

## Reviewer 2: Comments

### Major comments

---

*As far as I understood the CHORA data are provided on a daily basis. To me it is not clear whether the validation is performed on a daily or monthly basis:*

*I127: "Ozonesonde profiles are considered collocated to the satellite data when the station is inside the spatio-temporal grid-box of the TROPOMI tropospheric ozone data, here  $0.5^\circ \times 0.5^\circ$  and the same day". (daily comparison)*

*I246: "The monthly mean of tropospheric ozone retrievals ... from all three CHORA algorithms were calculated." (sounds like monthly data)*

**Ans:** Here we intended to say spatially collocated. Changed the sentence to "Ozonesonde profiles are considered spatially collocated to the satellite data when the station is inside the grid-box of the TROPOMI tropospheric ozone data, here  $0.5^\circ \times 0.5^\circ$ ." We provide daily data, but validations are conducted using monthly averages.

*Despite the first statement it seems to be on a monthly average: To a large extent the deviations in the subtropics are driven by the March soundings in Hanoi. As one explanation the authors mention "The challenge of limited ozonesonde measurements in March further complicates the validation." (L274) You would need at least 15-20 measurements per month to represent the monthly mean values adequately. This is not given at any sounding station. Because of that I recommend comparison on a daily basis.*

**Ans:** From June 2018 to December 2022, Hanoi has only 4 ozonesonde measurements in March which is too sparse to compare.

We appreciate your suggestion, but upon reviewing our data, particularly for subtropical stations, we noticed a lack of sufficient daily collocated data for validation. Therefore, we chose to rely on non-collocated data to ensure a more robust analysis. It's important to note that our daily TCO retrievals do not involve any smoothing. Seasonal variation of TCO with daily collocated data is shown in the supplement (Fig.S1).

*The description of the CPC algorithm (fixed pacific reference) states that it is designed for the  $20^\circ\text{S}$  to  $20^\circ\text{N}$  latitude band. However, throughout the manuscript it is applied from  $25^\circ\text{S}$  to  $25^\circ\text{N}$  and beyond e.g. figure 8 (comparison to Irene  $25.9^\circ\text{S}$ ). When looking at the operational CCD dataset, the data in the "winter hemisphere" are often corrupted due to the lack of convective clouds outside the  $15^\circ$  range. Here some more details should be added how the respective extension is achieved.*

**Ans:** Our aim was to assess the performance of the CPC algorithm in extending it to the subtropics and compare it with local cloud algorithms. While the CPC algorithm

follows the same methodology as the operational TROPOMI Level 2 product, it differs primarily in the threshold for cloud fraction ( $CF \leq 0.2$ ) used to select non-cloudy scenes, the number of days averaged, climatology, and grid resolution. However, we have chosen not to validate CPC beyond  $20^\circ$  latitude. The figures and text have been updated accordingly.

## Minor comments

---

*l 124: "and when the burst height of the sonde is below 15 km." replace the "and" by "or"? what is the reason for the second requirement? Of cause, if the burst height is below the 270 hPa altitude level the data are not useful but it does not matter the balloon bursts at 14.8 km*

**Revised:** replaced “and” by “or”

The sonde column is not processed when data gaps in the profiles are wider than 1.2 km or when the burst height of the sonde is below 15 km.

*p 5: The authors should decide whether the column between the surface and 270hPa is above (caption figure1) or below (l134) this pressure levels. From my point of view both have their justification: "above" (because the pressure gets higher) or "below" as the altitude is below the 270hPa level, but using both (within a few lines) is a bit confusing.*

**Revised:** Replaced “below” by “above” for the caption of Figure 1. The other sentence (**earlier l 134**) doesn’t exist anymore after updating the text.

