General comments

1. The study conducted by Tang et al. evaluated the impact of assimilating multispectral/joint versus TIR-only retrieval, column versus profile retrieval, and joint products versus single-spectral products separately, through a 15-day inversion in CAM-chem+DART system, based on MOPITT CO database. They evaluated the assimilation performance against both the assimilated data and independent data (TROPOMI, TCCON, NOAA CCGG, IAGOS, and WE-CAN). Overall, this study is convincing and of great interest to the community studying CO or satellite retrieval, within the scope of AMT.

Response: Thank you for spending time reviewing our work.

2. Readers would be benefit from understanding if such comparison remains consistent across other seasons and years. In section 7.4, the authors acknowledge that these results might not be representative for other seasons, and performing a long-term assimilation has high computational cost. Please consider a direct comparison of measurements, rather than their assimilation performance, to evaluate the potential of robustness in discrepancies and similarities across datasets over varying timeframes.

Response: This study focused on using assimilation to compare with other observations. We refer to previous direct comparisons of measurements, such as MOPITT and TCCON in Hedelius et al. (2019), and we have added a reference for MOPITT and TROPOMI comparisons, that we neglected previously, to Martínez-Alonso et al., (2020), along with the text:

> "TROPOMI CO data were compared to MOPITT CO in Martínez-Alonso et al., (2020). TROPOMI and MOPITT data show good agreement in terms of temporal and spatial patterns with global average biases <4% between all MOPITT CO column products (TIR, NIR and JNT) and TROPOMI. TROPOMI CO values were slightly lower than MOPITT in most regional comparisons."

Note that direct measurement comparisons of remotely sensed satellite observations with IAGOS in situ data are problematic since IAGOS data are at discrete pressures along a flight track without the vertical profile information needed for smoothing with a satellite instrument operator. However, comparisons at the specific IAGOS pressure levels CO distributions after data assimilation are straightforward.

Unfortunately, the number of experiments prohibits longer simulations, and as the reviewer noticed, we discussed this limitation in the text. Nevertheless, a previous study from our group performed longer simulations for one experiment that assimilated the MOPITT profile product for a whole year (Gaubert et al., 2016) and found that there is no significant seasonal change in the

performance of the CAM-chem+DART. We have added this statement to Section 7.4. We also added the following statement to Section 7.4:

"If observations of roughly the same quality/quantity are available in other years, the performance of the DA might be expected to be similar."

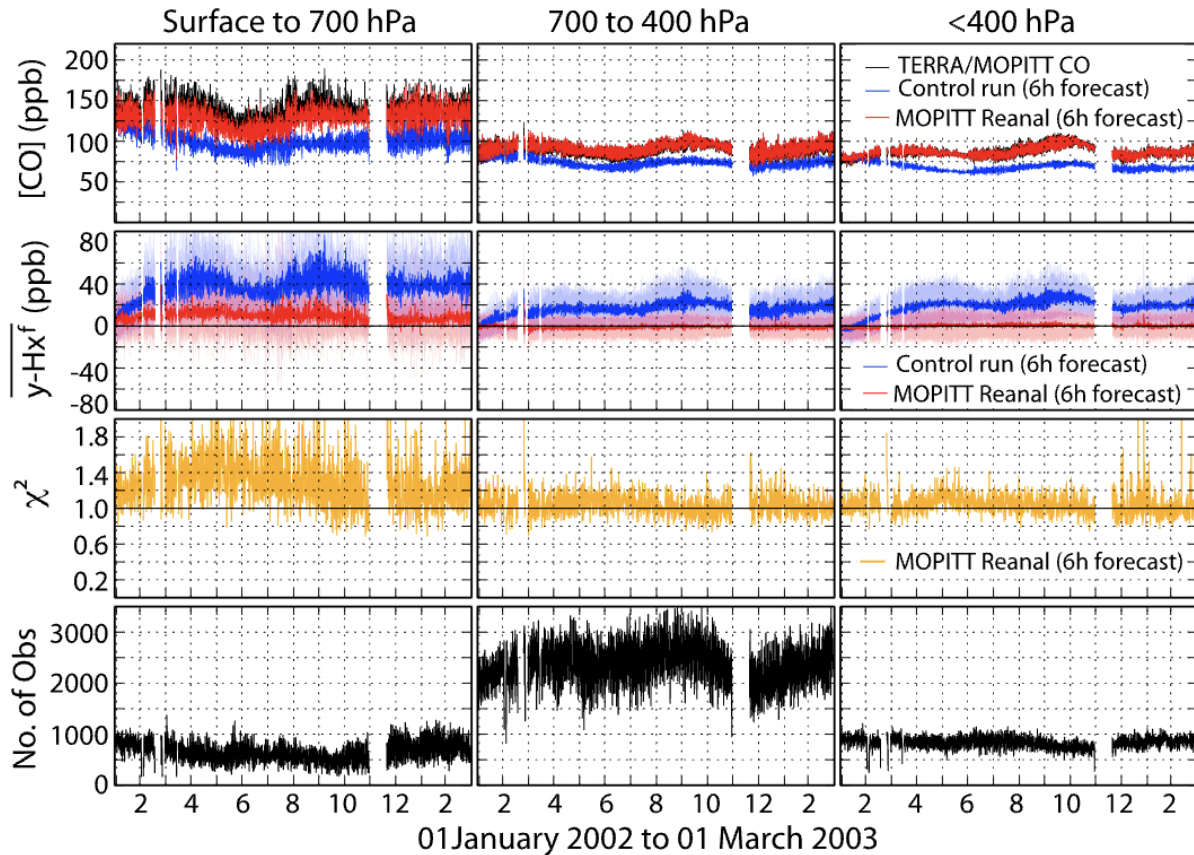Below is Figure 3 of Gaubert et al. (2016).

