The manuscript “Tropospheric ozone column dataset from OMPS-LP/OMPS-NM limb-nadir matching” by A. Orfanoz-Cheuquelaf et al. presents a new research tropospheric ozone product derived from Suomi NPP OMPS satellite measurements. The paper fits to the scope of the AMT journal. The paper is well structured and written. This dataset will be of interest to the atmospheric community especially if it will be combined with a similar dataset from SCIAMACHY as promised in this publication. I recommend this paper for publication after a revision. My comments are summarized before.

General Comment:

It is not clear why the dataset is limited to 2012-2018 time period. Considered that Suomi NPP OMPS is still an active instrument it would be desirable to extend this dataset and provided analysis at least untill the end of 2022. That would extend the overlap with TROPOMI substantially and provide community with the up-to-date information about the status of tropospheric ozone layer.

A: The reason is explicitly mentioned in the revised manuscript:

The OMPS-LP ozone profile time series based on L1 V2.5 data, which are used by both V2.6 and V3.3 retrievals, were found to exhibit a significant positive drift after 2018 (Kramarova et al., 2018). For this reason, only the data until 2018 are used to create the OMPS-LNM-TrOC dataset. (Lines 139-141)

A new version of IUP OMPS-LP profiles is being processed based on the improved L1 (V2.6) data that counts for the observed drift after 2018. Using the improved stratospheric data, the OMPS-LNM TrOC dataset will be reprocessed, extended to the present and will be subject of a later paper. (Lines 466-468)

Specific comments:

p. 3 line 85. You need to rephrase this statement “The charge-coupled device performs instantaneous measurements of the entire atmosphere”. I assume you meant that radiances are collected simultaneously spectrally and spatially over the FOV.

A: The text is now corrected as follows:

The pixel columns of the charge-coupled device observe the atmosphere vertically in 1 km steps with a field of view of 1.5 km for each detector pixel. The pixel rows resister the spectral distribution of the radiance at each tangent height. (Lines 83-85)

p.4 lines 101: Why did you include only odd-numbered spectral points and how does that relate to the temperature dependence? I assume you meant the temperature dependence of O3 cross-sections, right?

A: The line is corrected as follows:

This selection reduces the influence of the temperature weighting function within the fit procedure and makes the fit more stable. (Lines 102-103)

Details on the justification of this selection and illustration of the results are provided by Orfanoz-Cheuquelaf et al. (2021).

p.4 line104: if your intention is to name the instrument first and then the algorithm type, then it should be “OMPS-NM TOMS” rather than “OMPS L2”.

A: It is now defined as “NASA's product OMPS-NM L2 V2.1” (Line 109)

Section 2.2. I was a bit confused with version numbers and references. First you refer Arosio et al., 2018, then you stated that in this study you use v3.3, but didn't provide a reference or any explanation of how that v3.3 differ from Arosio et al., 2018. It is not clear what had happened in between. Are there v2.7, 3.0 etc.? Please, clearly explain which version you use and provide the correct references. If the refernence doesn’t exist, then explain how this v3.3 differ from what is described in Arosio at al., 2018.

p. 5 line 124: why can’t you trust “OMPS-LP ozone profile time series based on level 1 V2.5 data” after 2018? Please provide reference or explanation. And again please clearly explain if v3.3 uses v2.5 Level 1 data or not.

P5., line 127: I didn’t find any mentioning of IUP-OMPS v2.6 in Arosio et al. 2018. Are you including the right reference?

P5., line 131: Why do you see the improvements in v3.3? How does v3.3 differ from v2.6?

A: The whole section was reorganized considering answers to these questions (Lines 129-143):

This study uses OMPS ozone profiles version 3.3 (IUP-OMPS V3.3). Comprehensive validation of the ozone profiles and details about the retrieval can be found in Arosio et al. (2018) and Arosio (2019) for the previous retrieval version (here named IUP-OMPS V2.6). The main differences between V2.6 and V3.3 are in the usage of the spectral segments and normalization THs . Table 1 lists the TH ranges, respective spectral segments selected for the retrieval, TH used for the normalization, and the order of the polynomials used for V3.3. Figure 1 presents a comparison of ozone profiles from Microwave Limb Sounder (MLS) with IUP-OMPS V2.6 (left panel) and V3.3 (right panel) data. Both panels show relative differences between IUP-OMPS and the MLS L2 version 4.2 data as a function of altitude and latitude. Differences within ±10% are observed above 20 km for both IUP-OMPS versions. Below 20 km, the differences can reach ±30%. An overall reduction of the bias is found for IUP-OMPS V3.3, with main improvements in the lower tropical stratosphere and between 35 and 50 km. However, the bias increases between 30 and 35 km from 20°N to the south and below 20 km polewards of 60°S.

