Response to Referee Henk Eskes

We thank the reviewer Henk Eskes for the careful reading and the useful comments. Below we give the reviewer’s comment, our response, and the changed text in the manuscript. The page and line numbers refer to the mark-up version of the manuscript.

General remarks:

1. The reader may find the title and abstract confusing. It suggests that the authors are improving the operational TROPOMI retrieval product, while actually the paper discusses and upgrade of the retrieval product implemented at DLR. This should be made more clear at the top of the abstract and in the title. I suggest the authors make clear which retrieval product they refer to, for instance by calling it the DLR TROPOMI NO₂ European product. Maybe good to add the word "scientific" as opposed to the "operational" retrieval.

We have updated the title to “An improved TROPOMI tropospheric NO₂ research product over Europe” and referred the data to “the DLR TROPOMI NO₂ European product” throughout the paper.

2. The performance of the POLYPHEMUS model is not easy to judge. The resolution of this model is not very high, 0.2x0.3 degree, while TROPOMI has a resolution of 0.05 degree. The CAMS models for instance run at 0.1x0.1 degree. This resolution is in between the TM5-MP and TROPOMI. Would higher resolution produce a major further improvement? I would like to see a comparison between POLYPHEMUS and TROPOMI as extra figure. The POLYPHEMUS model seems to produce very low free tropospheric NO₂ concentrations. Is this expected to impact the retrieval?

The 0.2x0.3 degree is selected for POLYPHEMUS/DLR due to the feasibility in terms of runtime and affordability in terms of computing resources. We have analysed the effect of model resolution using a priori NO₂ profiles from LOTOS-EUROS, which constitutes one of the state-of-the-art atmospheric chemistry models used by the regional Copernicus Atmosphere Monitoring Service (CAMS, http://www.copernicus-atmosphere.eu). Figure R1 shows the differences in the tropospheric NO₂ columns retrieved by altering the model resolution for LOTOS-EUROS (0.1x0.1 degree - 0.2x0.3 degree) in February 2019. The increase of the spatial resolution from 0.2x0.3 to 0.1x0.1 degree improves the tropospheric NO₂ columns moderately by up to 5×10¹⁴ molec/cm² or 11% for polluted regions.
Figure R1. Differences in the tropospheric NO$_2$ columns retrieved by altering the model resolutions (0.1x0.1 degree - 0.2x0.3 degree) for LOTOS-EUROS a priori NO$_2$ profiles in February 2019.

We have included a new Fig. 11 in page 20 comparing the tropospheric NO$_2$ columns from TROPOMI measurements and POLYPHEMUS/DLR simulations (using the satellite averaging kernel) and included the description in pages 19 as “Figure 11 compares the tropospheric NO$_2$ columns from TROPOMI measurements and POLYPHEMUS/DLR simulations. The satellite averaging kernel, which describes the vertical sensitivity of measurements of NO$_2$ concentrations, is applied to reduce the systematic biases caused by unrealistic a priori profile information. From Fig. 11, POLYPHEMUS/DLR NO$_2$ columns are generally higher than satellite measurements, which can be partly related to the use of the TNO-MACC emission dataset (Denier van der Gon et al., 2010). An update of the POLYPHEMUS/DLR model using the more recent TNO-MACC II emission (Kuenen et al., 2014) is planned for the near future. We note here that the profile shape is of far more importance than the column bias for the interpretation of satellite retrievals.”
Despite the use of outdated emissions, the POLYPHEMUS/DLR is in general reliable, as indicated in a validation exercise in Fig. R2 comparing the near-surface NO$_2$ concentrations from POLYPHEMUS/DLR and ground-based EEA air quality monitoring stations (https://discomap.eea.europa.eu/map/fme/AirQualityExport.htm). The correlation coefficients are generally higher than 0.6 for polluted hot spots, and the biases are generally lower than 20 uq/m$^3$ or 38%.
Figure R2. Comparisons of near-surface NO$_2$ concentrations [µg/m$^3$] from POLYPHEMUS/DLR simulations and ground-based air quality monitoring stations (POLYPHEMUS/DLR results – ground-based observations) for Europe in 2018.
Figure R3. A priori NO\textsubscript{2} profiles from the chemistry transport models LOTOS-EUROS, POLYPHEMUS/DLR, and TM5-MP over Hamburg in Germany (53.55°N, 9.99°E) on 1 February 2019. The calculated clear-sky tropospheric AMF is given in the bracket next to each label in the legend. Normalized profiles (to the lowest values) are also shown on a logarithm scale.

We have compared a priori NO\textsubscript{2} profiles from LOTOS-EUROS, POLYPHEMUS/DLR, and TM5-MP over Hamburg in Fig. R3. POLYPHEMUS/DLR shows the largest surface layer NO\textsubscript{2} concentration (Fig. R3 left panel) and the steepest profile shape (Fig. R3 right panel), which yields the smallest tropospheric AMF. LOTOS-EUROS shows lower free tropospheric concentration and larger surface concentration than TM5-MP, and thus the tropospheric AMF is lower by 9%.

Figure R4. Differences in the tropospheric NO\textsubscript{2} columns retrieved using the POLYPHEMUS/DLR and LOTOS-EUROS a priori NO\textsubscript{2} profiles over Europe in February and August 2019.

Figure R4 compares the tropospheric NO\textsubscript{2} columns retrieved using POLYPHEMUS/DLR and LOTOS-EUROS a priori NO\textsubscript{2} profiles, where the use of LOTOS-EUROS reduces the tropospheric NO\textsubscript{2} columns by up to ~1\times10^{15} molec/cm\textsuperscript{2}, e.g. over the ocean partly due to difference in the shipping emissions.
3. It would be of interest to extend the comparisons with the operational TROPOMI product and also list differences with other (regional) retrievals like the TROPOMI-POMINO approach for Asia. Maybe in the form of a table listing the choices for albedo, cloud, stratosphere, a-priori for several retrievals. For the validation, Table 4, it would be nice if also the operational retrieval could be added (now there is only the comparison with the old DLR retrieval in brackets).

We have extended the comparison with different retrievals in Table 1 in page 9.

**Table 1. Overview of tropospheric NO$_2$ column retrievals.** See Table 2 for details of the chemistry transport models used to obtain the a priori NO$_2$ profiles for the DLR improved retrieval in this work.

<table>
<thead>
<tr>
<th>Reference</th>
<th>DLR reference retrieval</th>
<th>DLR improved algorithm</th>
<th>KNMI operational product</th>
<th>POMINO-TROPOMI product</th>
<th>ECCC product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
<td>Global</td>
<td>Europe</td>
<td>Global</td>
<td>China</td>
<td>Canada</td>
</tr>
<tr>
<td>Stratospheric correction</td>
<td>STREAM</td>
<td>DSTREAM</td>
<td>Data assimilation</td>
<td>Same as KNMI</td>
<td>Same as KNMI</td>
</tr>
<tr>
<td>Surface albedo</td>
<td>OMI LER climatology</td>
<td>TROPOMI GE_LER data</td>
<td>OMI LER climatology</td>
<td>MODIS BRDF</td>
<td>MODIS albedo</td>
</tr>
<tr>
<td>A priori NO$_2$ profile</td>
<td>TM5-MP</td>
<td>POLYPHEMUS/DLR</td>
<td>TM5-MP</td>
<td>GLOS-Chem</td>
<td>GEM-MACH</td>
</tr>
<tr>
<td>Cloud parameter</td>
<td>OCRA/ROCINN_CRB version 1.x</td>
<td>OCRA/ROCINN_CAL FESCO-S version 2.1</td>
<td>Recalculated cloud fraction with aerosol correction</td>
<td>Same as KNMI</td>
<td></td>
</tr>
<tr>
<td>Aerosol treatment</td>
<td>Implicit correction</td>
<td>Implicit correction</td>
<td>Implicit correction</td>
<td>Explicit correction</td>
<td>Explicit correction</td>
</tr>
</tbody>
</table>

We have included the validation results for the operational product in Tables 5 and 6 in pages 31-32.

**Table 5.** The Pearson correlation coefficient $R$, as well as the slope $S$ and intercept $I$ (in $1 \times 10^{15}$ molec/cm$^2$) obtained with the robust Theil–Sen estimator for the monthly TROPOMI tropospheric NO$_2$ product compared to MAX-DOAS data. Stations are ordered by increasing mean difference. Values for the DLR improved algorithm (DLRimp) are given, and the values for the DLR reference algorithm (DLRref) and the KNMI operational product (KNMIop) are reported for comparison.
Table 6. Similar as Table 5 but for the mean difference (MD, SAT-GB in $1 \times 10^{15}$ molec/cm$^2$), standard deviation (STD, in $1 \times 10^{15}$ molec/cm$^2$), and relative difference (RD, in %) for the monthly TROPOMI tropospheric NO$_2$ product compared to MAX-DOAS data.

<table>
<thead>
<tr>
<th>Station</th>
<th>DLRimp</th>
<th>DLRef</th>
<th>KNMlop</th>
<th>DLRimp</th>
<th>DLRef</th>
<th>KNMlop</th>
<th>DLRimp</th>
<th>DLRef</th>
<th>KNMlop</th>
</tr>
</thead>
<tbody>
<tr>
<td>Athens</td>
<td>0.85</td>
<td>0.87</td>
<td>0.81</td>
<td>0.65</td>
<td>0.42</td>
<td>0.45</td>
<td>0.33</td>
<td>0.25</td>
<td>0.33</td>
</tr>
<tr>
<td>De Bilt</td>
<td>0.63</td>
<td>0.72</td>
<td>0.80</td>
<td>0.42</td>
<td>0.25</td>
<td>0.77</td>
<td>1.69</td>
<td>1.31</td>
<td>-0.57</td>
</tr>
<tr>
<td>Thessaloniki_ciri</td>
<td>0.91</td>
<td>0.95</td>
<td>0.91</td>
<td>0.51</td>
<td>0.45</td>
<td>0.26</td>
<td>-0.80</td>
<td>0.14</td>
<td>1.23</td>
</tr>
<tr>
<td>Thessaloniki_lap</td>
<td>0.79</td>
<td>0.86</td>
<td>0.90</td>
<td>0.35</td>
<td>0.30</td>
<td>0.37</td>
<td>1.97</td>
<td>1.09</td>
<td>0.93</td>
</tr>
<tr>
<td>Bremen</td>
<td>0.85</td>
<td>0.71</td>
<td>0.70</td>
<td>0.58</td>
<td>0.44</td>
<td>0.40</td>
<td>-0.24</td>
<td>-0.20</td>
<td>0.79</td>
</tr>
<tr>
<td>Uccle</td>
<td>0.85</td>
<td>0.81</td>
<td>0.44</td>
<td>0.70</td>
<td>0.42</td>
<td>0.27</td>
<td>-0.52</td>
<td>-0.24</td>
<td>2.26</td>
</tr>
<tr>
<td>Cabaauw</td>
<td>0.75</td>
<td>0.67</td>
<td>0.43</td>
<td>0.41</td>
<td>0.21</td>
<td>0.33</td>
<td>1.08</td>
<td>1.24</td>
<td>1.98</td>
</tr>
<tr>
<td>Munich</td>
<td>0.57</td>
<td>0.72</td>
<td>0.63</td>
<td>0.35</td>
<td>0.39</td>
<td>0.42</td>
<td>1.44</td>
<td>0.29</td>
<td>1.35</td>
</tr>
<tr>
<td>Mainz</td>
<td>0.85</td>
<td>0.87</td>
<td>0.81</td>
<td>0.39</td>
<td>0.28</td>
<td>0.50</td>
<td>1.65</td>
<td>0.82</td>
<td>0.02</td>
</tr>
</tbody>
</table>

4. The word "improved" is used many times in the paper, not only for the final NO$_2$ result, but also for the retrieval aspects like cloud parameters, albedo, a-priori. To my opinion one should be a bit careful with this term. One can claim something is improved if there is a better match with independent observations, or if obvious biases are removed. This is not always clear. The main conclusion of the paper is an increase of the NO$_2$ columns which better agrees with MAXDOAS (smaller bias).

We have rephrased the word “improved” e.g. to “new” in line 24 and 500, “updated” in line 405, and “corrected” in line 299, “higher” in line 327.

5. The uncertainty estimates for the individual aspects of the retrieval: cloud fraction/pressure, albedo, profiles could be discussed in more detail. It seems the authors make use of numbers from De Smedt et al, instead of deriving typical uncertainties for GE-LER, ROCINN, POLYPHEMUS. It would be good to have some rough estimates of uncertainty reductions for the individual terms in the new versus old retrieval.
We have updated the uncertainty estimates in pages 26-27 as “The estimated parameter uncertainties considered in the AMF uncertainty budget include 0.0016 for GE_LER albedo (Loyola et al., 2020b), 75 hPa for POLYPHEMUS/DLR profile height (estimated from the profile height standard deviation), 0.05 for the OCRA cloud fraction, and 50 hPa ROCINN_CAL cloud pressure (Loyola et al., 2020a).”

We have included a Table 3 in page 25 showing the individual changes of each step (as suggested by the other reviewer) and the uncertainty reduction.

Table 3. Main settings for the step-by-step improvements and results of the tropospheric NO2 retrievals for the different steps of the updates for Munich (48.15N, 11.57E). The tropospheric NO2 columns (VCDtrop) are given in absolute values (molec/cm$^2$), and the percentage numbers in the brackets are changes relative to the reference case. The uncertainties in the tropospheric NO2 columns (VCDtropErr) are given in relative values.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Sect. 2.3 (Reference)</th>
<th>Sect. 4.1</th>
<th>Sect. 4.2</th>
<th>Sect. 4.3.1</th>
<th>Sect. 4.3.2 (Improved)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface albedo</td>
<td>OMI LER climatology</td>
<td>TROPOMI</td>
<td>TROPOMI</td>
<td>TROPOMI</td>
<td>TROPOMI</td>
</tr>
<tr>
<td>A priori NO2 profile</td>
<td>TM5-MP</td>
<td>GE_LER data</td>
<td>TM5-MP</td>
<td>POLYPHEMUS/DLR</td>
<td>POLYPHEMUS/DLR</td>
</tr>
<tr>
<td>Cloud parameter</td>
<td>OCRA/ROCINN_CRB</td>
<td>OCRA/ROCINN_CRB</td>
<td>OCRA/ROCINN_CRB</td>
<td>OCRA/ROCINN_CRB OCRA/ROCINN_CRB OCRA/ROCINN_CAL</td>
<td></td>
</tr>
<tr>
<td>VCDtrop</td>
<td>1.02 x 10$^{16}$</td>
<td>1.07 x 10$^{16}$</td>
<td>1.51 x 10$^{16}$</td>
<td>1.57 x 10$^{16}$</td>
<td>1.43 x 10$^{16}$</td>
</tr>
<tr>
<td>(Feb. 2019)</td>
<td>(+4.9%)</td>
<td>(+48.0%)</td>
<td>(+53.9%)</td>
<td>(+40.2%)</td>
<td></td>
</tr>
<tr>
<td>VCDtropErr</td>
<td>65%</td>
<td>62%</td>
<td>57%</td>
<td>56%</td>
<td>53%</td>
</tr>
<tr>
<td>(Feb. 2019)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VCDtrop</td>
<td>3.86 x 10$^{15}$</td>
<td>4.03 x 10$^{15}$</td>
<td>5.02 x 10$^{16}$</td>
<td>5.13 x 10$^{16}$</td>
<td>4.50 x 10$^{16}$</td>
</tr>
<tr>
<td>(Aug. 2019)</td>
<td>(+4.4%)</td>
<td>(+30.1%)</td>
<td>(+32.9%)</td>
<td>(+18.3%)</td>
<td></td>
</tr>
<tr>
<td>VCDtropErr</td>
<td>41%</td>
<td>41%</td>
<td>37%</td>
<td>36%</td>
<td>35%</td>
</tr>
<tr>
<td>(Aug. 2019)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. In our experience with the operational product the impact of the free troposphere is not negligible. This may be discussed in more detail.

