Referee #2
General Comments:

This study provides a comparative analysis by assimilating different MOPITT CO datasets to evaluate the advantages of assimilating multi-spectral satellite retrievals. This topic is important for better usage of MOPITT observations in the future. However, the manuscript focuses on the demonstration of assimilation results but lacks a further explanation for the possible causes of their large discrepancies, which reduces the importance of this manuscript. Furthermore, the content of this manuscript may need to be better organized. For example, the title of Section 4 is "Results". The audience may assume it is the major discussion section, however, it is followed by Sections 5-7. In addition, there are lots of typo errors. Overall, this is a good study with potentially important impacts on the application of satellite data. However, a significant improvement may need to be made before the paper can be considered for publication.

Response: Thank you. We agree with the reviewer and have re-organized the manuscript accordingly.

Specific Comments:

1. The authors raised three questions about the potential advantages of assimilating multispectral/joint retrievals. The developed and potential multispectral satellite retrievals are demonstrated in Table 1, including species of CO, O3, and CH4. More discussions are suggested to clarify the possible limitation of this analysis, i.e., whether the conclusion based on MOPITT CO is applicable to other instruments and species.

Response: As introduced in the introduction, regarding multispectral capabilities, MOPITT is a unique instrument as it retrieves total column amounts and vertical profiles of CO using both thermal-infrared (TIR) and near-infrared (NIR) measurements, and provides a multispectral TIR-NIR joint product. Therefore the MOPITT instrument is an ideal instrument to demonstrate the value of assimilating multi-spectral satellite retrievals of atmospheric composition and address these three questions. This study using MOPITT as a demonstration sheds light on the impacts of assimilating different satellite products of the same atmospheric composition. To address the reviewer's comment, we added the following statement to Section 7.4 where we discuss limitations:

"This study provides guidance for future work on the assimilation of multi-spectral satellite retrievals of atmospheric composition using MOPITT as a demonstration. However, whether the conclusions based on MOPITT CO are applicable to other species (e.g., CH4 and O3) needs further study."

2. The authors demonstrate dramatic discrepancies among different assimilations but lack a further explanation for the possible causes. For example, Figure 5: there are very large differences in the upper troposphere level. What are the possible causes and what are the possible impacts on the future assimilation of MOPITT CO data? Figure 10: Which assimilations have better performance in the upper troposphere, given their large discrepancies as shown in Figure 5? Section 6.1: The authors state "Assimilating profile products (Experiments (2) and (5)) tends to have a larger change to the emissions compared to only assimilating column products (Experiments (1), (3), and (4))". However, the change by assimilating Profile JNT data is very small, which cannot be described as "tends to have a larger change". The large discrepancy in Figure 13 by assimilating different datasets needs to be explained. Section 6.2: same as the above question, the change by assimilating Profile JNT data is almost ignorable, I don't understand why it could lead to better agreement with observations.

Response: Here we use ensembles to correct model priors with observed quantities. Localization is commonly used in ensemble-based data assimilation to address insufficient ensemble samples to ameliorate the spurious long-range correlations between the background and observations. In this study vertical localization is used. The main difference between profile assimilation and column assimilation is the involvement and treatment of the vertical information. For each MOPITT retrieval, profile products have multiple observations at different layers but their impacts are vertically localized around 100 hPa. Therefore not all vertical layers will be impacted. For the column, all vertical levels will be impacted by a single column value. In this case, if the mismatch is due to an underestimation of surface emissions rather than weak vertical transport, updating the upper tropospheric CO might lead to erroneous adjustments in CO abundance. In this study, the apparent contradiction between profile assimilation and column assimilation demonstrates the impacts of the inclusion and treatment of vertical information. Figure S4 shows the vertical profile increment for the experiments.