Figure 3. Time series of global assimilation diagnostics in observation space. For every assimilation step, the observations (and statistics) are averaged across the globe (horizontal) and over vertical sections: (left column) lower troposphere (from surface to 700 hPa), (middle column) middle troposphere (from 700 to 400 hPa), and (right column) upper troposphere (lower pressure than 400 hPa). (first row) The CO concentrations in ppb for MOPITT in black, 6-hourly forecast from the Control Run in blue and from MOPITT Reanalysis run in red. (second row) The innovations for both runs. (third row) The chi-square statistics of the forecasts from the MOPITT Reanalysis run. (fourth row) The number of observations (number of observation per 6 h) successfully assimilated in CAM-Chem and for which statistics are calculated.

3. Please clarify the treatment of observation error estimates. Did the inversions include the information of observation uncertainty? Can the observation uncertainty provided by MOPITT associated with measurements, effectively highlight the useful information from each dataset?

Response: The assimilations in this study include the information of observation uncertainty. The observation uncertainty provided by MOPITT is associated with measurements, and can effectively highlight the useful information from each dataset. Data assimilation requires observation errors associated with the quantity assimilated. MOPITT provides 3 types of uncertainty estimates: total error, measurement error, and smoothing error. Total error includes both measurement error and smoothing error. Since our observation operators include the smoothing by the MOPITT averaging kernels and the prior profiles, we only use the measurement error rather than total error provided by MOPITT for both column and profile products as smoothing error is already addressed by observation operators in the system. Specifically, for MOPITT profile products, measurement error is provided by the variable "MeasurementErrorCovarianceMatrix" while for MOPITT column products, measurement error is provided by the variable second column of the "RetrievedCOTotalColumnDiagnosticsDay". In addition, within the Data Assimilation Research Testbed, we use adaptive covariance inflation to adjust the total error (model and observations) using the given observation error as reference. Satellite observations are inversions themselves and are quite complex, data assimilation provides an accurate way to assess this information. We have added this information to Section 2.1:

"Data assimilation requires observation errors associated with the quantity assimilated. MOPITT provides 3 types of uncertainties/errors: total error, measurement error, and smoothing error in the products. Total error includes both measurement error and smoothing error. Since our observation operators include the smoothing by the MOPITT averaging kernels and the prior profiles, we only use the measurement error rather than total error provided by MOPITT for both column and profile products as smoothing error is already addressed by observation operators in the system. Specifically, for MOPITT profile products, measurement error is provided by the variable "MeasurementErrorCovarianceMatrix" while for MOPITT column products, measurement error is provided by the variable second column of the "RetrievedCOTotalColumnDiagnosticsDay". In addition, within DART, we use adaptive covariance inflation to adjust the total error (model and observations) using the given observation error as reference."

Specific comments

4. Figure 3: Please explain the criteria used to reject observations too far from the ensemble mean. Providing details on the threshold or methodology employed for this rejection would enhance the transparency of the assimilation process.

Response: We have included the following information in Section 2.4:

"Quality Checks are common in data assimilation as the algorithms are employed operationally for near real time forecasting. We use the standard option in DART to do such quality checks. The absolute value of the difference between the observed value and the prior ensemble mean estimate is divided by the expected value of this difference. That expected value is the square root of the sum of the specified observation error variance and the prior ensemble variance. If this ratio is greater than a threshold, the observation is not used. The threshold ratio used here is three which is commonly used for large tropospheric applications in DART."

5. Figure 5: The authors did a good job in effectively presenting the inversion performance in Figure 5, but did not discuss the 200 hPa results in the main text.

