Reviewer #2

Comment [2-1]: Li et al. present an interesting analysis of tropospheric ozone responses to ENSO, including a quantification of the effects of transport, chemistry and biomass burning emissions on ozone in the tropical Pacific. The authors utilise the GEOS-Chem model alongside satellite observations for analysis of 'present-day' conditions, as well as CMIP6 models to study projections of future ENSO-ozone relationships.

The scientific questions addressed fall well within the scope of ACP and I recommend publication after the following comments, alongside concerns raised by reviewer 1, are addressed.

Response [2-1]: 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 [2-2]: To make the research reproducible, more information on the methodology needs to be provided. Reviewer 1 has already raised the question of how future ENSO events are identified in the CMIP6 models. In addition, please specify what correlation and linear regression methods were used to calculate the respective coefficients.

Response [2-2]: Thank you for your suggestion. We have specified the calculation of the Niño3.4 index in Section2.2: "<u>This index is calculated as:</u>

$$Niño3.4\ Index = SST - \overline{SST}_{1981-2010} \quad (1)$$

<u>where SST is the monthly mean SST averaged over the Niño3.4 region (5°N–5°S, 170°W–120°W), $\overline{SST}_{1981-2010}$is the 1981–2010 climatological mean SST for the same month over the Niño3.4 region.</u>"

We have also clarified the calculation of $r_{\text{TCO-Niño34}}$ and $m_{\text{TCO-Niño34}}$ in Section2.5:

"The $r_{\text{TCO-Niño34}}$ and $m_{\text{TCO-Niño34}}$ for each grid are calculated as:

$$r_{X-Y} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \quad (2)$$

$$m_{X-Y} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} \quad (3)$$

Where X_i is the gridded monthly deseasonalised and detrended TCO, Y_i is the monthly Niño3.4 index. These metrics effectively normalize comparisons across models with differing climate variability backgrounds."

Comment [2-3]: The uncertainty in the OMI/MLS retrieval should be introduced in the methodology or discussed when evaluating GEOS-Chem.

Response [2-3]: Thank you for your suggestion. We have added this information to Section 2.1: "Previous study shows an excellent agreement between the TCO from OMI/MLS and those observed from ozonesonde, with a relatively small bias of about 5 DU. These discrepancies mainly arise from from stratosphere-troposphere separation errors (3-5 DU), and cloud contamination (~2 DU) mitigated by filtering scenes with reflectivity >0.3 (Ziemke et al., 2006)."

Reference:

Ziemke, J. R., Chandra, S., Duncan, B. N., Froidevaux, L., Bhartia, P. K., Levelt, P. F., and Waters, J. W.: Tropospheric ozone determined from Aura OMI and MLS: Evaluation of measurements and comparison with the Global Modeling Initiative's Chemical Transport Model, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/10.1029/2006JD007089, 2006.

Comment [2-4]: Explain the choice of boundaries for the west and east Pacific regions in more detail (line 246-247). Were any particular thresholds used or, maybe, do these

regions align with previous studies?

Response [2-4]: Thank you for pointing it out. We carefully selected the boundaries for the west Pacific (WP) and east Pacific (EP) regions based on two primary considerations: (1) consistency with the established Ozone ENSO Index (OEI) proposed by Ziemke et al. (2010), and (2) statistical robustness within our specific dataset. While our regional definitions are conceptually similar to Ziemke et al.'s OEI, we optimized the boundaries to ensure that approximately 90% of the grid points within the regions shows consistent rTCO-Niño34 with a low p-value.

We added the explanation to Section 3.1: "The selection of these regions is based on consistency with the Ozone ENSO Index (OEI) framework (Ziemke et al., 2010), which also ensures that ~90% of the grid points within the regions shows consistent $r_{\text{TCO-Niño34}}$ with a low p-value."

Reference:

Ziemke, J. R., Chandra, S., Oman, L. D., and Bhartia, P. K.: A new ENSO index derived from satellite measurements of column ozone, Atmospheric Chemistry and Physics, 10, 3711–3721, https://doi.org/10.5194/acp-10-3711-2010, 2010.

