https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

Improved NO_2 spectral fits for TROPOMI and OMI by removing wavelengths around 430 nm

Jos van Geffen[1], Henk Eskes[1], Maarten Sneep[1], Mark ter Linden[1], and J. Pepijn Veefkind[1,2]

[1] Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands
[2] Delft University of Technology (TUDelft), Delft, The Netherlands

Correspondence: Jos van Geffen (geffen@knmi.nl)

Abstract.

The Fraunhofer absorption feature at $430\,\mathrm{nm}$ influences the retrieval of nitrogen dioxide (NO_2) from measurements by satellite-based instruments such as the Tropospheric Monitoring Instrument (TROPOMI) and Ozone Monitoring Instrument (OMI). The width and depth of the feature in the measured spectrum are affected by rotational Raman scattering (RRS) throughout the atmosphere and by vibrational Raman scattering (VRS) in open water bodies. RRS, or the Ring-effect, is accounted for in the Differential Optical Absorption Spectroscopy (DOAS) retrieval of the NO_2 slant column density (SCD) by means of a scalable reference spectrum, which will not fully pick up the variation of the depth of the $430\,\mathrm{nm}$ feature with the solar activity cycle. It is not possible to account for VRS with a scalable reference spectrum, since VRS characteristics depend on several aspects, including the viewing geometry and the material dissolved in the water, such as chlorophyll. From detailed inspection of DOAS fit residuals, the difference between the measured and modelled spectra, it is clear that the $430\,\mathrm{nm}$ feature disturbs the NO_2 SCD retrieval.

In this paper we investigate the benefits of removing the wavelength range $428-433\,\mathrm{nm}$ from the DOAS retrieval. This "NO2-gap approach" reduces the SCD error and the RMS error of the fit over water bodies by $10-20\%$ and the fit residual for the remaining parts of the window improves. Over some land scenes, where the residual outside the $428-433\,\mathrm{nm}$ range looks very good, the SCD error and RMS error are reduced by $5-10\%$. For other areas the fit residual does not deteriorate by the NO2-gap approach. Over ocean waters the SCD is seen to decrease by a few percent, which leads to a decrease of the stratospheric NO_2 column of on average up to $-2\,\mu\mathrm{mol\,m^{-2}}$ in the tropics. Over land the change in SCD may be positive or negative by a few percent, which in combination with the change in the stratospheric column leads to changes in the tropospheric NO_2 column of on average $\pm 2\,\mu\mathrm{mol\,m^{-2}}$. These changes are too small to alter the general conclusions of the routine validation of TROPOMI data. Because of the improvement of the SCD error and systematic improvements over open water it has been decided to implement the NO2-gap approach in the forthcoming processor versions of TROPOMI and OMI.

1 Introduction

Tropospheric nitrogen dioxide (NO_2) is a key contributor to air quality issues, as it directly affects human health (WHO, 2003), it is an essential precursor for the formation of ozone (O_3) in the troposphere (Sillman et al., 1990), and it influences

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

OH concentrations and thereby shortens the lifetime of methane (CH_4; Fuglestvedt et al., 1999). Over remote regions with few to no sources, NO_2 is primarily located in the stratosphere, where it is involved in photochemical reactions with ozone, either by acting as a catalyst for ozone destruction (Crutzen, 1970; Seinfeld and Pandis, 2006; Hendrick et al., 2012) or by suppressing ozone depletion (Murphy et al., 1993).

Satellite measurements of NO_2 have provided valuable contributions to world-wide monitioring of air quality (see e.g. Levelt et al., 2018) and estimations of emissions of nitrogen oxides (NO_x = NO_2 + NO; see e.g. van der A et al., 2024). Such measurements usually provide total column amounts of NO_2 which need to be seperated into a stratospheric and a tropospheric contribution. This implies that measurements of NO_2 over large remote areas, such as the oceans, need to be as accurate as possible, not just for the sake of knowing stratospheric NO_2 concentrations but also to sufficiently accurately determine tropospheric NO_2 concentrations over polluted areas in order to, e.g., reliably monitor NO_2 emissions.

The first step in the NO_2 processing of satellite measurements usually is a Differential Optical Absorption Spectroscopy (DOAS; see Platt, 1994; Platt and Stutz, 2008) retrieval in a window in the visible wavelength range around $440\,\mathrm{nm}$ to determine the slant column density (SCD), that is: the total amount of NO_2 along the effective light path from Sun through atmosphere to satellite. This DOAS NO_2 retrieval takes into account absorption by other atmospheric species in the same wavelength window (in particular ozone, water vapour and the O_2-O_2 collision-complex), as well as absorption by liquid water in clear open water bodies.

On its way to the satellite instrument, solar light also may undergo Raman scattering, which leads to filling-in, widening and shifting of Fraunhofer lines in the measured radiance spectrum. In the atmosphere this is primarily rotational Raman scattering (RRS, a.k.a. the Ring effect: inelastic Raman scattering of incoming sunlight by N_2 and O_2 molecules; see Grainger and Ring, 1962; Chance and Spurr, 1997), which is accounted for in the DOAS retrieval by including a scalable reference spectrum determined from a reference solar irradiance spectrum. Vibrational Raman scattering (VRS) occurs in clear open water bodies (Vasilkov et al., 2002; Vountas et al., 2003; Joiner et al., 2004; Vountas et al., 2007; Dinter et al., 2015; Holtrop et al., 2021), while it is negligible in the atmosphere (Peters et al., 2014). Accounting for VRS in the DOAS retrieval by way of a scalable reference spectrum cannot be done (A. Richter, pers. comm., 2024), since VRS results in light coming into the NO_2 DOAS fit window from a broad wavelength range, and how much light is available in that spectral range depends on a) the incoming spectrum, i.e. on viewing geometry and cloudiness, and b) the type and concentration of absorbing substances in the water (e.g. chlorophyll, dissolved organic matter (DOM), etc.; e.g. Vountas et al., 2007; Dinter et al., 2015; Holtrop et al., 2021). Hence, VRS effects may be disturbing NO_2 retrievals over large ocean areas, which in turn will have an impact on the derived stratospheric NO_2 concentrations.

Close inspection of NO_2 DOAS fit residuals, i.e. the difference between the measured reflectance (which is the ratio between the earth radiance and the solar irradiance; see Sect. 2.2.1) and the modelled reflectance over clear-sky scenes, shows a significant structure around $430\,\mathrm{nm}$. Clearly, some absorption and/or scattering effects taking place along the light path are not accounted for, or at least not sufficiently.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

At 430 nm, the solar irradiance has a large peak related to Fraunhofer lines from iron (Fe, at 430.790 nm [1]) and calcium (Ca, at 430.774 nm), which are broadened by RRS and VRS. In addition, VRS widens two strong Ca+ Fraunhofer lines at 393.368 and 396.847 nm and shifts these to around 430 nm and higher (Peters et al., 2014; Dinter et al., 2015). It thus seems likely that VRS is a large contributor to the residual issue seen around 430 nm. Residuals may also shows broad-band structures above 430 nm, which are likely related to chlorophyll and/or other substances present in the ocean waters (Sect. 6.3). NO_2 slant column retrievals over clear-sky dry land, where VRS certainly does not play a role, may also show remaining structures in the fit residual around 430 nm in case the residual over the rest of the fit window is very small, which seems to indicate that accounting for RRS effects may not be fully accurate, though the impact of this on the resulting NO_2 values is less than when VRS also plays a role.

This paper uses Tropospheric Monitoring Instrument (TROPOMI) and Ozone Monitoring Instrument (OMI) NO_2 retrievals (Sect. 2) to investigate the issues around 430 nm (Sect. 3) and proposes as solution to disable a part of the fit window (Sect. 4). Sect. 5 discusses the impact of this solution on the stratospheric and tropospheric NO_2 columns, while some additional points are discussed in Sect. 6.

2 Data sources and data retrieval

2.1 Data sources

2.1.1 TROPOMI instrument and data versions

The Tropospheric Monitoring Instrument (TROPOMI; Veefkind et al., 2012) is the only instrument aboard ESA's Sentinel-5 Precursor (S5P) spacecraft, which was launched on 13 October 2017 into an ascending sun-synchronous polar orbit with an equator crossing at about 13:30 local time. TROPOMI provides measurements in four channels (UV, visible, NIR and SWIR) of various trace gas columns (such as NO_2, O_3, SO_2, HCHO, CH_4, CO), as well as cloud and aerosol properties. With its full swath width of about 2600 km, TROPOMI achieves global coverage each day, except for narrow strips between orbits of about 0.5° wide at the equator. Across-track, the swath is divided in 450 ground pixels (rows) and their size is 3.6 km at nadir and increases towards the edges; the largest pixels are about 14 km wide. Along-track, the pixel size initially was 7.2 km; as of 6 Aug. 2019 this is reduced to 5.6 km.

TROPOMI NO_2 data is available as of 1 May 2018 up to the present. This paper uses officially released offline (OFFL) and reprocessed (RPRO) data of collection 03 (processor versions v2.4.0 – v2.8.0, documented in van Geffen et al., 2022, 2025), as well as dedicated data made locally with a preliminary version of the upcoming processor v2.9.1. That new version includes the solution proposed in this paper and is operational as of the end of Nov. 2025 and it is going to be the basis of a full mission reprocessing scheduled to take place in 2027.

[1] Fraunhofer line wavelengths mentioned are taken from https://en.wikipedia.org/wiki/Fraunhofer_lines; last access: 20 November 2025

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

2.1.2 OMI instrument and data versions

The Ozone Monitoring Instrument (OMI; Levelt et al., 2006) is one of the instruments aboard NASA's EOS/Aura spacecraft, which was launched on 15 July 2004 into an ascending sun-synchronous polar orbit with an equator crossing at about 13:40 local time. OMI provides measurements in two channels (UV and visible) of various trace gas columns (such as NO_2, O_3, SO_2, HCHO), as well as cloud and aerosol properties. With its full swath width of about 2600 km, OMI achieves global coverage each day. Across-track, the swath is divided in 60 ground pixels (rows) and their size is 24 km at nadir and increases towards the edges; the largest pixels are about 150 km wide. Along-track, the pixel size is 13 km throughout the mission. Since June 2007 a part of the OMI detector suffers from a so-called row anomaly, which appears as signal suppression in the level-1b radiance data at all wavelengths (Schenkeveld et al., 2017), leading, e.g., to large uncertainties on the NO_2 data in the affected rows and hence these rows need to be skipped from the NO_2 analysis.

OMI NO_2 data is available as of 1 October 2004. Currently publicly available OMI NO_2 data version is collection 03, which was processed within the framework of the QA4ECV project (Boersma et al., 2018), covering data from Oct. 2004 up to March 2021. OMI collection 04 data is under development, based on the TROPOMI NO_2 data processor (ATBD in preparation: van Geffen et al., 2026). That version will include the solution proposed in this paper and it will be used for a full mission reprocessing.

2.2 Data retrieval

The NO_2 retrievals of TROPOMI and OMI use the three step approach that was introduced for the OMI NO_2 retrieval and named DOMINO (Boersma et al., 2007, 2011). The first step is a DOAS retrieval to determine the SCD, N_s – see Sect. 2.2.1 for details. Next, NO_2 vertical profile information from the TM5-MP chemistry transport model / data assimilation system that assimilates the SCDs is used to determine the stratospheric vertical column density (VCD), N_v^{strat}. Finally, the tropospheric VCD, N_v^{trop}, is determined using stratospheric and tropospheric air-mass factors (AMFs), which depend on surface albedo, surface pressure, cloud fraction, cloud pressure, the shape of the NO_2 vertical profile (not of the absolute concentration levels), and the viewing geometry of the satellite ground pixel in question. A description of the last two steps falls outside the scope of this paper; for details see van Geffen et al. (2020, 2025, 2026). Since the SCD depends strongly on the along-track and across-track variation in the solar and viewing zenith angles, it is often easier to consider the geometric column density (GCD), N_v^{geo}, defined as the SCD divided by the geometric AMF, which depends only on the viewing geometry.

To make the use of the TROPOMI data easier, a so-called `qa_value` (where 'qa' stands for 'quality assurance') is assigned to each ground pixel, which serves as an easy filter of the NO_2 observations. The usage of the `qa_value` is detailed in the Product User Manual (PUM; Eskes et al., 2024). For most applications, the recommended filter is `qa_value` > 0.75, which removes scenes with large cloud fractions (cloud radiance fraction > 0.5) and snow/ice. In this paper "cloud-free", "clear-sky" or just "clear" thus refers to `qa_value` > 0.75, while "cloudy" refers to $0.50 <$ `qa_value` < 0.75. Details of how the NO_2 `qa_value` is constructed are given in the ATBD (van Geffen et al., 2025, App. E).

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

2.2.1 NO_2 slant column retrieval

In the DOAS retrieval of the NO_2, the difference between the measured reflectance and a modelled reflectance is minimised. Details of the TROPOMI and OMI implementation are given in van Geffen et al. (2020, 2022, 2025, 2026); in short it is as follows.

The measured reflectance is the ratio between the radiance at the top of the atmosphere (I) and the the solar irradiance (E_0) measured by the same instrument:

$$R_{\mathrm{meas}}(\lambda) = \frac{\pi\, I(\lambda)}{\mu_0\, E_0(\lambda)} \tag{1}$$

where $\mu_0 = \cos(\theta_0)$ is the cosine of the solar zenith angle (SZA). For TROPOMI the daily irradiance is used, while for OMI the irradiance averaged over all the daily 2005 measurements is used.

