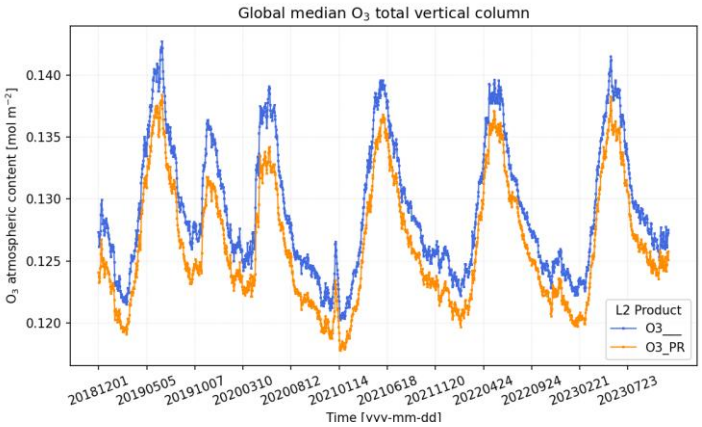
The authors very much appreciate the reviewers' insightful comments. These have been addressed point-by-point in the author replies and corresponding manuscript updates. Where reviewer comments strongly overlapped, a single author reply has been provided to both reviewers (indicated as such). Note that in between the initial manuscript submission and this updated version, a few additional ozonesonde stations and comparisons have become available to the authors. These have now been included, and tables and figures have been updated accordingly (minor changes, see redline version). This did however not affect the overall conclusions.

1. From Figure 12 (relative differences of total columns between S5P products), I think it is of importance to dig up the sources of differences in the western ocean out of South Africa. It could be either S5P profile measurements miss the biomass-burning signals due to low sensitivity or S5P total ozone retrievals are contaminated. Please give more descriptions.
   Author reply: The comparison provided in the manuscript does not involve the averaging kernel of the products, as the intention was to show qualitatively the consistency between the products. The consistency can also be seen in the timelines in the Figure below, illustrating the total vertical column of the GODFIT (blue line) and of the Ozone Profile product (orange line), in comparison since the beginning of the mission. There is an obvious bias between the two products over the whole period, which we intend to investigate. We are actually preparing another manuscript, which will be soon submitted, focusing specifically on the Ozone Profile algorithm itself.



   Regarding the area pointed out in the South Atlantic Ocean, we believe that this might be due to the different implementation of the climatology in the two products. Both products use the climatology Labow et al. (2015), however for the Ozone Profile a modified version of this climatology is used (as it is also described in the manuscript in section 2.2.3). In order to reduce the stratospheric influence in the troposphere, the original values of the climatology in the troposphere and upper atmosphere are replaced with the median values along the total ozone axis. This specific area is more driven by the troposphere, therefore a-priori differences might cause larger differences, which will be further investigated in the next manuscript on the retrieval algorithm itself. Manuscript update: Figure 12 was updated and a phrase in line 526 was added: "A slightly higher bias can be observed in the western ocean out of South Africa, which might be due to the difference in the climatology implementation between

the two products. This is under investigation and will be discussed in a follow-up manuscript."

2. Figure 5: the geophysical distribution of the mean sub-column layers looks very informative for both the lower troposphere and upper stratosphere, showing no apparent artefacts. However, the S5P product is expected to offer high spatiotemporal information compared to other processors (e.g., OMI, GOME). Therefore, I highly encourage you to provide an example of the daily (or a few days average) tropospheric ozone map zoomed in on a specific continent (e.g., Europe, Asia).
   Author reply: Only one day of data might contain quite some clouds, therefore we decided to provide a few days average (5 days) over Europe. We applied a cloud filter of 0.2 to look at cloud-free scenes. See manuscript update for description.
   Manuscript update: Figure 5 has been updated with the new figure and the caption has been accordingly updated. A description of the new figure has been added to the text: "In addition, Figure 5 (g) shows a five-day average in October 10-15 of the 0-6 km layer. The map data contains a cloud filter to only look at cloud-free scenes (cloud fraction below 0.2). The map clearly shows some regions with higher ozone levels in Eastern Europe and reduced columns over the Alps and the Pyrenees."

