Response to reviewer #2's comments

Journal: Atmos. Chem. Phys

Manuscript ID: acp-2022-592

Title: " O_3-precursor relationship over multiple patterns of time scale: A case study in Zibo, Shandong Province, China"

Author(s): Zheng et al.

Overall comment:

This manuscript describes a recent 5-month field campaign to better understand the NO_x-VOC sensitivity of ozone during the summer months in Zibo, China. The authors conducted 0-D box modeling with the MCM v3.3.1 near-explicit mechanism, and determined that the selected time scale for this modeling (i.e., daily, weekly, monthly, or campaign-wide) can affect the magnitude of the dependence of ozone formation on its precursors, with shorter time scales (i.e., daily) leading to a wider range of relative incremental reactivities (RIR). RIRs determine the expected effect of reducing certain emissions on the production of ozone, so are a useful tool in mitigating ozone pollution. The authors determine that the RIRs can vary within a city, due to local emissions, and can also vary within a season. This indicates that care should be taken when ascribing meaning to RIRs, as they can be dependent on the modeling parameters selected.

In general, I think this paper is worth publication in ACP. The main result, that averaging over several months blurs out some relevant chemical complexity, isn't that surprising, but the paper is well written, and it is good to see these results analyzed so completely and clearly. I think it will help the community choose time scales wisely.

Reply: We appreciate the professional and positive comments by the reviewer, and we have addressed the proposed concerns in below point-by-point, with revised text in red.

General comments:

Comment 1: What is the authors recommendation for time scale? On line 35, they state that "integrating multiple patterns of time scale is useful to derive reliable and robust O_3-precursor relationships". Do the authors not think that the daily time scale is the most accurate because of its detail? What would be the benefit of doing a weekly or monthly average instead?

Reply: Thanks for pointing this unclear statement. Of course, we agree that the daily time scale is the most accurate because it can provide more informative details among the four patterns of time scale. We recommend the narrower time scale (i.e., daily pattern in this study) in box modelling, as it can provide a more reliable and robust O3-precursor relationship, when considering the non-negligible variability among the four patterns of time scale. We have revised this sentence as below.

Line 40-41: "This implies that utilizing narrower time scale (i.e., daily pattern) is useful to derive reliable and robust O_3-precursor relationship."

Comment 2: A dilution parameter of 3/86400 s^{-1} was chosen based on its best fit to the data. Is this the only parameter that was tuned to fit the model to the data? Is it possible that the trends determined here could be affected by the selection of that dilution parameter?

Reply: In this study, the dilution rate is indeed the only parameter that was tuned to fit the model to the measured O_3 data. As described in **Text S1**, we have performed a stepwise sensitivity test to generate an optimal dilution rate for all non-constraint species for all simulation days. Many 0-D box model simulations include this dilution rate for all non-constraint species to avoid secondary species from building up to unreasonable levels (Bloss, et al., 2005; Wolf, et al., 2012; Wolf, et al., 2016), which is regarded as a technical model parameterization because of the very rare ideal stagnant conditions in the realistic atmsphere. As proposed in Wolf et al., (2016), the dilution rate (k_{dil}) is represented as a first order reaction in the box model. Currently we are unable to analyze how different k_{dil} values affect the trends in photochemical regime as it requires systematical model calculation effort, but this is worth for further investigation in the future. These have been incorporated into our manuscript as below.

Line S86-S93 in **Supplement Text S1**: "By comparing the modeled O_3 with observed O_3 for the three sites, we obtained an optimized k_{dil} of 3/86400 s^{-1}, and assigned it to all non-constraint species for all simulation days, which is the only model parameter that was tuned to fit the measured O_3 data. In general, this optimized k_{dil} is conductive to ensuring the rationality and comparability of model performance for all modeled days at the three sites, and it is also worth for further investigation about how different k_{dil} values affect the trends in photochemical regime.”

