

In this response file we have collected the point-by-point responses to the comments of all three referees, and describe the changes made to the text due to each point in detail. In addition to the changes due to responses, minor edits (language, typos) have been made to the manuscript.

Referee #1

The Virtanen et al manuscript entitled “A global perspective on CO₂ satellite observations in high AOD conditions” tries to evaluate the OCO-2 product from the point of view of aerosols by the combined analysis of the OCO-2 retrieved XCO₂ and AOD, collocated MODIS AOD, and TCCON retrieved XCO₂. This manuscript found a systematic difference between AOD retrieved by OCO-2 and MODIS, which can impact the retrieval of XCO₂. The dependence of the OCO-2 retrieved XCO₂ on the AOD difference was also found. But toward the future CO2M mission, in my opinion, besides simple data coverage variation with different AOD threshold, you should also discuss the quality control with the AOD threshold of 0.5.

Reply: We thank the reviewer for the insightful comments which helped us to improve the manuscript. We have revised the manuscript considering the comments and the questions presented.

Major comments:

You only analyzed the data coverage variation after adjusting the AOD threshold from 0.2 to 0.5. According to your analysis of the relationship between OCO-2 retrieved XCO₂ and AOD, the accuracy of the former significantly depend on the uncertainty of the later. You should also discuss how to conduct quality filter, or what the data quality will be for CO2M, which will use 0.5 as AOD threshold.

Reply: Clearly the discussion in section 4.5 was not sufficiently clear, as pointed out by all Referees. We have revised this section to better describe the objective of this part of the work and to clarify the issues raised. Please find a more detailed response in the specific comments below.

The specific comments are listed below:

P2L24: Change the order of the sentences “An essential monitoring component will be the Copernicus Anthropogenic CO₂ Monitoring Mission (Meijer et al., 2023).” And “While ground-based greenhouse gas measurements are mainly 25 available in developed countries – with limited coverage and representativeness – satellite-based XCO₂ information will be irreplaceable in areas where ground-based measurements are not made.”

Reply: We agree that this makes text more fluent and have changed the order.

(The line numbers in our replies refer to the latex difference pdf file.)

L24: We have changed

"An essential monitoring component will be the Copernicus Anthropogenic CO₂ Monitoring Mission (Meijer et al., 2023). While ground-based greenhouse gas measurements are mainly available in developed countries – with limited coverage and representativeness – satellite-based XCO₂ information will be irreplaceable in areas where ground-based measurements are not made."
to

"While ground-based greenhouse gas measurements are mainly available in developed countries – with limited coverage and representativeness – satellite-based XCO₂ information will be irreplaceable in areas where ground-based measurements are not made. An essential monitoring component will be the Copernicus Anthropogenic CO₂ Monitoring Mission (CO₂M; Meijer et al. (2023))."

P2L33: Check the format of the citations “(e.g., (Houweling et al., 2015; Crowell et al., 2019))”, should they be “first name et al. (year)”?

Reply: We agree.

L34: We have changed “(e.g. (Houweling et al., 2015; Crowell et al., 2019))” to “(e.g., Houweling et al. (2015); Crowell et al. (2019))”

P3L68: “In the collocated database we include only a limited selection of OCO-2 data fields, but the sounding ID in the combined daily files is equivalent with the original OCO-2 lite files, allowing addition of more data fields in an effective manner.” It makes me confused. Please further explain it. What is “sounding ID”? Why can you get “more data fields”?

Reply: Initially, only a few selected variables from the OCO-2 data were included in the collocated database (e.g., XCO₂, XCO₂_quality_flag), to save space. Later we wanted to test possible effects of additional variables, so the database was constructed in such a way that this could be effectively done. Sounding ID is a unique identifier for each observation.

We now realize that this is a technical detail on a level that is probably not relevant for readers and might just cause confusion. We have removed this part from the text.

L69: We have removed the sentence "In the collocated database we include only a limited selection of OCO-2 data fields, but the sounding ID in the combined daily files is equivalent with the original OCO-2 lite files, allowing addition of more data fields in an effective manner."

P3L70: “The aerosol parameters of the ACOS algorithm include five scatterers, two cloud types (water and ice), two tropospheric aerosol types and a stratospheric aerosol type (sulfate).” This may should be revised as “The aerosol products of the ACOS algorithm include parameters from five scatterers”. Is “five scatterers” refer to “two cloud types (water and ice), two tropospheric aerosol types and a stratospheric aerosol type (sulfate)” If so, this may should be revised as “five scatterers, which are two cloud types (water and ice), two tropospheric aerosol types and a stratospheric aerosol type (sulfate).”

Reply: This sentence was poorly formulated, as noted also by Referee #2. 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).”

P3L72: There are five aerosol types mentioned in the brackets. Why did you write “Two representative types of tropospheric aerosols”?

Reply: We have reformulated this.

L73: We have changed

“Two representative types of tropospheric aerosols (dust, smoke, sulfate aerosol, organic carbon, and black carbon) are drawn from a climatology based on location and time (Crisp et al., 2021).”

to

"The two most representative types of tropospheric aerosols out of five possible types (dust, sea salt, sulfate aerosol, organic carbon, and black carbon) are drawn from collocated 3-hourly aerosol fields from Goddard Earth Observing System Model, Version 5, Forward Processing for Instrument Teams (GEOS-5 FP-IT; see Crisp et al. (2021))."

P3L74: “Atmospheric Carbon Observations from Space (ACOS)”. Only use abbreviation is ok.

Reply: We now see that ACOS was defined already on line 43 and have removed the second explanation of the abbreviation.

P3L73: “From the large number of data products...”.

Reply: We have replaced “quantities” with “data products”.

P4L86: Add last accessed date after the website.

Reply: Done.

P4L106: What are “data fields which are not relevant for this study”?

Reply: This is also a technical detail related to the reduction of the size of the collocated database. As this was explained already on line 88 in section 2.2, we have now removed this part from line 106. The removed MODIS data fields include e.g. viewing geometry and algorithm performance parameters, which were not considered in this study.

L112: We have removed the sentence "The reduced daily MODIS files are first created starting from the L2 MODIS Aqua AOD files (MYD04) over land areas by removing pixels that do not have valid aerosol retrieval results and by dropping some data fields which are not relevant for this study."

P4L120: Why did you use latitude and longitude as the collocation criteria? Using this criterion, the distance between OCO-2 and TCCON is different at different latitudes.

Reply: It is true that latitude affects the sampling in this method. Initially, this was a crude sampling done as a first step when constructing a database, and the adapted approach is similar to earlier studies focusing on the validation of satellite data (e.g. Wunch et al., 2017). The original geolocation information was included in the database. The idea was that different refined sampling criteria (within the larger crudely sampled area) could be tested easily later. When testing smaller sampling areas around TCCON sites, it turned out that the finer sampling had little effect when considering the full global dataset including all 26 TCCON sites. Hence, we kept the original crude collocation criteria. For more detailed studies the sampling should be revisited.

P4L128: 0.1° is also the latitude and longitude threshold? Please describe it. Why did you not use the same spatial and temporal criteria with TCCON?

Reply: In the case of AERONET, the sampling criterion was a circle of 0.1 deg around the AERONET site, assuming pseudo-cartesian lat/lon coordinates. We have now clarified this in the text.

