Reply to the reviewer and editor:

We thank reviewers for their careful reading of the manuscript and their many constructive comments. We have adopted most of the suggestions. The original comments by the reviewer are in black font, and our replies are in blue.

This work investigated the reasons for the increase of surface ozone concentrations in perspectives of emission changes and meteorological variations. It's found that meteorological variation is the major driving force to the ozone increase in South China. The paper is informative and generally well organized. My major concern is the negative bias of TROPOMI data (~20% based on previous validations), especially over urban areas, and the effect to the results due to the bias.

Thanks for the positive comments. Our item-by-item responses can be found in below.

My detailed comments are listed below.

1. Line 126-127: not clear about what this means. Do you mean the difference between NOx, NOy, and NOz? Please specify this.

   The China National Environmental Monitoring Center measures the surface NO_2 using the chemiluminescence method, which quantifies the NO_2 concentrations by measuring the NO decomposed from NO_2, which may introduce a positive bias because NOz can also be decomposed to NO. You are right. We clarified it in the updated manuscript. We hope it is clear now.

2. Line 166-167: is the optimization done day by day, or on a monthly basis?

   The optimization was implemented by assimilating the daily TROPOMI NO_2 measurements to constrain the emission at the monthly scale. The scale factor is optimized for each month at each gridbox, as the original temporal resolution of MIX emissions is monthly. We have clarified this in the updated manuscript. The cost function in the optimization seeks to improve the overall agreement of daily TROPOMI and GEOS-Chem NO2 in a month.

3. Line 216-217: the anthropogenic VOC emissions in China has grown significantly from 2010-2017, driven by solvent use. I don't think it can be ignored.

   To further verify the assumption that anthropogenic VOC emission change can be ignored, we compared the GEOS-Chem simulation of HCHO with TROPOMI data. The results are added in Appendix D of the updated manuscript. The GEOS-Chem simulations

Baseline (2019) and 2020 Adjoint, which used the default MIX 2010 and updated MIX 2020 VOC emissions respectively, agree well with TROPOMI HCHO in Feb.-Mar. of 2019 and 2020, respectively (Fig. D1 in the updated manuscript). The correlation coefficients for the two comparisons are 0.878 and 0.874, and the mean bias error are around -24% and -27%, respectively. The negative bias is primarily contributed by the low (< 0.15 DU) HCHO columns (Fig. D1 in the updated manuscript), and it is in an acceptable range considering TROPOMI tends to overestimate the HCHO by 26% for HCHO column lower than around 0.1 DU (Vigouroux et al., 2020).This good agreement supports our assumption of ignoring the VOC emission change.

Fig. D1 also shows the GEOS-Chem tends to overestimate the HCHO VCD over urban regions. It indicates default MIX 2010 may overestimate the anthropogenic emissions in 2019, at least for HCHO, although Zheng et al. (2018) reported the anthropogenic emissions of non-methane volatile organic compounds (NMVOCs) increased by 11% during 2010-2017.

In addition, the uncertainty of the anthropogenic VOC emissions in this study can be smeared by the large biogenic source of VOC emissions.

4. Line 221-222: can you elaborate more on the HCHO uncertainties?

The satellite retrieval of HCHO using the differential optical absorption spectroscopy (DOAS) method has large uncertainty because of the low optical depth of HCHO. The total retrieval uncertainty is composed of the errors of slant columns and model errors (De Smedt et al., 2018). The errors of the slant columns are contributed by the characteristics of the measuring instrument and errors of the slant column fitting procedure in the DOAS method. The model errors are related to the inaccurate representation of the physical properties of the atmosphere and surface, such as the treatment of clouds and aerosols and the surface albedo (De Smedt et al., 2018). Vigouroux et al. (2020) reported that TROPOMI HCHO tends to be overestimated by 26% for HCHO column lower than 0.093 DU (2.5×10^{15} molecules cm^{-2}) and underestimated by 31% for HCHO column higher than 0.297 DU (8.0×10^{15} molecules cm^{-2}). We added this quantification of the TROPOMI HCHO uncertainty in the updated manuscript.

5. Figure 2. For (c), it seems that the VOC emission changes are mainly from biogenic or open biomass burning for Southeast Asia, instead of anthropogenic sources.

