

We are grateful to the reviewers for their thorough evaluations and constructive suggestions regarding the manuscript. Their comments have been carefully considered and have led to substantial improvements to the paper. A detailed response addressing each question or comment is provided below in *blue*. All modifications introduced in the revised version are indicated in *red*.

## Anonymous reviewer #1

This work builds on the previously reported LUT-COBRA algorithm (Theys et al., 2022) for volcanic SO<sub>2</sub> plume height retrievals, by applying it to the UV2 measurements from TROPOMI. Through theoretical calculations and sensitivity tests, the authors demonstrated that the TROPOMI UV2 measurements at shorter wavelengths (~305 nm) can provide better accuracy for SO<sub>2</sub> layer height than UV3 measurements. The authors also described the implementation of the LUT-COBRA algorithm with the TROPOMI UV2 band, and compared the retrieved plume heights with those from the operational TROPOMI SO<sub>2</sub> height algorithm (based on a machine learning technique) as well as those from infrared (IASI) and microwave limb (MLS) sensors. Overall, this is a well-written paper, and the topic should be of interest to the atmospheric science and remote sensing communities. While the technique has been previously described, the application to TROPOMI UV2 clearly shows improvement in the retrieved SO<sub>2</sub> heights. The paper can be further improved by addressing some technical points (see specific, mostly minor comments below) and I'd recommend minor revisions before the paper can be accepted for publication in AMT.

### Specific comments:

- Lines 140-142: In Jacobian calculations, the authors used LIDORT (instead of VLIDORT) and did not explicitly consider aerosols or rotational Raman scattering. Can the authors comment on the uncertainties in the retrieved SO<sub>2</sub> plume heights associated with these factors?

→ First of all, using LIDORT instead of VLIDORT has a large benefit in terms of computational time but introduces errors in TOA radiances due to the neglect of polarization, typically below 5% depending on the viewing geometry and spectral window, with smaller effects at shorter wavelengths (Lerot et al., 2014). These small differences in the radiative transfer mainly vary smoothly as a function of wavelength and largely cancel out when the SO<sub>2</sub> SOD is computed (as it is based on logarithmic intensity ratios), therefore suggesting only a minor impact on the retrieved plume heights.

Then, RRS arises from inelastic scattering, which partially fills in Fraunhofer and molecular absorption lines (Ring effect). As reported by Fedkin et al. (2021), including RRS modifies the TOA radiances by only a few percent. For the SO<sub>2</sub> LUT, this effect is therefore expected to be smaller than the uncertainties from ozone or aerosols. Note that the typical Ring effect signature is not only small compared to the SO<sub>2</sub> signal but it is also well represented in the set of clean spectra used to generate the covariance matrix, such that its interference in the spectral fit is very small.

The contribution of aerosols is a recognized limitation for UV SO<sub>2</sub> measurements (Theys et al., 2022). However, under most observation conditions, when the aerosol layer lies below the SO<sub>2</sub> plume, the associated errors remain moderate (< 15%), as noted by Yang et al. (2010). Neglecting aerosols could lead to errors of up to 0–5 km in the retrieved plume heights when the aerosol layer is collocated with the SO<sub>2</sub> plume, and we made this point clear in the revised version of the paper.

In summary, while these factors may introduce small to moderate uncertainties, they were deliberately omitted in this study to maintain a simple and controlled synthetic framework with convenient computational time, but it could be addressed in future algorithm developments.

The manuscript has been updated accordingly.

### **Changes in lines 144–149**

**Submitted version:** "To evaluate the performance of LUT-COBRA across different wavelength ranges (*i.e.*, spectral bands), a set of synthetic spectra was generated using the Linearized Discrete Ordinate Radiative Transfer (LIDORT) model (Spurr and Christi, 2019) under typical TROPOMI measurement conditions. Simulations were performed between 300 and 330 nm, with a spectral sampling of 0.065 nm for BD2 and 0.2 nm for BD3."

**Revised version:** "To evaluate the performance of LUT-COBRA across different wavelength ranges (*i.e.*, spectral bands), a set of synthetic spectra was generated using the Linearized Discrete Ordinate Radiative Transfer (LIDORT) scalar model (Spurr and Christi, 2019). Although LIDORT neglects polarization effects, these typically alter TOA radiances by less than 5% (Lerot et al., 2014; Escribano et al., 2019), mainly in a spectrally smoothed way, suggesting only a minor influence on the retrievals. Simulations were performed under typical TROPOMI measurement conditions between 300 and 330 nm, with a spectral sampling of 0.065 nm for BD2 and 0.2 nm for BD3."

### **Comments added in lines 151-154**

**Revised version:** "While these factors may introduce uncertainties of a few percent in the radiative transfer (Yang et al., 2010; Fedkin et al., 2021), and up to 5 km bias on the SO<sub>2</sub> LH for highly absorbing aerosols collocated with SO<sub>2</sub>, our aim was to maintain a simple and controlled synthetic framework."