The modelled reflectance is given by:

$$R_{\mathrm{mod}}(\lambda) = P(\lambda) \cdot \exp\left[-\sum_{k=1}^{n_k} \sigma_k(\lambda) \cdot N_{\mathrm{s},k}\right] \cdot \left(1 + C_{\mathrm{ring}} \frac{I_{\mathrm{ring}}(\lambda)}{E_0(\lambda)}\right) \tag{2}$$

where:

- $P(\lambda) = \sum a_m\, \lambda^m$ $(m = 0, 1, \ldots, n_p = 5)$ is a 5-th order polynomial, which accounts for spectrally smooth structures resulting from molecular (single and multiple) scattering and absorption, aerosol scattering and absorption, and surface albedo effects;
- $\sigma_k(\lambda)$ the reference spectrum (a.k.a. cross section) and $N_{\mathrm{s},k}$ the slant column amount of molecule $k = 1, \ldots, n_k = 5$ included in the fit: NO_2, O_3, water vapour (H_2O_{vap}), liquid water (H_2O_{liq}) and the $O_2{-}O_2$ collision complex, where all $\sigma_k(\lambda)$ have been convolved with the instrument spectral response function (ISRF, a.k.a. slit funtion) – see Fig. 1;
- C_{ring} is the Ring fit coefficient and $I_{\mathrm{ring}}(\lambda)/E_0(\lambda)$ the sun-normalised synthetic Ring spectrum (see Fig. 1), with I_{ring} constructed following Chance and Spurr (1997) from a reference irradiance spectrum, E_{ref} (Dobber et al., 2008);

which makes a total of 12 fit parameters. If including VRS in the DOAS fit with a scalable reference spectrum were possible, this could be done either as a σ_{VRS}-term in the summation term of Eq. (2) or as a non-linear term similar to the Ring-term with an I_{VRS} spectrum.

The DOAS retrieval then amounts to minimisation of the chi-squared merit function:

$$\chi^2 = \sum_{i=1}^{n_\lambda} \left(\frac{R_{\mathrm{resid}}(\lambda_i)}{\Delta R_{\mathrm{meas}}(\lambda_i)}\right)^2 \tag{3}$$

where $R_{\mathrm{resid}}(\lambda)$ is the so-called fit residual:

$$R_{\mathrm{resid}}(\lambda) = R_{\mathrm{meas}}(\lambda) - R_{\mathrm{mod}}(\lambda) \tag{4}$$

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

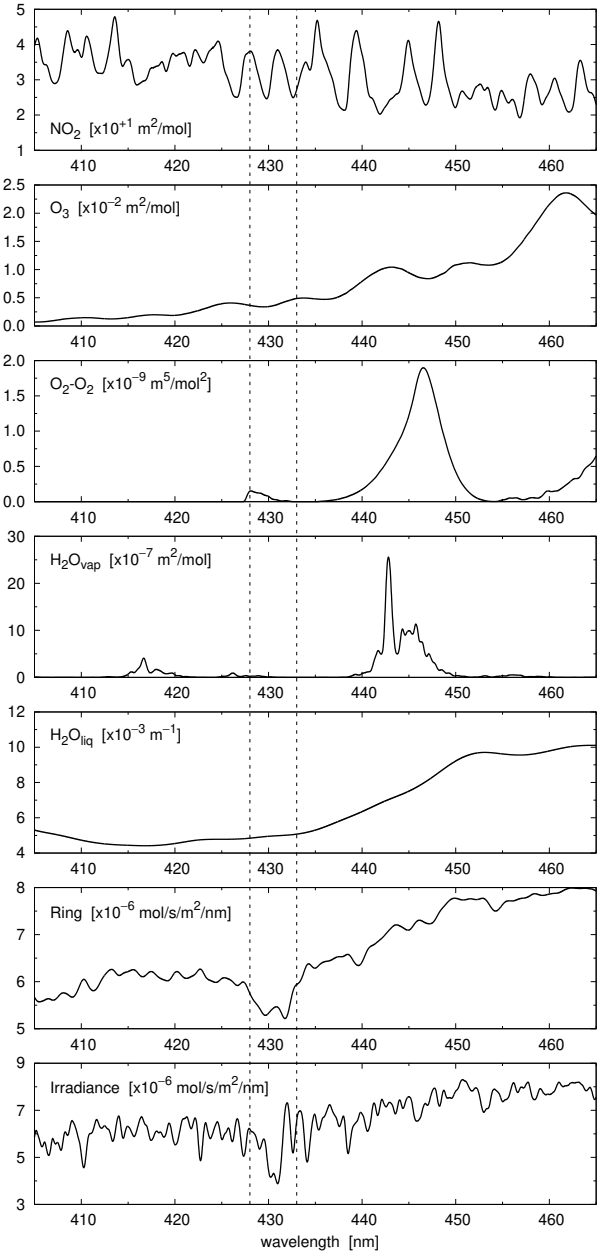

Figure 1. Convolved reference spectra used in the TROPOMI NO_2 SCD retrieval, Eq. (2), as well as the reference irradiance spectrum for detector row 225 within the fit window. The vertical dashed lines indicate wavelengths 428.0 and $433.0\,\mathrm{nm}$. A description of these spectra, including references, is given in the NO_2 ATBDs of TROPOMI and OMI (van Geffen et al., 2025, 2026).

n_λ is the number of wavelengths (spectral pixels) in the fit window ($405-465\,\mathrm{nm}$) and $\Delta R_{\mathrm{meas}}(\lambda_i)$ is the uncertainty on the measured reflectance, which depends on the precision of the radiance and irradiance measurements as given in the level-1b

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

product, i.e. on the signal-to-noise ratio (SNR) of the measurements (Kleipool et al., 2018; Ludewig et al., 2020). The χ^2-minimisation is performed with an Optimal Estimation (OE; based on Rodgers, 2000) routine. For TROPOMI n_λ is 304 or 305, for OMI n_λ is 287 or 288, depending on the row. Both χ^2 and the root-mean-square (RMS) error of the fit:

$$R_{\mathrm{RMS}} = \sqrt{\frac{1}{n_\lambda}\sum_{i=1}^{n_\lambda}\Big(R_{\mathrm{resid}}(\lambda_i)\Big)^2} \tag{5}$$

can be seen as a measure for the goodness of the fit, though both are continuous quantities, hence it is not possible to say where the separation between "good fit" and "bad fit" lies.

Spectral pixels are flagged in case they need to be removed from the fit due to problems encounted in producing the level-1b spectra (e.g. saturation, blooming, transients; Ludewig et al., 2020), the outlier removal routine included in the NO_2 DOAS retrieval (van Geffen et al., 2020; van Geffen et al., 2025, App. F), or deliberate removal of a section of the fit window, as suggested in this paper. The error on the reflectance of these flagged spectral pixels is set to 10^4 times the measurement: $\Delta R_{\mathrm{meas}}(\lambda_i) = 10^4 \times R_{\mathrm{meas}}(\lambda_i)$, as a result of which they do not contribute to the residual in the χ^2-minimisation, but the value of χ^2 will be lower due to the flagging, even if the rest of the residual would remain unchanged. In the computation of the RMS error in Eq. (5), however, the flagged pixels need to be skipped, lowering the n_λ therein.

If the DOAS retrieval does not converge or if the number of flagged spectral pixels is too large ($>$ 1/4-th of the fit window) or if the number of outliers found is too large (> 10), the `qa_value` is set to zero. The OE retrieval provides an estimate of the uncertainties on (precision of) the fit parameters; if the NO_2 uncertainty is large ($\Delta N_s > 33.0\,\mu\mathrm{mol\,m^{-2}} = 2\times10^{15}\,\mathrm{molec.\,cm^{-2}}$), the `qa_value` is set to 0.15.

2.2.2 NO_2 slant column uncertainties

The DOAS uncertainty on the NO_2 SCD, ΔN_s, provided by the OE routine is an estimate that depends on details of the fit. The spatial variability of the SCDs over a remote Pacific Ocean sector can be used as an independent statistical estimate of the random component of the SCD uncertainty. Zara et al. (2018) used this approach to compare OMI and GOME-2A NO_2 and formaldehyde SCDs retrieved by different retrieval groups within the QA4ECV project, as well as to compare the SCD error estimates following from the different DOAS fits over the years 2005 through 2015. van Geffen et al. (2020) provided an initial analysis of TROPOMI NO_2 SCD uncertainties based on the then available data versions (v1.2.0 & v1.3.0) up to 31 Jan. 2020. As part of an ongoing monitoring of the stability of TROPOMI NO_2 retrievals, the DOAS and statistical uncertainties were recomputed after the mission reprocessing with data version v2.4.0 and is continued for subsequent versions manually on an irregular basis. To monitor the OMI NO_2 stability, preliminary collection 4 data (cf. Sect. 2.1.2) has been analysed as well, from the first available data of 1 Oct. 2004 onwards up to end of 2024.

For each day the first available orbit with satellite (nadir viewing) equator crossings west of about $-135°$ is selected as Pacific Ocean orbit (if such an orbit is missing on a given day, that day is skipped from the analysis) and analysed in the latitude range $[-60° : +60°]$: N_s and ΔN_s values are averaged in $2° \times 2°$ grid cells, after which the average over these grid

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

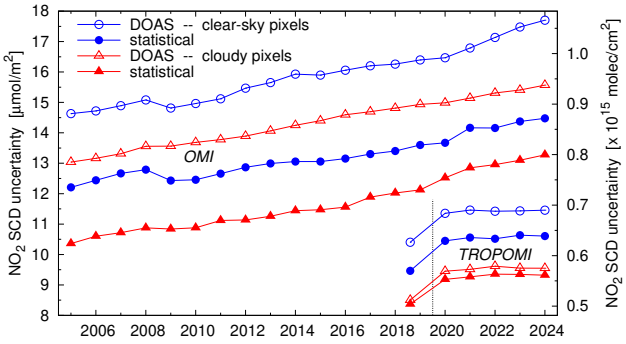

Figure 2. Yearly average DOAS (open symbols) and statistical (filled symbols) uncertainties for clear-sky (blue symbols) and cloudy (red symbols) pixels for OMI (top four curves) and TROPOMI (bottom-right four curves). In view of the TROPOMI pixel size reduction on 6 Aug. 2019 (indicated by the vertical dotted line), the first year cannot be a calendar year; instead the period 1 Aug. 2018 through 31 July 2019 is used.

cells of the ΔN_s gives the DOAS uncertainty, while the standard deviation of the N_s over these grid cells gives the statistical uncertainty. Results of this analysis, along with the monitoring of several other quantities, are presented on a set of web pages [2].

Fig. 2 shows a summary of the results: yearly average DOAS and statistical uncertainties for TROPOMI and OMI for clear-sky and cloud-free pixels separately. Per ground pixel the mean uncertainty of TROPOMI is significantly smaller the for OMI (due to the lower SNR of OMI) and on top of that the ground pixels are also much smaller. The pixel size reduction of TROPOMI on 6 Aug. 2019 has lead to an increase of the uncertainties of about $10\,\%$ (van Geffen et al., 2020). After this size reduction there is a small increase in the TROPOMI uncertainties of on average $0.03\,\mu\text{mol}\,\text{m}^{-2}$ ($= 1.7 \times 10^{12}$ molec. cm^{-2}) per year, while OMI uncertainties increase nearly five times faster: on average $0.14\,\mu\text{mol}\,\text{m}^{-2}$ ($= 8.2 \times 10^{12}$ molec. cm^{-2}) per year.

The TROPOMI data will be re-analysed after the forthcoming mission reprocessing (cf. Sect. 2.1.1) and the OMI data will be re-analysed once collection 4 data is finalised. These re-analyses will include the improvement in the NO_2 SCD retrieval proposed in this paper.

2.2.3 The Wald-Wolfowitz runs test on the fit residual

TROPOMI observations over lakes in Tibet under clear-sky and snow-free circumstances revealed tropospheric NO_2 columns markedly larger than in the surrounding area. Kong et al. (2023) attributed these enhanced columns to unknown NO_2 sources in the lakes. This prompted us to investigate NO_2 fit residuals in detail and we noticed, as reported by Labzovskii et al. (2024), that these residuals contain clear broad-band structures that are likely an indication that the NO_2 SCDs retrieved over these lakes are unreliable: some kind of absorber present in the water is clearly not accounted for in the modelled refectance. A tell-tale sign in this case was also that the water vapour fit coefficients have large negative values (down to $-1700\,\text{mol/m}^2$)

[2] See https://www.temis.nl/tropomi/no2scd/; last access: 20 November 2025

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

rather than positive values similar to those around the lakes (about $+200\,\mathrm{mol/m^2}$). The issue found over these Tibetan lakes is revisited in Sect. 6.1.

The discovery of structures left in the fit residual prompted us to implement a statistical test to try to signal for remaining low frequency structures in the fit residual: the Wald-Wolfowitz runs test, or 'runs test' for short (Barlow, 1989, Sect. 8.3.2) [3],

This test checks a randomness hypothesis based on the number of positive (k_p) and negative (k_n) values in the fit residual, i.e. based on the notion that an ideal fit residual is pure white noise, where a sequence of same-signed values is called a "run" and $n_\lambda = k_p + k_n$ the number of spectral points in the fit window. The number of expected runs, k_r^{exp}, the variance σ_r^2, and the deviation in terms of the standard deviation, R_D, are given by:

$$\begin{aligned} k_r^{\mathrm{exp}} &= 1 + (2\,k_p\,k_n)\,/n_\lambda \\ V(k_r^{\mathrm{exp}}) = \sigma_r^2 &= (2\,k_p\,k_n\,(2\,k_p\,k_n - n_\lambda))\,/\left(n_\lambda^2(n_\lambda - 1)\right) \\ R_D &= \left(k_r^{\mathrm{fit}} - k_r^{\mathrm{exp}}\right)/\sigma_r \end{aligned} \tag{6}$$

where k_r^{fit} is the number of runs in the fit residual. The deviation R_D has been defined here with a sign in order to make a distiction between fewer-than-expected ($R_D < 0$) and more-than-expected ($R_D > 0$) runs, i.e. to identify between low-frequency and high-frequency structures, respectively, in the fit residual. An additional quantity that proved to be useful is the length of the longest run, R_L, in the fit residual. See the ATBD (van Geffen et al., 2025) for further details and some examples.

An R_D that is $>+5$ or <-5 means that the number of runs is really in the tail of the distribution, while an R_L of, say, >20 is a significant fraction of the fit window. But since both R_D and R_L are continuous variables, it is not clear where to put a line between "good" and "bad" results. In addition, it is not certain that large R_D and R_L values actually mean that the retrieved NO_2 SCD is incorrect. And, vice versa, it is not certain that problems with the fit will always be picked up in the form of large R_D and/or R_L values. Still, both quantities can give useful information, which is why R_D and R_L are added to the TROPOMI NO_2 data product as *additional independent* information for the data user as of NO_2 data version v2.7.1.

