In this manuscript, the authors report on an improved version of the O2-O2 cloud retrieval for OMI and TROPOMI, aiming at consistent trace gas retrievals for the two instruments. The main improvements of the algorithm are the use of consistent a priori data, a larger O4 fitting window, treatment of the temperature dependence of the O4 cross-sections, destriping for OMI and an offset correction of the TROPOMI O4 columns. With these updates, the cloud retrieval results in more consistent cloud parameters between the two instruments than with the current implementations of the O2-O2 retrieval, in particular with respect to cloud pressure. Application to tropospheric NO2 columns also shows good consistency between the two instruments.

We sincerely thank Referee #3 for reviewing our manuscript and offering valuable, constructive comments. Our detailed responses to the specific comments are provided below.

The manuscript reports on an interesting study, which fits well into the scope of AMT. It is well written, has a good level of detail and is of relevance for the UV/vis satellite community. I, therefore, recommend it for publication after revisions.

General comments

Although I overall liked the paper and enjoyed reading it, there are several important aspects which in my opinion need to be addressed before publication.

1) Many aspects of the O4 retrieval have been changed, and it is not entirely clear, which one is the driver of the better agreement between the two instruments that the authors find. The use of the larger fitting window is not really motivated in the manuscript, the temperature correction is physical but was not used on either of the instruments in the past, so the main differences are the use of consistent a prioir information and the (not very satisfying) ad hoc offset correction. I suggest that the effect of the different changes is tested separately on a subset of data to clarify this point.

Thank you for your suggestion.

In the revised manuscript, we have added a table summarizing the impact of the main improvements in the BIRA-IASB cloud pressure retrieval compared to the OMCLDO2 approach.

We have added the following paragraph in section 4.1.1:

“Table 2 summarizes the impact of each improvement in our cloud pressure retrieval relative to OMCLDO2, as well as the overall differences between the two retrieval approaches, based on one day of OMI and TROPOMI data. Effects on cloud fraction are not included, as the implemented changes have minimal influence in this regard. The impact is generally much greater in scenes with low cloud fraction than in cloudy conditions, and it tends to be more pronounced over land than over ocean. Although this analysis does not account for the wavelength dependence introduced by using different fitting windows in the O2-O2 SCD retrieval, the impact of this effect is relatively minor.” (1-2% in the AMF calculated at 465 and 475nm)

Table 2: Impact of the key improvements in the BIRA-IASB cloud pressure retrieval compared to the OMCLDO2 approach, and the overall differences between the two retrievals. For each improvement, the BIRA-IASB retrieval is applied, and the cloud pressure difference resulting from the individual change is calculated. The analysis is based on one day of OMI (1 October 2004) and TROPOMI (1 October 2018) measurements, restricted to snow/ice-free pixels within the 50°S to 50°N latitude range.

	Cloud fraction[a]>0.2	0.05<Cloud fraction<0.2, Ocean	0.05<Cloud fraction<0.2, Land
O2-O2 SCD fitting[b]	OMI: 28±35hPa TROPOMI: -17±34hPa	OMI: -5±85hPa TROPOMI: -78±84hPa	OMI: -62±103hPa TROPOMI: -188±119hPa
Correction for OMI across-track variability	0±15hPa	1±28hPa	1±38hPa
TROPOMI O2-O2 SCD Offset correction: -0.08×10^{43} $molec^2$ cm^{-5}	-17±6hPa	-31±13hPa	-44±19hPa
Updated temperature correction	-17±5hPa	-25±11hPa	-34±16hPa
Updated surface albedo[c]	5±11hPa	22±34hPa	31±74hPa
Overall difference for OMI	20±36hPa	70±105hPa	26±126hPa
Overall difference for TROPOMI	-32±28hPa	-50±68hPa	-147±98hPa

[a] cloud fraction is from the BIRA-IASB retrieval

[b] the wavelength dependence due to the use of different fitting windows between the BIRA-IASB and OMCLDO2 retrievals is not accounted for.

