Author responses to comments of Anonymous Referee #1.

The authors would like to thank Anonymous Referee #1 for his/her detailed review of the manuscript. There are enough comments, especially concerning the satellite comparison study, that would benefit from some additional explanation and the authors feel it is appropriate for us to submit a revised manuscript. We will however follow the wishes of the Editor in this regard. If the explanations presented below are deemed sufficient then a revised manuscript may not be necessary.

Here we will attempt to respond to all concerns and answer all questions adequately.

General Comments:

1. The abstract does not mention that the aerosol extinction comparisons with the CALIPSO measurements were done after the in situ measurements were adjusted to ambient relative humidity. This should be mentioned.

This is an important detail. This has been added into the abstract in a revised manuscript.

2. (page 17194, line 25) Was there any noticeable difference in the measurements between the beginning and ending of the flights that may be traced to the change in the orientation of the inlet?

We looked at this early on and found that it was not an easy question to answer. The early parts of the profiles were conducted in the free troposphere whereas the measurements later in the profiles were down in the boundary layer. The aerosols at these different altitudes can be quite different with respect to amount, particle size and composition. Determining if there was any effect of a slightly changing inlet orientation is difficult under these conditions because the same aerosol is not being sampled during the two comparison periods. As discussed in the manuscript we oriented the inlet to be isoaxial during the boundary layer sampling legs where most of the aerosol resides. We did this intentionally because we wanted accurate measurements for comparison with the remote (i.e., satellite, sunphotometer) sensors. During the free troposphere legs early in the flight, the airplane flew in a 'nose-up' orientation by 1-2 degrees (as determined by pitch angle). This could conceivably affect the inlet transmission efficiency at higher altitudes, but it is unknown to what extent. So the answer to the reviewer's question is no, no noticeable differences were observed that could be attributed to inlet orientation.

3. (page 17198, line 20) Since g is used to adjust the wet and dry measurements, it would be helpful to show the equation that makes use of g to make these computations.

The gamma exponent for each valid flight segment was derived as described on page 17198, Lines 15-25, based on a power law fit of the three segment–average scattering vs. RH points. Once the gamma exponent has been determined for the given flight segment, the *f(RH)* was computed by using the equation given in Table 2. For the determination of *f(RH)*, RH_{wet} = 85 and RH_{dry} = 40 by definition. To scale the dry scattering values to ambient RH, we used the same general equation and the derived gamma exponent value. The equation used was

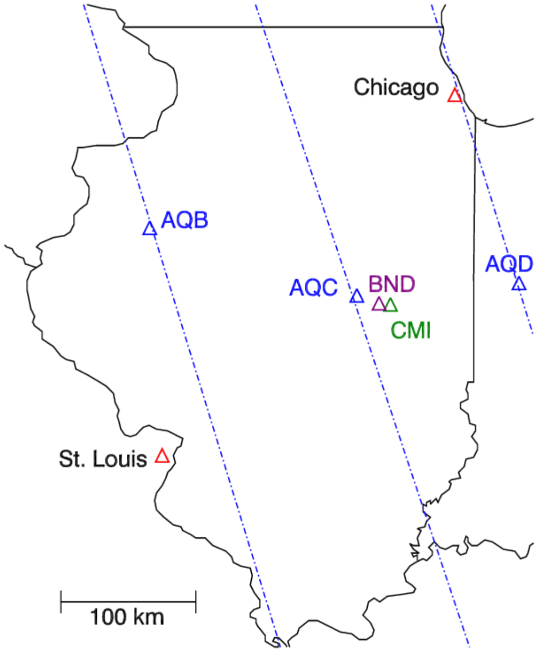
$\sigma_{sp(ambient)}/\sigma_{sp(dry)} = ((1-(RH_{ambient}/100))/(1-(RH_{dry}/100)))^{-\gamma}$

where RH_{dry} and $\sigma_{sp(dry)}$ = the RH and σ_{sp}, respectively, measured in the low-humidity (<40%) nephelometer.

An expanded discussion of how the adjusted scattering values were obtained is included in the revised manuscript in the Methods section.

