Response to Reviewer 1's Comments

We appreciate Dr. Maarten Krol for giving detailed, insightful comments. We address the comments below. The comments are shown in normal fonts, while our response to each comment is shown in **bold characters**. Specific changes are highlighted in the track-changes file submitted with the revised manuscript, and these changes and their corresponding line numbers are also provided below.

Reviewer: The comparison shows still large discrepancies with observations. For instance, the ozone profiles presented in Figure 6 show large discrepancies. Moreover, figure 5 shows only averages (bit unclear). Apparently, there is a wealth of information on e.g. diurnal variations of isoprene, but we do not get to see these. I would therefore strongly advise to make a more convincing comparison to observational data.

Response: We appreciate the reviewer's comment highlighting the discrepancies between the model results and observations, and the need for further comparisons with observational data, particularly regarding the chemistry. Indeed, modeling approaches inherently involve uncertainties, necessitating ongoing model development and refinement. The chemical initial and boundary conditions (IC & BC) in our simulations are derived from another global model (i.e. CESM-WACCM), which is commonly used to drive model simulations but could introduce uncertainties. Additionally, emissions inventories come with their uncertainties, and reanalysis meteorological datasets can also contribute to model-observation discrepancies. However, we made every effort to validate the model results against observations and to make the most of the valuable observational data obtained during TRACER and TRACER-AQ, which contributes to the length of this paper.

In the paper, the evaluation of the model chemistry is shown in Figures 4-6 in the manuscript and Figures S3-S6 in the Supplementary Materials, where we explore various observational datasets for a comprehensive evaluation of the model. We acknowledge the reviewer's point regarding the discrepancies in model-observation comparisons for the chemistry profile, particularly for O_3 in Figure 6, which is documented in the manuscript on Lines 414-419 as "*The LES domain D04 overestimates O_3 mixing ratios, while the mesoscale domain D02 underestimates them at both sites. The accuracy of reanalysis meteorological datasets, which are used to drive model simulations, in representing vertical profiles of key meteorological factors and their diurnal evolution may contribute to the comparison discrepancies. In addition, discrepancies between the model and observations may be partially attributed to the time lag between sonde measurements and simulation outputs. Minute-to-minute changes in wind speed and direction can significantly influence the dispersion of air masses and distribution of chemicals (Li et al., 2019), potentially affecting the accuracy of model-observation comparisons.*"

However, the spatial distribution of surface concentrations in Figure 4 shows better model-observation agreement, particularly within the LES domain. The statistical analysis presented in Table S1 shows smaller biases for surface concentrations of O_3 and NO_2 from the LES domain than the mesoscale D02. The comparison of isoprene and other VOCs (Figure 5) relies on modeled and observational data spanning from 8:00 to 16:00 CDT. However, model outputs are analyzed hourly while observations are available every 30 seconds. To address the reviewer's concern regarding diurnal variability, we include a new time series plot showing the diurnal variability of these VOCs in the Supplementary Material of the revised version (now Figure S5) for a more robust comparison. We add the following sentences to the revised manuscript on Lines 378-387: "*Figure S5 compares the simulated diurnal variability of four VOCs with observations from the PTR-MS. Isoprene is underestimated in the LES domain D04 in the morning, but increasing concentrations are simulated in the afternoon, reducing biases and aligning more closely with observations. The simulated diurnal pattern of HCHO in the LES domain D04 shows better agreement with observations. MEK is overestimated in the morning across both the LES and mesoscale domains, but modeled concentrations become more aligned with observations in the*

afternoon, with both observations and modeled results indicating high levels in the morning and decreasing concentrations throughout the day. Acetaldehyde, however, is underestimated in the afternoon in both domains, suggesting missing secondary chemical sources from lumped alkanes (Read et al., 2012). It should be noted that the PTR-MS measurement accuracy may be constrained by uncertainties in calibration standards and compound identification based on mass information only (Ammann et al., 2004)."