Reference:

Bloss, C.; Wagner, V.; Jenkin, M. E.; Volkamer, R.; Bloss, W. J.; Lee, J. D.; Heard, D. E.; Wirtz, K.; Martin-Reviejo, M.; Rea, G.; Wenger, J. C.; Pilling, M. J. Development of a Detailed Chemical Mechanism (MCMv3.1) for the Atmospheric Oxidation of Aromatic Hydrocarbons. *Atmos. Chem. Phys.* **2005**, *5* (3), 641–664. https://doi.org/10.5194/acp-5-641-2005.

Jenkin, M. E.; Wyche, K. P.; Evans, C. J.; Carr, T.; Monks, P. S.; Alfarra, M. R.; Barley, M. H.; McFiggans, G. B.; Young, J. C.; Rickard, A. R. Development and Chamber Evaluation of the MCM v3.2 Degradation Scheme for β-Caryophyllene. *Atmos. Chem. Phys.* **2012**, *12* (11), 5275–5308. https://doi.org/10.5194/acp-12-5275-2012.

Wolfe, G. M.; Marvin, M. R.; Roberts, S. J.; Travis, K. R.; Liao, J. The Framework for 0-D Atmospheric Modeling (F0AM) v3. 1. *Geosci. Model Dev.* **2016**, *9* (9), 3309–3319.

Comment 3: In general, some of the Supplemental Figures and Tables are not presented in the order in which they appear in the manuscript.

Reply: Thanks for your careful review of our manuscript, and we have checked and relocated these supplemental Figures and Tables in order.

Comment 4: On line 392, the authors state that the RIR_AVOC (and others) are increasing as the time scale gets narrower. But looking at Figure 6, it seems that only the mean is increasing. The median stays the same, implying that there are some more extreme values of RIR_AVOC in the daily model, that are getting averaged out as the time scale broadens. I think this is a different statement than saying that the RIR_AVOC value changes, and the authors should be more careful about making that distinction.

Reply: Thanks for the good comments. As mentioned by the reviewer in **Figure 6**, there are discrepancies of the RIRs trends depicted by mean and median respectively. To make it clear, we use the mean to describe the trends of RIRs in accordance with the method of averaging dataset into multiple time scales throughout this section. This has been incorporated into our manuscript as below.

Line 403-410: "As the time scale changed from wider (i.e., five-month scale) to narrower (i.e., daily scale) pattern, all three sites showed increases in the means of RIR_{AVOC} and $RIR_{alkenes*}$ as well as decreases in averaged RIR_{NOx}, whereas the averaged RIR of other precursors (i.e., BVOC, CO, alkanes and aromatics) did not vary obviously (see **Table S6**). Comparing with the O_3-VOC-NO_x sensitivity at the daily scale, the results obtained at the five-month scale underestimated O_3-AVOC sensitivity (indicated by averaged RIR values) by 48% (TZ), 66% (BJ), and 49% (XD), and overestimated O_3-NO_x sensitivity by 37% (TZ), 142% (BJ), and 144% (XD)."

Line 428-431: "Compared with the five-month pattern, it was further found that the averaged RIR_{NOx}/RIR_{AVOC} from other time scale patterns (i.e., monthly, weekly, and daily) were higher (12% to 20% for TZ; 38% to 153% for XD) or lower (21% to 65% for BJ) than that from five-month scale."

Comment 5: Section 3.7 describes the uncertainty analysis, but I believe the authors are conflating the mathematical terms for uncertainty (as analyzed by standard deviation of the averages), and the broader qualitative term for uncertainty (that some chemical species weren't included in the model). Figure 10 does not seem to be informative, it simply demonstrates that more data reduces the uncertainty, but that's well known. I would recommend removing that figure altogether, and focusing section 3.7 on what is missing in the model.

Reply: We agree with the reviewer, and we have removed the **Figure 10** in manuscript and relocated into supplement as Figure S14. To focus the uncertainty on what is missing in the model, we have simplified and rephrased the uncertainty analysis about averaging approach in **Section 3.7** of our manuscript as below.