Response: Thank you for pointing this out! We added the following statement to Section 4.2 where Figure 5 (now Figure 6) is discussed:

"At 200 hPa, the spatial distribution of CO difference caused by assimilation is small in Experiment (2) Profile JNT assimilation, followed by Experiment (5) Profile TIR and column NIR assimilation. On the contrary, for the other three experiments which do not involve profile assimilations, the spatial distribution of CO difference caused by assimilation is relatively large. I.e., assimilating MOPITT profile product(s) only slightly change 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 DA cannot represent."

6. Figure 9: The two assimilations with profile observations exhibit >1 line fit. Does this indicate assimilations with profile measurements tend to overestimate emissions or surface concentration?

Response: Even though the least square regression lines of the experiments involving profile assimilation seem to be above the 1:1 1ine while the least square regression lines of the experiments only involving column assimilation seem to be below the 1:1 line, the mean biases are positive for all the five experiments. This indicates that all five experiments, regardless of profile assimilation or column assimilation, tend to overestimate surface CO compared to CO observations from the NOAA Carbon Cycle Greenhouse Gases (CCGG) sites during July 31st, 2018 to August 14th. To avoid any confusion, we removed the least square regression lines from the figure. Nevertheless, as discussed in Section 6.1, the reviewer is right that the experiments involving profile assimilations tend to have higher values in emissions (Figure 13) and surface concentrations (Figure S15).

7. Line 175: Please clarify the approach to project MOPITT CO AK and prior to the model resolution.

Response: Instead of "In this case, the forward operators apply MOPITT averaging kernel and prior information to model CO field before comparing it to MOPITT products. The capability of assimilating MOPITT profile products is described in Barré et al., (2015)."

We rephrase to "We use the forward operators introduced in Barré et al., (2015), consisting of i) estimating the log of a pressure weighted partial column volume mixing ratio that corresponds to the MOPITT grid and ii) applying the MOPITT averaging kernel and prior information."

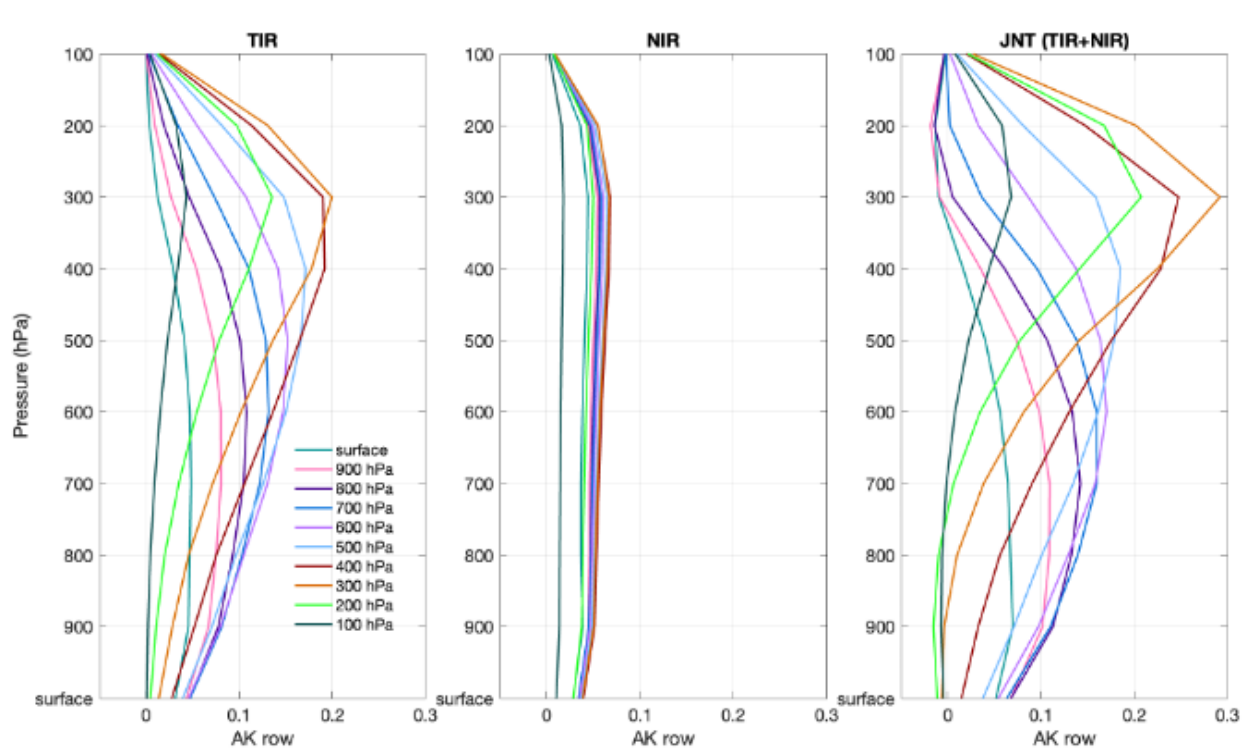We also added the new Figure 1 to show MOPITT CO AK.

