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.

Unfortunately, in Section 4, which describes the retrieval sensitivity and validation results, the authors decided to save space and joined tens of panels with thousands of curves into single plots. This makes the plots absolutely unreadable (Figs 9, 10, 11 and A2). Although the text is still well written, the reader cannot follow the discussion and verify the conclusions of the authors because no information can be read from the plots. In my opinion, the authors have to reconsider the way they present their results in Section 4 to make the manuscript suitable for publication. It should be analyzed which information is important and how to present it without making the plots look like squeezed together unresolved color spots. A rotation of the plots by 90 degrees is also not a really good idea. This makes the reader terribly difficult to look at the plots when reading the text.
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.

- The paragraph mentioning previous retrievals (lines 505–516, Section 4.7) should be moved to 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.
  Manuscript update: Lines 505-516 have been moved to the introduction, subject to technical corrections.
- Figure 1: the right box shows an endless loop between the radiative transfer and optimal estimation
  Author reply: It is not an endless loop in practice. Inside the box is written that it is until the convergence of eight iterations at maximum.
  Manuscript update: The following is added to the caption: "until convergence or the maximum number of eight iterations is reached."
- Line 125: for the pixel size after August 6, 2019 please indicate which size is crosstrack and which along-track.
  Author reply: (across-track x along-track)
  Manuscript update: This has been implemented in the text.

- Line 147: Usage of the CAMS data, with assimilated MLS profiles, and scaling them to OMPS total columns is expected to give a very good approximation for the ozone profile. In this respect the author should put a bit more focus to show that TROPOMI measurements increase the information content in comparison to a priori. Maybe it is already shown in plots in Section 4, they are, however, completely unreadable for me.
  Author reply: This is not the a-priori profile, which is described in section 2.2.3 and which comes from a modified version of the climatology references (Labow et al.). We use this ozone profile in the soft-calibration routine to compute the radiance expected from the best ozone profile shape we could get from that specific location. Then, we compare this estimate with a forward model calculation to compute the soft-calibration parameters.
  Manuscript update: This sentence was added: “Figure 6 clearly shows how the a-priori is smoothed by the measurements in the retrieved ozone. Moreover, the vertical sensitivity shown at the bottom indicates that the measurements add more information with respect to the a-priori in most of the vertical layers, with low-sensitivity areas depending on the latitude (as also shown in Figure 11).”
- Line 163: “computed by combining the black and light gray points of the same year” - There are no black points in the plot, I see two sets of grey points. It is unclear why there are two curves and how they are combined to get red points.
  Author reply: Yes, the image was not the one described and it has been updated with the correct one showing the points with different colours.
  Manuscript update: Now the points show the two different colours which refer to the two different datasets (OFFL and RPRO) combined to compute the total correction. This is also indicated in the plot legend.
- Section 2.2.2: it is unclear how the cloud fraction is included into the forward model.
  Author reply: The authors agree that a manuscript update is needed.
  Manuscript update: The following sentence is added to the manuscript: “In the forward model, clouds are represented as Lambertian reflecting surfaces which cover part of the ground pixel and are placed at cloud pressure. The cloud pressure and fraction are derived from the FRESCO algorithm using the oxygen A-band of TROPOMI at 760 nm. The cloud fraction and albedo are fitted at 330 nm during the retrieval of the ozone profile.”
- Line 179: “For all state vector elements, the OE method requires an a-priori value and its error estimates.” - I think ”a priori values and their error estimates” would be more correct
  Author reply: The authors agree with this suggestion.
  Manuscript update: The sentence has been updated in agreement.
- Line 199: “when the measurement errors are assumed to be uncorrelated” - it is unclear if it is the case in this study.
  Author reply: The error covariance matrix of the measurements is diagonal if there are no calibration errors. The diagonal then contains only the errors dominated by shot noise and are uncorrelated. In the current version of the retrieval algorithm, the non-diagonal part of the matrix is set to zero, therefore no calibration errors are considered.
  Manuscript update: “when” is replaced by “here as”; plus the phrase “dominated by shot noise and therefore” is added to specify our case.
- Eq. (2): Can this filter result in skipping real ozone profiles, which are strongly different from a priori, e.g. unexpected ozone loss?

Author reply: This filter has been applied to avoid unphysical deviation of the retrieval from the a-priori. It is not excluded that it can contain false positives, as any data screening would have. However, it can be seen in Figure 6 (in the initial submission) that the absolute difference between the ozone profile retrieval and a-priori is typically an order of magnitude smaller than this filter value, which would be equal to 10 on the current plot scale. It can also be seen from the third column in Appendix A1, that this filter is the one removing the least pixels.
Manuscript update: none