3 Investigation of the fit residual issue

When looking at areas with large R_D and R_L, large patches over the oceans stood out clearly, prompting an investigation of fit residuals there, which in turn led us to the issue that is the subject of this paper. Initial investigations were done with TROPOMI orbits from 5 June 2019, as for that date the input data necessary for local reprocessing was available, since it was used for earlier investigations into the Tibetan Lakes issue (Sect. 6.1). Local reprocessing of (sections of) orbits was necessary, as the DOAS fit residual is not part of the nominal level-2 NO_2 files, since including residuals would lead to very large files while they would be of little use to by for most data users.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

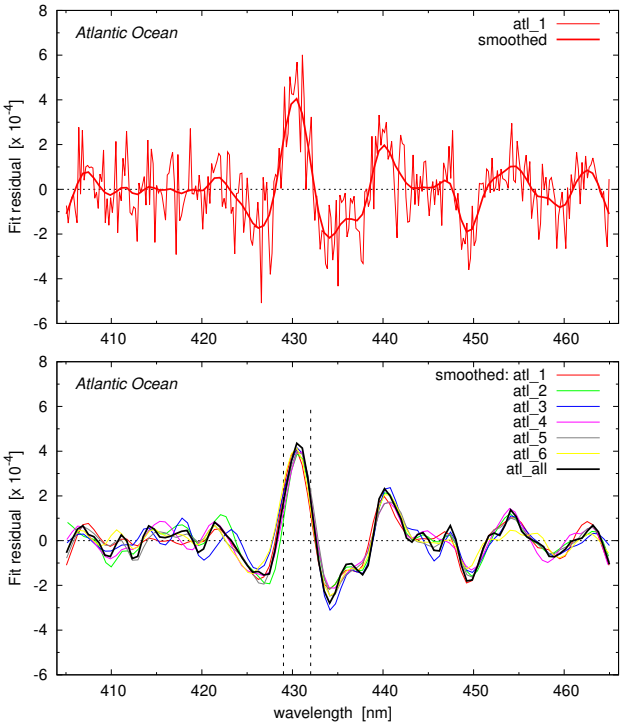

Figure 3. Examples of fit residuals of clear-sky pixels over the Atlantic Ocean ('atl' for short) that have large R_L and R_D. Top panel: residual of one example (thin red line) and the smoothed residual (thick red line). Bottom panel: smoothed residuals of six examples (thin lines) and a smoothed average of the six residuals (thick black line). Vertical dashed lines in the bottom panel indicate wavelengths 429.0 and 432.0 nm. Smoothing here and in other graphs is done with a natural spline without weights.

3.1 Atlantic Ocean

A section of orbit 08516 over the Atlantic Ocean ('atl' for short as id in figures and tables), with nadir latitudes $[+10° : +40°]$ was locally reprocessed so as to store the DOAS fit residuals and from this six clear-sky ground pixels with large R_L and large negative R_D were arbitrarily picked.

The top panel of Fig. 3 shows the fit residual from one example as a thin red line and a smoothed version of that as a thick red line. Results of the runs test for this fit residual are $R_D = -6.3$, i.e. really in the tail of the distribution, and $R_L = 33$ (at $\lambda = [432.28 : 436.58]$ nm).

The bottom panel of Fig. 3 shows smoothed residuals of the six examples as thin lines and a smoothed average residual ('all') as a thick line. All examples are around longitude $-40°$, latitude $+26°$; scanline and row numbers of these and of the other example pixels in this paper are listed in App. B.

[3] See also: https://en.wikipedia.org/wiki/Wald-Wolfowitz_runs_test; last access: 20 November 2025

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

Table 1. Residual RMS values over the full fit window (second column), the $[429:432]\,\mathrm{nm}$ range (third column), the remainder of the fit window (fourth column) and the Q^{430}_{RMS} ratio (fifth column) of the average over the six clear-sky examples of the Atlantic Ocean ('atl_all') in Fig. 3 and Western Australia ('aus_all') in Fig. 5.

	RMS $[\times 10^{-4}]$			
residual	full	[429:432]	rest	ratio
atl_all	1.73	3.90	1.53	2.54
aus_all	1.52	3.23	1.38	2.34

Clearly noticeable in the fit residuals is the large peak around $430\,\mathrm{nm}$. Both the RMS error and the χ^2 of the fit do not have suspiciously large values: those two quantities give no cause for alarm. Yet, the $430\,\mathrm{nm}$ peaks is clearly systematic and will have some impact on the NO_2 DOAS fit results.

To investigate the magnitude and occurence of this peak further, consider the ratio of the RMS of the peak – i.c. the wavelength range $[429:432]\,\mathrm{nm}$, indicated in the bottom panel of Fig. 3, which spans about 15 spectral pixels – and the RMS of the rest of the fit window:

$$Q^{430}_{\mathrm{RMS}} = \frac{R_{\mathrm{RMS}}(\lambda \in [429:432])}{R_{\mathrm{RMS}}(\lambda \notin [429:432])} \tag{7}$$

(Like the fit residual, this ratio is not part of the nominal level-2 files.) The values of the RMS elements in Eq. (7) and the full-window RMS error of the average residual ('atl_all') in Fig. 3 are given in Table 1; for the individual residuals the ratio varies between 2.3 and 2.8. The higher the ratio, the more pronounced the $430\,\mathrm{nm}$ peak stands out against the noise in the fit residual. Although RMS values are continuous quantities, a ratio $Q^{430}_{\mathrm{RMS}} \geq 2.0$ likely indicates there is a problem with the NO_2 SCD.

The top-left panel of Fig. 4 shows for the above mentioned orbit section the Q^{430}_{RMS} ratio as function of the cloud radiance fraction: by far most pixels with a ratio ≥ 2.0 are clear-sky pixels (left part of the panel). In fact, the ratio is large only for clear-sky pixels for which the full-window RMS error of the fit is small. For these clear-sky pixels, where the satellite sees the ocean waters, the RMS_430 ratio increases with increasing value of the liquid water fit coefficient (top-right panel of Fig. 4): the deeper the light reaches into the ocean waters, the larger the effect of VRS on the retrieval and therefore the more pronounced the peak around $430\,\mathrm{nm}$.

For pixels with higher cloud radiance fraction the RMS error of the fit is higher and the RMS_430 ratio is lower. In other words: for most cloudy pixels, any problem with the fit residual around $430\,\mathrm{nm}$ is hidden between the larger noise on the residual due the presence of those clouds.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

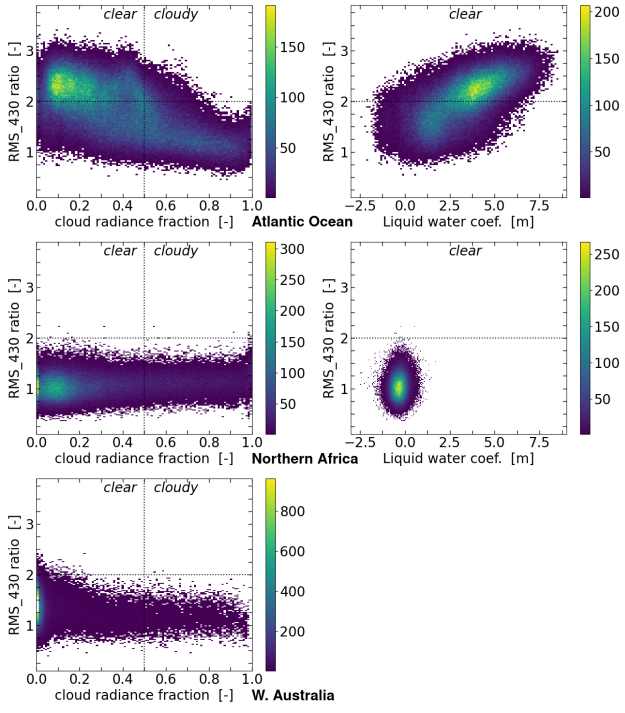

Figure 4. Scatter plots of the RMS_430 ratio (y-axis) as function of the cloud radiance fraction (left column) and the liquid water fit coefficient for clear-sky pixels (right column) over the Atlantic Ocean (top row), Northern Africa (middle row) and Western Australia (bottom row). Horizontal dotted lines indicate a ratio of 2.0. The vertical line in the left column panels indicates the separation between clear-sky (left) and cloudy (right) pixels. Colour bars in scatter plots like this show the number of occurrences.

3.2 Northern Africa

For comparison of what goes on over land, consider a section of orbit 08514 over Northern Africa, with nadir latitudes $[+10° : +30°]$, which is slightly smaller than the Atlantic Ocean section used above to avoid including the Mediterranean Sea, but otherwise covers the same latitudes.

As the middle row of Fig. 4 shows, there are only a few pixels with $Q^{430}_{\mathrm{RMS}} > 1.5$ (about 4% from the orbit section), while the full-window RMS has roughly the same range as for the Atlantic Ocean section. Apparently there is no issue around $430\,\mathrm{nm}$ over this area of land. But this does not hold everywhere over land.

3.3 Western Australia

When inspecting other land areas where $Q^{430}_{\mathrm{RMS}} > 2.0$ occurs in the 5 June 2019 orbits, several pixels stood out over Western Australia, though quite scattered. To look at this in more detail, consider a section of orbit 08510 with nadir latitudes $[-45.2° : -7.5°]$, which covers half of Australia west of about $+135°$, most of which is free of clouds.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

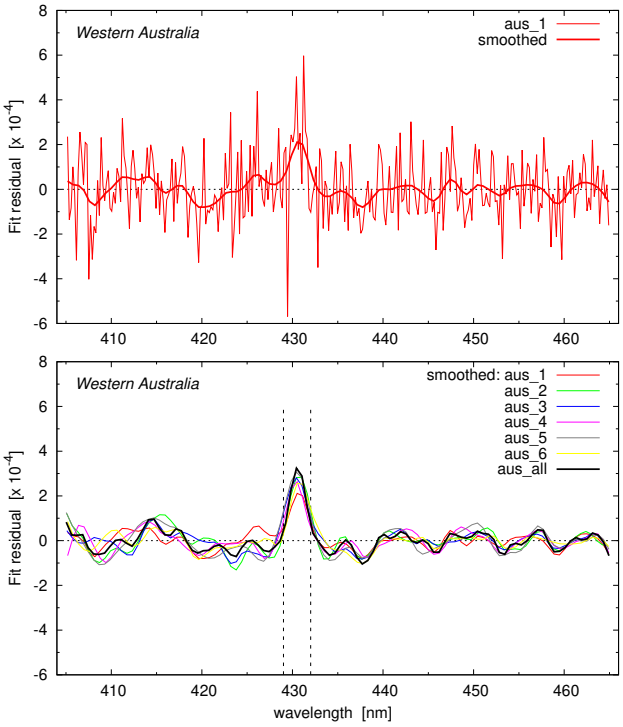

Figure 5. As Fig. 3 but for the clear-sky land pixels over Western Australia ('aus' for short) that have an RMS_430 ratio larger than 2.0.

The bottom-left panel of Fig. 4 shows the scatter plot of the RMS_430 ratio as function of the cloud radiance fraction: about 2% of the really cloud-free pixels (i.e. with a very small cloud radiance fraction) have a ratio larger than 2.0, with a maximum of 2.43; for the Atlantic Ocean section the maximum ratio found is 3.45. None of the Australian land pixels is signaled by the runs test, with $|R_D| < 3.5$ and $R_L < 20$.

Fig. 5 shows fit residuals of six ground pixels around longitude $+120°$, latitude $-27°$ with RMS_430 ratios around 2.3; the bottom row of Table 1 lists the values of the RMS elements of Eq. (7) for the average ('aus_all') over the six example residuals. There clearly is a peak around 430 nm that stands out above the noise in the fit residual, which definitely is not caused by VRS: the pixels are all classified as "shrubland". Perhaps this indicates that the effect of RRS is not completely accounted for in the fit.

3.4 RMS_430 ratio worldwide and change over time

Fig. 6 shows a map of the locations where $Q^{430}_{\rm RMS} \geq 2.0$ for all TROPOMI orbits of 5 June 2019. Almost all these locations are over open water; over land there are only scattered pixels, as mentioned in the preceeding sections. Cloudy pixels tend to have a low RMS_430 ratio and therefore do not show up in the map. The outer 22 (20) rows at the left (right) edge of the swath have a larger spectral uncertainty, which is reflected in larger SCD and RMS errors, as a result of which the RMS_430 ratio is somewhat lower there.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

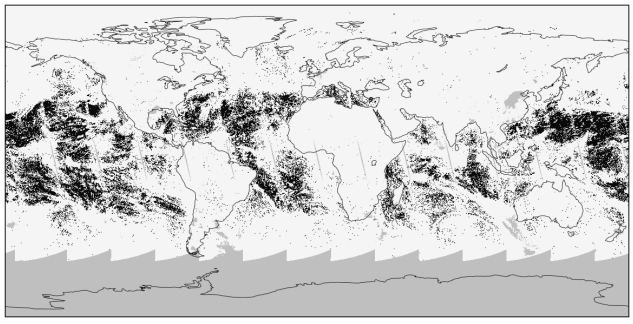

Figure 6. Map of ground pixels with $Q^{430}_{\mathrm{RMS}} \geq 2.0$ in black and the other pixels in white from all orbits of 5 June 2019. (Maps like this one are made with Panoply.)

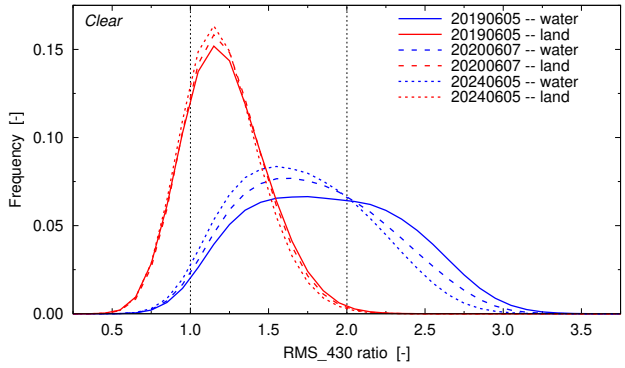

Figure 7. Frequency distribution of the RMS_430 ratio of all orbits of 5 June 2019 (solid lines), 7 June 2020 (dashed lines) and 5 June 2024 (dotted lines) for clear-sky pixels over water (blue) and land (red).

