Response to Reviewer 1's Comments

The reviewer's comments are written in blue and our responses are in black.

The authors ran WRF-Chem simulations over East Asia using three different emission inventories (EDGAR-HTAP v2, EDGAR-HTAP v3, and KORUS v5) and compared the model output to three sets of observations (routine monitoring data, airborne KORUS-AQ data, and ground-based KORUS-AQ data). They also ran sensitivity tests to doubling CO and VOC emissions and probed how the chemistry changed. Such comparisons are useful for model development, but I think there is opportunity for the paper to be strengthened in the following ways:

1. Currently, O_3 and NO_2 are treated separately in the comparison. I suggest adding a comparison of odd-oxygen ($O_x = O_3 + NO_2$) to probe whether the model issue is too much O_3 titration by NO or problems with the O_3 production regime.

→ We added analysis of Ox with surface observations in China and South Korea after P14 L24 (also see **Figure R1 and R2**). The diurnal patterns of Ox are well simulated with all emission inventories (**Figure R1**), showing similar issues that are previously discussed in section 3.2.

As the reviewer expected, underestimations in the model O_3 in YRD and NCP using EDV2 (light blue lines) disappeared when it was replaced by O_x, suggesting that the O_3 biases using EDV2 in the regions are caused by too much NO_x titration or inefficient O_3 formation in a NOx-saturated regime. For other regions and cases using EDV3 and KOV5, Ox plots highlight biases in the model NO_2 levels. In YRD, Ox overestimations correspond to NO_2 overestimations in Figure 4. Meanwhile, in SCG and SEC, there are Ox overestimations caused by O_3 overestimations, suggesting a potential VOC emission overestimation.

Detailed descriptions, along with Figure R1 and R2, are included in the revised manuscript and Supporting Information.

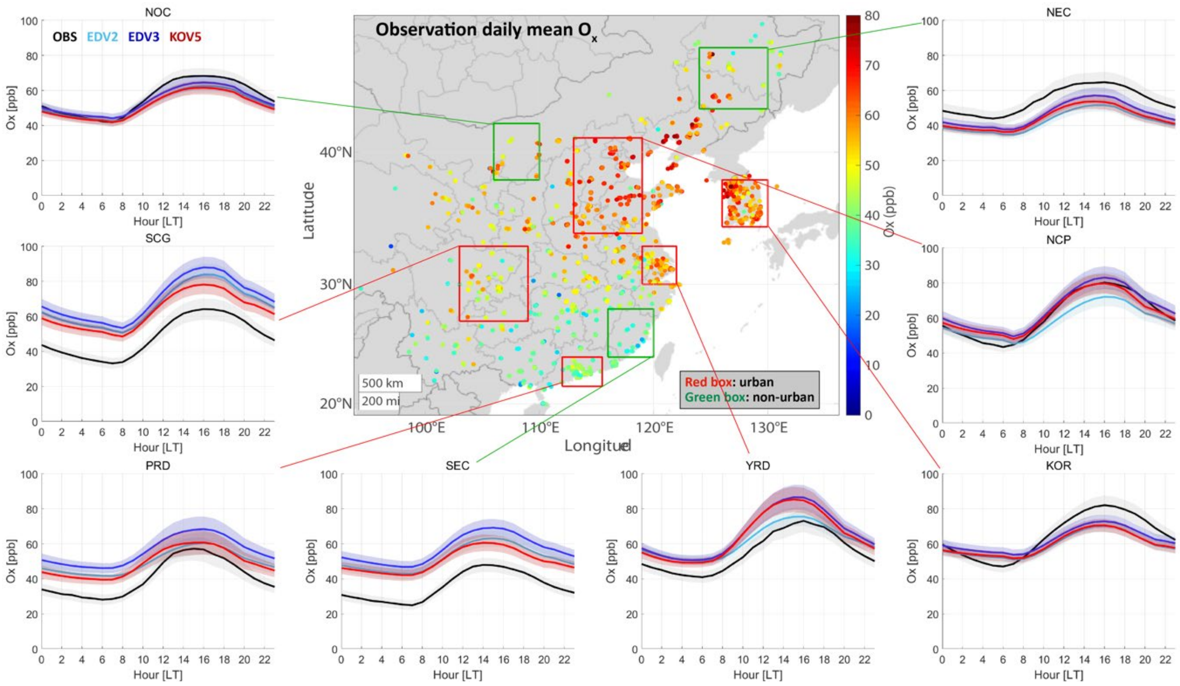

Figure R1. Averaged Ox concentrations from ground-based observations and model simulations over the areas that distinguish urban (red box) and non-urban (green box) region (central plot). Box-averaged diurnal cycle (solid lines) of Ox and 1/4 of standard deviations (filled area) from observations (black), EDV2 (sky blue), EDV3 (blue), and KOV5 (red) by local time are shown.

