

Final author reply to editor

Dear handling editor,

Based on both referees' comments we have made major revisions to the paper. These changes include:

1. Changing the title of the manuscript to “Examining TROPOMI formaldehyde to nitrogen dioxide ratios in the Lake Michigan region: implications for ozone exceedances”
2. Removing all text/figures/references/etc. relating to PAMS data
3. Re-processing the data composites, specifically:
 - a. Using the reprocessed TROPOMI NO₂ PAL dataset so that all data come from the same processor version
 - b. Removing the use of detection limit thresholds
 - c. Addressing the HCHO artifact over water through a bias-correcting approach
 - d. Using the same number of days for the TROPOMI weekday-weekend composites
4. Adding new 2-meter temperature composites from the NAM analysis dataset
5. Expanding on the discussions of FNR errors (e.g., citing Souri et al., 2023), the usage of the J20 thresholds, and comparisons to similar studies (e.g., Tao et al., 2022)
6. Moving some of the appendices to a supplemental information document along with new supplemental figures/tables

We believe that these changes have greatly added to the scientific content of the paper and look forward to another round of discussion, if needed.

Note: in the “tracked changes” version of the manuscript, **we highlight the major changes/additions in red text**. Some minor changes have not been highlighted in red text (otherwise most of the text would be in red text).

Our responses to the individual referee comments are below:

Response to referee #1

Referee Comment

Title: Ozone–NO_x–VOC Sensitivity of the Lake Michigan Region Inferred from TROPOMI Observations and Ground-Based Measurements

General comments

Acdan et al. use TROPOMI column density retrievals and PAMS surface concentration measurements of HCHO and NO₂ to study HCHO, NO₂, and ozone production sensitivities (indicated by FNRs) in the lake Michigan region. The authors carefully composite for typical ozone season days, ozone exceedance days, weekdays, and weekends. They identify a spatial heterogeneity in ozone chemistry sensitivity for Chicago metropolitan area and its surrounding region, where the metropolitan area remains VOC-sensitive. They find that changes in FNRs on ozone exceedance days indicate an increase in NO_x-sensitivity in NO_x-sensitive areas and an increase in VOC-sensitivity in VOC-sensitive areas. Connecting wind fields with lake breeze provides a nice illustration of a stronger lake breeze effect on higher-ozone days.

Overall, the paper provides important implications for ozone mitigation in the Lake Michigan Region. However, a few major issues should be addressed before I recommend publication.

Introductory comment

We thank referee #1 for providing thorough feedback on our manuscript. Based on both referees' comments, we have made major revisions to the paper, including:

7. Changing the title of the manuscript to “Examining TROPOMI formaldehyde to nitrogen dioxide ratios in the Lake Michigan region: implications for ozone exceedances”
8. Removing all text/figures/references/etc. relating to PAMS data
9. Re-processing the data composites, specifically:
 - a. Using the reprocessed TROPOMI NO₂ PAL dataset so that all data come from the same processor version
 - b. Removing the use of detection limit thresholds
 - c. Addressing the HCHO artifact over water through a bias-correcting approach
 - d. Using the same number of days for the TROPOMI weekday-weekend composites
10. Adding new 2-meter temperature composites from the NAM analysis dataset
11. Expanding on the discussions of FNR errors (e.g., citing Souri et al., 2023), the usage of the J20 thresholds, and comparisons to similar studies (e.g., Tao et al., 2022)
12. Moving some of the appendices to a supplemental information document along with new supplemental figures/tables

We believe that these changes have greatly added to the scientific content of the paper and look forward to another round of discussion, if needed.

Our responses to referee #1's specific comments are as follows:

(1) The uncertainties in using OMI threshold values to interpret TROPOMI FNRs are briefly mentioned. However, despite similar signal-to-noise ratios, TROPOMI still shows disagreement with OMI, and it would be good to see some discussion on how this difference, along with TROPOMI bias, may affect the study results.

Response

We agree that these topics should be addressed in the manuscript.

Changes to manuscript

We have added a subsection to the data & methodology section entitled “2.4 Analysis of data composites”. In this subsection we note that the primary method of analysis is through taking the difference between composite categories and providing a qualitative interpretation of what that means in terms of ozone chemistry sensitivity; we can then determine if the changes are due to changes in HCHO or NO₂ levels, or both.

Because we use the J20 threshold values as an additional analysis tool to interpret our TROPOMI derived FNRs, we also highlight some of the disagreements between TROPOMI and OMI data, as well as assumptions and uncertainties from the J20 study, and how these factors may affect the study results.

This new subsection can be found in the revised manuscript on lines 207–223.

(2) The unresolved biases and noises in FNR would be amplified as opposed to using HCHO and NO₂ columns individually. The equation for uncertainty propagation was not properly implemented for division in Appendix B. For division ($z=xy$): $\delta z = \sqrt{(\delta[x][x])^2 + (\delta[y][y])^2} |z|$. More discussion on retrieval errors can be found in Souri et al. 2022 (<https://acp.copernicus.org/preprints/acp-2022-410/>).

Response

We thank the referee for pointing out the error in the equation.

Changes to manuscript

We have removed the equation from appendix section and instead added a subsection to the data & methodology section entitled “2.1.1 Errors associated with FNRs derived from S5P TROPOMI data”. In this section, we reference the Souri et al. (2023) paper to provide a more detailed discussion of FNR errors and discuss what these errors imply for the FNRs we calculated for the Lake Michigan region. We reference Souri et al. (2023) eq. 15, which is the correct equation as the referee pointed out.

This new subsection can be found in the revised manuscript on lines 141–152.