### **Comments added in lines 483-484**

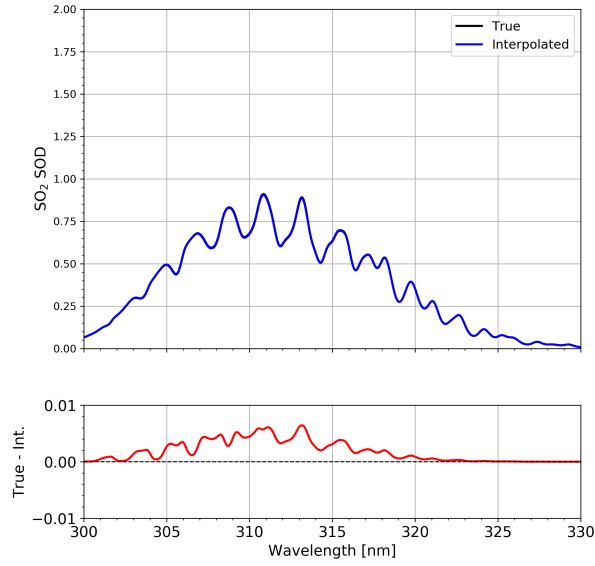
**Revised version:** "As discussed in Section 3.2, aerosols may introduce non-negligible additional uncertainties in the retrieved SO<sub>2</sub> heights when collocated with the plume."

- Lines 149-150 and Table 3: the interpolation error for Jacobians at relatively large SO<sub>2</sub> VCDs can be quite substantial especially at short wavelengths – can the authors comment on the selected SO<sub>2</sub> VCD nodes in the LUTs?

→ We thank the reviewer for raising this important point. To assess this, we performed a comparison between SO<sub>2</sub> slant optical depths (SOD) simulated at an intermediate value of 175 DU and obtained by interpolation from the LUT nodes at 150 and 200 DU. As shown in Figure 1, the interpolated spectrum reproduces the true SO<sub>2</sub> signal accurately under typical conditions, with differences below  $5 \times 10^{-3}$  (unitless). Although larger interpolation errors can occur under extreme conditions, they are expected to remain minor compared to other sources of uncertainty (*e.g.*, aerosols). A brief discussion on this aspect has been added in the revised manuscript.

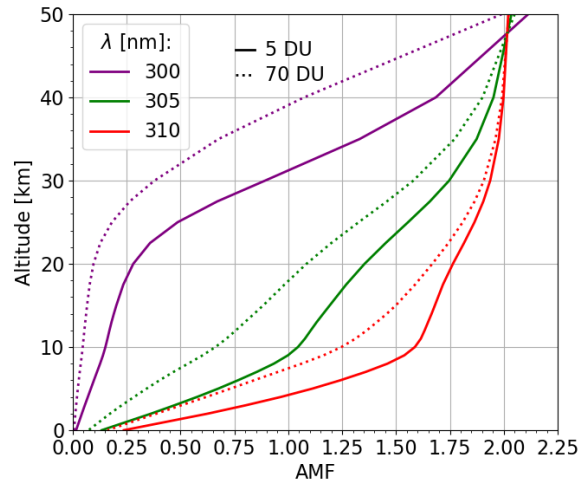
### **Comments added in lines 158-161**

**Revised version:** "For high SO<sub>2</sub> columns, a coarser sampling of VCD nodes was used to limit the total number of simulations. However, sensitivity tests indicate that interpolation errors in the SODs and Jacobians are negligible under these conditions, with differences of less than 0.5% in the radiative transfer, even at shorter UV wavelengths."



**Figure 1.** Comparison of  $\text{SO}_2$  SOD at 175 DU with the value interpolated from LUT nodes at 150 and 200 DU. Simulations were performed under the following conditions:  $\text{SO}_2$  LH = 10 km, SZA =  $10^\circ$ , VZA = RAA =  $0^\circ$ , albedo = 5%, surface height = 0 km, using US Standard atmospheric profiles (Anderson et al., 1986).

- Figure 5: is the apparent increase in error for SO<sub>2</sub> LH > 25 km due to the coarse resolution of LH nodes in the LUT?  
 → We chose to use a coarser grid at high altitudes to limit the number of simulations. A finer grid could be implemented in future versions of the algorithm and might slightly improve the results. However, this is not expected to significantly affect the findings, since the sensitivity to SO<sub>2</sub> height, although higher in BD2, remains generally lower in the stratosphere for UV measurements, as shown by the AMFs in Figure 2. No change in the revised manuscript has been made on this aspect.



**Figure 2.** SO<sub>2</sub> AMF distributions as a function of the altitude, for different wavelengths and SO<sub>2</sub> VCDs.

- Figure 6 and lines 239-240: some of the error sources are not completely independent (e.g., T prof. and Air prof., O<sub>3</sub> prof. and O<sub>3</sub> VCD). Not sure if quadrature summation is justified here.  
 → We thank the reviewer for this valuable comment. We acknowledge that some of the error sources are partially correlated, and therefore a strict quadrature summation may not be fully justified. Nevertheless, we chose this approach to provide a rough estimate of the overall uncertainty by combining all relevant sources. In practice, the resulting uncertainty should be interpreted as indicative rather than exact, and likely represents a conservative estimate of the total error. We have clarified this point in the revised manuscript.

