The paper brings up the temperature impact on anthropogenic emissions on line 146 – how important might this factor be?

Response [1-2]: Thank you for pointing this out. We agree that the discussion of the underestimation in simulated $m_{\Delta O3\text{-}\Delta Tmax}$ trends should be much strengthened. We find that the model's ability to capture the $m_{\Delta O3\text{-}\Delta Tmax}$ trends differs across regions, and the timing of the bias is

also not consistent. For example, the model underestimates $m_{\Delta O3\text{-}\Delta Tmax}$ from 1995 to 1999 mainly in the SWUS and NEUS regions, while overestimations for 2013 to 2017 are seen in the SEUS and Midwest, suggesting that these biases stem from different causes. In the revision, we have analyzed the regional model bias in $m_{\Delta O3\text{-}\Delta Tmax}$ in detail, and discuss three main sources of the underestimation of $m_{\Delta O3\text{-}\Delta Tmax}$ trends: bias in MERRA-2 temperature data in early periods, overestimations in anthropogenic emissions inventories, and model capability in parameterizing ozone-temperature response. We have added the following discussion to Section 3.2:

"The model shows a mean $m_{\Delta O3\text{-}\Delta Tmax}$ trend of -0.28 ppbv/K/decade over the CONUS that accounts for 42% of the observed trends of -0.67 ppbv/K/decade. Figure 5b also shows that the model's underestimation of $m_{\Delta O3\text{-}\Delta Tmax}$ trends is primarily attributed to an overestimation of $m_{\Delta O3\text{-}\Delta Tmax}$ from 2013 to 2017 and an underestimation from 1995 to 1999. The consistency between the observed and simulated $m_{\Delta O3\text{-}\Delta Tmax}$ trends also shows regional differences. As shown in Figure S8, the model reproduces the interannual variation of $m_{\Delta O3\text{-}\Delta Tmax}$ well in the Plains and Intermountain West regions, and also captures 65% of the observed trend in the NWUS. However, in other regions, the model only captures less than 50% of the observed $m_{\Delta O3\text{-}\Delta Tmax}$ trends, with either an overestimation in 2013-2017 or underestimation in 1995-1999.

Our GEOS-Chem simulation has successfully reproduced the observed long-term ozone trend averaged over the CONUS (-6.1 ppbv/decade in GEOS-Chem vs -6.5 ppbv/decade in observations) (Table S5). However, capturing the long-term trends in $m_{\Delta O3\text{-}\Delta Tmax}$ can be more challenging than that of ozone concentrations, as it involves the combined uncertainty in temperature data, simulated ozone concentrations, and the parameterization of ozone-temperature response. The underestimation of $m_{\Delta O3\text{-}\Delta Tmax}$ from 1995 to 1999 may be partly attributed to the larger bias in MERRA-2 temperature dataset compared to other periods (Figure S1), and such bias can propagate to the derivation of observed $m_{\Delta O3\text{-}\Delta Tmax}$ based on MERRA-2 dataset. Excluding the 1995, 1997, and 1999 records improve the model's ability in capturing observed $m_{\Delta O3\text{-}\Delta Tmax}$ trends in the CONUS (-0.46 ppbv/K/decade in GEOS-Chem vs -0.80 ppbv/K/decade, 58%). In particular, for the NEUS, Midwest, and SWUS, the model's ability to capture observed $m_{\Delta O3\text{-}\Delta Tmax}$ trends improves from 44%, 49%, and 23% to 83%, 66%, and 54%, respectively. The simulated ozone-temperature sensitivity for 2013–2017 shows an overestimation, particularly in the SEUS and Midwest regions (Figure S8). Christiansen et al. (2024) suggested that the CEDS inventory overestimates post-2010

anthropogenic NO_x emissions, especially in the eastern United States, which may lead to overestimation of ozone-temperature sensitivity in these regions. The GEOS-Chem model also misses several pathways in describing the responses of ozone to temperature, such as the responses in anthropogenic emission and land-atmosphere interaction through soil and vegetation. This will be discussed in detail in Section 4."