Comment [2-5]: Three years of data are used for the El Niño and La Niña conditions to minimize potential impacts from other climate modes. This is not the case for the 'normal' conditions, which are based just on 2013. Please clarify how other climate modes might impact the 'normal' input data or why this is unlikely to be an issue.

Response [2-5]: Thank you for pointing it out. We agree that using multi-year data to define the "normal" condition is preferable to enhance robustness of analysis, but this would further increase the computational cost as we already have 24 sensitivity experiments on global scale. We select the year 2013 as the reference year based on careful evaluation of index indicating major climate model. In

addition to theNiño3.4 index analyzed in the text, we also find that the Indian Ocean Dipole (IOD) index stayed below the ±0.5°C threshold and the annual-mean Arctic Oscillation (AO) index is close to zero, indicating this year is not suffered from significant climate variability. Thus, 2013 represents a suitable reference year for "normal" conditions.

We have added the following explanation to Section 2.3: "To reduce computational cost, we only choose year 2013 as the reference year for the Normal period. Evaluation of other climate indexes such as the Indian Ocean Dipole (IOD) index and the Arctic Oscillation (AO) index shows that the year 2013 is not suffered from significant climate variability."

Reference:

https://psl.noaa.gov/data/timeseries/month/data/dmi.had.long.data

https://psl.noaa.gov/data/timeseries/month/data/ao.long.data

Comment [2-6]: The 1997/98 El Niño event is discussed in some detail on line 401. I think a brief introduction to the event should accompany the first mention of 1997 ozone levels on lines 53-54.

Response [2-6]: We have followed your suggestion by adding the text below in Introduction: "A large response of up to 25 DU in tropospheric column ozone was observed over Indonesia during September–November 1997, the period experiencing exceptionally strong El Niño conditions and extreme fires and weather around the world (Page et al., 2002; Picaut et al., 2002), which is comparable to the annual mean level of local tropospheric ozone column (Ziemke and Chandra, 2003)."

Reference:

Page, S. E., Siegert, F., Rieley, J. O., Boehm, H.-D. V., Jaya, A., and Limin, S.: The

amount of carbon released from peat and forest fires in Indonesia during 1997, Nature, 420, 61–65, https://doi.org/10.1038/nature01131, 2002.

Picaut, J., Hackert, E., Busalacchi, A. J., Murtugudde, R., and Lagerloef, G. S. E.: Mechanisms of the 1997–1998 El Niño–La Niña, as inferred from space-based observations, Journal of Geophysical Research: Oceans, 107, 5-1-5–18, https://doi.org/10.1029/2001JC000850, 2002.

Ziemke, J. R. and Chandra, S.: La Nina and El Nino—induced variabilities of ozone in the tropical lower atmosphere during 1970–2001, Geophysical Research Letters, 30, https://doi.org/10.1029/2002GL016387, 2003.

Comment [2-7]: On figure 7, from what I can see, not all the models feature on all the subplots. Please explain whether the missing model values are beyond the scales or whether a particular variable was not available. Additionally, a brief explanation of the Taylor diagram could be included in the figure caption to help the reader. For example, clarifying the axes.

Response [2-7]: Thank you for your suggestion. We added the explanation to Figure 7 caption: "For Panels (a) and (e), the reference $r_{TCO\text{-}Niño34}$ is from the OMI/MLS observations. For other panels, the reference data is the MERRA-2 reanalysis fields used from GEOS-Chem simulation. The radian axis represents the spatial correlation coefficient between model result and reference data, the X-axis and Y-axis represent the standard deviation (normalized to observations). The distance between a model marker and the reference point ("REF") quantifies the root-mean-square error (RMSE). Thus, markers closer to "REF" indicate better overall performance. By design, the result would not be presented if the simulated values show negative correlation coefficient with the reference data."

