Response to the second round of review of the manuscript titled "*Explaining trends and changing seasonal cycles of surface ozone in North America and Europe over the 2000-2018 period: A global modelling study with NOX and VOC tagging*" by Ansari et al.

MS Number: egusphere-2024-3752

The reviewers' comments are shown in black while our point-by-point responses are shown in blue. All line numbers refer to the track-changes version of the revised manuscript.

Reviewer 1

The revised manuscript has incorporated reviewer suggestions and improved the analysis from *the* previous version. The results are presented in a more quantitative way with improved comparison between regions and removed redundancy. The manuscript has now included a Fourier analysis of the seasonal cycle.

I enjoyed reading the manuscript and I would recommend this for publication with one suggestion:

Perhaps the fourier analysis can be introduced in the methods, to give a clear introduction to the different parameters and their meaning. This may help in case the reader needs to refer back to the parameters, and also encourage others to use your analysis method.

We thank the reviewer for once again reviewing our revised manuscript and for their positive comment. We have now included the technical details about the Fourier Transform analysis in the section text S2 of the supplement and have referred the reader to this section in the main manuscript at lines L421-422.

Reviewer 2

Summary:
This paper is a revision of an earlier manuscript presenting a global modelling study with an innovative dual tagging analysis; the paper's focus is on surface ozone over North America and Europe, particularly with regard to the causes of long-term changes in mean concentrations and the seasonal cycle and its changes. In my view, this manuscript represents a very significant improvement over the previous version, and

addresses all issues that I raised in my first review, although some are still of concern. I recommend acceptance for publication when some remaining issues are addressed.

We once again thank the reviewer for seriously engaging with our work and providing detailed comments and feedback. We are glad that the reviewer appreciates the time and effort we put in in addressing their concerns in the first round of review and that they consider the updated manuscript as a very significant improvement.

Below I review the responses to the critical issues I identified in my previous review, discuss several remaining and new major issues, and list a few new minor issues.

Review of Critical Issues previously raised:

Section 3.1 is devoted to evaluation of the CAM-Chem model used in the chemical-transport simulations. The more robust evaluation now included in the paper is a significant improvement. I believe this section now gives a more comprehensive evaluation of the model performance, an approach that should be emulated by other studies presenting global modeling results.

Thank you for the positive comment. We agree that this is a more robust method of model-observation comparison, with less scope for visual misinterpretation.

Nevertheless, a few lingering issues remain, specifically:

Line 378 suggests that "… all HTAP tier 2 receptor regions for North America" were used for model evaluation. However, only 5 are included, while Fig. 3 of their reference (Galmarini et al., 2017) identifies 6 – Western Canada is excluded. I believe that this (at least below 60 deg N) is a particularly important region to include, both in the model evaluation and in the later discussion of results, since it provides a useful contrast with the more urbanized NW US and SW US inflow receptor regions lying to the south. Further, it may provide insight into reasons for the model overshoot of the maxima and undershoot of the minima for the earlier years of 2000-2006 in Eastern Canada. Please provide the comparison for Western Canada, or fully discuss the reason(s) for its being excluded. Importantly, please note that rural data are indeed available from this region - see Fig. 5 of Galmarini et al. (2017) and the figure that Owen R. Cooper included in his comment on this manuscript (https://doi.org/10.5194/egusphere-2024-3752-CC1).

The model evaluation in this study is performed against a gridded dataset derived from surface observations that are part of the TOAR-II database. This gridded dataset is ultimately based on the MDA8 O3 values reported only from rural stations. The rural/urban classification relies on the *type_of_area* station metadata from the original data providers. Unfortunately, we do not have any stations in Western Canada where

this metadata is included from the data providers and therefore they remain unclassified and are not included in the TOAR-II gridded dataset. Due to this reason, we do not discuss the Western Canada receptor region in the main text. However, given the importance of this region for background ozone studies, as pointed out by the reviewer, we have now included the modelled PSO time series with tagged contributions as well as the MDA8 O3 seasonal cycle change between the initial (2000-2004) and recent (2014-2018) periods for this region in the supplement (see Figures S7 and S8). These plots are based on the full sampling of the receptor region rather than co-sampling with TOAR observations due to their unavailability. We have now acknowledged this caveat in the manuscript in lines 396-401 and have also stressed the importance of including station metadata which allows selective evaluation in different contexts.

In my previous review I raised a subtle issue associated with model-observation comparisons such as the authors present in Section 3.1 and utilized in many such comparisons in the published literature, i.e., a degree of circular reasoning that results from models developed to agree with observations, so that such agreement cannot be taken as independent confirmation that models perform properly for the correct reasons. In their response to the reviews, the authors acknowledge a general concern with this issue, which they discuss in detail that generally discounts any large effects on their comparison. Some of this discussion is persuasive, but I could dispute other parts in further discussion; for example, nearly all persuasive model-measurement comparisons, whether for purposes of model "tuning" or evaluation of results, focus on metrics that are less sensitive to "nighttime ozone and avoid any large nighttime biases, which often arise due to improper simulation of the nighttime boundary layer". Rather than push this issue further, I suggest that the authors simply describe this issue and note their concern in one or two sentences in Section 3.1, perhaps at the end of the paragraph beginning on line 368.

We have now added the following text in lines L514-517.
"We note that agreement between models and observations does not in itself demonstrate that the models represent all processes correctly, since models are necessarily simplified representations of reality and can reproduce certain features for the "wrong" reasons. As Box (1976) succinctly put it, "all models are wrong, but some are useful"; our comparisons should therefore be viewed in this light."

I commend the authors for their illuminating discussion of the treatment of ship emissions in chemical transport models, now included in Section 3.1. In their response, the authors do attempt the first zeroth-order check that I suggested, and note a problem in comparing emissions in units of moles NOx/month with monthly average ozone concentrations in units of moles. However, the numbers derived in their comparison for July 2018 would be correct if the total global O3 attributed to ship NOx had a lifetime relative to total gross destruction of one month. This lifetime in the

summertime northern mid-latitudes is likely shorter than 1 month. For example, in July in the marine boundary layer (where the ozone derived from ship emissions is primarily formed) that lifetime is estimated to be ~10 days (see discussion in Mims et al., 2022 and their Figures S4 and S5). However, some of the ship derived ozone is transported to the cooler and drier free troposphere before destruction, so the effective overall lifetime for ship-derived ozone in summer at northern mid-latitudes might be best estimated as ½ month; thus this first zeroth-order check would indicate a model overestimate of ship derived O3 on the order of a factor of (7.04/4.25*2 =) 3.3. The authors also perform the second zeroth-order check that I suggested, and have included a discussion of this issue in Section 3.1 and the Supplement. Overall, I judge these responses to be an adequate response to this issue; however, I suggest inclusion of a discussion of the first zeroth-order check in the Supplement for interested readers to peruse, and in discussion of ship emissions in the paper, the authors should consider a likely overestimate of ship emissions by this factor.

