S5P/TROPOMI NO2 slant column retrieval: method, stability, uncertainties, and comparisons against OMI

by Jos van Geffen, et al.; amt-2019-471

Response to anonymous referee #2

RC The manuscript "S5P/TROPOMI NO2 slant column retrieval: method, stability, uncertainties, and comparisons against OMI" provides a detailed description of the DOAS slant column fitting for TROPOMI NO2 measurements and discusses important issues related to the method, instrument, and verification. The stability and uncertainty analysis is particularly interesting. The manuscript is well organized and written. I recommend publication after a few minor revisions.

We thank the referee for the kind words and for reading the manuscript in great details.

Changes to the manuscript are based on the comments and suggestions of three referees. In addition we have extended the data record of the paper by 3 months, which has lead to updating some figures and numbers, but has *not* affected the conclusions of the paper.

In the following we answer the specific comments of referee #2.

RC P4 L7 Is there a particular reason that the VIS band stops at 496 nm? While quite a few literatures numbered the TROPOMI VIS band until 500 nm.

Due to the so-called "spectral smile" the wavelength assignment across the swath to the spectral pixels is not constant. Regarding the wavelength range of Band 4 (given in nm):

- at the edges of the swath the wavelength ranges from 397.91 for row 0 and 397.79 for row 449 to 496.01 and 495.90, respectively
- at the centre of the swath, the start wavelength is highest for rows 222 and 223: 400.29, and ranges to 498.58

In other words: only the wavelength range 400–496 is available along all detector rows.

RC P4 L20 Please give the full name of ATBD.

Done.

RC P4 S2.1.2 Please give a short description of the saturation (perhaps also blooming) effect for TROPOMI, which is mentioned and analyzed in the result section and important for TROPOMI NO2 measurements.

A few lines are added to the end of Sect. 2.1.1 [P4,L19 in the revised text]:

> Over very bright radiance scenes, such as high clouds, the CCD detectors containing band 4 (Visible, e.g. used for NO_2 retrieval) and band 6 (NIR, e.g. used for cloud data retrieval) may show saturation effects (Ludewig et al., 2020), leading to lower-than-expected radiances for certain spectral (i.e. wavelength) pixels. In large saturation cases, charge blooming may occur: excess charge flows from saturated into neighbouring detector (ground) pixels in the row direction, resulting in higher-than-expected radiances for certain spectral pixels. Version 1.0.0 of the level-1b spectra contains flagging for saturation but not for blooming; version 2.0.0 will also have flagging for blooming (Ludewig et al., 2020).

RC P5 L12 What would be the implication of increasing the size towards the edge when comparing the OMI and TROPOMI measurements?

Given the difference in ground pixel size, OMI and TROPOMI measurements cannot be compared directly, not even when their orbits exactly overlap: OMI's ground pixels cover a larger geographical area than TROPOMI's, and this difference increases towards the swath edges. The only way to compare the two quantitatively is map their measurements on a common lat-lon grid, as done for Fig. 1 and 6.

RC P6 Table 1 I recommend to add the retrieval processor, namely TROPOMI (if no other name) vs QDOAS, before the date version. Please also add information of OMI/ OMNO2A (almost same as TROPOMI) in the title or as table footnote etc.

Processor names are added, as suggested. For details of the OMI/OMNO2A retrieval, a footnote refers to Sect. 3.4 (the text of which is somewhat extended, to cover both v1.2 and v2.0 of that processor), because otherwise the table would become unnecessarily complicated.

RC P9 L23 Please give rough numbers of the magnitude for OMI?

Actually the amplitude of the seasonal cycle in OMI's visible channel is comparable to TROPOMI's, as shown by Schenkeveld, et al. (2017) in their Fig. 34, as referee #1 pointed out correctly. The manuscript text has been adapted accordingly [P10,L3-5]:

> A similar seasonal variation of similar amplitude is seen in the wavelength calibration data of OMI's visible channel (Schenkeveld et al., 2017, Fig. 34). Both for TROPOMI and OMI this amplitude does not exceed scatter levels and is thus well within instrument requirements.

RC P11 Eq 9 The intensity offset term shall not be placed in the same parentheses with lamda.

Oops, misplaced parenthesis in Eq. (9) – sorry: has been corrected: $(I(\lambda) + P_{\mathrm{off}}(\lambda) \cdot S_{\mathrm{off}})$

RC P12 L10 Please give the full name of VCD.

You are right, this is the first time in the paper "VCD" is used – done.

RC P12 L15 What is the implication and how large is the impact of changing the LM solver to OE method with Gauss-Newton?

Test with both solvers in the wavelength calibration step, performed when setting up OE, have shown that both solvers essential give the same results and take up roughly the same cpu time. For the wavelength calibration OE has the advantage over LM that OE's solution is limited by the a-priori error value set on the wavelength shift (0.07 nm, i.e. 1/3-rd of the spectral sampling, thus guaranteeing that the shift will never be larger than the spectral sampling distance) while LM's solution is not restricted in any way. In addition adding the Ring effect in the wavelength calibration of the radiance using OE allows for the calibration to take place over a wider wavelength window than used in OMNO2A with LM. The text of Sect. 3.4 has been updated [P12,L23ff]:

> ...with the exception that χ^2 is minimised using a Levenberg-Marquardt (LM) solver, wavelength calibration is performed over part of the NO_2 fit window (409 – 428 nm), the 2005-average irradiance spectrum as reference, and an older ozone reference spectrum (van Geffen et al., 2015). Tests have shown that the LM and OE solvers essentially give the same fit results when used with the same settings.