4. (Table 3) It would be more helpful if there was a map that showed the locations listed in Table 3. This map could show these satellite tracks.

We originally decided that a table listing the profile locations was sufficient. But the referee's comment about adding the satellite tracks is reasonable. If the Editor agrees to the need for another figure, we have inserted a new figure into the revised manuscript showing the profile locations and recurring A-Train satellite tracks so the reader gets a better understanding where each type of profile was conducted.



5. (page 17201, line 15) Were there any systematic differences in the aerosol properties between the flights with and without satellite overpasses?

We made the same type of box/whiskers plots of the statistical aerosol distributions for all flights WITH and WITHOUT satellite overpasses. We did not put the statistics of these subsets in the paper because this issue was not a focus of our study and you have to draw the line somewhere, but Referee #1 has asked a legitimate question. One could guess that, since the satellite overflights were constrained operationally to occur on less cloudy days, the aerosol properties could be different during those drier conditions.

We found that the box/whiskers plots showed very little difference in both the extensive and intensive aerosol properties between subsets of flights conducted during and not concurrently with satellite overpasses. Median extinction coefficients and single-scattering albedos (at 550 nm) measured during the overpass flights at the 5 segment altitudes in the boundary layer all agreed to within 2 Mm^{-1} and 0.005, respectively, with the analogous measurements from the non-overpass flights. While this is admittedly a rough comparison it shows that there were not major differences between satellite and non-satellite flights. A brief discussion of satellite vs. non-satellite flights has been added in Section 3.3 in the revised manuscript.

6. (page 17202, line 7) What was the supermicron (large) particle response of the surface based in situ measurements? How much different was this response than the large particle response of the airborne instruments?

The submicrometer scattering fraction (R_{sp}) for the cumulative airborne measurements is shown in Figure 7h. The median values of R_{sp} for the different flight levels in the boundary layer are all between 0.85 and 0.90, with tails of the distributions extending to smaller R_{sp} values. At higher altitudes the distributions are broader and the median R_{sp} is smaller. Considering all AAO flight segments, the median R_{sp} is ~0.87. The surface-based R_{sp} values measured at BND during the flights showed a median value of ~0.84, which means that a slightly larger fraction of scattering by supermicron particles was observed at the surface compared with that measured by aircraft in the boundary layer. This issue was discussed in the manuscript (Page 17201, Line 21 to Page 17202, Line 15) and in Figures 5a and 5b. The comparisons in these figures are a bit different than the discussion above in that R_{sp} is not being directly compared but rather we are comparing the submicron and 'total' scattering coefficients between surface and lowest flight (flyby) level. The submicrometer measurements agree very well while the total scattering measurements (which presumably include a supermicron component) show less agreement, with the aircraft measurements showing lower total scattering coefficients than the surface measurements. In the manuscript we proposed that either some larger particles might have been excluded by the inlet or that there were real differences in the aerosol between the surface and the ~240m flyby altitude. It would be very difficult to know for sure which one of these influenced our measurements more.

7. (page 17204, line 20) In figure 7, dry scattering values were restricted to values above 3 Mm-1. How much better would the correlations in Figure 5 have been if this same restriction been used there also?

There would have been no difference. While we did not explicitly impose a 3 Mm^{-1} dry scattering threshold, the fact is that these were comparisons between surface and lowest flight level measurements, and all of the comparison periods showed dry scattering coefficients at 550 nm in excess of 3 Mm^{-1}.

8. (Page 17205, line 3 and Figure 7) The altitude dependence of the intensive parameters shown in Figure 7 are generally similar to that found in the Andrews et al. (2004, 2011) studies, with the exception of single scattering albedo. These earlier studies found that the single scattering albedo decreased with *altitude; however, this study over Bondville found single scattering albedo increased with altitude. Why the different behavior? Can this variation be attributed to uncertainties in measuring the low values of absorption at high altitudes?*