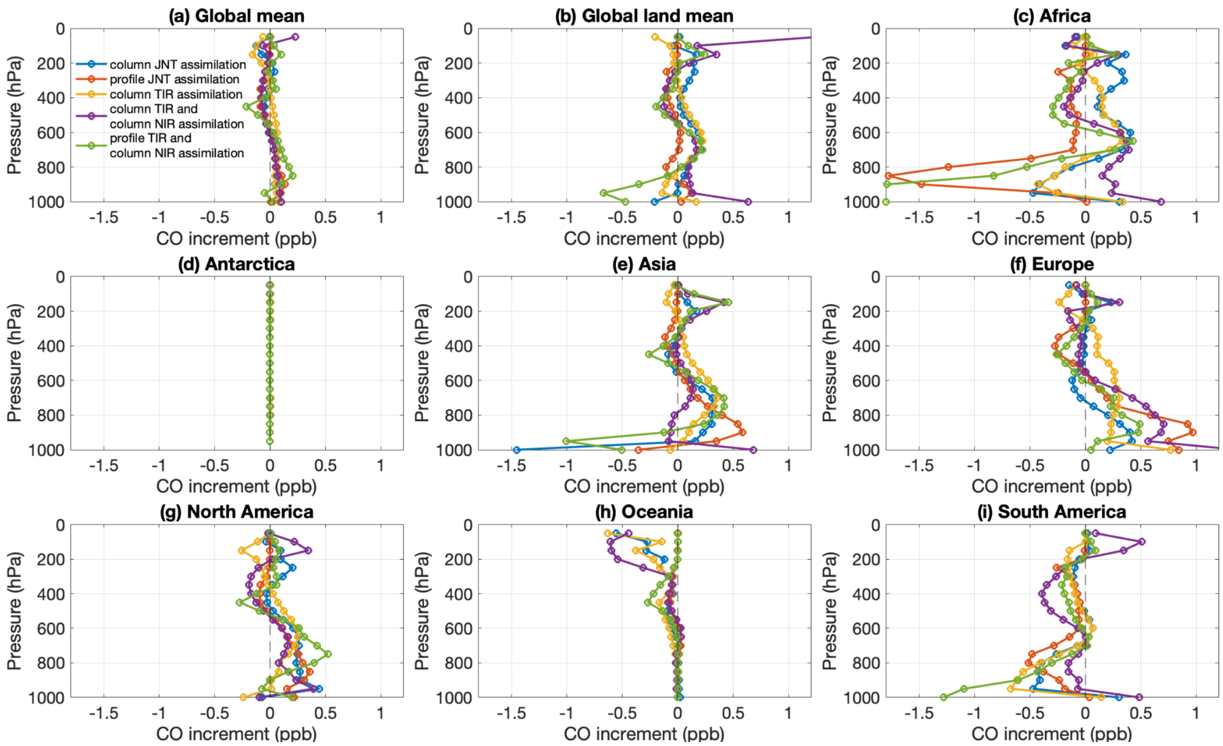

Figure S4. Vertical profile of the 15-day (July 31 - August 14, 2018) average increment in CO. Increment is calculated as analysis minus forecast.

We added more discussion in the manuscript. We added the following description of "vertical localization" in the manuscript. "*Localization is commonly used in ensemble-based data assimilation to address insufficient ensemble sample size. Since the correlation is expected to decrease as separation increases, it empirically reduces the impact of an observation on model state variable as a function of distance using the Gaspari–Cohn localization function. The spatial localization horizontal half width is 600 km and the vertical half width is 1200 m. The main difference between the profile and the column assimilation resides in the vertical localization. For each MOPITT retrieval, profile products have multiple observations at different layers but their impacts are vertically localized around 100 hPa. Therefore, not all vertical layers will be impacted. For the column data assimilation, there is no vertical localization in the column data assimilation except that the stratospheric (top 5) levels are not updated, as in the CO profile and meteorological DA. All vertical levels will be impacted by a single column value. In this case, if the signal is coming from an underestimation of surface emissions, correcting the upper troposphere might be wrong.*"

For Figure 5 (now Figure 6), we added "At 200 hPa, the spatial distribution of the CO difference caused by assimilation is smallest in Exp2-PJ, followed by Exp5-PT+CN. On the contrary, for the other three experiments which do not involve profile assimilations, the spatial distribution of the

CO difference caused by assimilation is relatively large. I.e., assimilating MOPITT profile product(s) only slightly changes CO values at 200 hPa whereas assimilating MOPITT column product(s) changes CO values at 200 hPa dramatically. This is expected as vertical distribution is often an advantage of profile DA that column can not represent."

For Figure 10 (now Figure 11), we added "Above 200 hPa, all five experiments overall agree better with IAGOS CO compared to the control run. However, experiments involving profile assimilation do not show obvious differences compared to experiments only involving column assimilation above 200 hPa."

For Section 6.1 and Section 6.2: The change by assimilating Profile JNT data is not small, and "tends to have a larger change" is the appropriate description as shown by Figure 12 (now Figure 13) (red line).