Maps for other days (not shown) look similar, with a seasonal variation in the overall North-South pattern over the oceans and an apparent small decrease in time of the number of pixels with $Q^{430}_{\mathrm{RMS}} \geq 2.0$. To investigate this further, Fig. 7 shows the frequency distribution of the RMS_430 ratio for cloud-free pixels of three selected days over water and over land separately.

From 5 June 2019 to 7 June 2020, the RMS_430 ratio decreases somewhat, clearly visible over water (blue lines) and less clearly over land (red lines). The reason for this decrease is likely the fact that due to the along-track pixel size reduction in August 2019 the uncertainties have increased (Sect. 2.2.2), i.e. the noise in the fit residual has increased, as a result of which the $430\,\mathrm{nm}$ peak stands out less clearly.

After that, the uncertainties of the NO_2 retrieval remain almost the same, as illustrated in Fig. 2, but Fig. 7 indicates that the $430\,\mathrm{nm}$ peak decreases in magnitude between 2020 and 2024, in particular over water but also over land. This is likely caused by the fact that in 2024 the Sun is more active than in 2020, as a result of which the $430\,\mathrm{nm}$ structure is less deep, in particular the calcium line there – see App. A for a brief investigation of the line depth over time. The irradiance and radiance are affected

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

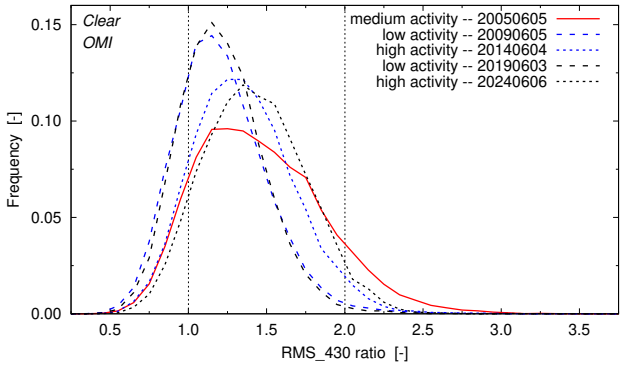

Figure 8. Frequency distribution of the RMS_430 ratio of clear-sky water pixels from OMI Atlantic Ocean orbits of five different years: 2005 with medium solar activity (solid red line), 2009 and 2019 with low solar activity (blue and black dashed lines), and 2014 and 2024 with high solar activity (blue and black dotted lines). Data affected by the row anomaly are filtered out; for consistency sake, the 2005 data have been filtered with the 2024 row anomaly flagging.

directly by the varying depth at that wavelength, while in addition to that the radiance over water is also indirectly affected by the change of the depth of the Ca+ lines around 395 nm which are shifted to around 430 nm by VRS (Sect. 1).

In other words: because the structure in the irradiance around 430 nm varies with the solar activity cycle, the effects of RRS are never really fully compensated around that wavelength, since the Ring reference spectrum (Sect. 2.2.1) is determined from a fixed irradiance reference spectrum, and over water things are made worse by the effects of VRS.

For ground pixels over clouds (not shown), the frequency distribution is slightly narrower than for the land pixels in Fig. 7, with a peak value at about the same RMS_430 ratio, also with a small narrowing over time and a small increase of the peak value.

3.5 OMI measurements

To investigate the RMS_430 ratio for OMI collection-4 measurements, Atlantic Ocean orbits similar to the TROPOMI one of 5 June 2019, where selected on or close to 5 June of 2005, 2009, 2014, 2019 and 2024.

Fit residuals of clear-sky pixels of the 5 June 2019 orbit (not shown) look very similar to those shown in Fig. 3, but on the whole the RMS_430 ratio is lower, the reason no doubt being that the noise on the OMI measurements is in general larger than on TROPOMI measurements: ratios above 2.5 are rare, most pixels have ratios well below 2.0, as can be seen from the red solid line in Fig. 8.

That figure also shows that the change over time of the RMS_430 ratio for OMI is quite different than for TROPOMI (cf. Fig. 7) in relation to the solar activity cycle: the higher activity of 2024 leads for TROPOMI to lower RMS_430 ratios than for the lower activity of 2020, while for OMI the higher activity of 2014 and 2024 (dotted lines in Fig. 8) leads to higher RMS_430 ratios than for the lower activity of 2009 and 2019 (dashed lines). The increase of the overall SCD error for OMI over time (Fig. 2) and thus of the RMS error, may be visible in Fig. 8 in the difference between the two low-activity (dashed)

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

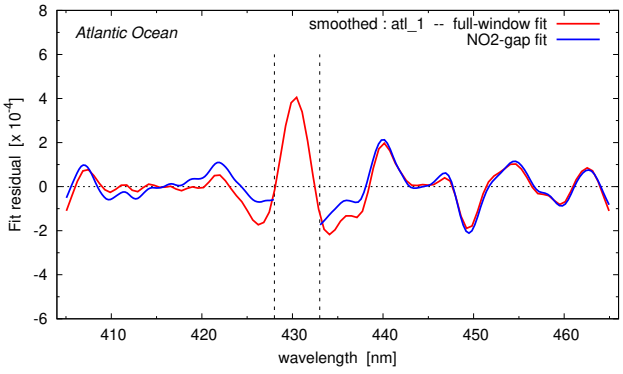

Figure 9. Smoothed fit residual of the Atlantic Ocean 'atl_1' example in Fig. 3 using the full fit window for the retrieval (red line) and with the NO2-gap fit (blue line). Vertical dashed lines indicate the wavelength range that is cut out: 428.0 to 433.0 nm.

lines as a small shift to lower RMS_430 ratios, whereas in the two high-activity (dotted) lines it seems to show up as an small shift to higher RMS_430 ratios; these differences, however, could also be due to differences in atmospheric circumstances.

The reason for the different behaviour of TROPOMI and OMI is not clear, but it may be related to the fact that for OMI NO_2 retrievals the 2005 average irradiance is used, while for TROPOMI the daily measured irradiance is used (cf. Sect. 2.2.1), as a result of which the effect of the 430 nm issue on the reflectance and hence on the fit residual is quite different for the two instruments.

For cloudy pixels (not shown), the situation is more like TROPOMI: at low activity, the RMS_430 ratios are on the whole a little larger than at high activity.

4 Dealing with the fit residual issue at 430 nm

From the above analysis it is clear that around 430 nm there is a systematic feature that will have some impact on the NO_2 SCDs, in particular over oceans but also lands pixels may be affected. Compensating for this issue within the DOAS fit is not trivial as the feature appears to vary over time. Using the RMS_430 ratio of Eq. (7) as a filter is not possible either, as that ratio is a continuous quantity, like the χ^2 and RMS error of the fit.

We therefore propose to cut the feature from the DOAS fit by ignoring wavelengths in the range $\lambda = [428 : 433]$ nm, named "NO2-gap fit" below. This window – slightly bigger than the range used in Eq. (7), to be sure to include the full feature – is indicated by the dashed vertical lines in Fig. 1. We recommend that this fix is implemented for all ground pixels, not just over ocean waters, to avoid introducing retrieval inconsistencies across land-water boundaries.

4.1 Individual pixel comparisons

Fig. 9 shows smoothed residuals of the Atlantic Ocean example from the top panel of Fig. 3 using the full-window fit (red line) and using the NO2-gap fit (blue line): especially on either side of the gap the residual is reduced. Table 2 lists for the six

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

Table 2. Changes in the main retrieval results between the NO2-gap and the full-window fit in percent for the six clear-sky Atlantic Ocean pixels of Fig. 3.

	NO2-gap minus full-window fit [%]			
atl_#	N_s	ΔN_s	R_{RMS}	χ^2
1	−5.59	−11.27	−16.03	−27.38
2	−5.32	− 8.74	−13.00	−22.79
3	−0.99	−10.13	−14.08	−24.82
4	−1.78	−10.94	−10.94	−27.06
5	−3.95	− 9.89	−13.74	−24.06
6	−4.07	−12.39	−17.54	−28.81

Table 3. As Table 2 but for the six clear-sky Western Australia pixels of Fig. 5.

	NO2-gap minus full-window fit [%]			
aus_#	N_s	ΔN_s	R_{RMS}	χ^2
1	+3.98	−6.24	−10.10	−17.57
2	−5.82	−7.27	−10.32	−20.05
3	−3.24	−3.24	−10.72	−18.24
4	−2.03	−6.89	−9.54	−19.56
5	−3.88	−7.98	−11.72	−20.49
6	−2.75	−6.29	−10.73	−17.48

Atlantic Ocean examples the change in the NO_2 SCD value, SCD error, RMS error and χ^2 of the fit. A reduction of the latter three – in these cases of 10% or more – is generally considered to indicate that the fit has improved and one then assumes that the resulting SCD value has become more reliable.

The NO2-gap fit also leads to improved fits for the Western Australian land pixels, as can be seen from the results listed in Table 3, which shows that in some cases the SCD value may increase due to the improvements. For the Northern Africa land pixels (not shown), changes in the SCD and RMS error are for by far most pixels no more than $\pm 2\%$; for clear-sky pixels the changes are less than for cloudy pixels.

4.2 Changes across the world and over time

Fig. 10 shows a map of the relative change in percent of the NO_2 SCD error due to the NO2-gap fit instead of the full-window fit for all orbits of 5 June 2019 – changes occur there where the RMS_430 ratio is large (cf. Fig. 6). Scatter plots of the changes in the NO_2 SCD error and GCD value are shown in Fig. 11 for clear-sky pixels over water and land and for all cloudy pixels,

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

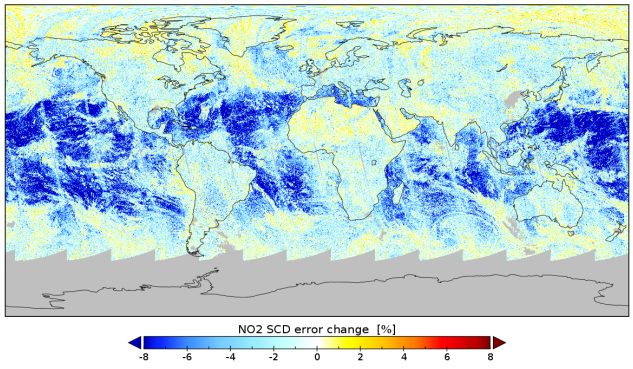

Figure 10. Map of the relative change in the NO_2 SCD error of the NO2-gap minus full-window fit for all orbits of 5 June 2019.

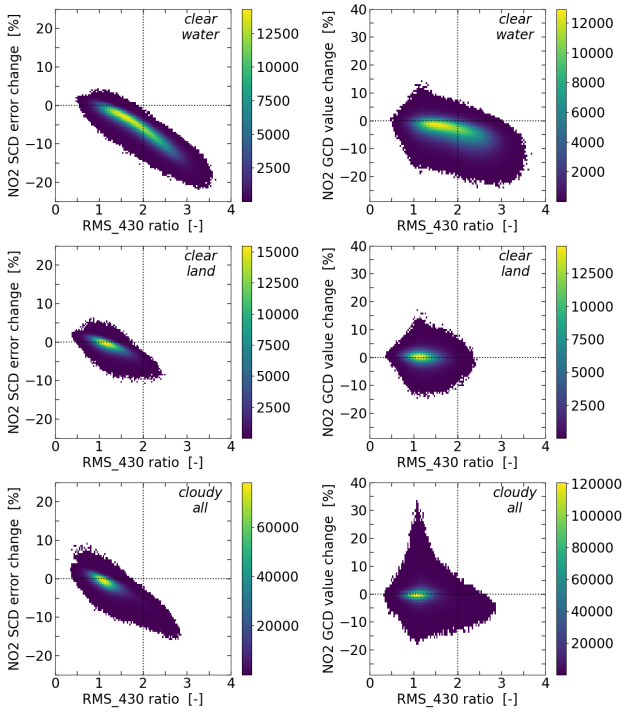

Figure 11. Scatter plots of the change in the NO_2 SCD error (left column) and GCD value (right column) of the NO2-gap minus full-window fit in percent for clear-sky pixels over water (top row) and land (middle row), as well as all cloudy pixels (bottom row), of all orbits of 5 June 2019 as function of the RMS_430 ratio.

separately, as function of the RMS_430 ratio. For pixels over water, the reduction of the SCD error is evident, while over land there is also a decrease of the SCD error for a lot of pixels but some pixels show a small increase as result of the NO2-gap fit. The GCD values show a small decrease over water, while over land they remain mostly the same. For cloudy pixels, the SCD

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

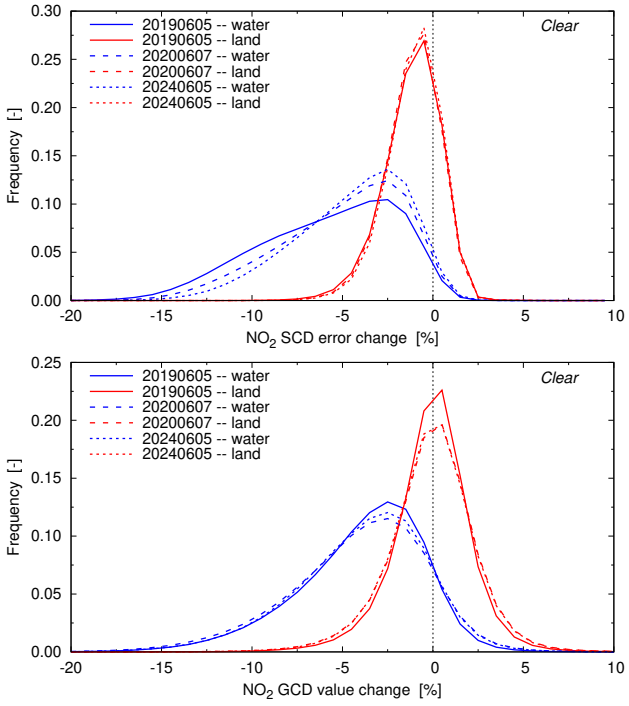

Figure 12. Frequency distribution of the NO_2 SCD error (top panel) and GCD value (bottom panel) of all orbits of 5 June 2019 (solid lines), 7 June 2020 (dashed lines) and 5 June 2024 (dotted lines) for clear-sky pixels over water (blue) and land (red).

error and GCD value changes are relatively small, where one has to keep in mind that for cloudy pixels with a cloud radiation fraction just above 0.5 part of the light still comes from the water or land surface.