The practical reason for different sampling for AERONET and TCCON is that the data were taken from existing collocation databases originally made for other purposes. For AERONET sampling we can afford a smaller sampling area due to the abundance of the data, while for TCCON we wanted to maximize the number of matches. In addition, the atmospheric lifetime of the species affects the choice of the co-location method. Different sampling criteria were tested also for AERONET, and we found that the results were not too sensitive to the sampling distance. Please see also the reply to comments for Referee #2 for more details.

L142: We have added the sentence “With the abundance of AERONET sites, we could afford a smaller sampling area than that used for TCCON data.”

P4L129: “The OCO-2 observations are not spatially averaged.” Not necessary for AERONET part.

Reply: This sentence refers to the comparison of OCO-2 total AOD component to AERONET AOD. Typically, in satellite vs AERONET comparison the satellite data around the AERONET site is spatially averaged and the AERONET data is temporally averaged around the satellite overpass time, to reduce the random noise (see e.g. Virtanen et al. 2018).

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.

L139: We have changed

"The OCO-2 observations are not spatially averaged."

to

"The OCO-2 observations within the 0.1° radius are included in the comparison individually (no spatial averaging). We note that the comparison statistics are typically affected by the spatial and temporal collocation parameters (see e.g. Virtanen et al. (2018)). Different sampling radii and time windows were tested with a subset of data, with minor effects on the results. With the abundance of AERONET sites, we could afford a smaller sampling area than that used for TCCON data."

P6L148: What is the specific definition of XCO₂ anomaly? How to calculate XCO₂ anomaly from the median XCO₂? The anomaly was calculated from all data, or for every year? After the calculation, have you removed the anomaly or how to use the anomaly? And you said “This is an alternative way to de-trend the data”, when did you use the LTC method and when did you use the XCO₂ anomaly? Please supplement this section with more information.

Reply: The first sentence in section 3.6, “The OCO-2 XCO₂ anomaly is calculated for each good quality OCO-2 pixel in the collocated dataset as the difference from the median XCO₂ value calculated within 500 km for the corresponding OCO-2 orbit.” defines the XCO₂ anomaly. It is calculated for each good quality pixel, as the difference of XCO₂ value from the spatially and temporally varying median. This reference median for each pixel is calculated from the good quality pixels on the same OCO-2 orbit, within 500 km from the pixel considered.

Since the values used to calculate the median are from the same orbit (within few minutes), using this anomaly is expected to remove both the trend and seasonal effects. The idea is that the yearly increase in CO₂ and the seasonal variation are large spatial scale effects which are captured by the 500 km portion of an orbit. When the median value is subtracted, the remaining ‘anomaly’ part is assumed to contain information on local sources and sinks, while the trend and seasonal effects are (presumably) removed. The XCO₂ anomaly is used as such (i.e. we do not attempt to ‘remove’ the anomaly from XCO₂ data).

The XCO₂ anomaly is used as a proxy of local CO₂ emission and compared with AOD data, to study the covariance of aerosols and CO₂ emissions. Since the anomaly data looks rather noisy when aggregated to global maps for five years, we do not show it as a map on the manuscript (see Fig. 1 a) below). Instead, XCO₂ anomaly statistics are shown in various tables in the manuscript (e.g. Table 2, Table A2, Table A3). In Table 2, the XCO₂ anomaly is larger in AOD quarter Q2 than in Q1, indicating a correlation between local XCO₂ enhancement (emissions) and MODIS AOD. Figure 1 b) below illustrates this. The XCO₂ anomaly analysis supports our findings in the manuscript and shows essentially the same results as obtained with ‘usual’ XCO₂ data, so they are not shown in more detail in the manuscript.

We have slightly extended the description of XCO₂ anomaly in section 3.6 to clarify these points.

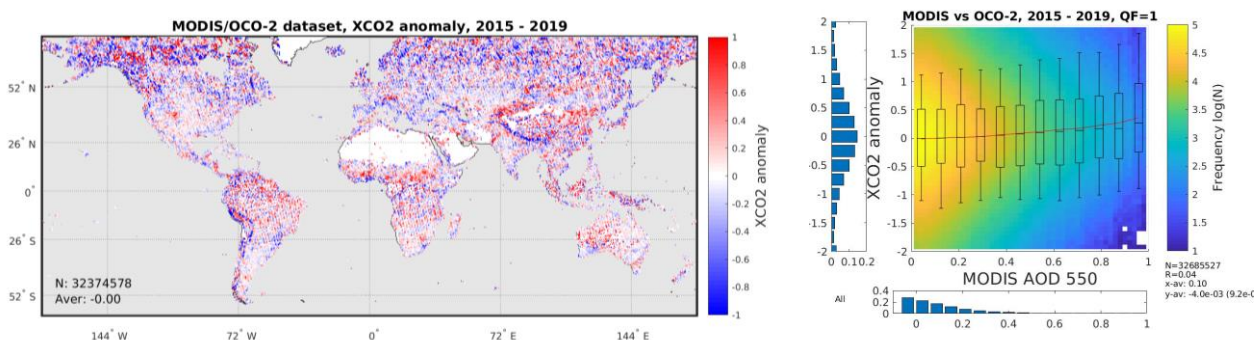


Figure 1. a) Local XCO₂ anomaly data [ppm] for five years aggregated to 0.5 deg grid cells. b) Correlation observed between XCO₂ anomaly and AOD.

L165: We have changed

"The OCO-2 XCO₂ anomaly is calculated for each good quality OCO-2 pixel in the collocated dataset as the difference from the median XCO₂ value calculated within 500 km for the

corresponding OCO-2 orbit. This is an alternative way to de-trend the data, instead of applying the simple LTC. Unlike LTC, the anomaly method also effectively de-seasonalises the data. It also allows to study the covariance of AOD values and local XCO₂ anomalies caused by possible CO₂ sources and sinks."

to

"The OCO-2 XCO₂ anomaly is calculated for each good quality OCO-2 pixel in the collocated dataset as the difference from a local, temporally varying median value. This median is calculated from the good quality pixels in the same OCO-2 orbit, within 500 km from the pixel considered. The idea is that the yearly increase in CO₂ and the seasonal variation are large spatial scale effects which are captured by the 500 km portion of an orbit. When the median value is subtracted, the remaining 'anomaly' part is assumed to contain information on local sources and sinks, while the trend and seasonal effects are removed. This is an alternative way to de-trend the data, instead of applying the simple LTC. Unlike LTC, the anomaly method also effectively de-seasonalises the data. It also allows to study the covariance of AOD values and local XCO₂ anomalies caused by possible CO₂ sources and sinks. While most of the results shown in this work have been processed with the linear trend correction, the corresponding XCO₂ anomaly results are also shown where appropriate to support the analysis."

P6L167: The dust loads may be identified by the areas, but "biomass burning aerosols increase AOD in central Africa and South-East Asia" lack of evidence.

Reply: We think that the high AOD values in these areas are often contributed to biomass burning, see e.g. van der Werf et al. (2010), Myhre et al. (2009), or Kinne (2019). As this is not in the focus of the manuscript, we have removed the reference to biomass burning aerosols.

References:

Van der Werf et al., 2010, www.atmos-chem-phys.net/10/11707/2010/

Myhre et al., 2009, www.atmos-chem-phys.net/9/1365/2009/

Kinne, 2019, <https://doi.org/10.1080/16000889.2019.1623639>

L190: We have removed the text " , biomass burning aerosols increase AOD in central Africa and South-East Asia,".