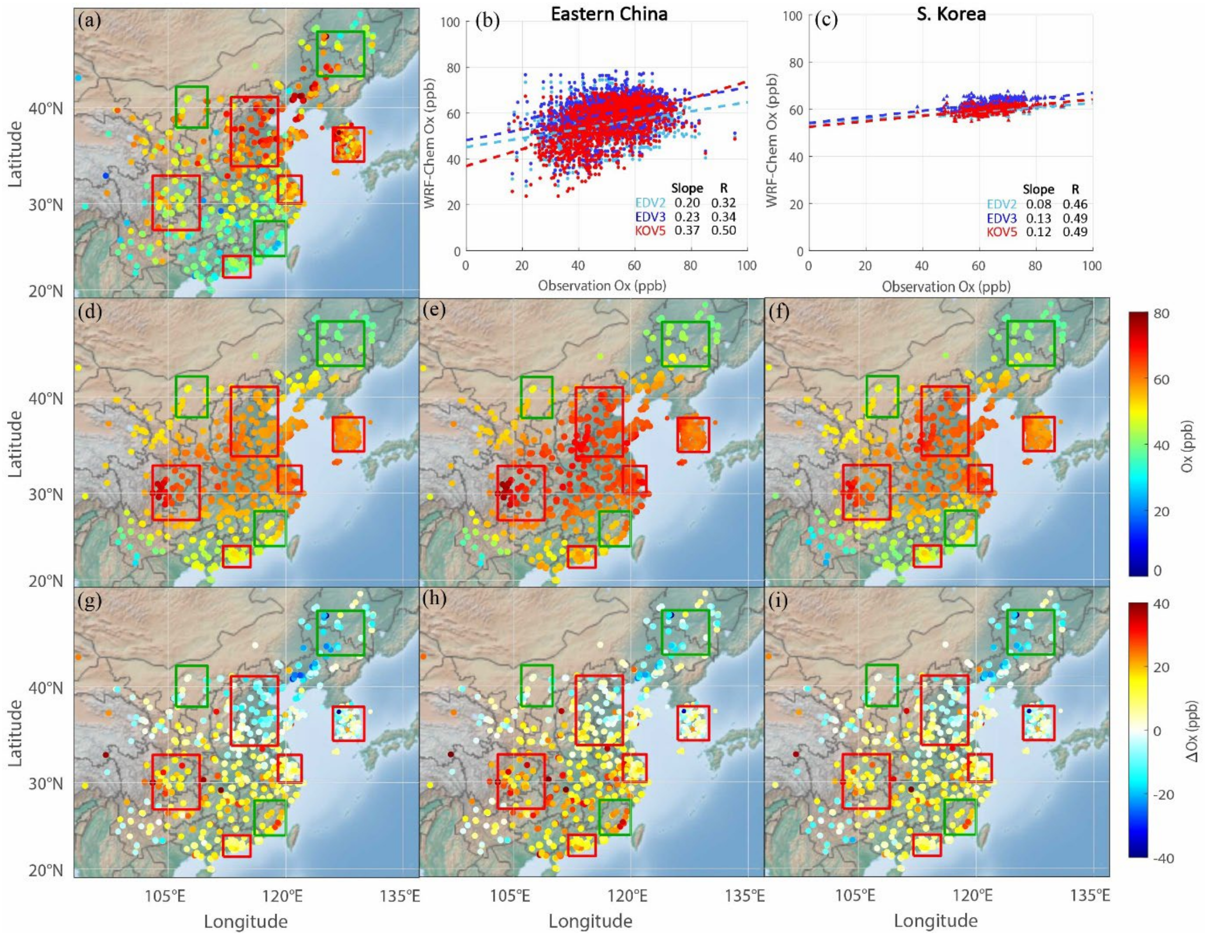

Figure R2. Comparison of (a) the campaign averaged ground-based maximum daily average of 8-hour Ox (MDA8 Ox) (unit: ppb) observations and WRF-Chem simulations with (d) EDGAR-HTAP v2 (EDV2), (e) v3 (EDV3), (f) KORUS v5 (KOV5) and (g, h, i) the differences between the observations and model results. The scatter plots comparing averaged observations and the three-emission-based WRF-Chem simulations (sky blue; EDV2, blue; EDV3, red; KOV5) are shown in (b) and (c) for Eastern China and South Korea, respectively.

2. On a related note, can you use the individual comparisons you've done of VOCs, NOx, and O_3 in different regions with each inventory to draw some conclusions about how biases in either or both NO_x or VOC emissions affect O_3 predictions / chemical regimes in the different regions? There is a little bit of this on pages 13-14, but more organized conclusions about this (especially with your sensitivity tests to doubling the CO and VOC emissions) would be very useful for future model interpretation and emissions inventory development. For example, on page 12, can you discuss why the differences in VOCs and NO_x in each of the inventories cause them to simulate O_3 differently?