(3) For weekday/weekend analysis, there seem to be many more days selected for weekdays (327 days) than weekends (132 days). Would the result be impacted by averaging over more days? It may be necessary to test if the same number of days were selected.

Response

We agree that it is best to test if the number of days selected impacts the results.

Changes to manuscript

We tested using only Tuesdays/Wednesdays in the weekday composites and Saturdays/Sundays in the weekend composites. Both composites were created with the same number of days over the 3-year period (114 days). The analysis of these new composites is largely the same as before. We replaced the old plots (created with unequal number of days) with the new weekday/weekend composites (created with equal number of days) in the revised manuscript (Figures 7–10 in the revised manuscript).

(4) It would be interesting to see how different regions are similar/different in changes on higher-ozone days. Is there any broader implication of this study on similar urban environments? For example, Tao et al. 2022 (<https://pubs.acs.org/doi/full/10.1021/acs.est.2c02972>) compare TROPOMI HCHO, NO₂, and FNRs on ozone exceedance days versus non-exceedance days and weekdays versus weekends, for summer 2018 over New York City.

Response

We agree that comparisons to similar studies increases the scientific content of this paper.

Changes to manuscript

We added a paragraph talking about how our study is similar to Tao et al. (2022), which can be found on lines 379–383 in the revised manuscript. We end the paragraph with the following to highlight potential broader implications:

“These similarities suggest that the results presented here are broadly applicable to other coastal urban environments with O₃ exceedance problems. Future work could investigate FNRs over Detroit, Michigan, and Los Angeles, California, to see if this implication is true.”
(lines 381–383 in the revised manuscript)

Specific comments

P3. Line 89. The definition of “typical O3 season days” and “exceedance days” could be moved from P3. Line 104 to here (or briefly mentioned), as they appear for the first time.

Response

We agree with this suggestion.

Changes to manuscript

We moved the definitions of the “ozone season” and “exceedance days” to the introduction when the terms first appear. (line numbers 103–107 in the revised manuscript)

P4. Line 116-122. It is controversial to conclude that the errors affecting HCHO and NO₂ retrievals can be canceled out rather than amplified by using their ratio. Please see more in the general comments.

Response

We agree with this comment.

Changes to manuscript

We added a subsection to the data & methodology section entitled “2.1.1 Errors associated with FNRs derived from S5P TROPOMI data” to provide a more detailed discussion of FNR errors. Please see our response to the related general comment above for more information.

P7. Section 2.2. What are the uncertainties in the PAMS surface measurements?

Response & changes to manuscript

To address the specific comments from Referee #2, we have decided to entirely remove the sections involving PAMS surface measurements and focus on analyzing TROPOMI and meteorological data composites.

P9. Line 233-235. This sentence on diurnal cycles seems confusing. Not sure how having diurnal information would make a difference in the current observations. What time of the day were the 6-day interval HCHO measurements collected? Or is it daily mean?

Response & changes to manuscript

To address the specific comments from Referee #2, we have decided to entirely remove the sections involving PAMS surface measurements and focus on analyzing TROPOMI and meteorological data composites.

P12. Line 309-311. May examine the mean temperatures for each composite and verify whether the higher-ozone days co-occur with hotter temperatures.

Response

We agree with this suggestion and believe it would add scientific content to the paper.

Changes to manuscript

We added 2-meter temperature data composite images for comparison. The data come from the NCEP North American Mesoscale (NAM) 12 km Analysis (ds609.0 | DOI: 10.5065/G4RC-1N9 | <https://rda.ucar.edu/datasets/ds609.0/>). These composites confirm our original suggestion that higher ozone days co-occur with hotter temperatures. These composites are Figures 2d–f and 7d–f and in the revised manuscript.

Technical comments

P2. Line 36. Could say “exceed the NAAQS” as O₃ is mentioned three times in this sentence.

Response

We agree with this suggestion.

Changes to manuscript

We removed “O₃” from before NAAQS so that the sentence does not sound too repetitive.

P6. Line 164. 15 individual monthly composites?

Response

Originally, we created 15 individual monthly composites (May-September = 5 months) for 2019–2021 (3 years) $\rightarrow 5 \times 3 = 15$ composites. We then combined them using a weighted average to get an overall 2019-2020-2021 mean O₃ season composite. This was originally done to reduce processing time while creating the composites.

However, to address comments by referee #2, we decided to re-process the TROPOMI data. During this re-processing, we made our compositing script more efficient, which allowed us to create a single 2019–2020–2021 mean O₃ season composite. We have confirmed that taking the weighted average of individual composites and creating a single composite of all the data produce the same average values.

Changes to manuscript

We have removed any mention of “individual monthly composites” from the revised manuscript and updated the data & methodology section to reflect the changes in the satellite data processing.

P6. Line 166. “Next, we ~~next~~ created”

Response

We thank the referee for catching this grammatical error.

Changes to manuscript

We have removed the second “next” as suggested.

Response to Referee #2

Acdan et al., 2022 leveraged the satellite observations of HCHO and NO₂ columns retrieved from TROPOMI radiance and a ground-based monitoring station to contrast the underlying ozone regimes in a region undergoing high ozone exceedances in different episodes such as weekday vs. weekends and ozone exceedance days vs. seasonal averaged values. They observed higher NO₂ columns over Chicago during high ozone exceedances, but its dominantly VOC-sensitive regime did not change due to apparent enhancements in HCHO columns. They observed the typical weekday/weekend tendencies in the former ozone studies. The PAM measurements revealed higher FNRs than those of TROPOMI due to differences in sampling time and inherit column-to-surface discrepancies (Jin et al., 2017). Unfortunately, the scientific content of the paper is really thin; there are artifacts associated with HCHO retrievals; some assumptions about the thresholds were not well thought out; the paper does not inform about the driving factors of the PAMS vs. the satellite discrepancies, and the time period of the case study (during the lockdown) is poorly chosen. The paper also has repetitive analyzes, such as recycling the spatial distributions of HCHO and NO₂ in the shape of histograms that do not provide new content (they could have been presented in SI). The paper clearly does not reach the ACP standard; thus, I recommend rejection.