3. I would like to recommend moving Section 2.3 to Section 3 as a part of validation meteorology, rather than retrieval meteorology.
   Author reply: The authors agree with this suggestion.
   Manuscript update: Section 2.3 has been inserted into Section 3 as Section 3.1 (only title marked in red).

4. Line 224: What is the threshold value applied for the quality of fit (cost function) in data screening?
   Author reply: 200
   Manuscript update: The following sentence is added "It is recommended to apply a screening to the retrieval values showing a cost function fc larger than 200."

5. Line 229-253: the description of the operational validation system is better to be placed in Appendix to ensure the consistency and conciseness in text.
   Author reply: The authors tend to disagree with this suggestion. The manuscript appendix is limited to tables and figures, hence having only this text added there would seem not to improve readability. In general, the MPC systems/services are quite complicated to understand their connection and functionalities, so for external people who are interested it can be useful to have a brief description/introduction readily available in this work.
   Manuscript update: none

6. This manuscript should be revised overall to enhance readability. Even when the full name of the term is occasionally abbreviated in this paper, there is frequent inconsistency in the use of abbreviations and full names, even when the abbreviation is more familiar to the reader than the full name. DFS, TOLNet, WOUDC, NDACC, AK, FWHM, GAW, SHADOZ.

Author reply: The authors have verified the appearance of all acronyms throughout the manuscript text. The authors however have to stick to the Copernicus journal policy of spelling out acronyms upon first usage.
Manuscript update: Double appearances of full terms and acronyms have been removed at those instances where this was believed to increase readability.

7. (a) Section 3.1 should be placed just before Section 3.2.4 or merged into Section 3.2.4 to ensure the consistency and conciseness in text.
Author reply: The authors tend to disagree with this suggestion. As explained at the beginning of Section 3, the authors are following the validation scheme developed in Keppens et al. (2015). Information content studies are an essential part of this scheme, and therefore situated in the QA approaches, before being discussed independently in Section 4 as well.
Manuscript update: none
(b) And, it is much better to re-organize Section 4 into three parts, Retrieval characteristics including all AK, DFS, FWHM, sensitivity…, Geophysical distribution and validation results (comparison results between S5P and ground reference, between S5P profiles and S5P column).
Author reply: The authors are following the validation scheme developed in Keppens et al. (2015), where this distinction is less strict. It is especially insightful to look at information content measures for the co-located data, as this allows interpreting the observed differences in terms of important retrieval characteristics. In order to nevertheless address the reviewer's request, the authors have added global sensitivity maps to Figure 6 (numbering of initial submission).
Manuscript update: Sensitivity maps are added to Figure 6. Reference to these plots has been added in Section 4.2.
(c) Correspondingly, Figures 9-11 should be re-organized to enhance the readability
Author reply: Both reviewers have expressed major concerns on the readability of Figures 9 to 11, and related Figure A2 in the Appendix. The authors agree with the reviewers' suggestions and have therefore made substantial changes to these Figures and their captions.
Manuscript update: (1) Only the first three columns of Figures 9 and 10 are kept, while the remaining columns are moved to the Appendix. (2) These first three columns are provided in portrait instead of landscape orientation. (3) The SZA colour coding has been removed from Figure 11, although it is maintained in Figure A2 (left column) in the Appendix for the expert reader. (4) A legend is added to Figure 11. (5) The latitude-dependent drift results of Figure A2 (right column) have been moved to a new Figure in the main text.

8. Multi-Figures should have captions (Figure 1.a, 2.b, 3.b).
Author reply: The authors agree and will update the relevant figures.
Manuscript update: The relevant multi-figure plots have been updated.