→ We provided **Table R1**, detailing NOx, TOL, XYL, biogenic isoprene emissions, and formaldehyde-to-NO_2 ratio (FNR) for each region and emission inventory to enhance the understanding of regional differences. We included some discussions in section 3.2, such as the descriptions of VOC-limited regime in NCP with low FNR (< 1). The higher emissions of TOL and XYL in EDV3 and KOV5 resulted in higher O_3 concentrations with the smaller biases than EDV2 (**Table R1**). In SCG and SEC, biogenic emissions exceeded TOL and XYL by up to the factor of 10 with all emission inventories.

Table R1. Comparison of total NOx, TOL, XYL, biogenic isoprene emissions in May, and formaldehyde-to-NO_2 ratio (FNR) for the KORUS-AQ campaign period for different emission datasets in each regional box. The MEGAN biogenic isoprene emissions are equally applied to all simulations using different emission data. (unit = mol/s for emissions)

Type	emissions	NCP	SCG	YRD	PRD	KOR(SMA)	NEC	NOC	SEC
NOx emission	EDV2	5967	1500	2366	1178	990(196)	987	688	590
	EDV3	5202	1654	1642	1091	1191(214)	876	597	662
	KOV5	3237	902	1166	607	886(191)	513	373	410
TOL emission	EDV2	140	56	84	47	27(6)	26	8	20
	EDV3	220	77	99	68	27(8)	40	9	36
	KOV5	403	106	234	155	98(26)	68	21	79
XYL emission	EDV2	84	34	51	28	15(4)	15	4	12
	EDV3	132	46	60	41	16(4)	24	6	22
	KOV5	133	35	79	52	41(9)	21	7	26
Biogenic isoprene emission		132	364	43	127	135(6)	106	23	310
FNR (14-16LT)	EDV2	0.25	1.31	0.19	0.52	0.53(0.19)	0.68	0.76	1.18
	EDV3	0.44	1.30	0.32	0.52	0.43(0.18)	0.93	0.94	1.33
	KOV5	0.72	2.33	0.48	1.00	0.71(0.22)	1.44	1.49	1.91

Interpreting O_3 biases using FNR is cautioned due to the complex interplay of VOC and NOx emissions and chemistry. Therefore, we added section 4 (discussion) with 4 additional sensitivity simulations (C5-C8 in **Figure R3**) as discussed in Kim et al. (2023), providing insights for O_3 bias correction in each region and city.

In SCG and SEC, the C5 case (50% anthropogenic VOC emission reduction only) exhibited the lowest O_3 biases, with a slight decrease in O_3 concentrations in the C4 case (50% NOx reduction only), implying the need to reduce VOC emissions (biogenic and/or anthropogenic emissions) (**Figure R3**). For the YRD and PRD, both NOx and VOC emissions should be reduced based on C6 case (50% NOx and VOC reduction), while C4 case (only NOx 50% reduction) increased O_3 bias.

We also compared the sensitivity simulations with 12 mega cities in China and South Korea (**Figure R4**). VOC 50% reduction (C5 case) improved O_3 and NO_2 simulations in Chengdu and Chongqing. The lowest biases of O_3 and NO_2 were achieved with 50% NOx and VOC reduction case (C6 case) for Shanghai, Nanjing, Guangzhou, Shenzhen, and Wuhan.

The detailed analysis, along with Figure R3 and R4, has been added to the revised manuscript.