*l 151: The product readme file recommends a qa-value greater than 70 (refrence ESA\_21b, [https://sentinel.esa.int/documents/d/sentinel/s5p-mpc-dlr-prf-o3\\_tcl\\_v02-06-01\\_2-7\\_20231129\\_signed](https://sentinel.esa.int/documents/d/sentinel/s5p-mpc-dlr-prf-o3_tcl_v02-06-01_2-7_20231129_signed)). This will probably not improve the validation with sondes.*

**Revised:** We followed the variable description in the ESA operational data file, which states that the "qa\_value is a continuous quality descriptor, varying between 0 (no data) and 1 (full quality data). It is recommended to ignore data with qa\_value < 0.5."

We recalculated the data using a qa-value threshold of greater than 0.7 and updated the figures (Figures 8 to 11 and 16) and the numbers in Table 1 accordingly.

*p5 and 6 I suggest to combine the descriptions of the operational ESA product and the CPC. The algorithm is basically the same except for the climatology and the daily vs 3 day averages.*

**Revised:** Modified the text. We explained the CCD algorithm in general and clarified that while CPC and the ESA product use the same methodology, they differ in terms of climatology, averaging, and grid resolution (**l 200-209, Section 4.4**).

*I 165 The CPC algorithm is described as being applied between 20°S and 20°N, but in Figure 13 data up to 25° South and North are shown (see also major comments).*

**Revised:** As mentioned earlier, the validations for CPC beyond 20° are omitted and updated the figures and text accordingly.

*I 216: This sentence is unclear: " ... by substituting the values of slope and medians of the ACCO ... "by what? and why replacing the median of the ACCO?*

**Revised:** Replaced “substituting” with “applying”. The Theil-Sen estimate of the intercept can be obtained from the linear regression of ACCO against cloud top pressure (now I 191)

*figure 6: "latitude band 2.4° S - 0.4° S and longitude band 26.8° E - 46.8° E" according to the previous description I expected a latitude range of  $\pm 1^\circ$  for a grid box of  $0.5^\circ$ , so  $1.25^\circ S + 1^\circ = 0.25^\circ S$  to  $2.25^\circ S$  but not  $2.4^\circ$ . The same holds for the longitudinal range. I assumed the centre of the grid box to be the respective starting point. How fast are the longitudinal ranges extended. Starting with  $\pm 5^\circ$ , what is the next step if the number of cloudy data is too low,  $\pm 10^\circ$ ?*

**Revised:** Updated the figure and the latitude and longitude range. Yes. Regarding the longitudinal expansion the next step would be  $\pm 10^\circ$ , with an increment of  $\pm 5^\circ$ .

*I245 I suggest to look for additional sonde data at <https://woudc.org/data/explore.php?lang=en> especially for the subtropical areas (Hong Kong "King's Park")*

**Revised:** Validated the results with ozone sonde data over King’s Park. Updated Table 1 and all figures (Figures 8 to 16) accordingly.

*I 271 as well as in many other places (figures 10 and 11) throughout the manuscript, the seasons are reverred to as Spring to Winter, according to our northern hemispheric seasons. I am not sure if this useful in the tropical to subtropical regions. I suggest to use the month instead, or local dry and wet season.*

**Revised:** We have updated the seasons in the text using boreal and austral terminology. The months corresponding to each season are specified in the figure captions (Figures 12 and 13).

*Figures 8 and 9 Can you include the operational O3\_TCL product in this figures? It is included in figure 13 so the analysis has already been done anyway.*

**Revised:** We added the operational O3\_TCL product (called CCD-ESA in the paper) to Figs. 8,9,10,11, and S1.

**figure 10:** *this is a very interesting figure - maybe you could make more use of it? For example, the median distance of deep convective clouds over Ascension Island is (3000km) more or less independent of the season. That means the deep convective clouds are over the South American continent.*

**Revised:** Figure 12 has been updated. Additionally, the station-by-station analysis for Ascension Island has been revised (**I 375 to I 388**).