We thank the reviewer for their insightful comment. Based on their suggestion, we have calculated an ship-tagged inferred ozone production rate by using an ozone lifetime of 0.5 months which yields a potential overestimation of ozone production from ships by a factor of 3.3. We have now included the details of our attempt at performing the first-order sanity check, its issues, and the potential fix as suggested by the reviewer, in Text S1 in the supplement in lines L17-62. We have also mentioned this in the manuscript in lines L481-485: "*We also performed a zero-order sanity check by comparing the inferred ozone production rate from ship NOx within the marine boundary layer of the northern hemisphere midlatitude region in the model with observational values. We found a potential overproduction of ozone by ships in the model by a factor of 3.3 when compared to the data from previous observational studies. We refer the reader to Text S1 in the supplement for a detailed discussion on these calculations.*"

• I am pleased that the issues of possible corruption of observational data have been resolved.

We once again thank the reviewer for pointing at the anomalous drop in stations which led us to track down this technical problem with data retrieval scripts. We are glad that it is sorted now.

• The discussion in Sections 3.2 and 3.3 now include more systematic and quantitative analysis approaches; it is now much improved, with the new Figures 5, 7, 11, and 13 adding significant additional information and clarity. However, several issues remain as detailed in many of the Major Issues below.

We are glad that the reviewer acknowledges the newer approach adopted for analyses of seasonal cycles and the additional spatial analyses based on local anthropogenic NOx

contribution to PSO in North America and Europe. We have addressed many of the newly raised *major issues* in the subsequent section.

• Section 4. Conclusion, Limitations and Future Outlook now provides a clear and reasonably concise summary of the new understanding of the atmosphere that has emerged from this study.

Thank you for the positive comment.

Major Issues:

1) I note that the current manuscript contains 17 figures composed of nearly 110 separate graphs, most with multiple traces that tend to overwhelm the reader. Some comments below suggest combining or revising manuscript figures and/or moving figures to the Supplement. Any changes the authors can make in this regard would improve the paper.

The increase in the number of figure panels is a direct consequence of the reviewers' earlier requests. For example, the figures showing PSO for North America and Europe (Figures 6 and 12) now have an extra column of panels to accommodate the reviewer's request to shorten the y-axis scales on the tagged contributions time series. We believe that the new figures look good and present the results in a clear manner. The addition of new spatial maps showing PSO-relevant local anthropogenic NOx emissions alongside their contribution to PSO and the total PSO for both continents (figures 5 and 11) were added based on reviewer 1's comments - and they seem to be satisfied with them. The second reviewer has also acknowledged in their earlier comment that they add "significant additional information and clarity". Similarly, the new figures showing the correlations between NOx emissions versus their contribution to PSO on a gridcell level for both continents (figures 7 and 13) were added based on this reviewer's request. This was a significant amount of added effort and as the reviewer notes, adds significant additional information and clarity, which we agree with. Therefore, we have decided against moving some of these figures into the supplement (which already has plenty of figures for other receptor regions).

2) In my first review I noted that the modeled contribution of local anthropogenic NOx to PSO decreases along western North America from the Southwestern US (with many large urban areas) to the Northwestern US (with few large urban areas); in this regard, I suggested that it would be useful to include the Western Canada receptor region in the analysis in Section 3.2 Ozone in North America. This inclusion would extend the western North American contrast to a region without large urban areas. The authors responded "We have not included discussion for Eastern Canada due to the unavailability of TOAR-II data from rural stations in this region which prevents model

evaluation and co-sampling for this region." This response is inaccurate; first, my comment referred to Western, not Eastern, Canada, and second, in fact there are TOAR-II data from rural stations in Western (as well as Eastern) Canada as I note in discussion above, and as is readily apparent in the authors' Figures 5 and 7 which include results for that Western Canada region. In my view, this western North American contrast is one of the more exciting prospects for improved understanding of surface ozone that could emerge from this manuscript, and should be exploited as fully as possible. It may help illuminate the comparison between the present work and recent observation-based studies that the authors discuss in the paragraph beginning on line 608. Please provide and fully discuss the model calculations for Western Canada as a 6th North American receptor region, or fully discuss the reason(s) for its exclusion.

We indeed made a typographical mistake in our earlier comment: we meant "Western" and not "Eastern" Canada. The reviewer is right in noting that results are available over this region as shown in Figures 5 and 7, however these are only modelled results and emissions, not observations. We do not discuss Western Canada in our study because we did not find any station categorized as rural over this region in the TOAR-II database. We would also like to point out that while this study stands on its own feet independently, it is submitted to a TOAR community special issue of ACP and one of its aims is to validate the TOAR-II observational database in its current form. The TOAR-II database relies on the *type_of_area* metadata from the original data providers and many of the data providers did not include this information, which led to only a single rural station in Western Canada which is so close to the US border that it gets engulfed into the NW US receptor region in the gridded observational dataset (see here: https://toar-data.fz-juelich.de/gui/v2/dashboard/?zoom=3&center=48.71,-87.14&country=CA&type_of_area=Rural&variable_id=5&data_start_date=2000-01-01T00:00&data_end_date=2018-12-31T11:59). We agree that there must be more rural stations in Western Canada but we did not have the appropriate metadata at the time of performing the study.

However, to address the reviewer's concerns, we have now added new figures for PSO and seasonal cycle change along with tagged contributions for Western Canada in the supplement; see figures S7 and S8. Here, due to lack of TOAR observations, the entire receptor region was sampled to produce the regional averaged metrics. We hope that these plots can provide value for those interested in understanding the contrast between Western Canada and Western US regions especially in terms of the role of background ozone.

3) I suggest that Figure 6 be replaced with simpler, more informative figures.

For discussion of observed and model-simulated total PSO (paragraphs beginning on lines 524 and 534) I suggest replacing the 5 graphs a, d, g, j, m and the potential 6th

graph for Western Canada with a two graph figure – one with all observed and the other with all model-simulated total PSO time series. Include the linear fits in both panels, and expand the ordinate to just the PSO range spanned by the time series (35 to 65 ppb?). For these graphs linear fits are appropriate, as the derived slopes quantify the scientifically interesting quantity of average PSO trends over the 6 receptor regions during the 2000-2018 period. Such a figure will allow improved visual comparisons of a) regional differences in PSO within the same graph, b) observed and model-simulated total PSO in side-by-side graphs, and c) the quality of the linear fits superimposed on the fitted data in each graph.