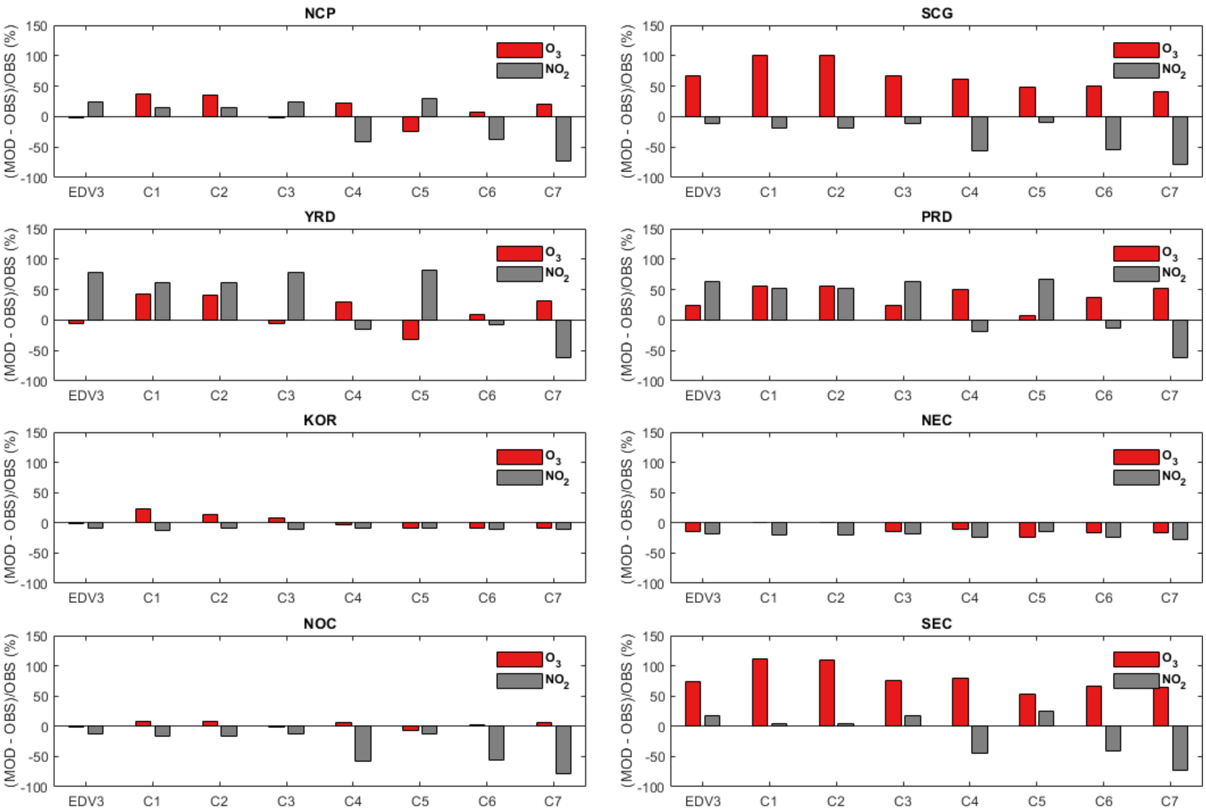

Figure R3. Comparison of relative biases ((Model-Observation)/Observation, unit=%) of daily O_3 and NO_2 at surface observation sites during the KORUS-AQ campaign period from sensitivity simulation (C1-7) with EDV3 in each region (NCP, SCG, YRD, PRD, KOR, NEC, NOC, and SEC). C1; EDGAR-HTAP v3 with double CO and VOC emission in China and South Korea, C2; EDGAR-HTAP v3 with double CO and VOC emission in China, C3; EDGAR-HTAP v3 with double CO and VOC emission in South Korea, C4; EDGAR-HTAP v3 with 50% NOx reduction in China, C5; EDGAR-HTAP v3 with 50% VOC reduction in China, C6; EDGAR-HTAP v3 with 50% NOx and VOC reduction in China, C7; EDGAR-HTAP v3 with 75% NOx reduction in China.