The cumulative profile median ω_0 in the Andrews et al (2004) study decreased with altitude throughout the profile, which is opposite to the small altitude dependence reported here. We need to point out, however, that the median ω_0 in this study (Fig. 7g) increased slightly only up to the top of the mixed layer and then remained more or less constant throughout the free troposphere up to the highest flight segment altitude. The altitude dependence of ω_0 over Oklahoma reported in the Andrews et al. (2011) paper was substantially less than for their 2004 paper. Figure 2d in Andrews et al. (2011) shows the plot of seasonal vertical profiles of median ω_0 for both sub-μm and total aerosol (which the paper called sub-7μm) size cuts. For winter for both size cuts, they observed a decrease in ω_0 with altitude, but for the sub-7μm data for the other seasons median ω_0 is either constant or shows a slight increase - much more like what we observed in this paper over BND.

Possible reasons for the differences are speculative. The Andrews et al. (2004) paper reported only submicrometer aerosol data, whereas their 2011 paper and this study include 'total' aerosol data in the statistics. Since there is a particle size dependence to aerosol light absorption, this could play a role in the discrepancy. Uncertainty at very low absorption levels could also play a role, especially if there is some bias at very low absorption levels. The differences between the median ω_0 measurements at low and high altitudes in this study (~0.025, or roughly a few percent relative difference) are certainly within the realm of measurement uncertainty given the estimated total measurement uncertainties in the nephelometer (~5-10%) and PSAP (~10-20%). Of course it is also possible that different meteorology and/or aerosols in the Oklahoma and Illinois regions could contribute to different aerosol optical properties. Whether we should even expect a similar altitude dependence of ω_0 over two different regions is difficult to know.

9. (page 17206, line 6) If the absorption lower limit threshold was raised to 1.0 Mm-1 instead of 0.2 Mm-1, how would the results in Figure 7 have changed?

We did not have an explicit absorption threshold of 0.2 Mm^{-1}. That was an approximate absorption threshold based on the dry scattering threshold of 3.0 Mm^{-1}. It is clear you need some type of threshold because with the low values you are trying to compare measurements that are dominated by noise. The question is where do you set the threshold for inclusion into the cumulative statistics? We chose a threshold of 3.0 Mm^{-1} dry scattering coefficient (@ 550 nm) because it eliminated the obviously noisy values while keeping the vast majority of measurements at all but the highest altitudes. If we had used an absorption coefficient threshold of 1.0 Mm^{-1} (@ 550 nm), that would imply an effective extinction threshold around 15 Mm^{-1}, and this would have eliminated essentially all of the free tropospheric data and a lot of the boundary layer data. If we did this the extensive aerosol optical property distributions would obviously be shifted to higher values. The intensive property distributions would reflect the aerosols present during higher aerosol conditions.

10. (page 17210, line 19) In making the adjustment to ambient relative humidity, was the instantaneous RH for each point, or was the mean or median RH used for each segment? Since the adjustment of scattering to ambient RH is nonlinear in RH, the procedure on how this is done is important.

We agree with this reviewer that the method used to adjust aerosol properties to ambient RH could under certain circumstances make a difference and needs to be clear to the reader. We did not use the instantaneous value of RH for the adjustment of scattering to ambient RH conditions… instead we used the segment average (i.e., mean) value. We felt this was acceptable for two reasons. First, the RH measured on most of our flight segments was not very high, and RH variability at lower RH does not cause large changes in the optical properties. Figure 7f shows that median RHs for flight levels in the mixed layer were 50-60% and were lower at higher altitudes. Only a small fraction of all flight segments show segment-average RH values at or above 70%, where the growth function becomes steeper. Secondly, RH variability was consistently small over the 5 and 10 minute horizontal segments flown over the course of the program. Considering all AAO flight segments, the average standard deviation (SD) of RH measurements during individual flight segments was 2% (with a SD of the SD values of +/-2%. That is, on average, each AAO flight segment showed a relatively small variability (expressed as the SD) of 2% RH. While there were a few flight segments conducted at higher RH that also showed a larger RH variability, this was not a commonly observed feature of the data set.

We have included an expanded discussion of how the adjustment of dry scattering values to ambient relative humidity was made in Section 3.2 in the revised manuscript.