Comparing to the operational data assimilation technique that applies actual meteorological fields, one general limitation of modified reference sector methods is the possible misinterpretation of tropospheric background column over clean regions as stratospheric column. To overcome this issue, (D)STREAM makes use of cloudy pixels with medium altitudes, which directly reflect the actual stratospheric column as the tropospheric column is mostly shielded. However, the broad-scale free tropospheric diffuse NO2 cannot be fully distinguished from stratospheric NO2 by (D)STREAM. We have included more discussions in page 7 regarding the tropospheric background as “Due to the use of cloudy pixels, which directly reflect the actual stratospheric column as the tropospheric column is mostly shielded,
STREAM reduces the bias caused by the free-tropospheric contamination or tropospheric background in the reference region, in comparison with other modified reference sector methods (Beirle et al., 2016). However, the broad-scale free tropospheric diffuse NO$_2$ might not be fully separated from the stratospheric NO$_2$, which is a general limitation of the modified reference sector methods.”

7. The paper mentions that the datasets are available upon request. A brief description of the product would be useful, e.g. are the input fields like GE-LER and OCRA/ROCINN cloud parameters included? Are averaging kernels included?

We have included an introduction of the product file in Data Availability as “The DLR TROPOMI NO$_2$ product are published as HDF version 5 files. For each ground pixel, the TROPOMI data product provides the retrieval results (e.g., the slant column, stratospheric column, and tropospheric column of NO$_2$), input information (such as the GE_LER surface albedo, POLYPHEMUS/DLR a priori NO$_2$ profiles, and OCRA/ROCINN cloud parameters), uncertainty estimate, processing quality flag, and averaging kernel.”

8. And finally, it would be nice if the authors can provide some recommendations for the future development of the (operational and scientific) TROPOMI NO$_2$ retrievals in the conclusions section!

We have extended the future works in page 35 regarding the POLYPHEMUS/DLR a priori profile, GE_LER surface albedo, and MA-DOAS validation.

Detailed remarks:

Title: Improved compared to what? The retrieval and reference should be clearly defined.

We have updated the title and texts accordingly. Please refer to major comment 1.

l12: Uncertainty strat column = 3.5 10^14 "for polluted conditions". This is a bit strange: the stratosphere does not have polluted conditions.

We have updated the expression in line 12 as “3.5×10$^{14}$ molec/cm$^2$ in case of significant tropospheric sources”.

l25: Decrease from 55 to 34 %: what is the 55% reference?

We have clarified the reference in line 26 as “The implementation of the algorithm improvements leads to a decrease of the relative difference from -55.3% to -34.7% on average in comparison with the DLR reference retrieval.”
At the end of section 6 it is shown that the comparison with MAXDOAS is affected significantly by the differences in the sensitivity profiles for MAXDOAS compared to TROPOMI. With kernel smoothing the remaining difference is about -20%. It would be interesting to mention this.

We have included the texts in line 26 as “When the satellite averaging kernels are used to remove the contribution of a priori profile shape, the relative difference decrease further to ~-20%.”

There is no reference to the data assimilation approach. It would be useful to explain a bit more this alternative approach.

We have included references in line 67 as “(Eskes et al., 2003; Dirksen et al., 2011)” and extended the description in line 69 as “Advantages of the data assimilation approach include a realistic error estimation and the capture of small-scale dynamical and chemical variability of stratospheric NO2.”

Improvements compared to what? To the previous DLR algorithm or compared to the operational algorithm, or both? Even at the end of the introduction it is still somewhat unclear which two retrievals are compared.

We have updated the sentence in 125 as “a number of improvements to the tropospheric NO2 retrieval over Europe are introduced for the DLR product.”

An intensity offset correction is used, while this is not done in the operational retrieval. Would be good to have a brief discussion of the impact of this intensity offset term. How do the slant columns compare with the operational algorithm?

Though the precise physical origin of such an intensity offset has not been fully discussed, the intensity offset correction is included in the DLR product to reduce the fitting residues. Figure R5. illustrates the effect of applying a linear intensity offset correction, which decreases the NO2 columns by up to 11% and the fitting residues by up to 23% mainly over the cloud-free ocean, likely due to the compensation of inelastic vibrational Raman scattering in water bodies (Vountas et al., 2003).
Figure R5. Difference in NO₂ columns (slant columns scaled by geometric AMFs) and fitting residues (retrieval root mean square, RMS) estimated with and without a linear intensity offset correction (with-without) on 5 February 2019.

Figure R6 compares the DLR and KNMI NO₂ slant columns, with similar spatial distribution to Fig. R5 mainly reflecting the differences of the intensity offset correction. The results are in agreement with van Geffen et al., 2020. We have referred to van Geffen et al., 2020 for further discussion in line 158.

Figure R6. Difference in NO₂ columns (slant columns scaled by geometric AMFs) from the DLR improved algorithm and KNMI operational product (DLR-KNMI) on 5 February 2019.

Sec 2.2 STREAM: How does STREAM distinguish the stratospheric background from a free tropospheric background? Please add some discussion on the free troposphere, which is supposed to be included in the tropospheric NO₂ column.

Please refer to major comment 6.

l182: "average bias of 1e13 molec/cm² with respect to the ground-based zenith-scattered light differential optical absorption spectroscopy (ZSL-DOAS) measurements" This is a very small number. Please provide the uncertainty range on this comparison. STREAM produces a somewhat larger column than the assimilation approach because it does not distinguish stratosphere from free troposphere.

We have included the uncertainty in line 193 as “an average bias of 1±8×10¹³ molec/cm²”.

Sec 3.1, description of DSTREAM: I am wondering if there is any interference between DSTREAM and the destriping? The directional part removes east-west
Biases. Does the destriping do something similar? Is destriping done before STREAM is calculated, or after?

The interference between DSTREAM and the de-striping is expected to be small.

We first have corrected the mistake in the label of Fig. 3 in page 11 as “The de-striping is not implemented here.” The correct figure without de-striping is shown in Fig. R7, which shows similar impact of local time changes across the orbit as the results with de-striping.

Figure R7. Similar as Fig. 3 in the manuscript but for measurements without de-striping correction.
Figure R8. NO$_2$ slant columns (scaled by geometric AMFs) averaged for clean regions between 20°S and 20°N with and without de-striping correction for orbit 6748 on 1 February 2019.

The de-striping correction is implemented before the (D)STREAM. Figure R8 shows the effect of applying the empirical de-striping correction. According to the box-car averaging method (Boersma et al., 2011), the total NO$_2$ columns from 60 adjacent viewing angles are averaged for every TROPOMI along-track array. In comparison to the operational de-striping relying on CTM, only relatively small-scale variation is removed, and the larger-scale directional part may not be captured.

I 257: Why is this latitude weighting introduced? Even though the diurnal cycle effect is smaller at the equator, I would assume it could still be modelled with STREAM? For instance, average slopes with viewing angle could be accommodated as a function of latitude also near the equator.
Figure R9. NO2 columns (slant columns scaled by geometric AMFs) for the eastern, central, and western segments of the orbit swath.

The main consideration of the latitude weighting is to reduce interpolation errors for low latitudes with less orbital overlap, as shown in Fig. R9 and explained in line 267.

I273: Can the difference with the model (3.5e14) be considered a true uncertainty estimate? Or is it a lower limit, e.g. because of the finite model resolution?

3.5e14 can be regarded as a lower limit due to the use of 0.75 degree resolution for the synthetic data.

I 290: Does the GE-LER approach also provide an uncertainty estimate?

In the near future an independent Neural Network will be trained and implemented for error estimation. We have updated the conclusion accordingly.

I290: How sensitive is the GE-LER to L1B calibration errors? How has the GE-LER been validated, e.g has it been compared with MODIS-based BRDF results? Please
add some more info for the reader to judge the performance of the GE-LER approach.

Large L1B calibration errors will significantly affect the GE_LER retrieval. For UV fitting window, like O₃, SO₂, and HCHO, the GE_LER data has been validated with representative products, such as OMI and GOME-2 LER climatologies. A first comparison of NO₂ GE_LER data has shown a general good agreement with S5P/KNMI and GOME-2 LER climatology, particularly for oceans and the polluted continents, as shown in Fig. R10 for example. We are currently composing a paper about an extensive validation of the GE_LER products.

![Figure R10. S5P/KNMI LER climatology at 460 nm and GE_LER retrieved from TROPOMI data (February 2018-2020) in the NO₂ window.](image)

I315: "The surface LER values from GE_LER are lower than the climatological OMI values by 0.03 on average" Is this a statement for February or August? It seems that the average differences in August are smaller than in February (by looking at the figures).

We have corrected the statement in line 331 as “The mean differences between the surface LER values from GE_LER and OMI climatological values are lower than 0.03.”

Fig. 7: Please comment on the very low albedo values over the Mediterranean compared to the OMI LER.

The very low albedo values could be due to errors in the input parameters, such as SCD or the constant term of DOAS polynomial. As compared to the OMI LER based on years of data, the GE_LER used in this study considered a short time period of data, as explained in line 334. We have included the expression in line 335 as “which makes GE_LER more likely affected by aerosol contamination or outliers in the input data”. Further analyses are required to explain the low values.
Sec 4.2: Please provide the definition of the POLYPHEMUS European domain (lat-lon domain boundaries).

We have included the domain in line 346 as “for Europe (34.0N-60.4N, 12.0W-40.2E) in this study”.

I339: "European TNO-MACC". Please specify the version and reference year.

We have included the version and reference year in line 356 as “European TNO-MACC (from 2011 with base year 2005)”.

I349: "The profile shape from POLYPHEMUS/DLR agree better with the MAX-DOAS measurements". The profiling capabilities of the MAXDOAS instruments are limited and often also quite uncertain. Part of this profile shape is a-priori defined. So, does it make sense to compare these profiles?

We have removed the comparison of profiles and updated the expression in line 367 as “In comparison with the tropospheric AMF calculated using the MAX-DOAS NO$_2$ profile, the bias reduces from 0.32 (48.5%) for TM5-MP to -0.04 (-6.0%) for POLYPHEMUS/DLR.”

Fig.10: It would be nice to also show a direct comparison of POLYPHEMUS tropospheric NO$_2$ column against TROPOMI (using the averaging kernels), e.g. also for February and August. Would it be possible to add these plots?

We have included the plot in pages 19-20. Please refer to major comment 2.

Sec. 4.2: Has POLYPHEMUS been compared to the CAMS regional modelling results? If so, please provide a brief summary of these comparisons. Please provide more detail on how the model has been validated in general, e.g. against observations of NO$_2$.

Please refer to the major comment 2 for the comparison between POLYPHEMUS/DLR and EUROS-LOTOS as well as the validation of POLYPHEMUS/DLR.

Sec. 4.2: Fig.9 indicates that POLYPHEMUS has basically very low concentrations of NO$_2$ in the free troposphere. Even though the free tropospheric column is small, this may have quite a significant impact on the AMF, and may in part be a reason why we see only red colors in Fig. 10. With larger free tropospheric NOx concentrations I would expect more blue colors especially in the more remote areas away from the main pollution hotspots. Please comment on this.
From Fig 9 left panel, POLYPHEMUS/DLR shows larger free tropospheric NO$_2$ concentration. Please refer to major comment 2 for further discussions.

**Fig.11: "comparisons with version 1.x " Please be more explicit. Do you show cf(2.1)-cf(1.x) or cf(1.x)-cf(2.1)? What is x in "1.x" in this case?**

We have updated the plot in page 21 with text “\textit{(version 2.1 \textendash\ version 1.x)}”. We have added the description “\textit{(1.0 and 1.1)}” in line 402.

**Fig.12: Why do you show a normalised CF instead of a mean CF?**

We have changed the Fig. 13 in page 22 to show the mean CF instead of a normalized one.

**Is Fig.12 consistent with Figs. 11 and 13 and the text? Fig.12 indicates smaller CF at the right side of the orbit in v2.1, while e.g. line 389 mentions an increase of the cloud fraction?**

We have updated the plot. Please refer to the previous reply.

**Fig.13: Again: do you show cf(2.1)-cf(1.x) or cf(1.x)-cf(2.1)?**

We have updated the Fig.14 in page 22 with “\textit{(version 2.1 \textendash\ version 1.x)}”.

**Fig.13: Is there any evidence that the new v2.1 cloud pressures are better than the old ones? Please discuss this in more detail.**

The biggest change for ROCINN between v1.x and version 2.1 is the replacement of the coarse MERIS surface albedo climatology used in v1.x with an actual surface albedo retrieval with the matching TROPOMI scene information in real time (GE_LER retrieval). Hence, shortcomings of a climatology (e.g. spatial resolution, short term temporal variations like snow/ice coverage, cloud contamination) are largely reduced by the GE_LER retrieval which is based on TROPOMI data themselves and performed at the same time with the TROPOMI measurements, hence it is more representative of the actual scene conditions than a climatology can be. Although the ROCINN retrieval scheme for cloud height and optical thickness / cloud albedo itself did not change, an improvement for those parameters is achieved indirectly through the improved surface albedo, particularly over scenes with rapidly varying surface conditions like snow/ice.

**Fig.14: How can I understand the reduction in NO$_2$ over the Po Valley? The cloud pressure does not seem to change much here.**
The reduction in NO\textsubscript{2} over the Po Valley is mainly due to a new check in the version 2.1 processor: if the cloud fraction is smaller than 0.1 and the height difference between the retrieved cloud height and the surface height is lower than 100 m, the scene is assumed to be cloud-free. We have included the explanation in page 23 as “The NO\textsubscript{2} reductions over the Po Valley are related to an additional check implemented in the version 2.1 processor: the pixel is assumed to be cloud-free for the almost clear-sky condition (cloud fraction < 0.1) with the retrieved cloud height very close to the surface height (difference < 100 m). This correction improves the data yield of the TROPOMI cloud products compared with other satellite cloud products; the performance of this correction under different surface conditions (dark, bright, snow, ice) or under presence of different types of low-level aerosols (fog, smoke, dust, ash) is under investigation.”

I407: Could you also provide the average % difference in the AMF in winter and summer for CAL vs CBR?

We have included the relative difference as “by more than 1×10^{15} molec/cm\textsuperscript{2} (18\%) for polluted regions in winter” in line 441 and “less than 5×10^{14} molec/cm\textsuperscript{2} (10\%) for summer” in line 443.

Fig. 16: Please be explicit what is shown: CRB-CAL or CAL-CRB

We have included the information (CAL cloud -CRB cloud) in the figure in page 25.

I419: Please also provide the % increase in winter and summer.

We have included the relative difference in line 454 as “enhanced by 2×10^{15} molec/cm\textsuperscript{2} (37\%) in winter and 8×10^{14} molec/cm\textsuperscript{2} (15\%) in summer”.

Figures 8, 10, 16 provide the contributions to Fig. 18. However, it would be good to also have the numbers for the contributions of the various terms (new STREAM, new profiles, GE-LER, CAL) to the increase. Could such numbers be provided for February and August?

We have included the individual changes of each step in Table 3. Please refer to major comment 5.

I427: I think it is good to refer to van Geffen 2020 here, who discuss this in detail. The 4.5e14 is similar/close to the 5.2e15 estimated for the operational retrieval.

Done.