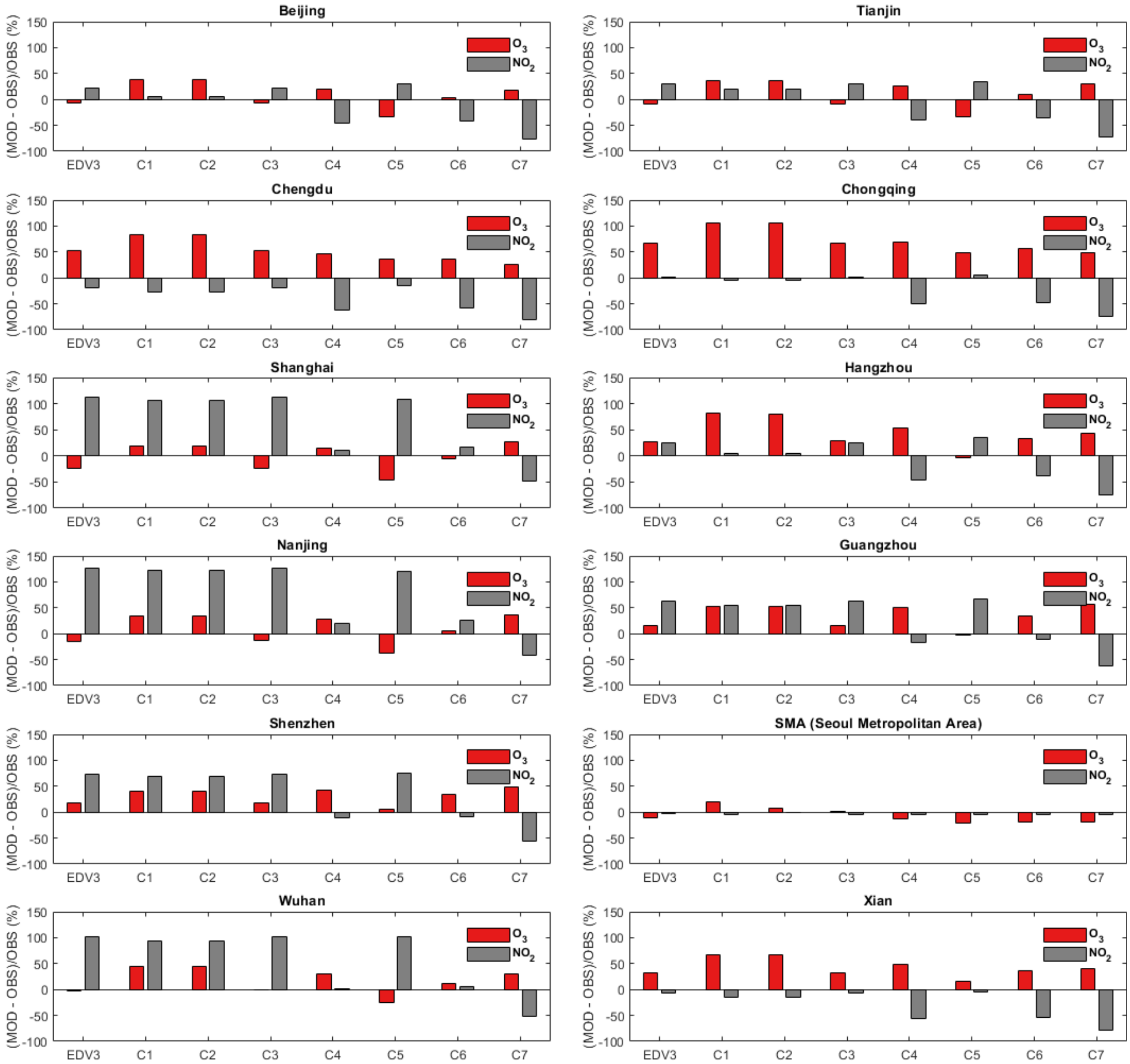

Figure R4. Same as **Figure R3** except that the region is changed to cities; Beijing (39.4-41.1N, 115.4-117.5E), Tianjin (38.55-40.25N, 116.7-118.1E), Chengdu (30.05-31.5N, 103-105E), Chongqing (28.15-32.25N, 105.3-110.2E), Shanghai (30.7-31.5N, 120.85-122E), Hangzhou (29.2-30.6N, 118.3-120.9E), Nanjing (31.2-32.65N, 118.35-119.25E), Guangzhou (22.55-24N, 112.9-114.05E), Shenzhen (22.4-22.9N, 113.7-114.65E), SMA (37.2-37.8N, 126.5-127.3E), Wuhan (29.95-31.4N, 113.65-115.1E), and Xian (33.65-34.75N, 107.65-109.9E).

3. Most of the conclusions in the manuscript are stated as "X is biased low with Y emissions inventory." These statements would be more useful to the atmospheric chemistry community writ large if those statements were extended to say, "X is biased low with Y emissions inventory, which has Z implications for our understanding of emissions/chemistry." For example, on page 20 line 12, can you add something to this sentence about the implications of having larger biases in the Transport case compared to the Local case? For a second example, on page 20, line 20, can you add something about the implications for NO_x emissions (based on O_3 being wrong but CO and HCHO being largely okay)? There are many other instances in the manuscript where this would be useful, but hopefully, the two examples I provided here are helpful illustrations.

→ To enhance discussions about causes of the model O_3 biases, we added a separate section of discussion about the chemical regimes in each region and city and the best way to reduce ozone biases accordingly, incorporating NOx emissions information. Please refer to our response to Reviewer's major comment 2. Furthermore, we added the sentences about analysis of Local and Transport case.

The excessive O_3 with double emissions in China is attributed to an overestimation of background O_3. We included Figure R5 to represent the overestimated O_3 from the downwind area (Yellow Sea) when CO and VOC emissions are doubled in China. This analysis is included in the revised manuscript. Furthermore, in section 4, causes of O_3 biases and directions to improvement for each region and cities are suggested in detail including Figure R3 and R4.