Reviewer: From a modelling point of view, the authors take a big step to run on high resolution (300 m). Apart from concluding that on this resolution the model performs better, there is unfortunately little analysis of the reasons why this is the case. Is this because the chemical contrasts get larger, causing more extreme chemical regimes (e.g. larger disruptions of chemistry because of high NOx emissions)? There are some hints that the boundary layer mixing is better resolved on 300 m resolution, but this aspect is not well worked out either.

Response: The differences between LES and mesoscale simulations could be mainly due to turbulence-chemistry interactions. We clarify this in the manuscript in lines 450-455 "*In general, the differences between LES and mesoscale simulations may largely stem from turbulence-chemistry interactions. These interactions can disrupt chemical processes by influencing the production of secondary compounds, which rely not only on NO_X emission variations but also on meteorological factors such transport flow rate and the mixing of primary pollutants. Due to the limited availability of observational data, we do not include model evaluations for turbulence parameters. However, the LES wind speed maps in Fig. 3 suggest higher TKE and stronger mixing of chemicals both horizontally and vertically, resulting in a more robust distribution of chemical species and their subsequent chemical reactions in each grid cell*."

Reviewer: In the end, the authors draw some conclusions that both isoprene is important (e.g. largest OH-reactivity) next to anthropogenic VOC emissions. VOC concentrations basically are not evaluated with observations. Moreover, I was shocked that existing NO_X inventories had to be scaled by ~0.2 to get meaningful results. This hardly receives attention in the abstract and conclusions. How does the model behave if the emissions of anthropogenic VOCs are off by a similar amount?

Response: We thank the reviewer for the comments. The availability of VOC measurements are limited and this data also has uncertainties. We made every effort in the study to validate the model results against VOC observations using available VOC data. The VOC evaluation is shown with a boxplot analysis (Fig. 5) and a new times series in Figure S5, which we discuss in lines 373-387 as "*To better understand O_3 formation, an accurate model interpretation of its VOC precursors is essential. We validate our modeled VOCs using measurements of key VOCs obtained by the PTR-MS at San Jacinto Battleground State Historical site (Fig. 5). The comparison shows that the model adequately reproduces observations, with results from the LES domain D04 more closely matching observed VOCs, particularly for acetaldehyde, isoprene, HCHO, methyl ethyl ketone (MEK), toluene and xylenes. Although acetone is significantly underestimated in the mesoscale D02 simulations, the LES results better reproduce its observed concentrations. Figure S5 compares the simulated diurnal variability of four VOCs with observations from the PTR-MS. Isoprene is underestimated in the LES domain D04 in the morning, but increasing concentrations are simulated in the afternoon, reducing biases and aligning more closely with observations. The simulated diurnal pattern of HCHO in the LES domain D04 shows better agreement with observations. MEK is overestimated in the morning across both the LES and mesoscale domains, but modeled concentrations become more aligned with observations in the afternoon, with both observations and modeled results indicating high levels in the morning and decreasing concentrations throughout the day. Acetaldehyde, however, is underestimated in the afternoon in both domains, suggesting missing secondary chemical sources from lumped alkanes (Read et al., 2012). It should be noted that the PTR-MS measurement accuracy may be constrained by uncertainties in calibration standards and compound identification based on mass information only (Ammann et al., 2004). We also acknowledge that the difference in measurement frequency (30 s) and model averaging time (1 h), as well as the representation of a*

model grid cell by a single measurement station, may contribute to further disparities between the modeled results and field measurements."