11. (page 17211, line 13) What was the mean and median extinction measured by the AAO sensors during the 35 cases where there were no CALIPSO retrievals? This may help answer whether the extinction was too low for the CALIPSO retrievals.

We have re-analyzed the CALIPSO extinction data for the 63 A-Train overpass flights. For 28 of the overpasses, CALIPSO retrieved aerosol extinction at some altitude within the 35-km satellite track length considered appropriate for comparison with the aircraft (these are the flights reported in the paper), while for 35 of them it did not. Inspection of CALIPSO browse images (using the vertical feature mask) revealed that for some of these 35 ‘no retrieval’ cases, CALIPSO detected aerosols farther away from the aircraft profile location but not within the comparison window even for clear sky conditions. For other flights clouds attenuated the lidar and obscured the viewing beneath, so it is unclear whether detectable levels of aerosols were present in the boundary layer in these cases.

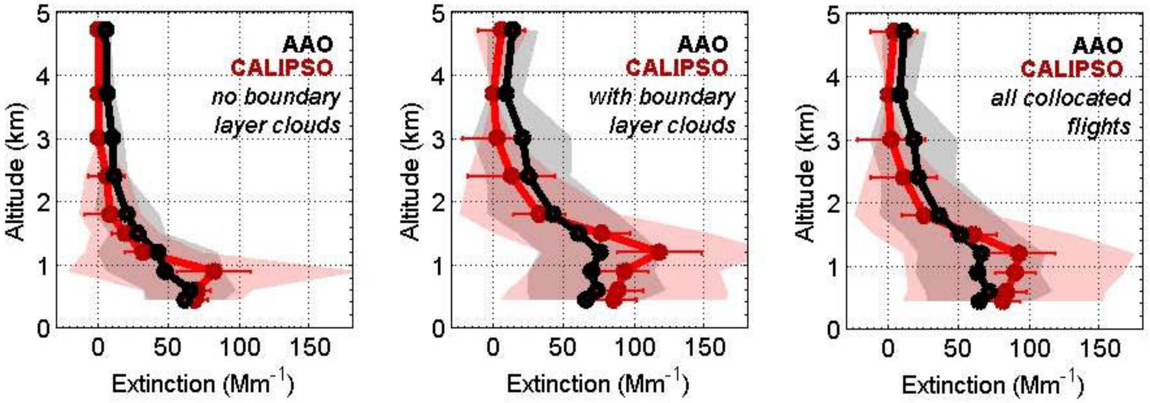
Looking at the AAO aircraft-measured extinction data, the 28 flights where CALIPSO also retrieved extinction showed in situ median aerosol extinction up to about 20% higher than the 35 flights where CALIPSO observed no aerosol. Medians of the distributions of 550 nm extinction for the 35 ‘no retrieval’ cases were typically in the 15-35 Mm^{-1} range for flight levels in the boundary layer, and Figure 13 suggests that CALIPSO would not have retrieved aerosol extinctions in this range a substantial fraction of the time. We have vertical profile plots of the AAO extinction medians that show this, but we do not feel that these add a lot to the paper. Instead we have included an expanded discussion, similar to that above, in the Section 3.3 explaining the differences in aircraft-measured extinction between the subset of 28 flights where aerosol extinction was observed by CALIPSO and the 35 collocated flights where it was not.

We hope to present details of the comparisons between satellite-derived and in situ extinction for specific case studies (i.e., individual profiles) in the next paper. That is where the causes for discrepancies between satellite and in situ measurements for specific profiles will be discussed in detail. The authors feel that the next paper would provide a better platform for understanding the discrepancies than this paper, which focused on the overall statistical agreement.

12. (page 17212, line 19) How many of these 28 AAO-CALIPSO cases had clouds? Do the AAO measurements keep a record of which of these cases had clouds? If so, why not do a separate comparison of only those cases that had no clouds as observed by the AAO in order reduce the impact of imperfect cloud masking?

The AAO data contain no direct information on clouds. The pilot was instructed to fly only if the forecast was for scattered clouds (cloud fraction <= 0.5). Unfortunately the forecast was not always correct and conditions sometimes deteriorated during the flight. If unbroken clouds were present at a given altitude range then those corresponding flight segments were skipped.