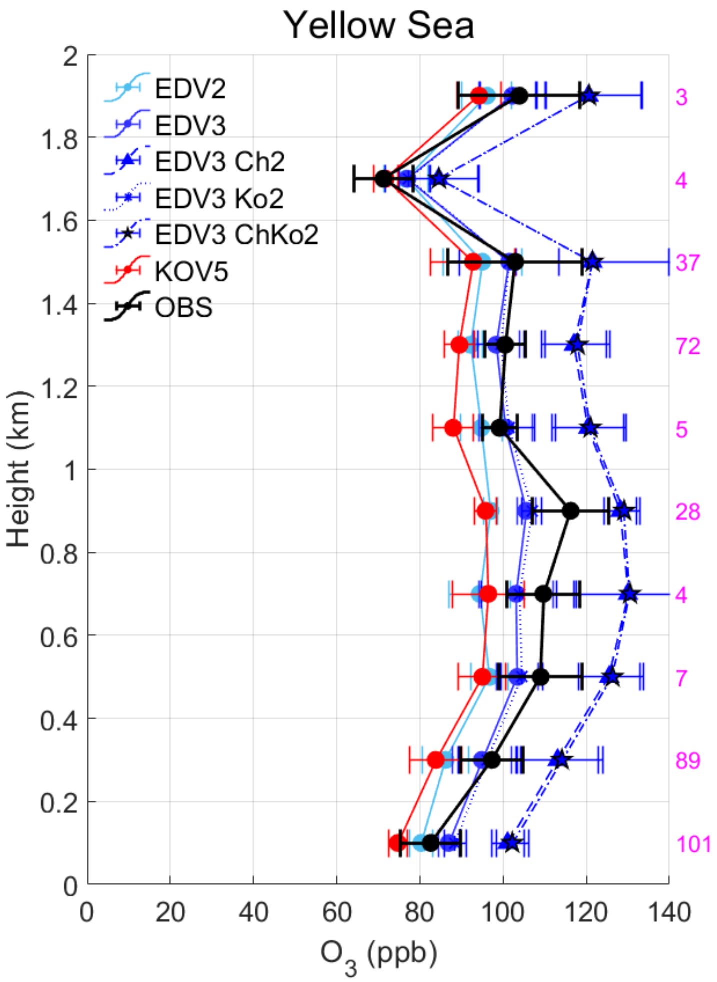

Figure R5. Vertically averaged O_3 from DC-8 (black), EDV2 (sky blue), EDV3 (blue), KOV5 (red), EDV3 with doubling Chinese CO and VOC emissions (dashed blue), EDV3 with doubling Korean CO and VOC emissions (dotted blue), and EDV3 with doubling Chinese and Korean CO and VOC emissions (dotted dashed blue) in Yellow Sea under 2 km height above ground level. The 1/2 of standard deviations are represented with whiskers in each 200m layer. The sample number is presented with magenta color on the right side of the plots.

4. The manuscript includes some contextualization of this work in the context of other emissions inventory comparisons (e.g., for CO on page 8). I think the paper would be strengthened by adding similar contextualization for the other comparisons (NO_x, O_3, VOCs, etc.), especially given how many model-measurement comparisons have been done to date with KORUS-AQ (and related) data.

→ Our objective is to systematically identify and summarize potential issue associated with anthropogenic bottom-up emission inventories, investigating their potential impact on O_3

simulations in East Asia. We included relevant previous model studies in section 1 (Introduction) as explained below.

"Many modeling studies are done during this period including validations of CTM results with various observations. Miyazaki et al. (2019) adjusted emission inventories using various satellite data sets and Model for Interdisciplinary Research on Climate with chemistry (MIROC-Chem) resulting in O_3 simulations improvement. Goldberg et al. (2019) reported underestimations of NOx emissions in South Korea including Seoul. Souri et al. (2020) also revealed the same issue in South Korea and analyzed sensitivity of O_3 formation to the NOx and VOC emission adjustments derived from inverse modeling. Tang et al. (2019) revealed negative bias of simulated CO concentrations in East Asia by utilizing satellite data and the Community Atmosphere Model with Chemistry (CAM-Chem). Choi et al. (2022) modified anthropogenic VOC emissions using satellite HCHO observations and inverse modeling method with the Goddard Earth Observing System with Chemistry (GEOS-Chem), which reduced O_3 and HCHO biases."

Specific comments:

Page 5, line 8: are the NMVOCs lumped or speciated?

→ It is lumped NMVOC. We added 'total' in front of 'non-methane volatile organic compound'.

Page 5, line 9: do you apply any scale factors for using 2010 emissions data in a 2016 simulation?

→ We did not use scale factors.

Page 5, line 17: What does 'specifically' mean here?

→ We intended to describe that it is speciated NMVOCs from EDGAR-HTAP v2. We will change 'specifically mapped EDGAR-HTAP v2 data' to 'speciated EDGAR-HTAP v2 VOC data' to avoid confusion.

Page 7, lines 1-2: should read "…toluene and less *reactive* aromatics…"

→ Thank you. We added 'reactive' in front of 'aromatics' in the revised manuscript.

Page 7, line 14: what species are you referring to that is larger in South Korea by 263%?

→ It's TOL as mentioned in previous sentence. We added 'of TOL' behind 'relative difference' to avoid confusion in the revised manuscript.

Page 8, line 1: add "respectively" after "(HCHO)"

→ Thank you for the comment. We included 'respectively' after '(HCHO)'.