Frequency distributions of the changes for clear-sky pixels in the NO_2 SCD error and GCD value are shown by solid lines in Fig. 12, along with the changes for the other two test days mentioned above. The changes in the SCD error due to the NO2-gap fit become a little smaller over time, along with what is seen for the RMS_430 ratio in Fig. 7, but on the whole there clearly is in improvement in the fit over water and a small improvement over land. The GCD value changes are more or less the same over time: a decrease over water and on average no change over land.

For the evaluation of the impact of the NO2-gap fit on the stratospheric and tropospheric columns, two full months where processed: July 2023 and January 2024 – see Sect. 5. The Pacific Ocean orbits of these two months can be used to check the impact on the slant column uncertainties discussed in Sect. 2.2.2. Given that there is quite some day-to-day variation, Table 4 lists the monthly average ratios of the uncertainties of the NO2-gap over the full-window fit, as well as the average SCD value over the same latitude range (the difference between the two months is related to the seasonal cycle in the uncertainties [4]).

As expected from the above reported changes in the SCD error, the DOAS uncertainty improves, in particular for clear-sky scenes. For those scenes the statistical uncertainty does not change much, but the NO2-gap fit appears to have an impact on

[4] See the graphs on the webpages referred to in footnote 2.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

Table 4. Monthly average ratio of NO2-gap over full-window fit results of Pacific Ocean orbits of the two TROPOMI test months.

pixel type	quantity	2023-07	2024-01
clear-sky	DOAS uncertainty	0.954	0.943
	statistical uncertainty	0.992	0.992
	average SCD value	0.964	0.957
cloudy	DOAS uncertainty	0.983	0.980
	statistical uncertainty	0.914	0.908
	average SCD value	0.993	0.990

the statistical uncertainty over cloudy scenes with changes up to -10%. It is not clear why the latter decrease occurs. As the bottom-right panel in Fig. 11 shows, the GCD does on average not change over clouds, but there is quite some scatter, even for low RMS_430 ratio. Perhaps the removal of the $430\,\mathrm{nm}$ peak gives better, more consistent results over clouds due to the incomplete handling of that peak by the Ring correction in the full-window retrieval.

4.3 OMI measurements

Inspecting OMI fit residuals of individual clear-sky Atlantic Ocean pixels with a relatively high RMS_430 ratio of the 2005 orbit without and with the NO2-gap leads to graphs similar to Fig. 9 and the changes in the SCD and RMS error are of the same order as those listed in Table 2. For pixels from the 2024 orbit the changes in the SCD and RMS error appear to be smaller, probably because the $430\,\mathrm{nm}$ peak is relatively less strong in 2024 than in 2005 due to the increased solar activity (Sect. 3.5) and/or the increased measurement uncertainties (Sect. 2.2.2).

An evaluation of the data for water pixels only in the latitude range $[-40^\circ : +40^\circ]$, which covers the area where the largest changes are expected to occur, and filtering out the rows suffering from the row anomaly (also for the 2005 data, to have equal viewing geometry coverage), reveals that for most pixels the NO_2 SCD error is reduced in the NO2-gap approach: the changes lie roughly between -12% and $+2\%$. For clear-sky pixels the distribution of the SCD error changes peaks at about -0.5% for the orbits of 2009 and 2019, i.e. when solar activity is low, with a frequency a 10th of the peak-frequency for a change of -5.5%. For the high-activity years 2014 and 2024 the peak of the distribution lies at about -1.5% and a frequency a 10th of the peak-frequency is found for a change of -9%. For the 2005 orbit, a year of medium solar activity, the peak lies at roughly the same change but the tail is somewhat longer: a frequency a 10th of the peak-frequency is found at -10%. In other words, for low-activity years the SCD error changes of clear-sky pixels have a narrower distribution than for high-activity years. The reverse is the case for cloudy pixels, and for these pixels the peaks lie between -1.5% and -0.5% for all years.

The changes of GCD values for these Atlantic Ocean pixels (which in general have low GCD values to begin with) have a wide distribution, roughly between -15% and $+10\%$ for most clear-sky pixels, with a peak at about -0.5%, with only a little difference between low- and high-activity years. For cloudy pixels the GCD appears to increase a little on average: the changes range from about -10% to $+15\%$, with peaks around $+1\%$ to $+2\%$.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

5 Impact on stratospheric and tropospheric columns

To study the impact of the NO2-gap approach on the final stratospheric and tropospheric NO_2 columns two months were selected: July 2023 and January 2024, so as to capture seasonal variation. Since the TM5 data assimilation was started from an existing distribution and the across-track stripe correction amplitude is an average over a period of seven days (van Geffen et al., 2025), it takes a few days to adjust (spin-up) to new approaches, hence we analyse the vertical column data starting at day 8 of each month; the stripe correction amplitude is determined and applied after the SCD and GCD are calculated, and it may be different for the full-window and the NO2-gap approaches. NO_2 column data were gridded per day using the HARP software [5] on $0.2° \times 0.2°$, which gives average values where orbits overlap.

5.1 Overall impact

Since most of the NO_2 over the oceans, away from emission sources, is found in the stratosphere, a change in the SCD in those regions will primarily lead to changes in stratospheric NO_2. For the vast majority of the grid cells the change of the stratospheric column is in the range -5% ($-2\,\mu mol\,m^{-2}$ or 1.2×10^{14} molec. cm^{-2}) to 0%; cf. the top panel of Fig. 13. For some individual grid cells the changes may be a bit larger, while for very few grid cells stratospheric NO_2 may increase a little. Most of the change is found in the latitude range $[-40° : +40°]$, the same range where the largest changes in SCD and SCD error occur.

Since tropospheric NO_2 column values over oceans are usually small and over land they can vary quite a bit, it is best just to consider absolute differences; cf. the bottom panel of Fig. 13. For by far most clear-sky water grid cells the tropospheric column decreases: average changes are between -2.0 and $+0.5\,\mu mol\,m^{-2}$, with a peak around $-1.0\,\mu mol\,m^{-2}$. Clear-sky land scenes and all cloudy scenes, on the other hand, show an overall increase of the tropospheric column: average changes are between -0.5 and $+2.0\,\mu mol\,m^{-2}$. Some individual grid cells show much larger (monthly average) changes though: between -20 and $+15\,\mu mol\,m^{-2}$ for clear-sky scenes, while for cloudy scenes the changes may range between -40 and $+40\,\mu mol\,m^{-2}$ (2.4×10^{15} molec. cm^{-2}).

On average we expect that the impact on tropospheric columns is small, since the data assimilation will remove average biases. The lower stratospheric columns above tropical seas are transported over land, as is particularly visible over Eastern Asia and over South-East USA and Mexico (top panel of Fig. 13), where tropospheric columns will increase since the SCD over land remains more or less the same. Since over polluted areas the AMF is much smaller than the stratospheric NO_2, tropospheric columns may increase more than the stratospheric columns decrease. The data assimilation with the NO2-gap retrieval results reduces the entire column over oceans, leading to a small decrease of the tropospheric columns there.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

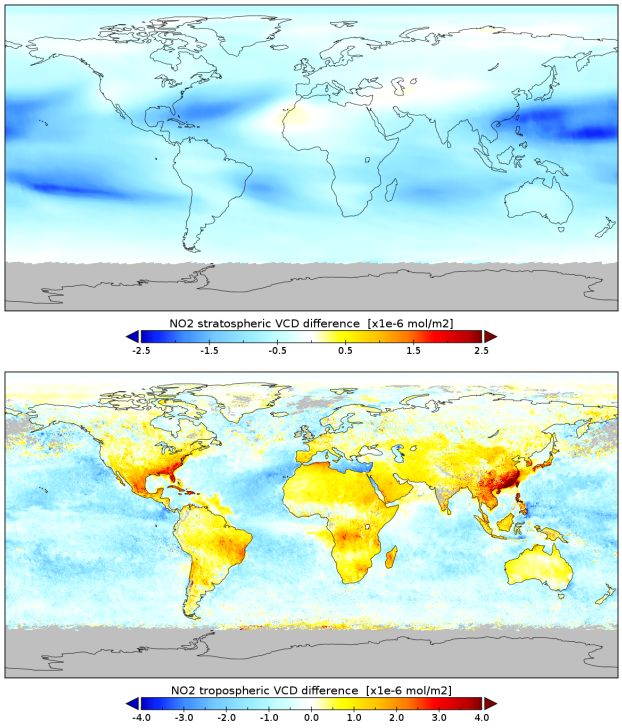

Figure 13. Average change NO2-gap minus full-window in the stratospheric (top panel; all pixels) and tropospheric (bottom panel; clear-sky pixels only) NO_2 vertical column for the test month July 2023. The two panels have different scale ranges, both in $\mu mol\,m^{-2}$, where $4\,\mu mol\,m^{-2}$ corresponds to 2.4×10^{14} molec. cm^{-2}.

5.2 Impact over validation stations

Routine validation of TROPOMI tropospheric, stratospheric and total column data with ground-based measurements is being carried out by the Validation Data Analysis Facility (VDAF[6]), with support from the S5P Validation Team (S5PVT), which issues Quarterly Validation Reports, such as Lambert et al. (2025).

The VDAF data of a given station could be compared to the new results by selecting the same ground pixel from the orbit files, as a repetition of the validation. But that probably does not provide clear answers, given that (a) ground-based measurements are not available on each day of the two months, (b) in view of the spin-up the first seven days of each months need to be skipped, and (c) a large day-to-day variation is observed in the validation results, which is larger than the differences discussed here. Instead it seems a better idea to considered the data from the full-window and NO2-gap retrievals at the VDAF locations in the gridded data.

[5] https://stcorp.github.io/harp/doc/html/; last access: 20 November 2025
[6] https://mpc-vdaf.tropomi.eu/; last access: 20 November 2025

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

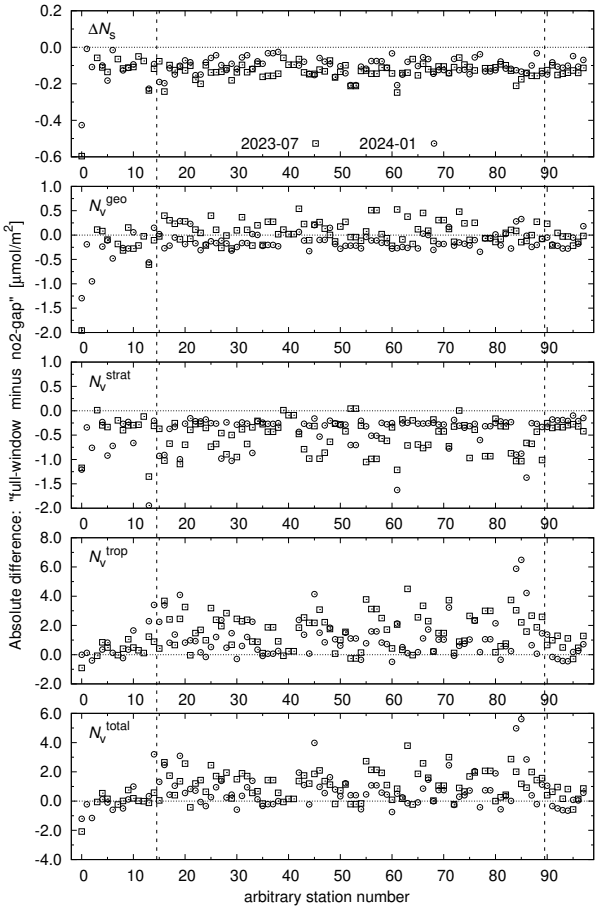

Figure 14. Average absolute change of, from top to bottom, the NO_2 slant column error (ΔN_s) and the geometric (N_v^{geo}), the stratospheric (N_v^{strat}), the tropospheric (N_v^{trop}) and the total (N_v^{total}) vertical columns between the NO2-gap and the full-window fit results for the two test months over the VDAF validation sites for the 15 stratospheric (left part of each panel), 75 total (middle part) and 8 tropospheric (right part) columns. Stations have been assigned numbers in alphabetical order per group; the groups are separated by vertical dashed lines.

The results of this comparison are shown in Fig. 14 in the form of monthly averages of daily absolute differences, where vertical dashed lines separate the three groups of stations for the validation of the stratospheric (15 locations; left), total (75 locations; middle) and tropospheric (8 locations; right) column; a few locations at high latitudes have no data in the local winter month.

For most locations, the average SCD error (ΔN_s; top panel) is reduced by about $0.1\,\mu mol\,m^{-2}$ (0.6×10^{13} molec. cm^{-2}) or about 1%; the standard deviation of the averaging is smaller than the average for most stations. These reductions are quite small, because the stations are located on land, where the impact of the NO2-gap approach is small. The largest change in ΔN_s occurs for the first stratospheric column station (Bauru, Brazil): -0.6 (-0.4) $\mu mol\,m^{-2}$ for the month 2023-07 (2024-01), or -5.8% and -3.7%, respectively.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

For that station, the average GCD column (N_v^{geo}; 2nd panel) goes down by 2.0 (1.3) µmol m^{-2} in 2023-07 (2024-01), or about -3.3% for both months, while the third stratospheric column station (Dumont d'Urville, Antartica) shows a decrease of N_v^{geo} by 1.3 µmol m^{-2} in 2024-01. For most other stations, the change in N_v^{geo} is between -1% and $+1\%$.

The average stratospheric column (N_v^{strat}; 3rd panel) decreases for almost all stations, from -0.2 to -1.0 µmol m^{-2} (0.6×10^{14} molec. cm^{-2}, or up to -2%), with a standard deviation lower than 0.3 µmol m^{-2}, i.e. the decrease of N_v^{strat} is fairly robust. The 14th stratospheric column station (St. Denis, Reúnion) shows the largest decrease: -1.9 µmol m^{-2} in 2024-01 (about -4.5%), followed by the 47th total column station (Mauna Loa, Hawaï): -1.6 µmol m^{-2} in 2024-01 (about -5.1%). As the routine validation mentioned above shows, TROPOMI slightly underestimates the stratospheric column, but those results are latitude dependent. with a possible small overestimation in the tropics where the NO2-gap approach is most prominent. It looks like the NO2-gap approach reduces the latitude dependency a little, but the impact on the biases is difficult to estimate in view of the large uncertainties in the validation comparisons.