Here, the reviewer is asking us to override their previous request (which was to modify the aspect ratio of the panels to a nearly square format, and to make appropriate changes in the graph to better show the tagged contributions time series, rather than completely getting rid of the tagged contributions), which we have already accommodated in the first revised version of the manuscript. We separated the total PSO time series from the tagged contributions panel which allowed us to curtail the y-axis to a lower value thereby allowing better representation of the various tagged contributions (we did not change the y-axis to a log-scale but now we have separately fitted exponential functions to the local NOx contributions to PSO and reported them in Table S10). Now, the reviewer is asking us to only retain two panels: one with all observed and the other with all modelled PSO time series for all receptor regions. This approach will seriously compromise the reader's ability to view the trend in total PSO as a sum of trends in various NOx- and VOC-tagged contributions for different regions, which is a key result made possible by our TOAST 1.0 tagging methodology (as also mentioned in lines L237-238) that we want to communicate through this figure.

For discussion of the local anthropogenic NOx contribution (paragraph beginning on line 541) I suggest a figure with a single graph illustrating the time series from all 6 North American receptor regions. The 0 to 40 ppb ordinate range in the present Figure 6 should be retained. However, here I (again) suggest that the time series of this contribution in each receptor region be fit to an exponential function, rather than the linear analysis the authors currently employ. The rationale for this suggestion is discussed fully below in the 4th Major Issue. Such a figure will allow improved visual comparisons of a) regional differences in this PSO component within the same graph, and b) the quality of the exponential fits superimposed on the fitted data.

We have decided against merging the results of different receptor regions into a single graph for the same reasons as noted above - and therefore retained the figure in its previous (first revised) form. However, we have addressed the reviewer's concern of fitting exponential functions to the local NOx contribution to PSO for each receptor region. We have presented these results in the supplement (Figure S16 and Table S10) and have added a few lines of text discussing these in the main manuscript (lines L593-602).

"Several previous observational-based studies have inferred the magnitude and temporal decline of local contributions to ozone in North America based on curve fitting the observed ozone time series data and have reported these magnitudes and e-folding times of the local ozone enhancements for various stations and regions (Parrish & Ennis, 2019; Derwent & Parrish, 2022; Parrish et al., 2025 among others). In order to facilitate a comparison with these observational studies, we also fitted an exponential function of the form shown in eq. 1 to our model-derived local anthropogenic NOx contributions to PSO for various receptor regions (see Figure S16) and have tabulated the derived e-folding times against those found in literature (see Table S10). Here, A represents the magnitude of local NOx contribution to PSO for the initial year (2000) in ppb and τ represents the e-folding time of these contributions. We find τ = ~25-38 years from the model and ~22 years from the literature for various US receptor regions.

$y = A\ exp\ (-\frac{1}{\tau}t)$ ..... (1)"

For discussion of other NOx tagged contributions and all VOC tagged contributions, the time series plots in the current Figure 6 provide little information, and the discussion of their trends is based on the derived linear trends from Table 2, while the plots are not directly discussed in the relevant paragraphs (lines 574-607). Thus I suggest those plots be moved to the Supplement.

I suggest a similar treatment of Figure 12.

It is possible that on a first read it may appear to the reviewer that the discussion of results is purely based on the derived linear trends from Table 2. However, in fact, the first author of this manuscript wrote the description entirely while looking at Figure 6 and only inserted the linear trends within the text later to comply with the TOAR guidelines. So, we can only say that the assertion that plots are not directly discussed in the paragraphs is highly subjective and based on the reviewer's unique perspective and interpretation. We find it unreasonable to shift the time series of NOx and VOC-tagged contributions to PSO - a unique result from our modelling technique - to be relegated to the supplement.

The same arguments apply for Figure 12. However, we have also fit exponential functions to the local anthropogenic NOx contribution to PSO for Europe (Figure S16 and Table S10).

4) In the paragraph beginning on line 534 the authors discuss linear trends in PSO quantified by means of linear fits to time series. They note regional differences in the trends derived from the observations, and comment that the "model generally captures these decreasing trends and the interannual variability reasonably well, though with some regional differences in magnitude ...." Additional discussion should quantify what is meant by "reasonably well" (i.e., quantify differences between modeled and observed

trends and provide correlation coefficient between them), and discuss likely reasons for regional differences in trends (e.g., are larger trends found in regions with larger local NOx emissions?).

We have now computed and reported the correlation coefficients between observed and modelled PSO values for all receptor regions considered along with the differences between observed and modelled PSO trends. We have added the following text in lines L553-556 in the manuscript: "*r-values between observed and modelled PSO are 0.89, 0.78, 0.89, 0.93, 0.93, and 0.95 and the difference in modelled and observed trends are -0.09 ppb/yr, -0.02 ppb/yr, 0.07 ppb/yr, -0.16 ppb/yr, and -0.17 ppb/yr for E Canada, NW US, SW US, NE US, and SE US, respectively. These regional differences in PSO trends are driven by regionally different local and remote contributions to PSO as revealed in Figure 6.*"

We have also added the following text for European PSO at lines L871-873:
"*The model generally also captures the interannual variability in PSO for Western Europe and C&E Europe successfully (r=0.75 and 0.69 respectively) and to a lesser extent in Southern Europe (r=0.37; Figure 12).*"

We have also updated Figures 6 and 12 where correlation coefficients (r values) have been printed over the left panels for each receptor region.

5) I must strongly argue that the most interesting science question to address with regard to the local anthropogenic NOx contribution to PSO is quantification of the average relative rate of decrease. A particular fractional decrease in local anthropogenic NOx emissions is expected to give a larger absolute decrease in a region of large emissions compared to a region of small emissions; however the relative decrease may well be similar in the two regions. Similar relative rates of decrease of local anthropogenic NOx contributions to PSO would suggest similar fractional emission reductions and similar photochemical environments in the different regions, even if the absolute emissions differ, and thereby causing the absolute local anthropogenic NOx contribution to PSO to differ between those regions. An exponential fit to a time series of a quantity provides a means to quantify the average relative rate of decrease of that quantity. This is the rationale for my strong recommendation that the time series of the local anthropogenic NOx contribution be fit to an exponential function, since it gives a quantification that is of much greater scientific interest than the quantification of the absolute rate of decrease provided by a linear fit. Alternatively, the authors could employ a different technique to quantify the average relative rate of decrease.