P6L172: I think the largest differences appear in Central Asia and South Asia.

Reply: Our definition of 'South-East Asia', as shown in Fig. A7, is rather loosely defined. It covers areas usually understood as South Asia, East Asia and parts of South-East Asia.

L195: We have changed

"The largest differences in AOD appear to be concentrated largely in the high AOD areas in South-East Asia, where OCO-2 AOD is lower than MODIS AOD."

to

"The largest differences in AOD appear to be concentrated largely in the high AOD areas in parts of Asia, where OCO-2 AOD is lower than MODIS AOD."

P6L173: "These positive difference values are related to the MODIS DT algorithm permitting small negative AOD values." Any references?

Reply: The negative MODIS AOD values are discussed in Sayer et al. (2014). This reference was given after a couple of sentences on line 176. We have moved the citation to line 173.

P7 Fig.1: The colorbar of panel (b) should be adjusted because 0 and invalid data are both white.

Reply: We have duly adjusted the color for missing data in Fig. 1 b).

P8L187: You should point out earlier.

Reply: The discussion on the wavelength issue is lengthy, which is why it was delayed until this point. We added a note on it at the beginning of section 4.1, with reference to later discussion.

L193: We have added the sentence "Note that the OCO-2 AOD is retrieved at 755 nm, while the MODIS AOD is obtained at 550nm; the effect of the wavelength difference will be discussed below."

P8L200: You acquired 770 nm AOD by averaging 675 nm AOD and 870 nm AOD. Can you really achieve that by the simple average? Please give more evidence or reference.

Reply: A simple average is not the optimal way to scale the wavelengths but should be acceptable in this case. Using scaling with Ångström exponent with a smaller subset of data did not change the results dramatically. Please see comments to Referee #2 for more detailed discussion.

We have added a note on this to section 3.3.

L144: We have added the sentence "While this simple approach may not be the most accurate, it is sufficiently accurate for our purposes. A more accurate method for the wavelength scaling using the Ångström exponent from AERONET was tested for a subset of data, and we did not find significant differences in the results."

P8L210: How did you define urban areas?

Reply: The urban areas are discussed from line 329 on, where references are given. We have added a note on this earlier to L210.

L244: We have added the note "(see Fig. A7 for definition of urban areas)".

P9L217: Please point out the spatial resolution of OCO-2 again.

Reply: We have added the OCO-2 resolution (approx 1x2km²) to the text here.

L254: We have added the text "(approximately 1×2 km²)".

P10L250: How did you handle with the AOD exactly equal to 0.2?

Reply: There were practically no data points with AOD exactly at 0.2. They are included in the larger AOD bin. We have changed the text to replace “AOD<0.2” by “AOD<=0.2”.

P13Figure 4 (b): You should distinguish 0 values and invalid values. Both of them are white.

Reply: We have duly changed the color for missing data.

P13L299: Table A3 or Figure A3?

Reply: The reference was incomplete. We have changed this to “(see Table A3)”.

P14L316: What does “the measured CO₂ absorption is divided into too short distance” mean?

Reply: The measured radiance at CO₂ absorption channels carries information of the total column load of CO₂ in the observed light path. The inversion of measured radiance signal to CO₂ column load (XCO₂) requires information on the light path length. The light path length is affected by aerosols: more aerosol means effectively more scattering, and a longer light path. If the retrieved AOD is incorrect, the light path length is also incorrect, and this directly affects the retrieved XCO₂ value. We have revised the text to state this more explicitly.

L375: We have replaced

"If the aerosol load is underestimated in the retrieval (Q2), the light path is also underestimated, and the measured CO₂ absorption is divided into too short distance, leading to overestimation of XCO₂. Similarly, if AOD is overestimated, the light path is also overestimated, causing underestimation of XCO₂."

by

"The top of atmosphere radiance measured by OCO-2 contains information on the total amount of CO₂ along the light path, and inversion of this information to XCO₂ values requires knowledge of the light path length, which is affected by aerosols. If the aerosol load is overestimated in the retrieval (Q4), the light path is also overestimated, and the measured CO₂ absorption is divided into too long distance, leading to underestimation of XCO₂. Similarly, if AOD is underestimated (Q2), the light path is also underestimated, causing overestimation of XCO₂."

P15L340: Again, if you acquire OCO-2 XCO₂ anomaly by calculating the difference between the OCO-2 XCO₂ data and the median value along 500 km orbit, how can the anomaly be used to de-seasonalize and de-trend the data? Please describe more details about this.

Reply: The XCO₂ anomaly values are used as such, instead of using them to ‘correct’ the XCO₂ values. It is assumed that the large scale seasonal and trend effects are removed when subtracting the 500 km median, and that the anomaly contains information on the XCO₂ variability in the local spatial scale, including local sources (and sinks). Please see the extensive reply to the previous similar comment above (reply to comment P6L148).

P15Table2: "TCCON(1)" should be revised as "OCO-2(1)"?

Reply: The naming refers to different subsets of OCO-2 data, hence 'global', 'urban' and 'TCCON'. To distinguish between the two data sources (OCO-2 and TCCON) for the collocated OCO-2 & TCCON dataset, we have used the numbers in parenthesis, (1) and (2). We have revised the Table caption to clarify this.

Table 2 caption: We have replaced "For collocated TCCON data two XCO₂ values are given, from OCO-2 (1) and from TCCON (2), respectively." with "Three datasets are used: global, urban and one collocated with TCCON. For the collocated TCCON data two XCO₂ values are given, from OCO-2 (labeled TCCON(1)) and from TCCON (labeled TCCON(2)), respectively."

P16L365: "OCO-2 seems to slightly overestimate XCO₂ for low AOD values, and underestimate at high AOD values."

-For panel (a) of Figure A8, you used MODIS AOD and XCO₂ difference (deviation of OCO-2 from TCCON). MODIS AOD and TCCON XCO₂ can be considered as references, so you can definitely get the statement about overestimate or underestimate of OCO₂ XCO₂ at different AOD.

Reply: We have revised the text.

L429: We have replaced "OCO-2 seems to slightly overestimate XCO₂ for low AOD values, and underestimate at high AOD values" with "OCO-2 slightly overestimates XCO₂ for low AOD values, and underestimates at high AOD values".

-For panel (c), it seems that the XCO₂ difference is close to zero when the AOD difference is in the range of 0 to 0.1, and when the AOD difference becomes higher or lower, the XCO₂ difference will be minus, which means the underestimate of OCO₂ XCO₂. It seems to be not totally consistent with the statement of Figure 5 "If the aerosol load is underestimated in the retrieval (Q₂), the light path is also underestimated, and the measured CO₂ absorption is divided into too short distance, leading to overestimation of XCO₂. Similarly, if AOD is overestimated, the light path is also overestimated, causing underestimation of XCO₂." Do you have any comments on it?

Reply: The discussion at Figure 5 was meant as the first impression and possible explanation of the results seen in Fig. 5 a), accompanied by the disclaimer that it is not easy to distinguish the two effects of co-emission of aerosols and CO₂ and the aerosol bias on XCO₂ retrieval. The discussion of the light path is a possible explanation, which turned out not to be consistent with the results obtained with collocated TCCON data, as correctly pointed out by the reviewer.