### **Changes in lines 250–253**

**Submitted version:** "Overall, the total systematic errors, obtained by combining the individual effects of each source of uncertainty in quadrature, are 1.91 km for the LH and 2.58% for the VCD."

**Revised version:** "Overall, the total systematic errors are estimated by combining the individual effects of each source of uncertainty in quadrature, yielding 1.91 km for the LH and 2.58% for the VCD. Since some contributions are partially correlated (e.g., temperature and air density profiles, or the O<sub>3</sub> profile and its VCD), these values should be interpreted as indicative and approximate estimates of the total error."

- Section 4.1: The construction of the weighted O<sub>3</sub> profiles was done to reduce the size of the LUTs, right? How does this affect O<sub>3</sub> profile related errors?  
 → Yes, the weighted O<sub>3</sub> profiles were designed to include the main variability of O<sub>3</sub> columns and profiles without requiring a full set of O<sub>3</sub> profile information in the LUT, effectively reducing its size. In practice, for a given O<sub>3</sub> VCD, profiles that deviate strongly from the others (and would thus introduce larger uncertainties) are weighted less in the mean profile. In other words, the weighted profiles are primarily built from the most probable O<sub>3</sub> profiles for a given

VCD, which indirectly reduces the associated uncertainty. Based on synthetic tests, a rough estimate of the impact on SO<sub>2</sub> retrievals is a reduction of the uncertainties by approximately 0.5–1 km. Overall, this provides a good compromise between LUT size and retrieval accuracy.

This has been clarified in the revised version of the manuscript.

### Comments added in lines 275-276

**Revised version:** "This was motivated by the need to limit the LUT size while adequately representing the ozone profile variability in the retrievals."

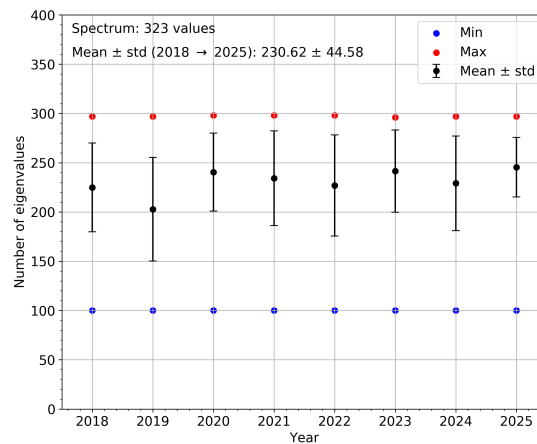
### Changes in lines 287-290

**Submitted version:** "Similarly, Figure 7b gives the obtained mean ozone distributions for various column densities. It can be seen that the O<sub>3</sub> peak shifts towards lower altitudes as its concentration increases, providing valuable information about the profile shape."

**Revised version:** "Figure 7b provides the calculated mean ozone distributions for various column densities. They thus correspond to the most probable profiles for such O<sub>3</sub> VCDs, and it can be seen that the O<sub>3</sub> peak shifts towards lower altitudes as its concentration increases, providing valuable information about the profile shape that helps reduce the associated uncertainties in the SO<sub>2</sub> retrievals."

- Lines 315-318: what is the typical number of eigen vectors used in equation 9? Is this done for all spectra (or just those with ill-conditioned covariance matrix)?

→ In practice, the covariance matrix is inverted using an eigenvalue decomposition with a fixed threshold ( $10^{-7}$ ), corresponding approximately to the precision of single-precision floating point (float32), which is the format of the input data (residuals). Eigenvalues smaller than this threshold are discarded to stabilize the inversion. This procedure is applied systematically to all spectra. The number of retained eigenvectors is therefore not fixed and depends on the conditioning of the covariance matrix. Figure 3 shows the typical number of eigenvalues retained for each year of TROPOMI data. On average, approximately 230 eigenvalues are kept out of 323 per spectrum. We have clarified this point in the revised manuscript.



**Figure 3.** Typical number of eigenvalues retained per spectrum for each year of TROPOMI data, using a fixed cut-off value of  $10^{-7}$  in the correction of the inverse covariance matrix.

### Changes in lines 333-335

**Submitted version** "We carried out some tests to identify the optimal value for a minimal error, and found a threshold of  $10^{-7}$  on the eigenvalues."

**Revised version:** "Tests were performed to identify the threshold minimizing the errors, and a value of  $10^{-7}$  on the eigenvalues was found, corresponding approximately to the precision of single-precision floating-point residuals. The procedure was thus applied systematically to all spectra, with typically less than 30% of eigenvalues discarded to stabilize the inversion, depending on the conditioning of the covariance matrix."