I433: The table in the paper of De Smedt is mentioned. Are these numbers used without any modification? Are these consistent with estimates for e.g. the GE-LER
and OCRA/ROCINN CAL uncertainties? I suggest to include the relevant numbers in the text! Some more discussion on the uncertainties related to GE-LER, OCRA and ROCINN would be very relevant. For clear sky the AMF uncertainty is estimated as 20%.

we have updated the uncertainty estimation. Please refer to major comment 5.

l469: "which is mostly explained by the relatively low sensitivity of spaceborne measurements near the surface, the aerosol shielding effect, and the gradient smoothing effect." This is not so clear. The retrieval accounts for the lower sensitivity at the surface, and with the new GE-LER these uncertainties are hopefully reduced. The aerosol shielding effect was discussed as implicitly accounted for via the cloud retrieval. So it is not clear if a bias should remain due to these effects.

We have reformulated the expression in page 30 as “The NO$_2$ levels are underestimated by 34% by TROPOMI with a standard deviation of 16%, which is likely explained by comparison errors (such as the gradient smoothing effect, the comparison choices, and the inherent difference in sensitivity), partly by the remaining impact of structural uncertainties in the satellite data (Boersma et al., 2016, such as the impact of the choice of the a priori NO$_2$ profiles and/or the albedo database assumed for the satellite AMF calculations), and by the different measurement types or the specific conditions of the validation sites.” Please see also the discussion to the last comment for the conclusion section.

Table 4: It would be great if also numbers for the operational product could be included.

We have included the operational validation. Please refer to major comment 3.

Figure 23: Nice to see that also the operational product is included!

Fig. 24: Nice to see this plot! The sensitivity profiles of MAXDOAS and satellite are very different, so it is good to demonstrate the impact on the comparison.

l500: "see Sec 5.2"? Section 5.2 discusses the uncertainties and does not discuss the kernels.

We have removed the expression in line 542.

Conclusions section: It would be relevant to comment on the differences between the new DLR retrieval (and inputs) and other NO$_2$ retrieval approaches (operational, NASA, POMINO, ECCC, BEHR, Bremen) and discuss possible recommendations following from this comparison of retrieval methods.
We have included the comparison in page 35 as “The TROPOMI NO₂ research product from DLR is a complement to the operational product due to the use of independent approaches for stratosphere-troposphere separation and AMF calculation. Comparing to the other regional TROPOMI NO₂ product, the DLR European retrieval reduces the potential biases introduced by using inputs from different instruments or climatologies and confirms the importance of applying more realistic input parameters with better resolution for AMF calculation.

Conclusions section: The bias in the updated retrieval against MAXDOAS is reduced from 55 to 34%, but is still substantial. Do the authors have an opinion what are the main retrieval aspects causing this difference? How can this gap between surface and satellite observations be closed? Something is said about this in section 6, but it would be interesting to discuss it again in the conclusions, and perhaps including recommendations for a way forward.

Part of the remaining biases are likely related to the spatial heterogeneity effect (e.g. Pinardi 2020; Goldberg 2019), limiting the comparison of a satellite pixel (even if smaller for TROPOMI than what we were used to before) to ground-based data having a few-to-some Km length line-of-sight sensibility. Tang et al., 2021 calculated the sub-grid variability from airborne campaign measurements over Korea and Los Angeles basin, finding between 15 and 20% impact for TROPOMI pixel size. Illustration of this effect around Uccle site, based on TROPOMI data itself, is discussed in Sect. 6, Fig 22, with impact up to 19% in summer.

We have included more recommendations in page 36 as “Further improvements in the ground-based validation include using the full MAXDOAS line-of-sight sensitivity and the intersect with the TROPOMI pixel(s) or having more ground-based instruments located within a TROPOMI pixel. More frequent ground-based measurements and measurements in more than one direction might better sample the temporal and spatial variability around the measurement sites (Richter and Lange, 2021; Dimitropopulou et al., 2020).”

References


An improved tropospheric NO₂ column retrieval algorithm for TROPOMI over Europe

Song Liu¹,*, Pieter Valks¹, Gaia Pinardi², Jian Xu¹,³, Ka Lok Chan¹, Athina Argyrouli¹,⁴, Ronny Lutz¹, Steffen Beirle⁵, Ehsan Khorsandi⁶, Frank Baier⁶, Vincent Huijnen⁷, Alkiviadis Bais⁸, Sebastian Donner⁵, Steffen Dörner⁵, Myrto Gratsea⁹, François Hendrick², Dimitris Karagkiozidis⁸, Kezia Lange¹⁰, Ankie J.M. Piters⁷, Julia Remmers⁵, Andreas Richter¹⁰, Michel Van Roozendael², Thomas Wagner⁵, Mark Wenig¹¹, and Diego G. Loyola¹

¹Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Methodik der Fernerkundung (IMF), Oberpfaffenhofen, Germany
²Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
³National Space Science Center, Chinese Academy of Sciences, Beijing, China
⁴Technical University of Munich (TUM), Department of Civil, Geo and Environmental Engineering, Chair of Remote Sensing Technology, Munich, Germany
⁵Max Planck Institute for Chemistry (MPI-C), Mainz, Germany
⁶Deutsches Zentrum für Luft- und Raumfahrt (DLR), German Remote Sensing Data Center (DFD), Oberpfaffenhofen, Germany
⁷Royal Netherlands Meteorological Institute (KNMI), De Bilt, the Netherlands
⁸Laboratory of Atmospheric Physics, Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece
⁹Institute for Environmental Research and Sustainable Development, National Observatory of Athens, Greece
¹⁰Institute of Environmental Physics (IUP-UB), University of Bremen, Bremen, Germany
¹¹Meteorological Institute (MIM), Ludwig-Maximilians-Universität München (LMU), Munich, Germany

*now at: School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China

Correspondence: Song Liu (Song.Liu@dlr.de), lius7@sustech.edu.cn

Abstract.

Launched in October 2017, the TROPOspheric Monitoring Instrument (TROPOMI) aboard Sentinel-5 Precursor provides the potential to monitor air quality over point sources across the globe with a spatial resolution as high as 5.5 km × 3.5 km (7 km × 3.5 km before 6 August 2019). The DLR nitrogen dioxide (NO₂) retrieval algorithm for the TROPOMI instrument consists of three steps: the spectral fitting of the slant column, the separation of stratospheric and tropospheric contributions, and the conversion of the slant column to a vertical column using an air mass factor (AMF) calculation. In this work, an improved DLR tropospheric NO₂ retrieval algorithm from TROPOMI measurements over Europe is presented.

The stratospheric estimation is implemented using the STRatospheric Estimation Algorithm from Mainz (STREAM), which was developed as a verification algorithm for TROPOMI and does not require chemistry transport model data as input. A directionally dependent STREAM (DSTREAM) is developed to correct for the dependency of the stratospheric NO₂ on the viewing geometry by up to 2 × 10¹⁴ molec/cm². Applied to synthetic TROPOMI data, the uncertainty in the stratospheric column is 3.5 × 10¹⁴ molec/cm² for polluted conditions in case of significant tropospheric sources. Applied to actual measurements, the...
smooth variation of stratospheric NO$_2$ at low latitudes is conserved, and stronger stratospheric variation at higher latitudes are captured.

For AMF calculation, the climatological surface albedo data is replaced by geometry-dependent effective Lambertian equivalent reflectivity (GE_LER) obtained directly from TROPOMI measurements with a high spatial resolution. Mesoscale-resolution a priori NO$_2$ profiles are obtained from the regional POLYPHEMUS/DLR chemistry transport model with the TNO-MACC emission inventory. Based on the latest TROPOMI operational cloud parameters, a more realistic cloud treatment is provided by a clouds-as-layers (CAL) model, which treats the clouds as uniform layers of water droplets, instead of the clouds-as-reflecting-boundaries (CRB) model, in which clouds are simplified as Lambertian reflectors.

For the error analysis, the tropospheric AMF uncertainty, which is the largest source of NO$_2$ uncertainty for polluted scenarios, ranges between 20% and 50%, leading to a total uncertainty in the tropospheric NO$_2$ column in the 30-60% range. From a validation performed with ground-based multi-axis differential optical absorption spectroscopy (MAX-DOAS) measurements, the improved DLR tropospheric NO$_2$ data shows good correlations for nine European urban/suburban stations with an average correlation coefficient of 0.78. The implementation of the algorithm improvements leads to a decrease of the relative difference from -55.3% to -34.7% on average in comparison with the DLR reference retrieval. When the satellite averaging kernels are used to remove the contribution of a priori profile shape, the relative difference decrease further to ∼-20%.

1 Introduction

Tropospheric nitrogen dioxide (NO$_2$) is an important atmospheric trace gas because of its contribution to the formation of tropospheric ozone, urban haze, and acid deposition (Charlson and Ahlquist, 1969; Crutzen, 1970; McCormick, 2013). NO$_2$ is a prominent air pollutant affecting the human respiratory system (World Health Organization, 2006). Substantial amounts of NO$_2$ are produced in the boundary layer from combustion processes. In Europe, cities regularly exceed the air quality standards for NO$_2$ (European Commission, 2017), with road transport as the largest contributor, ahead of the energy and the industry sectors (Crippa et al., 2018).

To monitor and quantify the NO$_2$ column, NO$_2$ measurements have been provided for more than a decade on a global scale and daily basis by European satellite instruments, such as Global Ozone Monitoring Experiment (GOME) (Burrows et al., 1999), SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) (Bovensmann et al., 1999), Ozone Monitoring Instrument (OMI) (Levett et al., 2006, 2018), and Global Ozone Monitoring Experiment-2 (GOME-2) (Callies et al., 2000; Munro et al., 2006, 2016), complementary to sparse measurements by ground-based instruments.

The satellite-based NO$_2$ dataset has been extended by the new generation TROPOspheric Monitoring Instrument (TROPOMI) (Veefkind et al., 2012) with an unprecedented spatial resolution of 5.5 km in the along-track direction (7 km before 6 August 2019) and 3.5 km in the across-track direction.

TROPOMI is a nadir-viewing push-broom imaging spectrometer aboard the European Space Agency (ESA) Sentinel-5 Precursor satellite, launched on 13 October 2017. The TROPOMI instrument measures the Earth’s backscattered radiance and extraterrestrial solar irradiance in the spectral range between the ultraviolet and the shortwave infrared. The spectral
resolution and sampling are 0.54 and 0.20 nm in the visible channel (400 - 496 nm) used for the detection of NO\textsubscript{2}. From a sun-synchronous polar orbit, TROPOMI provides trace gas measurements as well as cloud and aerosol properties with an ascending node equatorial crossing at \textasciitilde 13:30 local time. The swath width is \textasciitilde 2600 km in the direction across the track of the satellite, allowing daily global coverage.

NO\textsubscript{2} measurements from TROPOMI have been widely used for ground level concentration estimates (e.g. Cooper et al., 2020; Li et al., 2020) and emission estimates (e.g. Lorente et al., 2019; Beirle et al., 2019; van der A et al., 2020; Huber et al., 2020). The high spatial resolution and good data quality allow a detailed analysis of local distribution and evolution of NO\textsubscript{2} (e.g. Stavrakou et al., 2020; Goldberg et al., 2021; Georgoulas et al., 2020), which are particularly important and helpful during the COVID-19 pandemic (e.g. Bauwens et al., 2020; Liu et al., 2020a; Goldberg et al., 2020; Huang and Sun, 2020; Ding et al., 2020; Biswal et al., 2020; Koukouli et al., 2021).

Independent from the operational processing, the NO\textsubscript{2} retrieval algorithm for the TROPOMI instrument developed at DLR starts with the calculation of the slant column (the concentration integrated along the total light path through the atmosphere along the way from the sun to the satellite) from the TROPOMI reflectance spectra using the differential optical absorption spectroscopy (DOAS) method (Platt and Stutz, 2008). To determine the tropospheric NO\textsubscript{2} slant column, the stratospheric contribution is estimated and removed from the total slant column, after which both total and tropospheric slant columns are converted to vertical columns by the application of air mass factors (AMF).

The retrieval of tropospheric NO\textsubscript{2} columns from total column data requires an accurate stratospheric estimation, a procedure referred to as stratosphere-troposphere separation (e.g. Leue et al., 2001; Bucsela et al., 2006). One typical stratosphere-troposphere-separation algorithm is the modified reference sector method, which uses measurements over regions with negligible tropospheric NO\textsubscript{2} abundance to estimate the stratospheric NO\textsubscript{2} columns based on the assumption of longitudinally homogeneous stratospheric NO\textsubscript{2} fields. A more sophisticated approach used by the operational TROPOMI NO\textsubscript{2} retrieval relying on a chemistry transport model is data assimilation (Eskes et al., 2003; Dirksen et al., 2011), in which the three-dimensional distributions of NO\textsubscript{2} are regularly updated such that the modelled stratospheric NO\textsubscript{2} concentrations are in close agreement with satellite measurements for low-tropospheric contributions. Advantages of the data assimilation approach include a realistic error estimation and the capture of small-scale dynamical and chemical variability of stratospheric NO\textsubscript{2}. Compared to data assimilation, the modified reference sector method is in general simple and requires no additional model input. Therefore, the STRatospheric Estimation Algorithm from Mainz (STREAM) method (Beirle et al., 2016), which belongs to the modified reference sector methods, has been developed as a verification algorithm for TROPOMI, as a complement to the operational stratospheric correction based on data assimilation.

The quality of satellite tropospheric NO\textsubscript{2} measurements is strongly related to the tropospheric AMFs, which are determined with a radiative transfer model and depend on ancillary information such as surface albedo, vertical shape of the a priori NO\textsubscript{2} profile, cloud, and aerosol. The importance of these parameters in NO\textsubscript{2} retrievals has been recognized for OMI (e.g. Heckel et al., 2011; Lin et al., 2014; Laughner et al., 2018; Qin et al., 2019), GOME-2 (e.g. Valks et al., 2011; Lorente et al., 2018; Liu et al., 2019b, 2020c), and TROPOMI (e.g. Griffin et al., 2019; Liu et al., 2020b; Zhao et al., 2020; Ialongo et al., 2020; Tack et al., 2021).
The surface albedo has implications for satellite retrievals of aerosols, clouds, and trace gases including NO\textsubscript{2}. Most of current satellite NO\textsubscript{2} retrievals (Boersma et al., 2011; Liu et al., 2019b, 2020c; van Geffen et al., 2020b) rely on monthly Lambertian-equivalent reflectivity (LER) climatologies derived from satellite measurements such as OMI (Kleipool et al., 2008) and GOME-2 (Tilstra et al., 2017, 2019). However, this simple assumption of isotropic surface reflection can introduce a bias by up to 35\% in the NO\textsubscript{2} AMF calculation (Lorente et al., 2018). To account for the geometry-dependent surface scattering characteristics, previous works have applied measurements from the MODerate resolution Imaging Spectroradiometer (MODIS) dataset (e.g. Vasilkov et al., 2017; Qin et al., 2019). In this study we use a new algorithm developed at DLR to retrieve geometry-dependent effective LER (GE\_LER) in the VIS based on the full-physics inverse learning machine (FP\_ILM) technique (Loyola et al., 2020b). Compared to the typical climatological LER or the directionally dependent (DLER) data (Tilstra et al., 2021), the GE\_LER data represents better the actual surface conditions such as snow/ice scenarios based on each single TROPOMI measurements with a high spatial resolution. GE\_LER has been successfully applied in the retrievals of TROPOMI total ozone columns in the UV (Loyola et al., 2020b) and cloud parameters in the NIR (Loyola et al., 2020a) and is being used in the corresponding operational version 2.1 cloud products (see introduction below).