Introductory comment

We thank referee #2 for providing thorough feedback on our manuscript. Based on both referees' comments, we have made major revisions to the paper, including:

1. Changing the title of the manuscript to “Examining TROPOMI formaldehyde to nitrogen dioxide ratios in the Lake Michigan region: implications for ozone exceedances”
2. Removing all text/figures/references/etc. relating to PAMS data
3. Re-processing the data composites, specifically:
 - a. Using the reprocessed TROPOMI NO₂ PAL dataset so that all data come from the same processor version
 - b. Removing the use of detection limit thresholds
 - c. Addressing the HCHO artifact over water through a bias-correcting approach
 - d. Using the same number of days for the TROPOMI weekday-weekend composites
4. Adding new 2-meter temperature composites from the NAM analysis dataset
5. Expanding on the discussions of FNR errors (e.g., citing Souri et al., 2023), the usage of the J20 thresholds, and comparisons to similar studies (e.g., Tao et al., 2022)
6. Moving some of the appendices to a supplemental information document along with new figures/tables

We believe that these changes have greatly added to the scientific content of the paper and look forward to another round of discussion, if needed.

Our responses to referee #2's specific comments are as follows:

Major comments

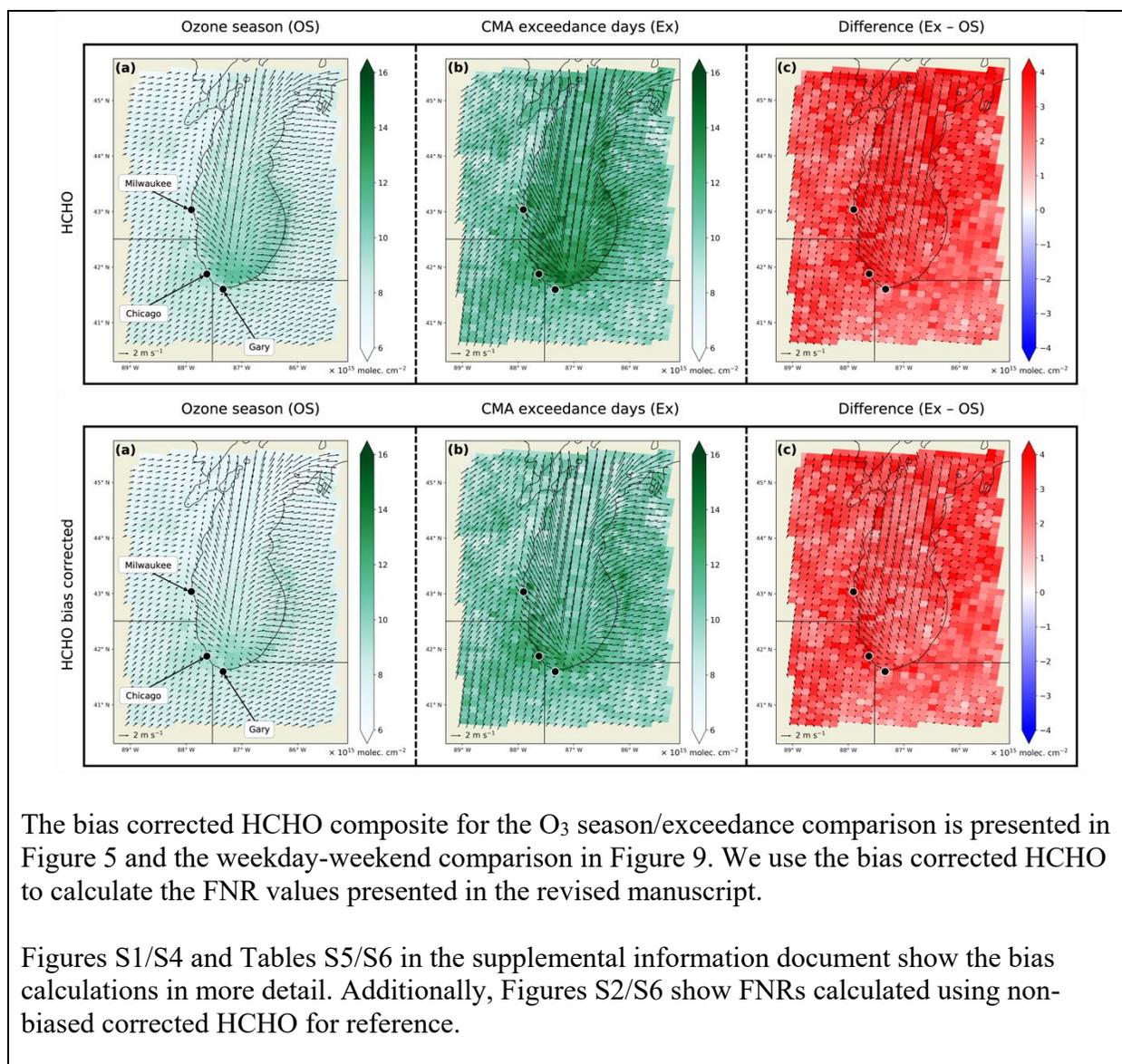
HCHO artifact: Figures 4 and 10 show elevated HCHO concentrations over Lake Michigan that are nonsensical. The surface albedo treatment in the TROPOMI HCHO retrievals most likely causes this artifact. The atmosphere cannot work in that way such that we see a sharp contrast in a relatively spatially homogenous compound like HCHO between land and water. The transport pattern shown in the draft indicates an outflow originating from the lake to the surrounding areas, so the lake will not act as a reservoir to accommodate the transported HCHO. As a result, the statistics regarding HCHO and the ratio (such as the percentage of each underlying ozone regime) are unrealistic. If the authors disagree with me, they should scientifically prove that such elevated HCHO values can prevail over the lake. Do you see the same tendency using a CTM model over the same area (e.g., <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2022JD037042>)? If yes, please break down the physiochemical processes to determine the major driver; I am very doubtful about the quality of TROPOMI HCHO over water especially lakes with complex surface albedo properties unresolved in 0.5x0.5o OMI albedo climatology used in TROPOMI HCHO retrieval.