Figure 1. Averaging kernel (AK) rows for MOPITT retrieval types TIR only, NIR only, and multispectral TIR+NIR. Global average of AKs during July and August 2018 are shown.

8. Line 195 and 196: the order of case 1 and 2 are opposite in the text and Figure 2.

Response: Thank you. We have updated Figure 2 (now Figure 3).

9. Line 211: Did the authors consider a spin-down timeframe? For observations after the assimilation timeframe while can still reflect emission signals?

Response: Regarding initial conditions, the improvement will theoretically improve the CO distribution for as long as the CO lifetime, which varies in space at the global scale (~1 or 2 months). Regarding emissions, Gaubert et al. (2023) showed that using posterior emissions can improve CO for years-long simulations by mitigating systematic errors in emission fluxes.

10. Line 320: Please clarify the definition of observation error variance, and if such error estimates have been incorporated into the assimilation system.

Response: The observation error variance has been incorporated into the assimilation system. We are not aware of a data assimilation system that would not use observation error variance. We use the measurement error rather than total error provided by MOPITT for both column and profile products as smoothing error is already addressed by observation operators in the system. Specifically, for MOPITT profile products, measurement error is provided by the variable "MeasurementErrorCovarianceMatrix" while for MOPITT column products, measurement error is provided by the variable second column of the "RetrievedCOTotalColumnDiagnosticsDay". This information has been included in the revised manuscript; also see response to comment 3.

11. Line 331: Please clarify the meaning of x in eq 4.

Response: "x" on the right of the equation is the modeled state. χ on the left of the equation is a greek letter chi, the Chi-squared test has been introduced in statistics during the 19th century. In the context of data assimilation, it has been introduced in Menard and Chang (2000; https://journals.ametsoc.org/view/journals/mwre/128/8/1520-0493_2000_128_2672_aoscto_2.0.co_2.xml).

12. Line 351: Experiment 2 and 5 are expected to assimilate similar information, with the major difference at 200 hPa, as indicated in Figure 5.

Response: As explained in Section 2.4 (line 207), the comparisons of experiment (2) and (5) will show the impacts of assimilating joint products (TIR+NIR) versus assimilating them separately for profile products. The two different approaches will impact the vertical distribution. In the profile DA experiment, there is an attempt to provide more freedom to the vertical distribution, but it is localized to lower levels. The column assimilation impacts all vertical layers and does not

include the vertical localization. Therefore column assimilation modifies the upper troposphere and the stratosphere more.

Response: Thank you for noticing. The numbers in the Figure are correct. We have updated the numbers in the text.

Response: Note that the weather spin-up is 10 days. We then update the emissions for a total of 35 days (July 11 – August 15), the unit of Tg/year is only a unit conversion. We discuss the update of annual emissions in Gaubert et al. (2023).

15. Section 7.4 How about temporal resolution? Can MOPITT optimize emissions up to daily scale based on the experiments here?

Response: The prior fire emission is at daily temporal resolution. As in previous studies (Gaubert et al., 2020, 2023), the daily flux is updated, but the increment is also applied to future days, with an exponentially decreasing weight and set to zero at 16 days.

Technical corrections

16. Table 1: Missing the labels of TEMPO spectral ranges and potential chemical species for geostationary satellites.

Response: Thank you. We have added missing information. 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)	AIRS (3.75–15.4) + OMI (0.27–0.5)	GIIRS (East Asia) (0.55–14.2) + TROPOMI (2.3–2.4)
(CO)	(O3)	(CO, O3)
IASI (3.6–15.5) + GOME2 (0.24–0.79)	TES (8.7–10.5) + OMI (0.27–0.5)	GEMS (East Asia) (0.3–0.5) + IASI (3.6–15.5)
(O3)	(O3)	(O3)
	GOSAT (0.75–15) + TES (8.7–10.5)	GEMS (East Asia) (0.3–0.5) + CrIS (3.9–15.4)
	(O3)	(O3)

CrIS (3.9–15.4) + GOSAT-2 (0.3–14.3)	TEMPO (N. America) (0.29–0.74) + IASI (3.6–15.5)
(CO, CH4)	(O3)
CrIS (3.9–15.4) + TROPOMI (2.3–2.4)	TEMPO (N. America) (0.29–0.74) + CrIS (3.9–15.4)
(CO, O3, CH4)	(O3)

17. Line 130: typo, an extra ";" after "CAM-chem".

Response: Thank you. We have corrected it.