The NOx inventory is scaled by ~0.2-0.3 after several model tests. In the paper, we use the 2017 version of the NEI inventory, which is a four-year offset compared to our simulation period. Previous studies also show NEI may often overestimate NO_X concentration, as we highlight in the study in lines 194 – 195: "*Our simulation results with the 2017 NEI show overestimated NO_X concentrations, consistent with findings in previous urban studies (Choi and Souri, 2015; Souri et al., 2016a; Silvern et al., 2019; Herrmann and Gutheil, 2022).*" This is why we scaled the NOx inventory. We acknowledge the reviewer's point that this was not adequately addressed in the abstract and conclusions. For clarity, we revise the scaling to indicate a 70-80% reduction in NO_X. Therefore, we revise the sentence in the conclusion in lines 687-696 as:

"*In this study, we use a high-resolution WRF-LES-Chem to investigate detailed atmospheric photochemistry during a high O_3 episode during TRACER-AQ 2021 in Houston. During this episode, observed surface O_3 significantly exceeded national standards by 20-50 ppb, particularly in the central areas of Houston. NO_X emissions from the anthropogenic NEI inventory are scaled down by a factor of ~0.2-0.3 to achieve more realistic concentrations in our domains, reflecting a 70-80 % reduction in NO_X emissions. By resolving turbulence and eddies at a scale relevant to the chemistry of reactive pollutants, the LES simulation at 300 m spatial resolution provides an improved representation of surface meteorology compared to the mesoscale WRF-Chem simulations, and more accurately captures the surface concentration of O_3, NO_X, and HCHO.*"

Also, we now include a statement on this reduction in this Abstract section on line 20 as "*NO_X emissions from the NEI anthropogenic inventory are reduced by 70-80 % to achieve realistic diurnal variability of NO_X concentrations.*"

Reviewer: In conclusion, the paper presents an interesting analysis of model data in terms of ozone production and chemical reactivity, but the analysis is based on poorly evaluated model results and fails to address the reasons why a high-resolution simulation would be required to analyze valuable data from measurement campaigns.

Response: We thank the reviewer for highlighting the concern about model performance. Although we made every effort in this study to validate the model results against available observations during TRACER and TRACER-AQ, and also generate more robust simulation results especially regarding surface concentrations of key chemical species, we acknowledge the model-observation discrepancies and the limitations of our study, which could be used as a motivation for future works to improve existing modeling efforts in atmospheric chemistry research. We outline the limitations of our study on Lines 721-729 as "*Furthermore, modeling approaches inherently involve uncertainties, necessitating ongoing model development and refinement. Modeling limitations, such as the accuracy of emissions inventories, as well as the chemical and meteorological IC & BC used to drive modeling, can introduce significant biases in chemical simulations. Using compatible high-resolution emissions inventories would better suit fine-resolution simulations, as they would more effectively capture microscale emission sources. Additionally, limitations in observational data, including measurement uncertainties and insufficient data coverage, constrain the ability to fully evaluate and improve model accuracy. However, the model-observation discrepancies identified in this study highlight key areas for future improvement of atmospheric modeling.*"

Reviewer: Referencing: it seems referencing is restricted to inner circle papers. There is a long history of research lines that address the effects of resolution on atmospheric chemistry that is totally missing.

Response: We appreciate the reviewer's observation of the cited articles. The choice of references used in this work is selected without bias to regional context or relationship with other scientists/researchers (inner circle papers); rather based on the application of such study to this

work. As rightly pointed out, many studies exist on the effects of resolution of atmospheric chemistry. We cite papers from this poll of studies that are relevant to our study but we acknowledge we cannot thoroughly cite all the papers. However, we now add citations including Danckwerts (1952), Mousavi et al. (1999), Krol et al. (2000), and Brasseur et al. (2023). Other citations are included in the study such as Ehhalt (1994), Fehsenfeld et al. (1998), Pusede and Cohen (2012), Klemm et al. (2000), and Blanchard and Fairley (2001).

Reviewer: Figure S1: Unclear how to interpret the bottom plots.

