Review of acp-2021-378
Anonymous Referee #1

Referee comment on "Comparative assessment of TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column measurements" by Isabelle De Smedt et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2021-378-RC1, 2021

Review of https://doi.org/10.5194/acp-2021-378, by De Smedt et al. (2021)

This study presents an evaluation of the formaldehyde (HCHO) product derived from the UV observations performed by the Sentinel-5 Precursor (S5P) TROPOspheric Monitoring Instrument (TROPOMI). First, the measurements are compared with the HCHO product from its predecessor, the Ozone Monitoring Instrument (OMI). It is shown that both HCHO products are in good agreement and that differences are partly attributable to the different treatment of the clouds between the two sensors. Moreover, TROPOMI allows the monitoring of HCHO tropospheric columns with a substantial reduction of the noise, as well as the detection of weaker HCHO signals at much shorter time scales than what is achieved with OMI. This is also illustrated with a case study; the detection of two shipping lanes in the Indian Ocean. Both OMI and TROPOMI HCHO products are then validated against a network of MAX-DOAS measurements. Although the satellite instruments show a similar underestimation of the high HCHO columns, TROPOMI represents an improvement for the low and medium HCHO columns and permits a validation at a daily time scale, in the closest vicinity of the MAX-DOAS stations. This assessment exercise is a complement to another validation of TROPOMI HCHO against ground-based FTIR measurements (Vigouroux et al., 2020), which are characterized by a maximum sensitivity in the free troposphere whereas the MAX-DOAS primarily probes the lowermost layers.

It is a comprehensive evaluation of the TROPOMI and OMI HCHO products, led by researchers with a long experience in retrieval of trace gases, and of HCHO in particular, from UV-Vis satellite sensors. The paper is well written and the structure is clear. To my point of view, the paper is a bit long with many items (23 figures and 5 tables in the main text!), but I assume an evaluation effort involving several large datasets is difficult to keep short. The analyses are robust and the assessment of OMI and TROPOMI against MAX-DOAS observations is convincing. TROPOMI’s performance at short temporal scales is very impressive and the paper demonstrates nicely the gain in precision compared with OMI. With the continuously growing use that is made of the TROPOMI datasets, such a comprehensive evaluation will undoubtedly be a valuable milestone for further users and developers of the TROPOMI HCHO product. I do not see any major issues with this study and I only have comments the authors can likely easily address. Therefore, I support publication after a minor revision step.
The vertical sensitivity of OMI and TROPOMI on individual scenes are unfortunately not discussed. It would have been very interesting to compare, e.g., the OMI and TROPOMI averaging kernels in typical observational conditions in order to know whether TROPOMI allows probing the tropospheric layers closer to the surface. Although the manuscript is already quite long, perhaps a discussion can be added in this regard.

It is clear that cloud filtering and cloud correction are crucial for ensuring satisfactory HCHO retrievals from the satellite sensors. On a reference sector, for a typical day, what is the proportion of OMI and TROPOMI measurements that are discarded because of the cloud coverage or that are affected by residuals clouds? For example, how many measurements fall into, respectively, the CF<0.1, 0.1<CF<0.4 and CF>0.4 categories? I also assume that the MAX-DOAS observations are affected similarly by the presence of clouds. Therefore, are the MAX-DOAS measurements filtered out or corrected the same way? I see on Figs 15 and 17 that the OMI and TROPOMI clear-sky columns (Nv_clear) are compared with the MAX-DOAS measurements (called Nv). Are these latter also calculated using clear-sky AMF? If no, does it affect the comparison with the satellite data?

Lines 241-245: Perhaps there is something I have misunderstood here, but it is relatively unclear to me why and how the OMI observations are daily gridded on a grid with a resolution as high as 0.05 x 0.05°. In the tropics, such a grid cell represents roughly a 5 x 5 km area, even smaller at higher latitudes, whereas the typical OMI pixel size is much larger than that. Hence, a single OMI pixel should overlap several grid cells. How do you attribute, on a daily basis, a value to such a small grid cell with observations having a footprint of 13 x 24 km in the best case? Without overlaps between the OMI observations taken the same day, I guess that oversampling techniques cannot be used. As well, what do you mean by "we require the region to be filled with at least 50% of valid grid cells"? How can you have a minimum of 2 OMI observations per 0.05 x 0.05° grid cell on a daily basis?

Lines 336-337: If I understood correctly, the random errors on the tropospheric vertical column is driven by the error on the slant column. However, I assume that the AMF calculation is also affected by uncertainties, which in turn would convolve with the SCDE and add to the total uncertainty on the vertical column. Is that accounted for?

Lines 529-539: Despite the underestimation, the results of the comparison with the BIRA-IASB instrument at Xianghe are actually more in line with the comparisons shown by Vigouroux et al. (2020) and the overall underestimation of the larger columns by TROPOMI. Perhaps it is worth to be mentioned.

Figure 20: Can you explain why the second ship line is detected more to the North by OMI? Is it because the OMI 2005-2009 dataset used here is an annual mean, whereas the TROPOMI data are selected between Dec. 2020 and Feb. 2021?

Figure S1, right panel: I am a bit puzzled to see that big cities in Western and Northern Europe (e.g., London, Paris, the Ruhr Basin, etc.) are not visible whereas cities located more to the South around the Mediterranean Sea clearly are (e.g., Marseille, Barcelona, etc.). Anthropogenic emissions of HCHO precursors should also be quite high in
Western and Northern Europe. Similarly, it looks like the annual HCHO columns are at least as high over the Sahara Desert, far from any direct sources, as over, e.g., England and most of Western Europe. Is there a latitudinal dependence of the retrievals that would introduce a bias?

**Technical comments/Typos**

- Line 32: assesses
- Line 49: a factor of 3
- Lines 56-61 and further in the text: There are several species that are not defined when they are first used. E.g., nitrogen oxides (NOx), carbon monoxide (CO), etc.
- Line 63: Because of
- Line 91: Do you mean a daily average on a global 20 x 20 km resolution grid or in a 20 km-radius area around a specific point?
- Line 97: requirements
- Lines 121-122: On board Aura
- Line 138: On board the S5P
- Line 273: 3-year averages
- Line 315: I think using “3-year” instead of “long-term” here is more appropriate.
- Figure 5: The color bar of the bottom panel could be better adapted; the differences are currently difficult to distinguish.
- Line 423: I would avoid using “HCHO emissions”, knowing that HCHO is mainly secondarily formed in the atmosphere.
- Line 445: see Table 2
- Line 448: 20 km radius?
- Line 503: Mexico, respectively.
- Line 523: I suggest using “relative” instead of “compared”
- Table 3, 4th row: Must be “±” instead of “+” in the two columns
- Line 645: Figure 20).
- Line 673: a factor of 3
- Line 678: 5-year OMI climatology
- Line 695: suggests
- Line 711: emissions
- Line 717: Delete one of the “molec. cm-2”
- Line 719: sources
- Line 769: “Owing to the sensitivity of TROPOMI”. I disagree. It is primarily a matter of spatial resolution and high spatial sampling, not of sensitivity. For example, it is said earlier that such shipping lanes could already be detected with GOME, combining many years of measurements.
- Several references are not up-to-date (e.g., still refer to papers in discussion whereas the final papers have already been published).