Review of "Satellite remote-sensing capability to assess tropospheric column ratios of formaldehyde and nitrogen dioxide: case study during the LISTOS 2018 field campaign", Johnson et al., AMTD, 2022.

Johnson et al. present a detailed statistical analysis of FNR observations with two different OMI products, and with TROPOMI. FNR, NO2 and HCHO satellite retrievals are validated against airborne measurements over the New-York area during summer 2018. It is well demonstrated in the paper that the noise of the HCHO satellite retrievals is the limiting factor for the FNR observations since the individual HCHO columns need to be averaged at poorer time and space resolution than NO2. The precision of the OMI HCHO observations does not allow for daily FNR observations at OMI native resolution. The OMI QA4ECV HCHO product is found to perform better than the OMI NASA HCHO product. TROPOMI offers an important improvement in the spatial and temporal resolution of HCHO and NO2 tropospheric columns, allowing for daily FNRs retrievals at TROPOMI native spatial resolution. The results are further improved by averaging TROPOMI observations on a larger spatial grid. Both NO2 and HCHO satellite products suffer from bias compared to aircraft observations. This study identifies an important positive bias over rural regions (lowest columns) for both species, and for OMI and TROPOMI products. However, the positive bias found for the TROPOMI products is reduced compared to OMI thanks to the better spatial resolution and lower noise. It is also demonstrated that the bias of the FNR satellite observations is much lower than the respective NO2 and HCHO biases. This is an important result that would deserve more discussion in the paper. The paper is well written, albeit a bit long and too detailed. The scientific approach is solid, however some points should be tested or clarified. I recommend publication in AMT after some revisions.

General comments

One concern is the small number of days that are available for the validation. Here the field campaign covers only a few days (8 days collocated with OMI, 12 days with TROPOMI). The statistical results are not always significant, especially for OMI. Studies on longer time period could improve the observed correlations, that are poor for HCHO.

- The paper could be improved by providing an information about the spatial and temporal resolution that might provide useful FNR observations with OMI (ex. monthly averaged data). How many observations are needed at minimum to reduce the noise at the level of the TROPOMI daily observations?
- It would be interesting to know if the HCHO observations with aircraft instruments are also noisier than the NO2 observations, and therefore also the limiting factor of suborbital FNR observations.
- In Table 2, I recommend adding a line providing the mean value +- the standard deviation of FNR, HCHO and NO2 for the aircraft, NASA OMI, QA4ECV OMI, and TROPOMI (0.15° and 0.05°).
- I recommend more tests on the selection of the data, that is currently at the edge of the statistical significance (see later).

One interesting result of the paper is that the errors in NO2 and HCHO columns tend to offset in the FNR observations. There might be good reasons for this, such as error cancellation. It is therefore important to use HCHO and NO2 products that have been retrieved with algorithms and auxiliary data as consistent as possible. This is an important message for the future TEMPO product.

- It would be good to discuss further what type of error might cancel out, or at least might reduce, when using NO2 and HCHO retrieved using consistent algorithms to derive FNR (surface albedo, cloud products, a priori profiles).
- I recommend adding a table providing a quick look at the auxiliary data used in the AMF calculations for the NASA, QA4ECV and TROPOMI products, and TEMPO.
- Discuss the different FNR biases with the level of consistency between NO2 and HCHO AMF settings.

The low HCHO correlations are also partly due to lower spatial variability of the HCHO distribution compared to NO2, also in the airborne measurements, over the time and domain of the study.

Selection of data:

- Filter row anomaly both for HCHO and NO2 products.
- The lower bound limits for HCHO and NO2 appear to bestrict, compared to the reported standard deviations of the bias. For HCHO, the bias std ranges from 9 to 5e15 molec.cm-2, while the lower limit has been set to -8e15. For NO2, bias std is about 4e15, while the lower limit has been set to -1e15 molec.cm-2. There is a possibility that a significant part of the negative values has been filtered out while it actually belongs to the normal distribution. The effect could be an artificial increase of the mean background values. Please test a lower bound limit for the data selection.
- At the spatio-temporal resolution of the study, OMI retrievals are clearly at their detection limit. Please consider testing a lower grid resolution (0.2°) for OMI.
- To increase the number of collocations, I would suggest testing a larger temporal window of 3h for the airborne retrievals.