9. Figure 9 and Figure 11. It is not necessary to perform the retrieval characteristics (DFS, sensitivity, FWHM, offset) at specific stations and specific reference dataset. I think a few orbit files are enough to specify the dependence of the retrieval characteristics on the geophysical parameters (SZA, VZA, cloud fraction…).

Author reply: This comment very much relates with point 7. It is especially insightful to look at information content measures for the co-located data, as this allows interpreting the observed differences in terms of important retrieval characteristics. In order to nevertheless address the reviewer's request, the authors have added meridian sensitivity maps to Figure 6 (numbering of initial submission).
Manuscript update: See point 7.

10. This paper strongly assures that the impact of sampling and smoothing errors on comparison results between S5P and reference are insignificant, thanks to the application of tight co-location criteria and AK smoothing of the reference observations. But, the substantial offset between nominal retrieval altitude and effective vertical retrieval altitude could introduce artificial features, right?
Author reply: The offset between the nominal retrieval altitude and the effective vertical retrieval altitude is also accounted for by averaging kernel smoothing, i.e., upon application of Eq. (3). This has been made clearer in the manuscript update.
Manuscript update: In Section 3.2.4 "Given that the effective vertical resolution of the satellite retrieval is significantly lower than the resolution of the retrieval grid (also see information content studies)..." has been replaced by "Given that the satellite retrievals show an effective vertical resolution and altitude registration that differs from the retrieval grid (also see Sections 3.1 and 4.3)..."

11. (a) Please add legends in Figure 11 for many lines (dashed, dotted, thin). Need to revise Figure 11.
Author reply: The authors agree with this suggestion.
Manuscript update: See update on Figures 9-11.
(b) It is hard to draw any insight on the data assessments as a function of SZA with the current way to put all individual profiles with different color-cording as a function of SZA. Maybe take a look at the mean difference/standard deviations for several SZA regimes (SZA < 40, SZA all, SZA > 60….)
Author reply: This suggestion is already taken into account for the sub-column plots in Figure 9 and 10, and by the overall update of Figures 9-11.
Manuscript update: See update on Figures 9-11.

12. In Section 4.7 and the conclusions (pages 585-595), the authors present a comparative analysis of data quality, comparing their TROPOMI operational ozone profiles with those from other research products (Zhao et al. 2020, Mettig et al. 2021; 2022, Malina et al., 2020), relying on literature assessment reports. Specifically, they assert that, "Apart from these exceptions (Malina et al., 2020 and Mettig et al. products), the operational product demonstrates comparable or lower uncertainty than the scientific products." However, it is not proper to draw definitive conclusions without conducting cross-validation using the same reference and validation criteria. With the difficulty in collecting other S5P ozone profile products for a long-term period, I recommend removing Section 4.7, except for the comparison results between S5P total column and S5P ozone profile-integrated column. Instead, provide a concise summary of other scientific products developed for delivering ozone profile information from TROPOMI measurements in the Introduction.

Author reply: Both reviewers have suggested moving at least part of section 4.7 (especially lines 505–516) to the introduction. The authors agree with this and have updated the manuscript accordingly. On the other hand, the authors prefer keeping the remainder of the section, including the discussion on the validation results of the scientific products. It is acknowledged that validation approaches cannot be fully matched, but this does not usually hamper comparative validation assessments in the literature. In order to take into account the reviewer's concern, the authors have therefore added a disclaimer to this discussion.

Manuscript update: Lines 505-516 have been moved to the introduction, subject to technical corrections. The following disclaimer has been added to Section 4.7 instead: "The operational TROPOMI ozone profile validation results obtained in this work are additionally compared with those of the scientific TROPOMI ozone profile retrieval algorithms that have been found in the literature (see Introduction). However, as the validation approaches for these products are not matched, this comparison should be considered with caution, and within their spatiotemporal validity."