Regarding the influence of anthropogenic emissions' response to temperature on ozone-temperature sensitivity, we classified this under model uncertainties and further elaborate on this in section 4: "Nevertheless, there is significant room for improving the ability in capturing the ozone-temperature relationship in the chemical transport model. The GEOS-Chem simulations do not account for the response of anthropogenic NO_x and VOCs emissions to temperature. Recent studies have shown that these emissions can increase simulated regional ozone-temperature sensitivity by up to 7% and 14% (Kerr et al., 2019; Wu et al., 2024). The parameterization of several temperature-dependent processes is limited or even missing in the model. For example, the dry deposition scheme used in this study lacks the temperature response of non-stomatal pathways (Clifton et al., 2020), which could introduce uncertainty in simulated $m_{\Delta O3\text{-}\Delta Tmax}$ particularly in vegetation-rich regions such as the southeastern United States. Additionally, according to the BDSNP scheme used in this study, soil NO_x emissions are modeled as an exponential function of temperature between 0 and 30 °C, remaining constant at temperatures above 30 °C. However, some studies have reported continuous increases in soil NO_x emissions at temperatures higher than 30 °C in regions such as California (Oikawa et al., 2015; Wang et al., 2021). The absence of other temperature-dependent natural emissions, such as soil Nitrous acid (HONO) (Tan et al., 2023), may also lead to an underestimation of ozone responses to extreme temperatures in the GEOS-Chem simulations. Uncertainties in the biomass burning emission inventory (Fasullo et al., 2022) limit the accuracy of ozone-temperature sensitivity simulations in fire-impacted regions, such as the mountainous western United States. The 50 km resolution of the model may not fully capture sub-grid meteorological variations, which can play an important role in reproducing extreme conditions at site-level scales. Our study demonstrates that ozone-temperature sensitivity is highly responsive to changes in emissions, emphasizing the importance of more accurate anthropogenic emissions inventory for interpreting the ozone-temperature relationship. Further efforts are needed to enhance the model's ability to capture long-term trends in the ozone response to

temperature (including underlying weather conditions and transport patterns), and to better unravel the mechanisms driving the observed ozone-temperature relationship, in particular the role of transport and ventilation.”

Reference:

Christiansen, A., Mickley, L. J., and Hu, L.: Constraining long-term NOx emissions over the United States and Europe using nitrate wet deposition monitoring networks, Atmospheric Chemistry and Physics, 24, 4569–4589, https://doi.org/10.5194/acp-24-4569-2024, 2024.

Clifton, O. E., Fiore, A. M., Massman, W. J., Baublitz, C. B., Coyle, M., Emberson, L., Fares, S., Farmer, D. K., Gentine, P., Gerosa, G., Guenther, A. B., Helmig, D., Lombardozzi, D. L., Munger, J. W., Patton, E. G., Pusede, S. E., Schwede, D. B., Silva, S. J., Sörgel, M., Steiner, A. L., and Tai, A. P. K.: Dry Deposition of Ozone Over Land: Processes, Measurement, and Modeling, Reviews of Geophysics, 58, e2019RG000670, https://doi.org/10.1029/2019RG000670, 2020.

Fasullo, J. T., Lamarque, J.-F., Hannay, C., Rosenbloom, N., Tilmes, S., DeRepentigny, P., Jahn, A., and Deser, C.: Spurious Late Historical-Era Warming in CESM2 Driven by Prescribed Biomass Burning Emissions, Geophysical Research Letters, 49, e2021GL097420, https://doi.org/10.1029/2021GL097420, 2022.

Kerr, G. H., Waugh, D. W., Strode, S. A., Steenrod, S. D., Oman, L. D., and Strahan, S. E.: Disentangling the Drivers of the Summertime Ozone-Temperature Relationship Over the United States, J. Geophys. Res. Atmos., 124, 10503–10524, https://doi.org/10.1029/2019JD030572, 2019.

Oikawa, P. Y., Ge, C., Wang, J., Eberwein, J. R., Liang, L. L., Allsman, L. A., Grantz, D. A., and Jenerette, G. D.: Unusually high soil nitrogen oxide emissions influence air quality in a high-temperature agricultural region, Nat Commun, 6, 8753, https://doi.org/10.1038/ncomms9753, 2015.

Tan, W., Wang, H., Su, J., Sun, R., He, C., Lu, X., Lin, J., Xue, C., Wang, H., Liu, Y., Liu, L., Zhang, L., Wu, D., Mu, Y., and Fan, S.: Soil Emissions of Reactive Nitrogen Accelerate Summertime Surface Ozone Increases in the North China Plain, Environ. Sci. Technol., 57, 12782–12793, https://doi.org/10.1021/acs.est.3c01823, 2023.

