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 NO_2 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

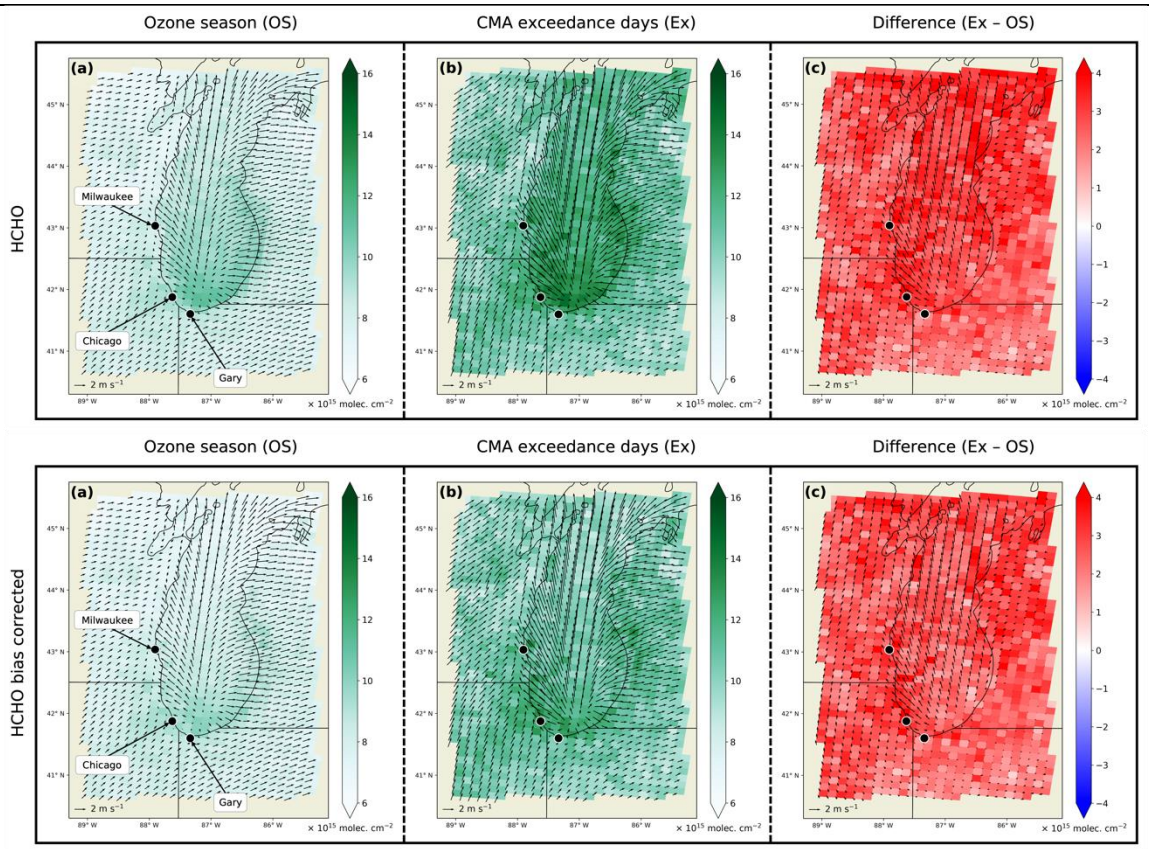
HCHO artifact: Figures 4 and 10 show elevated HCHO concentrations over Lake Michigan that are nonsensical. The surface albedo treatment in the TROPOMI HCHO retrievals most likely causes this artifact. The atmosphere cannot work in that way such that we see a sharp contrast in a relatively spatially homogenous 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):



The bias corrected HCHO composite for the O_3 season/exceedance comparison is presented in Figure 5 and the weekday-weekend comparison in Figure 9. We use the bias corrected HCHO to calculate the FNR values presented in the revised manuscript.

Figures S1/S4 and Tables S5/S6 in the supplemental information document show the bias calculations in more detail. Additionally, Figures S2/S6 show FNRs calculated using non-biased corrected HCHO for reference.

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 NOx 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 O_3 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 O_3 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 O_3 levels to precursors as opposed to O_3 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 O_3 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 O_3 exceedance day composite was created with more data, making it more statistically robust.

We used the 12 × 12 km^2 grid based on a sensitivity test of using a 4 × 4 km^2 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 O_3 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 NO_2, we provide the following explanation for higher NO_2 along the shoreline between Chicago and Milwaukee during exceedance days:

"The increased NO_2 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 NO_2 VCDs are higher for the whole domain during exceedance days (e.g., examining emissions inventories/datasets, looking for temperature dependent natural sources of NOx, 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 HO_2. 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 NOx-sensitive or VOC-sensitive regimes is irrelevant to the availability of free oxygen atoms. In NOx-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 NOx-sensitive conditions, yielding H2O2. In rich NOx regions, so much NOx is available that terminates OH/HO2 cycling (the ROx cycle) through NO2+OH. You need to involve the ROx-HOx cycle in this paragraph. It may also be advantageous to talk about OPEs (how much O3 is produced per NOx molecule), which vary from NOx-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 HOx-ROx cycle?

Response

Yes, Sillman (1995) did conclude that H_2O_2/HNO_3 was one of the most robust indicator ratios.

Changes to manuscript

We added a discussion referencing the above fact, but also acknowledging that H_2O_2 and HNO_3 levels/VCDs are not regularly measured/observed. We then transition to talking about $HCHO/NO_2$, 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 H_2O_2/HNO_3

Changes to manuscript

We added an additional discussion referencing the above fact, but also acknowledging that H_2O_2 and HNO_3 levels/VCDs are not regularly measured/observed. We then transition to talking about HCHO and NO_2 since they are measurable from space. These changes can be found on lines 69–78 in the revised manuscript.

L67. But NOy can provide information on how transported NOx 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 NO_2 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 O_3** chemistry sensitivities for Chicago, Illinois, U.S." so as to not confuse this with PO_3. (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 NO_2 data (https://data-portal.s5p-pal.com/products/no2.html), which is a harmonized dataset for NO_2 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 NO_2.

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 "-2cov(HCHO, NO2)/(HCHO×NO2)". 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 NO_2 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 NO_2 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 × 1015 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 NO_2 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 NO_2 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 O_3. 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 O_3 exceedance events (**Fig. 2f**), which lead to increased biogenic VOC emissions and thus increased O_3 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 NOx 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 NO_2 VCDs between the O_3 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 O_3 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 O_3 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 NO_2 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 O_3 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 appendices from the revised manuscript and created a new supplemental information document.