[c] comparison between use of OMI mode LER (Kleipool et al., 2008) and TROPOMI DLER (Tilstra et al., 2024)

2) A significant effort was taken to correct the row dependence of the OMI O4 columns, and the results are certainly an improvement. However, in the final products (Fig. 14), the row dependency of the cloud pressure is smoother but not really smaller than in the original OMI product. It also is different between OMI and TROPOMI. This is a concern as it points at systematic problems in the retrieval. Please discuss this result in more detail.

The average cloud pressure values in the figure contained an inconsistency, as the OMI BIRA-IASB values were calculated over all measurements, whereas the other cloud pressures were computed only for pixels with a cloud fraction greater than 0.05. We updated this figure using a consistent calculation method for all retrievals, and further separated the average cloud pressure values for both cloudy and low cloud fraction scenes, due to the notably different row dependence observed between these two scene types.

We have updated the paragraph in the revised manuscript:

"The mean cloud pressure values are generally higher over ocean compared to land. For cloudy scenes (cloud fraction > 0.2), OMI retrievals show a weak dependence on VZA for OMI, while TROPOMI retrieves slightly higher cloud pressures near nadir, with values increasing toward the edges of the swath. The mean cloud pressure values from two TROPOMI products are closely aligned on the west side of the swath, whereas on the east side of the swath, the BIRA-IASB cloud pressures are generally higher than those from OMCLDO2. However, no clear west-to-east trend is observed between the two OMI retrievals.

For scenes with low cloud fraction, cloud pressure retrievals exhibit significantly stronger across-track variation compared to cloudy scenes. In such cases, the OMI BIRA-IASB cloud pressure values are generally higher than those from OMCLDO2, while for TROPOMI, BIRA-IASB values tend to be lower. Most cloud pressure retrievals exhibit a consistent broadly across-track pattern, typically with lower values near nadir that decrease toward the edges of the swath. An exception is noted for the TROPOMI OMCLDO2 retrieval over land, where cloud pressures at nadir are slightly higher than those at the swath edges. Over ocean, the OMI and TROPOMI BIRA-IASB retrievals exhibit a consistent across-track behavior. Over land, however, discrepancies arise near the swath edges, likely caused by surface reflectance anisotropy effects resulting from slight differences in solar and viewing geometries between OMI and TROPOMI. Additionally, OMI OMCLDO2 cloud pressures exhibit pronounced across-track variability due to the lack of a de-striping correction, resulting in amplitudes exceeding 100 hPa in certain detector rows. The difference between the retrievals from 2004 and 2018 is minimal, indicating temporal stability in these features."