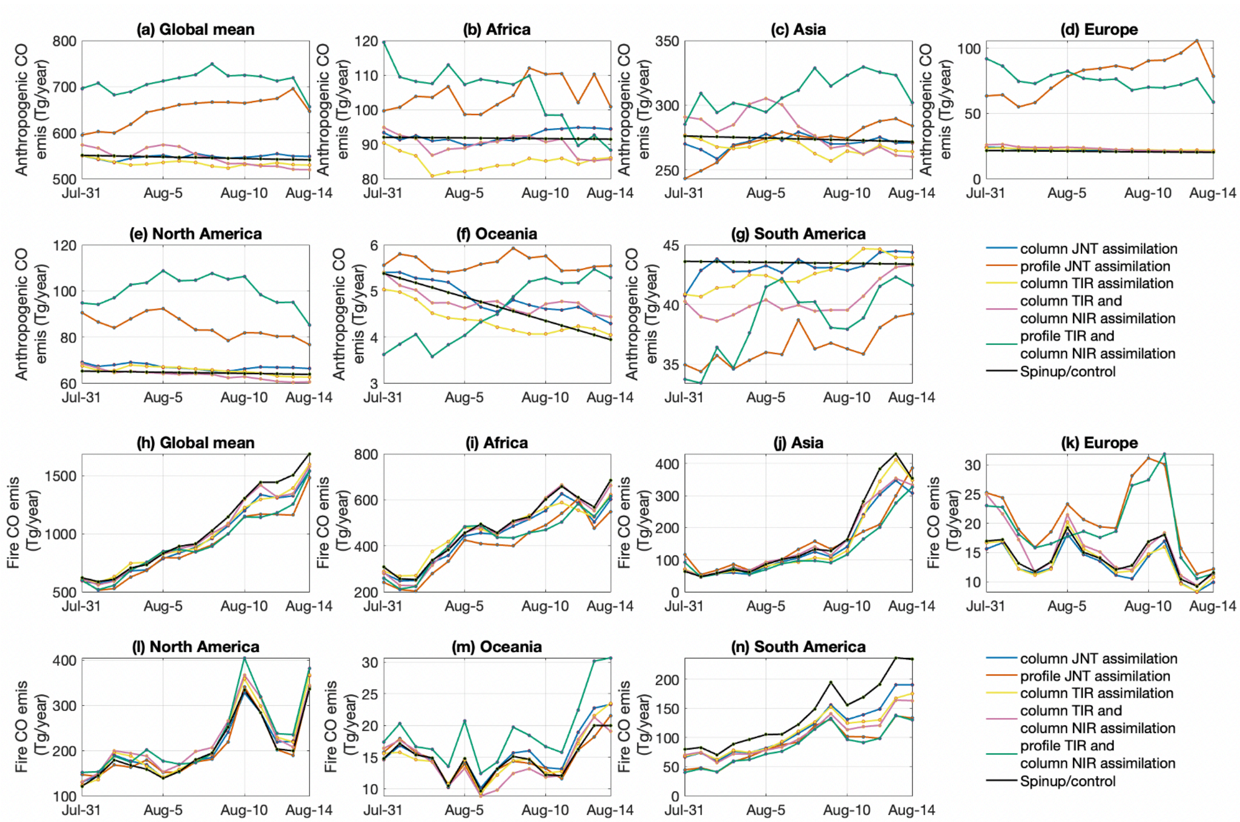

Figure 13. Updated (a-g) CAMS anthropogenic CO emissions and (h-n) FINNv2.4 fire CO emissions as a result of assimilating different MOPITT products. The emissions from the Spinup/control run are the unchanged original emissions of CAMS and FINNv2.4.

For Figure 13 (now Figure 14), we added "As shown previously, profile assimilation can out-perform column assimilations near the surface due to vertical localization. Different CO concentrations at and near the surface resulted in different emission updates between profile assimilation and column assimilation."

3. There are lots of typo errors. It seems the manuscript was not checked carefully. For example, Lines 421-423: "(1) Column JNT assimilation, (2) Profile JNT assimilation, (3) Column TIR assimilation, (4) Column TIR and column NIR assimilation, and (5) Profile TIR and column NIR assimilation". Why are the five experiments listed here? Line 568 in the conclusion section, the authors state: "Results were not improved compared to WE-CAN because …" Because what?

Response: Thank you for pointing this out. We have carefully revised the manuscript to remove the typos.