Response

We agree that the higher HCHO over Lake Michigan is likely an artifact.

Changes to manuscript

In the manuscript, we first present the HCHO composite (Figure 4) and acknowledge the lake artifact (including discussing the fact that HCHO is a relatively well-mixed gas with no sources over the water surface and the coarse OMI surface albedo climatology used in the retrieval). Then we apply a “bias correcting” procedure by assuming that the mean HCHO over land should be equal to the mean HCHO over water. To calculate the bias, we subtract the mean over land HCHO VCD from the mean over water HCHO VCD. Finally, we subtract this bias value from all water grid box values. As a quick example, below shows Figure 4 (non-bias corrected) and Figure 5 (bias corrected):



J20 assumptions: The analysis heavily relies upon the thresholds defined in J20, whose application for this case study is questionable. Two central problems exist 1) J20 thresholds are not intended for understanding the sensitivity of PO₃ to NO_x and VOC but rather for understanding the sensitivity of maximum peak in ambient O₃ concentrations to its precursors. Ambient O₃ levels can be largely impacted by physical processes such as dry deposition, transport, etc. These two sensitivities will not be the same. J20 thresholds are case-study specific and only applicable to their time period/location because the physical processes (i.e., transport, deposition, ...) can vary greatly from time to time. 2) J20 focused on OMI data possessing significant dispersions in HCHO columns (De Smedt et al., 2021) as opposed to those of TROPOMI. The spatial representation between these two sensors is also different. As the retrieval algorithm is a major source of error in the ratio, the fuzziness in J20 thresholds was induced by the errors in OMI that are largely different from those in TROPOMI. The authors must have re-calibrated J20 thresholds by establishing the same relationship between max O₃ and TROPOMI HCHO and NO₂ columns over their region of interest. Also, please avoid mixing up different thresholds from different studies looking at different things. For example, Schroeder et al. 2017 focused on aircraft observations that are not necessarily applicable to the columnar ratio. J20 studied ambient ozone concentrations instead of PO₃. Duncan et al. 2010 used a CTM realization subjective to assumptions made for chemical mechanisms and physical processes. Comparing these numbers is apple-to-orange.

Response

We agree that the use of the J20 thresholds deserve more thorough discussion within the manuscript. We still believe insights can be gained from using the J20 thresholds, particularly because we are primarily interested in providing a qualitative picture of the spatial differences in ozone sensitivity between O₃ season days and exceedance days and identifying the causes of those changes on the most polluted days.

We believe that re-calibrating the thresholds by establishing the same relationship between high O₃ and TROPOMI precursors is out of the scope of this work.

Changes to manuscript

In the introduction section, we specify that the J20 thresholds describe sensitivity of high O₃ levels to precursors as opposed to O₃ production (lines 86–92 in the revised manuscript).

We have expanded the discussion of the use of the J20 thresholds in a newly added section called “2.4 Analysis of data composites”. In this section, we mention that because the J20 thresholds describe high O₃ sensitivity, they are less robust. Then we discuss some of the differences between OMI and TROPOMI. We acknowledge that these differences impact our interpretations ozone chemistry sensitivity when we apply the J20 OMI-based thresholds to TROPOMI FNRs (lines 207–223 in the revised manuscript).

Finally, we write in the limitations section that our interpretations of the changes in O₃ chemistry sensitivity between composite categories are best viewed through a qualitative lens (lines 489–492 in the revised manuscript).

Additionally, we have removed the PAMS surface data section and no longer refer to other thresholds specifically when talking about results (e.g., Schroeder et al. 2017, Duncan et al. 2010).

PAMS's loneliness: The authors briefly showed the contrast between the columnar observations and the surface ones in Section 3.1. They came to the conclusion that various thresholds should be used to segregate chemical conditions using satellite vs. surface observations because they saw a large offset in the PAMS FNRs. This argument is oblivious to the fact that these two datasets look at two different areas, one at the surface layer and the other one within columns, so even if we assumed a universal threshold, the underlying chemical regime would be totally different between those two regions. See Jin et al., 2017 who carefully studied the column-to-surface conversion for different areas/times. The authors could have potentially applied a conversion factor to look at the same layer. Moreover, this section is fully detached from the rest of the study. How did PAMS data look like for the weekday/weekend and ozone exceedances days/normal warm days, i.e., the rest of the paper? What can we really learn from this point measurement that TROPOMI cannot offer? Just showing the ratio difference between the surface and the column is not new; it has been carefully studied in more detail by Jin et al., 2017 and Schroeder et al., 2017 with more suitable tools and data.