- Line 258: “The full width at half maximum (FWHM) of the AK corresponding to a given altitude is selected in this work as an indicator of the effective vertical resolution of the retrieved profile at this altitude” - Please add here a note, that this approach ignores the displacement of the AK maximum (which you treat then independently as offset).
  Author reply: We agree with this suggestion.
  Manuscript update: “The full width at half maximum (FWHM) of the AK corresponding to a given altitude is selected in this work as an indicator of the effective vertical resolution of the retrieved profile at this altitude, although it is determined independently of any vertical displacement of the kernel”
- Line 258: “the true, physical resolution” - if I understand the meaning of the sentence correctly, the comma needs to be deleted.
  Author reply: We agree with this suggestion.
  Manuscript update: Comma is deleted.
- Line 337: “retrieval differences and vertical sampling and smoothing differences” - the wording seems to be incorrect
  Author reply: We agree that this phrasing can be improved.
  Manuscript update: “retrieval differences including vertical sampling and smoothing differences”
- Figure 6: A relative difference should be provided in addition
  Author reply: The authors agree with this suggestion, but the absolute difference has been removed from the plot to avoid having a large figure. Only the relative difference is provided now.
  Manuscript update: The absolute difference has been replaced by the relative difference.
- Line 359: “cost function > 200” - This criterion was not properly described and justified in Section 2.2.
  Author reply: The manuscript has been updated in the section regarding data selection.
  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.”
- Figure 9 is unreadable. The panels are too small. There are too many panels in the plot. A rotation by 90 degrees makes the plot even less readable.
  Author reply: The authors agree with this suggestion.
  Manuscript update: See update on Figures 9-11.
- Line 388: “Yearly drifts are added to the temporal dependence plots in Figure 9.” - the meaning of this statement is unclear. Are you correcting some data for the drift?
  Author reply: The authors agree that this statement is dubious. The text has therefore been rephrased and extended.

Manuscript update: “Yearly drift values that are calculated from a linear fit are added to the temporal dependence plots in Figure 9. The two-sigma uncertainties on these values result from a bootstrapping technique with thousand samples (Efron and Tibshirani, 1986).”

- Figure 10 is not referenced, it is unclear what is the difference between Fig 9 and 10. Figure 10 is unreadable similar to Figure 9.
  Author reply: Because of their similarity, Figures 9 and 10 are kept together, although Figure 11 appears in the text before Figure 10.
  Manuscript update: See update on Figures 9-11. The discussion of these figures in the text has been updated accordingly.
- Figure 11: The scale of the figure does not enable the reader to estimate the values. The figure cannot definitely be read by people with color vision deficiencies. There are too many panels in one plot. The white dashed line is almost invisible. It cannot be read from the plot if the retrieval compares better than a priori.
  Author reply: The authors agree with this suggestion.
  Manuscript update: See update on Figures 9-11. It has additionally been stressed how the retrieval performs in comparison to the mean prior in the first paragraph of Section 4.4.
- Section 4.4: please state clearly whether the comparisons are done convolving the reference data with the averaging kernels of TROPOMI retrieval.
  Author reply: The authors agree with this suggestion.
  Manuscript update: The first sentence of Section 4.4 has been updated as follows (initial numbering): “Figure 11 contains all comparisons between TROPOMI ozone profiles and reference data, the latter AK-smoothed using Eq. (3), and corresponding statistics.”
- Line 431: “... has a mean bias below ±5-10% in the troposphere...” - It is absolutely impossible to see if the mean bias decreases in comparison to the a priori
  Author reply: The authors agree with this suggestion.
  Manuscript update: See update on Figures 9-11. It has additionally been stressed how the retrieval performs in comparison to the mean prior in the first paragraph of Section 4.4.
- Lines 447-450: It looks like this text refers to Figure 10. It is unclear why it is placed after the discussion of Fig. 11.
  Author reply: Because of their similarity, Figures 9 and 10 are kept together, although Figure 11 appears in the text before Figure 10.
  Manuscript update: See update on Figures 9-11. The discussion of these figures in the text has been updated accordingly.
- Figure A2 needs to be moved to the main text as it is discussed here
  Author reply: The authors agree with this suggestion.
  Manuscript update: See update on Figures 9-11.
- Line 466: “... on average the observed differences confirm...” - with exception of the stratospheric lidars this is valid only above 20 km. Thus, this statement is not suitable in general.
  Author reply: The authors agree that the “on average” at the beginning of the sentence may be confusing. It is now stressed that this only applies to the stratosphere.
  Manuscript update: Line 466 (initial numbering) has been updated as follows: “The chi-square plots in Figure 11 (third graph in each plot) demonstrate that

the observed differences confirm (chi^2 close to one) the combined ex-ante satellite and ground uncertainty estimates in the stratosphere on average, despite the appearance of large outliers."

- Line 466: Again, Figure A2 is discussed in the main text but placed in the Appendix
  Author reply: The authors agree with this suggestion.
  Manuscript update: See update on Figures 9-11.
- Line 481: "All requirements are summarized in Table 2, with the compliance of the operational TROPOMI ozone profile product added" - It is unclear how the values for the table are obtained. From the ozonesonde comparison, for example, it is difficult to understand why authors claim that the accuracy between 12 and 18 km is below 5 %.
  Author reply: The authors agree that this is hard to see from Figure 10 (in its initial formatting), which contains grey areas indicating the product requirements on the uncertainty. These should be more clearly visible by the update of Figures 9-11.
  Manuscript update: See update on Figures 9-11.
- Line 486: "This can be seen from Figure 10, with the black lines (average differences) being within the grey areas (SRD requirements)." - I cannot see anything from Figure 10 but this is most probably because of the quality of the figure.
  Author reply: The authors agree that this is hard to see from Figure 10 (in its initial formatting), which contains grey areas indicating the product requirements on the uncertainty. These should be more clearly visible by the update of Figures 9-11.
  Manuscript update: See update on Figures 9-11.
- Line 501: "... typically amounts to about 5%, meaning ..." - From the color scale used in the plot it is difficult to read whether the values in maxima exceed 5%.
  Author reply: The authors agree with this comment.
  Manuscript update: Figure 12 has been updated with extended colour scale.