The average tropospheric column (N_v^{trop}; 4th panel) of the MAX-DOAS stations – all of which lie on the Northern hemisphere – does not change much and the same holds for most of the stratospheric column stations. For the total column stations the change of N_v^{trop} has a quite varied distribution, with little to no change for some locations and changes $+4.0$ to $+6.0$ µmol m^{-2} (3.6×10^{14} molec. cm^{-2}) for other locations. Given that the tropospheric column values for some stations are small, relative changes can be large, but for most stations the relative change is less than $\pm 15\%$.

The average total column (N_v^{total}; 5th panel), i.e. the sum of the tropospheric and total column, shows changes similar to those in the N_v^{trop}.

All in all the changes in the vertical columns are on average small and are not likely to affect the main conclusions of the VDAF validation reports.

6 Discussion points

6.1 Tibetan lakes

As mentioned at the beginning of Sect. 2.2.3, unexpectedly large TROPOMI tropospheric NO_2 columns over Tibetan lakes, attributed by Kong et al. (2023) to sources in those lakes, are likely due to unreliable NO_2 slant column retrievals (Labzovskii et al., 2024), as indicated by the presence of broad-band structures in the fit residuals and large negative water vapour (H_2O_{vap}) coefficients over those lakes. For that study, orbit 08511 of 5 June 2019 was selected – the first orbit in June 2019 that has fully clear-sky pixels over two major lakes: Lake Siling and Lake Nam, but also other lakes in and around Tibet were investigated. For this paper, we focus on those two large lakes (located at about $4.5 - 4.7$ km altitude), as well as on Issyk-Kul in Kyrgyzstan (at about 1.6 km altitude), as these cover multiple TROPOMI ground pixels.

The top panel of Fig. 15 shows the fit residual of a pixel over Lake Siling. Like the residuals over the Atlantic Ocean, shown in Fig. 3, there are broad-band structures in the residual above about 430 nm, but there are striking differences: the peak around 430 nm that lead to the NO2-gap approach is much less pronounced here, the "downward wave top" between 430 and 440 nm is much deaper, and there is a strong negative "peak" at about 442.7 nm (indicated by an arrow). That peak can be attributed to the

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

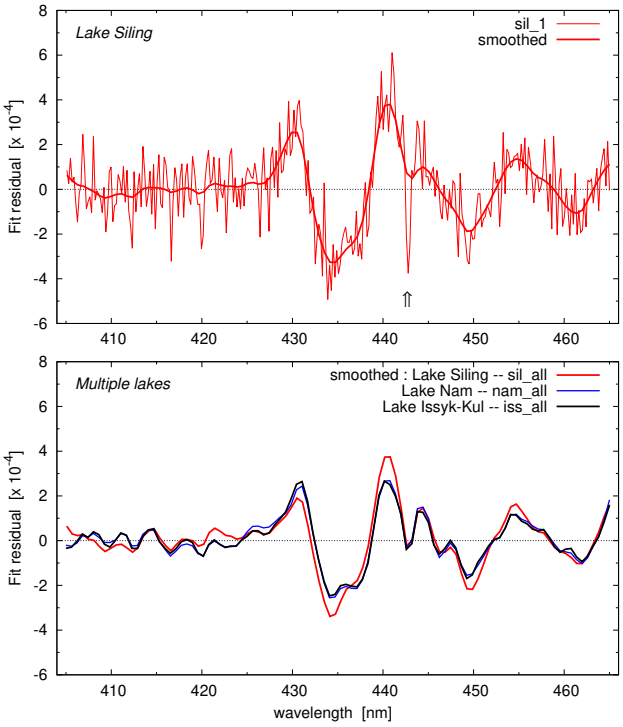

Figure 15. Top panel: example of a fit residual over Lake Siling (thin red line) and the smoothed residual (thick red line). The arrow at $442.7\,\mathrm{nm}$ points to an issue discussed in the text. Bottom panel: smoothed fit residuals of the average over six residuals over three lakes: Lake Siling (red line), Lake Nam (blue line) and Lake Issyk-Kul (black line) from the full-window fit; the blue and black lines almost exactly overlap.

fact that the DOAS fit has, in order to minimise the residual, among others resulted in a strongly negative H_2O_{vap} coefficient of $-1.419 \pm 0.280 \times 10^3\,\mathrm{mol/m^2}$ for this example, as opposed to an average of $+0.235 \times 10^3\,\mathrm{mol/m^2}$ outside the lake.

Both the broad-band structures and the large negative H_2O_{vap} coefficient were tell-tale signs to conclude that the higher GCD values over the lakes ($0.71\,\mu\mathrm{mol/m^2}$ on average), compared to surrounding values ($0.66\,\mu\mathrm{mol/m^2}$ on average; cf. the maps in Labzovskii et al., 2024), are probably due to unrealiable fit results rather than due to NO_2 emissions from the lakes: since stratospheric NO_2 is more or less constant over the short distances involved here, any elevation in the retrieved SCD value will end up as an enhancement of the tropospheric NO_2 column.

Residuals of other ground pixels over Lake Siling, Lake Nam and Lake Issyk-Kul look similar, with some difference in the details of the broad-band structures and the H_2O_{vap} values. The bottom panel of Fig. 15 shows for the three lakes the smoothed fit residual averaged over six ground pixels with surface classification "Water" – even in these smoothed versions, the downward peak related to the negative H_2O_{vap} coefficient can be seen.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

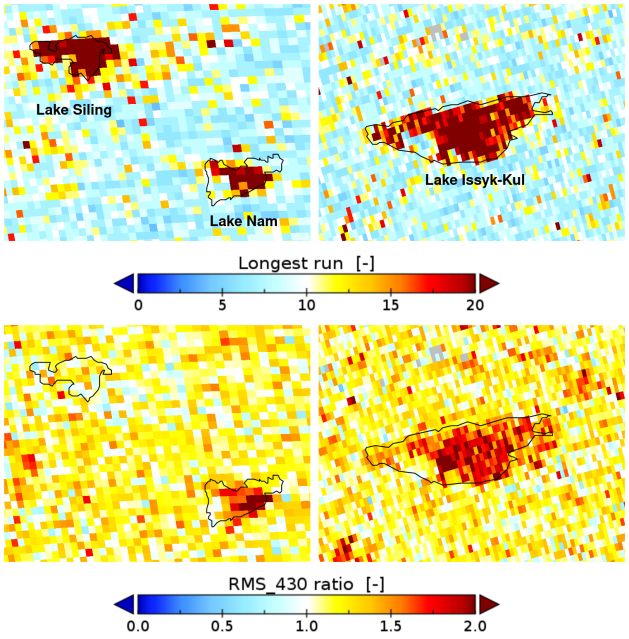

Figure 16. Maps of the longest run R_L (top row) and RMS_430 ratio (bottom row) for Lakes Siling and Nam (left column; centred at $+89.8°$, $+31.2°$) and Lake Issyk-Kul (right column; centred at $+77.4°$, $+42.5°$). Maps are $2°$ in north-south direction. Approximate lake contours are made by Panoply.

The broad-band structure between 430 and $440\,\mathrm{nm}$ is picked up by the runs test: $R_L > 20$, as is visible in the maps on the top row of Fig. 16, and $R_D < -5$ for most pixels. For Lake Siling, the RMS_430 peak is barely visible, while for Lake Nam and Lake Issyk-Kul part of the water pixels have $Q^{430}_{\mathrm{RMS}} \geq 2.0$ (red patches in the bottom row maps of Fig. 16).

The lower RMS_430 ratio over these lakes as opposed to the ratio over the Atlantic Ocean indicates that for the lakes the effects of VRS are less important and in some cases apparently absent. Liquid water fit coefficients over the Atlantic Ocean are usually around several metres (cf. top-right panel of Fig. 4), whereas for Lake Siling and Lake Issyk-Kul they are around $+1\,\mathrm{m}$, while Lake Nam has values around $-3.5\,\mathrm{m}$; the latter is another indication that the DOAS fit is not really going well here. Differences in the overall broad-band structures in the fit residual may be due to the fact that the characteristics of the material dissolved in the water of the lakes differs from what is dissolved in the oceans.

Looking at results for the three lakes when using the NO2-gap fit introduced above reveals that for Lake Siling the changes are small and erratic: for some pixels the RMS and SCD error and the GCD value go down by one or two percent, while for other lake pixels they go up a little. For Lakes Nam and Issyk-Kul, the RMS and SCD error appears to go down by five to ten percent, and the GCD value by one to four percent.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

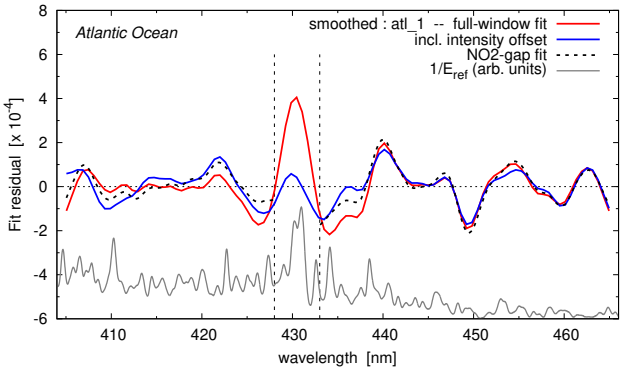

Figure 17. Smoothed fit residual of the Atlantic Ocean 'atl_1' example in Fig. 3 using the full fit window for the retrieval without (red line) and with (blue line) the IOC, and with the NO2-gap fit (black dotted line). The thin grey line at the bottom is $1/E_{\rm ref}$ in arbitrary units (cf. bottom panel in Fig. 1), indicative of the shape of the IOC term.

6.2 Intensity offset correction

Several DOAS applications include an intensity offset correction (IOC for short), constant or linear in wavelength, to improve the retrievals in some spectral ranges. This correction has been implemented in the TROPOMI $\mathrm{NO_2}$ processor in the form of an addional term to Eq. (2):

$$R_{\rm mod}(\lambda) = \cdots + \frac{\sum c_m \lambda^m}{E_0(\lambda)}, \qquad m = 0, 1, \ldots, n_{\rm off} \tag{8}$$

with c_m fit parameters; in most applications $n_{\rm off} = 0$ or 1. The option has, however, not been turned on, mainly because the precise physical origin of such an intensity offset is not known – it is thought to be related to instrumental issues (e.g. incomplete removal of stray light or dark current in Level-1b spectra, neither of which is deemed necessary for TROPOMI measurements) and/or atmospheric issues (e.g. incomplete removal of Ring spectrum structures and VRS in clear ocean waters); see, for example, Platt and Stutz (2008), Richter et al. (2011), Lampel et al. (2015), Boersma et al. (2018).

The blue line in Fig. 17 shows the smoothed fit residual in case the IOC is turned on in the full-window fit with $n_{\rm off} = 0$: the IOC clearly decreases the $430\,\mathrm{nm}$ issue and gives close to it a slightly better fit residual, quite similar to the NO2-gap fit (black dotted line); the shape of the IOC term is indicated by the thin grey line at the bottom of the panel. The change in the $\mathrm{NO_2}$ SCD value, SCD error, RMS error and χ^2 for the six Atlantic Ocean example pixels are listed in Table 5 – these changes are similar to but somewhat less than those for the NO2-gap fit listed in Table 2.

The IOC is proportional to one over the irradiance and the Fraunhofer peak in the irradiance is much narrower than the $430\,\mathrm{nm}$ residual issue, hence the IOC can never fully compensate for the issue. In addition to that, the $1/E_0(\lambda)$ spectrum has additional structures and shows a slope above $440\,\mathrm{nm}$, which may be introducing artifacts in the retrieval.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

Table 5. Changes in the main retrieval results in percent for the six clear-sky Atlantic Ocean pixels of Fig. 3 due to the intensity offset correction (IOC).

	IOC minus full-window fit [%]			
atl_#	N_s	ΔN_s	R_{RMS}	χ^2
1	−3.64	−12.76	−17.74	−26.36
2	−4.44	−7.37	−10.63	−16.46
3	−1.08	−11.58	−15.73	−23.83
4	−3.55	−8.29	−13.03	−18.29
5	−3.22	−12.06	−16.60	−24.96
6	−3.28	−11.35	−17.00	−23.89

All things considering, the NO2-gap fit seems to be a physically better justifiable approach than including the IOC in the fit to solve the $430\,\mathrm{nm}$ issue. Neither of the two approaches, however, are able to deal with the broad-band structures at higher wavelengths.

6.3 Chlorophyll and other material in the water

The NO2-gap fit was introduced to remove the systematic residual feature around $430\,\mathrm{nm}$. This does not deal with issues at higher wavelengths visible in the fit residuals (e.g. Fig. 9), which are related on the one hand to wavelength shifts caused by VRS effects (Peters et al., 2014; Dinter et al., 2015; Holtrop et al., 2021) and on the other hand by the presence in the water of, for example, chlorophyll and dissolved organic matter (DOM), which are known to absorb light in the visible region (Joiner et al., 2004; Cannizzaro et al., 2006; Vountas et al., 2007), and particulate matter that may be scattering UV radiation. Chrolophyll and DOM come in different flavours, each with their own slightly different reference spectrum (Taniguchi et al., 2021). It is unclear whether these structures remaining in the fit residual mean that the retrieved NO_2 SCD is affected, but the results of the lakes discussed in Sect. 6.1 suggest they may well be.

The broad-band structures visible in the fit residual themselves do not represent the missing reference spectrum, because the shape of the residual is the result of DOAS adjusting all fit parameters so as to minimise the residual, i.e. possibly using incorrect fit coefficients – cf. the downward peak in Fig. 15 due to the large negative water vapour coefficient. Note further that the reference spectra $\sigma_k(\lambda)$ are in the exponent of the modelled reflectance of Eq. (2), rather than directly in the residual.

Over relatively small areas like the lakes in Sect. 6.1, where one may assume the stratosphere to be more or less constant, one could try to reconstruct the missing reference spectrum by assuming that the fit coefficients found outside the lake are valid over the lake as well, with some sensible assumption of what the liquid water coefficient might be, although it is then unclear what one should assume for the polynomial coefficients. An attempt in that direction falls outside the scope of the present paper. In addition, the approach would not work over oceans, as there are no neighbouring values for the fit coefficients available.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

7 Concluding remarks

The first step in the process to determine NO_2 tropospheric and stratospheric columns from measurements by satellite-based instruments, such as TROPOMI, is the retrieval of the so-called slant column density (SCD) of NO_2 with a DOAS approach, in the case of TROPOMI NO_2 in the wavelength window $405-465\,\mathrm{nm}$. This retrieval step accounts for the presence of several absorption and scattering effects that occur along the light path from Sun through atmosphere to satellite and has proven to be quite successful and robust. Since TROPOMI measurements have a higher signal-to-noise ratio (SNR) and higher spatial resolution than earlier instruments, some hitherto weak or even unobserved features have been identified in TROPOMI retrievals.