We find that it is difficult to draw detailed conclusions from Fig. A8 c), because there are two dependencies mixed in the plot. First, the AOD difference between OCO-2 and MODIS depends on the MODIS AOD in a nontrivial way as shown in Fig. 2, with OCO-2 low bias at one end and high bias at the other. Second, the XCO₂ bias depends also on MODIS AOD. We find that Fig. A8 a) best describes what we can say about the XCO₂ bias, but nevertheless we wanted to show Fig. A8 c) for completeness.

It is certainly important to state these points explicitly in the manuscript. We have revised the discussion to section 4.3 to achieve this.

L374: We have replaced

"The striking connection between XCO₂ and the relative AOD values between the two instruments in Fig. 5 a) can potentially be explained by the light path length used in the ACOS full physics retrieval."

by

"As a first guess, the striking connection between XCO₂ and the relative AOD values between the two instruments in Fig. 5 a) could potentially be explained by the light path length used in the ACOS full physics retrieval."

L382: We have added the sentence: "However, for Q2 the interpretation turns out to be more complicated, when the reference XCO₂ data from TCCON is considered, as discussed below."

L430: We have added the text: "As with Fig. 7 c), the interpolation of Fig. A8 c) is complicated, since there are two aerosol related dependencies affecting the data. First, the AOD difference between OCO-2 and MODIS depends on the MODIS AOD in a nontrivial way as shown in Fig. 2, with OCO-2 low bias at one end and high bias at the other. Second, the XCO₂ bias depends also on MODIS AOD."

P18L388: The XCO₂ and AOD also show negative correlation in North America and Europe. Do you have any comments on this?

Reply: Here we wanted to highlight only the most obvious differences between the areas. In Fig. 8 c) the negative slope of XCO₂ as function of AOD is most obvious for North Asia, but the linear fit slopes are negative also for Europe and South America (for North America the slope is slightly positive). Looking at Fig. 4 b), we see that in North America there are areas of negative and positive correlation, but the data is noisy. For South America the correlation is weaker, and close to zero in large areas.

Actually, most of the areas have negative correlation (when looking at the matrix plots, not shown in the manuscript), with the exception of Africa and 'South-East' Asia.

We have reformulated the text.

L465: We have replaced

"Most of the areas show a positive correlation between XCO₂ and AOD, with the clear exception of Northern Asia, which is dominated by strong seasonal cycle of XCO₂. We also note that the northern areas are strongly undersampled in winter months due to snow cover and high SZA, which prevent MODIS aerosol retrievals."

by

"SE Asia and Africa show a positive correlation between XCO₂ and AOD, N America, S America, and Australia have slopes closer to zero, while Europe and in particular Northern Asia have negative slopes. The northern areas are dominated by strong seasonal cycle of XCO₂, and are strongly undersampled in winter months due to snow cover and high SZA, which prevent MODIS aerosol retrievals."

P19L395: “Satellite XCO₂ retrievals are known to have higher uncertainty in high aerosol conditions.” Where is this statement summarized from? From Figure A8 (a) and (b), it can be seen that the OCO-2 XCO₂ also has deviation from TCCON XCO₂ at low AOD.

Reply: This sentence refers to earlier research in the field; we have added references to Connor et al. (2016) and O’Dell et al. (2018). Figure A8 a) and b) are plotted using quality filtered OCO-2 data, where high AOD cases (OCO-2 AOD over 0.2) have already been removed. Unfiltered OCO-2 data would show a larger spread (see Fig. 2 below). However, it is true that there is considerable spread in the data also at the low AOD region.

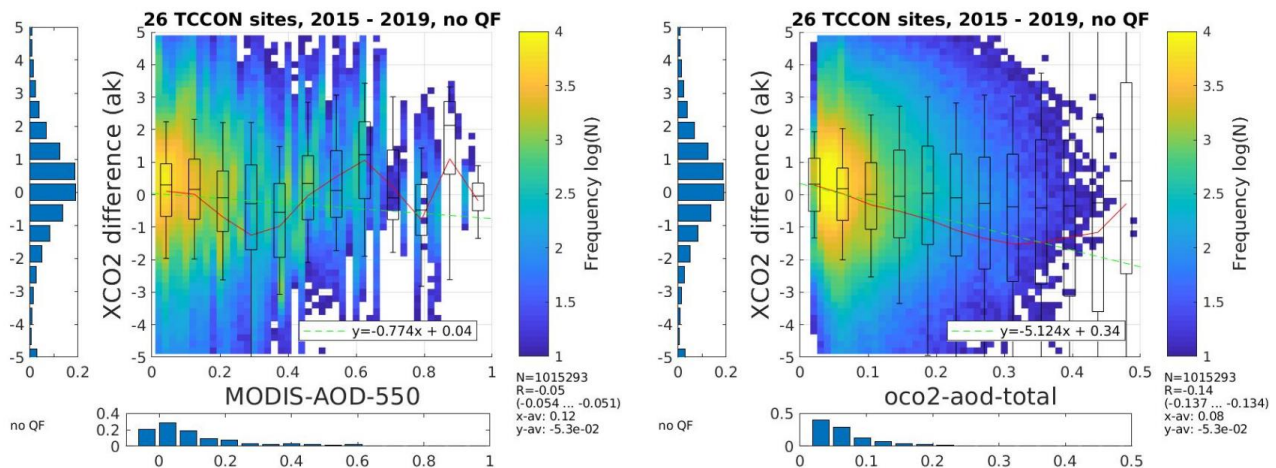


Figure 2. Same as Fig. A8 a) and b) in the manuscript, but without OCO-2 quality filtering.

L473: We have added the references “(Connor et al., 2016; O’Dell et al., 2018)”.

P21L427: For future CO₂M, you found the large increase of data coverage after adjusting AOD threshold from 0.2 to 0.5. After adjusting, there will be more underestimated and overestimated data when the AOD is below 0.5, i.e., the data in Q1 area you defined will include data in Q2 and Q4 areas. You should discuss how to conduct quality filter for CO₂M, or the data quality with the AOD threshold of 0.5.

Clearly the discussion in section 4.5 was not sufficiently clear, as pointed out by all referees. The idea here is that the coming CO₂M mission will have a dedicated aerosol instrument - Multi-Angle Polarimeter (MAP) – which is expected to allow XCO₂ retrieval at heavier aerosol conditions, with AOD threshold of 0.5.

Here we use the collocated, quality filtered OCO-2/MODIS dataset as a proxy for CO₂M data. This dataset includes high MODIS AOD pixels, although the OCO-2 quality filter including an AOD threshold of 0.2 has been applied. We assume that the OCO-2 quality filtering assures that the XCO₂ data is of good quality even for higher MODIS AOD cases, as CO₂M data is expected to be up to AOD of 0.5. We further assume that the MODIS AOD in the collocated dataset is representative of ‘true’ AOD and can be used to study the AOD threshold, even though the OCO-2 quality filtering has removed a large part of the original pixels.

With this collocated data set, we can test what is the effect of relaxing MODIS AOD threshold from 0.2 to 0.5. Note that this does not mean that we extend the OCO-2 coverage; the MODIS AOD threshold used here is an additional constraint on the quality filtered OCO-2 data.