The OMPS-LP ozone profile time series based on L1 V2.5 data, which are used by both V2.6 and V3.3 retrievals, were found to exhibit a significant positive drift after 2018 (Kramarova et al., 2018). For this reason, only the data until 2018 are used to create the OMPS-LNM-TrOC dataset. Currently, only measurements from the central slit are used to retrieve ozone profiles because of remaining calibration issues related to the measurements from the side slits. More information and technical details on OMPS-LP can be found in Kramarova et al. (2018), Arosio et al. (2018), and references therein.

Figure 1. what is the time period for the comparisons? 2012-2018? Please, specify.

A: This figure shows the average for 2013. It is now explicitly stated in the figure and the caption.

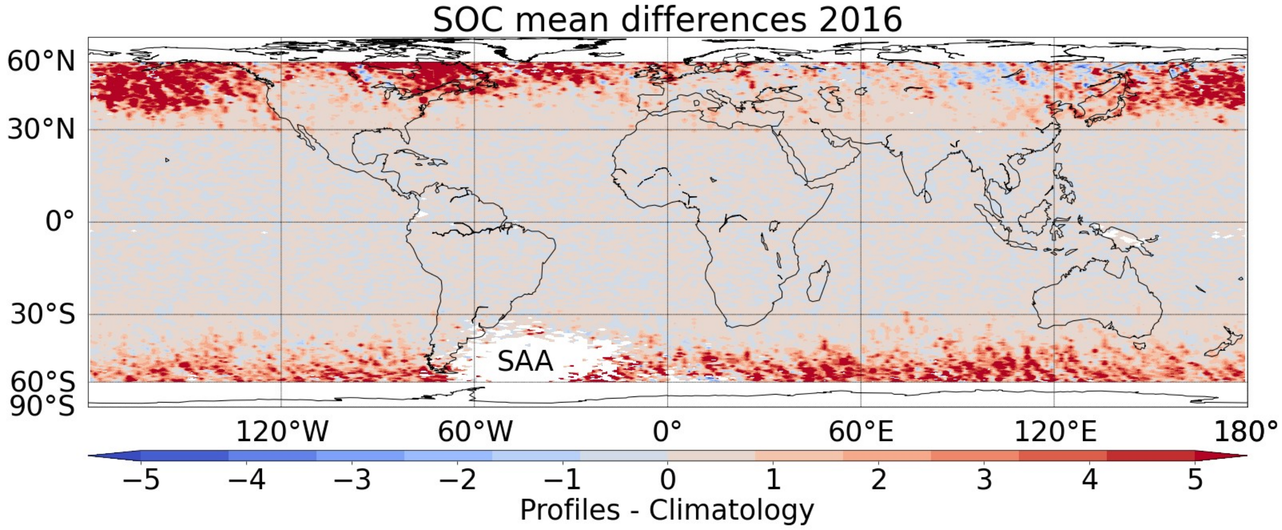
P.7 lines 162-163 I assume you meant that retrieved ozone is very noisy below 12.5 km.

A: It is clarified in the revised version of the manuscript that the retrieval uncertainty is larger:

“only ozone values above 12.5 km are considered because of a large retrieval uncertainty of the limb profiles below this altitude (Arosio et al., 2018)” (Lines 173-174)

Have you tried to use retrieved values instead of climatological? Does that change calculated SOC values in any significant way?

A: As suggested by the reviewer, we processed a sub-sample of the data set by using the retrieved values when the tropopause high is lower than 12 km (and higher than 8.5 km), instead of the climatology. Differences of around 5 DU in the SOC field are found at high latitudes when using the retrieved values below 12 km instead of the climatology (see figure below). Below 12 km, the vertical resolution of the profiles is, on average, larger than 10 km, which makes the retrieved values unreliable. The use of the climatology gives larger SOC values, leading to lower tropospheric ozone columns.



P.7 Lines 167. Are you deriving the cloud height from OMPS-LP? Then you should have mentioned how it's done in Sec. 2.2.