It would be good to better stress the specificities of this paper compared to the recent paper of Souri et al., 2022, which also compares OMI and TROPOMI NO2, HCHO and FNR errors over the US. (Characterization of Errors in Satellite-based HCHO / NO2 Tropospheric Column Ratios with Respect to Chemistry, Column to PBL Translation, Spatial Representation, and Retrieval Uncertainties)

Detailed comments

Abstract

Line25: “high spatiotemporal coverage”: please provide numbers, such as the native resolution of OMI and TROPOMI. I would rephrase “OMI and TROPOMI are capable of providing NO2 and HCHO daily global observation at native resolution of respectively … and …”. However, satellite observations are known to be affected by noise and biases, that limit the precision of FNR.

Line 25: “…, yet a recent study suggested ….”. This sentence is rather vague. Which study?

Line 30: Please specify the covered period.

Line 32: Please be clearer in the abstract with the term “suborbital”. This is not obvious for a general reader.

Line 49: Place replace large by larger biases.

Introduction

Line 95: please add the 2 following references: Wang et al, 2022; Harkey et al., 2015.

Line 100: the choice of references seems weird. It might be good to add references for NO2 and HCHO L2 products of each sensor, and not only for studies using both species together. The SCIAMACHY instrument is missing in the list.

Methods

Line 163: The OMI rows affected by the row anomaly should be filtered out in the HCHO product such as in the NO2 product. The reference sector method does not correct for the row anomaly, but for the stripes between the valid rows. Please rephrase (and check that the HCHO data are filtered correctly).

Line 204: Please explain what you mean by "iterative fitting algorithm" and "simultaneous fitting". To me, a DOAS fit is an iterative fit (least-squared fit).

Line 215: The QA4ECV fitting window is 328.5-359 nm, such as TROPOMI. For all HCHO products, please double check the retrieval intervals that are mentioned in the paper. Most of the recent retrievals use a fitting window larger than 328.5-346 nm.

Line 261: Please explicit the term SWs.

Results

Line 444: "Tropospheric columns NO2 concentrations", "tropospheric columns NO2 retrievals". Could be simplified to "Tropospheric NO2 columns" and homogenized throughout the paper.

Line 465: It should be emphasized here that TROPOMI offset for low columns is lower than OMI at the resolution of 0.05.

Line 491: add a reference to Verhoelst et al. 2021.

To our knowledge, the cited references do not report a high bias of NO2 for background values. But the studies were made with the previous version of the TROPOMI NO2 product. This should be clarified here.

Line 495. The comparison of TROPOMI NO2 Bias at 0.05 and 0.15° also clearly shows the spatial resolution effect on the background values (from negative to positive and similar to OMI NMB). Please mention this resolution effect.

Table2: Please add one line with the mean FNR, NO2 and HCHO columns and their standard deviations.

Figure 3: Please test different data selection as suggested in the general comments.

Line 517: The results are not so much in agreement with the study of Vigouroux, who reported indeed a high bias for the lowest columns, but for columns lower than 2.5e15 molec.cm-2. The TROPOMI bias ranges from 0 to negative values for columns larger than 5e15 molec.cm-2.

Line 527: Please also compare the bias standard deviation between OMI and TROPOMI.

Line 530: In De Smedt 2021, it is reported that the OMI HCHO offset is larger than for TROPOMI. But the reported bias are all negative for columns larger than 5e15 molec.cm-2. The conclusions of this study are therefore not completely in agreement with De Smedt et al. or with Vigouroux et al..

Line 594: I agree with the reasons for the poor HCHO correlation. Please add that they are also partly due to the low HCHO variability over the studied time and domain. A full year study would result in larger correlations.

High pollution case study: The added value of this section is not clear. As the paper is already long and detailed, I would suggest removing this section. If not removed, I then suggest to discuss the causes of higher NO2 columns and lower HCHO columns, such as surface temperature.

Common a priori sensitivity tests:

- It is not clear why the WRF-CMAQ profiles need to be scaled for the NASA OMI datasets, but not for the TROPOMI datasets.
- Figure 6: please explain in the legend what is the NASA OMI (scaled).
- Comparing Table 2 and Table 4, I can only see an improvement for TROPOMI at 0.05° resolution. The added value of this section is not clear, given the uncertainties in the WRF-CMAQ profiles.