13. Figure 2: what does '" the middle radiance bin" mean? Does it indicate the middle spectral pixel of 270-330 nm, at 290 nm? If so, please delete the middle radiance bin in part of describing Figure for the right panel of Figure 2.

    Author reply: In the soft-calibration routine, we compute the correction parameters as a function of several variables, among which also the radiance. The radiance is additionally binned in 20 bins, with each bin showing a particular atmospheric scene (first bins, for example, represent cloudy scenes). In the manuscript, we only give an example of the correction for the central radiance bin. The author agrees that this might cause some confusion as we only wanted to show the general trend of the radiometric correction, therefore the middle radiance bin from the description will be deleted. Detailed information about the correction can be found in the ATBD of the Ozone Profile.

    Manuscript update: the description of Figure 2 has been updated accordingly to the answer above, and adding multi figures names ("left" replaced with "(a)"; "right" replaced with "(b)"). From line 162, we also deleted the reference to "the middle radiance bin" and replaced "left" with "(a)", "right" with "(b)"

14. Line 545 "The main elements of the operational retrieval algorithm include the spectral pre-processing, which involves spectral/spatial regridding and wavelength/radiometric correction, a forward model, and an optimal estimation based inverse model."

    Author reply: The authors agree that the phrasing of this sentence can be improved, however the reviewer's suggestion is not fully correct as the spectral preprocessing is also one of the pre-processing steps. The authors suggest the following update.

    Manuscript update: The main elements of the operational retrieval algorithm include several pre-processing steps, the forward model, and the Optimal Estimation fitting based on the inverse model.

15. Line 553 The reference dataset used here includes WOUDC ozonesondes, TOLN tropospheric lidars, and NDACC stratospheric lidars.

Author reply: The authors agree that a rephrasing of this sentence may improve readability, but the reviewer's suggestion is not fully correct, as ozonesonde data does not only originate from WOUDC.
Manuscript update: "The latter are acquired by ozonesondes contributing to WMO's Global Atmosphere Watch, by tropospheric lidars from the Tropospheric Ozone Lidar Network, and by NDACC stratospheric lidars."

16. Line 508 delete "additionally applied to the MLS and to OMPS for intercomparison", And, connecting to the following sentence, like, "which applied for joint UV-IR retrievals from TROPOMI and CrIS.
Author reply: The authors agree with this suggestion.
Manuscript update: The text has been updated as follows: "Mettig et al. (2021) followed with the TOPAS (Tikhonov regularised Ozone Profile retrievAl with SCIATRAN) algorithm, which has also been used for joint UV-IR retrievals from TROPOMI and the Cross-track Infrared Sounder on the Suomi National Polar-orbiting Partnership (CrIS/Suomi-NPP)." Note that these lines (505-516) have been moved to the introduction.

17. Line 515 delete ",which is applied ~ data"
Author reply: The authors agree with this suggestion.
Manuscript update: This sentence part has been removed from the text. Note that these lines (505-516) have been moved to the introduction.

18. C17. Line 547 ranging from ➔ spanning
Author reply: We agree with this suggestion.
Manuscript update: "spanning"

19. Line 580. Please specify "Observed above" e.g. observed for the stratospheric ozone retrievals.
Author reply: We agree with this suggestion.
Manuscript update: "above" is replaced by "for the stratospheric ozone retrievals"

20. Line 584 Please provide any reference to "this agrees with the operational TROPOMI total ozone column retrieval". And, the long-term stability is commonly assured for other S5P L2 products?
Author reply: The authors agree with the suggestion to add a reference. On the other hand, it is hard to compare the TROPOMI (operational) ozone profile drift with the drift of other molecule retrievals. Only the O3_PR retrieval makes use of TROPOMI's UV spectrometer, which degrades differently (typically faster) than the other detectors within the instrument.
Manuscript update: After this statement, a reference to Garane et al. (2019) has been added to the text.