TROPOMI HCHO data cannot distinguish the anthropogenic emissions from biogenic and biomass burning sources for the Indochinese Peninsula in Southeast Asia because of the dense vegetation in this region. However, this study investigated the O_3 pollution in China, the Southeast Asia with the dense vegetation is out of our study domain. The impact of VOC emission bias in Southeast Asia on the surface O_3 pollution in China is negligible considering the lifetime of biogenic VOC is generally short. For example, the lifetime of the most prevalent biogenic VOC species, isoprene, is only 1-2 hours (Atkinson, 2000). For the populated urban regions in China, where the surface O_3 pollution exerts more significant health impacts, the anthropogenic source dominates the VOC emissions (Williams & Koppmann, 2007). We made this clarification in Section 3.1 of the updated manuscript.

6. Line 294: previous validations have found that TROPOMI NO_2 product has negative bias of ~20%, especially over urban areas (Verhoelst et al., 2021, Judd et al., 2020, Li et al., 2021). How will this bias affect the validation and your conclusions?

Verhoelst et al. (2021) reported that TROPOMI tropospheric NO_2 column displays a negative bias of -23% to -37% in clean to slightly polluted areas but reaching up to -51% over highly polluted regions compared to the ground-based measurements, but they did not consider the averaging kernel in the comparison. Judd et al. (2020) reported that the negative bias of TROPOMI NO_2 was reduced from 19%-33% to 7%-19% by recalculating the tropospheric air mass factor (AMF) using the averaging kernel to replace the priori NO_2 profiles. Li et al. (2021) also found the TROPOMI NO2 displays a negative bias up to -20% over polluted regions even with the AMF correction applied. In our study, we applied the AMF correction using the TROPOMI averaging kernel for both data assimilation and model evaluations. We expect a negative bias up to -20% of TROPOMI NO_2 even for urban regions. This low bias may cause an underestimation of NOx emission in our study.

To evaluate the influence of the underestimation of the NOx emissions on our main conclusions, we conducted sensitivity simulations with the anthropogenic NOx emissions increased by 20%. The validation of simulated NO_2 column against TROPOMI NO_2 would not be significantly affected since we are using TROPOMI NO_2 for both data assimilation and model evaluation, and the bias in TROPOMI data will be cancelled out. The impacts on the O_3 simulation would be a major concern. Fig. R1 below displays the comparison of MDA8 surface O_3 from the simulation 2020 Adjoint (Table 1 in the Manuscript) and the one with NOx emission increased by 20%. The relative difference is within 6% in North China and within 2% in South China. Overall, this difference is lower than the inter-annual variation of surface O_3 between 2019 and 2020 by one order of magnitude (Fig. 8 in the manuscript).

