

Referee report to the revised version of the “Five years of Sentinel-5p TROPOMI operational ozone profiling and geophysical validation using ozonesonde and lidar ground-based networks” manuscript by Arno Keppens et al.

The manuscript has been significantly improved with respect to the presentation quality. Most of my comments were addressed in a satisfactory way. However, some issues are still needed to be dealt with. The manuscript can be accepted for the publication in AMT after a minor revision. My detailed comments are provided below.

Detailed comments

- Lines 13-14: “vertical sensitivity” - it is not a common notation and should be defined before using.
- Line 20: “meridian dependence of its bias” - Whose bias is meant here, that of the sensitivity or of the tropospheric ozone?
- Line 66: “The same combination of TROPOMI UV and CrIS IR retrieval wavelengths has been exploited by ...” - “A similar” instead of “the same” would be more correct, as the wavelength ranges used by these two retrievals are quite different.
- Line 160: “Additionally, the CAMS ozone profiles are scaled to match the total ozone column derived from the OMPS total column data (Jaross, 2017).” - I am wondering why you do not use the total ozone column from TROPOMI instead (just for a curiosity, not as a requirements to change).
- Line 267: Could you please comment on the value of 200 for the cost function threshold. How did you come to this value?
- Lines 341-342: “... resolution and altitude registration that differs from the retrieval grid ...” - what does “altitude registration” mean here? You probably want to highlight what the AK peaks are not at nominal altitudes but this formulation seems quite confusing to me.
- Line 382: “... the a-priori is smoothed by the measurements ...” - this statement sounds extremely confusing. I am sure you agree, measurements cannot affect a priori in any way. Please reword.
- Line 473: “... an increase of the DFS ...” - Fig. A3 does not show any DFS, I suppose you refer to Fig. A2 here. From the sentence it is not clear if you refer to 6-12 km column then talking about DFS increase with SZA. Looking at Fig. A2 I see a much larger increase of DFS with SZA for the 12-18 km column (the third row from the bottom) than for the 6 - 12 km one.

- Line 473: "... an increase of the ... bias for the 6-12 km column with SZA" - In Fig. 3 I do not see any increase of the bias with SZA for any of the columns.
- Lines 473 - 475: "This correlation seems to be somewhat compensated for in the lowest column by increased atmospheric penetration of the sunlight at low solar zenith angles (0 to about 30°)." - I cannot understand which correlation you are talking about here and where you see it compensated.
- Lines 475 - 477: "Additionally, the bias is clearly negatively correlated with the surface albedo for the 6-12 km subcolumn, despite the latter's apparently slightly positive correlation with the retrieval DFS." - a similar correlation for the differences is seen for the 0-6 km column and a bit reduced for the 12 - 18 km column. DFS for 0 - 6 km column does not seem to show any correlation with albedo while this correlation for the 12-18 km is largest. In general this sentence does not seem to overview the full picture.
- Lines 481 - 482: "... while a negative drift is observed for the two subcolumns above (18-32 km)." - I see a negative drift only for 18 - 24 km but not for 24-32 km (numbers in the pot), are you still discussing Fig 10? By the way, in the caption of Fig. 10 it is not explicitly indicated which column belongs to which row. I understand it is the same as for Fig 9 but this still should be mentioned explicitly.
- Lines 519 - 520: "This can be seen from Figure 10, with the black lines (average differences) being within the grey areas (SRD requirements)." - This is not really visible in the plot, especially in the right column.
- Line 527: "The vertical retrieval grid is sampled at a resolution of 6 km or higher, ..." and Table 2: "Partially, as the vertical grid complies, ..." - I do not think it is correct to rate a sufficient sampling of the vertical grid as a partial compliance with respect to the vertical resolution. I agree it is required to have a vertical grid with a sufficient sampling but it has nothing to do with the measurement/retrieval capabilities.
- Line 536: "... observed in the western ocean out of South Africa ..." - Do you mean "in the Atlantic ocean western of South Africa"?
- Figure 9: suboptimal position of the text boxes in the lower right plots, the boxes strongly cover the plot contents. 50% quantile lines are often difficult to distinguish, another color, e.g. green, might help.
- Figure 11: The figure is still difficult to read. It should be stretched to occupy the full page width. Horizontal space between the sub-plots would be useful
- Figure A4: same as for Fig. 11

Technical corrections:

- Line 19: “The vertical sensitivity of the TROPOMI tropospheric ozone amount” - This sounds a bit weird to me. Maybe you should exchange “of” by “to” or “for”, or talk about sensitivity of the retrieval and not that of ozone amount...
- Line 268: “...for all 33 levels l combined...” - should “ l ” be separated by commas?
- Line 302: “...of up to 5 %, and except in the tropical upper troposphere,...” - should there be a comma after “and”?