It was relatively simple to count the number of cases with BL clouds (detected at 1/3 km resolution) in the proximal satellite track as reckoned by CALIPSO. These are the cases that are most likely to cause cloud contamination in CALIPSO 20 km and 80 km horizontal averages. Of the 28 flights where CALIPSO detected aerosols, the vertical feature mask images showed 20 cases with and 8 cases without BL clouds. The figure below shows the mean extinction profiles for the 8 cases without BL clouds, the 20 cases with BL clouds, and all 28 cases (the original Fig. 12a). The mean extinction for the cases without BL clouds is smaller below ~3 km for both instruments compared to the cases with BL clouds. The agreement below 1 km is better between AAO and CALIPSO for the case without BL clouds, except for the spike in CALIPSO extinction population at 0.9 km. The extinction browse images show granules on two flights that could have caused this spike in the mean extinction profile, although we'll have to dig deeper to figure out the cause. We suspect unidentified BL clouds given the altitude, and we hope to look into this further for the next paper on AAO-CALIPSO case studies. It is interesting that the spike at 1.2 km in the case with BL clouds is reduced significantly in the case without BL clouds. This supports our hypothesis that cloud contamination could be occurring, particularly at that altitude. We have added this figure to the revised manuscript and a discussion of the effects of BL clouds in the profiles.



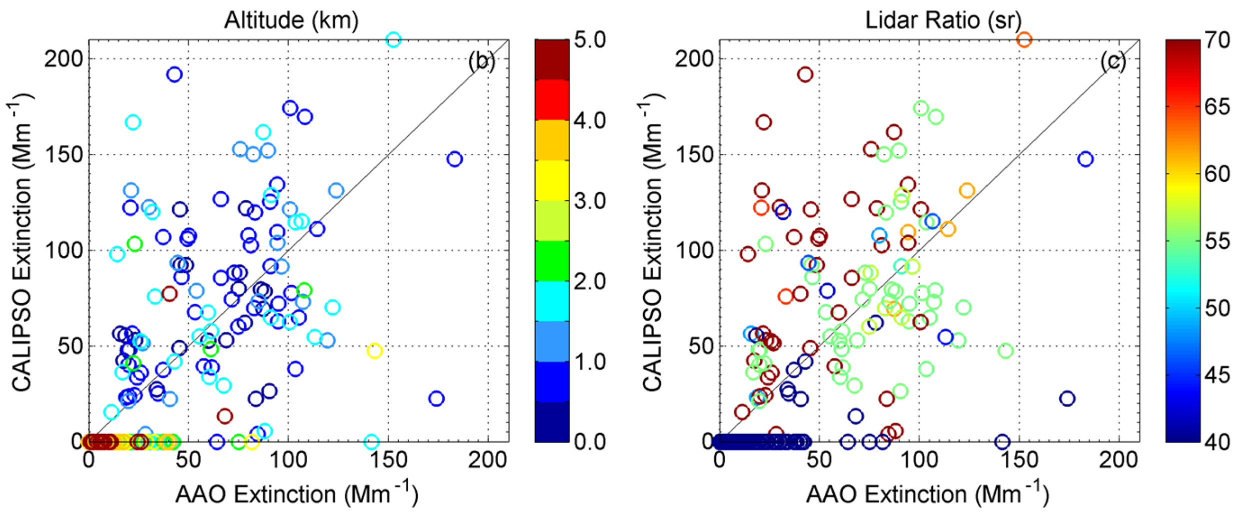
13. (page 17213, line 13) Instead of simply taking the closest 60 m range bin result from CALIPSO, why not have averaged the CALIPSO data over 120-180 m or so to avoid this? It seems strange to have a horizontal average that extends over 35 km, yet no vertical averaging to reduce the impacts of changes in the layer heights.

We initially analyzed the CALIPSO data both using some vertical averaging and also with no averaging, and our statistics did not change very much at all. We decided not to average vertically to avoid incorporating aerosol not observed by the aircraft into CALIPSO profile data. The reviewer has a good point that layer heights may change by 120 – 180 m over 35 km. Since the statistics did not change much we ended up inserting some text to that effect into the manuscript.