Response: The lower panel in Figure S1 is a wind rose diagram and describes how wind speed and direction are typically distributed at a particular location. In the revised manuscript, a legend for the lower panel is now included in the Supplementary Material (Fig. S1) to provide a clearer understanding of the figure, especially in relation to wind speed. A brief explanation is provided in the submitted manuscript on lines 157-164 as

"*Our model results driven by the two datasets show similar spatial distributions of the surface heat budget (Fig. S1). We also show a wind rose diagram which gives a succinct view of how wind speed and direction are typically distributed at a particular location. Observations from TCEQ indicate a prevailing northwesterly wind. The WRF-simulated winds driven by NCEP-FNL meteorology align more closely with the TCEQ observed northwesterly wind compared to HRRR. However, both models show weaker wind speeds than observed. This discrepancy may arise from uncertainties in the NCEP-FNL and HRRR datasets, which are used as meteorological IC & BC for representing near-surface winds in our modeling, as well as from the model's simplified physical parameterization, especially in the mesoscale domains.*"

Reviewer: Figure S5: Panels d-f not in the caption.

Response: The caption of Figure S5 (now Figure S6) is revised as "*Vertical cross sections of the temporal evolution of ozone profiles (ppb) from (a, d) the ground-based Lidar measurements from the NASA TRACER-AQ campaign (figure plotted from https://tolnet.larc.nasa.gov/), (b, e) model results from the mesoscale domain D02, and (c, f) model results from the LES domain D04. The upper left panel represents data from the University of Houston LMOL site (29.724° N, 95.3392° N) available from 12:30 UTC (September 8, 2021) to 00:00 UTC (September 9, 2021), and the lower left panel represents the La Porte TROPOZ site (29.667° N, 95.064° N) available from 00:00 UTC (September 8, 2021) to 00:00 UTC (September 9, 2021). All time units are in UTC to match with NASA LMOL measurements.***"**

Reviewer: Figure S6: Idem: lower panels are not explicitly mentioned in caption.

Response: The caption of Figure S6 (now Figure S7) is revised as "*Vertical cross sections of the temporal evolution (time in CDT) of (a, f) HCHO, (b, g) O_3, (c, h) water vapor, (d, i) temperature and (e, j) potential temperature profiles over the LES domain D04 with a focus on the land area. The upper panel represents simulation from the Mesoscale domain D02; the lower panel is for the LES domain D04.***"**

Other Comments

Reviewer: referencing is very limited. There are other groups that wrote papers about this....(e.g. the Ehhalt group in Julich)....

Response: We thank the reviewer for this reminder. Indeed, there is a wide pool of studies that discuss the NOx-radical cycling. Additional references are now being added including Ehhalt,

1994; Fehsenfeld et al., 1998; Pusede and Cohen, 2012; Klemm et al, 2000; Blanchard and Fairley, 2001. We modify Lines 78-83 as "*In a VOC-limited regime, an incremental increase in VOCs promotes the chemical production of peroxy (RO_2) and hydroperoxyl (HO_2) radicals, which enhances NO_X cycling and ultimately increases ambient O_3 levels (Ehhalt, 1994; Fehsenfeld et al., 1998; Pusede and Cohen, 2012; Ye et al., 2016, Zara et al., 2021). By contrast, in a NO_X-limited regime, increased emissions of NO_X enhance the photolysis of NO_2, thus producing free reactive oxygen atoms, which then react with O_2 to produce O_3, resulting in increased O_3 concentrations (Klemm et al, 2000; Blanchard and Fairley, 2001; Jin and Holloway, 2015).***"**

Reviewer: Again here, the intensity of segregation has a long line of history that needs to be credited, starting with Danckwerts, P. V.: The definition and measurement of some characteristics of mixtures, Appl. Sci. Res. A, 3, 279–296, https://doi.org/10.1007/BF03184936, 1952. a

Response: We thank the reviewer for these suggestions. Additional references, Danckwerts, 1952 and Mousavi et al., 1999 are now added to the manuscript. We revise Lines 104-106 as "*Previous studies suggest that turbulence may lead to reductions of chemical reaction rates in the boundary layer, as a result of negatively correlated concentrations and vertical velocities (Danckwerts, 1952; Mousavi et al., 1999; Krol et al., 2000; Li et al., 2016; Brasseur et al., 2023)***"**

Reviewer: This information is duplicated above.