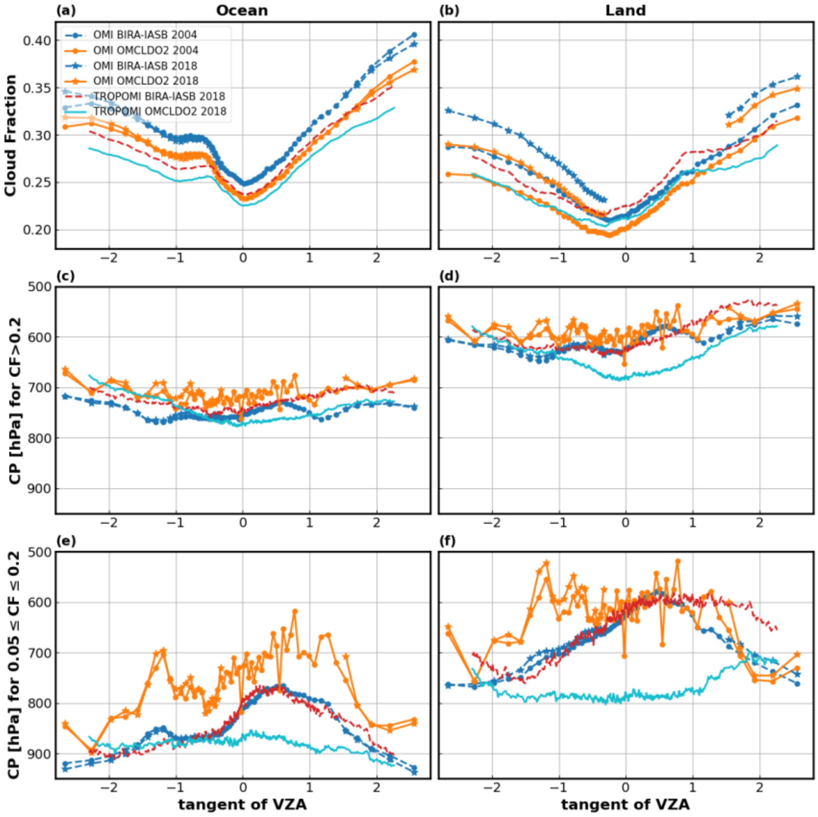

Figure 14: Across-track dependence of monthly mean cloud retrievals from OMI and TROPOMI for October, based on the measurements between 50°S and 50°N latitudes. The left panels present results over ocean, while the right panels show results over land. The analysis includes cloud fraction (top), cloud pressure in cloudy scenes (middle), and cloud pressure in low cloud fraction scenes (bottom). The legend details the sensors, retrieval methods, and time periods included in the analysis.

3) While the results of the improved algorithm are compared in detail with those of the existing O2-O2 algorithm, the real alternative to the new algorithm is the existing FRESCO cloud algorithm. In contrast to the O2-O2 data stored in TROPOMI NO2 files, the FRESCO results are used by the current TROPOMI NO2 and HCHO retrievals, and this is what the new algorithm should (also) be compared to! I think that all comparisons dealing with cloud fractions and cloud pressures need to include the currently used TROPOMI cloud algorithm.

Thank you for the comment. This analysis was conducted but is not included in the manuscript. We compared the two TROPOMI O2-O2 products not only with FRESCO but also with OCRA/ROCINN, the operational cloud product used for cloud correction in the TROPOMI HCHO and SO2 retrievals. While these comparisons are relevant, including them in detail would significantly lengthen the manuscript and potentially divert focus from our primary research objectives.

For the NO2 comparison presented in the manuscript, we included the FRESCO cloud correction, as it is used in the operational TROPOMI NO2 retrieval. From this perspective, we have included a brief discussion comparing our TROPOMI O2-O2 retrieval with the cloud product used in the operational NO2 retrieval.

In the revised manuscript, we have added a paragraph in section 4.2 and included two figures in appendix to compare the pixel-by-pixel difference of cloud fraction and cloud pressure:

"As shown in Fig. A4, the 440 nm cloud fractions agree well with the results from our O2-O2 retrieval. In contrast, the cloud fractions derived from the O2-O2 (477nm) and O2-A band measurements exhibit larger discrepancies, and these differences are particularly evident for low cloud fractions, with differences more pronounced over land. For cloudy scenes, there is generally good agreement in cloud pressure retrieval between our O2-O2 results and FRESCO (Figure A5). However, in scenes with low cloud fractions, the figures show substantial scatter. FRESCO tends to retrieve relatively lower cloud pressure for high cloud cases over ocean, and slightly higher over land."