Line 480-485: "**Figure S14** shows the distributions of the standard deviations for OH reactivity (k_{OH}) or concentration of O_3 precursor groups at three averaged patterns of time scale at the three sites. As the time scale changed from wider (i.e., five-month scale) to narrower (i.e., weekly scale) pattern, the uncertainty (indicated by the average, median and 25%-75% quantile) decreased accordingly."

Comment 6: The RIR_CO is presented in many figures but never discussed. Does CO play a relevant role in the O_3 formation here? If not, why is it presented in these figures?

Reply: Overall, CO played a relatively limited role in comparison with other major categories of O_3 precursors (e.g., NO_x, AVOC and BVOC) at the three sites. Therefore, those RIR_CO figures without discussion in supplement have been removed directly, while we have added some descriptions for RIR_CO plots in manuscript (see below).

Line 316-318: "In addition, the RIR_{CO} values at the three sites suggested its limited role in O_3 formation at the three sites, compared with other major categories of O_3 precursors."

Line 403-407: "As the time scale changed from wider (i.e., five-month scale) to narrower (i.e., daily scale) pattern, all three sites showed increases in the means of RIR_{AVOC} and $RIR_{alkenes*}$ as well as decreases in averaged RIR_{NOx}, whereas the averaged RIR of other precursors (i.e., BVOC, CO, alkanes and aromatics) did not vary obviously (see **Table S6**)."

Specific comments:

Comment 1: Line 19 and 91: "integrating" → "integrates"

Reply: Corrected. (Line 22 and 95)

Comment 2: Line 21: "multiple-site" → "multiple sites"

Reply: Corrected. (Line 24)

Comment 3: Line 26: The authors state that the RIRs are "consistent with time scale", but the manuscript demonstrates that they have different magnitudes. Do they mean consistent with regard to sign (i.e., all positive or all negative)? This should be stated more clearly. Same comment for line 33, which describes consistency in the photochemical regimes, when I think the authors mean consistency in the sign, but not the magnitude.

Reply: Thanks for careful review of our manuscript, and we have rephrased the two sentences as below.

Line 27-31: "It was found that the relative incremental reactivity (RIR) of major precursor groups (e.g., anthropogenic volatile organic compound (AVOC), NO_x) was overall consistent in the sign along with time scales changed from wider to narrower (four patterns: five-month, monthly, weekly, and daily) at each site, though the magnitudes of RIR varied at different sites."

Line 36-39: "It was further found that the campaign-averaging photochemical regimes showed overall consistency in the sign but non-negligible variability among the four patterns of time scale, which was mainly due to the embedded uncertainty in model input dataset when averaging individual daily pattern into different timescales."

Comment 4: Line 58: Describing the 0-D model as "advanced" makes it sound like it is more complex than the regional scale air quality models on line 55. But I would say that the benefit of the 0-D model is its relative simplicity (despite the larger MCM mechanism), which allows the kind of in-depth chemical analysis that the authors do here.

Reply: We agree with the reviewer's comments that the regional scale air quality models are more sophisticated than 0-D box model, and we have revised this statement in our manuscript as below.

Line 62-65: "Unlike the complicated 3-D air quality models, the 0-D box model is an observation-based model that implemented with gas-phase chemical mechanism, and has been widely used to diagnose O_3-precursor relationship in various locations (Liu et al., 2021a; Sun et al., 2016; Tan et al., 2019b; Xue et al., 2014a; Yu et al., 2020a)."

Comment 5: Line 112: The Thermo Scientific 42i measures total NO_x or NO. Can you describe more fully how the measured NO_x was separated into NO and NO2? Do you expect any uncertainty there to make a difference to your modeling results?