Response: We thank the reviewer for this observation. The initial simulation period on line 155 is used to describe the sensitivity test for the meteorological parameters (NCEP and HRRR). For clarity, we include the information on line 178-182 to describe the simulation length for all domains in the final simulation. We modify the sentence slightly for clarity on Lines 176-180 as "*In this study, simulations for the mesoscale domains (D01 – D02) span from September 6, 2021 (00:00 UTC) to September 9, 2021 (00:00 UTC), which corresponds to Houston's local time from September 5, 2021 (19:00 Central Daylight Time, CDT) to September 8, 2021 (19:00 CDT), including the first 20 hours as spin-up time. The output from D02 is used to drive parallel simulations for D03 and the LES domain, covering the period from 11:00 to 21:00 UTC (06:00 – 16:00 CDT) on September 8, 2021.***"**

Reviewer: Maybe include as a main finding: seems to me relevant that emissions need to be scaled back by 0.2/0.3?

Response: We now include sentences regarding this scaling in the Conclusion and Abstract sections to integrate this information as part of the main findings.

On Line 20 in the Abstract section, this is modified as "*NO_X emissions from the NEI anthropogenic inventory are reduced by 70-80 % to achieve realistic diurnal variability of NO_X concentration.***"**

On Line 681 in the Conclusion as "*NO_X emissions from the anthropogenic NEI inventory are scaled significantly by a factor of ~0.2-0.3 to achieve more realistic concentrations for the chemistry in our domains, indicating a 70-80 % reduction in NO_X emissions.***"**

Reviewer: I think this is not the Production of Ozone but the net budget of O_X, since loss terms are included for NO_2 also. Anyhow, this is confusing.

Response: We appreciate the reviewer for this observation. The definition of net ozone production in equation (1) is primarily based on radical chemistry that either enhances or suppresses ozone production rates. Reactions of NO with RO_2 or HO_2 are essential daytime photochemical ozone formation pathways. NO_2 consumption leads to the loss of OH (a radical terminating reaction), thus influencing the production of HO_2/RO_2 radicals, and hence ozone formation rates. These reactions consume NO_2, a precursor to O_3.

Reviewer: Here, it remains unclear how the model is sampled. If you sample instantaneously at high resolution, you become sensitive to the influence of resolved eddies that are visible in panel (g) and (j). So, you need to take some "time average" over e.g. one convective overturn timescale (~15 min.)...