Response

We agree that the differences between surface and column ratios are better investigated in Jin et al. (2017) and Schroeder et al. (2017). We also agree that these analyses are detached from the rest of the paper.

Changes to manuscript

We have decided to entirely remove the sections involving PAMS surface measurements and focus on analyzing TROPOMI and meteorological data composites.

Covid-19 time period and re-gridding: The study aimed to diagnose the chemical conditions for emission regulations; I wonder why the authors chose the covid-19 period when there were unusual disruptions in the emissions. What we can potentially learn from these ratios may not be applicable for a regular year. Also, an important advantage of using TROPOMI lies in its high spatial resolution. It is disappointing that the authors picked a 12x12 km² resolution for their analysis, while TROPOMI offers more spatial variance within this grid.

Response

We chose to use the 3-year period between 2019 and 2021 so that the O₃ exceedance day composite was created with more data, making it more statistically robust.

We used the 12 × 12 km² grid based on a sensitivity test of using a 4 × 4 km² grid, which produced a very noisy HCHO composite and a resulting noisy FNR composite. Using the less noisy HCHO composite on the coarser grid was preferable since we wanted to assess the general spatial patterns in FNR values and ozone sensitivities.

Changes to manuscript

In the newly added section “2.4 Analysis of data composites”, we acknowledge that the 3-year period includes years impacted by the COVID-19 pandemic (lines 207–223 in the revised manuscript). Throughout the paper, we discuss any evidence of a pandemic signal in our composites by looking at individual years (e.g., lines 286–288 and lines 441–444 in the revised manuscript; Figures S3 and S5 in the supplemental information). We do note, however, that the spatial patterns in FNR values and the associated ozone sensitivities appear generally consistent among all 3 individual years. In the summary and conclusions section, we also point to another study (Jing and Goldberg, 2022, <https://doi.org/10.1016/j.apr.2021.101313>) which found that meteorology (and not just NO_x emissions decreases alone) explains much of the differences between O₃ production in Chicago in 2020 and the preceding years (lines 526–528 in the revised manuscript).

In the satellite data processing section, we provide the explanation given above for why we use a grid that is coarser than the TROPOMI pixel footprint (lines 169–171 in the revised manuscript).

The inability to explain the differences in concentrations: One of the potentially interesting tendencies observed from TROPOMI NO₂ is the larger NO₂ concentrations over Chicago in high ozone exceedances. This certainly deserves a more thorough discussion using EPA surface monitoring network, bottom/top-down emissions, or available CTMs. Another possible explanation that could have been easily vetted was to study the fraction of the number of weekdays/weekends for this episode. In terms of HCHO, the authors could use parametrized isoprene emissions (e.g., MEGAN) to potentially single out the biogenic contributions. There are also well-established studies performing a temperature-dependency adjustment to minimize the meteorological effect (e.g., Shen et al., 2019). Explaining tendencies adds value to the paper, not mapping out the data.

Response

While we agree that the suggestions above for future work would provide great context to our findings, we believe most of them are out of the scope of this specific work.

Changes to manuscript

For NO₂, we provide the following explanation for higher NO₂ along the shoreline between Chicago and Milwaukee during exceedance days:

“The increased NO₂ VCDs on exceedance days found along coastline between Milwaukee and Chicago can be partially explained by the stronger convergence of the wind field, which concentrates emissions originating in these areas along the southwestern shore of Lake Michigan.” (lines 281–284 in the revised manuscript)

Additionally, we write:

“Further research is needed to determine why NO₂ VCDs are higher for the whole domain during exceedance days (e.g., examining emissions inventories/datasets, looking for temperature dependent natural sources of NO_x, etc.).” (lines 284–286 in the revised manuscript)

For HCHO, we have added 2-meter temperature composites (Figure 2d-f) that provide evidence for our suggestion that higher temperatures lead to more biogenic HCHO emissions (and thus higher HCHO VCDs) on exceedance days.

Repeatability: The manuscript repeats the same tendencies observed from spatial distribution maps by plotting histograms which can be moved to the SI. You can briefly mention whether the differences are statistically significant in one or two sentences. This task could also be better executed by taking a different part of the distribution, like what was done beautifully by Lin et al., 2015 (<https://www.nature.com/articles/ncomms8105>). In general, two things can degrade the quality of a paper: i) repeating what other people have already done and ii) repeating the same results with a different presentation (aka fillers). There are many aspects pertaining to the analysis that deserves deeper analysis. More in-depth studies can be found related to this region's ratio and chemistry (e.g., Abdi-Oskouei et al.).

Response

We thank referee #2 for their suggestions regarding the K-S testing and histogram plots.

Changes to manuscript

We have removed all histogram plots from the manuscript. Additionally, we have adopted the K-S testing procedure by Lin et al. (2015) as suggested. We added a subsection called “2.4.1 Significance testing” to describe the methodology (lines 224–232 in the revised manuscript). We describe the K-S test results for each variable/composite in their relevant sections and provide a summary of the results in Tables S3 and S4 in the supplemental information.

Specific Comments:

L50. You mentioned two regimes, but you will define three ones.

Response

We thank referee #2 for pointing out this mistake.

Changes to manuscript

We restructured the introduction paragraphs talking about ozone chemistry regimes. Additionally, we added more details regarding ozone production in general to address this comment and many of the following comments as well. The paragraph in which the regimes are discussed can be found on lines 50–60 in the revised manuscript.

L53. HO₂ needs to be accounted too.