And in Sect. 3.2.1 [P9, L19ff]:

> For the $I(\lambda)$ calibration a 2nd order polynomial as well as a term representing the Ring effect are included: the model function used for the radiance wavelength calibration is a modified version of Eq. (2); including the Ring effect allows for a wavelength calibration to be performed across the full fit window. For the $E_0(\lambda)$ calibration the Ring term is obviously excluded. The a-priori error of the wavelength shift is set to 0.07 nm, 1/3-rd of the spectral sampling in the NO_2 wavelength range, so as to ensure that w_s will not exceed the spectral sampling distance.

RC P14 L17 & P15 L20 Following the previous question, is the difference between TROPOMI and OMNO2A (Fig 3c and 3f) mainly resulted from the difference in the wavelength calibration window (I assume not the mathematic solver)?

As mentioned at the previous question, the solver is not likely to be the reason for the differences seen in the panels. The difference in the wavelength calibration window indeed has a large impact on the results; this has been added to the text [P14, L27-30]:

> Differences in results of the OMNO2A and TROPOMI processor are likely mainly due to differences in the wavelength calibration: TROPOMI's radiance wavelength calibration includes a correction for the Ring effect, which allows the use of a larger calibration window (in casu the NO_2 fit window; viz. Sect. 3.2.1), while OMNO2A's calibration window is necessarily limited (viz. Sect. 3.4).

Other differences between the two are minor (e.g. the sampling of the cross-sections, the precise selection of first and last wavelengths in the fit window); providing this much detail in the paper is not necessary,

RC P15 L6 Previous introduction has mentioned that TROPOMI measurements suffering from saturation are filtered out. But the residual saturation is still affecting the retrieval (even not strongly) in Fig 3d-e. Is there any recommendation of further removing the effect during retrieval or data using?

Spectral pixels flagged in the L1B spectra as saturated are removed from the fit, and in v1.2-v1.3 of the TROPOMI processor ground pixels with more than 3 such flags are not processed (the 405–465 fit window has about 305 spectral pixels).
Ground pixels with 1 to 3 flags will give more or less normal-looking SCDs but are likely to give markedly higher SCD error levels. In addition, ground pixels next to saturation cases may suffer from blooming, which possibly affects the SCDs and certainly will increase the SCD error.
As mentioned in Sect. 3.2 ground pixels with large SCD error values are flagged in the final data product as unreliable. Filtering on the SCD error was *not* done for the data in Fig. 3.
In the forthcoming v2.1 processing, the L1B flagging will be improved and the NO2 algorithm will use an outlier removal, leading to more spectral pixels being removed from the fit. As a result of this, we can allow for more spectra to enter the SCD retrieval and we can expect cases around saturation effects to give us more reliable SCDs, also for previously discarded ground pixels, albeit with somewhat elevated SCD error levels.

RC P25 L3 Please give the full name of VRS (introduced already in Sect 3).

Done.

RC P26 L13 Even the air pollution is reduced in China, it is still very possible to see columns of a few 10e16 molec/cm2 (not optically thin), particularly in Winter. With this pollution level, the boundary AMF will also show spectral features, and this nonlinearity will contribute a few percent difference to the retrieval during pollution episode. Please rewrite the statements and perhaps also provide the example for Winter for China (at the moment is July for Africa).

The example given refered to the orbit with the highest number of high NO2 values in any of the orbits from that month, July 2018, which turns out to be over Africa.
In Jan. 2019 the highest GCDs are indeed found over China (those over Africa are much lower than in July). The top 5 in highest number of ground pixels with GCD $>$ 300 (using unit μmol/m^2 for GCD and SCD error here for brevity) is:

- ◦ 1609 for orbit 06580 of 20 Jan.: highest GCD is 512 ± 14; highest SCD error is 18
- ◦ 1052 for orbit 06566 of 19 Jan.: highest GCD is 620 ± 15; highest SCD error is 15
- ◦ 1045 for orbit 06495 of 14 Jan.: highest GCD is 581 ± 12; highest SCD error is 17
- ◦ 577 for orbit 06637 of 24 Jan.: highest GCD is 701 ± 16; highest SCD error is 16
- ◦ 519 for orbit 06466 of 12 Jan.: highest GCD is 549 ± 16; highest SCD error is 16

None of these values seems to be exceptionally high and reason to worry.
The text has been adapted accordingly, replacing the July 2018 example: Jan. 2019 is indeed a better example, as the discussion in this section is about high NO_2 in China) [P27,L29ff]:

> ... it is currently unlikely to encounter NO_2 concentrations that are not optically thin in the TROPOMI data, except in a few individual pixels.
> NO_2 concentration over China are highest in winter. In Jan. 2019, for example, the highest GCD found over China is 701 ± 16 μmol/m^2 in orbit 06637 (24 Jan.), which has 577 pixels (0.05% of the 1204367 pixels with a successful retrieval) with a GCD exceeding 300 μmol/m^2; 73 pixels have a GCD values exceeding 400 μmol/m^2. Orbit 06580 (20 Jan.) has in that month the largest number of pixels with a GCD exceeding 300 μmol/m^2, namely 1609, with a peak value 512 ± 14 μmol/m^2; 256 pixels have a GCD values exceeding 400 μmol/m^2.

RC P28 L3 Do you mean case QDOAS case 4?

Corrected: it's case 3 [P29,L20]; the "case 3" mentioned a few lines down [P29,L25] has to be "case 2" (leftovers from earlier manuscript version; sorry).