

Response to Referee #1 – “A comparison of carbon monoxide retrievals between the MOPITT satellite and Canadian High-Arctic ground-based NDACC and TCCON FTIR measurements” by Ali Jalali et al.

We would like to thank Referee #1 for their helpful comments. Here we address their review, with their comments in green and our responses indented in black.

General comments:

The manuscript compares MOPITT CO v8 retrievals from 2006-2019 with TCCON and NDACC FTIR observations from the PEARL Eureka site. MOPITT observations within 110km and for 24h intervals are compared. The authors thoroughly describe the methodologies for comparing space-based total column retrievals with those ground-based networks. Beside the pixel-to-pixel bias the work examines and compares MOPITT NIR, TIR and NIR-TIR retrievals with the ground-based site and provides comprehensive of statistics. The results and improvements of v8 are thoroughly discussed and described.

- consider to mention that the fundamental and first overtone of CO are relevant in the TIR/NIR spectral region

We have added the following comment to Sec. 3 (FTIR Instrument) on p12,10. “The FTIR measures CO absorption spectra from either the fundamental band (NDACC) or first overtone band (TCCON).”

- do I get this right: each channel has four pixels, and each pixel in a channel sounds a certain part of a single CO line? if so does that mean that the inner pixel record transmittance from the line and the outer pixel from the line wing? why not sensing multiple CO lines?

The four MOPITT pixels provide spatial information not spectral information. They are arranged linearly in the orbit “along-track” direction. As the instrument is a gas correlation radiometer, the spectrum is not measured directly.

Specific comments:

Page3, line 2:

Consider to include the MAPS mission aboard the space shuttle (e.g. Reichle 1999)

We have made this update to p3, l2:

From: “Over the past two and a half decades, CO has been measured from space using a suite of nadir sounders. One of the earliest satellite-based instruments that measured CO was the Interferometric Monitor for Greenhouse Gases (IMG) (1996) (Wang et al., 1998) which collected eight months of data in 1996-1997.”

To: “Over the past four decades, CO has been measured from space using a suite of nadir sounders. The earliest instruments that measured CO were Measurements of Air Pollution from Satellites (MAPS) which flew on the Space Shuttle in 1981, 1984 and 1994 (Reichle et al., 1999) and the Interferometric Monitor for Greenhouse Gases (IMG) (Wang et al., 1998) which collected eight months of data in 1996-1997.”

Reference:

Reichle, H. G., et al.: Space shuttle based global CO measurements during April and October 1994, MAPS instrument, data reduction, and data validation, J. Geophys. Res., 104, 21 443–21 454, doi:10.1029/97JD03299, <https://doi.org/10.1029/97JD03299>, 1999.

p3,18:

Does NDACC only measure absorption of solar radiation or does the network also measure emission from the atmosphere?

The NDACC FTIR instruments only measure solar absorption spectra.

p4,21: Maybe add one sentence why $\log(\text{vmr})$ is fitted?

This following sentence has been added on p4,22. “Compared to retrievals of VMR, the $\log(\text{VMR})$ -based retrieval algorithm improves retrieval convergence and yields fewer profiles with unphysically small VMR values (Deeter et al., 2007).”

Reference:

Deeter, M. N., Edwards, D. P., and Gille, J. C. (2007), Retrievals of carbon monoxide profiles from MOPITT observations using lognormal a priori statistics, *J. Geophys. Res.*, 112, D11311, doi:10.1029/2006JD007999.

p4,23: One or two words explaining the ,outer' and ,inner' pixels meaning (line centre vs line wing?). Also see general comment above.

As mentioned above, the MOPITT pixels provide spatial information. This has been clarified in the text as follows.

From: “Each channel’s detector is comprised of a four-pixel linear array, where 1 and 4 are the outer pixels and 2 and 3 are the inner pixels of the array.”