Wu, W., Fu, T.-M., Arnold, S. R., Spracklen, D. V., Zhang, A., Tao, W., Wang, X., Hou, Y., Mo, J., Chen, J., Li, Y., Feng, X., Lin, H., Huang, Z., Zheng, J., Shen, H., Zhu, L., Wang, C., Ye, J., and Yang, X.: Temperature-Dependent Evaporative Anthropogenic VOC Emissions Significantly Exacerbate Regional Ozone Pollution, Environ. Sci. Technol., https://doi.org/10.1021/acs.est.3c09122, 2024.

Wang, Y., Ge, C., Garcia, L. C., Jenerette, G. D., Oikawa, P. Y., and Wang, J.: Improved modelling of soil NOx emissions in a high temperature agricultural region: role of background emissions on NO2 trend over the US, Environ. Res. Lett., 16, 084061, https://doi.org/10.1088/1748-9326/ac16a3, 2021.

Comment [1-3]: Line 155 suggests that the simulations have only been spun-up for one month; it's also unclear what initial condition is used (i.e. consistent with what year of meteorology and emissions). The manuscript needs to justify that the short spin-up time does not impact the results and that the simulation is at steady-state with the emissions. The authors cite the short lifetime for O3 in the boundary layer (and longer aloft) (lines 155-157). However, given that they do not parse how much of the surface O3 in the simulations is locally produced vs transported (regionally, intercontinentally, from the stratosphere), and that one of the important temperature-sensitive drivers is PAN, the 1 month simulation spin-up is not necessarily sufficient. The authors should test this for the year of maximum difference from the initial conditions (i.e. if the initial condition is consistent with 1990 emissions, then perform this sensitivity simulation for 2021): a global simulation that is spun-up for 6 months prior to generating the boundary conditions for the July simulation (to verify that changing transport of ozone and ozone precursors do not impact the results). These results should be included in the SI to justify the approach used here, and, in the unfortunate case that the results are impacted by the spin-up time, the authors would need to perform longer spin-ups for all their simulations.

Response [1-3]: Thank you for pointing it out. We agree that increasing the model simulation spin-up time to 6 months or longer is more reasonable. However, our study includes 17 simulations with different configurations at a resolution of 0.5°(latitude) × 0.625°(longitude), with three simulations are conduced biennially in 1995-2017. Re-running all the simulations will be a significant challenge to time and computational sources. We follow your suggestion to validate the reliability of our simulation experiments with a relatively short spin-up time. For this purpose, we first performed an 8-month global simulation starting from January 2017, which provided boundary conditions for high-resolution simulations in June and July (so that the spin-up time for global simulation is 6-month). The initial conditions for the high-resolution (0.5°x0.625°) simulations were obtained by interpolating the spin-up of the 6-month global simulation onto the high-resolution grid. We compared the surface ozone concentrations and ozone-temperature sensitivity between this long

spin-up time simulation and the BASE simulation with 1-month spin-up in Figure S2. The results show that the differences between the simulations with 1-month and 6-month spin-up times have only minor impacts on ozone concentrations and $m_{\Delta O3\text{-}\Delta Tmax}$. The average differences between the two simulations were only 0.3% for ozone concentrations and 2.3% for $m_{\Delta O3\text{-}\Delta Tmax}$, with extremely high spatial consistency (r > 0.99). This confirms that using a 1-month spin-up period does not affect our analysis and conclusions. Nevertheless, we acknowledge longer spin-up time should be taken in future modeling studies.

We have added the following content to the main text to demonstrate the reliability of our experiments in section 2.4: "To demonstrate this, we conducted an additional set of experiments, starting with a global simulation at 2°×2.5° resolution from1st January 2017 to 1st August 2017. The global simulation on 1st June 2017 was then interpolated into the high-resolution nested grid to drive the high-resolution simulation from 1st June 2017 to 1st August 2017. A comparison of surface MDA8 ozone concentrations and ozone-temperature sensitivity between the two sets of simulations is shown in Figure S2. We find that the differences between the simulations with 1-month and 6-month spin-up times had only minor impacts on ozone concentrations and $m_{\Delta O3\text{-}\Delta Tmax}$. The average differences between the two simulations were only 2.3% for ozone concentrations and 0.3% for $m_{\Delta O3\text{-}\Delta Tmax}$, with high spatial consistency (r > 0.99). This confirms that using a 1-month spin-up time for the simulation should not affect the analysis and conclusions. However, a longer spin-up time is favorable for generating global chemical fields when sufficient computational resources are available."