In the manuscript we did not address the effect of different MODIS AOD thresholds on the XCO₂ retrieval quality. Figure 3 below shows such assessment: We find that in this case the MODIS AOD threshold did not have any significant effect on the XCO₂ retrieval quality. But as mentioned, in these cases we use the quality filtered OCO-2 data, which should be of sufficiently good quality. Having said that, we need to clarify that the purpose of this study was not to improve the capability of OCO-2 to deal with higher aerosol loads, but to find how increasing the AOD threshold affects the coverage.

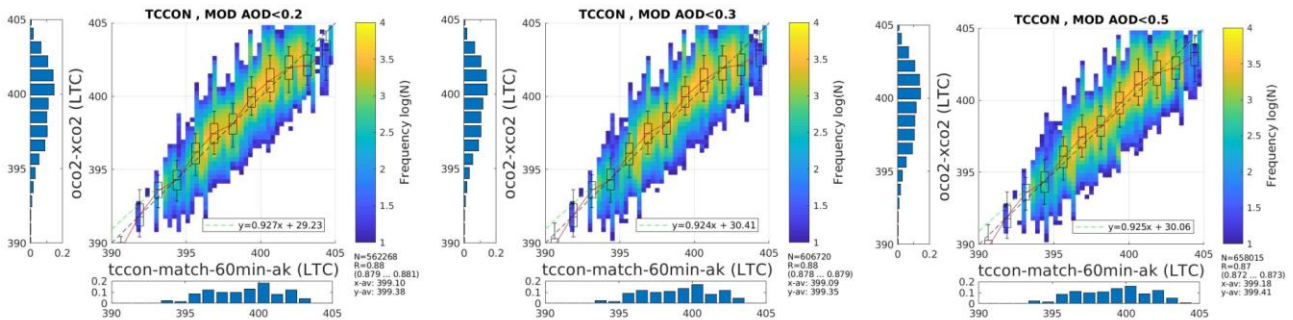


Figure 3. Effect of MODIS AOD threshold on XCO₂ comparison between OCO-2 and TCCON. The OCO-2 quality flag has been applied.

We do not propose replacing the OCO-2 quality filtering with a single MODIS AOD threshold. The OCO-2 quality filter is based on 32 different tests, of which the total AOD threshold of 0.2 is only one component. Replacing this with a MODIS AOD threshold, even as stringent as 0.2, would result in reduced XCO₂ quality, as shown in Figure 4 below.

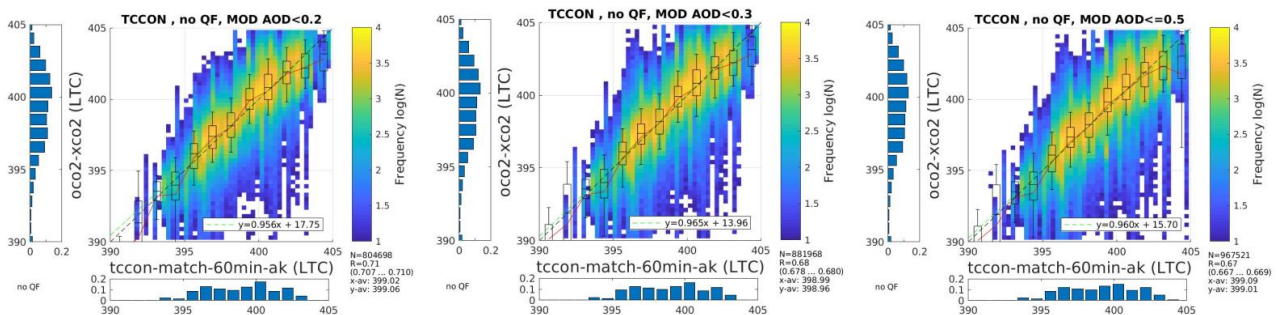


Figure 4. Effect of MODIS AOD threshold on XCO₂ comparison between OCO-2 and TCCON, when the OCO-2 quality flag has not been applied.

In conclusion, here we treat the quality filtered OCO-2 data as a proxy of the coming CO₂M data, which can be further filtered by using MODIS AOD thresholds. Our main finding, then, is that if CO₂M can handle AODs up to 0.5, this will increase coverage, particularly in urban areas, compared to a case where AOD only up to 0.2 could be allowed.

We have clarified the text in section 4.5 to bring these points explicitly out.

L477: We have added the sentence: “The CO₂M mission will include a dedicated aerosol instrument, the Multi-Angle Polarimeter (MAP), and is expected to be better equipped to deal with high aerosol conditions.”

L480: We have added the text:

“In this section we estimate how the selected AOD threshold affects the coverage of satellite XCO₂ retrievals, in particular in urban areas with high co-emission of aerosols and CO₂.”

Here we use the collocated, quality filtered OCO-2/MODIS dataset as a proxy for CO2M data. This dataset includes high MODIS AOD pixels, although the OCO-2 quality filter including an AOD threshold of 0.2 has been applied. We assume that the OCO-2 quality filtering assures that the XCO₂ data is of good quality even for higher MODIS AOD cases, as CO2M data is expected to be up to AOD of 0.5. This assumption is supported by a comparison of quality filtered OCO-2 XCO₂ data against TCCON, where additional collocated MODIS AOD thresholds had minimal effect on the retrieval quality (not shown). We further assume that the MODIS AOD in the collocated dataset is representative of 'true' AOD and can be used to study the AOD threshold, even though the OCO-2 quality filtering has removed a large part of the original pixels. With this collocated data set, we can test what is the effect of relaxing MODIS AOD threshold from 0.2 to 0.5. We emphasise that this does not mean that we extend the OCO-2 coverage (or propose to relax the OCO-2 AOD threshold); the MODIS AOD threshold used here is an additional constraint on the quality filtered OCO-2 data."

L504: We have added the text:

"In conclusion, here we have used the quality filtered OCO-2 data as a proxy of the coming CO2M data, which can be further filtered by using AOD thresholds from collocated MODIS data. We find that if CO2M can handle AODs up to 0.5, this will significantly increase the coverage, in particular in the urban areas, compared to a case where AOD only up to 0.2 could be allowed. We also find that due to the correlation found between AOD and XCO₂, including data with higher AOD increases the mean XCO₂ values, especially for the urban pixels."

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.

L86: We have added the sentence: "While the global AOD patterns are somewhat different between DT and DB, we find that the global statistics and conclusion regarding the connection to XCO₂ retrievals are largely the same."

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.

L71: We have replaced

"The aerosol parameters of the ACOS algorithm include five scatterers, two cloud types (water and ice), two tropospheric aerosol types and a stratospheric aerosol type (sulfate). Two representative types of tropospheric aerosols (dust, smoke, sulfate aerosol, organic carbon, and black carbon) are drawn from a climatology based on location and time (Crisp et al., 2021)."

by

"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). The two most representative types of tropospheric aerosols out of five possible types (dust, sea salt, sulfate aerosol, organic carbon, and black carbon) are drawn from collocated 3-hourly aerosol fields from

Goddard Earth Observing System Model, Version 5, Forward Processing for Instrument Teams (GEOS-5 FP-IT; see Crisp et al. (2021))."

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.

L107: We have added the reference "Sinyuk et al. (2020)".

L108: We have added the reference "(Giles et al., 2019)".

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.

L124: We have added the sentence “We note that although both data products are cloud screened, possible mutual cloud contaminated pixels can cause erroneous high AOD values, which may affect the obtained correlation coefficients.”

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.