Response: The model-observation comparison in Figure 3 is instantaneously sampled. This is done due to the timescale of the TCEQ observation data (hourly averages). Below we provide a figure with the attempt to compare averaged model results and TCEQ observations between 12:00 – 15:00 CDT, which shows similar distributions of meteorological factors as our previous analysis. We add a sentence to clarify the sampling details and this finding on Line 324-325 as *"Similar to the instantaneously sampled horizontal distributions of the modeled meteorological factors, the hourly averaged horizontal distributions show comparable results."*
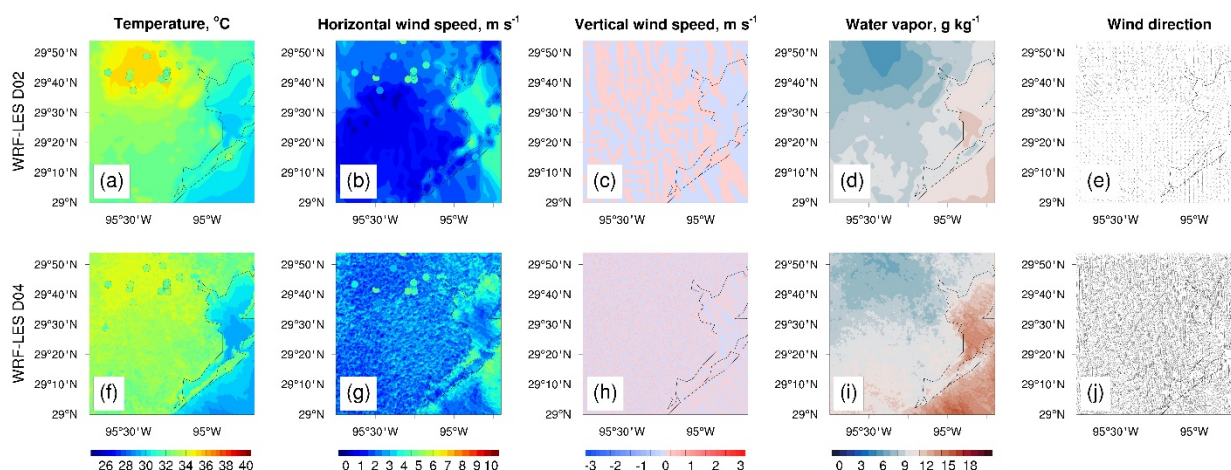
Reviewer: This is strange. Photolysis is expected to peak around midday...so this remark leads to confusion.

Response: We thank the reviewer for the observation. We are not suggesting peak photolysis occurs in the late afternoon, but rather that photolysis is more pronounced in the afternoon compared to the morning hours. We now revise the sentence by removing "late" on line 350. We clarify this as *"In the afternoon, the photolysis of NO_2 is enhanced, thus resulting in reduced NO_2 concentrations."*

Reviewer: again, now it sounds like there is more light in the late afternoon.

Response: Also, we now clarify the sentence in lines 363-365 *"Isoprene, a major precursor to daytime O_3 production, shows higher concentrations in the late afternoon (Fig. 4), as the biogenic emissions are temperature-dependent (Rinne et al. 2002; Guenther et al., 2006; Zeng et al., 2023)".* **We do not imply more lights in the afternoon.**

Reviewer: What do the bars represent? I guess diurnal variability....

I additionally think the readers would appreciate a comparison of time series. Now a lot of information is missing...

Response: We appreciate the reviewer's suggestion. The bars represent the range of VOC values, from minimum to maximum, based on observations and model output at the San Jacinto site between 8:00 and 16:00 CDT. Additionally, we have included a time series comparison in the Supplementary Material (new Fig. S5) and discuss the time series plot in the manuscript on lines 380-389 as *"Figure S5 compares the simulated diurnal variability of four VOCs with observations*

from the PTR-MS. Isoprene is underestimated in the LES domain D04 in the morning, but increasing concentrations are simulated in the afternoon, reducing biases and aligning more closely with observations. The simulated diurnal pattern of HCHO in the LES domain D04 shows better agreement with observations. MEK is overestimated in the morning across both the LES and mesoscale domains, but modeled concentrations become more aligned with observations in the afternoon, with both observations and modeled results indicating high levels in the morning and decreasing concentrations throughout the day. Acetaldehyde, however, is underestimated in the afternoon in both domains, suggesting missing secondary chemical sources from lumped alkanes (Read et al., 2012). It should be noted that the PTR-MS measurement accuracy may be constrained by uncertainties in calibration standards and compound identification based on mass information only (Ammann et al., 2004)."

Reviewer: mm, of course, that is what a sonde does. I think you mean horizontal drift.

Response: Yes, we imply a horizontal drift of the sonde at higher altitudes. We also clarify this on line 406 as "*The trajectories of sondes can drift horizontally from their launch sites as they reach higher altitudes.*"

Reviewer: ??? unclear sentence.