Response

We agree that more information should be provided regarding the ozone production chain reactions.

Changes to manuscript

We included more details about ozone production reactions in the introduction, including those involving HO₂. This can be found on lines 43–48 in the revised manuscript.

L54. What type of non-linear chemistry? Please elaborate.

Response

We agree that we can elaborate further.

Changes to manuscript

We included more details about ozone production reactions and chemistry sensitivities in the introduction. This can be found on lines 43–60 in the revised manuscript.

Line 54-55. The definition of NO_x-sensitive or VOC-sensitive regimes is irrelevant to the availability of free oxygen atoms. In NO_x-sensitive conditions, PO₃ is reduced due to decreased [NO][RO₂] and [NO][HO₂] because all terms are reduced. [RO₂] and [HO₂] are efficiently removed in NO_x-sensitive conditions, yielding H₂O₂. In rich NO_x regions, so much NO_x is available that terminates OH/HO₂ cycling (the RO_x cycle) through NO₂+OH. You need to involve the RO_x-HO_x cycle in this paragraph. It may also be advantageous to talk about OPEs (how much O₃ is produced per NO_x molecule), which vary from NO_x-sensitive (high OPE) to VOC-sensitive (low OPEs) conditions.

Response

We agree that we should be more descriptive when talking about ozone production chemistry sensitivities.

Changes to manuscript

We included more details about ozone production reactions and chemistry sensitivities in the introduction. This can be found on lines 43–60 in the revised manuscript. However, we do not talk about OPEs because we do not want to make the introduction section too long.

L55-60. Jin and Holloway, 2015 are not the founders of chemical condition labels. Please use a better reference, such as Sillman et al., 2002 or Duncan et al., 2011.

Response

We thank referee #2 for this suggestion.

Changes to manuscript

We adjusted the references for the chemical sensitivity labels, which can be found on lines 50–60 in the revised manuscript.

L61-62. But didn't he conclude that H₂O₂/HNO₃ was the most viable indicator fully describing the HO_x-RO_x cycle?

Response

Yes, Sillman (1995) did conclude that H₂O₂/HNO₃ was one of the most robust indicator ratios.

Changes to manuscript

We added a discussion referencing the above fact, but also acknowledging that H₂O₂ and HNO₃ levels/VCDs are not regularly measured/observed. We then transition to talking about HCHO/NO₂, the indicator we use in this work. These changes can be found on lines 69–78 in the revised manuscript.

L62. HCHO is not a proxy for VOC concentrations. It is a proxy for VOC reactivity.

Response

We thank referee #2 for pointing out this wrong use of terminology.

Changes to manuscript

We have changed the wording to say “VOC reactivity” on line 74.

L68. We shouldn't rule out the importance of H₂O₂/HNO₃.

Response

We thank referee #2 for bringing up the importance of H₂O₂/HNO₃

Changes to manuscript

We added an additional discussion referencing the above fact, but also acknowledging that H₂O₂ and HNO₃ levels/VCDs are not regularly measured/observed. We then transition to talking about HCHO and NO₂ since they are measurable from space. These changes can be found on lines 69–78 in the revised manuscript.

L67. But NO_y can provide information on how transported NO_x from far areas can affect local PO₃. I don't think it's necessarily a weakness.

Response

We thank referee #2 for pointing this out.

Changes to manuscript

We have changed the phrasing from “more useful” to “another useful” so that it does not seem like we are suggesting that HCHO/NO_y is not a useful indicator:

“Building upon Sillman's work, Tonnesen and Dennis (2000) found that HCHO/NO₂ (“FNR” for the rest of this paper) is **another useful indicator** of ozone–NO_x–VOC sensitivity since HCHO and NO₂ have similar lifetimes (on the order of hours).” (lines 75–77 in the revised manuscript)

L79-81. What do you mean by avoiding? They ignored the critical fact that PO₃ is not equal to O₃. O₃ can easily get impacted by meteorology and dry deposition, which are not informed by the ratio. Please rewrite this part.

Response

We agree that further discussion about the J20 study and the use of thresholds in our study is warranted.

Changes to manuscript

Please see our response to referee #2's major comment about the J20 thresholds above for the changes we made to the manuscript.

Table1. These thresholds do not define the regimes you defined earlier. They are not directly related to PO₃. What is the definition of VOC-sensitive from an ambient O₃ concentration perspective? You should carefully describe the assumption J20 made and its major limitations.

Response

We agree that further discussion about the J20 study and the use of thresholds in our study is warranted.

Changes to manuscript

Please see our response to referee #2's major comment about the J20 thresholds above for the changes we made to the manuscript. Additionally, we have changed the table caption to say: "J20 FNR threshold values indicating different **high O₃** chemistry sensitivities for Chicago, Illinois, U.S." so as to not confuse this with PO₃. (line 96 in the revised manuscript)

L85. This time period is during the lockdown. How informative is the case study for a normal year?

Response

We believe that our results are still applicable to other years. As mentioned above in our response to the major comment, the same general pattern of FNR values and inferred ozone chemistry sensitivities is seen in individual composites for 2019, 2020, and 2021.

Changes to manuscript

We highlight any differences in the composites among the individual years. Please see our response to referee #2's major comment about the COVID-19 period for more details.

L101. Why do you need both versions?

Response

We thank referee #2 for this question; we have re-made the composites (see below).

Changes to manuscript

We re-made our composites using S5P PAL TROPOMI NO₂ data (<https://data-portal.s5p-pal.com/products/no2.html>), which is a harmonized dataset for NO₂ from 2018–2021 using the same processor version (thus removing the discontinuity).

