Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC2
https://doi.org/10.5194/amt-2022-264-RC2, 2022

Comment on amt-2022-264

Anonymous Referee #3

Referee comment on "Validation of Sentinel-5P TROPOMI tropospheric NO_2 products by comparison with NO_2 measurements from airborne imaging, ground-based stationary, and mobile car DOAS measurements during the S5P-VAL-DE-Ruhr campaign" by Kezia Lange et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2022-264-RC2, 2022

Referee comments to amt-2022-264 titled, 'Validation of Sentinel-5P TROPOMI tropospheric NO2 products by comparison with NO2 measurements from airborne imaging, ground-based stationary, and mobile car DOAS measurements during the S5P-VAL-DE-Ruhr campaign'

This manuscript uses DOAS data collected in September 2020 in a polluted region in western Germany from airborne, ground, and car-based instrumentation to validate the set of TROPOMI L2 NO2 products (both research and operational). The airborne datasets are first validated by the ground and car-based systems which then justifies the airborne use for validating TROPOMI. This paper fits the scope of AMT and will be valuable as a validation dataset for the TROPOMI NO2 product. However, before publishing, this manuscript requires some minor technical corrections/clarifications as detailed below but more reflection toward conclusions drawn about improvements in the S5P PAL product and the impact of clouds. Detailed comments below.

More significant comments:

- Most of the results in this work are too heavily based on the slope of the regression, which is not representing the complete behavior of the validation activity. Table A1 has at least median difference in % which actually in some cases contradicts the results of the slope (e.g., having a 21% higher column from TROPOMI as a median from S5P PAL V02.03.01.). Consider more in-depth analysis based on statistics other than slope for all intercomparisons.
- Some conclusions drawn in section 6 are either overgeneralized or not quite technically correct. These comments do not specify lines in the test but more so in general comments that need to be kept in mind when adding to and editing the analysis based on the suggestions below:
  - With the data presented, conclusions about the S5P PAL product are only stated as an improvement. This is an overgeneralized conclusion, and the authors should do

some more detailed analysis from other statistics. Some of this is already done with discussion of the lower lobe results but it is missing discussion on the higher lobe. Additionally, with the loss in precision, some users may find this result more detrimental than having a predictable low slope and this is not commented upon in the results, abstract, or conclusions.

  - The main reason concluded about the improved PAL product is due to the cloud correction. It is stated that all these changes are due to more ‘realistic’ cloud pressures or more ‘realistic’ cloud corrections but this more ‘realistic’ outcome is not demonstrated in this region. Therefore, these conclusions cannot be stated unless they are proven with the data available in that specific region (e.g., could look at imagery from satellites or other creative sources and reflect on what it should be in reality). In fact, removing the cloud correction all together (Figure 9b) shows that the massive improvement in slope is something else removed from the cloud correction as this is the best result in terms of conserving precision (correlation) and a higher slope.
  - Conclusions drawn about the cloud pressure in some cases seems to not be interpreted correctly as written. For example, discussions from line 555-563 talk about the low lobe. (1) It is stated that cloud pressures are too low, but looking at imagery online there seems to be zero clouds seen by VIIRS on this afternoon, so cloud pressures shouldn’t be low to start with. (2) Aerosols are also pointed at as a potential cause, but the sensitivity results in Figure A2 show that the impact of aerosols would not be large enough to create this bias in this lobe.
- Technical comments in relation to the AirMAP retrieval that need more justification or clarification.
  - It is said that the reference VCD in the troposphere for AirMAP is 1e15. One of the MAX-DOAS retrievals has a different value of 1.5E15 but they are referred to as similar. It is different by 50% rather than similar. Please clarify these difference or explain them.
  - Can the reference value be justified with any other data from this work? (i.e., What does the CAMS model say the tropospheric amount is?)
  - Could this reference assumption be the cause for a low offset between the car DOAS systems and the airborne dataset?
  - Can the authors justify why a 1km box profile used if CAMS analysis is available for these flights to provide a profile shape and what that assumption impact may be in the results? A 1 km box profile assumes that NO2 is well mixed through that 1km boundary layer which has been demonstrated as not the case with in situ measurements from aircraft near strong sources (which is the case here in many of these flights). (e.g., https://doi.org/10.1002/2015JD024203 and https://doi.org/10.1525/elementa.2020.00163). This paper also shows the impact of AMFs based on assuming a 1km box vs an urban profile atmos-meas-tech.net/3/475/2010/
  - Line 291: ‘Surfaces with different brightness introduce artefacts in the maps of NO2’. The impact isn’t necessarily an artifact at the SCD stage. This is caused by the brighter surface increasing sensitivity in the lower parts of the atmosphere meaning a higher slant column if NO2 is present (if there is not any or minimal NO2 then this spatial pattern will not show up in the slant column). It only becomes an artifact if the surface reflectivity assumption in the AMF calculation doesn’t account for this accurately.