Response: The sentence is now rephrased on lines 419-421 as "*In addition, the parameterization of land surface schemes used to estimate surface sensible and latent heat fluxes can differ between land and near-water surfaces.*"

Reviewer: compared

Response: We change "comparing" to "compared" on line 426.

Reviewer: But what you actually say is that the turbulent variability is random, and model and observation cannot be compared at a single time?

Response: What we mean here is that there is a time lag between the sonde observations and the model output. For example, Galveston sonde measurements are taken at 12:47 CDT, while the closest model output is at 13:00 CDT. This time difference may impact the accuracy of the model-observation comparison.

Reviewer: the

Response: "the" is now added on line 439.

Reviewer: unreadable sentence

Response: The sentence refers to an additional figure that compares the results from the mesoscale and LES domains. For clarity, this sentence is now modified on Lines 445 – 447 as "*Additional inter-domain comparisons of the diurnal evolution of vertical profiles for HCHO, O₃, temperature, potential temperature and water vapor are shown in Fig. S7 of the Supplementary Materials.*"

Reviewer: Figure use (the outdated) ppbv, while text uses ppb. Use ppb throughout.

Response: We thank the reviewer for this suggestion. We change the unit ppbv to ppb in Figure 9, and other instances in the manuscript text.

Reviewer: I guess the mixing volume (= high PBL height) in the afternoon plays also an important role...e.g. fig. 7

Response: We appreciate the reviewer's input on the PBL daytime mixing. This is also included in the manuscript on lines 540-542 as "*Notably, NO_X levels are generally low from late morning into the afternoon, highlighting the diurnal variability of NO_X emission sources as well as the oxidation of NO_2 by OH and the fast dissociation of NO_X. Additionally, the higher PBL in the afternoon results in a larger mixing volume for NO_X, which further reduces daytime NO_X levels*."

Reviewer: reiterate here ...by more than xxx %

Response: We further reiterate the overestimation of the NO_X emission on lines 543-544 as "*Although the scaled NEI anthropogenic emissions inventory more reasonably captures the diurnal patterns, further improvements are needed as it still overestimates NO_X by more than 30 % in Houston.*"

Reviewer**:** This is a rather easy way of concluding. Some more depth in understanding may be obtained by e.g. running without isoprene...

Response: We thank the reviewer for this helpful comment. We agree that having a different simulation without isoprene may provide more depth to the analysis of their role in ozone production. We didn't consider this in the current study due to the computational demand for running different LES chemistry scenarios. Also, we don't want to further lengthen the paper. However, we include this in the future direction for the study in our conclusion on lines 719-721 as "*Given the findings from this study, it is important to further quantify the role of isoprene and anthropogenic VOCs on O_3 formation by adopting a scenario-based approach with and without their emissions. This will provide further evidence of the impacts of BVOCs and HRVOC in O_3 chemistry.*"

Reviewer: Conclusion is not in line with the big role of isoprene you mention above. Again, scenarios without isoprene or anthropogenic emissions could provide meaningful extra evidence.

Response: We thank the reviewer for this observation. We discuss the important role of isoprene in the original manuscript on lines 671-673 as "Our results further underline the role of BVOCs in contributing to enhanced O_3 production, consistent with finidngs from several recent studies (Leong et al., 2017; Ma et al., 2022)." However, we now modify the concluding remarks on lines 715 – 717 as "*Our study concludes the important role of both anthropogenic and biogenic emissions in the production of RO_2 and HO_2 radicals, and how they may influence the enhancement or inhibition of O_3 formation, depending on the NOx-O_3-VOC sensitivity regimes*".

Finally, we also highlight future works that may further assist in quantifying the role of isoprene and anthropogenic VOCs through a scenario-based chemistry test. We add this on lines 719-721 as "*Given the findings from this study, it is important to further quantify the role of isoprene and anthropogenic VOCs on O_3 formation by adopting a scenario-based approach with and without their emissions. This will provide further evidence on the impacts of BVOCs and HRVOC in O_3 chemistry.*"