We still had to use V1 and V2 of HCHO data product because no harmonized HCHO product exists that contains our entire study period (processor changed version in July 2020). However, the changes between versions of the HCHO product are not as drastic as the changes for NO₂.

L115. Errors in AMFs also contribute to the total error.

Response

We thank referee #2 for pointing this out.

Changes to manuscript

We have removed the end of the sentence so that it states:

“The total uncertainty in HCHO tropospheric vertical column density retrievals is currently estimated to be between 30–60 % in polluted conditions”. (lines 137–138 in the revised manuscript)

L120. What assumption did they make to say that? The surface albedo and aerosol effects can vary between 340 and 440 nm.

Response

We thank referee #2 for the question.

Changes to manuscript

We have revised our discussion of errors by adding a subsection to the data & methodology section entitled “2.1.1 Errors associated with FNRs derived from S5P TROPOMI data”. In this section, we reference Souri et al. (2023) to provide a more detailed discussion of FNR errors and discuss what these errors imply for the FNRs we calculated for the Lake Michigan region.

This new subsection can be found in the revised manuscript on lines 141–152.

L121. The correlated term should be “ $-2\text{cov}(\text{HCHO}, \text{NO}_2)/(\text{HCHO}\times\text{NO}_2)$ ”. So if HCHO and NO₂ retrievals are positively correlated, they will only reduce the total relative errors when either NO₂ or HCHO values are low. The correlated term will likely be small in polluted areas where HCHO and NO₂ are elevated.

Response

We thank referee #2 for pointing this out.

Changes to manuscript

We have revised our discussion of errors by adding a subsection to the data & methodology section entitled “2.1.1 Errors associated with FNRs derived from S5P TROPOMI data”. In this section, we reference Souri et al. (2023) to provide a more detailed discussion of FNR errors and discuss what these errors imply for the FNRs we calculated for the Lake Michigan region.

This new subsection can be found in the revised manuscript on lines 141–152.

L128. I am not sure if I agree with the discussion about SNR. SNR has a specific definition related to the instrument specifications and the observed radiance. HCHO retrieval is inherently inferior because its optical depth (despite being higher than NO₂) is located in the UV range where Rayleigh scattering and O₃ absorption prevail, resulting in a less robust spectral fitting.

Response

We agree with this comment.

Changes to manuscript

We have re-worded the phrasing to be more accurate & specific:

“However, HCHO has an optical density that is an order of magnitude smaller than that of NO₂ because the spectral band its retrieval is derived from is in the UV range where Rayleigh scattering and ozone absorption occur (De Smedt et al., 2018). As a result, individual HCHO retrievals are noisier than NO₂ retrievals.” (lines 156–159 in the revised manuscript)

L135. The detection limit is sensor/retrieval specific; those studies are not applicable. Why not use TROPOMI studies? De Smedt et al. 2021 say 3×10^{15} molec.cm⁻² for TROPOMI, which is an improvement of a factor of 2 compared to OMI.

Response

We thank referee #2 for pointing this out. However, we no longer view detection limit filters as necessary because we are compositing the data on longer time scales, and other studies do not employ them when using TROPOMI data (e.g., see <https://pubs.acs.org/doi/full/10.1021/acs.est.2c02972>, <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020EF001665>, <https://acp.copernicus.org/articles/23/1963/2023/>)

Furthermore, the minimum value of all HCHO composites presented is greater than the 3×10^{15} molec.cm⁻² detection limit reported by De Smedt et al. (2021). To our knowledge, no published paper exists that reports a detection limit for TROPOMI NO₂ tropospheric vertical column density.

Changes to manuscript

We removed the use of detection limit filters during the re-processing of the data.

L135. Also, I am unsure if I agree that the SNR is the same between OMI and TROPOMI. What does the literature say? When comparing SNRs, we should account for the footprint, so you have to normalize it by pixel size.

Response

We agree that comparing SNRs should consider footprint size.

Changes to manuscript

Because we are no longer using detection limit thresholds, we remove any mentions of OMI and TROPOMI having similar SNRs.

L154. Why do you degrade TROPOMI spatial variance by upscaling it to 12x12 km² when it provides higher spatial information?

Response

We thank referee #2 for the question.

Changes to manuscript

Please see our response the major comment about the choice of using a coarser grid.

L174. What are the weights? The spatial response function?

Response

We thank referee #2 for these questions.

Changes to manuscript

During the re-processing of our data, we made our compositing script more efficient, which allowed us to create singular 2019–2020–2021 mean composites as opposed to monthly/yearly ones. We have confirmed that taking the weighted average of individual composites and creating a single composite of all the data produce the same average values.

Section 2.2. Please provide the errors associated with PAM measurements. Also, because TROPOMI captures one snapshot, can we rely on monthly-averaged samples from in-situ measurements? Large diurnal variability is associated with HCHO and NO₂, which is not resolved in PAMS.

Response

We thank referee #2 for the question and suggestion.

Changes to manuscript

As mentioned in our response to the major comment above, we have removed all sections regarding PAMs measurements.

L258. Those thresholds are not necessarily related to satellites. So I don't think you should put all of them in one basket.

Response

We thank referee #2 for the suggestion.

Changes to manuscript

As mentioned in our response to the major comment above, we have removed all sections regarding PAMs measurements.

L289. Some hypotheses based on previous works?

Response

We thank referee #2 for the question.

Changes to manuscript

Please see our response to the major comment regarding higher NO₂ on exceedance days.

