Responses to the reviewer 3

The review comments and our responses are coloured in blue and black texts, respectively. The changes in the manuscript corresponding to the comments are highlighted in yellow. The line and figure numbers refer to the revised manuscript.

In this study, Hoque et al. evaluated the global distribution of formaldehyde (HCHO) simulated by the CHASER v4.0 model against satellite, aircraft, and ground-based observations. Studies evaluating the global distribution of volatile organic compounds (VOCs) from models have been limited, and observations from space have the potential to help filling this gap. The investigations presented here add new insights, nevertheless, the manuscript in its current form has some limitations.

We thank the reviewer for the insightful comments, which has helped improving the manuscript quality.

The authors have already published a paper on the global distribution of HCHO from the CHASER model comparing with satellite data and MAX DOAS [Hoque et al., Atmos. Chem. Phys., 2022]. Sensitivity simulations analyzing the roles of different emissions have also been published there. In this case, a clear and detailed discussion is required (at the end of the introduction) on the main findings of that paper, the research gap, and the novelty of this new study.

Response: We have added the following texts in the revised manuscript.

L93-103 : Hoque et al. (2022) validated CHASER-simulated NO_2 and HCHO against OMI and MAX-DOAS observations for 2017. CHASER showed good skills in reproducing the OMI- (spatial correlation ® = 0.74) and MAX-DOAS- (temporal correlation R> 0.80) observed HCHO abundances. The study found that biomass burning contributes ~50% to the HCHO levels observed at the site in Thailand. However, the limitations of the study are: (1) Simulated HCHO partial column and profile were evaluated against MAX-DOAS observation on a seasonal scale only, (2) Model sensitivity studies were site-specific, thus providing no global statistics on emission contribution, and (3) Satellite observations were used as supporting datasets; thus the model-satellite comparison has not been comprehensive. This study utilizes multi-satellite (TROPOMI and OMI) HCHO observations, different NO_x emission inventories, aircraft measurements, and daily and diurnal MAX-DOAS data to provide robust and comprehensive statistics on the model HCHO simulations.

Abstract: l.19: "CHASER reproduced the observed….", which observational data you are referring to?

Response: We have revised the sentence as follows:

CHASER reproduced the TROPOMI-observed global HCHO spatial distribution with a spatial correlation (r) of 0.93 and a negative bias of 7%.

Introduction

"ozone production regime can be determined". In this context, you are referring to your past study. References where this type of approach was proposed [Martin et al., GRL, 2004] and later applied [Duncan et al., Atmos. Environ., 2010] should also be cited. Additionally, l.57-58: Several satellite-based observations have been used to evaluate the model simulation of HCHO by Chutia et al., [Environ. Poll., 2019]

Response: We thank the reviewer for the suggestion. We have included the references in the revised manuscript.

l.59-65: I agree that higher resolution TROPOMI may provide new features at finer resolution (3.5 km x 5.5 km). But how does it help your study running model at roughly 300 km x 300 km? Satellite data also seems to be averaged to the same grid resolution as the model, although that is not described in detail.

Response: We have added the following texts in response to this comment

L196-205 TROPOMI observations are averaged spatially and temporally to the CHASER grid (T42) daily, leading to horizontal representativeness errors. However, the random horizontal representativeness errors are in the order of 5-10%, which is lower than the individual retrieval error of the satellite observations (Boersma et al., 2015). If the model horizontal resolution is increased by 50% (i.e., simulated at a horizontal resolution of 1.4º × 1.4º), the change in HCHO abundances is less than 6% (Fig S1 and Table S1 in supplementary information). The vertical sensitivity of the satellite retrievals is the most relevant source of representativeness error (Boersma et al., 2015). The current study utilizes the TROPOMI AK information to minimize the representativeness error. Therefore, the horizontal representative error will likely affect the results less than other error sources, such as uncertainties in satellite retrieval, emission inventories, and model chemical mechanisms.

l.72-74: Mention what has been learned from these studies, possibly the quantitative role of anthropogenic emissions.