- Table 4 and Section 4.2: all examples given here appear to be under relatively small or moderate SZAs. Can the authors present some results for higher SZAs, where the O<sub>3</sub> profile effect could be more significant?

→ We understand the reviewer's concern. All cases analyzed with our TROPOMI algorithm in this study indeed correspond to moderate SZAs (approximately 20–45°). These were intentionally selected to limit the influence of ozone and ensure robust SO<sub>2</sub> height retrievals from BD2. While we acknowledge that ozone effects may become more significant at higher SZAs, the primary objective of this work is to demonstrate the development of LUT-COBRA and assess the added value of BD2 compared to BD3 for SO<sub>2</sub> plume height retrievals. The investigation of higher SZA case studies is therefore left for future work and will be addressed using a larger TROPOMI dataset. However, note that additional synthetic sensitivity tests were performed in BD2 for high SZAs, as summarized in Table A1 of the paper. Despite slightly higher uncertainties, these results suggest that scenarios with higher SZAs would lead to conclusions consistent with those presented in Section 4.2. No changes were made to the revised manuscript regarding this point.

- Figure 8: there appears to be a general tendency of pixels with lower VCDs to have lower LHs. Is there an explanation for this?

→ This is a good point. The fact that low SO<sub>2</sub> VCDs have lower LHs is partly real, as SO<sub>2</sub> from Raikoke was mostly injected high in the atmosphere (Vernier et al., 2024; and references therein). No change in the revised manuscript has been made on this aspect.

- Figure 10c: the BD2 retrievals appear to be quite noisy for the eastern branch of the plume – how confident are the authors in the retrieved heights here?

→ The eastern branch (between 37–41°N, 17.5–22°E) corresponds to a thin and diluted filament with low SO<sub>2</sub> columns (on average ~3.5 DU), which naturally leads to noisier pixels. This increased scatter reflects a reduced sensitivity at low SO<sub>2</sub> loadings rather than a systematic bias. Despite the noise, the mean SO<sub>2</sub> layer height retrieved over this region (10.75 km) is in very good agreement with the IASI-C estimate (mean peak height of 10.45 km). The mean retrieval uncertainty for BD2 in this area is about 1.1 km, which supports the statistical robustness of the averaged height estimates. Furthermore, we tested different data selection criteria and found that using slightly higher SO<sub>2</sub> VCD thresholds (e.g., 2.5 or 3 DU) leads to reduced noise while yielding consistent mean plume heights, confirming the results. We are therefore confident that BD2 provides physically meaningful height information for this plume branch, even though individual pixels are noisy. We have updated the manuscript to address this point.

#### **Changes in lines 381–384**

**Submitted version:** "In this region, the scatter on BD2 LH is quite strong (due to the low SO<sub>2</sub> column amount). However, the mean LH and VCD retrieved from BD2 are 10.75 km and 3.53 DU, respectively, while IASI-C estimates a mean peak height of 10.45 km and a column density of 2.32 DU."

**Revised version:** "In this region, the scatter in BD2 LH is quite strong due to the low SO<sub>2</sub> column amount, reflecting a reduced sensitivity. However, the mean plume height and column remain well constrained over the region, with BD2 estimates of 10.75 km and 3.53 DU, respectively, while IASI-C indicates a mean peak height of 10.45 km and a column density of 2.32 DU."

- Figures 11 and 14: it is a bit difficult to directly compare IASI with TROPOMI using these maps. The profile plots (Figures 12 and 15) are more informative for these two cases.

→ The reviewer's remark is appreciated. However, the map-based representations in Figures 11 and 14 were included to provide a spatial overview of the SO<sub>2</sub> plume, while the profile plots in Figures 12 and 15 complement these by showing the vertical distribution in more detail. We believe that presenting both types of results offers a comprehensive view,

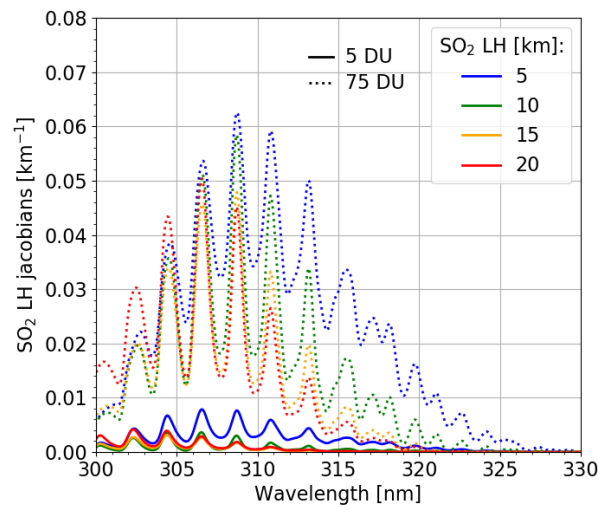
and readers can use the profile plots for more direct comparison between IASI and TROPOMI. Accordingly, we did not modify the manuscript for this point.

