Response to Anonymous Referee #2.

We thank the reviewer for the constructive comments and for providing suggestions to improve the manuscript. Our responses are shown in blue and added text is shown in italics.

This study investigates the impact of model biases in CO on the ability to simulate oxidants observed during the KORUS-AQ campaign. In particular, it explores the effects of assimilating MOPITT CO as well as using different emissions inventories. The question of how improving simulated CO alters other aspects of atmospheric chemistry is an interesting one, and the study applies state-of-the art tools to address it. The methodology is rigorous, but I have some suggestions, described in the general and specific comments, for how the results could be presented more clearly.

General Comments

The study includes quite a few model simulations, both with and without MOPITT assimilation and different prior or posterior anthropogenic and biogenic emissions inventories. Further simulations with different biogenic emissions are also discussed in the supplemental material. While I understand the need for multiple sets of simulations to explore different ways of addressing CO bias, it can be difficult for the reader to keep track of what is included in each one and which lines to compare on figures (e.g. Fig. 9) showing many different model simulations. I suggest that the discussion in Section 6 (and some of the figures) would be easier to follow if that section were divided into two parts, with one part focused on the impact of MOPITT data assimilation, and the other part focused on the role of the emissions inventories.

The manuscript devotes as substantial amount of text to describing aspects of the model and assimilation system that do not seem directly tied to the discussion of the results. Some of this information, such as Section 4.2, could be condensed and/or moved to the Supplement or Appendix.

Appendix B contains a considerable amount of information and analysis, but I did not find it referenced in the main text.

The evaluation of NOx and other species in Section 2 of the supplement, and Fig. S2 in particular, is quite relevant to the interpretation of the results. It would be helpful to include some of this material in the main text.

Thanks, as suggested, we divided the section 6 in two. The section 6 is now: ***Comparison of anthropogenic emission estimates*** and section 7 is: ***Evaluation of the simulated vertical profiles against ARIAs and KORUS-AQ***.

We completely rewrote the section 5.2, and added references to Fig. S2 and Appendix B. the evaluation of NOx is now included in Figure 4 and is discussed in the text.

There are new sentences that refer to appendix B, In section 3.1:

"A summary of the model references is presented in Table 1. We have made some additional changes for this study, presented in Appendix B. In particular, we updated the heterogeneous uptake coefficient of HO2 and its coefficient."

In Sect. 5.2, we added:
“Despite the update of the HO2 heterogeneous uptake reaction and coefficient presented in appendix B, […]”
“A lower value of the HO2 heterogeneous uptake coefficient than the one used here (γ=0.1) might produce better results by reducing the HO2 sink (see Appendix B).”
“While the errors in NOx are important, the low CH2O points to a missing source, which could be due to an underestimation of CH4 as well as NMVOCs (Appendix B).”

Specific Comments

Line 62: Is “pollution ozone” the same as “ozone pollution”?
We changed it to ozone pollution.

Line 460: Why is met data assimilation used in one set of runs and nudging in another set? Does this lead to transport differences between the two sets of runs?
The assimilation is done using an Ensemble Adjustment Kalman Filter, that requires to run an ensemble of simulations, thus the assimilation is much more computationally expensive to perform.
The two assimilation runs perform well as described in Fig. S3, S4, S5, and overall provide a better meteorology than the nudging runs. Here is an update of table S1, now Table S2 with the Control-run added:

Table S2: Overall statistics for Wind Speed, H2O, and Temperature.

Wind Speed	simulations	RMSE	r	Bias
	Control-Run	8.44	0.67	1.02
	prior-d-0.24	9.21	0.61	1.32
	prior-d-0.48	8.96	0.63	1.22
	prior-d-0.72	8.87	0.63	1.11
	prior-m-0.24	8.89	0.64	1.32
	prior-m-0.48	8.61	0.65	1.01
	prior-m-0.72	8.56	0.66	0.92
H_2O	simulations	RMSE	r	Bias
	Control-Run	2124.7	0.91	-123.4
	prior-d-0.24	2990.14	0.86	766.62
	prior-d-0.48	2570.17	0.89	498.83
	prior-d-0.72	2459.9	0.89	396.89
	prior-m-0.24	2429.31	0.89	32.67
	prior-m-0.48	2191.41	0.91	-68.4
	prior-m-0.72	2126.97	0.91	-98.73
Temperature	simulations	RMSE	r	Bias
	Control-Run	2.65	0.99	1.09
	prior-d-0.24	2.95	0.98	1.09

prior-d-0.48	**2.88**	**0.98**	**1.03**
prior-d-0.72	**2.85**	**0.99**	**1.01**
prior-m-0.24	**2.88**	**0.98**	**0.96**
prior-m-0.48	**2.81**	**0.99**	**0.93**
prior-m-0.72	**2.79**	**0.99**	**0.94**

We believe an entire study could be dedicated to this interesting topic of research.

Table 2: The simulation names should be made more informative/intuitive. For example, only one is called "Posterior" even though multiple runs use Posterior emissions.
We updated the Table 2 and simulation names as follows. We also made it clear in the text, and contrast those in section 6 which compare the emissions and section 7 which compares a suite of deterministic runs forced by different anthropogenic CO inventories.

