Referee #2

I understood that this manuscript studies the possibility of CO2 retrieval in high AOD pixels by the AOD threshold change.

Because of the disadvantage of spatial coverage for CO2 satellites, this study will contribute to enhancing the global coverage of observation data for CO2 satellites. Although the purpose of the manuscript is acceptable, the detailed analysis and results are not so clear. For details..

Reply: We thank the referee for the valuable comments. We have revised the manuscript accordingly, to make the analysis and results clearer.

1) Section 2: The study used the Dark Target (DT) algorithm to concentrate the urban surface. However, the DT algorithm have large uncertainty of AOD over land surface as compared to the ocean surface. For the detailed analysis of retrieval uncertainty related to the AOD, retrieval results over land surface are carefully handled. Do you have the same results when AOD from MODIS uses the Deep-blue algorithm?

Reply: As discussed in the manuscript, we have repeated the analysis using MODIS Deep Blue (DB) algorithm results, and the results are shown in the Appendix. While the global patterns of AOD (difference) are somewhat different than with MODIS Dark Target (DT), we find that the global statistics and conclusion regarding the connection to XCO2 retrievals drawn from the data are largely the same. We have added a note on this in Section 2.2. MODIS DB results are shown in Fig. A2 and Fig. A6 in the Appendix.

2) L72-L73: For the OCO-2 AOD retrieval, two representative types selection is confused. Does this sentence mean that spatio-temporal variated climatological types are selected for the AOD retrieval? In addition, Is the AOD from OCO-2 hard to consider the 'case dependent' aerosol types?

Reply: This sentence was poorly formulated, as noted also by Referee #1. We have clarified the sentence, and it now reads: “The aerosol parameters of the ACOS algorithm include five scatterers, which are two cloud types (water and ice), two tropospheric aerosol types and a stratospheric aerosol type (sulfate).”

3) L99: For the AERONET AOD reference, Eck et al. (1999) is too old to explain the Version 3. Giles et al. (2019) or Sinyuk et al. (2020) are more suitable.

Giles, David M., et al. "Advancements in the Aerosol Robotic Network (AERONET) Version 3 database–automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements." Atmospheric Measurement Techniques 12.1 (2019): 169-209.

Sinyuk, Alexander, et al. "The AERONET Version 3 aerosol retrieval algorithm, associated uncertainties and comparisons to Version 2." Atmospheric Measurement Techniques 13.6 (2020): 3375-3411.

Reply: We thank the referee for pointing out this omission. We have checked for more recent references and added them to the manuscript.

4) Section 3.1: For the collocation, I can't find the cloud screening method before the grinding. The AOD retrieval products are very important to the cloud screening before the analysis, although respective AOD retrieval algorithm have there own cloud masking method.

Reply: Both OCO-2 and MODIS data are cloud screened, using their own respective methods, before the collocation. The MODIS Collection 6 cloud mask used in aerosol retrieval is described in Levy et al. (2013) and references therein and has been continuously developed for aerosol retrieval purposes. The MODIS quality flag is further used to reduce possible cloud contamination e.g. by thin cirrus. OCO-2 data are cloud-screened before the Level-2 retrieval. The cloud screening is described and validation shown in Taylor et al. (2016).

We hope that the combined use of cloud masks from both instruments helps mitigate cloud contamination issues. However, we are aware that cloud screening is a trade-off between coverage and possible residual clouds contaminating the retrieval results. High AOD outliers seen in the data may be caused by insufficient cloud screening.

The effect of cloud screening on the collocated dataset is indicated in the manuscript in section 3.1 (and Table A1 and Fig. A1), but not explicitly discussed. We have added a brief note on the effect of cloud screening (and possible residual clouds) in section 3.1.

References:

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu, N. C.: The Collection 6 MODIS aerosol products over land and ocean, Atmospheric Measurement Techniques, 6, 2989–3034, https://doi.org/10.5194/amt-6-2989-2013, 2013.