A: It is now mentioned in Lines 125-128:

Clouds in the instrument field of view are detected using the Color Index Ratio (CIR) concept described in Eichmann et al. (2016). The ratio between two radiances at wavelengths with weak ozone absorption (754 and 868 nm), called the Color Index (CI), is calculated for every TH. The CIR is defined as the ratio of the CI at two neighbouring THs. For CIR higher than 1.08, the tangent height is marked as cloudy.

P.7 line 184 Since the cloud fraction is one of the key factors in producing TrOC, it would be good if you describe how the cloud fraction is calculated in Sec. 2.1 when you talk about TOC.

R: The cloud information is now detailed: "The cloud fraction information is obtained from the operational OMPS-NM L2 product V2.1 from NASA (Jaross, 2017). The retrieval of effective cloud fraction is made using the Mixed Lambert Equivalent Reflectivity model, using a weak ozone absorption wavelength, 331.2 nm for most conditions, and 360 nm for large SZAs and high amounts of ozone (Seftor and Johnson, 2017)." (Lines 104-107).

Figure 2. All labels should be explained. What does "Pixel n^{o} 110 and FOV n^{o} 14" or "State n^{o} 85" mean? Also, it might be better to say: "The red points mark the footprints of tangent points (TPs) of the limb observations".

A: The caption was modified as follows:

Example of the matching between OMPS-NM and OMPS-LP observation scenes. The red points mark the footprints of tangent points (TPs) of the limb observations. The grid cells represent the ground pixels of OMPS-NM. The yellow boxes indicate the nadir ground pixel averaged to obtain the TOC assuming the nadir pixels are cloud-free. The orange box marks the exact match between the OMPS-NM ground pixel number 110 (along-track) in position 14 (across-track) with the OMPS-LP observation number 85 (state).

Table 2. I am not sure I understand the source of error titled "Tropospheric ozone increase". First I thought it represents the error in apriori, but then I found you have another entry for "O3 and T apriori". Please, explain how the changes in tropospheric ozone affect the retrieved TOC.

A: It is clear now that it is indeed an error in the a-priori, but only in the tropospheric ozone profile.

Table 2. What is the threshold for "enhanced aerosol"?

A: Details can be found in Orfanoz-Cheuquelaf et al., 2021, as it is now mentioned in the paper (Line 2013).

The following text is from Orfanoz-Cheuquelaf et al. 2021:

We generated synthetic radiances for different aerosol scenarios using SCIATRAN V4.2 with the aerosol parameterization from LOWTRAN (Kneizys et al., 1988; Shettle and Fenn, 1979). From these radiances, the LER albedo was re trieved and used in the WFFA retrieval. The synthetic radiances were calculated with a total ozone of 325 DU, solar zenith angles of 59.88 and 27.02° (chosen from real values of OMPS-NM ground pixels), visibility of 2 km, and surface albedos of 0.05 and 0.2. The different types of boundary layer aerosols are maritime, rural, tropospheric, and urban. One case with extreme volcanic stratospheric aerosol loading was included.

p.9 line 218. I assume you mean to include a reference after "…are quantified using synthetic retrievals and extensively discussed".

A: is included " in the above mentioned study". (Line 232)

Figure 3. What did you mean by "syst. bias"? Biases in measured radiances or biases in retrieved ozone?

A: It is corrected now to "Retr. Bias", referring to the retrieval bias and also corrected in the equation.

Section 4.2. and Fig. 3. In my view several types of errors (T, P, albedo and TH) are highly correlated and it's hard to isolate contributions from them. Limb scattered radiances are directly proportional to atmospheric air density. T and P are used to calculate the density, and therefore these errors are correlated.

A: The responses in ozone profiles to these errors are indeed correlated, the errors in P, T, and TH themselves not. We see no reasons why we cannot investigate influence of each of this error independently and then sum them up in the Gaussian sense.

I feel that TH error is perhaps the leading source of uncertainties in limb scattering measurements. I am not sure what uncertainties in TH you are assuming to get the TH error. I believe that the TH error has two components systematic and random.

A: The error in TH used for the error estimation, i.e. ±100 m, is based on the information from the NASA team and corresponds to the best current knowledge. Contrary to the random error, a systematic error in TH is easily identified by validation studies. For now, no indications for a systematic TH error were found.

