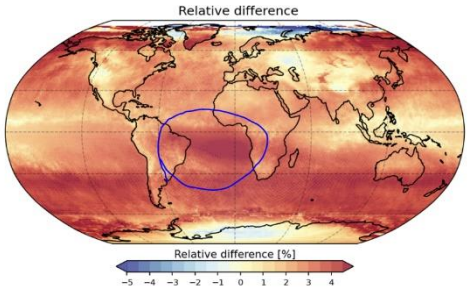
Overall Comments:

Keppens et al. provide a comprehensive description of TROPOMI operational ozone retrievals and retrieval characteristics (sensitivity, smoothing error, averaging kernels), along with validation results spanning five years of their product. In general, I recommend this manuscript for publication after the following comments being addressed.

Both the retrieval (Section 2) and validation procedures (Section 3) outlined in this paper are already part of the operational process. Therefore, most of my major concerns are focused on the validation results (Section 4).



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.
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 artifacts. However, S5P product is expected to offer high spatiotemporal information compared to other processors (e.g., OMI, GOME). Therefore, I highly encourage 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).
3. I would like to recommend moving Section 2.3 to Section 3 as a part of validation meteorology, rather than the retrieval meteorology.
4. Line 224 What is the threshold value applied for the quality of fit (cost function) in data screening?
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.
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.
7. 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. 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). Correspondingly, Figures 9-11 should be re-organized to enhance the readability.
8. Multi-Figures should have captions (Figure 1.a, 2.b, 3.b).
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…).

10. This paper strongly assured 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?
11. Please add legends in Figure 11 for many lines (dashed, dotted, thin). And need to revise Figure 11. 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….)
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 Metting 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 an 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.
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.
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.
15. Line 553 The reference dataset used here includes WOUDC ozonesondes, TOLN tropospheric lidars, and 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.
17. Line 515 delete “,which is applied ~ data”
18. C17. Line 547 ranging from ➔ spanning
19. Line 580. Please specify “Observed above” e.g. observed 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?