**Figure 11:** *include the 270 hPa level in the figure*

**Revised:** Included and updated the caption of the figure (now labelled as Figure 13) accordingly.

**Discussion of individual sounding stations.**

*For some stations it might be nice to have figure of the differences per month, however including this in the manuscript might extend it too much. How about adding such figures to a supplement?*

**Revised:** We value your suggestion. We think the figure (Fig.11) is important for displaying the seasonal variations in terms of differences and supporting the validation process. To maintain clarity and provide important insights, we have decided to keep the figure in the main text.

**I 315ff (Paramaribo)** *How about local sources, biomass burning might also occur in the vicinity of Paramaribo?*

**Ans:** According to Peters et al. (2004), during the dry season (Aug- Nov) the amount, frequency, and timing of ozone pollution events at Paramaribo depend mostly on the occurrence and intensity of fires over Africa, the lofting of plumes to the free troposphere by convection, and subsequent transport by the tropical easterlies. The transport of biomass-burning plumes from South America itself is less frequent in this period.

**I 343-346 (Kuala Lumpur)** *The CPC compares better to the sondes than the local cloud reference algorithms. Does this imply that for the local reference algorithms the reference area might be too small and should be increased by default. How large are the local reference areas for a typical dry and wet season? According to figure 10 the mean distance is very small (~200km) only for the Nov-January region the distance is slightly larger.*

**Ans:** As mentioned in the paper, even though the local cloud reference sector used in both CLC and CLCT algorithms belongs to the Pacific sector, the fixed large Pacific cloud reference sector ensures the inclusion of more high-reflective clouds and a comparatively less polluted background, which benefit the CPC retrievals over Kuala

Lumpur. The local reference areas are selected automatically whenever more than 50 cloudy scenes are found, which is the case for Kuala Lumpur. Increasing the local reference area too much by default would reduce the sensitivity of ACCO measurements.

The Figure of cloud distances from the station (Figure 12) is now updated.

**Figure 13:** *maybe a longitudinal plot might be more telling. The fixed reference region algorithms (CCD-ESA and CPC) compare best to the sondes in the reference region (70°E to 170°W, Fiji, Samoa and Kuala Lumpur), outside the reference region the CLC algorithms might show their capabilities.*

**Ans:** Even though Fiji and Samoa are in the fixed Pacific reference sector, local cloud algorithms show similar or better agreement with ozonesondes compared to the CPC and CCD-ESA algorithms. Only over Kuala Lumpur do the CPC and CCD-ESA algorithms show better agreement.

*I 481: this sensitivity study might be interesting, parts of it might be shown in a supplement.*

**Ans:** The sensitivity study was conducted using non-gridded data. Unfortunately, I am unable to include this in the supplement at this time.

*I 500 shift figure 13 to here.*

**Ans:** Shifted Figure 13 (now Figure 16) to the end of the sentence. Changed the sentence to, “The operational TROPOMI tropospheric ozone data (CCD-ESA) consistently exhibit a positive bias across all ozonesonde stations, which is maximum among all satellite datasets (Fig.16) (now **I 542**).

## Technical comments

---

*I 58 "... in the time frame of 2017 -2022" even if this was the original plan, meanwhile it's 2024 S5P is still operational and we get very nice results every day. I suggest to replace by " ...the timeframe after 2018"*

**Revised:** Changed accordingly. “...providing information and services on air quality, climate forcing, and ozone layer in the time frame after 2018.”

*I 84/85 "3.5 x 7.5 km<sup>2</sup> (across-track x along-track), was further refined to 5.5 x 3.5 km<sup>2</sup>". exchange the updated resolution to 3.5 x 5.5 km<sup>2</sup> to have across-track x along-track.*

**Revised:** Changed accordingly. “..., originally of 3.5×7 km<sup>2</sup>, was further refined to 5.5×3.5 km<sup>2</sup> (across-track × along-track) on 6 August 2019.”

*l 475 "significantly small", what does this mean? significant or small or significant but small?*

**Revised:** That exact sentence doesn't exist anymore after updating the text. (l 487-489)