## Anonymous reviewer #2

The authors presented an improved SO<sub>2</sub> layer height and vertical column densities retrieval algorithm for TROPOMI's second spectral band (BD2) demonstrated stronger SO<sub>2</sub> absorption sensitivity than the traditional (BD3) approach and plume height estimates consistent with independent IASI and MLS measurements. The paper was organized in three sections, introduction of TROPOMI and SO<sub>2</sub> retrieval methodology, theoretical bases of LUT-COBRA and sensitivity tests, and showing main results for various examples of volcanic events. Because my background is in atmospheric science and I work on modelling, my comments mainly focus on the results and their comparison with satellite observations (Section 4.2).

- Throughout the analysis of the SO<sub>2</sub> height retrievals from BD2 and BD3 over eruption cases, it is observed that SO<sub>2</sub> heights are generally lower in BD2 than in BD3 for common pixels. It seems the authors implicitly explain that BD2 reduces over-sensitivity to noise and brings the retrievals closer to IASI and MLS. I might be missing this somewhere in the manuscript, but a clearer explanation of the higher values of the BD3 values would be helpful. Do these cause artificially inflated heights in BD3?

→ We thank the reviewer for this comment. The differences between both bands are mainly observed for high-altitude SO<sub>2</sub> plumes (above approximately 15 km), where they typically range between 1 and 3 km. At such altitudes, BD3 loses a large part of its sensitivity to SO<sub>2</sub> layer height, as shown by the AMFs (Figure 2) and Jacobians (Figure 4). As a result, the retrieval becomes poorly constrained in height and may converge to different solutions that reproduce the measured signal for a given SO<sub>2</sub> VCD. At lower altitudes, the differences between BD2 and BD3 are on the order of a few hundred metres and are small compared to the retrieval uncertainties. We emphasize, however, that the precise origin of the small difference between BD2 and BD3 at low altitudes remains unclear and may benefit from further investigation in the future, including the impact of aerosols on the retrievals.



**Figure 4.** SO<sub>2</sub> LH Jacobians for different column densities and plume heights.

- With the improved sensitivity of BD2, could the authors comment on whether the retrievals provide any insights even indirectly in the vertical extent or thickness of the volcanic plume from the lower to the upper layers?

→ This is an interesting point. Actually, our LUT-COBRA approach retrieves only two parameters: the SO<sub>2</sub> vertical column density and peak height (*i.e.*, the altitude of maximum SO<sub>2</sub> concentration). No direct information on the vertical extent of the plume is provided. The Jacobians are computed assuming a prescribed vertical SO<sub>2</sub> profile with a fixed shape and width, and thus do not contain sufficient information to constrain the plume thickness. In practice, the algorithm derives a plume height that best reproduces the measured radiances/SODs and therefore, no robust or quantitative inference on the vertical extent of the plume can be determined from the retrievals, even indirectly.

No changes have been made to the manuscript on this point.

- In the conclusion, the authors mention that the operational product fails to reproduce SO<sub>2</sub> heights at the tropical tropopause level, and that the reason for this remains unclear. With respect to temperature inversion and ozone gradients in the tropical tropopause layer, it would be useful if authors could comment on whether these factors might help explain that failure to reproduce SO<sub>2</sub> heights at this level.

→ We thank the reviewer for this comment. While temperature inversions and ozone gradients in the tropical tropopause layer can slightly influence the AMFs, the FP\_ILM algorithm already accounts for typical variations in ozone and atmospheric profiles in its training phase (Hedelt et al., 2019). Therefore, these factors are secondary, and the main limitations of the product at the tropopause are the limited vertical sensitivity of BD3 and low SO<sub>2</sub> columns condition.

The FP\_ILM algorithm of Hedelt et al. (2019) has been trained primarily for relatively large SO<sub>2</sub> columns ( $\geq 15$  DU). Below this threshold, the derived plume heights are less reliable, as retrieval errors exceed 2 km for 15–20 DU and only decrease for higher columns. Note that in this study, we applied a QA threshold of 0.5, as recommended by Hedelt et al. (2023), but the suggested criterion on the LH validity flag, designed to exclude pixels with SO<sub>2</sub> columns of 15–20 DU (and thus higher uncertainties in the plume height), was not used here to keep enough pixels and facilitate the comparison with LUT-COBRA.

A discussion of these points has been added in the revised manuscript.

#### **Changes in lines 373–377**

**Submitted version:** Figure 10 presents SO<sub>2</sub> maps derived from TROPOMI BD3 observations using LUT-COBRA and the operational product, applying the recommended data selection criteria (Hedelt et al., 2023). The two algorithms agree well, displaying similar patterns as well as comparable layer heights and column densities.

**Revised version:** Figure 10 presents SO<sub>2</sub> maps derived from TROPOMI BD3 observations using LUT-COBRA and FP\_ILM. As recommended by Hedelt et al. (2023), a quality assurance threshold of 0.5 was applied to the operational product, while the suggested LH validity flag, designed for pixels with high SO<sub>2</sub> columns (typically above 15 DU) and retrieval errors exceeding 2 km, was omitted to preserve enough pixels and allow a meaningful comparison with LUT-COBRA. Overall, the two algorithms agree well, displaying similar patterns as well as consistent layer heights and column densities.