As an example of how prevalent this might be, we can look at Figure 13. For in situ extinction values above 50 Mm^{-1}, CALIPSO retrieved a valid extinction value ~95% of the time. The only bin where CALIPSO clearly missed a high-extinction layer was in the 140-150 Mm^{-1} bin, where it retrieved only 2 of the 3 segments that fell into that bin. Closer inspection of this one missed case revealed that CALIPSO retrieved an extinction value near what was measured by the aircraft in the adjacent vertical range bin. So in this one case, we can clearly say that vertical averaging of the CALIPSO data would have helped to detect a layer sampled by the aircraft. As shown in Figure 13, this does not appear to be a problem that was observed frequently. It is possible that CALIPSO also occasionally missed an extinction retrieval in the 60-70, 70-80, and 80-90 Mm^{-1} bins, but these are not clear cut cases and could be due to detection limit or cloud contamination issues as much as aerosol layers changing height.

14. (page 17213) The CALIPSO retrievals of aerosol extinction depend heavily on the aerosol type that was assigned to the CALIPSO measurements and the resulting lidar ratio used in the computations. It would be helpful to summarize what the aerosol type was chosen in these 28 cases. This could help determine if the CALIPSO bias may have been due in part to an inappropriate assignment of aerosol type and/or lidar ratio.

There is a population of large (>100Mm^{-1}) CALIPSO extinction values, typically at lower altitudes, that do not correspond with equally large AAO extinction values. These could represent the comparisons where the cloud screening process should have eliminated these satellite data from consideration but did not. Additionally, this population of large CALIPSO extinction values relative to AAO could result from either misidentification of aerosol type or incorrect assignment of lidar ratio, or both. The lidar ratio (ratio of extinction to backscatter) is required by the CALIPSO algorithms to retrieve extinction from backscatter and varies by aerosol type. We have made a figure of the AAO vs. CALIPSO extinction coded by lidar ratio. This figure could be added as another panel to Figure 12, as shown in the figure below. Many of these higher CALIPSO extinction values were derived using high lidar ratios (dark red circles in the idar ratio plot).



Aerosol types identified by CALIPSO are shown in Table 4 for the 28 collocated flights. Most aerosol samples in this regime (46%) are classified as polluted dust with a lidar ratio of 55 sr. However, 39% of samples are classified as either polluted continental or smoke aerosol types, both having lidar ratios of 70 sr. If this lidar ratio is too large for aerosol in these cases the retrieved extinction will be overestimated and could contribute to the large CALIPSO extinction values since errors in retrieved extinction are roughly proportional to errors in the assumed lidar ratio. Indeed, Anderson et al. (2000) measured lidar ratio values at Bondville of 64 ± 4 sr which, if representative of aerosol measured during these 28 flights, could account for some of the discrepancy in extinction below 1.5 km.

Table 4. Aerosol types identified by CALIPSO.

Aerosol Type	Lidar Ratio (sr)	Percent of Samples (%)
Dust	40	15
Polluted dust	55	46
Polluted continental	70	24
Smoke	70	15

With this in mind, the reviewer's comment certainly has merit that if the lidar ratio is too large it could add a systematic bias to the mean CALIPSO extinction. Aerosol samples classified as polluted dust and smoke tend to compare best with AAO extinction perhaps meaning a lidar ratio near 55 sr might be more appropriate for this region and this is consistent with the crude estimate above. We should always bear in mind that CALIPSO extinction biases due to lidar ratio can occur for two reasons, either the type was misclassified or the lidar ratio for the type is wrong.

Anderson, T. L., S. J. Masonis, D. S. Covert, R. J. Charlson, and M. J. Rood (2000), In situ measurement of the aerosol extinction-to-backscatter ratio at a polluted continental site, J. Geophys. Res., 105(D22), 26,907–26,915, doi:10.1029/2000JD900400.

This discussion, including the revised figure, table, and reference are included in the revised manuscript.

