$2^{nd}$  review of "Explaining trends and changing seasonal cycles of surface ozone in North America and Europe over the 2000-2018 period: A global modelling study with NO<sub>X</sub> and VOC tagging" by Ansari et al.

MS Number: egusphere-2024-3752

### **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. 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. 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).

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.

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 O<sub>3</sub> 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  $\sim 10$  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  $O_3$  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.

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

### **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.
- 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 6<sup>th</sup> North American receptor region, or fully discuss the reason(s) for its exclusion.

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 6<sup>th</sup> 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.

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 4<sup>th</sup> 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.

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.

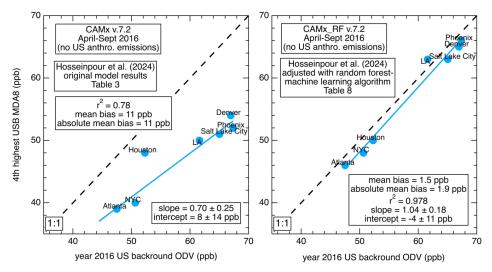
I suggest a similar treatment of Figure 12.

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?).

- 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.
- 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.

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. 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 4<sup>th</sup> 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 the 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. 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.
- 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 O<sub>3</sub>. 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).
- 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.)

Region	y0 (ppb)		A1 (ppb)		Φ1 (radians)		A2 (ppb)		Φ2 (radians)	
	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

Paragraph beginning on line 649: Please more simply and clearly quantify the "tendency for overestimation"; e.g. the 2<sup>nd</sup> 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. 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.

8) To more fully inform the readers (and this reviewer) the mathematical definition and the physical significance of  $\varphi_1$ , 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)  $A1*\sin(\chi+\varphi_1)$ . In this approach, when  $\varphi_1$  is zero, the peak of the fundamental is at  $\pi/2$  radians, which corresponds to 1/4 of the year or roughly the end of March. Importantly, a larger value of  $\varphi_1$  gives an <u>earlier</u> (not later) peak; e.g. if  $\varphi_1=\pi/2$  radians the peak is on January 1. If the authors followed this approach, then their discussion of derived values of  $\varphi_1$  is incorrect, because that discussion assumes a larger value of  $\varphi_1$  gives a <u>later</u> 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  $\varphi_1$  after subtracting  $2\pi$ , so that more negative  $\varphi_1$  values correspond to later peaks. This is valid since the phase angle repeats after it advances by  $2\pi$ ,

- 9) In my judgement the most interesting feature of the  $\varphi_1$  values is that for all but one receptor region in North America and Europe, both the modeled and observed  $\varphi_1$  values fall within  $\pm$  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.
- 10) Line 728: Please specify that Figure S2 shows <u>modeled</u> seasonal cycle envelopes. It would be illuminating to include a similar figure showing <u>observed</u> seasonal cycle envelopes. 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.
- 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).
  - Lines 741-742: Note that the shifts in  $\varphi_1$  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.
  - Lines 773-774: A φ<sub>1</sub> shift from 5.40 to 5.01 radians actually indicates a shift of the seasonal maximum by 23 days, but from <u>spring</u> towards <u>summer</u>. 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 1<sup>st</sup> and 2<sup>nd</sup> 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.

- 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.
- 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.
- The 2<sup>nd</sup> sentence in the 2<sup>nd</sup> 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 NO<sub>2</sub> 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. NO<sub>2</sub> 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 NO<sub>2</sub> directly flowing from foreign regions, and NOx reservoir species transported from foreign regions.

#### **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.
- 2) In lines 403 and 413 the y0 parameter is described as representing annual average MDA8 O<sub>3</sub> 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.
- 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.
- 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.
- 5) Line 430: Modify final phrase to "..., except SE Europe and RBU."
- 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.
- 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.
- 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.
- 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 1<sup>st</sup> line, p-values (shown in parentheses) on 2<sup>nd</sup> line, and 95% confidence intervals on 3<sup>rd</sup> and 4<sup>th</sup> lines with all negative signs appearing on correct lines.
- 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.
- 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.
- 12) Line 644-645 state "The phase φ₁ 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).
- 13) Line 646 and elsewhere: "Tables S2-S6" should be "Tables S1-S5".
- 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."
- 15) In Table 3, certainly only one decimal place in the entries is statistically justified.
- 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.

## New references not included in the manuscript under review or in my previous review

- 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, <a href="https://doi.org/10.1016/j.atmosenv.2023.120145">https://doi.org/10.1016/j.atmosenv.2023.120145</a>, 2024.