To address the reviewer's request, we have now fitted exponential functions of the form $y = A\,exp\,(-\frac{1}{\tau}t)$ to our model-derived (i.e. tagged) local anthropogenic NOx contribution to regional PSO for all the nine receptor regions considered in this study.

These fits are presented in Figure S16 in the supplement along with the fit parameters **A** (initial year contribution on the curve, in ppb) and τ (e-folding time, in years). We have also included an additional table in the supplement Table S10 wherein we have compared the model-derived τ values from this study to the observations-derived τ values in previously published studies. We find a broad agreement in τ values from this study and published literature, i.e., smaller τ values for North American regions (~25-38 years from the model and ~22 years from observations) suggesting a relatively faster decline and larger values for European regions (~37-63 years from the model and ~37-44 years from observations) suggesting a slower decline in local contributions to ozone in Europe.

We have added extra text describing this in the manuscript in section 3.2.1 in lines L571-580:

"Several previous observational-based studies have inferred the magnitude and temporal decline of local contributions to ozone in North America based on curve fitting the observed ozone time series data and have reported these magnitudes and e-folding times of the local ozone enhancements for various stations and regions (Parrish & Ennis, 2019; Derwent & Parrish, 2022; Parrish et al., 2025 among others). In order to facilitate a comparison with these observational studies, we also fitted an exponential function of the form shown in eq. 1 to our model-derived local anthropogenic NOx contributions to PSO for various receptor regions (see Figure S16) and have tabulated the derived e-folding times against those found in literature (see Table S10). Here, A represents the magnitude of local NOx contribution to PSO for the initial year (2000) in ppb and τ represents the e-folding time of these contributions. We find τ = ~25-38 years from the model and ~22 years from the literature for various US receptor regions.
$y = A\ exp\ (-1/\tau\ t)$ ..... *(1)"*

And in section 3.3.1 in lines L896-901:
"As for North America, we have also fitted exponential curves (based on eq. 1) to the local anthropogenic NOx contributions to PSO in European regions in order to facilitate the comparison of the e-folding time (τ) with observationally-derived values in published literature (see Figure S16 and Table S10). We find a broad agreement with the observationally-derived values in that they are larger than those for North America (~37-63 years from the model and ~37-44 years from observations), suggesting a relatively slower decline in local contributions in Europe."

6) The authors' have added a very useful paragraph (beginning on line 609) that compares their results to the recent observation-based studies of Parrish et al. (2025). That comparison focuses on the apparent difference between the relatively small (<6 ppb in recent years) local anthropogenic enhancement to Ozone Design Values (ODVs) found in the observation-based study and the larger (~16 ppb) local anthropogenic NOx

contribution to average PSO found in the authors' model-based analysis. Despite some differences (discussed by the authors) between the two quantities derived to characterize the local anthropogenic contribution, it is clear that the authors have identified an important quantitative difference between the results of the two analyses.

We agree that this is an interesting and important observation and thank the reviewer for pointing us earlier to look deeper into such comparisons.

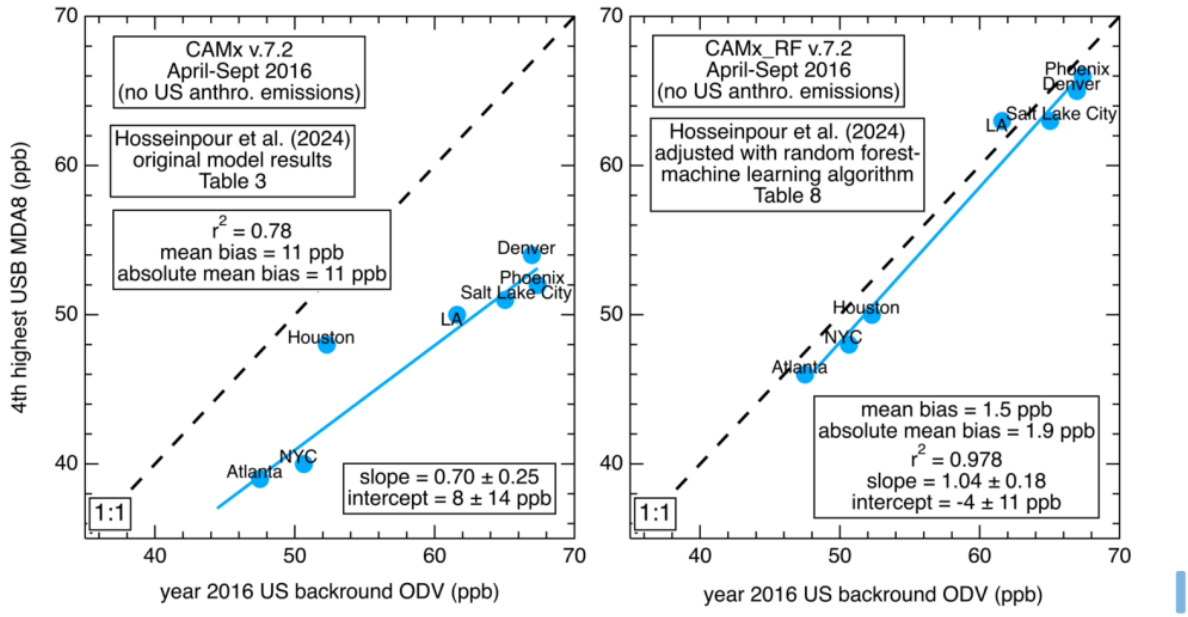
The authors suggest that perhaps the fundamental cause of the difference in results is that the observational-based method may systematically underestimate the full impact of local anthropogenic emissions, thereby overestimating the background contribution. However, they do not provide robust, quantitative analysis to support this suggestion. Importantly, Parrish et al. (2025) and references therein thoroughly discuss multiple lines of quantitative analysis to demonstrate that their observation-based approach does provide quantitatively accurate estimates of the ODV contributions from background ozone (US background ODV) and the local anthropogenic enhancement to ODVs.

Parrish et al. (2025) and references therein rely on deriving long-term trends for US background ozone which they describe by fitting a parabolic curve of the form $\mathbf{a + bt + ct^2}$ over long-term historical time-series data (1987-2018) and an additional exponential term of the form $\mathbf{A\ exp(-t/\tau)}$ to describe the changes in local anthropogenic enhancements. The idea of fitting the parabolic curve finds justification from the results of Parrish et al. (2020) where ozone data from only "background" (west-coastal, MBL, alpine) stations is considered. This methodology implicitly assumes that the local anthropogenic contribution will not be present in the ozone recorded at these sites. However, the local anthropogenic emissions from North America and Europe can be expected to contribute to some fraction of this background ozone too. To ascertain this, we sampled the PSO and the contribution of local anthropogenic NOx to PSO over the model grid cells corresponding to Trinidad Head and Mace Head stations. We have tabulated these values in Table S11. Indeed we found significant local (i.e., North American for Trinidad Head and European for Mace Head) anthropogenic NOx contributions to PSO sampled at these so-called background sites. In recent years (2014-2018) we find the local contribution to PSO at Trinidad Head to range between 4-6.6 ppb. Counting this portion of ozone as background would inevitably lead to an overestimation of background ozone and an underestimation of local enhancement. If this portion of ozone is added to the statistically-inferred local enhancement, it would bring the values reported in Parrish et al. (2025) much closer to our tagging-based results (i.e., around 12ppb vs 16 ppb).