Reply: Thank you for the good comments. The overestimated NO_2 by chemiluminescence technique is always a challenging problem, therefore some studies also designated "NO_2" as "NO_y-NO". Both gaseous HNO_3 and organic nitrates can result in interferences on NO_x measurement by chemiluminescence technique, and they are typically found in some polluted urban environments. For example, gaseous HNO_3 contributed approximately most to total nitrate (particle-phase and gas-phase nitrate) through gas-particle partitioning in summer (Ryota et al., 2022; Uno et al., 2017), while the contribution of organic nitrates to the total particle nitrate decreased as the $PM_{2.5}$ loading increased (Ge et al, 2022). In addition, the study of Xu et al. (2013) suggested that the overestimation of NO_2 by the molybdenum converter is limited in areas with fresh NO_x emission sources, while such interference is more significant in rural or remote areas due to large amount of oxidized nitrogen in aged air mass. Considering significant fresh NO_x emissions nearby the three selected sites in Zibo, we believe that the interference of NO_2 measurement from chemiluminescence method should be limited in this study. Nevertheless, it is still meaningful to perform more in-depth study on NO_x measurement uncertainty in box model simulation in the future, particularly when both traditional and accurate NO_x measurement are available, as the accuracy of NO_x measurement is essential in determing the photochemical regime. These discussions have been incorporated into our manuscript as below.

Line 502-507: "Besides, both gaseous HNO_3 and organic nitrates can result in interferences on NO_x measurement by chemiluminescence technique, which may arise uncertainty in our box modelling (Ge et al., 2022; Uno et al., 2017; Xu et al., 2013). Since accurate NO_x measurement is essential in determining the photochemical regime, more in-depth studies on NO_x measurement uncertainty in box model simulation are required in the future."

Reference:

Ge D, Nie W, Sun P, et al. Characterization of particulate organic nitrates in the Yangtze River Delta, East China, using the time-of-flight aerosol chemical speciation monitor[J]. Atmospheric Environment, 2022, 272: 118927.

Steinbacher M, Zellweger C, Schwarzenbach B, et al. Nitrogen oxide measurements at rural sites in Switzerland: Bias of conventional measurement techniques[J]. Journal of Geophysical Research: Atmospheres, 2007, 112(D11).

Uno I, Osada K, Yumimoto K, et al. Seasonal variation of fine-and coarse-mode nitrates and related aerosols over East Asia: synergetic observations and chemical transport model analysis[J]. Atmospheric Chemistry and Physics, 2017, 17(23): 14181-14197.

Xu Z, Wang T, Xue L K, et al. Evaluating the uncertainties of thermal catalytic conversion in measuring atmospheric nitrogen dioxide at four differently polluted sites in China[J]. Atmospheric environment, 2013, 76: 221-226.

Nojiri R, Osada K, Kurosaki Y, et al. Variations in gaseous nitric acid concentrations at Tottori, Japan: Long-range transport from the Asian continent and local production[J]. Atmospheric Environment, 2022, 274: 118988.

Comment 6: Line 116: More information about the Zibo Eco-Environmental Monitoring Center is needed here. Is this data publicly available? What instrumentation was used? Or is there a reference to this site?

Reply: Sorry for the confusing statement, and Zibo Eco-Environmental Monitoring Center and our group are responsible for the rountine operation of these monitoring sites in Zibo city. In this study, the meteorological parameters (i.e., temperature, relative humidity, UV-A solar radiation, precipitation, wind speed, and wind direction) were monitored by standard instruments following the Chinese meteorological monitoring regulation (GB/T 35221-2017). The dataset of the three sites in this study can be available after its publication, and we have added a relevant reference to these sites (Li et al., 2021).

Line 113-116: "Following the Chinese meteorological monitoring regulation (GB/T 35221-2017), we continuously monitored the meteorological parameters (i.e., temperature, relative humidity, UV-A solar radiation, precipitation, wind speed, and wind direction) at the three sites (Li et al., 2021)."

Reference:

Li, K., Wang, X., Li, L., Wang, J., Liu, Y., Cheng, X., Xu, B., Wang, X., Yan, P., Li, S., Geng, C., Yang, W., Azzi, M. and Bai, Z.: Large variability of O_3-precursor relationship during severe ozone polluted period in an industry-driven cluster city (Zibo) of North China Plain, J. Clean. Prod., 316, 128252, doi:https://doi.org/10.1016/j.jclepro.2021.128252, 2021.