Same is with Pressure. If sea/surface pressure is off in GEOS assimilation that would produce a systematic bias in P profiles.

A: We are not aware of any reports about a systematic error in pressure data from GEOS. Unknown errors cannot be included in the error budget.

In some conditions the two errors will cancel each other, in other cases they will amplify the resulting error.

A: This is true, this is why the errors are summed up in the Gaussian sense.

Wouldn't error in albedo be a systematic error? Albedo depends on absolute calibrations and TH accuracy.

A: No, the error in albedo is predominantly random as it is retrieved for each measurement independently. Uncertainties of the aerosol scattering has the largest contribution to the albedo error. There are no indications of a dominating systematic error neither in the calibration nor in TH nor in the albedo retrievals.

I am confused with how the errors are sorted by systematic and random.

A: They are sorted based on the best knowledge about the error nature and in some cases are educated guesses. For more details on the classification of the errors, please refer to Arosio et al. 2022.

Section 4.3. Again you are ignoring the systematic part of the error in TPH which might be much larger than the error in vertical resolution you quoted here.

A: We found no indication of a systematic error when calculating TPH. There is no way to account for unknown errors in the error budget calculation. However, there is certainly a systematic part related to the choice of the TPH definition, which shall be taken into account when comparing different data sets (see efforts in TOAR II), rather then in the uncertainty budget itself.

P.12, lines 279-281. Since you are considering only nadir pixels and illuminating cloudy pixels, you end up with the stripes of data along the orbital tracks. The statement that "you are binning data" in my view misrepresents the reality. Perhaps, it would be better to say that you map your sparce measurement on to the regular grid.

A: The text is modified as: "the OMPS-LNM TrOC data were mapped onto a regular daily grid of 0.5° × 1.5° (latitude/longitude) from 60°S to 60°N." (Lines 295-296)

Figure 5. Please, specify how these anomalies were calculated. Did you subtract the long-term annual averages or seasonal averages? Also please, add a) and b) signs to each panel.

A: "The anomalies were computed by subtracting the long-term mean from all data in the tropics." (Line 332). The figure was updated, including adding labels a and b.

Figure 6. Be more specific. I assume SOC is from OMPS-LP as well as could top height, right? Is the surface reflectivity from OMPS LP or OMPS NM?

A: The figure is now more specific, as well as the text.

p.15 lines 316-318: In the text you need to add references to the figures you meant “The SOC anomalies (Fig.5b) show lower values over the Pacific and Atlantic, matching the band of high TrOC (Fig 4). This feature is not evident in the TOC anomalies (Fig. 5a)”

A: Done (Lines 332-334)

P.15 lines 322-324. How did you calculate the surface reflectivity? Is it from OMPS LP or OMPS NM?

A: Now mentioned in the text “OMPS-NM” (Line 345).

P.16, lines 340-346. Are these weighted averages? How many adjacent pixels were used/considered? What is the range of distances between the station and co-located OMPS? How was the temporal averaging applied?

A: It is mentioned that “OMPS-LNM data from the grid cell enclosing the launch site and all immediately adjacent grid cells were averaged (without any weighting) to create the collocated OMPS dataset. The temporal averaging included OMPS-LNM data from the day of the ozonesonde launch, one day before, and one day after the launch.” (Lines 362-364)

P.16, line 360. How was the bias calculated? Did you calculate the mean for each instrument and then estimated the bias? Please, explain.