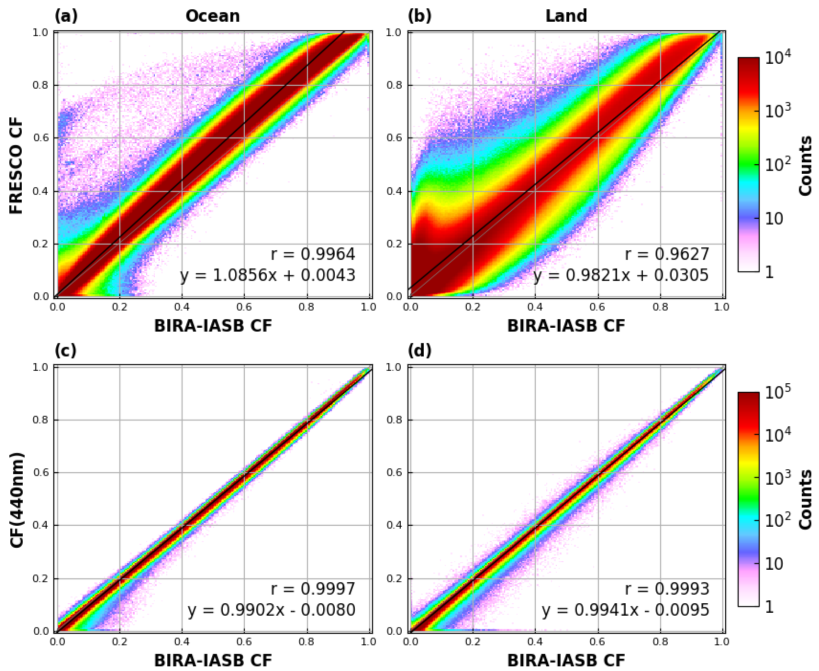

Figure A5: Similar to Figure 12, but showing a comparison between our TROPOMI O2–O2 cloud fraction retrievals and FRESCO (top panel), as well as the 440 nm cloud fraction (bottom panel) used in the operational TROPOMI NO2 retrieval.

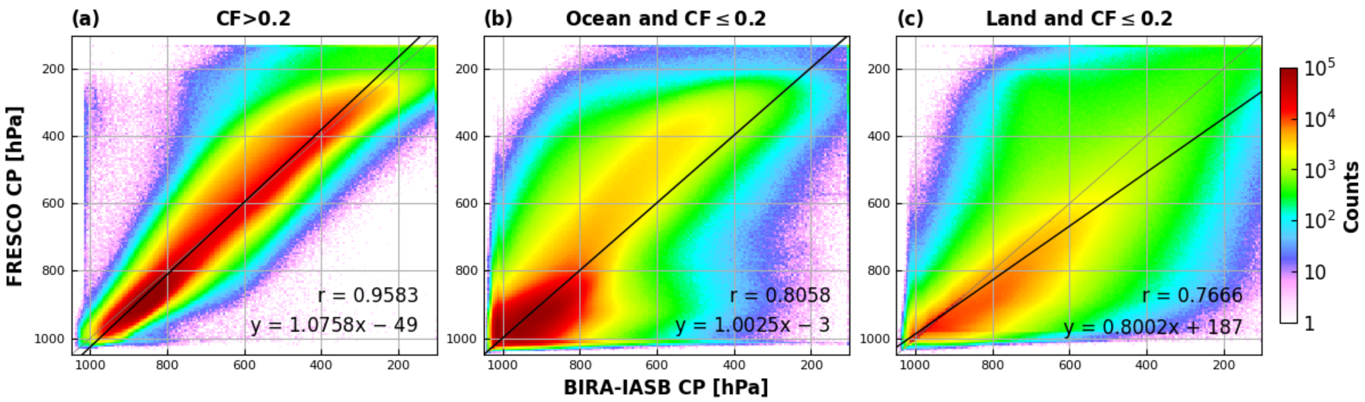

Figure A6: Similar as Figure 13, but for comparison of TROPOMI cloud pressure retrievals between BIRA-IASB O2-O2 and FRESCO.

4) I'm not comfortable with the discussion of the effect of using different cloud products on the NO2 tropospheric columns. On the one hand, the way the numbers are discussed in the text focuses on the largest differences, which are driven by the TROPOMI O2-O2 algorithm which has to my knowledge never been used for tropospheric NO2 retrievals. In that sense, the importance is exaggerated, and in my opinion, differences of a few percent are small considering all the other uncertainties in the retrievals. A more appropriate message would therefore be "cloud correction does not make a big difference for NO2 tropospheric retrievals under the assumptions made here".

