We would like to thank the reviewers for their comments. Below are responses to the individual comments. Responses are in blue, and revised text in the manuscript is in red.

Per the editor, one reviewer has raised the following concern regarding competing interests:

I believe that funding from Phillips 66 - an oil company - for TNS and AGR is a competing interest (conflict of interest) despite the authors not stating so. Phillips 66 is absolutely not a neutral funder, they are - via the American Petroleum Institute - actively lobbying against lowering ozone standards, so I don't see how this is not a major conflict of interest.

The authors disagree with the reviewer that the funding from the Phillips 66 Company constitutes a conflict of interest on the part of some of the authors. We have updated the funding acknowledgements to the following (new text in red):

"TNS and AGR received funding from the Phillips 66 Company. AGR also received funding from NASA HAQAST. Funding organizations have not dictated the topic or content of this work nor have they had any editorial role. The views expressed in this paper are those of the authors and do not necessarily represent the view or policies of the U.S. Environmental Protection Agency, the Phillips 66 Company, or NASA."

Reviewer 3:

I was not a reviewer for the original draft, and thus will only make overarching comments.

I found the acronyms of the model runs to be very confusing and not at all intuitive. For that reason, this manuscript was very hard to follow. I highly recommend the authors spell the words out rather than use acronyms. NAT=Natural, LINTL=Long-range international CANMEX=Canada-Mexico USB =U.S. Background, USA=U.S. Anthropogenic (which is different than USA the country, so confusing!) Please spell out these acronyms.

We have updated the acronyms as listed below and no longer use the acronyms in the text. There are many resulting changes in tables, figures, captions, and the text which are not detailed here.

USB becomes US Background

USA becomes US Anthropogenic

INTL becomes International

LINTL becomes Long-range international

CANMEX becomes Canada+Mexico

STRAT becomes Stratospheric

USB_NOSTRAT becomes non-Stratospheric US Background

There is a mismatch between the annual/seasonal averages of MDA8 and how the MDA8 is regulated which is the 4th highest concentration (essentially 99th percentile of annual concentrations). This is not a relevant comment in all aspects of the manuscript, however, this is particularly relevant for Tables 2 & 3 and Figures 2-5: the U.S. anthropogenic component, ~10 ppb, which makes sense for an annual average, but not relevant for a 99th percentile day. If I had reviewed the original manuscript, I would've had you plot the 99th percentile of Figure 2 and insert/discuss between Figures 2 & 3. That is still my recommendation, although I realize I did not review this in Round 1, so incorporating this comment may take a lot of effort. I will ultimately leave that decision up to the Editor. At the absolute minimum it would be important to acknowledge several times throughout the manuscript that the annual average is not relevant to the current NAAQS, especially near the text referring to Tables 2 & 3, Figures 2-5 (more than a brief mention in Line 764 Track Changes version at the end of the manuscript).

We have added two new figures similar to what is suggested by the reviewer. The new figures show the fourth highest simulated MDA8 O_3 in the base case along with the contributions from each of the O_3 components on the same day for the PA (Figure 4) and EQUATES (Figure 7) simulations. Additional discussion of these results has also been added to the manuscript.

The biases in the Spring and Summer (Lines 814-816 Tracked Changes version) seem like they warrant further investigation, especially since O3 is regulated based on the 4th highest value which typically occurs in May-Sept. I think Lines 817-832 suggests stratospheric influence is underestimated and missing particulate nitrate photolysis, but I think there may be other reasons. There is no mention of O3 produced from lightning and soil in the Discussion section, and I'm sure both have large uncertainties. Also no mention of O3 deposition uncertainties, or ozone production efficiency uncertainties. All of these uncertainties need to be discussed.

We have added the following to the discussion to mention lightning and soil NOx (new text in red):

"The biases associated with long-range international may be misattributed due to the difficulty of the regression model formulation to isolate stratospheric influences from other natural sources such as lightning and soil NO_x, wildfires, and biogenic VOC emissions, all of which have a high degree of uncertainty."

We have also added the following to the discussion about uncertainties in processes other than emissions that affect ozone (new text in red):

"Additional future work could take a process-oriented approach rather than the source-oriented approach described here. A process-oriented approach would focus on how different physical and chemical processes (deposition, transport, photochemical activity, etc.) relate to biases in O_3 simulations. The role of uncertainties in O_3 deposition and in O_3 production efficiency across various chemical regimes could be examined in a more process-focused analysis."

See also the response to the other reviewer's comment #3 where additional discussion of the role of natural emissions, including lightning and soil NOx, has been added.

Reviewer 4:

This study highlights biases as seen in US background ozone and its contribution to total ozone as predicted by two sets of CMAQ simulations, using a model-measurement fusion model utilizing a multivariate least square regression approach. The authors highlight that such a data fusion approach leverages the strengths of different data sources to reduce modeled US background (USB) ozone (O3) biases. The reduction in model USB O3 biases is critical, given the NAAQS for ozone will become more stringent, increasing the importance of background O3 (and its sources) to total O3. The manuscript is quite comprehensive and acceptable for publication, given the following main conceptual comments are addressed:

Comments:

1) Abstract + Lines 61-63: "We extend the bias correction method to estimate biases in separate components of USB O3. Separating the USB O3 components provides new insights into the inferred CTM error in USB O3 that was not possible when USB O3 was treated as a lumped quantity". Comments from the erstwhile reviewer(s) to make the abstract more quantitative have been addressed to some extent. However, the findings (or its summary) about 'source-specific US background O3 biases' (or quantitatively how much of them are corrected by this work?) are lacking in the abstract, given the authors propose this manuscript as an update to Skipper et al., 2021 in that specific direction only. Please consider adding 1-2 sentences on the main findings on source-specific USB O3 biases upfront in the abstract. Currently, the variation in USB O3 bias seasonally and with model resolution is quantitatively mentioned in the abstract. Summarize source-specific bias findings as discussed in the conclusions, in the Abstract as well.