Response: We have revised the sentences as follows:

HCHO products have been used to infer changes in the global HCHO levels during the COVID-19 pandemic-led shutdown (Level et al., 2022; Souri et al., 2021; Su et al., 2021), demonstrating the role of anthropogenic emission on global HCHO variability.

Section 2: Model

Anthropogenic emission is representative of which year. How has it been varied for different simulation years (2019, 2020)?

Response: we have revised the sentences as follow:

L129-137Anthropogenic NO_x emissions for 2018 are obtained from the HTAP_v3 inventory (Crippa et al., 2023). Other anthropogenic emissions are taken from the HTAPv2.2 for 2008 and the biomass burning emissions from MACC-GFAS (Inness et al., 2013). The monthly soil NO_x emissions derived from Yienger and Levy (1995) are constant each year. Biogenic emissions of VOCs are obtained from a process-based biogeochemical model: the Vegetation Integrative Simulator for trace gases (VISIT) (Ito and Inatomi, 2012). VISIT is a part of the CHASER modeling framework and incorporates the biogenic flux estimate scheme of Guenther et al. (1997) (Ito et al., 2022). The global isoprene emissions in VISIT and CAMS global biogenic emission inventory (Sinderolova et al., 2022; based on MEGANv2.1) are 400 and 450 TgC/yr, respectively.

l.111-112- Seems ambiguous. The reanalysis data might not have provided an emission inventory. Maybe there is some inventory from the same or similar project.

Response: We have revised the sentences as provided in the earlier response.

l.113- The VISIT model is used here for estimating the flux of biogenic VOCs. How do these estimates compare to other widely applied MEGAN model (Guenther et al., Atmos. Chem. Phys., 2006) based inventories? ECMWF's CAMS has made freely available inventory for biogenic emissions, which may be used for comparison.

Response: We have revised this section and incorporated additional information of the biogenic flux.

VISIT is a part of the CHASER modeling framework and incorporates the biogenic flux estimate scheme of Guenther et al. (1997) (Ito et al., 2022). The global isoprene emissions in VISIT and CAMS global biogenic emission inventory (Sinderolova et al., 2022; based on MEGANv2.1) are 400 and 450 TgC/yr, respectively.

Table 1: Some simulations are missing from this list, like in which biogenic / biomass-burning is switched OFF.

Response: We have revised Table 1. And included all the simulations used in the study

l.196: Check and correct "TOGO" to "TOGA"

Response: We have corrected the typo error.

Results

Results from two different years appear identical in Figure 1. In the text also, there is no significant discussion on interannual differences. Values of correlation and RMSE have turned out to be the same between the analysis for 2 years separately. I suggest combining and discussing averages of both years for better statistics and reducing extra figures. This will further make this analysis consistent with follow-up results (e.g., Figures 2, 3), where mean is presented instead of year-wise segregation. The size of the figures can be enhanced.

Response: We thank the reviewer the comment. We also agree with the reviewer's perspective. We have revised (enlarged) Figure 1, but decided to retain the comparison for individual years. The reasons are-

(1) We believe, it is important to demonstrate the model-satellite agreement in both years to ensure consistency of our model.

(2) We have used one year of simulation in the sensitivity studies. Thus, information on the model performance in the individual years is important for supporting the results.

Figure 2 and other results: you are referring to different regions of the world. These need to be marked clearly on the global distribution map (Figure 1). Make bigger figures and define the regions on them.

Response. We thank the reviewer for this important suggestion. We have included a new figure (Fig.S3) in the supplementary information showing the region of interests. Moreover, we have redefined the regions of C-Africa, Europe, and India to be consistent with earlier studies with CHASER. We have also recalculated the statistics for these regions and made

appropriate corrections in the manuscript. There redefinition did ‘not change the statistics and discussion significantly.

l.293 (and throughout the manuscript), be careful to always mention if you are referring to "spatial" or "temporal" correlations while reporting r values.