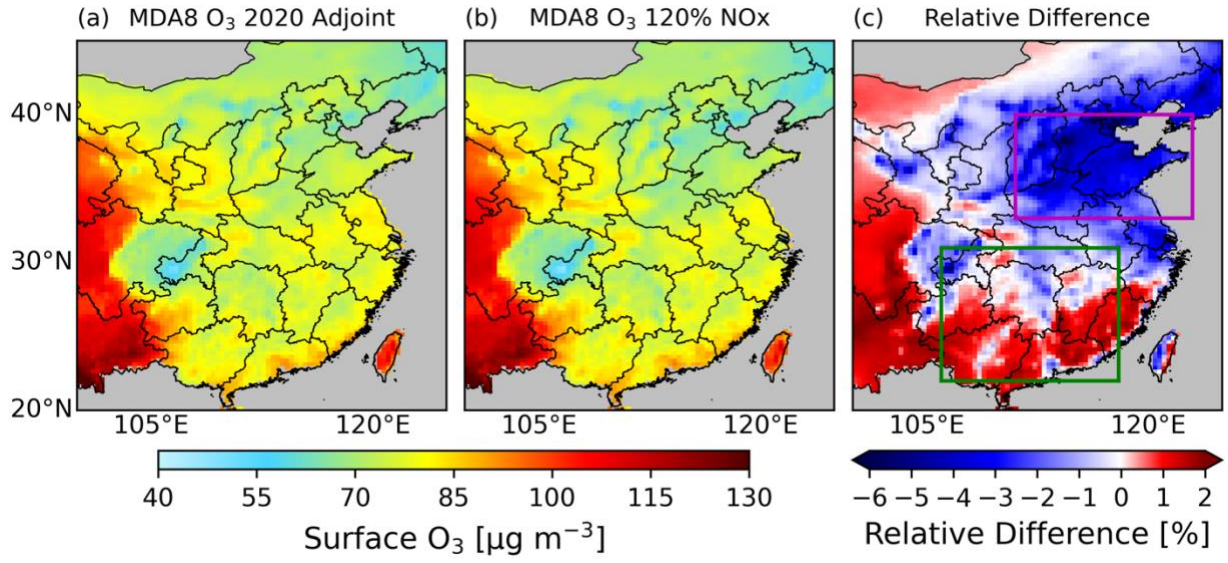

Figure R1. (a) GEOS-Chem mean MDA8 O_3 at 9 m above the surface under standard temperature and pressure (273.15 K, 101.325 kPa) from 2020 Adjoint simulation (Table 1 in the

manuscript); (b) Same as (a) but with anthropogenic NOx emission increased by 20%; (c) Relative difference between (b) and (a), which is calculated as dividing the difference of (b) minus (a) by (a).

We further investigated how this difference affect the validation of O_3 simulation against ground measurements. Fig. R2 shows the scatter plot for comparing the surface MDA8 O_3 over South China from ground measurements versus GEOS-Chem simulations 2020 Adjoint and the one with NOx emission increased by 20%. Over South China, the model performance of surface O_3 simulation after increasing NOx emission by 20% is similar to that of default simulation, with mean bias error (MBE) decreased by 40.2% (0.553 μg m^{-3}) but root mean square error (RMSE) increased by 5.51% (0.484 μg m^{-3}). Over North China, the underestimation of the surface O_3 becomes more pronounced as NOx emission increased by 20% (Fig. R3). The negative mean bias increased from -12.620 μg m^{-3} to -15.624 μg m^{-3} by 23.8%, but the slope of the regression line increased from 0.336 to 0.372. Overall, the impact of increasing NOx emission by 20% on the surface O_3 simulation is not significant.