#### **Changes in lines 477–480**

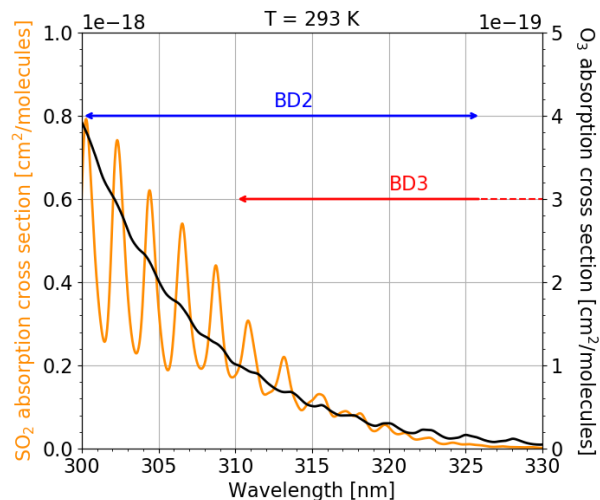
**Submitted version:** "The cases of Ruang and Ulawun also show that our SO<sub>2</sub> height retrievals outperform the operational product. The latter apparently fails to reproduce the SO<sub>2</sub> height at tropical tropopause level. The reason for this is unclear."

**Revised version:** "The cases of Ruang and Ulawun also show that our SO<sub>2</sub> height retrievals outperform the operational product. The latter apparently fails to reproduce the SO<sub>2</sub> height at the tropical tropopause level. The main limitations likely stem from the intrinsically low vertical sensitivity of BD3 at these altitudes and the decreased sensitivity of the FP\_ILM algorithm for lower SO<sub>2</sub> columns."

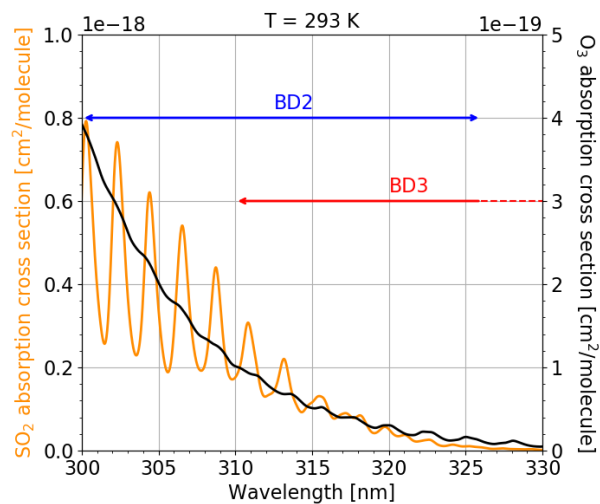


**Additional changes in the revised manuscript:**

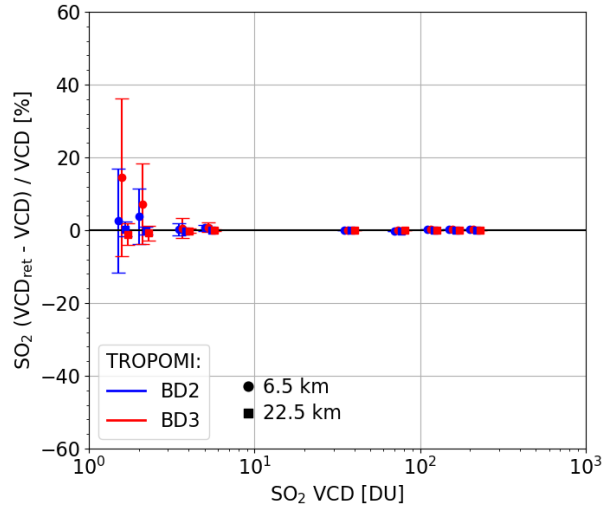
- Consistent use of the term *Jacobians* (uppercase) throughout the manuscript.
- Improved visual formatting of equations, including corrected spacing and typography of quantities (e.g., VCD, LH) and indices.
- Figure 1a (see below): correction of units in y labels:  $\text{cm}^2/\text{molecules}$   $\rightarrow$   $\text{cm}^2/\text{molecule}$ .
- Figure 4b (see below): correction of an error in the standard deviation at each point.
- Figure 7 (see below): the profiles in (b) were not represented on the same altitude grid as those in (a).