Taylor, T. E., O'Dell, C. W., Frankenberg, C., Partain, P. T., Cronk, H. Q., Savtchenko, A., Nelson, R. R., Rosenthal, E. J., Chang, A. Y., Fisher, B., Osterman, G. B., Pollock, R. H., Crisp, D., Eldering, A., and Gunson, M. R.: Orbiting Carbon Observatory-2 (OCO-2) cloud screening algorithms: validation against collocated MODIS and CALIOP data, Atmos. Meas. Tech., 9, 973–989, https://doi.org/10.5194/amt-9-973-2016, 2016.

5) Section 3.3 and more: For the collocation, the author has to check the consistency of data variability within the collocation range (spatially and temporally). Could you provide the reference for spatial and temporal collocation ranges?

Reply: We are aware of the collocation mismatch uncertainty related to comparing point-like AERONET observations to snapshot type satellite aerosol data over larger areas. The sampling distance used for satellite data and temporal averaging window size for AERONET data affect the comparison metrics (see e.g. Virtanen et al. 2018).

In this work, we content ourselves with a limited comparison with AERONET, using a simple sampling strategy as described in Section 3.3. This is mainly done to confirm the results we see in the comparison with MODIS: that the OCO-2 AOD has a slope around 0.3 with respect to MODIS AOD, and a slope around 0.4 when accounting for the wavelength difference. Based on Virtanen et al., 2018 (e.g. Figure S8 in the Supplement), a sampling distance of 0.1 degree and temporal window of one hour seem reasonable. We have also tested AERONET comparison with a subset of data using different collocation parameters. We note that the MODIS AOD product, used as the reference for OCO-2 here, has been extensively validated elsewhere (e.g. Levy et al. (2013)), and we do not intend to repeat the effort here with the limited collocated dataset.

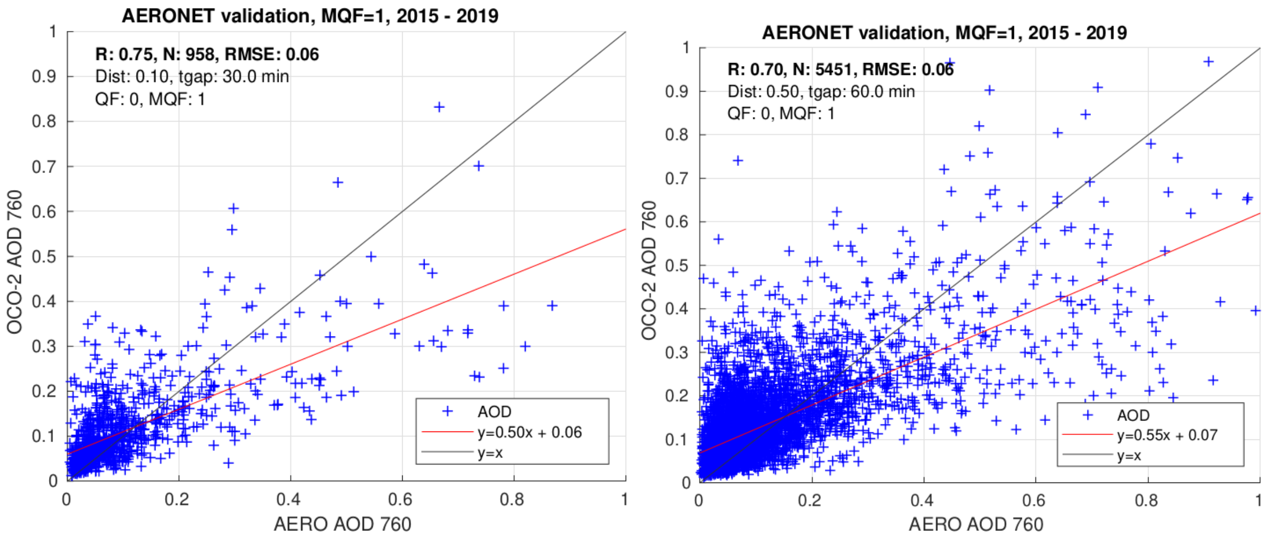

Figure 1. Effect of different collocation criteria on the AERONET comparison. The results differ only slightly when using a) distance of 0.1 degree for averaging OCO-2 data around the AERONET site and a temporal averaging window of ±30 min centered at the OCO-2 overpass time for AERONET data; or b) distance of 0.5 degree and a temporal averaging window of ±60 min.