The varying sensitivity of the satellite to NO\textsubscript{2} at different altitudes is considered in the tropospheric AMF calculation using vertically resolved box-AMFs and a priori NO\textsubscript{2} profiles. Typically prescribed by a chemistry transport model, the importance of applying a priori NO\textsubscript{2} profiles with sufficiently detailed resolution has been addressed (e.g. Russell et al., 2011; McLinden et al., 2014; Kuhlmann et al., 2015; Boersma et al., 2016; Laughner et al., 2016), particularly for TROPOMI with a small pixel size (Griffin et al., 2019; Liu et al., 2020b; Ialongo et al., 2020; Tack et al., 2021). Routine simulations of tropospheric trace gases and aerosols have been provided by POLYPHEMUS/DLR since 2014 with a spatial resolution of 0.2° × 0.3° (latitude, longitude) covering Europe and parts of North Africa. POLYPHEMUS/DLR is an air quality modelling platform operated at DLR based on the POLYPHEMUS chemistry transport model (Mallet et al., 2007) coupled to the Weather Research and Forecasting (WRF) model (Skamarock et al., 2008) with the TNO-MACC emission inventory (Denier van der Gon et al., 2010). It has been further developed within the PASODOBLE project for sensitivity studies of the mountainous Black Forest region (Bergemann et al., 2012) and to cover urban areas in southern Bavaria (Khorsandi et al., 2018). It uses the TNO-MACC emission inventory (Kuenen et al., 2014). Daily model forecasts are freely available via DLR Geospatial Web Services (http://wdc.dlr.de/cgi-bin/produkt_4d_w?).

The NO\textsubscript{2} retrieval is affected by the presence of clouds, because high clouds shield underlying parts of the atmosphere, and low clouds can enhance the NO\textsubscript{2} absorption due to cloud albedo and multiple scattering if they are below or at the same height as the NO\textsubscript{2} layer (Martin et al., 2002; Kokhanovsky and Rozanov, 2008). The operational cloud retrieval for TROPOMI is implemented using Optical Cloud Recognition Algorithm (OCRA) and Retrieval Of Cloud Information using Neural Networks (ROCINN). In addition to the retrieval product based on the assumption that clouds are simple Lambertian reflecting surfaces, referred to as Clouds-as-Reflecting-Boundaries (CRB) model, a more sophisticated set of cloud products is provided by OCRA/ROCINN, which considers clouds as optically uniform layers of scattering liquid water spherical particles, referred to as Clouds-as-Layers (CAL) model. The more realistic CAL model is regarded as the preferred method, particularly for small TROPOMI ground pixels and for low clouds (Compernolle et al., 2021). With an updated OCRA/ROCINN processor
version 2.1 processor for OCRA/ROCINN in operation since August 2020, new features such as the application of GE_LER to describe the surface albedo have been added in OCRA/ROCINN (Loyola et al., 2020a).

The satellite NO$_2$ data has been widely validated by comparison with correlative ground-based multi-axis differential optical absorption spectroscopy (MAX-DOAS) measurements (e.g. Celarier et al., 2008; Irie et al., 2008; Ma et al., 2013; Pinardi et al., 2014, 2015; Drosoglou et al., 2017, 2018; Chan et al., 2020; Pinardi et al., 2020). MAX-DOAS measures the vertically resolved abundances of atmospheric trace species in the lowermost troposphere (Hönninger et al., 2004; Wagner et al., 2004; Wittrock et al., 2004; Heckel et al., 2005). Based on the scattered sky light under different viewing directions, high NO$_2$ sensitivity close to the surface is obtained for the smallest elevation angles, whereas measurements at higher elevations provide information on the rest of the column.

In this paper, a number of improvements to the tropospheric NO$_2$ retrieval over Europe are introduced for the DLR product. To estimate and remove the stratospheric contribution, an improved STREAM algorithm is developed and evaluated by applying it to synthetic TROPOMI data and actual satellite observations. To calculate the tropospheric AMFs, the surface albedo is described by the GE_LER data consistently in both NO$_2$ and cloud retrievals; a priori NO$_2$ profiles are obtained from the regional POLYPHEMUS/DLR chemistry transport model; the OCRA/ROCINN_CAL cloud model from the new version 2.1 OCRA/ROCINN processor is used for cloud correction.

In Sect. 2, we introduce the reference algorithm at DLR for the TROPOMI NO$_2$ retrieval, which is based on an improved algorithm originally designed for GOME-2 (Liu et al., 2019b) and adapted for TROPOMI measurements with optimization related to the specific instrumental aspects. In Sects. 3 and 4, we improve the stratosphere-troposphere separation and the tropospheric AMF calculation, respectively. In Sect. 5, examples of applying the retrieval algorithm to TROPOMI measurements are shown, and the error estimate is discussed. In Sect. 6, we show a comprehensive validation of the TROPOMI NO$_2$ data using ground-based MAX-DOAS observations in Europe.

2 DLR reference retrieval for TROPOMI NO$_2$ measurement

2.1 DOAS slant column retrieval

Applied to the backscattered spectra measured by TROPOMI, the DOAS fit (Platt and Stutz, 2008) is a least-squares inversion to isolate the trace gas absorption from the background processes, which are approximated by a fifth-order polynomial $P(\lambda)$ at wavelength $\lambda$:

$$\ln \left[ \frac{I(\lambda) + offset(\lambda)}{I^0(\lambda)} \right] = - \sum_g S_g \sigma_g(\lambda) - \alpha_R R(\lambda) - P(\lambda).$$  (1)

The measurement-based term is defined as the natural logarithm of the measured earthshine radiance spectrum $I(\lambda)$ divided by the daily solar irradiance spectrum $I^0(\lambda)$. The DOAS fit is performed in the 405-465 nm wavelength range for consistency with other NO$_2$ retrievals from TROPOMI (van Geffen et al., 2020a) and the heritage instrument OMI (van Geffen et al., 2015; Zara et al., 2018).
The spectral effect from the absorption of a species \( g \) is determined by the fitted slant column density \( S_g \) and associated absorption cross-section \( \sigma_g(\lambda) \):

- NO\(_2\) absorption at 220K from Vandaele et al. (2002);
- ozone (O\(_3\)) absorption at 228K from Brion et al. (1998);
- water vapor (H\(_2\)O\(_{vap}\)) absorption at 293K from Rothman et al. (2010), rescaled as in Lampel et al. (2015);
- oxygen dimer (O\(_4\)) absorption at 293K from Thalman and Volkamer (2013);
- liquid water (H\(_2\)O\(_{liq}\)) absorption at 297K from Pope and Fry (1997), smoothed as in Peters et al. (2014).

The contribution of the rotational Raman scattering to the measured spectrum, namely the Ring effect (Grainger and Ring, 1962; Solomon et al., 1987), is treated as a pseudo absorber, by means of an additive Ring reference spectrum \( R(\lambda) \) and a scaling coefficient \( \alpha_R \) as fitting parameter. A linear intensity offset correction \( \text{offset}(\lambda) \) is fitted as an additional effective cross-section to correct for the stray light in the spectrometer, the inelastic scattering in the ocean, and remaining calibration issues in the level 1 data (Liu et al., 2019b). See additional discussions regarding intensity offset in van Geffen et al. (2020a).

The TROPOMI level 1b version 1 spectra are analysed using the QDOAS software developed at BIRA-IASB (Fayt et al., 2011; Danckaert et al., 2017). Figure 1 shows examples of the TROPOMI NO\(_2\) slant columns over Europe in February and August 2019, where large NO\(_2\) hotspots can be identified. Note that the slant columns are scaled by geometric AMFs to correct for the angular dependencies of TROPOMI measurements.

The NO\(_2\) slant columns from single orbits show an across-track striping pattern, a well-known feature of observations of push-broom spectrometers such as OMI (Boersma et al., 2011) and TROPOMI (van Geffen et al., 2020a), which is likely caused by the viewing zenith angle (VZA) dependency of the spectral calibration and detector sensitivity (Boersma et al., 2018). To
reduce the systematic stripes, a de-striping correction is applied to the TROPOMI NO\textsubscript{2} slant columns, which amplitude is calculated empirically (Boersma et al., 2011) based on the daily averaged across-track variability of NO\textsubscript{2} slant columns over clean regions between 20°S and 20°N. The magnitude of the NO\textsubscript{2} de-striping correction is up to $1 \times 10^{14}$ molec/cm\textsuperscript{2} and is stable over time (not shown), in agreement with the operational TROPOMI de-striping that relies on the chemistry transport model data (van Geffen et al., 2020a).

### 2.2 Stratosphere-troposphere separation

The stratospheric NO\textsubscript{2} component is estimated using the STREAM method (Beirle et al., 2016). Belonging to the modified reference sector methods, STREAM uses total NO\textsubscript{2} column measurements over clean and remote regions as well as over clouded scenes with negligible tropospheric columns. STREAM calculates weighting factors for each satellite pixel to define the contribution of total columns to the stratospheric estimation: potentially polluted pixels are weighted low, cloudy observations with medium cloud heights are weighted high, and the weights are further adjusted in case of large biases in the tropospheric residues. Depending on these weighting factors, stratospheric NO\textsubscript{2} fields are derived by a weighted convolution of the total columns using convolution kernels, which are wider at lower latitudes to account for the low longitudinal variability assumption of stratospheric NO\textsubscript{2} and narrower at higher latitudes to reflect the stronger natural variations.

STREAM was developed as a verification algorithm for the TROPOMI instrument, as a complement to the operational stratospheric correction based on data assimilation (van Geffen et al., 2020b). STREAM has been successfully applied to the NO\textsubscript{2} measurements from GOME, SCIAMACHY, OMI, and GOME-2 (Beirle et al., 2016; Liu et al., 2019b) with the advantage of requiring no model input. In contrast to previous modified reference sector methods which normally apply a conservative masking approach (flagging pixels as either clean or polluted and skipping the latter for stratospheric estimation) and hardly use information over continents, STREAM introduces an improved treatment of polluted and cloudy pixels by defining weighting factors for each satellite pixel. Due to the use of cloudy pixels, which directly reflect the actual stratospheric column as the tropospheric column is mostly shielded, STREAM reduces the bias caused by the free-tropospheric contamination or tropospheric background in the reference region, in comparison with other modified reference sector methods (Beirle et al., 2016). However, the broad-scale free tropospheric diffuse NO\textsubscript{2} might not be fully separated from the stratospheric NO\textsubscript{2}, which is a general limitation of the modified reference sector methods.

Stratospheric NO\textsubscript{2} columns from STREAM differ by up to $3 \times 10^{14}$ molec/cm\textsuperscript{2} as compared to results from data assimilation and other modified reference sector methods, within the general uncertainties of stratosphere-troposphere separation (Beirle et al., 2016; Boersma et al., 2018). The STREAM stratospheric NO\textsubscript{2} columns show an average bias of $1 \pm 8 \times 10^{13}$ molec/cm\textsuperscript{2} with respect to the ground-based zenith-scattered light differential optical absorption spectroscopy (ZSL-DOAS) measurements (Compernolle et al., 2020).
Figure 2. Tropospheric NO$_2$ columns from the reference algorithm over Europe in February and August 2019. Only measurements with cloud radiance fraction less than 0.5 are included.

2.3 AMF calculation

The conversion between the slant column $S$ and the vertical column $V$ is implemented by division with an AMF $M$:

$$V = \frac{S}{M}. \quad (2)$$

Given the small optical depth of NO$_2$, $M$ can be derived as:

$$M = \frac{\sum_l m_l(b)x_l c_l}{\sum_l x_l}, \quad (3)$$

where $m_l$ is the box-AMFs in layer $l$, $x_l$ is the partial column from the a priori NO$_2$ profile, and $c_l$ is a correction coefficient to correct for the temperature dependency of the NO$_2$ cross section (Boersma et al., 2004; Nüß et al., 2006; Bucsela et al., 2013). $m_l$ is a function of model inputs $b$, which include TROPOMI measurement geometries, surface albedo, and surface pressure. The box-AMFs $m_l$ values are calculated at 437.5 nm (near the mid-point wavelength of fitting window 405-465 nm), as recommended by Boersma et al. (2018), using the linearised vector code VLIDORT version 2.7 (Spurr, 2006). The light path in the troposphere is affected by scattering on air molecules as well as cloud and aerosol particles, and therefore the tropospheric AMF calculation depends on surface albedo, a priori NO$_2$ profiles, and cloud properties. Table 1 summarises the parameters used in the AMF calculation in several research products and the operational product. Figure 2 shows the tropospheric NO$_2$ columns retrieved from the reference algorithm over Europe in February and August 2019. A large amount of NO$_2$ is located in the troposphere for industrialised and urbanised areas (see Fig. 1).

In the reference algorithm, the surface albedo is described by a monthly climatology based on four years (2004-2007) of OMI LER measurements at 440 nm (Kleipool et al., 2008) with a similar overpass time and viewing conditions as TROPOMI. The surface albedo for each TROPOMI pixel is calculated by an area-weighted tessellation of the OMI monthly averaged surface albedo maps ($0.5^\circ \times 0.5^\circ$) and a linear interpolation in time to the measurement day.
Table 1. Parameters used to calculate tropospheric AMFs. Overview of tropospheric NO\textsubscript{2} column retrievals. See Table 2 for details of the chemistry transport models used to obtain the a priori NO\textsubscript{2} profiles for the DLR improved retrieval in this work.

<table>
<thead>
<tr>
<th>Reference</th>
<th>DLR Retrieval</th>
<th>DLR Improved Algorithm</th>
<th>KNMI Operational Product</th>
<th>POMINO-TROPOMI Product</th>
<th>ECCC Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
<td>Global</td>
<td>Europe</td>
<td>Global</td>
<td>China</td>
<td>Canada</td>
</tr>
<tr>
<td>Stratospheric correction</td>
<td>STREAM</td>
<td>DSTREAM</td>
<td>Data assimilation</td>
<td>Same as KNMI</td>
<td>Same as KNMI</td>
</tr>
<tr>
<td>Surface albedo</td>
<td>OMI LER climatology</td>
<td>TROPOMI GE_LER data</td>
<td>OMI LER climatology</td>
<td>MODIS BRDF climatology</td>
<td>MODIS albedo</td>
</tr>
<tr>
<td>A priori NO\textsubscript{2} profile</td>
<td>TM5-MP</td>
<td>POLYPHEMUS/DLR TM5-MP</td>
<td>GEOS-Chem GEM-MACH</td>
<td>Recalculated cloud fraction with aerosol correction</td>
<td>Same as KNMI</td>
</tr>
<tr>
<td>Cloud parameter</td>
<td>OCRA/ROCINN_CRB version 1.x</td>
<td>OCRA/ROCINN_CAL FESCO-S</td>
<td>Same as KNMI</td>
<td>Same as KNMI</td>
<td></td>
</tr>
<tr>
<td>Aerosol treatment</td>
<td>Implicit correction</td>
<td>Implicit correction</td>
<td>Implicit correction</td>
<td>Explicit correction</td>
<td>Explicit correction</td>
</tr>
</tbody>
</table>

Daily TM5-MP vertical NO\textsubscript{2} profiles (Williams et al., 2017) simulated at a global 1° × 1° resolution are used as a priori NO\textsubscript{2} vertical profiles due to the operational advantage, as summarized in Table 2. The a priori profiles are determined for the satellite overpass time and interpolated to the center of the TROPOMI pixel based on four nearest neighbour TM5-MP cell centers.