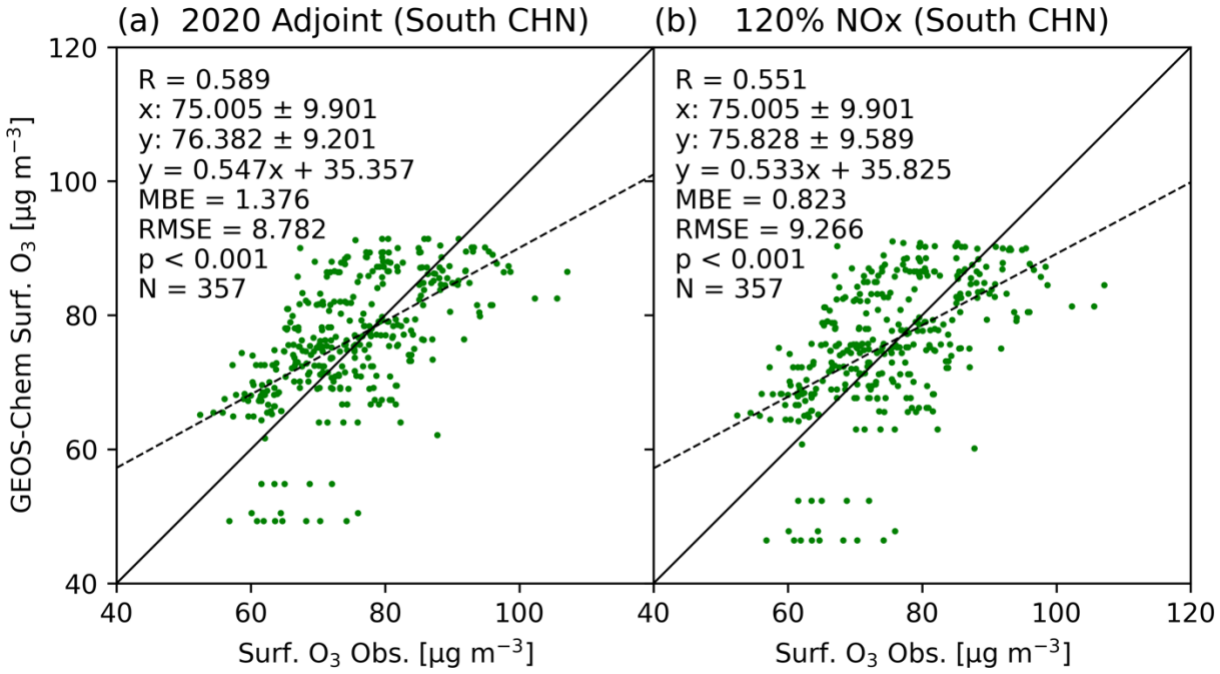

Figure R2. Scatter plot for comparing the surface MDA8 O_3 from GEOS-Chem simulations and ground measurements over South China (green box in Fig. R1). (a) GEOS-Chem simulation 2020 Adjoint versus ground measurements in 2020 Feb.-Mar. (b) The simulation same as 2020 Adjoint but with anthropogenic NOx emission increased by 20% versus ground measurements in 2020 Feb.-Mar.

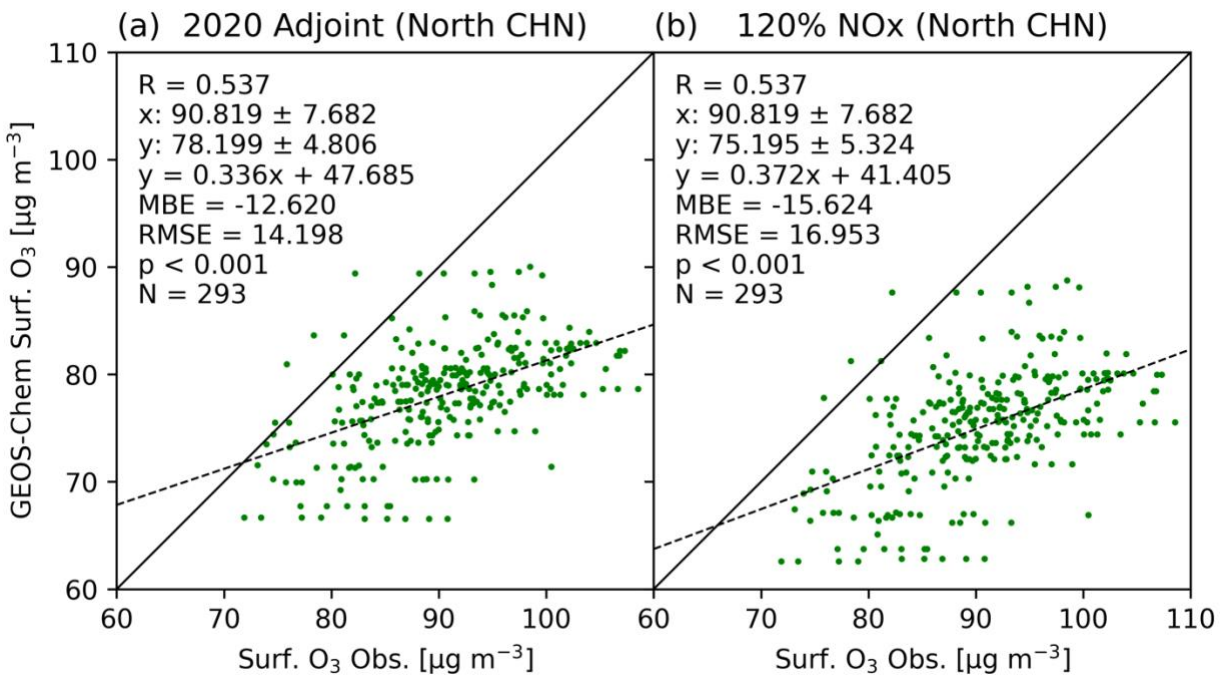

Figure R3. Same as Fig. R2 but for North China.