We have added discussion on the spatial and temporal collocation ranges to Section 3.3.

Reference:

Virtanen, T. H., Kolmonen, P., Sogacheva, L., Rodríguez, E., Saponaro, G., and de Leeuw, G.: Collocation mismatch uncertainties in satellite aerosol retrieval validation, Atmos. Meas. Tech., 11, 925–938, https://doi.org/10.5194/amt-11-925-2018, 2018.

6) Section 4: Did only use the Quality flag value from OCO-2? Why don't you use the quality flag for another satellite platform (MODIS AOD)?

Reply: We did use the quality flag from MODIS as well. We systematically removed the lowest quality MODIS pixels (MODIS quality flag 0), as discussed in Section 2.2. The OCO-2 quality flag is discussed more, since the focus is more on the AOD threshold used in that flag. We have added clarification of the quality flags used in the text in Section 2.2 and Section 4.

In most of the results' subsections, the paper did not have sub summary or sub-conclusion. For this reason, it may be difficult to a connection among the results.

Reply: We agree that the purpose of different subsections and the conclusions drawn from each section were not described clearly enough. We have added more discussion to the text, including conclusions to section 4.1, 4.2, and 4.3, and introductory parts to sections 4.2 and 4.4.

7) L184-L185: How to eliminate the cloud contamination? Cloud contamination affects the high AOD, and it affects the high correlation between OCO-2 and MODIS, when both algorithms have cloud contamination.

Reply: As discussed above at point 4), we make an effort to screen for clouds. Cloud contamination is one of the likely reasons for high AOD outliers for each instrument. It is possible that simultaneous cloud contamination of both instruments affects the correlation data. However, the vast majority of data is at the low AOD range, as seen in Fig. 2. Studying the effect of residual clouds in more detail would require some reference data of the actual cloud cover and is beyond the scope of this work. We have added a note of the effect of possible simultaneous cloud contamination to the observed correlations to Section 4.1.

8) L194: I don't agree with the spectral conversion based on the MERRA-2. To use this method, the author has to analyze the intercomparison between MERRA-2 angstrom exponent and AERONET angstrom exponent.

Reply: We realize that the use of MERRA-2 Ångström exponent for the spectral conversion involves high uncertainty. This is why we chose to compare the AOD results at their original wavelengths, 550 nm for MODIS and 755 for OCO-2, for most of the paper, and address the wavelength difference respectively. As we discuss in the manuscript, we expect that part of the difference seen between the instruments is due to the wavelength difference. MERRA-2 data was chosen, because it is readily available for the full global collocated MODIS/OCO-2 dataset.

We use the spectral conversion with MERRA-2 data merely to get a rough estimate on the effect of the wavelength difference on the AOD difference. This is done only in a statistical sense for the global dataset, understanding that the high uncertainties involved with the scaling do not allow for a more detailed comparison. The main conclusion drawn from this is that while the slope of OCO-2 AOD against MODIS AOD is ~0.3 before spectral scaling, it is ~0.5 after the scaling, i.e. the wavelength difference explains part, but not all, of the difference.

We agree that the spectral conversion using AERONET data has much lower uncertainty. To support the MERRA-2 exercise, we compared OCO-2 AOD at 755 nm with the AERONET AOD data at different wavelengths in Fig. A5 in the manuscript. The AERONET results confirm that the OCO-2 AOD (at 755 nm) slope against AERONET AOD is ~0.3 when using 550 nm and ~0.5 when using average of 675 nm and 870 nm for AERONET.