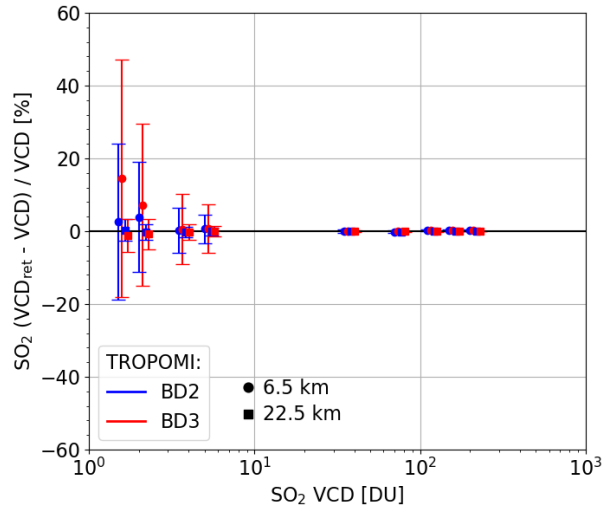
**Figure 5. Submitted version:** Absorption cross sections (ACS) of SO<sub>2</sub> (in orange) from Bogumil et al. (2003) and O<sub>3</sub> (in black) from Serdyuchenko et al. (2014), for a temperature (T) of 293K. For illustration purposes, the data are convolved using a Gaussian function with a standard deviation of 0.5 nm.



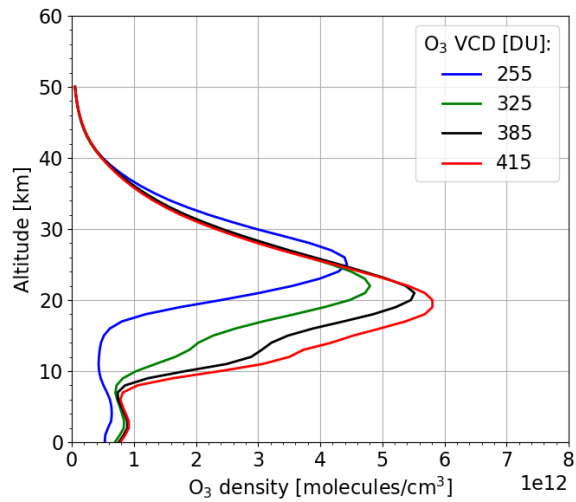
**Figure 6. Revised version:** Absorption cross sections (ACS) of SO<sub>2</sub> (in orange) from Bogumil et al. (2003) and O<sub>3</sub> (in black) from Serdyuchenko et al. (2014), for a temperature (T) of 293K. For illustration purposes, the data are convolved using a Gaussian function with a standard deviation of 0.5 nm.



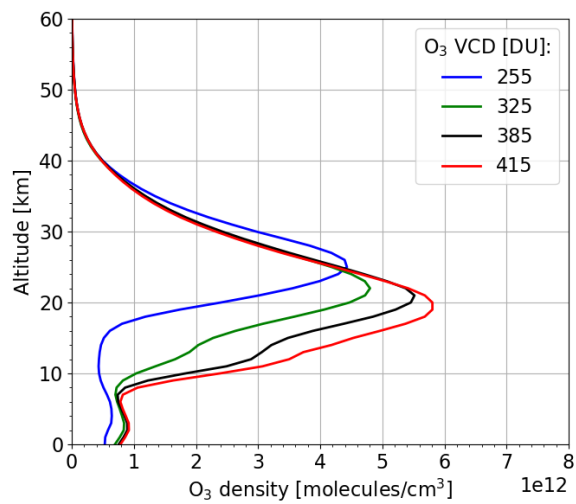
**Figure 7. Submitted version:** Relative difference between the retrieved and expected  $\text{SO}_2$  VCDs, expressed as a percentage, and shown as a function of the input parameters. Retrievals are performed using synthetic spectra with added noise, in BD2 (305–320 nm, blue) and BD3 (310–326 nm, red). Circles and squares indicate the corresponding  $\text{SO}_2$  LHs. Error bars represent the standard deviation associated with each retrieval. Note that the data are slightly offset along the  $x$ -axis for clarity.



**Figure 8. Revised version:** Relative difference between the retrieved and expected  $\text{SO}_2$  VCDs, expressed as a percentage, and shown as a function of the input parameters. Retrievals are performed using synthetic spectra with added noise, in BD2 (305–320 nm, blue) and BD3 (310–326 nm, red). Circles and squares indicate the corresponding  $\text{SO}_2$  LHs. Error bars represent the standard deviation associated with each retrieval. Note that the data are slightly offset along the  $x$ -axis for clarity.