Response: We thank the reviewer for the important comment. We have ensured consistency in the correlation reporting throughout the manuscript.

l.300-301: This needs some supporting analysis/discussions. Either compare the emission inventory used here with other estimates or discuss if the model is underestimating particularly near urban centers (so to attribute to anthropogenic) but performing better in remote / vegetated areas.

Response: We have included the following sentences to address this comment

L335-338 NMVOC emissions from these sources (i.e., vehicular exhaust, solvent usage, and transport) are considered in the HTAPv2.2 inventory (Crippa et al., 2023). Although CHASER considered HCHO production from the degradation of anthropogenic VOCs, it is likely underestimated, resulting in a lower simulated winter-time HCHO column in this region.

Table 3 and other places: Are the temporal correlations derived from the mean seasonal cycle (12 points)? It is advisable to use all data (daily values over 2 years) to comment on temporal correlations. Or to discuss both ways. This is a "model evaluation paper" and these details are important.

Response: Yes, the temporal correlation has been calculated from the seasonal variation. We have included the temporal correlation estimated from the daily values in Table S2(supplementary information).

Table 4 and l.450: here also, clearly write if these are spatial correlations, seasonal (or daily). Check and make this aspect clear throughout the manuscript.

Response: We have included appropriate changes in the manuscript to address this comment.

l.471-472: here also check from MEGAN model-based emissions.

Response: The OMI-based top-down estimate is based on the MEGAN model. We have revised the sentence as follows:

L517-519 In CHASER, annual isoprene emissions over Amazonia are 67 Tg/yr, consistent with the OMI-based top-down estimates of 70 Tg/yr, estimated using apriori emissions from

MEGAN (Stavrakou et al., 2015).

Figure 5: I did not get the rationale behind enhancing anthropogenic emissions by a factor of 3. While HCHO was underestimated in reference simulation, now with this change the levels are equally (or more) overestimated over China, US, Africa, America (also see table 5). What has been achieved in terms of model performance?

Response: We thank the reviewer for the comment. We perturbed the anthropogenic VOC emissions to assess it's effect on the model-satellite comparison. Multiple simulation was performed and we found that, the perturbation effect is relevant when the anthropogenic VOC emissions are increased at least three-fold. Thus, we selected the lowest value (i.e., three-fold increase). If the perturbed simulation improves the model-satellite agreement, it can be interpreted as underestimated anthropogenic VOC emissions in standard simulation. Section 3.3 has been revised to address the reviewer comments.

l.503-504: No, the MBE values have increased! Check and revise/strengthen this whole section l.503-518. Also, reconsider tuning the simulation design itself (in place of 3 times more emissions)

Response: We have revised section 3.3 following the reviewer's earlier comment.

Section 3.4: This is an important aspect. Errors in the NOx emissions could have impacted model performance, especially in regions like South and Southeast Asia where inventories have greater uncertainties. Your simulations show that the model driven by older inventory shows lower bias in HCHO. Do you conclude that NOx emissions are overestimated in the new inventory? I did not find a clear assessment out of this important exercise.

Response: We have revised this section as follows:

L610-628The standard HCHO columns in India, China, and Southeast Asia are approximately 10–20% lower than the OLNE estimates (Fig.6(c)). In fact, those differences are consistent with changes in the regional OH estimates (Fig.6(d)). This finding implies that the changes in the NO_x emissions estimates have affected the OH and HCHO abundances in these regions. Satellite data assimilation results reported by Miyazaki et al. (2017, 2020) indicate that NO_x emissions in India have increased by 30% since 2008, whereas NO_x emissions in China have declined since 2011 (Liu et al., 2016). Over E-China (Fig. 6(a &b)), the standard simulations reduce the absolute annual mean difference between OLNE and TROPOMI of 3×10^{15} molecules cm^{-2} to 1×10^{15} molecules cm^{-2}, which is consistent with the lower NO_x emissions in this region in the updated inventory (Fig . S8). Over India and SE-Asia, the standard OH concentrations are ~40% lower (Fig.6(d)) than the OLNE estimates, resulting in lower HCHO columns. The lower standard HCHO columns can be linked to the increasing NO_x emissions in these regions (Fig.S8); however, the magnitude of

the change in the NO_x emissions for these regions in the updated inventory is likely overestimated.