To further support this analysis, we have now repeated the analysis using a subset of collocated MODIS/OCO-2/AERONET data. Here we have used the AERONET Ångström exponent to do the spectral conversions. Here we also use a slightly different approach than in Fig. A5, in that we average the OCO-2 data within 0.125 deg around the AERONET site and average the AERONET data within 1 hour of the overpass time. From Fig. 2 below we see that comparison of OCO-2 AOD to AERONET AOD scaled to 550 nm and to 760 nm, respectively, gives similar results (in terms of the slope) as in Fig. A5 of the manuscript. The small difference in the numbers may be explained by the smaller subset of data used here.

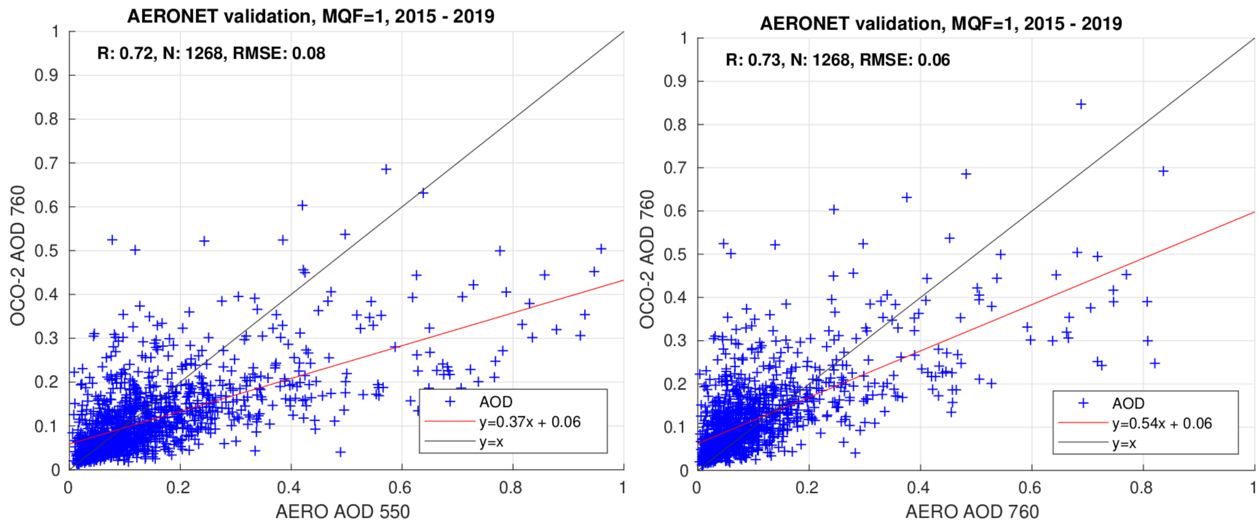

Figure 2. Comparison of OCO-2 AOD at 760 nm against AERONET AOD at 550 nm (left) and against AERONET AOD at 760 nm (right). The AERONET AOD values are scaled to different wavelengths using the AERONET Ångström exponent.

Figure 3 below shows a comparison similar to Fig. A4 in the manuscript, where OCO-2 AOD is compared to MODIS AOD at 550 nm, first using the original data at wavelength of 760 nm and then using OCO-2 AOD scaled to 550 nm, this time using the Ångström exponent from AERONET. Naturally, here this is done only for the subset of data collocated with AERONET. The results agree strikingly well with those shown in Fig. A4 in the manuscript. The slope obtained here is 0.31 and 0.33 in Fig. A4 when using the original wavelengths, and 0.47 in here and in Fig. A4 when using the OCO-2 AOD scaled to 550 nm.