Comment 7: Line 125: "Tenax GR" → "Tenax GR cartridges"

Reply: Corrected. (Line 121)

Comment 8: Lines 128-137: It would be helpful to describe what differences might be expected between the FID and FID/PID instruments? Have they been cross-checked and calibrated?

Reply: Thank you for the good comments, and another reviewer also pointed out similar issue. In **Methods Section**, we have briefly discussed the differences between the GC-FID and GC-FID/PID instruments. Unfortunately, we did not conduct the inter-comparison between the GC-FID and GC-FID/PID instruments at the same site, as these VOC instruments were separately deployed at the three different sites for continuous routine operation, which is very difficult to relocate and maintain them in one site due to practical reasons. Nevertheless, these VOC instruments at the three sites are commercial instruments, and were regularly maintained and calibrated by standard gases with 55 VOC species from the same cylinder (Linde Co., Ltd, USA). Therefore, we assume the VOC datasets at the three sites are overall reliable and suitable for subsequent analysis in this study. These discussions have been incorporated into our manuscript as below.

Line 116-131: "Two online GC systems (gas chromatography–flame ionisation detector, GC-FID, Thermo Scientific GC5900) were deployed at TZ and BJ respectively to measure VOC species. For C_2-C_5 VOCs, desorption and separation were performed using a GC with pre-concentration on a combination of two columns, followed by a FID detector. For C_6-C_{12} VOCs, air sample was pre-concentrated on Tenax GR cartridges and subsequently separated by chromatographic column, then detected by another FID detector. Similarly, one online system (gas chromatography–flame ionisation detector/photoionisation detector, GC-FID/PID, Syntech Spectras GC 955-615/815) was deployed at XD site. For C_2-C_6 VOCs, the hydrocarbons were concentrated on a Tenax GR carrier, then thermally desorbed and separated on a DB-1 column, and finally detected by FID and PID detectors. For C_6-C_{12} VOCs, the air sample was concentrated on a Carbosieves SIII carrier at 5°C, then thermally desorbed and separated on a combination of two columns, and FID and PID detectors were employed for subsequent detection. These systems measured 55 VOC species at a 1-h resolution, and more detailed descriptions can be found elsewhere (Chien, 2007; Jiang et al., 2018; Xie et al., 2008)."

Line 134-137: "Unfortunately, we did not conduct the inter-comparison between the GC-FID and GC-FID/PID instruments at the same site due to practical reasons, as these VOC instruments were separately deployed at the three different sites for continuous routine operation."

Comment 9: Line 139: Were the 5-point calibration standards from 5 separate standard cylinders? Or dilution one cylinder. If the second option, does that affect the accuracy of the calibration at all?

Reply: In our campaign, we selected the later (i.e., dilution from one cylinder) to perform the 5-point calibration in our campaign, and we have added this information into our manuscript as below. According to **Table S2**, the correlation coefficient of five-point calibration (i.e., 2, 4, 6, 8, 10 ppbv) for the 55 VOC species were nearly 0.9990, thus we assume the impact on accuracy of the calibration by the second option is relatively limited.

Line 137-141: "To ensure the quality assurance / quantity control (QA/QC) of online VOC measurement, two five-point calibrations (i.e., 2, 4, 6, 8, 10 ppbv, dilution from one cylinder) for standard gases with 55 VOC species (Linde Co., Ltd, USA) were carried out in May and August of 2019 at the three sites."

Comment 10: Line 155: Define F0AM.

Reply: Done. (Line 155-156)

Comment 11: Line 174: "due to significant miss" → "due to significant missing data". Additionally, which data was missing? It is hard to see in Figure 2 what is missing. Is the cause of the missing data related to the fact that these time periods seem to overlap to unusual (i.e., not very diurnal) patterns in the O_3, such as Jun 3-6, Jul 5-8, and Aug 9-15 in figures S6-8?