The spectrum of the incoming solar light has a number of Fraunhofer lines. In the atmosphere rotational Raman scattering (RRS; a.k.a. the Ring effect) leads to filling-in, widening and shifting of these lines, which is accounted for in the DOAS retrieval by way of a reference spectrum (I_{ring}). Over open, clear water bodies the light that reaches the satellite may also be influenced by vibrational Raman scattering (VRS) in the water, the characteristics of which depend on, for example, the viewing geometry and the material that is present in the water (such as chlorophyll), as a result of which it appears not possible to compensate for VRS in the DOAS retrieval by way of a scalable reference spectrum.

Close inspection of TROPOMI fit residuals, the difference between the measured and DOAS-modelled spectrum, revealed a distinct peak around $430\,\mathrm{nm}$, which is associated with Fraunhofer lines. Apparently, this feature is not compensated for completely by I_{ring}, while over water VRS may further strengthen the $430\,\mathrm{nm}$ peak in the residual. The lower the overall noise level of the measurements (i.e. the higher the SNR), the more the systematic feature shows up, potentially leading to incorrect NO_2 SCD values. In addition, the depth of the $430\,\mathrm{nm}$ Fraunhofer line depends in the solar cycle: the less active the Sun is, the deaper the line and hence the larger the influence on NO_2 retrieval results.

The $430\,\mathrm{nm}$ peak primarily occurs over open water bodies, where both RRS and VRS play a role, but also for some more scattered ground pixels over land, such as Australian shrublands. The residual of the latter ground pixels outside the peak wavelengths is close to zero, which makes the peak stand out clearly.

As a way to prevent the $430\,\mathrm{nm}$ peak from influencing the NO_2 SCD retrieval results, we investigated the impact of excluding the wavelength range $428-433\,\mathrm{nm}$ from the DOAS fit. Over open water bodies this NO2-gap approach leads, on average, to a $10-20\%$ decrease of the SCD error and of the RMS error of the fit – a reduction of these two is generally considered to indicate a better fit with a more reliable SCD value – and the NO_2 SCD is reduced by a few percent. For some land pixels the approach may lead to a reduction of the SCD and RMS error of $5-10\%$, while the SCD may decrease or increase a few percent. For land pixels where the $430\,\mathrm{nm}$ peak is not prominent, the NO2-gap fit does not alter the results significantly.

OMI retrievals over the Pacific Ocean also show a decrease of the retrieval errors in the NO2-gap fit, but the improvements are smaller than for TROPOMI, which is likely due to the lower SNR of OMI, as a result of which the $430\,\mathrm{nm}$ peak is less pronounced.

The largest decreases of the SCD are seen over the oceans, where most of the NO_2 will be in the stratosphere, as a result of which the stratospheric NO_2 column decrease as well, by up to $-2\,\mu\mathrm{mol\,m^{-2}}$ on average in the tropics. A change in both

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

the SCD and the stratospheric VCD lead to a decrease of the tropospheric NO_2 column for part of the water scenes while for some land scenes the tropospheric column may increase; on average these changes are not very big ($\pm 2\,\mu mol\,m^{-2}$), but for individual clear-sky location the changes may as large as $\pm 20\,\mu mol\,m^{-2}$.

The NO2-gap approach reduces the problem directly related to the $430\,nm$ peak, linked in part to "pure" VRS effects on Fraunhofer lines, i.e. effects solely related to scattering in water. It does not deal with the broad-band features at wavelengths above that in the fit residual related to material dissolved in the water, such as chlorophyll, which also affects VRS characteristics. The impact of chlorophyll, dissolved organic matter and other substances in water bodies on NO_2 retrievals – clearly shown by the problem with NO_2 retrievals over some Tibetan lakes, where VRS effects seem to be small – is difficult to assess and requires dedicated studies.

Note that both the fit residual and the RMS_430 ratio used in this paper are not part of the nominal level-2 NO_2 files; they are only available in special (re)processing exercises. The results of the runs test, introduced earlier when investigating retrievals over Tibetan lakes, are still valuable in the NO2-gap approach and therefore remain available in the nominal orbit files.

The solution of the $430\,nm$ peak issue presented in this paper is implemented in the forthcoming TROPOMI and OMI NO_2 processor versions, as mentioned in Sect. 2.1. The solution is also relevant for the NO_2 retrieval of recently launched and future missions with high spatial resolution, such as Sentinel-5 (the first of which was launched on 12 August 2025), CO2M and TANGO. It may further be worthwhile to investigate whether the $430\,nm$ issue also occurs in NO_2 retrievals of preceeding low spatial resolution missions, such as GOME-2, SCIAMACHY and GOME.

Appendix A: Depth of the 430nm Fraunhofer structure

The structure in the irradiance around $430\,nm$ discussed in this paper is caused by iron and calcium absorption lines, as mentioned in the Introduction (Sect. 1). The depth and width of some Fraunhofer lines in the solar spectrum is known to depend on the activity of the Sun, in particular calcium lines (B. Van den Oord, pers. comm., 2025; see e.g. Marchenko et al. (2024), Chatzistergos et al. (2024), Srinivasa et al. (2025) and references therein), while other lines vary little if at all over time.

As a quick check of this characteristic, the depth of the $430\,nm$ structure is defined as illustrated in the top panel of Fig. A1: the distance between the lowest irradiance value in the structure and the point above that (black stars) along a linear fit (grey) through the spectral points in the wavelength ranges $[424.0 : 428.0]$ and $[432.0 : 436.0]\,nm$.

The evolution over time of this depth is shown in red (left axis) in the bottom panel of Fig. A1, derived from TROPOMI irradiance measurements every 225-th orbit (about once every 15 days), starting with orbit 02818 of 30 April 2018 (the first publicly available TROPOMI irradiance), along with the solar activity in terms of the sunspot number (blue line, right axis, increasing downward): the higher the sunspot number, i.e. the more active the Sun, the less deep is the $430\,nm$ structure. The depth of the two $Ca+$ Fraunhofer lines near $395\,nm$ (involved in effects of VRS; cf. Sect. 1) vary in a similar way with solar activity (not shown).

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

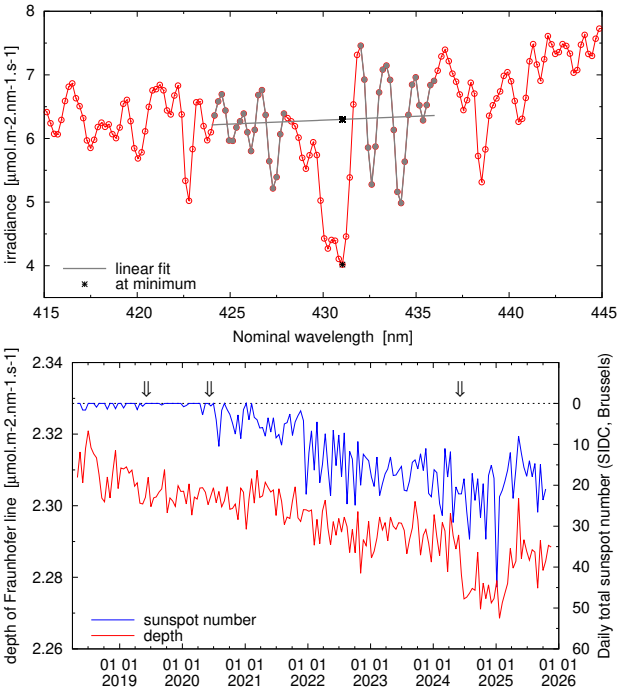

Figure A1. Top panel: definition of the depth of the Fraunhofer line structure around 430 nm using the TROPOMI irradiance of orbit 08518 of 5 June 2019. Bottom panel: change over time of that depth for selected irradiance measurements (red line, left axis) and the daily sunspot number for the same days (blue line, right axis). The arrows at the top point to the dates used in Sects. 3.4 and 4.2. Further details are given in the text of App. A. Source of the sunspot number data: WDC-SILSO, Royal Observatory of Belgium, Brussels.

Appendix B: Indices of example ground pixels

Tables B1 and B2 provide a list of the scanline and row numbers of the example ground pixels used in this paper, labelled by a three-letter identifier.

Author contributions. JvG conducted the research described in this paper and is responsible for the text, which has been read and approved by all co-authors. HE is responsible for the AMF and VCD steps and the final data product. MS and MtL implemented and tested the retrieval code in the TROPOMI processor and performed some dedicated runs. JPV is involved in retrieval issues and is the PI of TROPOMI.

Competing interests. The authors declare that they have no conflict of interests.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

Table B1. Scanline and row number of Atlantic Ocean ('atl') and Western Australia ('aus') ground pixels used in the examples of 5 June 2019; orbit numbers are given between parenthesis in the header line.

	Atlantic 'atl' (08516)		Australia 'aus' (08510)	
#	scan	row	scan	row
1	1583	367	755	209
2	1591	329	757	297
3	1604	426	766	264
4	1613	264	776	284
5	1633	146	777	141
6	1660	199	777	219

Table B2. As Table B1 but for the lake pixel used in Sect. 6.1.

	Lake Siling 'sil' (08511)		Lake Nam 'nam' (08511)		Lake Issyk-Kul 'iss' (08511)	
#	scan	row	scan	row	scan	row
1	1687	376	1670	393	1873	200
2	1688	376	1670	394	1874	205
3	1688	377	1671	393	1875	193
4	1689	374	1671	394	1875	208
5	1689	375	1671	395	1876	197
6	1690	376	1672	391	1876	215

Data availability. Standard TROPOMI NO_2 collection 03 data (v2.4.0 and onwards) are available via ESA's public data hub (https://dataspace.copernicus.eu/); data product DOI: https://doi.org/10.5270/S5P-9bnp8q8. OMI/QA4ECV NO_2 collection 03 (v1.1) data are available via the TEMIS portal at https://www.temis.nl/airpollution/no2.php. OMI NO_2 collection 04 data will be released publicly in due time. Data produced specifically for this paper is available upon request. Sunspot Number data source is WDC-SILSO, Royal Observatory of Belgium, Brussels, at https://www.sidc.be/silso/datafiles.

Acknowledgements. The authors would like to thank the following people: Andreas Richter and Piet Stammes on general retrieval issues, Folkert Boersma and Benjamin Leune on NO_2 data issues, Erwin Loots and Emiel van der Plas on level-1b issues, Bert van den Oord on solar activity issues, Sander Niemeijer for assisting special processing.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

Sentinel-5 Precursor is a European Space Agency (ESA) mission on behalf of the European Commission (EC). The TROPOMI payload is a joint development by ESA and the Netherlands Space Office (NSO). The Sentinel-5 Precursor ground-segment development has been funded by ESA and with national contributions from The Netherlands, Germany, and Belgium. This work contains modified Copernicus Sentinel-5P TROPOMI data (2018-2025), processed in the operational framework or locally at KNMI.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

References

Barlow, R. J.: Statistics: a guide to the use of statistical methods in the physical sciences, John Wiley & Sons, New York, ISBN: 978-0-471-92295-7, 1989.

Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., Van der A, R. J., Sneep, M., Van den Oord, G. H. J., Levelt, P. F., Stammes, P., Gleason J. F., and Bucsela, E. J.: Near-real time retrieval of tropospheric NO_2 from OMI, Atmos. Chem. Phys., 7, 2013-2128, https://doi.org/10.5194/acp-7-2103-2007, 2007.

Boersma, K. F., Eskes, H. J., Dirksen, R. J., Van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y. and Brunner, D.: An improved retrieval of tropospheric NO_2 columns from the Ozone Monitoring Instrument, Atmos. Meas. Tech., 4, 1905-1928, https://doi.org/10.5194/amt-4-1905-2011, 2011.

Boersma, K. F., Eskes, H. J., Richter, A., De Smedt, I., Lorente, A., Beirle, S., van Geffen, J. H. G. M., Zara, M., Peters, E., Van Roozendael, M., Wagner, T., Maasakkers, J. D., van der A, R. J., Nightingale, J., De Rudder, A., Irie, H., Pinardi, G., Lambert, J.-C. and Compernolle, S.: Improving algorithms and uncertainty estimates for satellite NO_2 retrievals: Results from the Quality Assurance for Essential Climate Variables (QA4ECV) project Variables (QA4ECV) project, Atmos. Meas. Tech., 11, 6651-6678, https://doi.org/10.5194/amt-11-6651-2018, 2018.

Cannizzaro, J. P. and Carder, K. L.: Estimating chlorophyll a concentrations from remote-sensing reflectance in optically shallow waters, Rem. Sens. Environment, 101, 13-24, https://doi.org/10.1016/j.rse.2005.12.002, 2006.

Chance, K. V. and Spurr, R. J. D.: Ring effect studies: Rayleigh scattering, including molecular parameters for rotational Raman scattering, and the Fraunhofer spectrum, Appl. Optics, 36, 5224-5230, https://doi.org/10.1364/AO.36.005224 1997.

Chatzistergos, T., Krivova N. A. and Ermolli, I.: Understanding the secular variability of solar irradiance: the potential of Ca II K observations, J. Space Weather Space Clim., 14, 9, 24 pp., https://doi.org/10.1051/swsc/2024006, 2024.

Crutzen, P. J.: The influence of nitrogen oxides on the atmospheric ozone content, Quart. J. R. Meteorol. Soc., 96, 320-325, https://doi.org/10.1002/qj.49709640815, 1970.

Dinter, T., Rozanov, V.V., Burrows, J.P. and Bracher, A.: Retrieving the availability of light in the ocean utilising spectral signatures of vibrational Raman scattering in hyper-spectral satellite measurements, Ocean Science, 11, 373-389, https://doi.org/10.5194/os-11-373-2015, 2015.

Dobber, M., Voors, R., Dirksen, R., Kleipool, Q. and Levelt, P.: The high-resolution solar reference spectrum between 250 and 550 nm and its application to measurements with the Ozone Monitoring Instrument, Solar Physics, 249, 281-291, https://doi.org/10.1007/s11207-008-9187-7, 2008.