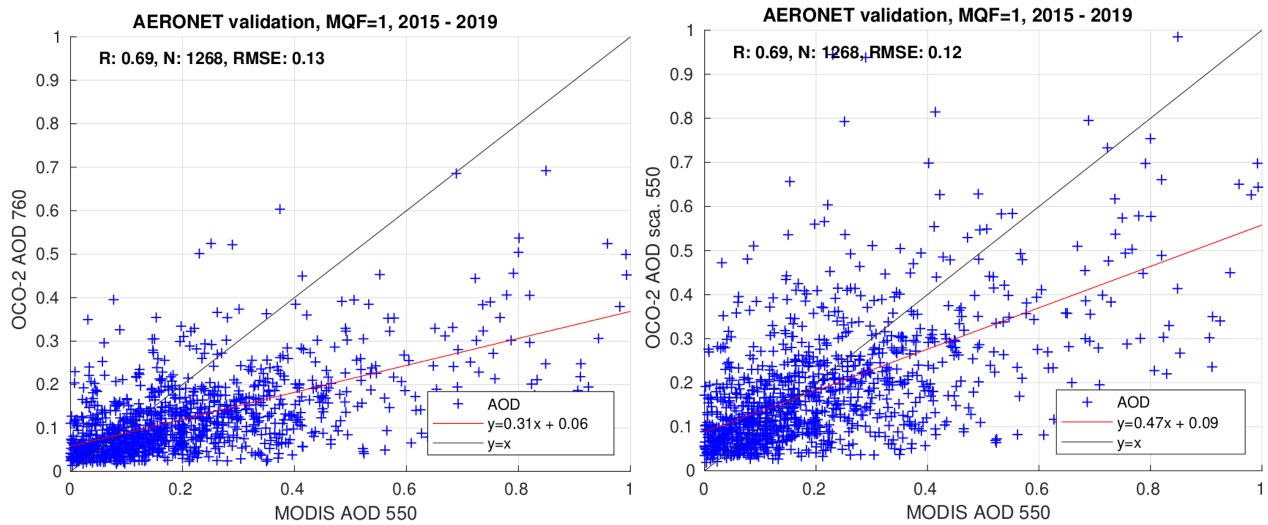

Figure 3. Comparison of OCO-2 AOD to MODIS AOD using the original wavelengths (left), and using AERONET Ångström exponent to scale OCO-2 AOD to 550 nm (right).

Based on this further analysis, we suggest that the use of MERRA-2 Ångström data for the spectral conversion is sufficient for the limited purpose of assessing the effect of wavelength difference to the AOD comparison. We have added a note summarizing this further analysis to Section 4.1.

9) Section 4.2: I clearly don't know the purpose of this section. Only quality checking? or making the threshold of AOD to define the high AOD?

Reply: This section has two purposes: 1) it continues the AOD comparison between OCO-2 and MODIS in a statistical sense (while section 4.1 concentrated on the spatial differences); 2) it assesses the AOD threshold of 0.2 currently used in OCO-2 retrievals to designate the good quality retrievals. While the first point can be considered as ‘quality checking’, the second point aims at discovering how well the current OCO-2 quality filtering works from the point of view of aerosols (there are other contributing factors, not addressed here). The division to four ‘AOD quarters’ in the density scatter plots shows the cases where the AOD-based quality filter works well (Q1, where both instruments agree that AOD is low, and Q3, where instruments agree that AOD is high), and where it could be improved (Q2, where high AOD cases may be included to good quality data, and Q4 where low AOD cases may be unnecessarily removed).

We have added a brief introduction to section 4.2 to make the purpose of the section more explicit.

In addition, for the AOD quality checking, gridded dataset is not adequate. If you use the gridded AOD data, the author has to make a finer resolution.

Reply: Figure 2 shows clearly the distribution of AOD differences between the two instruments. The main findings here are (1) the distribution of data in the four AOD quarters, and (2) the bias of OCO-2 AOD as function of MODIS AOD (with similar behavior observed in comparison against AERONET in the Appendix). This is shown both as a linear fit, and as binned means with respect to MODIS AOD bins. These results do not change when higher resolution is used, as shown in Fig. 4 below, which replicates Fig. 2 a) in the manuscript with higher resolution (AOD bin 0.005).

The large spread of the data reflects the fact that the ACOS algorithm is not an aerosol retrieval algorithm. This was added to the text.

Most of the statistics, such as the correlation coefficient, averages, linear fits, and fraction of data in different quarters are calculated from the original data points, not from the aggregated grid cells. Only the average OCO-2 AOD values for each MODIS AOD bin (dashed red line) is calculated using binned data.