To: “Each channel’s detector is comprised of a four-pixel linear array oriented along-track, where 1 and 4 are the outer pixels and 2 and 3 are the inner pixels of the array.”

p5,33 and Fig.1:

Just for confirmation, so the weighted average of all pixels means that each pixel’s average within the 30day/100km was multiplied by 0.25 and then summed?

No, all CO measurements from all pixels with their corresponding uncertainties (within 110 km radius and 30 days) are used to calculate the weighted average. Equations 4.20 to 4.23 from Bevington 1969 are used to calculate the weighted average.

Bevington, P. R., *Data Reduction and Error Analysis for the Physical Sciences*, 336 pp., McGraw-Hill, 1969.

p5, Fig1:

Is a joint TIR-NIR retrieval possible over water? does the signal then basically only come from the TIR channel?

Yes, as mentioned on p4,30 MOPITT NIR retrieval is only made over land. Therefore, for pixels over water, the joint TIR-NIR retrieval information is coming from TIR only.

p9,4:

Maybe add one sentence on how well the assumption for the small-area approximation is fulfilled at the Eureka site (to justify the chosen 1° radius criteria).

The 1° is consistent with the ~ 100 km x 100 km region assumed as sufficiently small by Hedelius2019 in their work. This has been clarified as follows on p9,4:

From: “...based on the assumption that over a small enough area (1° radius)...”

To: “...based on the assumption that over a small enough area (~ 100 km x 100 km or 1° radius, used here)...”

p9,16 and Fig. 4:

It is mentioned (in p13,30) that DOFS represent the information content of the retrieval, so why do high DOFS lead to large ΔX_{CO} ?

Does it mean that the priors are already close to the true values?

There are a few factors that could be influencing to the larger ΔX_{CO} for higher DOFS. First, there are a lower number of retrievals with higher DOFS and these make up are less than 5 percent of total retrievals. Second, the larger ΔX_{CO} is due mainly to pixel 1 (which has the largest systematic bias).

It may be counterintuitive, but it is not unexpected that increasing DOFS will be associated with larger retrieval biases. As DOFS decreases, retrieved profiles are more heavily weighted by the a priori. So, as DOFS decreases, retrievals for all four pixels will tend toward the a priori, and any inter-pixel biases will go to 0. DOFS is a useful index for the information content (or weighting of the measured radiances in the retrieval), not 'retrieval quality'. Information content and retrieval bias are really separate concepts.

Are low DOFS retrievals trustworthy fits?

Please clarify the meaning of DOFS for the present work.

The MOPITT v8 user guide recommended not to filter any data due to characteristics of the retrieval averaging kernels like DOFS. As mentioned in the paper, we did not use DOFS as a filter criterion.

DOFS is a retrieval parameter output provided for each of the datasets and is defined on p13,30-31.

p10, Fig. 4:

Why are the SZA values ranging from 60-120°? I would expect it from 60-90°.

The MOPITT TIR measurements do not require sunlight so the instrument provides measurements in these channels up to SZA of 120° (for the data set we examined within 110 km of PEARL).

Also check p16,5 which designates daytime measurements (SZA < 90°).

This is correct. To be consistent with the previous study by Buchholz2017, we only used daytime measurements when identifying coincident measurements for the comparison.

Note that the figure caption says the RMS is represented in blue when it actually is yellow.

This has been fixed in the paper. The RMS symbols are green (olive).

p12,20:

Is there a reason why NDACC is not using 4 times daily (6 hourly) NCEP data?

NDACC uses the daily profiles at 1200 GMT that NCEP generates for all NDACC sites as described in the metadata file located here: https://www-air.larc.nasa.gov/pub/NDACC/PUBLIC/meta/ncep/ncep_2022.pdf.

This has been clarified in the text as follows:

“The National Centers for Environmental Protection (NCEP) provides daily temperature and pressure profiles at 1200 GMT interpolated to the geographical location of NDACC stations; those for Eureka are used in the retrieval (<https://www-air.larc.nasa.gov/missions/ndacc/data.html?NCEP=ncep-list>).”