In the presence of clouds, the AMF calculation adopts the independent pixel approximation (Cahalan et al., 1994):

\[ M = \omega M^{cl} + (1 - \omega) M^{cr}, \]

where \(M^{cl}\) represents the AMF for completely cloudy sky and \(M^{cr}\) for completely clear sky. \(M^{cl}\) and \(M^{cr}\) are derived with Eq. (3) with \(M^{cl}\) mainly relying on the cloud pressure (height) and the cloud albedo (optical depth). The cloud radiance fraction \(\omega\) is derived from the TROPOMI cloud fraction \(c_f\):

\[ \omega = \frac{c_f I^{cl}}{(1 - c_f) I^{cr} + c_f I^{cl}} \]

with \(I^{cl}\) and \(I^{cr}\) representing the radiances for cloudy and clear scenes, respectively. \(I^{cl}\) and \(I^{cr}\) are calculated using the LIDORT version 3.6 model (Spurr et al., 2001), depending mostly on TROPOMI viewing geometries, surface albedo, and cloud albedo.

The operational TROPOMI cloud parameters from the OCRA/ROCINN algorithms (Lutz et al., 2016; Loyola et al., 2018) with clouds treated as opaque Lambertian surfaces (CRB model) are applied for the cloud correction. OCRA derives the cloud fraction by separating a spectral scene into cloudy contribution and cloud-free background. With the OCRA cloud fraction and
Table 2. Summary of the chemistry transport model specifications.

<table>
<thead>
<tr>
<th></th>
<th>TM5-MP</th>
<th>POLYPHEMUS/DLR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Huijnen et al., 2010; Williams et al., 2017)</td>
<td>(Mallet et al., 2007; Bergemann et al., 2012; Khorsandi et al., 2018)</td>
</tr>
<tr>
<td>Spatial resolution</td>
<td>1°×1° (latitude, longitude) 0.2°×0.3° (latitude, longitude)</td>
<td></td>
</tr>
<tr>
<td>Vertical resolution (&gt;150 hPa)</td>
<td>~18 layers</td>
<td>20 layers</td>
</tr>
<tr>
<td>Meteorological fields</td>
<td>ECMWF 3 h</td>
<td>RACM for trace gases (Stockwell et al., 1997)</td>
</tr>
<tr>
<td>Tropospheric chemistry</td>
<td>Modified CB05 (Williams et al., 2013)</td>
<td>SORGAM-SIREAM for aerosols (Debry et al., 2007; Schell et al., 2001)</td>
</tr>
<tr>
<td>Anthropogenic emission</td>
<td>MACCity (Granier et al., 2011)</td>
<td>TNO-MACC (Denier van der Gon et al., 2010; Kuenen et al., 2014)</td>
</tr>
<tr>
<td>Advection</td>
<td>Slopes scheme (Russell and Lerner, 1981)</td>
<td>third-order direct space-time scheme with a Koren-Sweby flux limiter</td>
</tr>
<tr>
<td>Convection</td>
<td>ECMWF</td>
<td>WRF (Skamarock et al., 2008)</td>
</tr>
<tr>
<td>Diffusion</td>
<td>Holtslag and Boville (1993)</td>
<td>second-order Rosenbrock method (Verwer et al., 2002)</td>
</tr>
</tbody>
</table>

---

the surface albedo from the MERIS black-sky climatology (Popp et al., 2011) as inputs, ROCINN calculates the cloud pressure and cloud albedo by comparing the measured and simulated sun-normalised radiances in and around the O$_2$ A-band in the NIR.

The original OCRA takes the spectral information from the UV-VIS-NIR part (320 – 800 nm) and transforms the radiances of three predefined spectral ranges to three-color reflectances (RGB: red, green, and blue region of the spectrum). The cloud-free background maps are calculated for each of these three colors. For the TROPOMI application, the OCRA color space approach is applied with two colors (GB) using the UV-VIS (350 – 495 nm) spectra to avoid the spatial misalignment between the UV-VIS and NIR footprints.

Based on the Lambertian cloud assumption, OCRA/ROCINN_CRB tends to retrieve a cloud height (at the optical centroid of the cloud rather than the cloud top) close to the surface altitude for low cloud fraction (Compernolle et al., 2021). The Lambertian cloud assumption is also applied in the Fast Retrieval Scheme for Clouds from the O$_2$ A-band (FRESCO) algorithm (Koelemeijer et al., 2001; Wang et al., 2008). FRESCO for Sentinel (FRESCO-S) (Wang and Sneep, 2019) is implemented as a support product for the TROPOMI operational NO$_2$ processing (van Geffen et al., 2020b). FRESCO-S retrieves the cloud fraction and cloud pressure from the reflectance in and around the O$_2$ A-band. The cloud albedo is assumed to be 0.8, as opposed to OCRA/ROCINN, where cloud albedo is retrieved.
Figure 3. VZA dependency of TROPOMI total NO\textsubscript{2} columns (compared to nadir) at different latitudes in January 2019. The VZA is defined negative for observations on the west side of the orbit swath. The de-striping is not implemented here.

3 New Stratosphere-troposphere separation

STREAM was originally designed for TROPOMI and optimized for OMI within TROPOMI verification activities (Beirle et al., 2016). STREAM consists basically of two steps: the definition of weighting factors for each satellite pixel and the application of weighted convolution. To identify potentially polluted areas, a climatology of tropospheric NO\textsubscript{2} columns is derived in this study using TROPOMI NO\textsubscript{2} measurements from 2018-2019, instead of using SCIAMACHY NO\textsubscript{2} measurements from 2003-2011 as in the original STREAM. Based on the pollution weight, as well as the cloud weight and tropospheric residue weight, STREAM estimates stratospheric fields for individual orbits using a weighted convolution on 0.5\degree × 0.5\degree grid pixels.

As a result of wide swath (~2600 km), local time differences across a TROPOMI swath are considerable at high latitudes, and the NO\textsubscript{2} measurements show dependency on VZA (directly related to local time) due to the diurnal variation of stratospheric NO\textsubscript{2} (Dirksen et al., 2011; Belmonte Rivas et al., 2014). Figure 3 shows the total NO\textsubscript{2} columns measured by TROPOMI in January 2019 for different latitudes as a function of VZA. The impact of local time changes across the orbit is up to \(2 \times 10^{14}\) molec/cm\(^2\) at the swath edge for latitudes higher than 50\degree, in agreement with estimations for OMI measurements (Beirle et al., 2016).

In the following, the concept of a directionally dependent STREAM (DSTREAM) is introduced to estimate the stratospheric NO\textsubscript{2} column (Sect. 3.1). The performance of STREAM and DSTREAM is analyzed using synthetic TROPOMI NO\textsubscript{2} data (Sect. 3.2), and both algorithms are applied to TROPOMI measurements (Sect. 3.3).
3.1 DSTREAM

To correct for the VZA dependency of stratospheric NO$_2$, the DSTREAM is developed, which divides the orbit swath into western (VZA from $\sim-66^\circ$ to $\sim-30^\circ$), central (VZA from $\sim-30^\circ$ to $\sim30^\circ$), and eastern (VZA from $\sim30^\circ$ to $\sim66^\circ$) segments. Note that the VZA is defined negative for observations on the west side of the orbit swath throughout the study. For each of the orbit swath containers, the original STREAM is applied based on data from the respective orbit swath segment. For each individual satellite pixel with a VZA, a directionally dependent stratospheric NO$_2$ column $V_s^{\text{dir}}$ is parameterized using a linear interpolation on the DSTREAM grid results estimated using the eastern, central, and western segments of the orbit swath. As the VZA dependency is negligible for low latitudes (from Fig. 3), and the interpolation error may increase for $V_s^{\text{dir}}$ due to less orbital overlap, the final stratospheric NO$_2$ column $V_s$ is calculated as the weighted mean in dependence on latitude $\theta$:

$$V_s = \cos^2(\theta)V_s^{\text{ori}} + \sin^2(\theta)V_s^{\text{dir}}.$$  \hspace{1cm} (6)

By this method, the stratospheric NO$_2$ from the original STREAM $V_s^{\text{ori}}$ is applied for equator, and the VZA dependency is captured for polar regions with significant orbital overlap.

3.2 Application to synthetic data

The performance of the original STREAM and the improved DSTREAM for TROPOMI is evaluated with simulated NO$_2$ fields from the IFS(CB05BASCOE) experiment (Huijnen et al., 2016). Particularly advantageous for stratospheric studies, the IFS(CB05BASCOE) model is particularly advantageous for stratospheric studies due to the extension of the tropospheric chemistry module in the Integrated Forecast System (IFS) is extended with the stratospheric chemistry from the Belgian Assimilation System for Chemical ObsErVations (BASCOE) system. The STREAM and DSTREAM are applied to the synthetic TROPOMI total NO$_2$ columns, and the estimated stratospheric NO$_2$ columns are compared with the a priori truth (stratospheric fields from model). See Liu et al. (2019b) for more details on constructing synthetic total and stratospheric NO$_2$ columns using IFS(CB05BASCOE) data.

Figure 4 displays the synthetic total NO$_2$ columns from IFS(CB05BASCOE), the modelled stratospheric columns, and the estimated stratospheric columns from STREAM and DSTREAM on 1 January and 1 August 2019. The overall latitudinal and seasonal dependencies are reflected in the stratospheric fields from STREAM and DSTREAM. Smaller structures in the synthetic total columns and the modelled stratospheric columns at high latitudes, caused by the diurnal variation of stratospheric NO$_2$ across the orbital swath, are aliased into the troposphere by STREAM but captured by DSTREAM. The average difference between the estimated and a priori results is $4 \times 10^{14}$ molec/cm$^2$ for STREAM and $3.5 \times 10^{14}$ molec/cm$^2$ for DSTREAM with improvements mainly for latitudes higher than $50^\circ$. 
Figure 4. Synthetic total NO$_2$ columns, a priori stratospheric NO$_2$ columns from IFS(CB05BASCOE), and estimated stratospheric NO$_2$ columns from STREAM and DSTREAM on 1 February and 1 August 2019.
Figure 5. Total NO₂ columns and stratospheric NO₂ columns estimated using STREAM and DSTREAM, as measured by TROPOMI on 1 February and 1 August 2019.

3.3 Application to TROPOMI measurements

Applying STREAM and DSTREAM to TROPOMI data, Fig. 5 shows the total columns from TROPOMI and the estimated stratospheric fields on 1 February and 1 August 2019. For both months, the STREAM and DSTREAM show similar global patterns of stratospheric NO₂. The stratospheric and tropospheric contributions over polluted regions are successfully separated due to the use of clean and cloudy measurements at the same latitude where the tropospheric column is shielded. The smooth background at low latitudes is conserved, and the stronger variations of stratospheric NO₂ at higher latitudes are captured, e.g. in the polar vortex on 1 February.
Figure 6. VZA dependency of TROPOMI stratospheric NO₂ columns estimated using STREAM and DSTREAM (compared to nadir) at different latitudes in January 2019. The VZA is defined negative for observations on the west side of the orbit swath.

Figure 6 shows the TROPOMI stratospheric columns estimated using STREAM and DSTREAM for different latitudes as a function of VZA in January 2019. The VZA dependency by up to $2 \times 10^{14}$ molec/cm² caused by the local time difference in Fig. 3 is captured by DSTREAM, particularly for high latitudes. The overestimation on the west side of swath edge and the underestimation on the east side in the STREAM results are improved corrected in DSTREAM.

4 Improved AMF calculation

4.1 Surface albedo

Surface albedo is an important parameter for an accurate retrieval of trace gas columns and cloud properties. The sensitivity of backscattered radiance to the boundary layer NO₂ is strongly related to the surface albedo, especially over polluted areas. In this study, the surface albedo is described using GE_LER retrieved by the FP_ILM algorithm (Loyola et al., 2020b). Unlike conventional approaches (Rodgers, 2000; Doicu et al., 2010), FP_ILM is a machine learning based approach consisting of a training phase wherein an inverse function is derived from synthetic data generated with a radiative transfer model and an operational phase wherein the inverse function is applied to measured spectra. The FP_ILM algorithm has been employed to retrieve ozone profile shapes and sulfur dioxide layer heights from GOME-2 and TROPOMI (Xu et al., 2017; Efremenko et al., 2017; Hedelt et al., 2019).

Combining the DOAS equation Eq. (1) and the conventional forward model, our forward problem can be formulated as an approximation of the DOAS-fitted slant column density ($SCD$) and the DOAS polynomial ($P$) using the forward model ($F$)
with the solar/satellite viewing geometry (\(\Theta\)), effective surface scene pressure (\(p_e\)), and surface albedo (\(A_s\)):

\[
\{SCD, P\} = F(\Theta, p_e, A_s).
\] (7)

During the training phase, synthetic TROPOMI spectra in the 405 - 465 nm range are simulated by LIDORT version 3.6 (Spurr et al., 2001) in conjunction with the smart sampling technique (Loyola et al., 2016). The cloud impact is considered in the simulations with the use of the effective surface scene pressure \(p_e\), which depends on OCRA cloud fraction, ROCINN.CR.B cloud pressure, and surface pressure (Loyola et al., 2020b) which is provided in the new version 2.1 processor for OCRA/ROCINN (see Sect. 4.3.1). The aerosol influence is not considered. The DOAS fitting is applied to the simulated spectra using the consistent DOAS settings as introduced in Sect. 2.1. The simulation results from Eq. (7) are grouped as inputs to a multi-layer neural network, and the neural network is trained to learn the inverse function. In the operational phase, GE.LER is generated using the trained neural network and the DOAS results from the measured spectra. An additional polynomial fitting is subsequently included to account for the bidirectional reflectance distribution function (BRDF) effect.

For consistency with the \(\text{NO}_2\) retrieval, the GE.LER retrieval is performed for every single ground pixel using the same TROPOMI spectrum and DOAS configurations. Global maps are generated from the GE.LER retrievals under clear-sky conditions (OCRA cloud fraction small than 0.05) and updated on a daily basis on a time window between one and four weeks depending on cloudiness. In contrast to the OMI LER climatology, the GE.LER data relies on the measurements from the TROPOMI instrument itself with an improved higher spatial resolution (0.1° × 0.1°) and better characterizes the actual surface conditions, particularly for snow/ice scenarios.

Figure 7 compares the climatological OMI LER data and GE.LER data for February and August 2019. The surface LER values from GE.LER are lower than the climatological OMI values by 0.03 on average.

The mean differences between the surface LER values from GE.LER and OMI climatological values are lower than 0.03. The improved spatial resolution for GE.LER enables a better representation of surface features. Larger differences by more than 0.2 are found in winter over snow/ice regions such as Russia and the Alps, because GE.LER captures the actual snow/ice conditions. The GE.LER values are higher by up to 0.05 over the North Sea, due to the use of only one month of TROPOMI data compared to the multiple years for OMI climatology, which makes GE.LER more likely affected by aerosol contamination or outliers in the input data.

In the near future, an improved aerosol screening based on TROPOMI aerosol index data will be implemented in the GE.LER algorithm. An independent Neural Network will be trained and implemented for error estimation.

Figure 8 shows the monthly average differences in the tropospheric \(\text{NO}_2\) columns retrieved using the climatological OMI LER and the TROPOMI GE.LER in February and August 2019. An effect is noticed mainly in winter under polluted conditions. Consistent with the LER changes in Fig. 7, the general reduced surface LER from GE.LER results in a decrease in the tropospheric AMF and thus an increase in the calculated tropospheric \(\text{NO}_2\) column by up to \(3 \times 10^{15}\) molec/cm\(^2\). A reduction by up to \(1 \times 10^{15}\) molec/cm\(^2\) is found for snow/ice coverages and aerosol scenes.
Figure 7. OMI surface LER climatology at 440 nm (Kleipool et al., 2008) and GE_LER retrieved from TROPOMI data in the NO₂ retrieval window (405 - 465 nm) over Europe in February and August 2019.