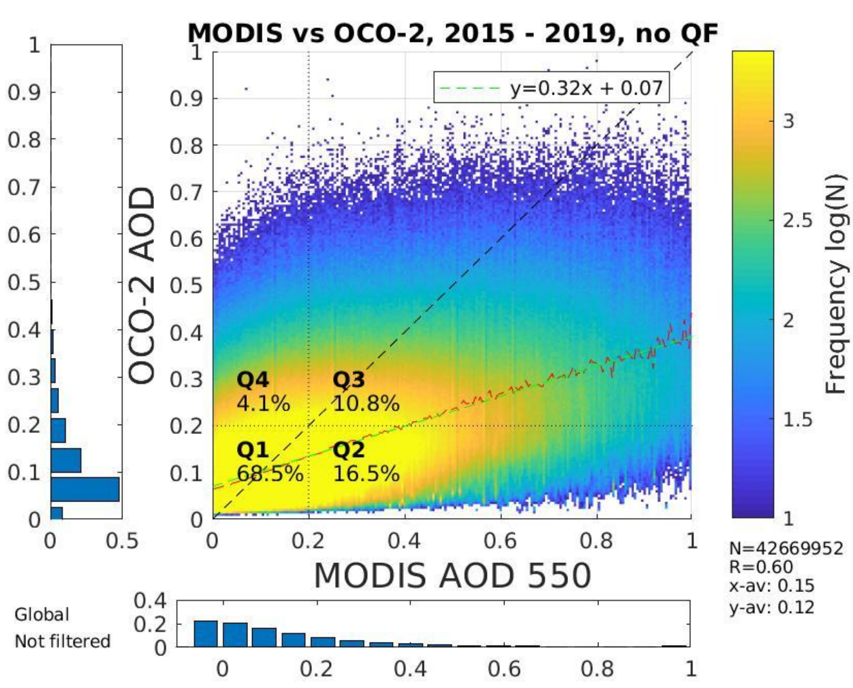

Figure 4. Density scatter plot of collocated AOD from OCO-2 and MODIS; same as Figure 2 a) in the manuscript, but using higher resolution (AOD bin 0.005, whereas Fig. 2 a has bin size 0.02). Grid cells with less than 10 points are not shown.

10) L236: Do you have references or analyzed results? I agree with the cloud contamination. However, the effect of ice aerosol component is not clear. Please include some back-up result.

Reply: As discussed in Section 2.1, the OCO-2 total AOD component includes contributions from water and ice particles. Preliminary comparison indicates that these scatterer types are more dominant in the low MODIS AOD/high OCO-2 AOD part of the AOD matrix, as indicated by Fig. 5 below. However, confirmation of this would require a more detailed look in the OCO-2 retrieval algorithm, which is beyond the scope of this work.

We have added a note to the text.