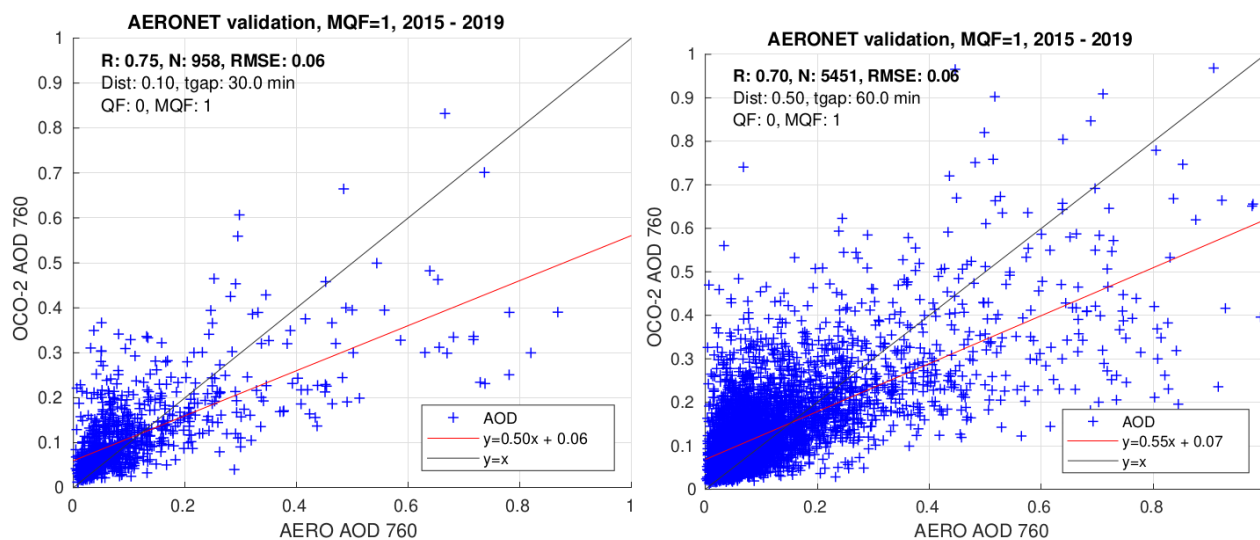


Figure 1. Effect of different collocation criteria on the AERONET comparison. The results differ only slightly when using a) distance of 0.1 degree for averaging OCO-2 data around the AERONET site and a temporal averaging window of ± 30 min centered at the OCO-2 overpass time for AERONET data; or b)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

L140: We have added the text “We note that the comparison statistics are typically affected by the spatial and temporal collocation parameters (see e.g. Virtanen et al. (2018). Different sampling radii and time windows were tested with a subset of data, with minor effects on the results.”

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.

L94: We have added the text “Note that the MODIS quality flag is systematically applied throughout the results in this paper, while the use of OCO-2 quality flag varies. In the rest of the paper, when the use of quality flag or quality filtering is discussed, this refers to the OCO-2 quality flag.”

L189: We have replaced “(using the OCO-2 quality flag)” by “(the MODIS quality flag is always applied; here we use also the OCO-2 quality flag)”.

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.

L266: We have added the text: “To conclude, in this section we considered the differences between collocated the MODIS DT AOD product and the OCO-2 total AOD component. We find that the AOD difference depends on region. OCO-2 tends to overestimate the aerosol load in regions with low MODIS AOD. More important for the XCO₂ retrievals, OCO-2 tends to severely underestimate AOD in the high MODIS AOD regions (including areas with high anthropogenic emissions), which may have an effect on the XCO₂ retrievals in these regions.”

L272: We have added the text: “In this section we will compare the OCO-2 total AOD component to MODIS AOD statistically for the full collocated dataset using e.g. density scatter plots. Specifically, we address the question of how well the OCO-2 quality filtering works from the point of view of aerosols. The OCO-2 quality filter uses an AOD threshold of 0.2, among several other tests, to remove heavy aerosol conditions. We use collocated MODIS AOD data to assess the performance of the OCO-2 AOD filter.”

L341: We have added the text: “To conclude this section, we have found that the quality filtered OCO-2 data contains a large fraction of data with high MODIS AOD, potentially affecting the XCO₂ retrieval quality. These data are more frequent in densely populated areas with high aerosol and CO₂ emissions. Hence, for monitoring anthropogenic CO₂ emissions with satellites, it is crucial that the high AOD cases are carefully detected and treated in the satellite retrievals.”

L446: We have added the text: “To conclude this section, we find that there is a linear relation between OCO-2 XCO₂ and MODIS AOD (Fig. 6 a). We also find a linear relation between the OCO-2 XCO₂ bias and MODIS AOD (Fig. A8 a). We also find a relation between the AOD difference between OCO-2 and MODIS and the OCO-2 XCO₂ values, as shown in Fig. 8 a). Aerosols are related to OCO-2 XCO₂ retrievals in two ways: there is a real correlation between XCO₂ and AOD, due to co-emission of aerosols and CO₂. There is also an aerosol related bias in the OCO-2 retrievals, which acts in opposite direction than the co-emission but with smaller magnitude, thus partly masking the co-emission effect. However, we are unable to directly relate the AOD difference observed between OCO-2 and MODIS to the XCO₂ difference observed between OCO-2 and TCCON for the quality filtered data. This is due to the non-trivial AOD difference observed between OCO-2 and MODIS, further complicating the entanglement caused by the two competing aerosol effects.”

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.

L212: We have added the text “We note that possible mutual cloud contamination of collocated data points could lead to erroneous high AOD values for both instruments, possibly leading to higher correlation values than without cloud contamination. However, data from each satellite is cloud screened with their respective cloud masks, and the vast majority of data is in the low AOD region, reducing the probability of large bias.”

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.

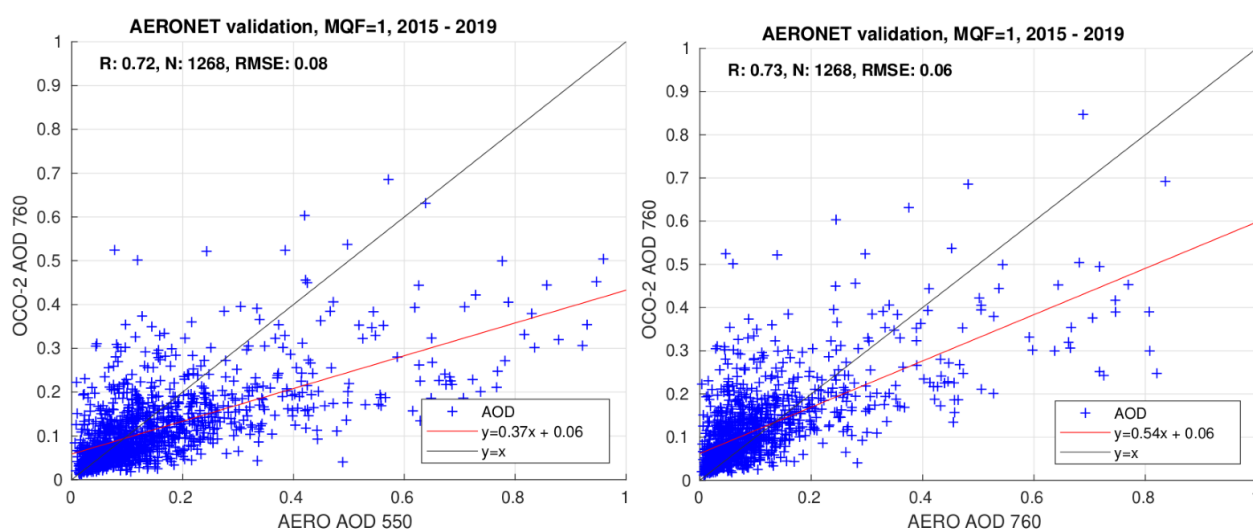


Figure 2. Comparison of OCO-2 AOD at 760 nm against AERONET AOD at 550 nm (left) and against AERONET AOD at 760 nm (right). The AERONET AOD values are scaled to different wavelengths using the AERONET Ångström exponent.