On the other hand, the data are monthly averages, and for individual scenes, I expect much larger differences. Therefore, I'd suggest to add histograms of differences to the mean values for a more balanced view.

Thank you for the suggestion. Indeed, the effects of cloud correction on the NO2 retrieval is generally small over the polluted regions. It is challenging to compare pixel-by-pixel differences between OMI and TROPOMI due to their differing spatial sampling. Even when measurements are taken over similar locations, the effects of cloud correction can vary as a result of differences in solar and viewing geometries between the two sensors. Instead of a direct comparison, we have included maps in the revised appendix showing the maps of the average cloud correction effects for each approach, and results show that the cloud correction effect is generally within ±20%.

We have rephrased this paragraph in the revised manuscript:

"The impact of cloud corrections on the NO2 AMF is generally within ±20% (see Figure A5). All corrections exhibit a systematic positive bias over ocean, whereas over land, the effect is comparatively minor, except in tropical regions, where most cloud products introduce a negative bias. As shown in Fig.18(a), the average impact of cloud correction for the selected polluted regions ranges from -6% to 11%, depending on the cloud product applied. The AMF difference resulting from various cloud corrections can exceed 10%. However, these values remain well below the typical uncertainty associated with tropospheric NO2 AMFs.\citep{Boersma2004}."

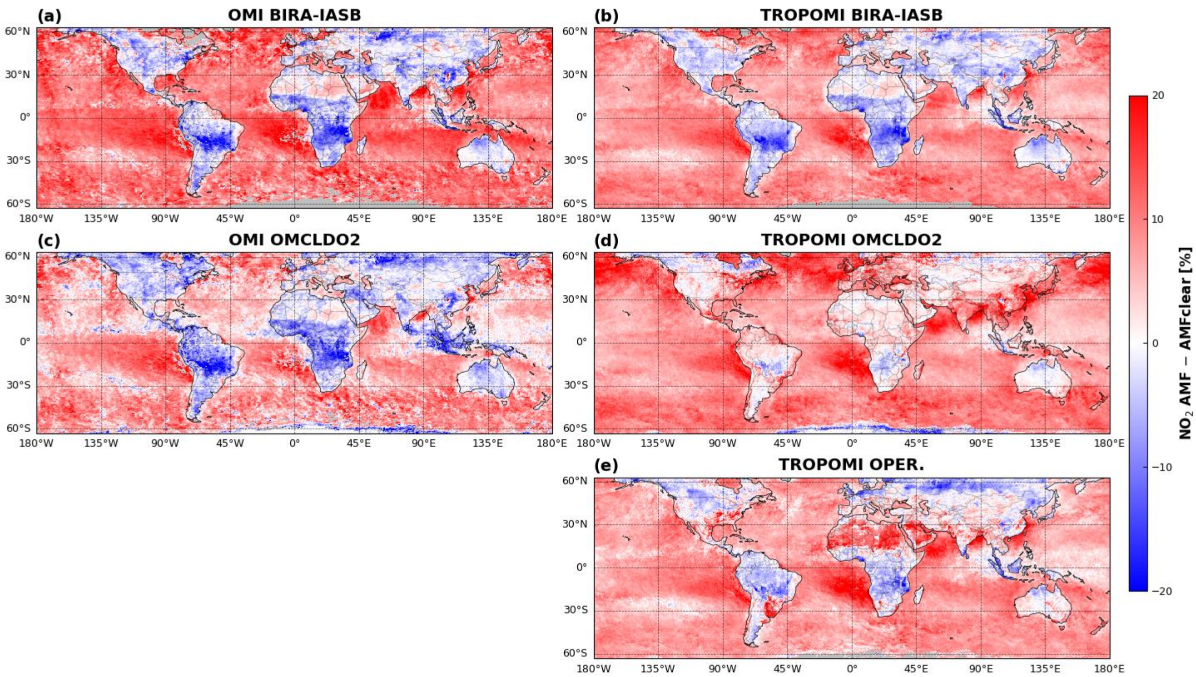

Figure A5: Monthly average impact of cloud correction on the tropospheric NO2 retrievals from OMI (left) and TROPOMI for October 2018. The NO2 AMF is calculated using identical retrieval settings, differing only in the cloud correction approach applied. Data are gridded at a spatial resolution of 0.5°, including only observations with a cloud radiance fraction below 0.5.