Reply: We have corrected the above statement as "due to some missing data". Indeed, we cannot perform simulation for some individual days due to some missing data as it requires a complete 24 h dataset for model input, which leads to limited modelling days [n=100 (TZ), n=81 (BJ), n=114 (XD)] as shown in **Figures S6-S8** for daily scale pattern.

Comment 12: Line 214: In this analysis, was the 10% change in X a 10% increase or a 10% decrease? Does it matter which is selected, since O_3 has a non-linear response to NO_x in particular?

Reply: It is assumed as 10% decrease in X for RIR calculation. We believe that it does not matter either 10% increase or 10% decrease, as the calcuated RIR reflects a relative change of O_3 production rate to the change in X. In addition, we further assess the influence of choosing different hypothetical changes (i.e., 5%, 10%, and 15%) on RIR values, thus box model sensitivity test was performed with the above three scenarios. **Figure 1** (in below) shows the model-derived RIR values under three hypothetical changes, using averaged diurnal pattern of five-month time scale as model input. In general, the RIR values of O_3 precursor groups and RIR_{NOx}/RIR_{AVOC} ratios were overall consistent under different hypothetical changes. Since "10% decrease" was widely employed in the previous studies (Lyu et al., 2016; Wang et al., 2017a, 2018b), we applied 10% as hypothetical change in our RIR calculation for consistency.

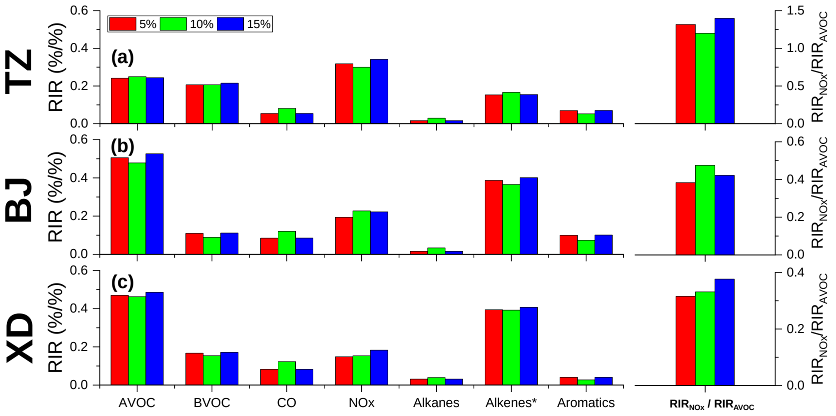

Figure 1. The RIR values of O_3 precursor groups and RIR_{NOx}/RIR_{AVOC} at different hypothetical changes (i.e., 5%, 10%, and 15%) using diurnal average of five-month pattern as model input at the three sites.

Reference:

Lyu, X. P., Chen, N., Guo, H., Zhang, W. H., Wang, N., Wang, Y. and Liu, M.: Ambient volatile organic compounds and their effect on ozone production in Wuhan, central China, Sci. Total Environ., 541, 200–209, doi:10.1016/j.scitotenv.2015.09.093, 2016.

Wang, Y., Wang, H., Guo, H., Lyu, X., Cheng, H., Ling, Z., Louie, P. K. K., Simpson, I. J., Meinardi, S. and Blake, D. R.: Long-term O_3-precursor relationships in Hong Kong: Field observation and model simulation, Atmos. Chem. Phys., 17(18), 10919–10935, doi:10.5194/acp-17-10919-2017, 2017.

Wang, Y., Guo, H., Zou, S., Lyu, X., Ling, Z., Cheng, H. and Zeren, Y.: Surface O_3 photochemistry over the South China Sea: Application of a near-explicit chemical mechanism box model, Environ. Pollut., 234, 155–166, doi:10.1016/j.envpol.2017.11.001, 2018.

Comment 13: Line 228: Do the authors have a hypothesis for why the wind speeds were so different at the different sites? Were they at equal altitudes above the ground? Could it have been an instrument issue?