Figure 8. Differences in the tropospheric NO₂ columns retrieved using the climatological OMI LER and GE_LER over Europe in February and August 2019. Only measurements with cloud radiance fraction less than 0.5 are included.
4.2 A priori NO\textsubscript{2} profiles

To account for the varying sensitivity of the satellite to NO\textsubscript{2} at different altitudes, the POLYPHEMUS/DLR simulations (Mallet et al., 2007) with a spatial resolution of 0.2° × 0.3° (latitude, longitude) and a temporal resolution of 1 h are applied for Europe (34.0°N-60.4°N,12.0°W-40.2°E) in this study. Compared to the reference algorithm using TM5-MP a priori NO\textsubscript{2} profiles, it can be expected that the improved resolution of POLYPHEMUS/DLR is better able to capture accurately the local NO\textsubscript{2} distribution, particularly for regions with large heterogeneity and variability.

As summarized in Table 2, the meteorological parameters are provided by Weather Research and Forecasting (WRF) Version 3.5 daily forecasts with a 30 km × 30 km spatial resolution, initialized by daily Global Forecast System (GSF) global forecast from National Oceanic and Atmospheric Administration (NOAA). The original POLYPHEMUS model is an assembly of several Eulerian and Gaussian models for handling passive tracers, photochemistry, and aerosol dynamics (Mallet et al., 2007) and is developed based on the chemistry transport model Polair3D (Boutahar et al., 2004). In this study, the Regional Atmospheric Chemistry Modeling (RACM) chemical mechanism (Stockwell et al., 1997) is applied along with the Size-REsolved Aerosol Model (SIREAM) and Secondary ORGanic Aerosol Model (SORGAM) size-resolved aerosol model (Debry et al., 2007; Schell et al., 2001). The anthropogenic emissions are extracted from the European TNO-MACC (from 2011 with base year 2005) emission inventory (Denier van der Gon et al., 2010; Kuenen et al., 2014). With a spatial resolution of 7 km × 7 km, TNO-MACC defines 10 source categories including road transport and shipping and uses source sector-specific data in a harmonized way. Biogenic emissions from soils are computed as proposed in Simpson et al. (1999). Lightning emissions are not considered.

Figure 9 shows the TM5-MP and POLYPHEMUS/DLR a priori NO\textsubscript{2} profiles over Munich in Germany (48.15°N, 11.57°E) on 5 February 2019, with the calculated clear-sky tropospheric AMFs also reported. POLYPHEMUS/DLR shows a higher surface layer NO\textsubscript{2} concentration and yields a tropospheric AMF that is reduced by 0.36 (36.7%). Figure 9 additionally shows the low-tropospheric NO\textsubscript{2} profile derived from the ground-based MAX-DOAS data (Chan et al., 2020) and the tropospheric AMF calculated using the MAX-DOAS NO\textsubscript{2} profile as a priori information (assuming a constant profile shape for the high troposphere). With a typical horizontal sensitivity of a few kilometers, the MAX-DOAS profile shows large amounts of NO\textsubscript{2} in the lower troposphere (Irie et al., 2011; Wagner et al., 2011). The profile shape from POLYPHEMUS/DLR agree better with the MAX-DOAS measurements than TM5-MP, with the tropospheric AMF bias improving from 0.32 (48.5%) to -0.04 (-6.0%). In comparison with the tropospheric AMF calculated using the MAX-DOAS NO\textsubscript{2} profile, the bias reduces from 0.32 (48.5%) for TM5-MP to -0.04 (-6.0%) for POLYPHEMUS/DLR.

Figure 10 shows the monthly average differences in the tropospheric NO\textsubscript{2} columns retrieved using TM5-MP and POLYPHEMUS/DLR a priori NO\textsubscript{2} profiles in February and August 2019. POLYPHEMUS/DLR uses the TNO-MACC emission database, which generally shows higher total NO\textsubscript{2} emissions at much higher spatial resolution in comparison with the global MACCity inventory. Further tests to investigate the sensitivity of model NO\textsubscript{2} profiles have concluded that the horizontal resolution and the representation of the tropospheric boundary layer have the largest influence (not shown). The generally steeper profile shape from POLYPHEMUS/DLR (see Fig. 9) increases the retrieved tropospheric NO\textsubscript{2} columns by more than 2 × 10\textsuperscript{15} molec/cm\textsuperscript{2}.
Figure 9. A priori NO$_2$ profiles from the chemistry transport models TM5-MP and POLYPHEMUS/DLR and the low-tropospheric NO$_2$ profile derived from the ground-based MAX-DOAS measurements over Munich in Germany (48.15°N, 11.57°E) on 5 February 2019. The calculated clear-sky tropospheric AMF is given in the bracket next to each label in the legend. Normalized profiles (to the lowest values) are also shown on a logarithm scale.

Figure 10. Differences in the tropospheric NO$_2$ columns retrieved using the TM5-MP and POLYPHEMUS/DLR a priori NO$_2$ profiles over Europe in February and August 2019. Only measurements with cloud radiance fraction less than 0.5 are included.

for pollution hot spots, e.g. regions with large population or heavy industry in the Benelux, northern Italy, and western Turkey, as well as highways with intense road traffic in northern Spain, southern France, and western Germany.

Figure 11 compares the tropospheric NO$_2$ columns from TROPOMI measurements and POLYPHEMUS/DLR simulations. The satellite averaging kernel, which describes the vertical sensitivity of measurements of NO$_2$ concentrations, is applied
Figure 11. Tropospheric NO$_2$ columns from TROPOMI measurements (retrieved with the POLYPHEMUS/DLR a priori NO$_2$ profiles) and POLYPHEMUS/DLR simulations (using the satellite averaging kernel) over Europe in February and August 2019. Only TROPOMI measurements with cloud radiance fraction less than 0.5 are included.

To reduce the systematic biases caused by unrealistic a priori profile information. From Fig. 11, POLYPHEMUS/DLR NO$_2$ columns are generally higher than satellite measurements, which can be partly related to the use of the TNO-MACC emission dataset (Denier van der Gon et al., 2010). An update of the POLYPHEMUS/DLR model using the more recent TNO-MACC$_{II}$ emission is planned for the near future. We note here that the profile shape is of far more importance than the column bias for the interpretation of satellite retrievals.

4.3 Cloud correction

4.3.1 New processor for OCRA/ROCINN

With a new version 2.1 OCRA/ROCINN processor in operation since August 2020 (Loyola et al., 2020a), improvements are introduced to the OCRA/ROCINN retrievals. The background maps used in the OCRA cloud fraction determination are calculated based on one year of TROPOMI data (April 2018 - March 2019) instead of the previously used OMI measurements, and the spatial resolution improves from $0.2^\circ \times 0.4^\circ$ to $0.1^\circ \times 0.1^\circ$. In the pre-processing step, a TROPOMI-based VZA dependency correction is applied instead of using the OMI measurements.
Figure 12. OCRA cloud fractions from the version 2.1 processor (left) and comparisons with version 1.x (right) over Europe for orbit 6939 on 14 February 2019. Only measurements with $0 < \text{cloud fraction} \leq 0.4$ are included.

For ROCINN, the static MERIS surface albedo climatology (Popp et al., 2011) is replaced with the dynamic on-line GE_LER retrieval in the ROCINN fitting window, which derives the surface properties directly from TROPOMI itself on a daily basis (see Sect. 4.1). The co-registration between UV-VIS band (used by OCRA) and NIR band (used by ROCINN) is optimized with a look-up table containing the fraction of overlapping area between the target and source pixels, and the cloud properties are provided in both spectral bands. Over snow/ice surfaces indicated by the European Centre for Medium Range Weather Forecast (ECMWF) dataset, ROCINN retrieves effective cloud pressure and cloud albedo values assuming a cloud fraction of 0. These values—the effective scene pressure and scene albedo values, which are added in the version 2.1 processor—are applied in the NO\textsubscript{2} AMF calculation when the difference between the scene pressure and the surface pressure is less than 2% and the observation is considered to be nearly cloud-free (van der A et al., 2020).

Figure 12 compares the OCRA cloud fractions from the version 1.x (1.0 and 1.1) and 2.1 processors for orbit 6939 on 14 February 2019. The update of cloud-free background maps increases the cloud fractions by more than 0.1 for large cloud fraction values and reduces the values by more than 0.1 for snow/ice covers, e.g., over the Alps and the Ore Mountains. Figure 13 shows the OCRA cloud fractions as a function of VZA. Mainly due to the improved updated VZA correction, the overestimation of cloud fractions, particularly at the east side of the orbit swath, are corrected by more than 0.03 for the new version 2.1 processor.

Figure 14 compares the ROCINN_CRB cloud pressures from the version 1.x and 2.1 processors for orbit 6939 on 14 February 2019. The cloud pressure differences are generally small for optically thin clouds with small cloud fractions. Due to the enlarged OCRA cloud fractions for relatively thick clouds in Fig. 12, the new ROCINN shows increased cloud pressures for large cloud fractions. The decreased cloud pressures e.g., over the Adriatic Sea is related to the reduction of surface albedo. Similar variations are observed for ROCINN_CAL cloud top and base pressures.

Figure 15 shows the tropospheric NO\textsubscript{2} columns for orbit 6939 on 14 February 2019 and the effect of upgrading the processor for OCRA/ROCINN processor from version 1.x to version 2.1. The tropospheric NO\textsubscript{2} columns reduce by more than $5 \times 10^{14}$.
Figure 13. VZA dependency of OCRA cloud fractions from the version 1.x and 2.1 processors (normalized to nadir) in February 2019. Only measurements with $0 < \text{cloud fraction} \leq 0.4$ are included. The VZA is defined negative for observations on the west side of the orbit swath.

Figure 14. ROCINN_CRB cloud pressures from the version 2.1 processor (left) and comparisons with version 1.x (right) over Europe for orbit 6939 on 14 February 2019. Only measurements with $0 < \text{cloud fraction} \leq 0.4$ are included.

molec/cm$^2$ for the edge of the swath, such as the Po Valley, Rome and Naples in Italy, and reduce by up to $3.5 \times 10^{14}$ molec/cm$^2$ for snow/ice scenarios, for instance the Ore Mountains. For polluted areas with optically thicker clouds (cloud fraction larger than 0.15 and cloud pressure larger than 700 hPa), e.g. northern Germany and the Benelux, the tropospheric NO$_2$ columns
increase by more than $1 \times 10^{15}$ molec/cm$^2$, because the increase in cloud fraction (and thus cloud radiance fraction) makes the retrieval less sensitive to the NO$_2$ below the cloud. The NO$_2$ reductions over the Po Valley are related to an additional check implemented in the version 2.1 processor: the pixel is assumed to be cloud-free for the almost clear-sky condition (cloud fraction < 0.1) with the retrieved cloud height very close to the surface height (difference < 100 m). This correction improves the data yield of the TROPOMI cloud products compared with other satellite cloud products; the performance of this correction under different surface conditions (dark, bright, snow, ice) or under presence of different types of low-level aerosols (fog, smoke, dust, ash) is under investigation.

### 4.3.2 CAL cloud model

The cloud correction in our TROPOMI NO$_2$ retrieval is improved using the CAL model from the ROCINN cloud algorithm (Loyola et al., 2018). The CAL model, which regards the clouds as optically uniform layers of light-scattering water droplets, is more representative of the real situation than the CRB model, which treats the clouds as idealized Lambertian reflectors with zero transmittance. The CAL model considers the multiple scattering of light inside the cloud and the contribution of the atmospheric layer between the cloud bottom and the ground.

Figure 16 presents the box-AMFs for clear and cloudy sky calculated using the CRB and CAL cloud models over Munich in Germany (48.15°N, 11.57°E) on 5 February 2019. The cloud pressures and the calculated tropospheric AMFs are also reported. Compared to the clear-sky box-AMFs, the cloudy-sky values increase above the cloud layer (albedo effect) and decrease below the cloud layer (shielding effect). The CRB-based cloud retrieval generally shows a cloud height (pressure) close to the altitude of the middle (Ferlay et al., 2010; Richter et al., 2015), because CRB neglects the oxygen absorption within a cloud layer (Vasilkov et al., 2008) and misinterprets the smaller top-of-atmosphere reflectance as a lower cloud layer (Saiedy et al., 1967).
Figure 16. Box-AMFs for clear and cloudy sky using the ROCINN_CRB and ROCINN_CAL cloud models over Munich in Germany (48.15°N, 11.57°E) on 5 February 2019. The calculated tropospheric AMF is given in the bracket next to each label in the legend. The ROCINN_CRB cloud top pressure is shown as a blue horizontal dotted line, and the ROCINN_CAL cloud top and base pressures are shown as brown horizontal dotted lines.

Compared to the CRB-based cloud correction, the use of CAL model considers the sensitivities inside and below the cloud layers and increases the tropospheric AMFs by 0.09 (13.2%) for Munich.

Figure 17 presents the monthly average differences in the tropospheric NO\textsubscript{2} columns retrieved using the ROCINN_CRB and ROCINN_CAL cloud models in February and August 2019. The use of CAL cloud correction decreases the tropospheric NO\textsubscript{2} columns by more than $1 \times 10^{15}$ molec/cm\textsuperscript{2} (18%) for polluted regions in winter, when most of the NO\textsubscript{2} concentrations are located at the surface (as shown in Fig. 9) and the cloud fractions are generally larger due to the seasonal variation of clouds. The effect is less than $5 \times 10^{14}$ molec/cm\textsuperscript{2} (10%) for summer.

5 TROPOMI tropospheric NO\textsubscript{2} measurements

5.1 Examples of TROPOMI tropospheric NO\textsubscript{2} measurements

Table 3 summarises the tropospheric NO\textsubscript{2} retrieval results for the step-by-step updates listed above. Figure 18 shows the tropospheric NO\textsubscript{2} columns over Europe retrieved using the improved algorithm in February and August 2019. The tropospheric NO\textsubscript{2} columns are higher than $5 \times 10^{15}$ molec/cm\textsuperscript{2} over urban and industrial areas in winter, such as the Po Valley, Germany’s Ruhr region, the Benelux, South-East England, and Turkey’s Marmara region. City-size polluted regions, e.g. around Paris,
Figure 17. Differences in the tropospheric NO$_2$ columns retrieved using the ROCINN_CRB and ROCINN_CAL cloud models over Europe in February and August 2019. Only measurements with cloud radiance fraction less than 0.5 are included.

Table 3. Main settings for the step-by-step improvements and results of the tropospheric NO$_2$ retrievals for the different steps of the updates for Munich (48.15°N, 11.57°E). The tropospheric NO$_2$ columns (VCD$_{trop}$) are given in absolute values (molec/cm$^2$), and the percentage numbers in the brackets are changes relative to the reference case. The uncertainties in the tropospheric NO$_2$ columns (VCD$_{tropErr}$) are given in relative numbers.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sect. 2.3</td>
<td>OMI LER climatology</td>
<td>TM5-MP</td>
<td>OCRA/ROCINN_CRB version 1.x</td>
<td>1.02 × 10$^{16}$</td>
<td>3.86 × 10$^{15}$</td>
<td>65%</td>
<td>37%</td>
</tr>
<tr>
<td>Sect. 4.1</td>
<td>TROPOMI GE_LER data</td>
<td>TROPOMI GE_LER data</td>
<td>OCRA/ROCINN_CRB version 1.x</td>
<td>1.07 × 10$^{16}$</td>
<td>4.03 × 10$^{15}$</td>
<td>62%</td>
<td>41%</td>
</tr>
<tr>
<td>Sect. 4.2</td>
<td>TROPOMI GE_LER data</td>
<td>TROPOMI GE_LER data</td>
<td>OCRA/ROCINN_CRB version 1.x</td>
<td>1.51 × 10$^{16}$</td>
<td>5.02 × 10$^{16}$</td>
<td>57%</td>
<td>37%</td>
</tr>
<tr>
<td>Sect. 4.3.1</td>
<td>TROPOMI GE_LER data</td>
<td>TROPOMI GE_LER data</td>
<td>OCRA/ROCINN_CRB version 2.1</td>
<td>1.57 × 10$^{16}$</td>
<td>5.13 × 10$^{16}$</td>
<td>56%</td>
<td>36%</td>
</tr>
<tr>
<td>Sect. 4.3.2</td>
<td>TROPOMI GE_LER data</td>
<td>TROPOMI GE_LER data</td>
<td>OCRA/ROCINN_CAL version 2.1</td>
<td>1.43 × 10$^{16}$</td>
<td>4.59 × 10$^{16}$</td>
<td>53%</td>
<td>35%</td>
</tr>
</tbody>
</table>
Madrid, Rome, Athens, and Moscow, are captured by the TROPOMI NO\textsubscript{2} measurements. NO\textsubscript{2} emissions over the shipping routes, e.g. the maritime connection between the Iberian Peninsula and North Africa, as well as emissions over the highways, e.g. the main East-West thoroughfare in Austria, are also detected.