We have now added this quantitative evidence in the manuscript text at lines L656:
"To ascertain this claim, we sampled the model output from the grid cells corresponding to these background stations (Trinidad Head for North America and Mace Head for Europe)

and calculated the site-specific PSO and local anthropogenic NOx contributions to PSO. These are reported in table S11 in the supplement. As expected, we found that a significant portion of PSO at these background sites contains contributions from local NOx emissions. For 2014-2018, we find the local contribution to PSO at Trinidad Head grid cell to be 4.0-6.6 ppb, which if added to the statistically-inferred local enhancement in SW US by Parrish et al. (2025) (6 ppb) would bring their values much closer to our findings (16 ppb).”

For example, there is strong evidence that zeroing out anthropogenic emissions in at least some global modeling systems lead to underestimates of the US background ODV; Parrish et al. (2025) discuss the work of Hosseinpour et al. (2024), who report the 4th highest US background (USB) MDA8 ozone concentration across the US from a global modelling system (CMAQ-CAMx) different from that used by the authors. The graph at left below shows that the model gives results that are biased (mean absolute bias = 11 ppb) from those of Parrish et al. (2025), much as the authors site for the current study. However when the model output is adjusted by a machine learning algorithm that regresses observed ozone on the simulated background and anthropogenic ozone fields, much improved agreement (mean absolute bias = 1.5 ppb) emerges, as the graph at right shows. It should be emphasized that this comparison is based on two completely independent analyses.



Given the “apples vs. oranges” aspects of comparing the overall results from the authors analysis with the observation-based approach, it is unlikely that a definitive comparison can be given in the current manuscript.

It is interesting to note this raw as well as bias-corrected ozone attribution result from a different modelling system. However, this example cited by the reviewer pertains to a

different kind of attribution system wherein local anthropogenic emissions are zeroed-out to infer the background concentrations. Such a large emission perturbation, in effect, fundamentally changes many aspects of the atmospheric chemistry within the model which no longer reflects real-world conditions where all emission sources are at play. Such perturbation-based estimates often lead to underestimation of background concentrations (see, Prather 2007; Grewe et al., 2010; Ansari et al. 2019 for further discussions). Here, in our tagging-based system, there are no perturbations to emissions and all emission sources are active at their realistic magnitudes when ozone formed from different sources is separately tagged and tracked within a non-perturbed, realistic, chemical environment reflecting real-world atmospheric conditions. We have already discussed this in the Introduction section (lines L111-127). A part of the "underestimation" of US background ozone in Hosseinpour et al. (2024) can be attributed to these nonlinear changes in atmospheric chemistry but the rest of it could also be realistic (i.e., the modelled background ozone isn't underestimated but rather the observationally-derived background ozone is overestimated as demonstrated in our previous response).

However, two straight forward comparisons would be informative and should be included:

• First, Parrish et al. (2025) and papers cited therein have shown that local anthropogenic enhancements of surface ozone in North American regions have decreased exponentially with a time constant of 21.8 ± 0.8 years, and they utilize fits to this characteristic temporal change as the basis for quantifying the local anthropogenic enhancement to Ozone Design Values (ODVs). Derwent and Parrish (2022) report similar exponential time constants 18 ± 4 years over the United Kingdom and 37 ± 11 years over continental Europe. Comparison of the present model-derived results for the temporal evolution of the local anthropogenic NOx contribution to those observationally derived results would be quite useful; hence, my continued insistence that the authors include exponential fits to the temporal evolution of the local anthropogenic NOx contribution over both North America and Europe.

As discussed in the previous response to major comment 5, we have now fitted exponential functions to our model-derived local anthropogenic NOx contributions to PSO time series for all the nine receptor regions considered in this study. These are shown in Figure S16 and Table S10.

• Second, the observation-based studies also show that the background contributions vary in a manner well-described by a quadratic polynomial (Equations 1 and 3 of Parrish et al., 2025) over the period of this study; this polynomial provides a means to calculate the overall change in background concentrations over that period. The authors report linear fits to all NOx-tagged NOx-contributions; these linear trends also provide a

means to calculate the overall change in concentrations of all NOx-tagged species over that period. The sum of the changes of all background species (foreign anthro. NOx, natural NOx, and ship NOx) would provide an estimate for the total background O3. This latter quantity should be compared with that from Parrish et al. (2025) (and from other published estimates of observation-based estimates of background ozone over this period, if the authors wish).

We have now tabulated the background contributions to PSO (i.e., Foreign anthropogenic NOx, natural NOx, and ship NOx contributions and their sum) for SW US in Table S12 in the supplement. Similar to Parrish et al. (2025) and previous observational studies we find a steady increase in background contributions to PSO in this region. We also fitted a quadratic curve of the form $a + bt + ct^2$ to the background contributions similar to Parrish et al. (2025). This curve fitting leads to parameter values of $a=26.6\pm0.94$ ppb, $b=0.35\pm0.24$ ppb/yr, and $c=0.0005\pm0.013$ ppb/yr^2.