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.

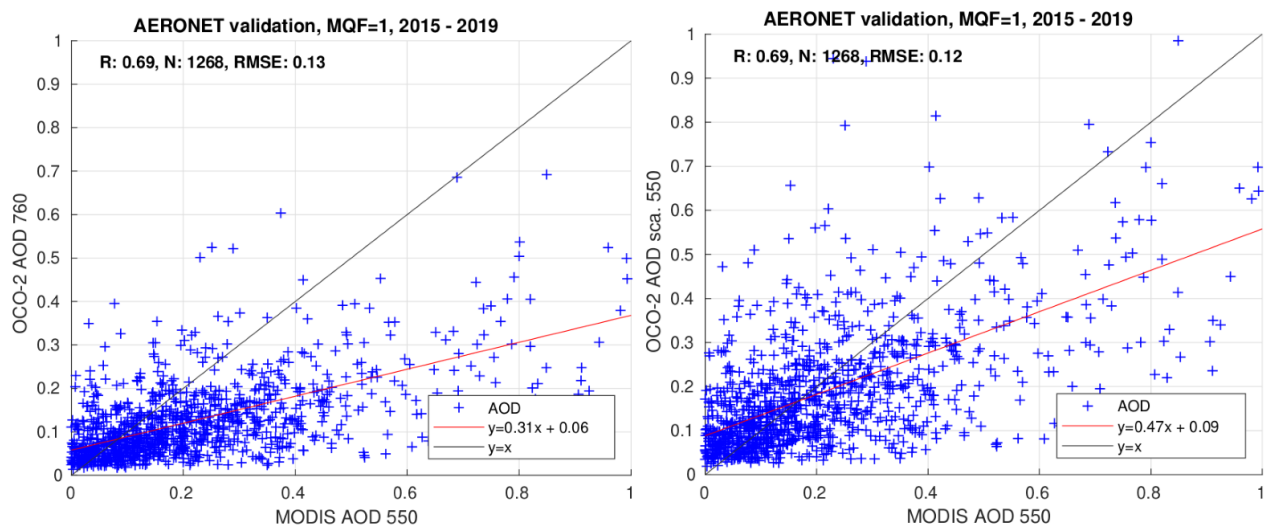


Figure 3. Comparison of OCO-2 AOD to MODIS AOD using the original wavelengths (left), and using AERONET Ångström exponent to scale OCO-2 AOD to 550 nm (right).

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.

L228: We have added the text “The use of MERRA-2 data potentially induces high uncertainty to the spectral conversion. We use this method merely to get a rough estimate of 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. The spectral conversion was repeated with a smaller subset of data using Ångström exponent from AERONET, and the results largely agreed with the global dataset.”

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. (See point 6 above for the description of changes made in the text.)

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.

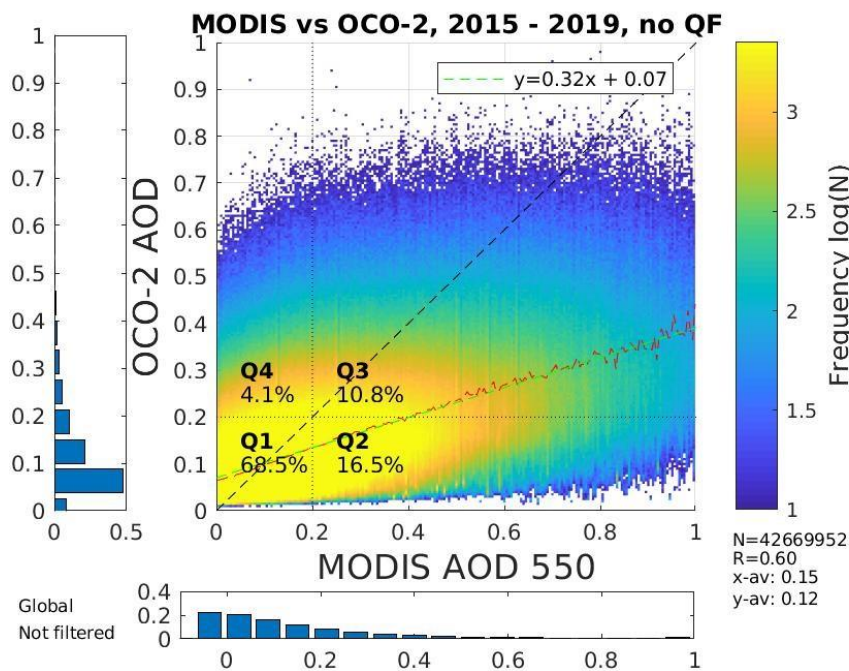


Figure 4. Density scatter plot

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.

L293: We have added the text “The large spread of the data reflects the fact that the ACOS algorithm is not optimized for AOD retrieval, as discussed above. Considering this, the obtained correlation with MODIS AOD can be considered acceptable.”

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.

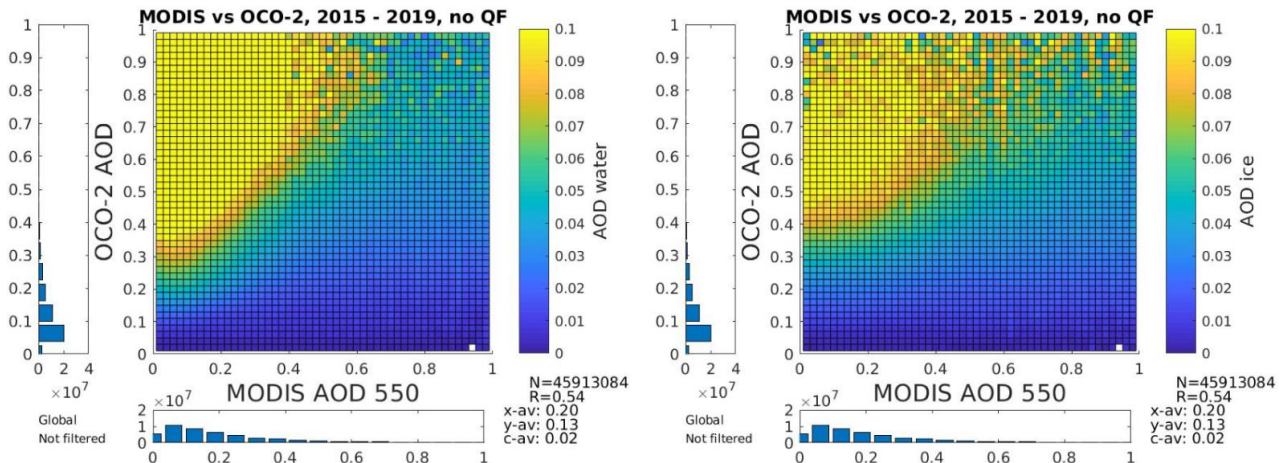


Figure 5. Contributions of water and ice components to the OCO-2 total AOD.