Figure 19 compares the tropospheric NO\textsubscript{2} columns retrieved using the reference algorithm and the improved algorithm over Europe in February and August 2019. The tropospheric NO\textsubscript{2} columns are on average enhanced by $2 \times 10^{15}$ molec/cm\textsuperscript{2} (37\%) in winter and $8 \times 10^{14}$ molec/cm\textsuperscript{2} (15\%) in summer mainly due to the combined effect of the improvements in the AMF.
calculation. Larger differences by more than $3 \times 10^{15}$ molec/cm$^2$ are noticed in polluted regions, such as London, Paris, and the Po Valley, as well as shipping lanes, e.g. in the Mediterranean Sea.

5.2 Uncertainty estimates

Derived by uncertainty propagation (Boersma et al., 2004), the overall uncertainty on the tropospheric NO$_2$ column is directly related to the main retrieval steps, which are performed independently and assumed to be uncorrelated. The slant column uncertainty, estimated following a statistical method (Boersma et al., 2007) based on the spatial variability in the slant columns over the Pacific Ocean ($20^\circ$S-$20^\circ$N, $160^\circ$E-$180^\circ$E), is on average $4.5 \times 10^{14}$ molec/cm$^2$, in agreement with van Geffen et al. (2020a). The uncertainty in the stratospheric columns is $3.5 \times 10^{14}$ molec/cm$^2$ for polluted conditions based on the daily synthetic data (see Sect. 3.2) and $1 \times 10^{14}$ molec/cm$^2$ for monthly averages.

The tropospheric AMF calculation, which is the largest source of NO$_2$ uncertainty for polluted scenarios (Lorente et al., 2017), is mainly dependent on surface albedo, a priori NO$_2$ profile, cloud fraction, and cloud pressure, as introduced in Sect. 2.3 and 4. The tropospheric AMF uncertainties are calculated based on uncertainty propagation (Boersma et al., 2004) and typical uncertainties of each parameter (De Smedt et al., 2018, Table 8 therein). The uncertainty contribution from the a priori NO$_2$ profile is practically described by a parameter referred to as profile height, defined as the altitude (pressure) below which resides 75% of the integrated NO$_2$ profile (De Smedt et al., 2018). The estimated parameter uncertainties considered in the AMF uncertainty budget include 0.0016 for GE_LER albedo (Loyola et al., 2020b), 75 hPa for POLYPHEMUS/DLR profile (estimated from the profile height standard deviation), 0.05 for the OCRA cloud fraction, and 50 hPa ROCINN cloud pressure (Loyola et al., 2020a).

Figure 20 shows the estimated tropospheric AMF uncertainties due to the errors in the surface albedo, cloud pressure, and a priori NO$_2$ profile. The uncertainty contribution from the a priori NO$_2$ profile is practically described by a parameter referred to as profile height, defined as the altitude (pressure) below which resides 75% of the integrated NO$_2$ profile (De Smedt et al., 2018). As the satellite measurements are normally filtered for cloud radiance fraction smaller than 0.5 or cloud fraction smaller than $\sim$0.2, the uncertainties related to the cloud fraction are generally smaller than 15% (not shown). From Fig. 20, larger uncertainties are found for small albedo values and for scenarios with large albedo biases such as new snow/ice coverage. The uncertainties due to the cloud pressure and a priori NO$_2$ profile can be up to 70% when the cloud is located below or within the NO$_2$ layer, particularly for thick clouds at low altitudes and for polluted situations (large profile heights).

The presence of aerosols can affect the sensitivity to tropospheric NO$_2$, depending on the particle properties and the NO$_2$ and aerosol vertical distribution (Martín et al., 2003; Leitão et al., 2010). The aerosol effect is not explicitly corrected in this study assuming that the effective cloud parameters from OCRA/ROCINN have partly accounted for the effect of aerosols on the light paths (Boersma et al., 2004, 2011). In comparison to the simple CRB-based cloud correction, which can not fully describe the effects inherent to aerosol particles (Chimot et al., 2019), the use of CAL cloud correction considers the sensitivities inside and below the cloud/aerosol layers and reduces the AMF errors by more than 10% (Liu et al., 2020c).

Note that the use of averaging kernel, which describes the vertical sensitivity of measurements of NO$_2$ concentrations, can remove the uncertainty contributed by the a priori NO$_2$ profile for applications such as data assimilation and validation study.
Figure 20. Tropospheric AMF uncertainties related to the surface albedo, cloud pressure, and a priori NO\textsubscript{2} profile errors. By default, the surface pressure is 1050 hPa, the surface albedo is 0.05, the profile height is 840 hPa, the cloud pressure is 840 hPa, the cloud fraction is 0.2. The definition of profile height is given in the text.

(Eskes and Boersma, 2003). Therefore, for a typical polluted scene, the tropospheric AMF uncertainty is estimated to be 20% for mostly clear sky and 50% in the presence of clouds, leading to a total uncertainty in the tropospheric NO\textsubscript{2} columns in the 30-60% range.

6 TROPOMI tropospheric NO\textsubscript{2} validation

The validation of the improved TROPOMI tropospheric NO\textsubscript{2} columns is based on ground-based MAX-DOAS measurements from nine stations in Europe. Table 4 provides the information about the stations, most of which are characterised by urban or suburban polluted conditions with heavy traffic and industrial emissions. For the validation of TROPOMI measurements, the satellite data from January 2018 - June 2020 are filtered for clouds (cloud radiance fraction less than 0.5), and the closest valid pixel within 20 km of the stations is compared to the ground-based MAX-DOAS data, which are linearly interpolated to the TROPOMI overpass time if original data exist within 1 h.

Figure 21 shows the time series and scatter plot of the comparison of the daily and monthly means between the improved TROPOMI tropospheric NO\textsubscript{2} columns and the ground-based MAX-DOAS measurements in Uccle. The monthly mean values from the TROPOMI and MAX-DOAS measurements show similar seasonal variations in the tropospheric NO\textsubscript{2} column. Figure 21 includes the statistical information on the Pearson correlation coefficient as well as the slope and intercept obtained with
Table 4. An overview of MAX-DOAS stations contributing to the TROPOMI tropospheric NO\textsubscript{2} validation in this study. More details on the QA4ECV datasets can be found at http://www.qa4ecv.eu/ecvs.

<table>
<thead>
<tr>
<th>Station</th>
<th>Location</th>
<th>Institute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Athen</td>
<td>38.05°N, 23.86°E</td>
<td>IUPB</td>
<td>QA4ECV dataset</td>
</tr>
<tr>
<td>Bremen</td>
<td>53.10°N, 8.85°E</td>
<td>IUPB</td>
<td>QA4ECV dataset</td>
</tr>
<tr>
<td>Cabauw</td>
<td>51.97°N, 4.93°E</td>
<td>KNMI</td>
<td>Vlemmix et al. (2010)</td>
</tr>
<tr>
<td>De Bilt</td>
<td>52.10°N, 5.18°E</td>
<td>KNMI</td>
<td>Vlemmix et al. (2010)</td>
</tr>
<tr>
<td>Mainz</td>
<td>49.99°N, 8.23°E</td>
<td>MPIC</td>
<td>QA4ECV dataset</td>
</tr>
<tr>
<td>Munich</td>
<td>48.15°N, 11.57°E</td>
<td>LMU</td>
<td>Chan et al. (2020)</td>
</tr>
<tr>
<td>Thessaloniki_ciri</td>
<td>40.56°N, 22.99°E</td>
<td>AUTH</td>
<td>Drosoglou et al. (2017), QA4ECV dataset</td>
</tr>
<tr>
<td>Thessaloniki_lap</td>
<td>40.63°N, 22.96°E</td>
<td>AUTH</td>
<td>Drosoglou et al. (2017), QA4ECV dataset</td>
</tr>
<tr>
<td>Uccle</td>
<td>50.80°N, 4.36°E</td>
<td>BIRA-IASB</td>
<td>Gielen et al. (2014), Hendrick et al. (2014), Dim-itropoulou et al. (2020)</td>
</tr>
</tbody>
</table>

Figure 21. Daily and monthly mean time series and scatter plot of TROPOMI (SAT) and MAX-DOAS (GB) tropospheric NO\textsubscript{2} columns (closest valid pixel within 20 km of Uccle). Results are shown for the improved satellite retrieval algorithm.

The robust Theil–Sen estimator (Sen, 1968; Vigouroux et al., 2020). A correlation coefficient of 0.85, a slope of 0.70, and an intercept of $-0.52 \times 10^{15}$ molec/cm\textsuperscript{2} are derived when comparing the monthly mean values.

Figure 22 presents the daily and monthly mean absolute and relative differences of TROPOMI and MAX-DOAS measurements in Uccle. The differences are generally within $1 \times 10^{16}$ molec/cm\textsuperscript{2} with a mean difference of $-2.6 \times 10^{15}$ molec/cm\textsuperscript{2}.
Figure 22. Daily (grey dots) and monthly mean (black dots) absolute and relative TROPOMI (SAT) and MAX-DOAS (GB) time series differences for the Uccle station. Results are shown for the improved satellite retrieval algorithm. The histogram of the daily differences is also given, showing the mean and median difference. The total mean values of absolute and relative monthly differences are given in the bottom-right panel.

The NO$_2$ levels are underestimated by 34% by TROPOMI with a standard deviation of 16%, which is mostly explained by the relatively low sensitivity of spaceborne measurements near the surface, the aerosol shielding effect, and the gradient smoothing effect. These effects are often inherent to the remaining impact of likely explained by comparison errors (such as the gradient smoothing effect, the comparison choices, and the inherent difference in sensitivity), partly by the remaining impact of structural uncertainties (Boersma et al. (2016), such as the impact of the choice of the a priori NO$_2$ profiles and/or the albedo database assumed for the satellite AMF calculations), and by the different measurement types or the specific conditions of the validation sites.

To analyse the gradient smoothing effect for Uccle, TROPOMI measurements for 2018-2020 are aggregated based on an area-weighted tessellation to a resolution of 0.01° × 0.01°, and the systematic variation in tropospheric NO$_2$ columns between the satellite pixel location and the ground-based station position is shown in Fig. 23, following the method from Chen et al. (2009); Ma et al. (2013); Pinardi et al. (2020). From Fig. 23, the smoothing effect is largest for summer (up to 19%), as the NO$_2$ gradients are large due to the shorter lifetime, in agreement with Ma et al. (2013). For the Uccle site, which is located south of Brussels at a distance of ~6 km from the city center, the tropospheric NO$_2$ columns increase by up to 4% outwards until 6 km due to the influence of the surrounding emission sources during summer and autumn. This effect is additionally influenced by the seasonal wind pattern, particularly for winter, when the wind is blowing in the direction of the site from north (Dimitropoulou et al., 2020).
Figure 23. Normalized tropospheric NO₂ columns at increasing colocation radii for the Uccle station, estimated using seasonal mean TROPOMI data in 2018-2020.

Table 5. The Pearson correlation coefficient R, as well as the slope S and intercept I (in $1 \times 10^{15}$ molec/cm$^2$) obtained with the robust Theil–Sen estimator for the monthly TROPOMI tropospheric NO$_2$ product compared to MAX-DOAS data. Stations are ordered by increasing mean difference. Values for the DLR improved algorithm (DLRimp) are given, and the values for the DLR reference algorithm (DLRref) and the KNMI operational product (KNMIop) are reported for comparison.

<table>
<thead>
<tr>
<th>Station</th>
<th>R</th>
<th>S</th>
<th>I</th>
<th>R</th>
<th>S</th>
<th>I</th>
<th>R</th>
<th>S</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DLRimp</td>
<td>DLRref</td>
<td>KNMIop</td>
<td>DLRimp</td>
<td>DLRref</td>
<td>KNMIop</td>
<td>DLRimp</td>
<td>DLRref</td>
<td>KNMIop</td>
</tr>
<tr>
<td>Athens</td>
<td>0.85</td>
<td>0.87</td>
<td>0.81</td>
<td>0.65</td>
<td>0.42</td>
<td>0.45</td>
<td>0.33</td>
<td>0.25</td>
<td>0.33</td>
</tr>
<tr>
<td>De Bilt</td>
<td>0.63</td>
<td>0.72</td>
<td>0.80</td>
<td>0.42</td>
<td>0.25</td>
<td>0.77</td>
<td>1.69</td>
<td>1.31</td>
<td>-0.57</td>
</tr>
<tr>
<td>Thessaloniki_ciri</td>
<td>0.91</td>
<td>0.95</td>
<td>0.91</td>
<td>0.51</td>
<td>0.45</td>
<td>0.26</td>
<td>1.69</td>
<td>1.31</td>
<td>-0.57</td>
</tr>
<tr>
<td>Thessaloniki_lap</td>
<td>0.79</td>
<td>0.86</td>
<td>0.90</td>
<td>0.35</td>
<td>0.30</td>
<td>0.37</td>
<td>1.97</td>
<td>1.09</td>
<td>0.93</td>
</tr>
<tr>
<td>Bremen</td>
<td>0.85</td>
<td>0.71</td>
<td>0.70</td>
<td>0.58</td>
<td>0.44</td>
<td>0.40</td>
<td>-0.24</td>
<td>-0.20</td>
<td>0.79</td>
</tr>
<tr>
<td>Uccle</td>
<td>0.85</td>
<td>0.81</td>
<td>0.44</td>
<td>0.70</td>
<td>0.42</td>
<td>0.27</td>
<td>-0.52</td>
<td>-0.24</td>
<td>2.26</td>
</tr>
<tr>
<td>Cabauw</td>
<td>0.75</td>
<td>0.67</td>
<td>0.43</td>
<td>0.41</td>
<td>0.21</td>
<td>0.33</td>
<td>1.08</td>
<td>1.24</td>
<td>1.98</td>
</tr>
<tr>
<td>Munich</td>
<td>0.57</td>
<td>0.72</td>
<td>0.63</td>
<td>0.35</td>
<td>0.39</td>
<td>0.42</td>
<td>1.44</td>
<td>0.29</td>
<td>1.35</td>
</tr>
<tr>
<td>Mainz</td>
<td>0.85</td>
<td>0.87</td>
<td>0.81</td>
<td>0.39</td>
<td>0.28</td>
<td>0.50</td>
<td>1.65</td>
<td>0.82</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Figure 24. Box and whisker plot of the daily biases and spread of the differences between the TROPOMI (SAT) and MAX-DOAS (GB) data. Results for the reference algorithm, the improved algorithm, and the operational product are compared. Stations are ordered by increasing mean difference. The mean differences are represented by crosses. The median differences are represented by vertical solid lines inside the boxes, which mark the 25 and 75% quantiles. The whiskers cover the 9-91% range of the differences.