Page 8, line 13: I think it would be clearer to say "For all emission inventories…" rather than "With all emission…"

→ Thank you for the comment. We changed "With all emission inventories" to "For all emission inventories" in the updated version of manuscript.

Page 8, lines 14-17: It was hard for me to figure out which simulations correspond to which numbers in these sentences. Reword to clarify?

→ We changed the sentence to "we conducted two additional model simulations using EDGAR-HTAP v3 that shows lowest bias of O3 concentrations compared to DC-8 than

EDGAR-HTAP v2 and KORUS v5 over the SMA (mean bias = EDV2: -16.9, EDV3: -14.2, KOV5: -18.1 ppb)" adding bias information between parentheses in the revised manuscript.

Page 10, line 19: What kind of interpolation method was used?

→ The linear interpolation method is used for the vertical interpolation. We added "using linear interpolation method" after "vertically interpolated to the aircraft data" in the revised manuscript.

Page 11, lines 11-19: Can you say something about how much these meteorological biases affect your comparisons? For example, how much would temperature-dependent evaporative VOC emissions change over these temperature ranges?

→ Both anthropogenic and biogenic VOC emissions can be affected by air temperature (Huang et al., 2022; Song et al., 2019). In this response, we could calculate the impact of temperature on biogenic VOC emissions. The isoprene emission in MEGAN is calculated following the equation below (Guenther et al., 2006).

$$Emission = [EF][\gamma][\rho]$$

$$\gamma_{Temp} = E_{opt} \cdot \frac{C_2 \cdot \exp(C_1 \cdot x)}{C_2 - C_1 \cdot (1 - \exp(C_2 \cdot x))}$$

$$x = \frac{\left(\frac{1}{T_{opt}} - \frac{1}{T}\right)}{0.00831}$$

EF is emission factor (mg m^{-2} h^{-1}). ρ is normalized ratio. γ is an emission activity factor that can vary for different conditions such as leaf area index, temperature, vegetation type, leaf age, soil moisture, and canopy environment. E_{opt} and T_{opt} are empirical coefficients. C_1 and C_2 are constants. We calculated isoprene (ISO) emission sensitivity to temperature bias at each SYNOP station by changing T. The negative temperature biases resulted in reduced isoprene emissions in South Korea (**Figure R6**). However, as discussed in 3.3.1, ISO is still overestimated for all regions.

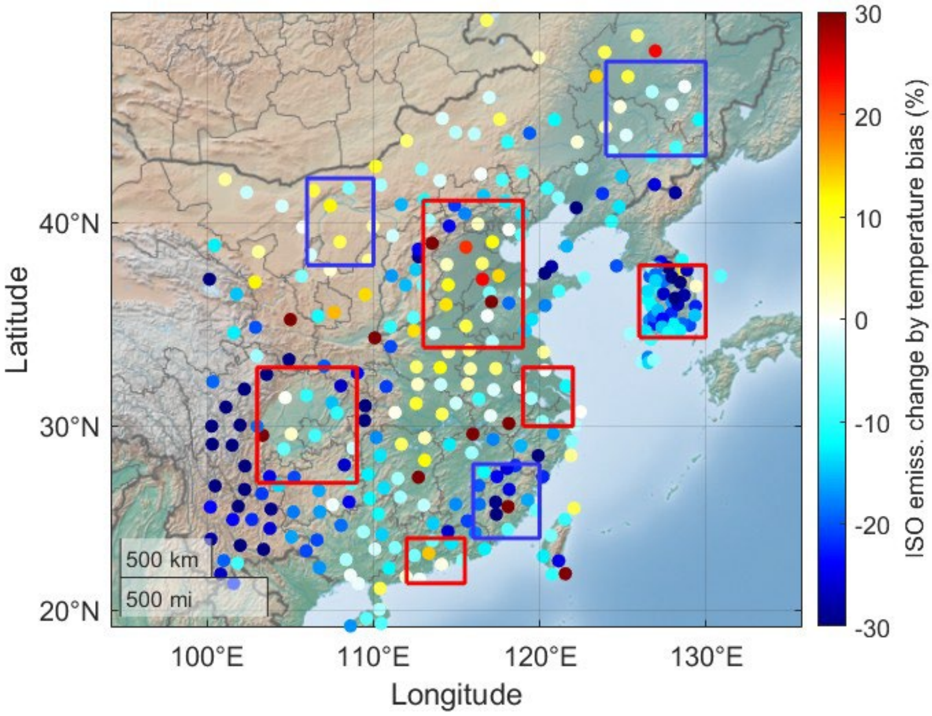

Figure R6. Relative isoprene (ISO) emission change from the temperature bias at the surface (unit = %).

Page 12, line 11: "all emissions inventories" instead of "all emissions"

→ Thank you for the comment. We changed "all emissions" to "all emission inventories".

Page 17, line 3: ISO definition should be moved earlier to where it's first used.