We have added the following lines in the manuscript at lines L662-665:
"To facilitate better comparison with previous observational studies, we have also fitted a quadratic curve of the form a + bt + ct2, where t represents time in years starting at 2000, similar to Parrish et al. (2025), to the background contribution (sum of foreign anthropogenic NOx, natural NOx, and shipping NOx) to PSO for SW US (see table S12 in the supplement). We obtain parameter values of a=26.6±0.94 ppb, b=0.35±0.24 ppb/yr, and c=0.0005±0.013 ppb/yr^2 as compared to observationally derived values of a=71.5±0.8 ppb, b=0.07±0.13 ppb/yr, and c=-0.015±0.005 ppb/yr^2 in Parrish et al. (2025)"

7) The authors' have added very useful analyses (Section 3.2.2, 3.2.3, 3.3.2 and 3.3.3) that use Fourier analysis to quantify the ozone seasonal cycle and its contributions in both North American and European receptor regions. This analysis and associated discussion are quite informative; however to follow those discussions I had to construct a table that collected material from Tables S1-S9 (see below). I suggest that or a similar table be included in the manuscript. (Tables S1-S9 of the Supplement could then be reduced to a single table including the results of the 5 year periods for receptor regions not included in this table; tables of the results for individual years are not needed.)

	y0 (ppb)		A1 (ppb)		Φ1 (radians)		A2 (ppb)		Φ2 (radians)	
Region	obs	model	obs	model	obs	model	obs	model	obs	model
Eastern Canada	36.9	37.7	5.9	9.9	4.82	5.36	1.9	2.1	3.4	1.6
NW US	40.8	44.1	5.9	7.1	5.16	5.38	1.0	1.7	2.3	1.7
NE US	39.5	41.5	9.3	14.9	5.24	5.53	1.6	3.1	2.9	1.4
SW US	48.5	52.3	11.3	10.7	5.44	5.47	1.4	1.9	2.3	2.1
SE US	41.8	44.4	8.0	11.4	5.36	5.53	3.3	3.8	2.6	2.2
Western Europe	35.4	34.7	8.6	11.1	5.05	5.20	1.8	1.9	3.4	3.2
Southern Europe	41.2	42.4	11.6	12.4	5.39	5.45	1.8	2.7	2.2	2.3
C&E Europe	38.1	36.6	11.3	15.2	5.24	5.39	1.8	2.2	2.6	2.3
SE Europe	39.9	47.4	10.4	12.5	5.55	5.60	2.8	2.7	1.4	1.7
NW US 2000-2004	41.4	43.5	6.5	9.0	5.22	5.43	1.1	1.8	2.4	1.9
NW US 2014-2018	40.6	43.9	5.2	5.3	5.18	5.33	0.3	1.1	1.9	1.3
NE US 2000-2004	40.4	41.1	11.9	20.0	5.40	5.59	1.3	3.9	2.4	1.2
NE US 2014-2018	38.3	41.1	6.7	9.3	5.01	5.44	2.2	2.3	3.4	1.7
W. Europe 2000-2004	35.8	34.1	10.1	13.2	5.18	5.32	1.6	1.5	3.0	2.7
W. Europe 2014-2018	35.9	35.3	8.0	9.5	5.09	5.18	1.2	1.9	3.5	3.1
S. Europe 2000-2004	40.8	42.1	13.0	15.2	5.43	5.48	1.6	2.4	2.1	2.1
S. Europe 2014-2018	41.8	42.8	10.6	9.9	5.41	5.47	1.7	2.6	2.2	2.3

We thank the reviewer for their positive comment. We have now included a more concise table within the main manuscript (Table 2) which includes the Fourier parameters for the 9 receptor regions for 19-year and 5-year averaged seasonal cycles.

Paragraph beginning on line 649: Please more simply and clearly quantify the "tendency for overestimation"; e.g. the 2nd sentence could read: "The model generally captures these mean levels, though with a tendency for overestimation of 0.7 - 2.6 ppb in the eastern and 3.3 - 3.8 ppb in the western regions." The following 2 sentences could then be eliminated, and the final 2 sentences of the paragraph eliminated since they are mostly speculative.

We have now shortened this paragraph and used the sentence suggested by the reviewer at line L693. However, we have not removed the subsequent statements as we believe they can be very valuable for further analyses in future studies.

The paragraphs beginning on lines 659, 670, and 681 should be reviewed for similar opportunities for simplifying and clarifying the quantification and discussion, and removing speculative statements, unless quantitative analysis is added to support the speculation. (See the next Major Issue in this regard). The final 3 paragraphs of this section compare the Fourier analysis with Figure 8 and give an overall summary; they strike me as largely speculative, without firm quantitative analysis. I suggest shortening and clarification.

We have removed and softened some of these supposedly speculative statements and mentioned the need for further assessment via perturbation experiments which could

be done in future studies. However, more fundamentally, we believe it's important to point the reader to the features in the tagged contributions when interpreting the Fourier parameters. We also performed some arbitrary tests, for example, by artificially lowering local contribution numbers in some regions which modified the shape of the simulated ozone seasonal cycle and affected the resulting fourier parameters which confirmed many of our suggestive statements in the text. However, we did not include these little arbitrary tests in the text because we believe that these suggestive statements currently included in the text aren't very far-fetched and should be obvious to the reader. We have made similar modifications in section 3.3.2 for European seasonal cycle discussion.

8) To more fully inform the readers (and this reviewer) the mathematical definition and the physical significance of φ_1, the phase of the fundamental harmonic (not really of the annual cycle, but close if A2 < <A1) must be more fully explained. In the authors' reference (Parrish et al., 2016), the first term included in Fourier Analysis for the fundamental harmonic is (in the authors' notation) A1*sin(c + φ_1). In this approach, when φ_1 is zero, the peak of the fundamental is at p/2 radians, which corresponds to ¼ of the year or roughly the end of March. Importantly, a larger value of φ_1 gives an earlier (not later) peak; e.g. if φ_1 = p/2 radians the peak is on January 1. If the authors followed this approach, then their discussion of derived values of φ_1 is incorrect, because that discussion assumes a larger value of φ_1 gives a later peak. However, I imagine it would be possible to do the Fourier analysis with a negative sign rather than a positive sign in the fundamental term; if the authors followed this approach, then their discussion is correct. A full discussion of the approach actually followed is required, and the discussion corrected if necessary.

If the authors followed the approach of Parrish et al. (2016) then it may be clearer to give values of φ_1 after subtracting 2p, so that more negative φ_1 values correspond to later peaks. This is valid since the phase angle repeats after it advances by 2p,

We thank the reviewer for pointing out this inconsistency with the phase angles. We had utilized the *fourier_info* function of the NCAR Command Language (NCL) to perform the fourier transforms for the seasonal cycles (https://www.ncl.ucar.edu/Document/Functions/Built-in/fourier_info.shtml). On a closer look, we found that this function returns the phases not in terms of phase angles in radians but in terms of the actual abscissa (x-coordinate) where the peaks occur for each harmonic. In our case, the supplied time series are of length 12 each, representing the 12 months, and therefore the phases can be reported within the range of 0-12 months. For example, a phase of 5.0 would mean the end of May while a phase of 2.5 would mean mid-March. We have now clarified this in the manuscript and added a new text section in the supplement Test S2. We have also revised the figures and tables to replace φ1 and φ2 with p1 and p2 respectively which are expressed in months rather than radians.