Expected FNR information from TEMPO:

- What is the expected signal ratio of Tempo compared to TROPOMI for NO2 and HCHO? Can we expect an improvement of the HCHO noise?
- It would be interesting to show the diurnal variation of NO2 and HCHO from the TEMPO simulations.
- Line 777 and figure 7b and 7c. Not clear if retrieved OMI and TROPOMI are shown (line 777) or only synthetic TEMPO data averaged at the different spatial resolutions. It should be possible to show real data for OMI and TROPOMI in 2020.

Conclusion

Line 831: Please comment on the spatial and temporal resolution allowed by the OMI datasets. This is important for trend studies.

Line 860. The statements made on the new version of the NASA OMI HCHO product appear to be optimistic. The SNR of the retrievals is primarily determined by the SNR of the instrument. Please be more cautious, especially since no publication can support the statements.

Line 866-867: This does not seem so clear in the paper that "using the WRF-CMAQ-predicted a priori information, resulted in highly accurate retrievals of FNRs". All L2 products used in the study also results in median biases lower than 0.5 for FNRs.

Line 871-872: This sentence is misleading. The need for accurate shape factors is not only for OMI retrievals. It should be even more important for TROPOMI and TEMPO because of their finer spatial resolution.

References

Verhoelst, T., Compernolle, S., Pinardi, G., Lambert, J.-C., Eskes, H. J., Eichmann, K.-U., Fjæraa, A. M., Granville, J., Niemeijer, S., Cede, A., Tiefengraber, M., Hendrick, F., Pazmiño, A., Bais, A., Bazureau, A., Boersma, K. F., Bognar, K., Dehn, A., Donner, S., Elokhov, A., Gebetsberger, M., Goutail, F., Grutter de la Mora, M., Gruzdev, A., Gratsea, M., Hansen, G. H., Irie, H., Jepsen, N., Kanaya, Y., Karagkiozidis, D., Kivi, R., Kreher, K., Levelt, P. F., Liu, C., Müller, M., Navarro Comas, M., Piters, A. J. M., Pommereau, J.-P., Portafaix, T., Prados-Roman, C., Puentedura, O., Querel, R., Remmers, J., Richter, A., Rimmer, J., Rivera Cárdenas, C., Saavedra de Miguel, L., Sinyakov, V. P., Stremme, W., Strong, K., Van Roozendael, M., Veefkind, J. P., Wagner, T., Wittrock, F., Yela González, M., and Zehner, C.: Ground-based validation of the Copernicus Sentinel-5P TROPOMI NO2 measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global networks, Atmos. Meas. Tech., 14, 481–510, https://doi.org/10.5194/amt-14-481-2021, 2021.

Wang, P.; Holloway, T.; Bindl, M.; Harkey, M.; De Smedt, I. Ambient Formaldehyde over the United States from Ground-Based (AQS) and Satellite (OMI) Observations. Remote Sens. 2022, 14, 2191. https://doi.org/10.3390/rs14092191

Harkey, M.; Holloway, T.; Oberman, J.; Scotty, E. An evaluation of CMAQ NO2 using observed chemistry-meteorology correlations. J. Geophys. Res. Atmos. 2015, 120, 11775–11797. [CrossRef]

Souri, A. H., Johnson, M. S., Wolfe, G. M., Crawford, J. H., Fried, A., Wisthaler, A., Brune, W. H., Blake, D. R., Weinheimer, A. J., Verhoelst, T., Compernolle, S., Pinardi, G., Vigouroux, C., Langerock, B., Choi, S., Lamsal, L., Zhu, L., Sun, S., Cohen, R. C., Min, K.-E., Cho, C., Philip, S., Liu, X., and Chance, K.: Characterization of Errors in Satellite-based HCHO / NO2 Tropospheric Column Ratios with Respect to Chemistry, Column to PBL Translation, Spatial Representation, and Retrieval Uncertainties, Atmos. Chem. Phys. Discuss. [preprint], https://doi.org/10.5194/acp-2022-410, in review, 2022.