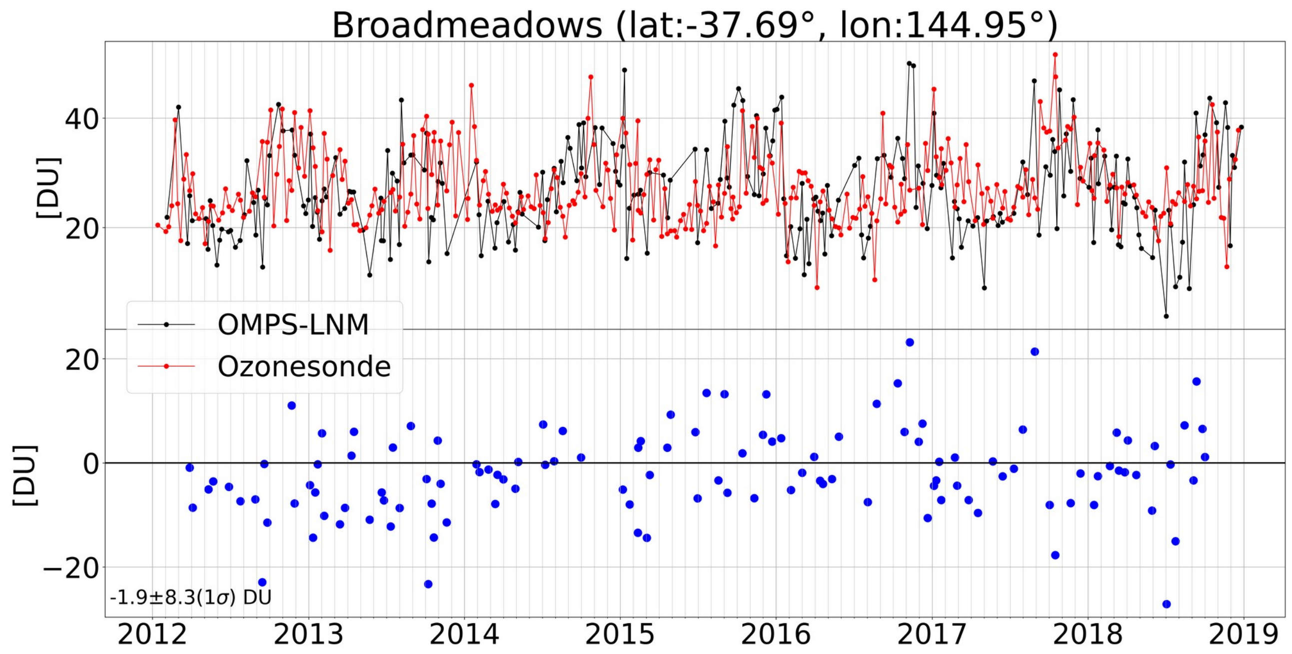
A: “mean difference (average of OMPS-LNM minus Ozonesonde time series)” (Lines 376)

Figure 7. Are you showing individual measurements or weekly/monthly averages. Please, specify that in the figure caption.

A: New caption: “Time series of tropospheric ozone column from individual ozonesonde measurements (red) and daily averaged collocated OMPS-LNM (black) data for three selected sites.”

P.20, line 420. From figure 7 at Broadmeadows it seems that OMPS LNM overestimates the seasonal cycle compared to sonde as well.

A: We compared the two seasonal cycles in more detail and did not find a significant difference between the two. The following figure shows the time series and the absolute difference between them.



P.20, conclusions, lines 433-435: I disagree with the conclusion that “no seasonality in the differences”. There is a clear seasonal bias in SH between LNM and OMI/MLS shown in Fig. 9.

A: The text has been modified as “with a seasonal differences of up to 10 DU in the extratropics. Nevertheless, a good agreement in the long-term variability is observed.” (Lines 456-457)

Minor comments:

3 line 70. Should it be “OMPS comprises of three instruments…”

A: To our best knowledge, “OMPS comprises three instruments” is the correct expression. (Line 69)

P6., line 143. You defined PV above, but never defined PVU.

A: Done (Line 153)

Figure 6. It should be “over the Pacific Ocean from OMPS”

A: Done

References:

Arosio, C., Rozanov, A., Malinina, E., Eichmann, K.-U., Von Clarmann, T., and Burrows, J. P.: Retrieval of ozone profiles from OMPS limb scattering observations, Atmospheric Measurement Techniques, 11, 2135–2149, https://doi.org/10.5194/amt-11-2135-2018, 2018.

Arosio, C., Rozanov, A., Gorshelev, V., Laeng, A., and Burrows, J. P.: Assessment of the error budget for stratospheric ozone profiles retrieved from OMPS limb scatter measurements, Atmospheric Measurement Techniques, 15, 5949–5967, https://doi.org/10.5194/amt-15-5949-2022, 2022.

Orfanoz-Cheuquelaf, A., Rozanov, A., Weber, M., Arosio, C., Ladstätter-Weißenmayer, A., and Burrows, J. P.: Total ozone column from Ozone Mapping and Profiler Suite Nadir Mapper (OMPS-NM) measurements using the broadband weighting function fitting approach (WFFA), Atmospheric Measurement Techniques, 14, 5771–5789, https://doi.org/10.5194/amt-14-5771-2021, 2021.