9) In my judgement the most interesting feature of the φ_1 values is that for all but one receptor region in North America and Europe, both the modeled and observed φ_1 values fall within ± 0.3 radian (or 17 days) of a mean value of 5.37 (or -0.91) radians, which corresponds to a seasonal maximum of the fundamental on Julian day 144 or May 24. The discussion of this quantity might best further emphasize this close regional and model-observation agreement, before discussing the relatively small differences.

With the updated definition of the phases, this broader point about model-observation agreement is still true but it now means that the agreement is even closer, within ± 0.3 months, i.e., 9 days of a mean value of 5.37 months or June 12th.

10) Line 728: Please specify that Figure S2 shows modeled seasonal cycle envelopes. It would be illuminating to include a similar figure showing observed seasonal cycle envelopes.

Figure S2 in fact shows the observed seasonal cycle envelopes. The caption was misleading - we have corrected it now, and have also included a similar envelope figure for modelled data (Figure S3). We have now changed the text in the manuscript accordingly, "*(see Figure S2 and S3 for observed and modelled seasonal cycle envelopes over the entire period).*"

To my eye, there are evident, but small, seasonal cycle changes, with significant variability about consistent systematic changes. Thus, I would expect difficulty in quantifying the systematic changes, and this difficulty should be carefully considered before making firm conclusions. In this regard, Figure S2 indicates that seasonal changes appear to be clearer and more systematic in NE US compared to SW US. Thus, it may make sense to give a clear, statistically significant analysis of NE US first, and then address the SW US second.

We agree with the reviewer that NE US shows a distinct and clearer shift in the seasonal cycle over the two decades than other sub-regions. Therefore, we have moved the discussion of NE US before the NW US discussion in the text and Figures 9 and 10 have been flipped. We have also made other small tweaks to the text, for example, referring the reader to the envelope plots to fully appreciate the variability in these 5-year averaged changes. Additionally, we have now included regionwise ASCII files containing observed and modelled monthly mean MDA8 O3 seasonal cycles along with their tagged contributions for each year along with the manuscript.

11) Section 3.2.3 is well organized, but I think the discussion could be simplified and clarified, and in a few places corrected; in particular:
• Line 731: Inclusion of parameters of the Fourier analysis of the 5-year averaged periods

should be included in a table in the manuscript for the two example regions, as suggested in Major Issue 7).

We have now included a new table (Table 2) in the manuscript which shows the Fourier parameters for observed and modelled seasonal cycles for the entire, initial, and recent periods for the receptor regions discussed in the manuscript.

• Lines 741-742: Note that the shifts in φ_1 of 0.04 and 0.10 radians correspond to shifts of only 2.3 and 5.8 days, respectively – quite small shifts. And as noted in Major Issue 8) the peak of the fundamental shifts in the opposite direction from the phase shifts.

Based on the clarified definition of the phases used in our analyses, these shifts mean 0.04 and 0.10 months, i.e., 1.2 and 3 days respectively which are indeed very small. We have noted this in the text at line L842. However, the sign of these peaks is in the same direction of the shifts in time as discussed earlier in our response to major comment 8.

• Lines 773-774: A φ_1 shift from 5.40 to 5.01 radians actually indicates a shift of the seasonal maximum by 23 days, but from spring towards summer. Those values correspond to peak value shifting from Julian Day 143 to 165 or May 23 to Jun 14. I do not see how this is consistent with Figure 10. An explanation is required (perhaps in the Supplement) so that the reader can fully follow the discussion. An example showing how the 1st and 2nd harmonics combine to approximate the seasonal cycle in the two 5-year periods in the NE US would be quite helpful to include in the Supplement.
Given the issues identified above, I suggest that this Section be completely rethought, with the concluding paragraph revised as needed.

With the clarified definitions of the phases, a φ_1 (now, p1) shift from 5.40 to 5.01 refers to a summer-to-spring backward shift from 12 June to 1 June. We have now discussed the Fourier Transform process in the supplement in Text S2.

12) I have not attempted to critically review Section 3.3 as carefully as I did Section 3.2. The discussion in these sections is similarly organized for both continents. Please seek to include any manuscript improvements made to the former sections in the latter sections where appropriate. And please similarly review all major and minor comments that refer to Section 3.2 when revising Section 3.3.

We have implemented many of the changes suggested in the comments to European analyses where relevant. For example, we have also fitted exponential functions to European receptor regions (see Figure S16 f-i; lines L947-952). We have also computed correlation coefficient r for observed vs modelled PSO for European receptor regions which are annotated in Figure 12 and mentioned in the text at lines L933-934. We have also included new text on the exponential fits to the local anthropogenic NOx contributions to PSO in European regions in order to facilitate the comparison of the

e-folding time (τ) with observationally-derived values in published literature (see Figure S16 and Table S10). We have also reported local contributions at a background site in Europe (Mace Head at the west coast of Ireland) in Table S11. We have also removed/clarified some of the speculative statements in section 3.3.2 which relate Fourier parameters to tagged contributions.

13) Section 3.4 raises an entirely new area of discussion that raises new questions in my mind, specifically:
• Its introduction is somewhat confusing. I suggest changing the phrase "in these regions" to "in the receptor regions", assuming this is correct.

We have now changed the text accordingly.

• The 2nd sentence in the 2nd paragraph is also confusing. "It is noteworthy that this NO2_FOREIGN, locally recovered from foreign ozone titration, is separately tagged in our modelling system than the NO2 directly flowing from foreign regions (which we do not discuss here)." First, "… separately tagged in our modelling system than …" is not clear to me. Second, it raises the question of what exactly is and what is not included in the tagging. NO2 directly flowing from foreign regions is generally considered to be small due to the short lifetime of NOx in the troposphere, but what about PAN and other organic nitrates? They have been considered reservoirs of sequestered NOx that can be transported over intercontinental distances in the free troposphere. However, it is not clear to me how the model treats ozone produced by foreign NOx transported as an organic nitrate to a receptor region, where it produces NOx after release from the reservoir species.

I suggest that the authors remove Section 3.4 from this paper, which is already quite long, and then more fully discuss the NOx-tagging system in the Introduction or Section 2.1 so that the reader is aware of issues such as tagging of NO2 directly flowing from foreign regions, and NOx reservoir species transported from foreign regions.

We have decided to retain this section because we believe it provides a crucial piece of knowledge on the reasons for the increase in wintertime ozone in NAM and EUR, especially the distinction between actual increase in the inflow of foreign Ox and the weakened titration due to a drop in local NOx. However, based on the reviewer's suggestion we have now further improved the text to make it clearer and have also added further details in section 2.1:
"The TOAST system differentiates NO2 into two distinct chemical families: NOy and Ox, with separate tracers for NO2 as members of each of these families. NO2 as a member of the NOy family tracks NOx which is directly emitted or produced in the atmosphere (e.g. by lightning), while NO2 as a member of the Ox family tracks NO2 which is formed chemically through reactions of NO with either ozone or peroxy radicals and subsequently

undergoes photolysis to ultimately form ozone. Further details are given in Butler et al. (2018)."