→ Agreed. We first defined isoprene as ISO in line 24 of page 13 in the revised manuscript.

Page 19, lines 23-24: can you use the biases calculated during the local case to draw some conclusions about the emissions inventory over China?

→ We added ", which implies that the insufficient local emissions of O_3 precursors in the emission inventories are much important that the Chinese emissions." after "15.5-18.2 ppb" in the revised manuscript to clarify local VOC emission issues to the low model O_3 concentrations in South Korea.

Title: Unclear what 'precursor' refers to here. Is it O_3 and HCHO precursors? If so, perhaps rephrasing it as "ozone, formaldehyde, and their precursors" would be clearer.

→ Agreed. We changed the title.

References

Choi, J., Henze, D. K., Cao, H., Nowlan, C. R., Abad, G. G., Kwon, H.-A., Lee, H.-M., Oak, Y. J., Park, R. J., Bates, K. H., Massakkers, J. D., Wisthaler, A., and Weinheimer, A. J.: An Inversion Framework for Optimizing Non-Methane VOC Emissions Using Remote Sensing and Airborne Observations in Northeast Asia During the KORUS-AQ Field Campaign, *J. Geophys. Res. Atmos.*, 127, e2021JD035844, https://doi.org/10.1029/2021JD035844, 2022.

Kim, S.-W., Kim, K.-M., Jeong, Y., Seo, S., Park, Y., and Kim J.: Changed in surface ozone in South Korea on diurnal to decadal timescales for the period of 2001-2021, *Atmos. Chem. Phys.*, 23, 12867-12886, https://doi.org/10.5194/acp-23-12867-2023, 2023.

Goldberg, D. L., Saide, P. E., Lamsal, L. N., de Foy, B., Lu, Z., Woo, J.-H., Kim, Y., Kim, J., Gao, M., Carmichael, G., and Streets, D. G.: A top-down assessment using OMI NO2 suggests an underestimate in the NOx emissions inventory in Seoul, South Korea, during KORUS-AQ, *Atmos. Chem. Phys.*, 19, 1801-1818, https://doi.org/10.5194/acp-19-1801-2019, 2019.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), *Atmos. Chem. Phys.*, 6, 3181-3210, https://doi.org/10.5194/acp-6-3181-2006, 2006.

Huang, J., Yuan, Z., Duan, Y., Liu, D., Fu, Q., Liang, G., Li, F., and Huang, X.: Quantification of temperature dependence of vehicle evaporative volatile organic compound emissions from different fuel types in China, *Sci. Total Environ.*, 813, 152661, https://doi.org/10.1016/j.scitotenv.2021.152661, 2022.

Miyazaki, K., Sekiya, T., Fu, D., Bowman, K. W., Kulawik, S. S., Sudo, K., Walker, T., Kanaya, Y., Takigawa, M., Ogochi, K., Eskes, H., Boersma, K. F., Thompson, A. M., Gaubert, B., Barre, J., and Emmons, L. K.: Balance of Emission and Dynamical Controls on Ozone During the Korea-United States Air Quality Campaign From Multiconstituent Satellite

Data Assimilation, *J. Geophys. Res. Atmos.*, 124, 387-413, https://doi.org/10.1029/2018JD028912 , 2019.

Song, C., Liu, B., Dai, Q., Li, H., and Mao, H.: Temperature dependence and source apportionment of volatile organic compounds (VOCs) at an urban site on the north China plain, *Atmos. Environ.*, 207, 167-81, https://doi.org/10.1016/j.atmosenv.2019.03.030, 2019.

Souri, A. H., Nowlan, C. R., Abad, G. G., Zhu, L., Blake, D. R., Fried, A., Weinheimer, A. J., Wisthaler, A., Woo, J.-H., Zhang, Q., Chan Miller, C. E., Liu, X., and Chance, K.: An inversion of NOx and non-methane volatile organic compound (NMVOC) emissions using satellite observations during the KORUS-AQ campaign and implications for surface ozone over East Asia, *Atmos. Chem. Phys.*, 20, 9837-9854, https://doi.org/10.5194/acp-20-9837-2020, 2020.

Tang, W., Emmons, L. K., Arellano Jr, A. F., Gaubert, B., Knote, C., Tilmes, S., Buchholz, R. R., Pfister, G. G., Diskin, G. S., Blake, D. R., Blake, N. J., Meinardi, S., DiGangi, J. P., Choi, Y., Woo, J.-H., He, C., Schroeder, J. R., Suh, I., Lee, H.-J., Kanaya, Y., Jung, J., Lee, Y., and Kim, D.: Source Contributions to Carbon Monoxide Concentrations During KORUS-AQ Based on CAM-chem Model Applications, *J. Geophys. Res. Atmos.*, 124, 2796-2822, https://doi.org/10.1029/2018JD029151, 2019.