Reply: The wind speeds at the three sites were measured following Chinese meteorological monitoring regulation (GB/T 35221-2017), thus we assume the obtained dataset of wind speeds should be reliable and highly unlikely due to instrument error. In addition, the altitudes above the ground of the three sites are nearly same. We guess such site-to-site difference of wind speeds may be associated with the discrepancies from local meteorological field, given the relatively long distance (i.e., more than 50 km) among the three sites.

Comment 14: Line 234: How was the scaling done?

Reply: Specifically, the geographical coordinates, date and time were initialized into the TUV model to derive photolysis rates and solar radiation. We obtained the scaling factor by comparing the observed with modeled solar radiation, and used this scaling factor to scale the TUV model-derived photolysis rates (Lyu et al., 2019; Lyu et al., 2016).

Line 163-167: "Specifically, the geographical coordinates, date and time were initialized into the TUV model to derive photolysis rates and solar radiation. We obtained the scaling factor by comparing the observed with modeled solar radiation, and used this scaling factor to scale the TUV model derived photolysis rates."

Reference:

Lyu, X., Wang, N., Guo, H., Xue, L., Jiang, F., Zeren, Y., Cheng, H., Cai, Z., Han, L. and Zhou, Y.: Causes of a continuous summertime O_3 pollution event in Jinan, a central city in the North China Plain, Atmos. Chem. Phys., 19(5), 3025–3042, doi:10.5194/acp-19-3025-2019, 2019.

Lyu, X. P., Chen, N., Guo, H., Zhang, W. H., Wang, N., Wang, Y. and Liu, M.: Ambient volatile organic compounds and their effect on ozone production in Wuhan, central China, Sci. Total Environ., 541, 200–209, doi:https://doi.org/10.1016/j.scitotenv.2015.09.093, 2016.

Comment 15: Line 243: "quantity" → "quantify"

Reply: Corrected. (Line 259)

Comment 16: Line 253: See my general comment #2. Isn't model performance also due to the selection of the dilution parameter that maximizes model performance?

Reply: Indeed, dilution rate is the only parameter that was tuned to maximize box model performance. We obtained an optimized dilution rate of 3/86400 s^{-1} through sensitivity test, and assigned it to all non-constraint species for all simulation days and the three sites (see details in **Text S1** and **reply to Commet 2**).

Comment 17: Line 287: Why is there no section dedicated to discussing the full campaign time scale? I only see three sections (monthly, weekly, and daily), but later on the authors discuss the four different time scale analyses.

Reply: The RIR results from five-month (full campaign) pattern of time scale have been comprehensively discussed in the **Section 3.6** by comparing it with other patterns of timescale.

Comment 18: Line 289 – 295: This is helpful information for understanding how the model was run, and should be moved to the methods section. This is also true for lines 314-319.

Reply: We agree with the reviewer's comments, and these two parts have been relocated in the **Methods Section** in our manuscript as below.

Line 224-232: "In this study, the O_3 precursors were divided into four major categories, including anthropogenic VOC (AVOC), biogenic VOC (BVOC, only isoprene in this study), CO and NO_x (Tan et al., 2019b). AVOC was further divided into three subcategories: alkanes, aromatics and alkenes* (the asterisk denotes anthropogenic alkenes, excluding isoprene in this study) (Yu et al., 2020a). As mentioned, RIR method was applied mainly to evaluate the O_3-NO_x-VOC sensitivity and determine the photochemical regimes among four patterns of time scale. Thus, we calculated the RIR values of major precursor groups (i.e., AVOC, BVOC, CO, NO_x, alkanes, alkenes* and aromatics) to further quantify the O_3-precursor relationship."

Line 233-238: "In general, O_3 formation chemistry is usually classified into three regimes (i.e., VOC-limited, transitional and NO_x-limited) (He et al., 2019; Wang et al., 2018). In this study, RIR_{NOx}/RIR_{AVOC} (the ratio of two RIR values) was used as a metric to classify the photochemical regimes (Li et al., 2021). Specifically, RIR_{NOx}/RIR_{AVOC} value of less than 0.5 was defined as VOC-limited regime, greater than 2 as NO_x-limited regime, and from 0.5 to 2 as transitional regime (see **Text S2** and **Table S4**) (Li et al., 2021)."