As stated in the abstract:

"Correlation among different US background O_3 components can increase the uncertainty in the estimation of the source-specific adjustment factors."

And as stated in Section 4:

"The seasonality of inferred long-range international bias highlights a key uncertainty in correlative bias attribution. The biases associated with long-range international may be misattributed due to the difficulty of the regression model formulation to isolate stratospheric influences from other natural sources such as lightning and soil NO_x, wildfires, and biogenic VOC emissions, all of which have a high degree of uncertainty. Stratospheric O_3 is expected to have similar temporal and spatial patterns to long-range international, with contributions being higher in spring and at high elevations. It is suspected that the regression model formulation may be assigning a negative bias in long-range international to make up for missing stratospheric O_3 that has a similar pattern to long-range international while at the same time assigning a high bias for natural to reallocate some of stratospheric O_3 that is present in natural to long-range international instead."

For these reasons, reporting quantitatively in the abstract the bias attributed by the regression model approach to natural and long-range international contributions to US background O_3 would imply a much greater level of certainty in these numbers than is actually the case. We prefer to keep these in the main text where the quantitative results can be placed in the proper context.

2) Lines 143-145: “The EQUATES hemispheric simulations therefore include losses of O3 over seawater that are not present in the PA hemispheric simulations which could affect O3 transported over the Pacific in particular.”
Can you elaborate more on which source-specific US background O3 bias (for example, LINTL? Or some other scenario is more relevant to this?) is this effect the most here (describe using Labels in Table S2)?

The labels in Table S2 were requested to be removed from the text by the other reviewer. We have revised the text as follows:

“One potential source of differences is updates to halogen chemistry introduced in CMAQ v5.3 (Sarwar et al., 2019). These updates in the EQUATES simulations enhance halogen-mediated O_3 losses, which are strongest over the oceans. These losses are most relevant for O_3 contributions (natural and anthropogenic) that are transported long-distances over oceans.”

3) Lines 174-185: “For hemispheric-scale simulations, biogenic VOC emissions are from the Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1) (Guenther et al., 2012). The PA simulations additionally replace MEGAN emissions with emissions from the Biogenic Emission Inventory System (BEIS) (Bash et al., 2016) over North America (US EPA, 2019a).. The EQUATES MEGAN emissions are obtained from a compilation by Sindelarova et al. (2014). Soil NOx emissions for the PA hemispheric simulations are also from MEGAN with replacement by BEIS soil NOx over North America. Soil NOx emissions for the hemispheric EQUATES simulations are from a dataset by the Copernicus Atmosphere Monitoring Service (CAMS, 2018) based on methods by Yienger and Levy (1995). Lightning NO emissions for both the PA and EQUATES hemispheric simulations are from monthly climatology obtained from the Global Emissions Initiative (GEIA) and are based on Price et al. (1997). Lightning NOx was not included in the PA continental-scale simulations, while lightning NOx for the EQUATES continental-scale simulations is calculated using an inline module in CMAQ (Kang et al., 2019).”

In the above manuscript text, the authors highlight the differences in emission inputs for the PA versus EQUATES simulations. Can the authors enunciate/comment more on the impact of these differences on the model ozone biases, as per the scope of the manuscript, say: 1) with (Continental-scale EQUATES) and without Lightning NOx (Continental-scale PA), 2) Different biogenic VOC and NOx (Soil) emissions used in different simulations. For instance, impact of difference in Soil NOx emission inputs on USB O3 bias might be inferred from time series analysis in growing season (say, JJA) in agricultural regions such as, Central Valley and US Midwest.

The impact (or possible impact) of these differences (biogenic emission inputs being different between simulations) needs to be linked more clearly to the findings in the results and/or Conclusions. Also, a small typo here: North America (US EPA, 2019a). (please review and correct of any other such typos throughout the manuscript)

We have added the following to discuss the potential role of the different biogenic VOC, soil NOx, and lightning NOx emission inputs:

“The annual average of simulated US background O_3 for the hemispheric-scale (108 km resolution) and continental-scale (12 km resolution) modeling was slightly higher for the EQUATES simulations (32-33 ppb) than for the PA simulations (30-31 ppb). The differences are not explainable by the updated chemical mechanism used in EQUATES because the most relevant updates (halogen-mediated O_3 loss) tend to reduce O_3 at the northern mid-latitudes (Sarwar et al. 2019; Appel et al., 2021). The difference is also not likely due to anthropogenic emissions outside of the US, which are similar between the two sets of simulations. So, the higher US background O_3 in EQUATES likely relates to differences in the natural emissions. The EQUATES simulations used MEGAN for biogenic emissions throughout the entire Northern Hemisphere while the PA simulations used BEIS for biogenic emissions in North America and MEGAN elsewhere. The two hemispheric model configurations also used different sources for soil NO_x emissions (see Section 2.1) which could contribute to differences in US background O_3. Lightning NO_x emissions were the same in EQUATES and PA hemispheric-scale simulations, but the continental-scale PA simulations did not include lightning in the continental domain. Given that US background O_3 in both the EQUATES and PA 12 km continental-scale simulations are 1 ppb lower than their northern hemispheric counterparts, the differences in US background O_3 in the continental-scale simulations is more likely driven by the large-scale background inherited through the lateral boundary conditions than from differences in lightning NO_x configurations."