Table 6. Similar as Table 5 but for the mean difference (MD, SAT-GB in $1 \times 10^{15}$ molec/cm$^2$), standard deviation (STD, in $1 \times 10^{15}$ molec/cm$^2$), and relative difference (RD, in %) for the monthly TROPOMI tropospheric NO$_2$ product compared to MAX-DOAS data.

<table>
<thead>
<tr>
<th>Station</th>
<th>MD</th>
<th>STD</th>
<th>RD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DLRimp</td>
<td>DLRref</td>
<td>KNMIop</td>
</tr>
<tr>
<td>Athens</td>
<td>-1.6</td>
<td>-3.5</td>
<td>-3.1</td>
</tr>
<tr>
<td>De Bilt</td>
<td>-2.0</td>
<td>-3.5</td>
<td>-2.1</td>
</tr>
<tr>
<td>Thessaloniki_ciri</td>
<td>-2.2</td>
<td>-3.4</td>
<td>-3.8</td>
</tr>
<tr>
<td>Thessaloniki_lap</td>
<td>-2.4</td>
<td>-3.7</td>
<td>-3.5</td>
</tr>
<tr>
<td>Bremen</td>
<td>-2.6</td>
<td>-3.7</td>
<td>-3.1</td>
</tr>
<tr>
<td>Uccle</td>
<td>-2.6</td>
<td>-4.5</td>
<td>-4.3</td>
</tr>
<tr>
<td>Cabauw</td>
<td>-3.2</td>
<td>-4.7</td>
<td>-2.9</td>
</tr>
<tr>
<td>Munich</td>
<td>-3.4</td>
<td>-4.6</td>
<td>-3.3</td>
</tr>
<tr>
<td>Mainz</td>
<td>-4.4</td>
<td>-5.6</td>
<td>-5.1</td>
</tr>
</tbody>
</table>

Similar figures as Fig. 21 and 22 for the improved and reference algorithms are gathered in Fig. S1 - S4 in the Supplement for all the stations. Figure 24 shows an overview of the daily differences between satellite and ground-based data for the improved
Tables 5 and 6 summarize the monthly comparisons of TROPOMI and MAX-DOAS measurements. High correlations are observed for the improved algorithm for all the stations with an average correlation coefficient of 0.78. The impact of the algorithm improvements leads to a decrease of the mean absolute difference in urban/suburban conditions from $-4.13 \times 10^{15}$ molec/cm$^2$ to $-2.71 \times 10^{15}$ molec/cm$^2$ and relative difference from -55.3% to -34.7%. The largest absolute bias ($-5.6 \times 10^{15}$ molec/cm$^2$ in Mainz) is reduced to $-4.4 \times 10^{15}$ molec/cm$^2$ (relative bias from -60% to -40%), while the smaller absolute bias ($-3.4 \times 10^{15}$ molec/cm$^2$ in Thessaloniki_ciri) is reduced to $-2.2 \times 10^{15}$ molec/cm$^2$ (relative bias from -54% to -34%). The largest reduction is found for Athens (-27% reduction from the reference to improved algorithm).

Smaller biases are found for the improved algorithm, not only in comparison with the reference algorithm but also compared to the operational product in Fig. 24 and Table 6, particularly for Athens, Thessaloniki_ciri, Thessaloniki_lap, Uccle, and Mainz. The relative biases ranging from -26 to -45% in Table 6 are lower than those reported by validation exercises for the operational TROPOMI product, where the NO$_2$ levels are normally found to be underestimated by the TROPOMI instrument by 30% to 50% for polluted conditions (Dimitropoulou et al., 2020; Verhoelst et al., 2020; Wang et al., 2020). These results are not directly comparable to results e.g. obtained by Dimitropoulou et al. (2020), as they use a more elaborated ground-based dataset with several pointing directions and specific area-weighted pixel selections in the MAX-DOAS line-of-sight. Note that the operational tropospheric NO$_2$ columns have been increased with an upgrade of the NO$_2$ processor (version 01.04.00) since 29 November 2020 due to the improved handling of cloud pressures (Eskes et al., 2020).

To investigate the impact of satellite a priori NO$_2$ profiles on the comparison, the satellite averaging kernel (see Sect. 5.2) is used to relate the MAX-DOAS retrieved NO$_2$ profiles to satellite column measurements by calculating the smoothed MAX-DOAS columns as:

$$ V_{GB,smoothed} = \sum_l AK_{SAT,l} \times x_{GB,l}. $$

The smoothed MAX-DOAS NO$_2$ columns $V_{GB,smoothed}$ are derived by convolving the layer ($l$)-dependent daily profile $x_{GB,l}$ (expressed in partial columns and interpolated to the satellite overpass time) with the satellite averaging kernel $AK_{SAT,l}$.

Figure 25 shows the original and smoothed comparisons of satellite and MAX-DOAS data for the Munich station. The use of the averaging kernel smoothing reduces the MAX-DOAS columns and thus improves the agreement between the satellite and MAX-DOAS columns. When the satellite averaging kernels are used to remove the contribution of the a priori NO$_2$ profile shape, the mean absolute difference reduces from $-3.4 \times 10^{15}$ molec/cm$^2$ to $-1.9 \times 10^{15}$ molec/cm$^2$, and the relative difference reduces from -39% to -23%.

7 Conclusions

The DLR retrieval algorithm developed for TROPOMI NO$_2$ measurements follows a three-step scheme. To calculate the NO$_2$ slant columns, a 405-465 nm fitting window is applied in the DOAS fit for consistency with other NO$_2$ retrievals from OMI and TROPOMI. Absorption cross-sections of interfering species and a linear intensity offset correction are applied. The striping
Figure 25. Daily (dots) and monthly mean (circles) time series of TROPOMI and MAX-DOAS tropospheric NO\textsubscript{2} columns for the Munich station. Results for the original comparisons and the smoothed comparisons are reported.

The stratospheric NO\textsubscript{2} component is estimated using the STREAM method, which requires no chemistry transport model data as used in data assimilation and provides an improved treatment of polluted and cloudy pixels comparing to other modified reference sector methods. An improved DSTREAM method is used to correct for the VZA dependency of stratospheric NO\textsubscript{2} for high latitudes, which is related to the local time changes across the orbit. DSTREAM divides the orbit swath into three segments, applies the original STREAM to data from each of the segments, and calculates the stratospheric NO\textsubscript{2} column based on VZA and latitude from the satellite measurement. Applied to synthetic TROPOMI data, constructed using the IFS(CB05BASCOE) model fields, the estimated stratospheric NO\textsubscript{2} columns from the original STREAM and the improved DSTREAM show good consistency with the a priori truth. Applied to actual TROPOMI measurements, STREAM and DSTREAM successfully separate the stratospheric and tropospheric fields for polluted regions. The VZA dependency of stratospheric NO\textsubscript{2} which amounts up to $2 \times 10^{14}$ molec/cm\textsuperscript{2} at high latitudes is captured by DSTREAM.

In the tropospheric AMF calculation, the surface albedo from the monthly OMI LER climatology is replaced by the TROPOMI GE\_LER data, which is consistently applied in both NO\textsubscript{2} and cloud retrievals. GE\_LER in the NO\textsubscript{2} fitting window is retrieved using the machine learning based approach FP\_ILM with inputs from the DOAS fitting. In comparison with the climatological LER values from previous satellite missions, the GE\_LER data relies on the real-time measurements from...
the TROPOMI instrument itself with an improved spatial resolution of $0.1^\circ \times 0.1^\circ$. Therefore, GE\_LER better characterizes the actual surface conditions with an impact on the tropospheric NO$_2$ columns by up to $3 \times 10^{15}$ molec/cm$^2$ under polluted conditions.

Mesoscale-resolution a priori profiles ($0.2^\circ \times 0.3^\circ$), obtained from the regional POLYPHEMUS/DLR chemistry transport model based on the European TNO-MACC emission inventory, provide a better description of the spatial variability in the NO$_2$ fields for Europe. Compared to the currently used TM5-MP profiles, the POLYPHEMUS/DLR profiles generally show higher surface NO$_2$ concentrations, which reduce the tropospheric AMFs and thus enhance the tropospheric NO$_2$ columns by more than $2 \times 10^{15}$ molec/cm$^2$ for polluted regions.

The presence of clouds is considered using the TROPOMI operational cloud retrieval algorithms OCRA/ROCINN. In a new version 2.1 processor, OCRA separates a spectral scene (in the UV-VIS wavelength range) into cloudy contribution and cloud-free background using TROPOMI-based background maps ($0.1^\circ \times 0.1^\circ$) instead of OMI-based ones, and ROCINN applies the surface albedo from the GE\_LER data in the TROPOMI NIR instead of a static climatology. The overestimation of cloud fractions at the swath edge is corrected. Larger differences in cloud fractions and cloud pressures are found for relatively thick clouds, which affect the tropospheric NO$_2$ columns by more than $1 \times 10^{15}$ molec/cm$^2$. In the tropospheric AMF calculation, the CRB model from ROCINN, in which clouds are idealized Lambertian reflectors, is replaced with the CAL model, in which clouds are represented by uniform layers of water droplets. CAL is more representative of the real situation and preferred for small TROPOMI ground pixels and for low clouds. The application of CAL cloud parameters considers the sensitivities inside and below the cloud layers and reduces the tropospheric NO$_2$ columns by more than $1 \times 10^{15}$ molec/cm$^2$ for polluted regions.

The uncertainty in the NO$_2$ slant columns is $4.5 \times 10^{14}$ molec/cm$^2$, derived from the spatial variability over the Pacific Ocean. The uncertainty in the stratospheric columns is $3.5 \times 10^{14}$ molec/cm$^2$ for polluted regions based on daily synthetic TROPOMI data. The tropospheric AMF uncertainty is estimated to be 20% for mostly clear sky and 50% in the presence of clouds, leading to a total uncertainty in the tropospheric NO$_2$ column in the 30-60% range.

Validation of the improved TROPOMI tropospheric NO$_2$ columns is performed by comparisons with ground-based MAX-DOAS measurements. The validation is illustrated for nine European stations with urban/suburban conditions. The improved data shows a similar seasonal variation in the tropospheric NO$_2$ columns as the MAX-DOAS measurements with an average correlation coefficient of 0.78. Compared to the reference data, the improved algorithm shows a significant improvement with absolute differences decreasing from $-4.13 \times 10^{15}$ molec/cm$^2$ to $-2.71 \times 10^{15}$ molec/cm$^2$ on average and relative differences from -55.3% to -34.7%. When the satellite averaging kernels are used to remove the contribution of a priori NO$_2$ profile shape, the absolute difference at the Munich station reduces from $-3.4 \times 10^{15}$ molec/cm$^2$ to $-1.9 \times 10^{15}$ molec/cm$^2$, and the relative difference reduces from -39% to -23%.

The TROPOMI NO$_2$ research product from DLR is a complement to the operational product due to the use of independent approaches for stratosphere-troposphere separation and AMF calculation. Comparing to the other regional TROPOMI NO$_2$ product listed in Table 1, the DLR European retrieval reduces the potential biases introduced by using inputs from different instruments or climatologies and confirms the importance of applying more realistic input parameters with better resolution for AMF calculation.
In the future, the spectral effect of extending the fitting window to 490 nm will be analysed, when the pixel blooming is better treated in a future update of the TROPOMI level 0-1b processor. An update of the POLYPHEMUS/DLR model using more recent TNO-MACC emission is planned. The operational OCRA/ROCINN cloud parameters will be compared with other cloud products such as FRESCO-S and MICRU. The interpretation of the cloud product for aerosol-dominated scenes and the impact on the NO$_2$ retrieval algorithm will be further investigated. Aerosol contamination will be removed in the GE_LER retrieval using TROPOMI aerosol index data. An independent Neural Network will be trained and implemented for error estimation.

The NO$_2$ data quality will be further analysed using data from additional ground-stations covering different pollution conditions and data from validation campaigns with independent instruments. Further improvements in the ground-based validation include using the full MAX-DOAS line-of-sight sensitivity and the intersect with the TROPOMI pixel(s) or having more ground-based instruments located within a TROPOMI pixel. More frequent ground-based measurements and measurements in more than one direction might better sample the temporal and spatial variability around the measurement sites (Richter and Lange, 2021; Dimitropoulou et al., 2020).

The NO$_2$ retrieval algorithm can be adapted for new instruments and missions, such as the polar-orbiting Sentinel-5 and geostationary Sentinel-4 missions, which offer new perspectives for monitoring NO$_2$ with a fast revisiting time and a high spatial resolution and provide information on atmospheric variables in support of European policies.

Data availability. The TROPOMI NO$_2$ datasets used in the study are available upon request. The DLR TROPOMI NO$_2$ product are published as HDF version 5 files. For each ground pixel, the TROPOMI data product provides the retrieval results (e.g., the slant column, stratospheric column, and tropospheric column of NO$_2$), input information (such as the GE_LER surface albedo, POLYPHEMUS/DLR a priori NO$_2$ profiles, and OCRA/ROCINN cloud parameters), uncertainty estimate, processing quality flag, and averaging kernel.

Author contributions. SL and PV developed the retrieval framework. SL processed the data, analysed the results, and contributed to the GE_LER processing and MAX-DOAS validation. SL and GP analysed the MAX-DOAS validation results. JX processed the GE_LER data. AA and RL provided expertise regarding the OCRA/ROCINN-based cloud correction. KLC contributed to uncertainty analysis. SB contributes to the STREAM development and improvement. EK and FB provided the POLYPHEMUS/DLR model data. VH provided the IFS(CB05BASCOE) model data. AB, KLC, SD, SD, MG, FH, DK, KL, AP, JR, AR, MVR, TW, and MW provided the MAX-DOAS measurements. DL is the main developer of the TROPOMI cloud products based on the CAL model and the GE_LER retrievals used in this study. SL, PV, JX, and GP wrote the paper with comments from all authors.

Competing interests. The authors declare that they have no conflict of interest.
Acknowledgements. This work is funded by the DLR/DAAD Research Fellowships - Postdocs programme (57478192). The TROPOMI NO\textsubscript{2} data generated at DLR are used in the S-VELD project, which is financed by the Federal Ministry of Transport and Digital Infrastructure to analyse the effect of traffic emission on air quality in Germany (grant no. 19F2065). Part of the results discussed in this paper were achieved within the JOSEFINA project funded by the Bavarian State Ministry for Environment and Consumer Protection. The POLYPHMEMUS/DLR model was developed with support by the 7\textsuperscript{th} EU Framework Program (PASODOBLE project grant 241557). We acknowledge the Belgian Science Policy Office (BELSPO) for supporting part of this work through the PRODEX programme B-ACSAF project. Part of this work is carried out in the framework of the S5p Validation Team (S5PVT) AO projects NIDFORVAL (ID #28607, PI G. Pinardi, BIRA-IASB). Part of this work is supported by the DFG Major Research Instrumentation Programme (grant no. INST 86/1499-1 FUGG). We are grateful to QA4ECV for the generation of harmonized MAX-DOAS datasets. We thank EU/ESA/KNMI/DLR for the provision of the TROPOMI/S5P level 1 products. We thank DLR colleagues for developing the Universal Processor for UV/Vis Atmospheric Spectrometers (UPAS) system used for generating level 2 products from TROPOMI. This paper contains modified Copernicus Sentinel data processed by DLR.
References


Huang, G. and Sun, K.: Non-negligible impacts of clean air regulations on the reduction of tropospheric NO\textsubscript{2} over East China during the COVID-19 pandemic observed by OMI and TROPOMI, Sci. Total Environ., 745, 141 023, 2020.


Loyola, D., Pedergnana, M., and García, S. G.: Smart sampling and incremental function learning for very large high dimensional data, Neural Netw., 78, 75–87, 2016.