L311. Does an increase in biogenic VOC always lead to higher O₃? I think you are trying to say here about the relationship between O₃ and increased temperature. See Figure 8 at <https://pubs.acs.org/doi/full/10.1021/cr5006815>. You shouldn't rule out the effect of RO₂NO₂. Can you show the 2m air temperature difference too?

Response

We are not saying that increases in biogenic VOC always lead to higher O₃. We are pointing out that on Chicago exceedance days, TROPOMI HCHO VCDs are higher, suggesting higher biogenic VOC emissions.

Changes to manuscript

We rephrased the sentence to:

“Because positive differences occur over the entire domain, the higher HCHO abundances are likely due to increased temperatures during O₃ exceedance events (Fig. 2f), which lead to increased biogenic VOC emissions and thus increased O₃ production in regions with VOC-sensitive chemistry (Sillman and Samson, 1995).” (lines 327–330 in the revised manuscript)

We provided 2-meter temperature composites as Figures 2d–f.

L365. This is a generic tendency you will observe in any city worldwide. As NO_x dilutes far from the sources, the chemical condition becomes less VOC-sensitive.

Response

We thank referee #2 for the comment.

Changes to manuscript

We have not made any changes based on this comment.

L409. I don't understand the connection between HCHO and thermal gradients. Why don't we look into air temperature from a model?

Response

We thank referee #2 for the question/suggestion.

Changes to manuscript

We provided 2-meter temperature composites as Figures 2d–f.

L410-414. If this is true, why is HCHO larger over the lake than the land? See my major comment.

Response

We thank referee #2 for the question.

Changes to manuscript

Please see our response to the major comment regarding the over water HCHO artifact.

Figure 8. I'm surprised by the KS test saying that the distributions of NO₂ are statistically different. How many times have the tests been done? Are they done on the total distribution or a specific part of it? Please see the analysis nicely done at <https://www.nature.com/articles/ncomms8105>. I really don't see them being too different.

Response

We thank referee #2 for the questions and suggestion.

Changes to manuscript

We have removed all histogram plots from the manuscript. Additionally, we have adopted the K-S testing procedure by Lin et al. (2015) as suggested. We added a subsection called “2.4.1 Significance testing” to describe the methodology (lines 224–232 in the revised manuscript). We describe the K-S test results for each variable/composite in their relevant sections and provide a summary of the results in Tables S3 and S4 in the supplemental information. Our new results indicate a significant difference in NO₂ VCDs between the O₃ season and exceedance days.

Figure8. What do we learn from these histograms that were not presented in the previous plots? I feel like the authors repeat the same tendencies. It really doesn't add new information.

Response

We thank referee #2 for this comment.

Changes to manuscript

We have removed all histogram plots from the manuscript.

L465. This is too speculative, given the HCHO artifact. Also, how sure are we that isoprene emissions behave similarly in two episodes?

Response

We thank referee #2 for this comment.

Changes to manuscript

Our new results (Figure 9) clearly show higher HCHO over the land and water in the southern part of the domain on weekends. In addition to providing a hypothesis for why this might be happening, we say future research is needed to find causes for our finding here.

L563. What do you mean by saying that ozone production occurs throughout the day? There is no production at nighttime.

Response

We meant to say that O₃ sensitivity to precursors can change hourly, which is not captured by the once daily data provided by TROPOMI.

Changes to manuscript

We have rephrased the sentence to say:

“However, the sensitivity of O₃ levels to NO_x and VOCs can change as the atmospheric concentrations of these gases change on shorter timescales (e.g., hourly).” (lines 483–485 in the revised manuscript).

Last paragraph in conclusion: Please always provide aspects that your analysis has focused on. Your study did not quantify the temporal representation errors to gauge the importance of TROPOMI vs. GEO satellites. This paragraph is just a filler with no relevance to the results.

Response

We thank referee #2 for the comment. While we agree that we did not quantify the temporal representation errors to gauge the importance of TROPOMI vs. GEO satellites, we believe it is important to mention the how the upcoming geostationary satellites will provide future opportunities to conduct FNR research with new datasets.

Changes to manuscript

We have shortened this last paragraph to:

“Future geostationary satellite instruments, such as the NASA Tropospheric Emissions: Monitoring of Pollution (TEMPO) set to launch in 2023 (Zoogman et al., 2017) and the ESA SENTINEL-4 set to launch in 2024 (Gulde et al., 2017), will make measurements of HCHO and NO₂ in hourly intervals over the United States and Europe, respectively. The datasets produced by these instruments will provide researchers with new opportunities to explore the viability of using satellite-derived FNRs to infer surface ozone–NO_x–VOC sensitivity at unprecedented spatiotemporal scales.” (lines 530–534 in the revised manuscript)

Editorial Comments:

L33. Longer than what?

Response

We thank referee #2 for the question.

Changes to manuscript

We have rephrased the sentence to say:

“Acute exposure to elevated O₃ levels can cause respiratory problems (e.g., asthma attacks) while chronic exposure can lead to premature death from respiratory and circulatory system illnesses...” (lines 32–34 in the revised manuscript)

L116. Please use the right symbol for times instead of x.

Response

We thank referee #2 for the suggestion.

Changes to manuscript

We have replaced all instances of “x” with the correct symbol “×”.

L117. Molec. is better over mol. Please remake all figures and apply this to the text. Mol can be wrongly interpreted as mole.

Response

We thank the referee for the suggestion.

Changes to manuscript

We have replaced all instances of “mol” with the correct symbol “molec.”

Appendixes could be moved to SI.

Response

We thank the referee for the suggestion.

Changes to manuscript

We have removed all appendixes from the revised manuscript and created a new supplemental information document.