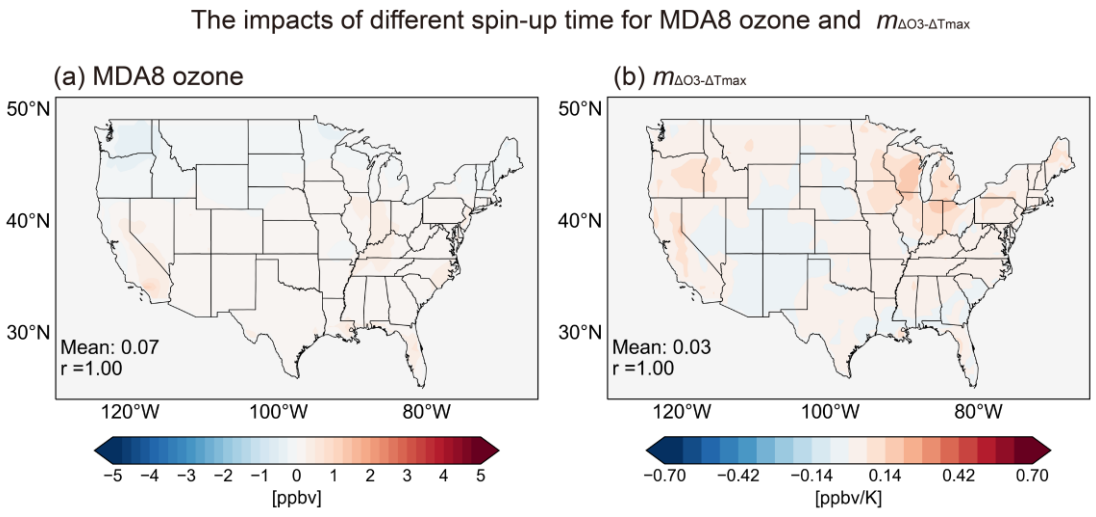

Figure S2. The impacts of different spin-up time for MDA8 ozone and $m_{\Delta O3\text{-}\Delta Tmax}$. The difference

between BASE and Long spin-up time simulations in (a) MDA8 ozone and (b) $m_{\Delta O3\text{-}\Delta Tmax}$. The correlation coefficients (r) between the simulations and mean values for the CONUS sites are shown inset.

Comment [1-4]: 3. Figure 3a and associated text: Can the authors explain the dip and rebound in the O3-T relationship from 1996-2004? Is this trend present in all regions?

Response [1-4]: We find significant spatiotemporal variability in ozone-temperature sensitivity trends across different regions. The overall declining trend in the CONUS ozone-temperature sensitivity slowed down or even reversed during 1996-2004, primarily due to an increase in ozone-temperature sensitivity across several regions, including the SEUS, Plains, and Midwest regions. Fu et al. (2015) pointed out that the increase in the SEUS region from 1990 to 2000 was mainly driven by changes in meteorological conditions, and this meteorological effect may have extended to other parts of the eastern United States during 1996-2004. In contrast, the western regions (the Intermountain West and SWUS) were not affected. We have added the following discussion in Section 3.1: "<u>However, we notice an increase of $m_{\Delta O3\text{-}\Delta Tmax}$ in 1990-2000 for the SEUS region and in 1999-2005 for the Plains region (Figure S4). The increase in ozone-temperature sensitivity in these two regions explains the $m_{\Delta O3\text{-}\Delta Tmax}$ plateau in CONUS during the 1996-2004 period. Fu et al. (2015) attributes the increase ozone-temperature sensitivity in 1990-2000 in the SEUS to variations in regional ozone advection tied to climate variability. This further underscores the significant influence of climate variability on $m_{\Delta O3\text{-}\Delta Tmax}$ trends.</u>"

Reference:

Fu, T.-M., Zheng, Y., Paulot, F., Mao, J., and Yantosca, R. M.: Positive but variable sensitivity of August surface ozone to large-scale warming in the southeast United States, Nature Clim Change, 5, 454–458, https://doi.org/10.1038/nclimate2567, 2015.

Comment [1-5]:

Line 94: "derive" seems inappropriate since the authors did not produce the MERRA2 product. I suggest "use" would be more accurate

Line 129: language “is capable of”

Line 151: “gas” should be plural

Line 180: language: replace “with both in” to “at both the”

Line 312: language: replace “propose by” with “theory using”

Line 314: language: “GEOS-Chem model simulates no”

Line 318: language: replace “neglectable” with “negligible”

Response [1-5]: Thank you for pointing it out. We have corrected them accordingly.