L283: We have added the text “These two AOD components are included in the ACOS retrieval to account for possible residual cloud contamination, while the MODIS aerosol retrieval does not have corresponding elements. Preliminary study shows elevated water and ice AOD values at low MODIS AOD values, but a more detailed study, beyond the scope of this work, would be required to confirm this.”

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.

L291: We have replaced the text "Pearson correlation coefficient for the unfiltered data is 0.60, reducing to 0.52 for the filtered data, which indicates that the sampling is biased for the quality filtered data (higher MODIS AOD values remain in the collocated data set). Note that in the collocated dataset the MODIS data is often the limiting factor (Table A1), already removing data over bright surfaces and in proximity of clouds. Applying the OCO-2 quality filter further reduces the collocated data to 56% of the original collocated data points. Most, but not all, of this reduction can be contributed to removing the high AOD cases."

by

"The Pearson correlation coefficient for the unfiltered data is 0.60, reducing to 0.52 for the data filtered with the OCO-2 quality filter. The large spread of the data reflects the fact that the ACOS algorithm is not optimized for AOD retrieval, as discussed above. Considering this, the obtained correlation with MODIS AOD can be considered acceptable. Note that in the collocated dataset the MODIS data is often the limiting factor (Table A1), already removing data over bright surfaces and in proximity of clouds. Applying the OCO-2 quality filter further reduces the collocated data to 56% of the original collocated data points. Note that only 15% of the original data is removed by the total AOD threshold of 0.2, while 29% are removed by other quality tests. The lower correlation coefficient of the quality filtered dataset reflects the imbalance between OCO-2 and MODIS in the AOD distribution of data points removed by the OCO-2 quality filter."

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 XCO₂ 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.

L365: We have added the note: "(see Fig. 2 for the number of data)".

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.

(See point 6 above for the description of changes made in the text.)

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.

(See our reply to Referee #1 for the changes made to the text.)

Referee #3

This paper focuses on the collocation of OCO-2 and MODIS data, and analyzes the relationship between CO₂ retrievals from OCO-2 and AOD retrievals from both OCO-2 and MODIS. The authors demonstrate that errors in AOD retrievals affect XCO₂ retrievals, and also show that excluding data points with moderate AOD (0.2-0.5) excludes many areas with high XCO₂. As a result, the authors recommend relaxing the MODIS AOD threshold to 0.5.

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

We see that our main message in section 4.5 was not stated clearly enough. We do not actually recommend relaxing the AOD threshold to 0.5 for the (OCO-2) XCO₂ retrievals; we simply wanted to test the effect of using the looser threshold (planned to be used with CO₂M) on the coverage in different environments. We have modified the abstract and conclusions to emphasize our main results, and section 4.5 to clarify the purpose of the aerosol threshold exercise.

L20: Added sentence "This crucial for monitoring anthropogenic CO₂ emission, considering the observed co-emission of aerosols and CO₂."

L512-523: We changed

"There is also co-variability between the AOD difference and the retrieved XCO₂ values, most strikingly visible for the unfiltered dataset: overestimation of AOD in the OCO-2 retrieval leads to lower XCO₂ values, and underestimation of AOD leads to higher XCO₂ values. We have also found evidence of real covariance of AOD and XCO₂, which is partly masked by the aerosol effect on the XCO₂ retrieval. This covariance is presumably at least in part due to co-emission of anthropogenic CO₂ and aerosols."

to

"The observed difference depends on location and conditions, but on average OCO-2 tends to overestimate at low aerosol loads and underestimate at higher AODs. We have found evidence of covariance of AOD and XCO₂, which is presumably at least in part due to co-emission of anthropogenic CO₂ and aerosols. There is also co-variability between the AOD difference and the retrieved XCO₂ values, most strikingly visible for the unfiltered dataset: overestimation of AOD in the OCO-2 retrieval is observed at lower XCO₂ values and underestimation of AOD at higher XCO₂ values. Comparison with TCCON reveals a weak but statistically significant dependence of the XCO₂ bias on the AOD, such that at high AOD OCO-2 tends to underestimate XCO₂. This aerosol bias acts in the opposite direction than the observed covariance between AOD and XCO₂, partly masking the correlation. However, disentangling the effects of real covariance and aerosol bias is not straightforward, and we were not able to directly connect the observed AOD difference between MODIS and OCO-2 to the XCO₂ difference observed between TCCON and OCO-2."

L527: Added sentence: "In the light of the correlation found between AOD and XCO₂, the AOD threshold affects also the average XCO₂ values of the quality filtered data."

Major comments

The authors note (in lines 181-185) that there is low correlation between MODIS AOD and OCO-2 AOD in Australia, the Sahel, the Western US, and Central Asia, using MODIS Dark Target observations. However, these areas seem to include bright land surfaces like large deserts and snowy mountain ranges, so perhaps using MODIS Deep Blue would be more appropriate for such areas. The Western US and Sahel also show high correlation between XCO2 and MODIS AOD in Figure 4b. Can this be explained by poor quality MODIS Dark Target observations? Would using MODIS Deep Blue for these areas change the analysis?

Reply: Indeed, the areas of low correlation listed in the manuscript all have bright surfaces, as noted by the Referee. This might be one contributing factor to the observed correlations. We have added this to the text. However, as for MODIS Deep Blue algorithm, the same areas have low AOD correlations, as shown in Fig. A2 c). We have added a more explicit reference to this in the text.

L189: We have changed

"High AOD areas due to anthropogenic aerosol emissions are seen in particular in parts of Asia, biomass burning aerosols increase AOD in central Africa and South-East Asia, and elevated aerosol loads due to dust are seen over various desert areas around the globe."

to

"High AOD areas due to anthropogenic aerosol emissions are seen in particular in parts of Asia and elevated aerosol loads due to dust are seen over various desert areas around the globe."

Section 4.5: The authors show that increasing the AOD threshold to 0.5 will increase the fraction of acceptable data, but do not show or discuss how this will affect the quality of XCO2 retrievals. It seems that improving the quality of XCO2 retrievals at higher aerosol loads remains an open challenge -- the authors should state this explicitly.

Reply: Clearly the discussion in section 4.5 was not sufficiently clear, as pointed out by all referees. The idea here is that the coming CO2M mission will have a dedicated aerosol instrument - Multi-Angle Polarimeter (MAP) – which is expected to allow XCO2 retrieval at heavier aerosol conditions, with AOD threshold of 0.5. Here we use the MODIS AOD data collocated with good quality OCO-2 retrievals, which includes high MODIS AOD pixels (although the OCO-2 quality filter including an AOD threshold of 0.2 has been applied). With this collocated data set, we can test what is the effect of relaxing MODIS AOD threshold from 0.2 to 0.5.

We have clarified the text in section 4.5 to bring these points explicitly out. Please see our reply to Referee #1 for a more detailed answer.

Minor comments/technical corrections

Line 299: Change (A3) to (see Fig. A3)

Reply: The reference was incomplete. We have changed this to “(see Table A3)”.