Eskes, H. J., van Geffen, J. H. G. M., Boersma, K. F., Eichmann K.-U.. Apituley, A., Pedergnana, M., Sneep, M., Veefkind, J. P. and Loyola, D.: S5P/TROPOMI Level-2 Product User Manual Nitrogen Dioxide, Report S5P-KNMI-L2-0021-MA, version 4.4.0, 2024-11-08, ESA, https://sentiwiki.copernicus.eu/web/s5p-products (last access: 20 November 2025), 2024.

Fuglestvedt, J. S., Berntsen, T., Isaksen, I. S. A., Mao, H., Liang, X.-Z. and Wang, W.-C.: Climatic forcing of nitrogen oxides through changes in tropospheric ozone and methane, Atmos. Environ., 33, 961-977, https://doi.org/10.1016/s1352-2310(98)00217-9, 1999.

Grainger, J. F. and Ring, J.: Anomalous Fraunhofer line profiles, Nature, 193, 762, https://doi.org/10.1038/193762a0, 1962.

Hendrick, F., Mahieu, E., Bodeker, G. E.. Boersma, K. F., Chipperfield, M. P., De Mazière, M., De Smedt, I., Demoulin, P.. Fayt, C., Hermans, C., Kreher, K., Lejeune, B., Pinardi, G., Servais, C., Stübi, R., Van der A, R., Vernier, J.-P. and Van Roozendael, M.: Analysis

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

of stratospheric NO_2 trends above Jungfraujoch using ground-based UV-visible, FTIR, and satellite nadir observations, Atmos. Chem. Phys., 12, 8851-8864, https://doi.org/10.5194/acp-12-8851-2012, 2012.

Holtrop, T., Huisman, J., Stomp, M., Biersteker, L., Aerts, J., Grébert, T., Partensky, F., Garczarek, L. and Woerd, H. J. V.: "Vibrational modes of water predict spectral niches for photosynthesis in lakes and oceans," Nat Ecol Evol., 5, 55-66, https://doi.org/10.1038/s41559-020-01330-x, 2021.

Joiner, J., Vasilkov, A. P., Flittner, D. E., Gleason, J. F. and Bhartia, P. K.: Retrieval of cloud pressure and oceanic chlorophyll content using Raman scattering in GOME ultraviolet spectra, J. Geophys. Res., 109, D01109, 13 pp., https://doi.org/10.1029/2003JD003698, 2004.

Kong, H. and Lin, J. and Zhang, Y. and Li, C. and Xu, C. and Shen, L. and Liu, X. and Yang, K. and Su, H. and Xu, W.: High natural nitric oxide emissions from lakes on Tibetan Plateau under rapid warming, Nat. Geosci., 16, 474-477, https://doi.org/10.1038/s41561-023-01200-8, 2023.

Kleipool, Q., Ludewig, A., Babić, L., Bartstra, R., Braak, R., Dierssen, W., Dewitte, P.-J., Kenter, P., Landzaat, R., Leloux, J., Loots, E., Meijering, P., van der Plas, E., Rozemeijer, N., Schepers, D., Schiavini, D., Smeets, J., Vacanti, G., Vonk, F. and Veefkind, J. P.: Pre-launch calibration results of the TROPOMI payload on-board the Sentinel-5 Precursor satellite, Atmos. Meas. Tech., 11, 6439-6479, https://doi.org/10.5194/amt-11-6439-2018, 2018

Labzovskii, L.D. and van Geffen, J. and Liu, M. and van der A, R. and de Laat, J. and Leune, B. and Eskes, H. and Lin, X. and Ding, J. and Richter, A.: NO_2 satellite retrievals biased by absorption in water, Nat. Geosci., 17, 972-975, https://doi.org/10.1038/s41561-024-01545-8, 2024.

Lambert, J.-C., Keppens, A., Compernolle, S., Eichmann, K.-U., de Graaf, M., Hubert, D., Langerock, B., Sha, M. K., van der Plas, E., Verhoelst, T., Wagner, T., Ahn, C., Argyrouli, A., Balis, D., Chan, K. L., Coldewey-Egbers, M., De Smedt, I., Eskes, H., Fjæraa, A. M., Garane, K., Gleason, J. F., Granville, J., Hedelt, P., Heue, K.-P., Jaross, G., Koukouli, M.-L., Loots, E., Lutz, R., Martinez Velarte, M. C., Michailidis, K., Pseftogkas, S., Nanda, S., Niemeijer, S., Pazmiño, A., Pinardi, G., Richter, A., Rozemeijer, N., Sneep, M., Stein Zweers, D., Theys, N., Tilstra, G., Torres, O., Valks, P., van Geffen, J., Vigouroux, C., Wang, P. and Weber, M.: Quarterly Validation Report of the Copernicus Sentinel-5 Precursor Operational Data Products, #28: April 2018 – August 2025, S5P MPC Routine Operations Consolidated Validation Report series, Issue 28.00.00, 227 pp., 2025-09-26, https://mpc-vdaf.tropomi.eu/index.php/nitrogen-dioxide/ (last access: 20 November 2025), 2025.

Lampel, J., Frieß, U. and Platt, U.: The impact of vibrational Raman scattering of air on DOAS measurements of atmospheric trace gases, Atmos. Meas. Tech., 8, 3767-3787, https://doi.org/10.5194/amt-8-3767-2015, 2015.

Levelt, P. F., van den Oord, G. H. J., Dobber, M. R. Dobber, Mälkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O. V. and Saari, H.: The Ozone Monitoring Instrument, IEEE Trans. Geosci. Rem. Sens., 44, 1093-1101, https://doi.org/10.1109/tgrs.2006.872333, 2006.

Levelt, P. F., Joiner, J., Tamminen, J., Veefkind, J. P., Bhartia, P. K., Stein Zweers, D. C., Duncan, B. N., Streets, D. G., Eskes, H., van der A, R., McLinden, C., Fioletov, V., Carn, S., de Laat, J., DeLand, M., Marchenko, S., McPeters, R., Ziemke, J., Fu, D., Liu, X., Pickering, K., Apituley, A., González Abad, G., Arola, A., Boersma, F., Chan Miller, C., Chance, K., de Graaf, M., Hakkarainen, J., Hassinen, S., Ialongo, I., Kleipool, Q., Krotkov, N., Li, C., Lamsal, L., Newman, P., Nowlan, C., Suleiman, R., Tilstra, L. G., Torres, O., Wang, H. and Wargan, K.: The Ozone Monitoring Instrument: overview of 14 years in space, Atmos. Chem. Phys., 18, 5699-5745, https://doi.org/10.5194/acp-18-5699-2018, 2018.

Ludewig, A., Kleipool, Q., Bartstra, R., Landzaat, R., Leloux, J., Loots, E., Meijering, P., van der Plas, E., Rozemeijer, N., Vonk, F. and Veefkind, J. P.: In-flight calibration results of the TROPOMI payload on-board the Sentinel-5 Precursor satellite, Atmos. Meas. Tech., 13, 3561-3580, https://doi.org/10.5194/amt-13-3561-2020, 2020.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

Marchenko, S. V., Ludewig, A., Criscuoli, S., Al Moulla, K., Choudhary, D. P., DeLand, M. T., Kopp, G., Loots, E., van der Plas, E. and Veefkind, P.: Sun-as-a-star spectral line variability in the 300-2390 nm wavelength range, Astrophysical Journal, 977, 17 pp, https://doi.org/10.3847/1538-4357/ad888f, 2024.

Murphy, D. M., Fahey, D. W., Proffitt, M. H., Liu, S. C., Chan, K. R., Eubank, C. S., Kawa, S. R. and Kelly, K. K.: Reactive nitrogen and its correlation with ozone in the lower stratosphere and upper troposphere, J. Geophys. Res.. 98, 8751-8773, https://doi.org/10.1029/92JD00681, 1993.

Peters, E., Wittrock, F., Richter, A., Alvarado, L. M. A., Rozanov, V. V., and Burrows, J. P. Liquid water absorption and scattering effects in DOAS retrievals over oceans, Atmos. Meas. Tech., 7, 4203-4221 https://doi.org/10.5194/amt-7-4203-2014, 2014.

Platt, U.: Differential Optical Absorption Spectroscopy (DOAS), in: Air monitoring by spectroscopic techniques, ed. M. W. Sigrist, Chemical Analisys Series, 127, 27-76, Wiley, New York, 1994.

Platt, U. and Stutz, Z.: Differential Optical Absorption Spectroscopy, Principles and Applications, Springer, Heidelberg, Germany, 597 pp., https://doi.org/10.1007/978-3-540-75776-4, 2008.

Richter, A., Begoin, M., Hilboll, A. and Burrows, J. P.: An improved NO_2 retrieval for the GOME-2 satellite instrument, Atmos. Meas. Tech., 4, 1147-1159, https://doi.org/10.5194/amt-4-1147-2011, 2011.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific Publishing, ISBN: 978-9-810-22740-1, https://doi.org/10.1142/3171, 2000.

Schenkeveld, V. M. E., Jaross, G., Marchenko, S., Haffner, D., Kleipool, Q. L., Rozemeijer, N. C., Veefkind, J. P., and Levelt, P. F.: In-flight performance of the Ozone Monitoring Instrument, Atmos. Meas. Tech., 10, 1957-1986, https://doi.org/10.5194/amt-10-1957-2017, 2017.

Seinfeld, J. H. and Pandis, S. N.: Atmospheric Chemistry and Physics - From Air Pollution to Climate Change (2nd Edition), John Wiley & Sons, ISBN: 978-1-118-94740-1, 2006.

Sillman, S., Logan, J. A. and Wofsy, S. C.: The sensitivity of ozone to nitrogen oxides and hydrocarbons in regional ozone episodes, J. Geophys. Res., 95, 1837-1851, https://doi.org/10.1029/JD095iD02p01837, 1990.

Srinivasa, A., Sreedevi, A, Raju, K. P., Nagaraju, K., Singh, J., Karuppath, N., Devendran, P., Kumar, T. R. and Kumaravel, P.: Variations of the Ca ii K line profile parameters with solar latitude and time observed from Kodaikanal Solar Observatory, Mon. Not. R. Astron. Soc., 541, Issue 4, August 2025, 3320-3330, https://doi.org/10.1093/mnras/staf1163, 2025.

Taniguchi, M. and Lindsey, J. S.: Absorption and fluorescence spectral database of chlorophylls and analogues, Photochem. Photobiol., 97, 136-165, https://doi.org/10.1111/php.13319, 2021.

van der A, R. J., Ding, J. and Eskes, H. J.: Monitoring European anthropogenic NOx emissions from space Atmos. Chem. Phys., 24, 7523-7534, https://doi.org/10.5194/acp-24-7523-2024, 2024.

van Geffen, J. H. G. M., Boersma, K. F., Eskes, H. J., Sneep, M., ter Linden, M., Zara, M. and Veefkind, J. P.: S5P/TROPOMI NO_2 slant column retrieval: method, stability, uncertainties and comparisons with OMI, Atmos. Meas. Tech., 13, 1315-1335, https://doi.org/10.5194/amt-13-1315-2020, 2020.

Van Geffen, J. H. G. M., Eskes, H. J., Compernolle, S., Pinardi, G., Verhoelst, T., Lambert, J.-C., Sneep, M., ter Linden, M., Ludewig, A., Boersma, K. F. and Veefkind, J.P.: Sentinel-5P TROPOMI NO_2 retrieval: impact of version v2.2 improvements and comparisons with OMI and ground-based data, Atmos. Meas. Tech., 15, 2037-2060. https://doi.org/10.5194/amt-15-2037-2022, 2022.

van Geffen, J. H. G. M., Eskes, H. J., Boersma, K. F. and Veefkind, J. P.: TROPOMI ATBD of the total and tropospheric NO_2 data products, Report S5P-KNMI-L2-0005-RP, version 2.9.1, 2025-10-27, KNMI, De Bilt, The Netherlands, https://sentiwiki.copernicus.eu/web/s5p-products (last access: 20 November 2025), 2025.

https://doi.org/10.5194/egusphere-2025-5836
Preprint. Discussion started: 12 February 2026

van Geffen, J., Eskes, H., Boersma, F., ter Linden, M. and Veefkind, P.: Algorithm Theoretical Basis Document for the OMI L2 OMNO2A Data Processor Collection 4. Report AURA-OMI-KNMI-L2-0300-SD, version 4.0.0.x, KNMI, De Bilt, The Netherlands, in review, *landing page of DOI to be published*, 2026.

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J.F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H. and Levelt, P. F.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Rem. Sens. Environment, 120, 70-83, https://doi.org/10.1016/j.rse.2011.09.027, 2012.

Vasilkov, A. P., Joiner, J., Gleason, J. and Bhartia, P. K.: Ocean Raman scattering in satellite backscatter UV measurements, Geophys. Res. L., 29, 1837-1840, https://doi.org/10.1029/2002GL014955, 2002.

Vountas, M., Richter, A., Wittrock, F. and Burrows, J. P. Inelastic scattering in ocean water and its impact on trace gas retrievals from satellite data, Atmos. Chem. Phys., 3, 1365-1375, https://doi.org/10.5194/acp-3-1365-2003, 2003.

Vountas, M., Dinter, T., Bracher, A., Burrows, J. P. and Sierk, B.: Spectral studies of ocean water with space-borne sensor SCIAMACHY using Differential Optical Absorption Spectroscopy (DOAS), Ocean Science, 3, 429-440, https://doi.org/10.5194/os-3-429-2007, 2007.

World Health Organisation: Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide, World Health Organisation, Bonn, Germany, EUR/03/5042688, 98 pp., https://iris.who.int/handle/10665/107478 (last access: 20 November 2025), 2003.

Zara, M., Boersma, K. F., De Smedt, I., Richter, A., Peters, E., Van Geffen, J. H. G. M., Beirle, S., Wagner, T., Van Roozendael, M., Marchenko, S., Lamsal, L. and Eskes, H.J.: Improved slant column density retrieval of nitrogen dioxide and formaldehyde for OMI and GOME-2A from QA4ECV: intercomparison, uncertainty characterization, and trends, Atmos. Meas. Techn. 11, 4033-4058, https://doi.org/10.5194/amt-11-4033-201, 2018.