Minor Issues:

1) Figures S3 and S4 present scatterplots for the parameter values derived from the Fourier analyses. The derived r values annotated in the figures quantify how well the model reproduces the interannual variability in the parameter values in the respective regions. It would be useful to also give the r value for the entire 95 (North America) or 76 (Europe) set of values; this (generally significantly larger value) would quantify how well the model reproduces both the spatial variability and the interannual variability of the respective parameter throughout all regions on each continent. From inspection of the figures the model performance for some parameters appears to be quite impressive indeed. Note that the caption to Figure S4 should give 76 (not 95) as the number of markers.

Thanks for this suggestion. We have now included the correlation coefficient *r* values for the entire 95 and 76 data points for NAM and EUR respectively in the updated figures. We have also corrected the names of the phase parameters (p1 and p2) and their units (months).

2) In lines 403 and 413 the y0 parameter is described as representing annual average MDA8 O3 derived from detrended data. However, since the authors derive values for only a single year, no detrending has been performed, and the y0 parameters thus represent actual annual averages, and thus, still include the interannual variability. I suggest removing the references to "detrended" data.

Thanks for pointing this out. *y0* was calculated for each year before detrending the full 19-year data that was used to perform the fourier decomposition. We have therefore removed these references to detrended data.

3) Unless the authors have a particular reason for including Tables S2-S9, I suggest they be removed, or at least shortened to only the summary values spanning the multi-year periods.

We have now included the fourier parameters for 19-year and 5-year averaged periods in a new table (Table 2) in the main manuscript. However, we have retained the tables S1-S9 in case readers are interested in a quantitative characterization of individual year seasonal cycles.

4) Line 396 : I suggest that the correlation coefficients at the annual average timescale be explicitly stated (i.e., 0.34 to 0.95). Add a similar statement to the paragraph for Europe beginning on line 428 and for the Belarus & Ukraine region on line 451.

We have now included the r values for annual averaged timescale for all these regions in the text.

5) Line 430: Modify final phrase to "…, except SE Europe and RBU."

We have modified the text accordingly.

6) Lines 531-32: This statement should be more forcefully stated, something like "… the observed PSO levels consistently exceeded the WHO guideline (31 ppb) throughout the study period by at least 10(?) ppb". Similarly for European regions on line 827.

We have now modified this sentence to: "*Crucially, across all North American regions, the observed PSO levels consistently exceeded the WHO long-term guideline (31 ppb) by at least 10 ppb throughout the study period.*"

7) Line 532 and elsewhere: When measured or modeled ozone concentrations are compared to the WHO guideline, it should be specified that it is the "WHO long-term guideline" that is being referenced.
We have modified all instances of "WHO *guideline*" to "WHO *long-term guideline*" in the text.

8) Line 535: Upon the first occurrence of the authors' Quantitative quote of a trend (e.g., (-0.19 (0.01) [-0.32, -0.06] ppb/yr), please define the 4 numbers given.

We have now added the following description after the first occurrence: "*[here and henceforth the trends are reported in the following format (trend (p-value) [95% confidence lower limit, 95% confidence upper limit])]*"

9) Table 2 should include the value of the derived trend (i.e., the most probable value of the trend) even if the 95% confidence intervals include zero. The table would be clarified if the column spacing were adjusted so that in each table entry the linear trend appears on the 1st line, p-values (shown in parentheses) on *the* 2nd line, and 95% confidence intervals on 3rd and 4th lines with all negative signs appearing on correct lines.

The entries in Table 2 follow the format as described by the reviewer.

10) Lines 553-57: This sentence requires clarification; it refers to "year-to-year variations in local emissions", which may be taken to indicate interannual variability. However, I certainly expect (and from the discussion the authors seem to agree) that the temporal correlation is largely driven by systematic decreases in local NOx emissions over the 19 year study period.

We have now included the point about the systematic decline in the sentence: "..year-to-year variations in local emissions (i.e., their systematic decline) significantly drive the variability (decline) in total PSO levels."

11) Line 568 states that: "… reductions in local NOx emissions translate directly and proportionally to reductions in the ozone …". It is clear that the translation is direct, but there is no analysis to show that it is proportional (i.e., linearly related). Unless this proportionality can be demonstrated and the proportionality constant quantified, the phrase "and proportionally" should be removed.

Thanks for picking up this subtle point. We have removed the phrase "*and proportionally*" from this sentence.

12) Line 644-645 state "The phase φ_1 indicates the timing of the annual peak, with numerically larger values typically corresponding to a later peak in the year …." This is not correct; larger phase angle values always correspond to an earlier peak in the year. Please see discussion in Major Point 8).
After the clarified definition of the phase, this statement is valid. We have changed all instances of φ1 and φ2 to p1 and p2, respectively, to avoid any confusion.

13) Line 646 and elsewhere: "Tables S2-S6" should be "Tables S1-S5".

Now corrected.

14) Lines 737-738: Discussion could be made more accurate, viz. "The observed annual mean ozone (y0) decreased slightly from 41.4 ppb to 40.6 ppb, while the modeled y0 increased slightly from 43.5 ppb to 43.9 ppb), slightly increasing the positive bias noted earlier."

We have modified this sentence accordingly.

15) In Table 3, certainly only one decimal place in the entries is statistically justified.

We have updated the entries in Table 3 (now, Table 4) to only one decimal place.

16) Line 1057: I suggest strengthening – perhaps end the sentence with "… in both regions exceeds the long-term WHO guideline by wide margins over the entire study period.

We have modified this sentence accordingly.

References:

Mims, C.A., D.D. Parrish, R.G. Derwent, M. Astaneh and I.C. Faloon (2022), A conceptual model of northern midlatitude tropospheric ozone, Environ. Sci.: Atmos., 2, 1303-1313, DOI: 10.1039/d2ea00009a. Hosseinpour, F., Kumar, N., Tran, T., and Knipping, E.: Using machine learning to improve the estimate of U.S. background ozone, Atmospheric Environment. 316. 120145, https://doi.org/10.1016/j.atmosenv.2023.120145, 2024.

Prather Michael J 2007Lifetimes and time scales in atmospheric chemistryPhil. Trans. R. Soc. A.3651705–1726
http://doi.org/10.1098/rsta.2007.2040