Comment 19: Line 396: How are these underestimation percentages being calculated?

Reply: These underestimation percentages were calculated by comparing the averaged RIR from five-month scale with the mean of RIRs from daily patterns, and all data were summarized in **Table S6**. This has been incorporated into revised manuscript in below.

Line 403-410: “As the time scale changed from wider (i.e., five-month scale) to narrower (i.e., daily scale) pattern, all three sites showed increases in the means of RIR_{AVOC} and $RIR_{alkenes*}$ as well as decreases in averaged RIR_{NOx}, whereas the averaged RIR of other precursors (i.e., BVOC, CO, alkanes and aromatics) did not vary obviously (see **Table S6**). Comparing with the O_3-VOC-NO_x sensitivity at the daily scale, the results obtained at the five-month scale underestimated O_3-AVOC sensitivity (indicated by averaged RIR values) by 48% (TZ), 66% (BJ), and 49% (XD), and overestimated O_3-NO_x sensitivity by 37% (TZ), 142% (BJ), and 144% (XD).”

Line 428-430: “Compared with the five-month pattern, it was further found that the averaged RIR_{NOx}/RIR_{AVOC} from other time scale patterns (i.e., monthly, weekly, and daily) were higher (12% to 20% for TZ; 38% to 153% for XD) or lower (21% to 65% for BJ) than that from five-month scale.”

Comment 20: Line 530: While it is true in this case that all time scales yielded the same information for ranking the top-10 VOC contributors, do the authors expect it would be true in all cases? Unless they have done that analysis, I would recommend changing the language here to just describe these results, and not try to make this broad statement about all models.

Reply: We agree with the reviewer’s comments, and have removed this broad statement “This demonstrates that datasets with wider pattern of time scale can still produce an accurate RIR ranking / prioritization for VOC control” in this revision.

Comment 21: Line 546: “difference” → “differences”

Reply: Corrected. (Line 553)

Comment 22: Table 1: This is a summary of the most relevant 0-D box models, correct? In that case it should be stated as "Summary of relevant published 0-D box model studies", so it doesn't imply this is every box model ever published.

Reply: Table 1 is indeed a summary of the most relevant 0-D box models, and we have corrected the Table 1 caption as "Summary of relevant published 0-D box model studies".

Comment 23: Text S1: This isn't ever reference in the main text, and doesn't add any new information. I would delete it.

Reply: We have deleted this part in Supplement.

Comment 24: Line S88: This URL should be a cited reference instead.

Reply: Done. (Line 100 in Supplement)

Comment 25: Table S2: State in the caption what the asterisk next to "alkenes" refers to. Additionally, does "non-listed in box model" mean that they were measured but not modeled? Or modeled but not measured?

Reply: We have added explanation of alkenes* in the caption of Table S2. Additionally, "non-listed in box model" represents that some VOCs were measured but not simulated in box model, as they are not included in the MCMv3.3.1 chemical mechanism. These have been added in the caption of Table S2 as below.

Table S2. Summary of the correlation coefficient of five-point calibration (i.e., 2, 4, 6, 8, 10 ppbv) for the 55 VOC species during the May and August of 2019 at the three sites in Zibo city. Alkenes* denotes anthropogenic alkenes, excluding isoprene in this study. "Non-listed in box model" represents ten measured VOC species that cannot be simulated in box model.

Line 155-159: "In this study, the box model (based on the Framework for 0-D Atmospheric Modeling, F0AM) (Wolfe et al., 2016) was applied and constrained by the mean diurnal profiles of meteorological data (i.e., temperature, relative humidity, and photolysis rates), 4 inorganic gases (i.e., SO_2, CO, NO, and NO_2), and 45 speciated VOCs (in MCMv3.3.1 species list; see **Table S3**)."