In E-USA and W-USA (Table S3), the standard simulation reduces the MBE by 26% and 12%, respectively. The reduction in MBE and RMSE values in Africa and South America is less than 10%. Therefore, NO_x emission uncertainties mainly affect the HCHO simulations in India and SE Asia.

Tables and figures coming afterward often has data shown in previous figures and tables. Review them carefully and combine them whenever possible. Like, instead of comparing 1 simulation then another, you may put them in same table as reference, simulation1 and 2;

Response: Redundant tables has been moved to the supplementary information.

CHASER and TROPOMI are already compared (Section 3.1). Then there is an extra section comparing CHASER, TROPMI with OMI. Better to combine and strengthen the discussion.

Response: We have revised the subsection headings to avoid confusion.

Fig 8: When emission is increased (OLNE), why HCHO is reduced over Chiba and Kasuga, what is the underlying chemistry?

Response: This has been discussed in the manuscript as follows:

L812-817 Although the bias between OLNE and standard simulations for Chiba and Kasuga is ~4%, the absolute difference is ~1×10^{15} molecules cm^{-2}. NO_x emissions in Japan have not changed markedly since 2005 (Miyazaki et al., 2017). The differences between the simulations are observed during the summer when isoprene emissions are expected to peak (Hoque et al., 2018a). Because the OH estimates over Japan are similar for both simulations (Fig. 6(d)), the differences are likely related to the interaction between isoprene and NO_x inventories.

Outside Japan also, there have been MAX-DOAS measurements. This paper being a global model evaluation, comparison over other regions of the world should also be added. If systematic data is not available, mean values may be compared (see Table 2 of Oomen et al., Atmos. Chem. Phys., 2024). Authors themselves have also published observations from another station in South Asia [Hoque et al., SOLA, 2018]

Response: We thank the reviewer for the comment. We are currently unable to include more A-SKY sites for the following reasons:

(1) Most of the A-SKY sites outside Japan were established after 2020, which is beyond the temporal limit of the current study. Moreover, the retrieved quantities for these sites are being investigated in detail in a separate project.

(2) Due to the mountainous terrain, we excluded the Pantnagar site (i.e., Hoque et al., SOLA, 2018).

However, we have included a comparison with the reported MAX-DOAS values by Oomen et al (2024) as shown in Figure 9.

Minor comments

Check the consistency of r values between l.352 and in Table 3.

Response: We have made appropriate changes in the manuscript

The data selection criteria for TROPOMI have not been discussed.

Response: We have added the following text in the revised manuscript

L194-195The filtering criteria of the TROPOMI datasets are as follows: quality assurance value (QA)>0.6, solar zenith angle <70º, cloud fraction < 0.3, AMF > 0.1, and surface reflectivity <0.2.

Line 387, line no. 467 - Figure no. is missing

Response: The figure has been included which is FigS6

In Figure S3, the correlation between TROPOMI and CHASER HCHO columns is marked as r=1 (blue text).

Response: We have revised the figure, which is FigS4 in the revised manuscript

Figure numbering should be corrected. There is no figure 4.

Response: We have ensured consistency in the Figure numbers and the corresponding texts.

Table 7 , last column name should have been 'r-value (CHASER vs. OMI)'

Response: We have revised Table 7

Line 781, slope values inconsistent with the slope in figure 9.

Response: We have ensured consistency in the Figure and texts.