Table 2: Summary of the simulations. The Nudging (GEOS) refers to a CAM-Chem deterministic run with specified dynamics, using a nudging to GEOS-FP analysis winds and temperatures (see supplement). Aside from the DART simulations (first 2 rows), all the simulations have the same initial conditions and the same nudging and only change by their anthropogenic CO emissions inputs.

Simulation name	***Meteorology***	***Emissions (prior)***
Control-Run	*Assimilation (DART)*	*Prior (CEDS-KORUS-v5)*
MOPITT-DA	*Assimilation (DART)*	*Optimized (CEDS-KORUS-v5)*
CAM_Kv5	*Nudging (GEOS)*	*Prior (CEDS-KORUS-v5)*
CAM_HTAP	*Nudging (GEOS)*	*Prior (HTAP v2)*
CAM_MOP	*Nudging (GEOS)*	*Posterior (CEDS-KORUS-v5)*
CAM_MOP-Bio	*Nudging (GEOS)*	*Posterior (CEDS-KORUS-v5) + MEGANx2 (see SI)*
CAM_TCR-2	*Nudging (GEOS)*	*Posterior (TCR-2, HTAP v2)*
CAM_CAMS	*Nudging (GEOS)*	*CAMS (CAMS-GLOB-ANTv3.1)*

Line 557: Does this mean it's the old RMSE minus the new one, or something else?
Yes, it does, but relative to the old RMSE. We updated the text in the figure captions:
"The relative error change is the opposite of the difference in Root Mean Square Error relative to the Control-Run (i.e., (Control-Run-MOPITT-DA)/Control-Run). Thus, a positive relative error change means an improvement compared to the Control-Run."

Lines 591-592: Please elaborate on this statement.
The paragraph has been updated and the sentence now reads:
"It suggests that errors in NOx and related chemistry drive the underestimation of HO2 and of the sum of OH and HO2 (HOx). Overall, HOx is underestimated, and OH is fairly well simulated. This suggests that the CO chemical sink alone cannot explain the CO underestimation during the campaign."

Line 598: I expect that biases in NOx, PAN, etc. can strongly impact the comparison of the simulated oxidants, so I think this point deserves more discussion in the body of the text. Also, Fig. S2 shows that the MOPITT-DA makes H2O2 worse, and this should also be mentioned.

We completely rewrote the section, added NOx comparison to the figure 4 and mentioned the figure S2 with and discussed the H2O2 issues. See also the responses to Reviewer 1.

Line 659: Can you speculate on why the emissions are underestimated?
The end of the paragraph now reads:
"More work should be dedicated to check whether the assumptions made on the prior estimates impact the retrieved emissions. This includes improving the regional distribution and scaling up the baseline prior CO emissions, but also how much the model uncertainties in the OH chemical sink impact the CO inversions (e.g., Müller et al., 2018). A comparison of the amount of residential coal burning emissions in bottom-up inventories could help in understanding the discrepancy and quantifying potential offsets (Chen et al., 2017; Cheng M., et al., 2017; Zhi et al., 2017, Benish et al., 2020)."

Line 698: Aren't VOCs also higher in this simulation? If so, how do you attribute the ozone production to CO rather than other VOCs?
VOCs are changed in the MOPITT-DA and in the CAM_MOP-Bio (aka CAM-Chem-Bio). The 5 simulations CAM_CAMS, CAM_Kv5 (DART-CAM-Chem prior), CAM_MOP (DART-CAM-Chem Posterior), CAM_HTAP (Prior TCR-2) and CAM_TCR-2 (posterior TCR-2) only differ by their CO anthropogenic emissions and share the same meteorology. The O3 is further improved with additional biogenic VOCs that fuel ozone production, as discussed in the following section.

Line 746: The definition of the tags needs to be explained somewhere.
In the end of section 4.6, we added:
We included artificial CO tracers or "CO tags", to track the anthropogenic contribution from different geographic area sources (e.g. Gaubert et al., 2016).

Line 777: Does this mean there is a stratospheric intrusion reaching all the way to the surface? Or just more stratospheric/upper tropospheric influence somewhere in the profile?
We believe you refer to L767, Weather phase 2 of the campaign. The CAM-chem stratospheric ozone tracer suggests enhancements of stratospheric O3 in the free troposphere, disconnected from the enhancement observed at the surface / in the boundary layer.

Fig. 9: It would be helpful to include panels with the CO profile here so readers don't have to keep referring back to Fig. 8.
We believe that the Fig. 8 should be placed close enough to Fig. 9 in the manuscript.

Editorial Comments
Line 204: "efficient" or "effective"?

Thanks, we changed efficient to effective.

Line 596: Sentence needs rewording
The new sentence now reads:
"In the lower part of the atmosphere, the oxidation of additional CO leads to more effective ozone production and no changes above, consistent with observations."

Lines 652-654: Sentence needs rewording.
The new sentences now read:
"Our inversion suggests an underestimation of bottom-up emission inventories for China. The agreement between the posterior emissions for Central China is better than for the bottom-up inventory (Fig. 6)."

Line 766: It would be clearer to avoid the parentheses and say lower OH and higher ozone.
The new sentence reads:
"This anomaly is reflected through lower OH and higher O_3 between 800 hPa and 400 hPa (Figure 9)."