Finally, the results in Fig. 18b are quite surprising when comparing the cloud corrections TROPOMI-BIRA and operational – with the exception of Eastern China, they always have opposite sign! I think that warrants more discussion than just referring to “sampling effects”. How can we trust cloud corrections at all, if two different, state of the art approaches lead to opposite effects?

Fig. 18b shows the NO2 VCDs, each normalized by the mean value of the five retrievals in the same region. This normalization is applied due to the large variability in NO2 VCD values among different regions, and it helps to better highlight the relative differences caused by various cloud correction methods for each region.

A clear inverse correlation is observed between AMF corrections and NO2 VCD retrievals for the three TROPOMI products (since VCD = SCD/AMF), as well as for the two OMI retrievals. The correlation between OMI and TROPOMI data is slightly weaker, primarily due to differences in sampling.

We have rephrased the paragraph in the revised manuscript:

“The NO2 VCDs generally exhibit an inverse relationship with the effect of cloud corrections, particularly for the three TROPOMI retrievals. This correlation is somewhat weaker when comparing OMI and TROPOMI, which may be influenced by sampling differences between the two sensors.”

Detailed comments:

L28: … most are partially covered by clouds.

Done

L31: wasn't the IPA assumption already used in Martin et al., 2002?

Yes, we included it, as the IPA assumption was already used in Martin et al., 2002.

L37: Although most people will know OMI, I think a very short introduction of OMI and in the next paragraph TROPOMI is needed here

We have added the following paragraph for OMI:

"The Dutch-Finnish-built OMI, a key payload aboard the NASA Aura spacecraft, is a nadir-viewing, wide-swath, push-broom imaging spectrometer designed for daily global monitoring of tropospheric composition."

And also, for TROPOMI:

"TROPOMI, aboard the Sentinel-5P platform, is the first Copernicus mission dedicated to atmospheric monitoring, providing daily city-scale measurements for air quality assessment, ozone and UV radiation monitoring, and climate observation and forecasting."

L38: Not sure if a trace gas with an exponential mixing ratio profile should be called "well mixed". What about "with a known vertical distribution"?

Thanks for the suggestion and corrected it.

L40: "collision-induced absorption" sounds a bit strange as the absorption of the photon is not caused by the collision of the molecules, just the collisional complex has an absorption cross-section which has larger values at this wavelength.

Yes, we rephrased it as: The OMCLDO2 cloud product uses satellite measurements of the O_2–O_2 collision complex absorption feature centered at 477 nm.

L44: There is an imbalance in the discussion of OMI and TROPOMI here, as for OMI, cloud fraction retrieval is not explained

We expand upon the description of the OMI cloud retrievals as follows:

"Both algorithms use the independent pixel approximation, which characterizes pixel reflectance as a weighted combination of cloudy and clear-sky parts. This approach enables the determination of the effective cloud fraction rather than the geometric cloud fraction. Due to limited spectral information in the O_2–O_2 absorption band and the RRS process that are used for cloud pressure retrieval, the cloud is modeled as a Lambertian reflector with a fixed albedo of 0.8. Consequently, only the pressure level of this Lambertian cloud is retrieved."

L99: "consistently sized rows" – not clear what is meant with this

We rephrased this sentence:

"The swath is divided into 77 to 450 across-track rows, with the binning factor adjusted to ensure similar spatial size for each row."