In addition, we have changed the original sentence on p12,19.

From: “The mean outputs from Whole Atmosphere Chemistry Climate Model (WACCM) version 4 between 1980-2020 are used for the a priori VMR profiles (Marsh et al., 2013) and daily temperature and pressure profiles from the National Centers for Environmental Protection (NCEP) interpolated to

the geographical location of PEARL are used in the retrieval (<ftp://ftp.cpc.ncep.noaa.gov/ndacc/ncep/>).”

To: “The mean outputs from Whole Atmosphere Chemistry Climate Model (WACCM) version 4 between 1980-2020 are used for the a priori VMR profiles (Marsh et al., 2013). The National Centers for Environmental Protection (NCEP) provides daily temperature and pressure profiles at 1200 GMT interpolated to the geographical location of NDACC stations; those for Eureka are used in the retrieval (<https://www-air.larc.nasa.gov/missions/ndacc/data.html?NCEP=ncep-list>).”

p13, Sec.4:

Consider to include a sentence which states that the total column averaging kernel is computed by a total column operator (converts profile concentrations to a column concentration, see Deeter 2002).

This sentence has been added to the paper on p13,14. “The CO total column averaging kernel is calculated from the profile averaging kernel matrix and total column operator (Deeter, 2002).”

Deeter, M. (2002): Calculation and Application of MOPITT Averaging Kernels, https://www.acom.ucar.edu/mopitt/avg_krnls_app.pdf, last accessed 17 Sept. 2022.

p18,3:

what is the motivation for 500 grid points per MOPITT layer? just to be dense enough?

Yes, this was increased to 500 points to ensure that there were sufficient grid points per MOPITT layer for the interpolation.

p18,13:

What is the motivation for selecting a critical difference of 80hPa? is it proposed somewhere (in Kerzenmacher et al. 2012?)

When Buchholz2017 applied the method from Kerzenmacher et al. (2012), they used a critical distance of 20 hPa for most locations but increased this value to 50 hPa for high latitude and altitude stations such as PEARL. We have used a slightly higher value to increase the number of profiles included in the comparison dataset.

p18,27:

Consider to explain why MOPITT retrievals are in log space (see previous comment p4,21)

This information was added on p4,21. See above.

p19,,11:

If TCCON method IV from Hedelius is similar to the method used for NDACC (more consistent?), why not using it instead of method II?

We chose to use methods consistent and similar to previous studies for each network in our study.

p20, Eq.(10):

Why is the number of measurements N the same for the MOPITT and FTIR (because of the averaged MOPITT measurements?)

Yes, this is correct.

p22,5:

is the larger correlation to NDACC caused by the TIR interval? is the Fig. 9 plot also available for MOPITT NIR?

Figure 9 presents an example pair of correlation plots for the NDACC and TCCON comparisons. The plots for all pixel and channel combinations are given in the supplementary document. The range of correlation values for NDACC and TCCON with the MOPITT TIR channel can be seen in

panel (b) of Figures 11 and 12, respectively. The average for all pixels is fairly consistent between NDACC and TCCON with a greater spread in values for NDACC comparisons.

p29, Sec. 6.4:

likely different co-location and filter criteria across studies are responsible for some of the difference (as you mention three times more V8 comparisons than V6 in Buchholz). Consider mentioning the effect of coincide criteria w.r.t. Hedelius 2019 in paragraph p29,12.

This sentence has been added to p30,1. “This could also be influenced by the looser coincidence criteria used by Hedelius2019 for high latitude stations ($4^\circ \times 8^\circ$ versus 1° radius used here).”

Technical corrections:

p31, Table 3:

Is the entry for the `V7 NOAA all sites` correlation coefficient a typo?

No, in Table 3 of Deeter et al. 2019, under V7N the total column correlation is 0.04.

Write text in equations upright in order to discriminate it from variables.

This has been done.