**Figure 9. Submitted version:** Mean O<sub>3</sub> profiles generated for different O<sub>3</sub> VCDs.



**Figure 10. Revised version:** Mean O<sub>3</sub> profiles generated for different O<sub>3</sub> VCDs.

## References

- Anderson, G., Clough, S., Kneizys, F., Chetwynd, J., and Shettle, E.: AFGL Atmospheric Constituent Profiles (0.120km), Optical Physics Division from Air Force Geophysics Laboratory, p. 46, 1986.
- Bogumil, K., Orphal, J., Homann, T., Voigt, S., Spietz, P., Fleischmann, O. C., Vogel, A., Hartmann, M., Kromminga, H., Bovensmann, H., Frerick, J., and Burrows, J. P.: Measurements of molecular absorption spectra with the SCIAMACHY pre-flight model: instrument characterization and reference data for atmospheric remote-sensing in the 230–2380 nm region, *Journal of Photochemistry and Photobiology A: Chemistry*, 157, 167–184, [https://doi.org/10.1016/S1010-6030\(03\)00062-5](https://doi.org/10.1016/S1010-6030(03)00062-5), 2003.
- Escribano, J., Bozzo, A., Dubuisson, P., Flemming, J., Hogan, R. J., C.-Labonnote, L., and Boucher, O.: A benchmark for testing the accuracy and computational cost of shortwave top-of-atmosphere reflectance calculations in clear-sky aerosol-laden atmospheres, *Geoscientific Model Development*, 12, 805–827, <https://doi.org/10.5194/gmd-12-805-2019>, 2019.
- Fedkin, N. M., Li, C., Krotkov, N. A., Hedelt, P., Loyola, D. G., Dickerson, R. R., and Spurr, R.: Volcanic SO<sub>2</sub> effective layer height retrieval for the Ozone Monitoring Instrument (OMI) using a machine-learning approach, *Atmospheric Measurement Techniques*, 14, 3673–3691, <https://doi.org/10.5194/amt-14-3673-2021>, 2021.
- Hedelt, P., Efremenko, D. S., Loyola, D. G., Spurr, R., and Clarisse, L.: Sulfur dioxide layer height retrieval from Sentinel-5 Precursor/TROPOMI using FP\_ILM, *Atmospheric Measurement Techniques*, 12, 5503–5517, <https://doi.org/10.5194/amt-12-5503-2019>, 2019.
- Hedelt, P., Theys, N., Romahn, F., Wagner, T., Lambert, J.-C., Loyola, D., Stein Zweers, D., Dehn, A., and Zehner, C.: S5P Mission Performance Centre Sulphur Dioxide [L2\_SO2\_\_\_\_\_], Tech. Rep. S5P-MPC-BIRA-PRF-SO2, Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Deutsches Zentrum für Luft und Raumfahrt (DLR), and European Space Agency (ESA), product version V02.06.01, issue 2.8, 2023.
- Lerot, C., Van Roozendaal, M., Spurr, R., Loyola, D., Coldewey-Egbers, M., Kochenova, S., van Gent, J., Koukouli, M., Balis, D., Lambert, J., Granville, J., and Zehner, C.: Homogenized total ozone data records from the European sensors GOME/ERS-2, SCIAMACHY/Envisat, and GOME-2/MetOp-A, *Journal of Geophysical Research: Atmospheres*, 119, 1639–1662, <https://doi.org/10.1002/2013JD020831>, 2014.
- Serdyuchenko, A., Gorshlev, V., Weber, M., Chehade, W., and Burrows, J. P.: High spectral resolution ozone absorption cross-sections – Part 2: Temperature dependence, *Atmospheric Measurement Techniques*, 7, 625–636, <https://doi.org/10.5194/amt-7-625-2014>, 2014.
- Spurr, R. and Christi, M.: The LIDORT and VLIDORT Linearized Scalar and Vector Discrete Ordinate Radiative Transfer Models: Updates in the Last 10 Years, in: *Springer Series in Light Scattering*, edited by Kokhanovsky, A., pp. 1–62, Springer International Publishing, ISBN 978-3-030-03444-3 978-3-030-03445-0, [https://doi.org/10.1007/978-3-030-03445-0\\_1](https://doi.org/10.1007/978-3-030-03445-0_1), series Title: Springer Series in Light Scattering, 2019.
- Theys, N., Lerot, C., Brenot, H., van Gent, J., De Smedt, I., Clarisse, L., Burton, M., Varnam, M., Hayer, C., Esse, B., and Van Roozendaal, M.: Improved retrieval of SO<sub>2</sub> plume height from TROPOMI using an iterative Covariance-Based Retrieval Algorithm, *Atmospheric Measurement Techniques*, 15, 4801–4817, <https://doi.org/10.5194/amt-15-4801-2022>, 2022.
- Vernier, J.-P., Aubry, T. J., Timmreck, C., Schmidt, A., Clarisse, L., Prata, F., Theys, N., Prata, A. T., Mann, G., Choi, H., Carn, S., Rigby, R., Loughlin, S. C., and Stevenson, J. A.: The 2019 Raikoke eruption as a testbed used by the Volcano Response group for rapid assessment of volcanic atmospheric impacts, *Atmospheric Chemistry and Physics*, 24, 5765–5782, <https://doi.org/10.5194/acp-24-5765-2024>, 2024.
- Yang, K., Liu, X., Bhartia, P. K., Krotkov, N. A., Carn, S. A., Hughes, E. J., Krueger, A. J., Spurr, R. J. D., and Trahan, S. G.: Direct retrieval of sulfur dioxide amount and altitude from spaceborne hyperspectral UV measurements: Theory and application, *Journal of Geophysical Research: Atmospheres*, 115, 2010JD013982, <https://doi.org/10.1029/2010JD013982>, 2010.