For the conclusion on the impacts of emission change and meteorological variation on the surface O_3 pollution, which is calculated from the relative difference of various model sensitivity simulations, the influence of underestimation of NOx emission is not significant. The reason is that we updated the NOx emission using TROPOMI NO_2 data for each simulation and the impacts of underestimation of NOx emission would be partly cancelled out. We also conducted the sensitivity simulations by increasing the anthropogenic NOx emission by 20% for both 2019 Baseline and 2020 Adjoint and computed their relative difference. The variation of the relative difference is within 4% compared to that of default simulations. The main conclusion that inter-annual meteorological variations have a larger impact than emission reductions on the aggravated surface O_3 pollution in China during the lockdown period of COVID-19 pandemic still holds considering the low bias in the TROPOMI NO_2 data. We noted this part in Section 4.1 in the updated manuscript.

References for reviewer comments:

Verhoelst, T., Compernolle, S., Pinardi, G., Lambert, J.-C., Eskes, H. J., Eichmann, K.-U., et al. (2021). Ground-based validation of the Copernicus Sentinel-5P TROPOMI NO2 measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global networks. Atmospheric Measurement Techniques, 14(1), 481–510. https://doi.org/10.5194/amt-14-481-2021

Judd, L. M., Al-Saadi, J. A., Szykman, J. J., Valin, L. C., Janz, S. J., Kowalewski, M. G., et al. (2020). Evaluating Sentinel-5P TROPOMI tropospheric NO2 column densities with airborne and Pandora spectrometers near New York City and Long Island Sound. Atmospheric Measurement Techniques, 13(11), 6113–6140. https://doi.org/10.5194/amt-13-6113-2020

Li, M., McDonald, B. C., McKeen, S. A., Eskes, H., Levelt, P., Francoeur, C., Harkins, C., He, J., Barth, M., Henze, D. K., Bela, M. M., Trainer, M., de Gouw, J. A., and Frost, G. J.: Assessment of Updated Fuel-Based Emissions Inventories Over the Contiguous United States Using TROPOMI NO2 Retrievals, Journal of Geophysical Research: Atmospheres, 126, e2021JD035484, https://doi.org/10.1029/2021JD035484, 2021.

References for our replies:

Atkinson, R. (2000). Atmospheric chemistry of VOCs and NOx. *Atmospheric Environment, 34*(12-14), 2063-2101. 10.1016/s1352-2310(99)00460-4

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., et al. (2018). Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project. *Atmospheric Measurement Techniques, 11*(4), 2395-2426. 10.5194/amt-11-2395-2018

Judd, L. M., Al-Saadi, J. A., Szykman, J. J., Valin, L. C., Janz, S. J., Kowalewski, M. G., et al. (2020). Evaluating Sentinel-5P TROPOMI tropospheric NO2 column densities with airborne and Pandora spectrometers near New York City and Long Island Sound. *Atmospheric Measurement Techniques, 13*(11), 6113-6140. 10.5194/amt-13-6113-2020

Li, M., McDonald, B. C., McKeen, S. A., Eskes, H., Levelt, P., Francoeur, C., et al. (2021). Assessment of Updated Fuel-Based Emissions Inventories Over the Contiguous United States Using TROPOMI NO2 Retrievals. *Journal of Geophysical Research-Atmospheres, 126*(24). 10.1029/2021jd035484

Verhoelst, T., Compernolle, S., Pinardi, G., Lambert, J. C., Eskes, H. J., Eichmann, K. U., et al. (2021). Ground-based validation of the Copernicus Sentinel-5P TROPOMI NO2 measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global networks. *Atmospheric Measurement Techniques, 14*(1), 481-510. 10.5194/amt-14-481-2021

Vigouroux, C., Langerock, B., Aquino, C. A. B., Blumenstock, T., Cheng, Z. B., De Mazière, M., et al. (2020). TROPOMI-Sentinel-5 Precursor formaldehyde validation using an extensive network of ground-based Fourier-transform infrared stations. *Atmospheric Measurement Techniques, 13*(7), 3751-3767. 10.5194/amt-13-3751-2020

Williams, J., & Koppmann, R. (2007). Volatile Organic Compounds in the Atmosphere: An Overview. In *Volatile Organic Compounds in the Atmosphere* (pp. 1-32).

Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., et al. (2018). Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions. *Atmospheric Chemistry and Physics, 18*(19), 14095-14111. 10.5194/acp-18-14095-2018