L120: Please give full version number of lv1 product

Add: OMI processor version: 2.0.8.4/24861; TROPOMI: processor version: 2.1.0.25042

L140: …can implicitly correct… => … can implicitly correct part of the …

Corrected

Figure 1: 10E-46 on y-axis label

The O_2-O_2 absorption cross-section values are scaled by 10^{46}, as indicated by the y-axis title ($\times 10^{46}$ cm^5 $molecule^{-2}$).

Figure 2: Which cloud fraction is used for (c) and (f)?

BIRA-IASB cloud fraction. Added it.

L225: Were the data also filtered for sun glint?

Not in this case. We investigated this by analyzing data within a selected latitude range that excludes regions affected by sun glint and found the results to be consistent with those obtained using our current approach. In addition, including a larger dataset helps improve the signal-to-noise ratio.

L259: ... at the 465 … => … at 465 …

Corrected

L265. Add horizontal resolution of the model

Done (with horizontal resolution of 0.75°x0.75°)

L312: I do not understand, how this reduces interpolation errors, and also did not find any additional explanation in Wang et al, 2020. Can you please explain?

Linear interpolation is commonly used in the LUT-based approach. The O2-O2 SCDs exhibit a strong nonlinear dependence on VZA/SZA for high geometry angles, which can lead to relatively large interpolation errors. In contrast, the O2-O2 VCDs show only a weak dependence on geometry, resulting in smaller uncertainties due to interpolation.

L327: … proportionally to with the … => … proportionally with …

Corrected

L375: … slightly highly … => … slight high …

Corrected (slightly high)

L418: … the algorithm is not sensitive to the high … => … the algorithm is not sensitive under high …

Corrected

Figure 10: There are clear orbital structures in the western part of the TROPOMI image and in many places of the OMI images, both for cloud fraction and cloud pressure. Please discuss.

We have added:

"Over low-latitude ocean regions, distinct orbital structures are visible in the western part of the satellite swath, primarily due to sun-glint effects. These effects can lead the O2–O2 cloud retrieval to overestimate cloud fraction and potentially result in artificially low cloud pressure values."

L481: Isn't that just a geometric effect, that apparent cloud fraction increases over broken clouds under slant view if clouds are 3d-objects?

Yes, there is a geometric effect. The apparent cloud fraction tends to increase under slant viewing angles due to vertical extent of clouds. However, 3D cloud structures may introduce both negative (due to shadowing) and positive biases in the cloud fraction retrieval.

Based on synthetic analysis of a scene containing a 1D cloud layer (Yu et al., 2022), the retrieved cloud fraction increases with SZA/VZA, primarily due to increased scattering by clouds along the longer light path.

We rephrased this:

"The cloud fraction values increase towards the edges of the swath, likely due to factors such as geometric effects and enhanced cloud scattering along the slant path."

L489: Shouldn't we expect a bias in OMI cloud fractions as the pixels are larger than those of TROPOMI?

In principle, TROPOMI observations include more cloud-free pixels than OMI due to TROPOMI's higher spatial resolution. However, we do not expect a bias in the average cloud fraction between the two sensors, as the cloud fraction retrieval is based on the Independent Pixel Approximation (IPA) assumption, and the mean value does not depend on pixel size.

L524: I think it would be worthwhile to point out, why you used the 0.2 threshold – I assume that's because this corresponds to 50% cloud radiance fraction, and therefore is the limit used in the tropospheric NO2 discussion. This point is then relevant when discussing figures such as Figure 16a, where large discrepancies are found in exactly the range of values relevant for cloud correction in trace gas algorithms.

Yes, and we added a sentence in section 4.1.1:

"Tropospheric NO2 retrieval from satellite measurements typically excludes pixels affected by clouds. Such pixels are defined as having a cloud radiance fraction greater than 0.5, which corresponds to an effective cloud fraction of approximately 0.15–0.2."