Minor comments:

- When referring to the spatial resolution of TROPOMI as 3.5 km x 5.5 km, please specify that this is at nadir.
- Line 74. Mention what version Verhoelst et al. validated to be consistent with this analysis and the other mentioned publications.
- Line 94: the conclusion of 'low bias' is prematurely stated (before showing any results). Recommend just removing 'low' from the sentence.
  - Figure 2 is mentioned before Figure 1. Consider reordering figures to reflect this or consider combining Figures 1 and 2 for a more helpful side-by-side comparison.
- Line 159: capitalized Ozone Monitoring Instrument
- Line 179-181: The sentence about V02.04.01 should either clearly state that this analysis does not include this product or should be removed.
- Lines 173-177: The following sentence needs references: 'Other factors that could contribute to the underestimation are the low spatial resolution of the used a priori NO2 profiles from the TM5-MP global chemistry transport model, the use of the OMI LER climatology given on a grid of 0.5° x 0.5° for the AMF and cloud fraction retrieval in the NO2 fit window and the GOME-2 LER climatology used for the NIR-FRESCO cloud retrieval given on a grid of 0.25° x 0.25° measured at mid-morning.'
- Line 189: add the spatial resolution of the CAMS global analysis
- Line 198-199: The sentence referring to 15% increases needs a reference.
- Line 308: define quantitatively what polluted means for this statistic.
- Equation 5 seems to be the same as equation 4. Is it needed?
- Consider making a table of all the various information of the retrievals for the AirMAP, car, and stationary DOAS retrievals as the sections get repetitive and there are small differences in places that are hard to keep straight.
- Are there references for all the individual car or ground-based systems? If so, please add in the sections that describe them.
- The MAX-DOAS measurement truck is different from the rest in that it measures in the UV rather than the visible wavelengths of the other retrievals. Is it realistic for their AMFs to be the same as the other systems?
- Line 415-416. The SCD of the reference for this DOAS instrument seems quite large considering the statements that the AMFs for a zenith DOAS retrieval are about 1.3. Is this off by an order of magnitude or are the measurements just in a densely polluted area for the reference?
- Line 449-451. Is there a reference for the tropospheric NO2 product from Pandora that can be added to this section? This is the first publication I have seen use that product.
- Line 501: Is it +/- 1 hour or 30 minutes? The rest of the paper seems to reflect 30 minutes.
- Line 576: Before this line, it says that the criterion for comparison is the same as Judd et al. 2020 but at this location the authors should specify that this criterion (filtering for delta CS less/greater than 50 hpa) is the opposite of the filter applied by Judd et al. to avoid confusion. Bonus suggestion: it could be nice to have a comparison of what the results look like for the points with delta CS less than 50 hPa?
- Line 604-606: 'This behavior is different from the small impact that we observed for changing the a priori NO2 profile information from TM5 to CAMS for the OFFL V01.03.02 dataset'. The change seems to be on the same order of magnitude rather than different.
- Line 660. Saying cloud fractions are always lower than 0.14 contradicts from other examples in the text. (e.g., saying it was on average 0.21 in line 128).
- Line 667: it is stated that on average TROPOMI is lower than air map but there are no averages reported in the manuscript.