Comment [2-8]: On line 182 you introduce the ensemble member for the CMIP6 models and identify a different one is used for UKESM1-0-LL. What is the potential impact of using a different ensemble member on the results? The explanation of

‘r1i1p1f1’ may provide too much detail if there are no substantial implications of using that particular ensemble members. If there are implications, please highlight them.

Response [2-8]: Thank you for pointing it out. We selected “r1i1p1f1” as the primary ensemble member to maximize the number of models in our multi-model ensemble, following the common practice in CMIP6 model analyses (Wang et al., 2022). According to the UKESM documentation, the difference between “f1” and “f2” versions is that historical stratospheric aerosol properties are updated in v6.2.0 (f2) to remove errors in some years. This change does not significantly affect ozone simulations or other key variables analyzed in our study (e.g., SST, large-scale circulation). The use of these ensemble members is consistent with other studies using CMIP6 output for ozone analyses (Skeie et al., 2020; Sun and Archibald, 2021). We prefer to provide this information to help the readers to reproduce the results.

Reference:

Skeie, R. B., Myhre, G., Hodnebrog, Ø., Cameron-Smith, P. J., Deushi, M., Hegglin, M. I., Horowitz, L. W., Kramer, R. J., Michou, M., Mills, M. J., Olivié, D. J. L., Connor, F. M. O., Paynter, D., Samset, B. H., Sellar, A., Shindell, D., Takemura, T., Tilmes, S., and Wu, T.: Historical total ozone radiative forcing derived from CMIP6 simulations, npj Clim Atmos Sci, 3, 32, https://doi.org/10.1038/s41612-020-00131-0, 2020.

Sun, Z. and Archibald, A. T.: Multi-stage ensemble-learning-based model fusion for surface ozone simulations: A focus on CMIP6 models, Environmental Science and Ecotechnology, 8, 100124, https://doi.org/10.1016/j.ese.2021.100124, 2021.

Wang, H., Lu, X., Jacob, D. J., Cooper, O. R., Chang, K.-L., Li, K., Gao, M., Liu, Y., Sheng, B., Wu, K., Wu, T., Zhang, J., Sauvage, B., Nédélec, P., Blot, R., and Fan, S.: Global tropospheric ozone trends, attributions, and radiative impacts in 1995–2017: an integrated analysis using aircraft (IAGOS) observations, ozonesonde, and multi-decadal chemical model simulations, Atmos. Chem. Phys., 22, 13753–

13782, https://doi.org/10.5194/acp-22-13753-2022, 2022.
https://ukesm.ac.uk/cmip6/variant-id/

Comment [2-9]:

Adding an explanation to the figure captions of the boxes highlighting specific regions in some of the figures (e.g., Fig. 2 g) would provide more clarity.

There are missing 'the', 'a', and 'an' articles throughout the text, e.g. line 15: "Here, we evaluate the GEOS-Chem model…". I assume these will be addressed during the copy-editing stages.

Similarly, occasionally there are mistakes associated with the plural or singular. For example, on line 59: "Mechanisms contributing to the ozone-ENSO response has been examined…". I again assume these will be addressed during the copy-editing stages.

Line 47: "featured by" is unnecessary in the sentence starting "The key response is"

Line 186: should 'forces' be 'forcings' in this sentence?

Lines 190 and 192: rephrase awkward phrasing of "perform interactively tropospheric chemistry" and "perform interactively stratospheric ozone chemistry" to improve readability. For example, change to "simulate tropospheric chemistry interactively".

Line 247: "showing a significant but contrasting ozone response"

Lines 288-290. I suggest splitting in two the sentence starting "The simulated regional mean" for better readability.

Line 324: Remove the unnecessary "by" in "estimated by from the corresponding sensitivity experiments" in the figure caption.

As you state you are not using a threshold for significance, I suggest changing the wording on line 375 from significant to substantial.

I suggest moving the text on limitations (Lines 591 – 597) to earlier in the Discussion and Conclusions section, so that you have a stronger ending focusing on the key results and their implications.

Response [2-9]: Thank you for your suggestion. We have corrected them accordingly.