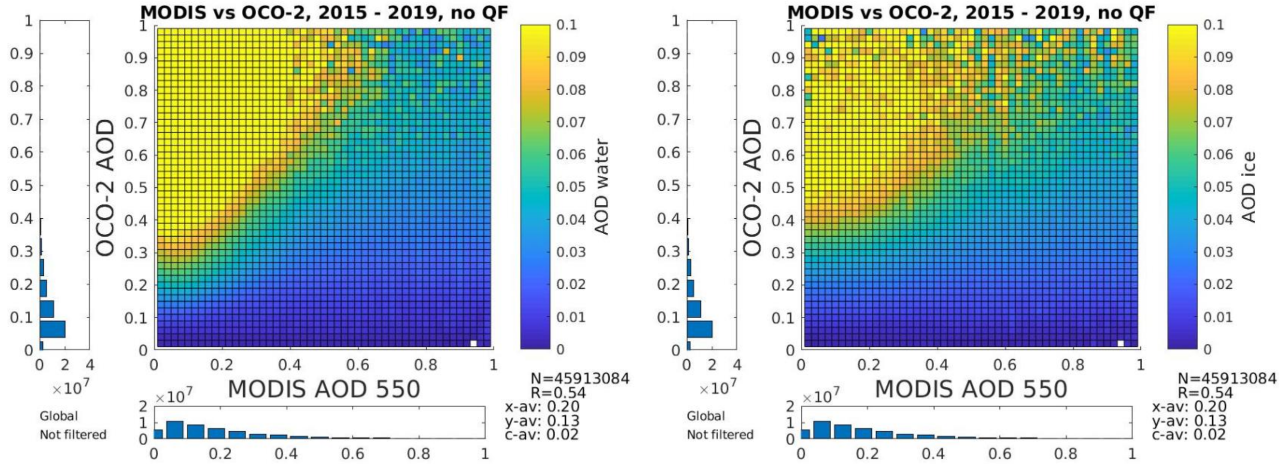

Figure 5. Contributions of water and ice components to the OCO-2 total AOD.

11) L240: Based on the statistical results and figures, this paragraph is not clear. The statistical score is possible to change due to the large number of data under low AOD grids. Statistical score change is not efficient in explaining the quality change of datasets.

Reply: We have reformulated the paragraph. The smaller slope of the linear fit and the lower correlation coefficient are a natural consequence of removing all data points with OCO-2 AOD over 0.2, while a large fraction of the high MODIS AOD pixels remain in the dataset after applying the OCO-2 quality filter. The slope in the quality filtered dataset is rather meaningless, and we merely wanted to point out that for the AOD comparison we need to use the full dataset. However, we think that showing both panels in Fig. 2 helps the reader to understand the effect of OCO-2 quality filtering, which is used in the XCO2 retrieval.

12) L293: Do you have reference?

Reply: Lines 292-294 in the manuscript: "Figure 4 b) shows the correlation between MODIS AOD and OCO-2 XCO 2 for 1◦ × 1◦ grid cells. We see particularly high correlation values for the Sahel region, parts of South-East Asia, and Western USA."

We are not sure what the Referee means here. On line 293 we simply describe the results shown in Fig 4 b), without any reference to previous literature.

13) Figure 5: Showing the number of data in each bin as adding figure.

Reply: The number of data for each AOD bin is shown in Fig. 2 in the manuscript (on logarithmic scale). We have added a note on this to the text describing Fig. 5.

14) Section 4.4: So, from this section, does the author think that the AOD affects the XCO2 retrieval? How to be quantitatively separate the effects between the AOD effect and real XCO2 enhancement?

Reply: As discussed in Section 4.3, it is difficult to disentangle the effects of real correlation between aerosols and CO2, and an aerosol bias in the XCO2 retrieval. Our conclusion is that there is a small bias in satellite XCO2 caused by aerosols, such that in heavy aerosol conditions the XCO2 is biased low. This acts to partly mask the true correlation between AOD and XCO2, but the bias has considerably smaller magnitude than the observed co-emission. We have reformulated the text to make this clearer.

15) Section 5: I am confused about whether the AOD threshold change is acceptable.

For focusing on the comparison between XCO2 and MODIS AOD, the moderate AOD condition will make it possible to estimate the accurate XCO2 value. However, it is just the data based on the AOD from MODIS.

The AOD difference between OCO2 and MODIS is partially due to the AOD retrieval limitation by the OCO2.

In this case, high AOD conditions from OCO2 have high uncertainty. From this study, is this case can be clarified?

Reply: Clearly the discussion in section 4.5 was not sufficiently clear, as pointed out by all referees. The purpose here was to demonstrate that the higher AOD threshold of 0.5, planned to be used in the coming CO2M mission (which will be better equipped to deal with higher aerosol concentrations), will bring considerable enhancement to the coverage. This is important in the light of the discovered co-emission of aerosols and CO2, since otherwise omitting high AOD areas might lead to biases in the XCO2 (source) data. Note that in Section 4.5 the MODIS AOD threshold is applied to data which has already been filtered using the OCO-2 quality filtering. We are unable to, and do not attempt to, judge if a higher (OCO-2) AOD threshold could be used with OCO-2 instrument.

We have extended the discussion in Section 4.5 to address the points raised by the referees and to clarify our message. Please see the more detailed answer in our reply to Referee #1.