L579: Why is weighting with cloud fractions a good idea? I think that this artificially makes results look better, as high cloud fractions show better agreement (Figure 13) and weighting will bias everything towards large cloud fractions.

From a technical perspective, the weighted average provides a more appropriate metric for cloud pressure comparison than the simple average. When the dataset are grouped by the average TROPOMI cloud fraction within each grid cell (< 0.2 and > 0.2), the low cloud fraction cases exhibit greater scatter. Nevertheless, the overall conclusions remain largely unchanged, as demonstrated below. Furthermore, Section 4.1.4 investigates the relationship between cloud pressure and cloud fraction, indicating that the consistency between the two BIRA-IASB O2-O2 cloud pressure retrievals nearly independent on cloud fraction.

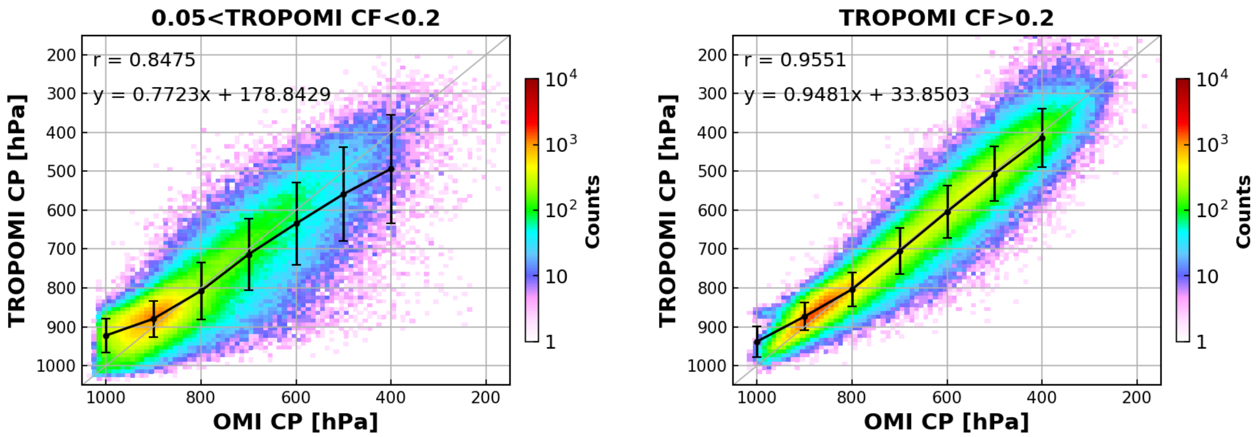

Figure: Similar to Figure 19(d), but showing data are grouped based on the average TROPOMI cloud fraction: < 0.2 (left panel) and > 0.2 (right panel).

Figure 17: remove "are considered"

Done

Figure 19: Figure headers "select data" => "selected data"

Done

Figure 20: axis labels "refletance" => "reflectance"

Done

Reference

Kleipool, Q. L., Dobber, M. R., de Haan, J. F., & Levelt, P. F. (2008). Earth surface reflectance climatology from 3 years of OMI data. *Journal of Geophysical Research Atmospheres*, *113*(18), 18308. https://doi.org/10.1029/2008JD010290;CTYPE:STRING:JOURNAL

Tilstra, L. G., De Graaf, M., Trees, V. J. H., Litvinov, P., Dubovik, O., & Stammes, P. (2024). A directional surface reflectance climatology determined from TROPOMI observations. *Atmospheric Measurement Techniques*, *17*(7), 2235–2256. https://doi.org/10.5194/AMT-17-2235-2024,

Yu, H., Emde, C., Kylling, A., Veihelmann, B., Mayer, B., Stebel, K., & Van Roozendael, M. (2022). Impact of 3D cloud structures on the atmospheric trace gas products from UV–Vis sounders – Part 2: Impact on NO_2 retrieval and mitigation strategies. *Atmospheric Measurement Techniques*, *15*(19), 5743–5768. https://doi.org/10.5194/amt-15-5743-2022