Technical Comments:

4. Line 32: "vertical localization" is frequently used in the manuscript, but the meaning is unclear.

Response: We added the following description of "vertical localization" in the manuscript. "Localization is commonly used in ensemble-based data assimilation to address insufficient ensemble sample size. The error in the relation between an observation and a state variable is expected to increase as correlation between them decreases. Localization empirically reduces the impact of an observation on model state variable as a function of distance since correlation is expected to decrease as separation increases. The spatial localization horizontal half width is 600 km and the vertical half width is 1200 m. The main difference between the profile and the column assimilation resides in the vertical localization. For each MOPITT retrieval, profile products have multiple observations at different layers but their impacts are vertically localized around 100 hPa. Therefore not all vertical layers will be impacted. For the column data assimilation, there is no vertical localization in the column data assimilation except that the stratospheric (top 5) levels are not updated, as in the CO profile and meteorological DA. All vertical levels will be impacted by a single column value. In this case, if the mismatch is due to an underestimation of surface emissions rather than weak vertical transport, updating the upper tropospheric CO might lead to erroneous adjustments in CO abundance."

5. Table 1: what species are observed by the Geostationary satellites?

Response: We have updated Table 1. Please see below.

Table 1. Developed and potential multispectral satellite retrievals. Shown in the table are satellites, their NIR and/or TIR spectral ranges (in µm), and potential chemical species from the multispectral retrievals.

Morning Overpass	Afternoon Overpass	Geostationary
MOPITT (2.3 & 4.7) (CO)	AIRS (3.75–15.4) + OMI (0.27–0.5) (O3)	GIIRS (East Asia) (0.55–14.2) + TROPOMI (2.3–2.4) (CO, O3)
IASI (3.6–15.5) + GOME2 (0.24–0.79) (O3)	TES (8.7–10.5) + OMI (0.27–0.5) (O3)	GEMS (East Asia) (0.3–0.5) + IASI (3.6–15.5) (O3)
	GOSAT (0.75–15) + TES (8.7–10.5) (O3)	GEMS (East Asia) (0.3–0.5) + CrIS (3.9–15.4) (O3)
	CrIS (3.9–15.4) + GOSAT-2 (0.3–14.3) (CO, CH4)	TEMPO (N. America) (0.29–0.74) + IASI (3.6–15.5) (O3)
	CrIS (3.9–15.4) + TROPOMI (2.3–2.4) (CO, O3, CH4)	TEMPO (N. America) (0.29–0.74) + CrIS (3.9–15.4) (O3)

6. Lines 155-156: "DART assimilates observations and produce the analysis, an ensemble of optimized initial conditions"; then in lines 164-166, it seems the emissions are also updated. What are the actual objectives to optimize? While references are provided, it would be better if the authors could clarify the assimilation methodology more clearly to facilitate the readers.

Response: We have updated the following statement:

"To assimilate meteorology and chemical observational data, an ensemble of 30 CAM-chem simulations with different initial conditions and emissions to generate the forecast ensemble at a given time. DART assimilates observations and produce the analysis, an ensemble of optimized initial conditions (see details in Gaubert et al., 2016)."
to

"Here, we use the Ensemble Adjustment Kalman Filter approach (EAKF; Anderson, 2001, 2003). The forecast ensemble is generated by 30 CAM-chem simulations with different initial conditions and emissions. The assimilation is performed using DART and produces an ensemble of optimized initial conditions and emissions, as described in Gaubert et al. (2023). Specifically the state vector includes CO initial conditions, and CO emission fluxes that are ascribed to fires and anthropogenic sources."

7. Line 210: typo error: “run” at the beginning of this line.

Response: This is not a typo. “The experiment runs start on July 16th 2018 and are initialized with the spin-up/control run.”

8. Line 213: it could be better to explain the cause of the higher fraction at the beginning of the assimilation here.

Response: We added the following statement to Section 2.4:

“Systematic errors are larger at the beginning of the spinup, explaining the higher rejection rate. As the assimilation proceeds and the forecast bias is reduced, the rejection rate goes down.”

9. The repeated usage of full experiment names, e.g., “Experiment (2) Profile JNT assimilation and Experiment (5) Profile TIR and column NIR assimilation” made the discussion hard to read. The authors may consider abbreviated names.

Response: Thank you for the suggestion. We have added the following abbreviated names and updated the manuscript accordingly. Please see the revised manuscript for details.

(1) Column JNT assimilation (Exp1-CJ);
(2) Profile JNT assimilation (Exp2-PJ);
(3) Column TIR assimilation (Exp3-CT);
(4) Column TIR and Column NIR assimilation (Exp4-CT+CN);
(5) Profile TIR and Column NIR assimilation (Exp5-PT+CN).

10. The Conclusion section needs better organization. It is not good to list 9 main findings in the Conclusion.

Response: We agree with the reviewer and rephrased the conclusions accordingly. Specifically, we removed a few conclusions that are less important and merged the rest to 5 main conclusions.