15. (page 17215, line 19) This sentence should be changed to reflect that the surface measurements could be used, on average, to estimate column RFE. However, an instantaneous retrieval of RFE using surface properties could be in error.

This sentence has been changed to reflect the reviewer's concern.

16. (page 17215, line 25) The behavior of the AAO-CALIPSO extinction comparison as a function of cloud cover was not shown in the paper. If the authors have indeed examined this behavior as a function of cloud cover to make this statement, why not show a graph that illustrates the behavior as a function of cloud amount? What was the behavior with no clouds present?

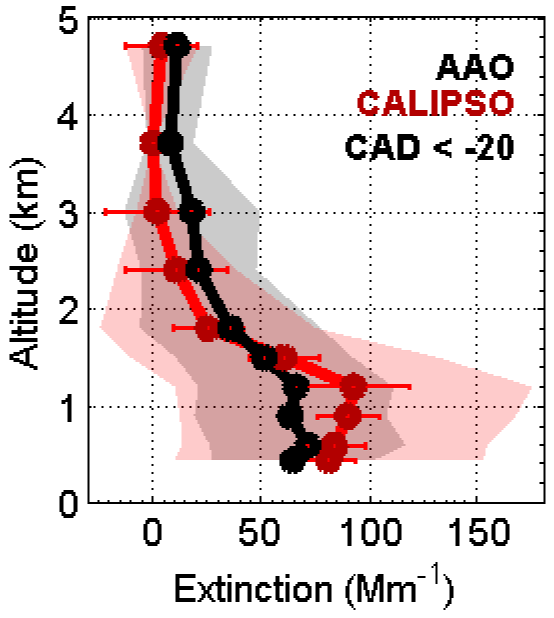
We did not specifically look at the agreement between CALIPSO and AAO extinction as a function of cloud cover; this is more of a feeling we got from working with the data. We propose to simply remove the line, "As clouds became thicker…the agreement generally worsened", and save that analysis for the next paper. Otherwise, we agree with the reviewer that a graph illustrating worse agreement is justified. As for the behavior with no clouds present, we could re-run the analysis to remove overlying clouds and boundary layer (1/3 km resolution) clouds though we think this would spell trouble. Removing layers with boundary layer clouds would likely wipe out almost all of our cases because these clouds are pervasive and the aerosol layers are mostly detected at 20 km to 80 km resolution, meaning if we removed aerosol layers containing boundary layer clouds we would remove 20 – 80 km of data we would otherwise use!

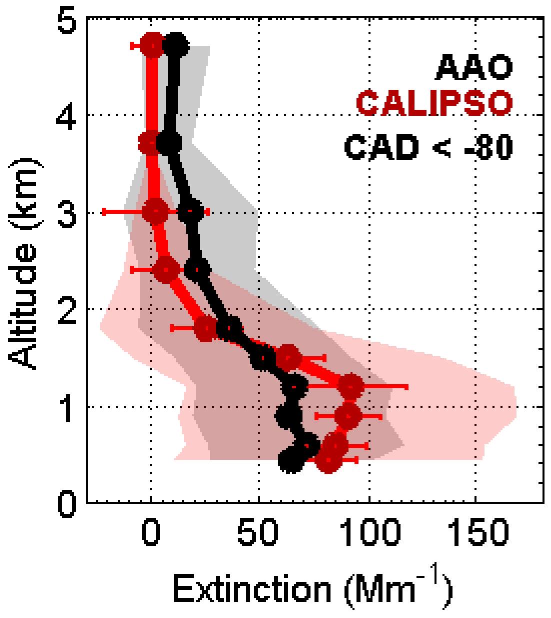
17. (page 17216, line 19) A CAD score of -20 seems a bit too high and may lead to cloud contamination. If the authors indicate that the differences between CALIPSO and AAO aerosol extinction measurements may have been due to cloud contamination, could they have repeated this using a lower CAD score? See the article by Yang et al. in Atmospheric Research (2012, vol. 116, 134-141) regarding the impact of the CALIPSO CAD score on observations of aerosol properties near clouds. This could have an impact on the level of agreement between the CALIPSO and AAO measurements.

The CAD score constraint is meant to remove cloud misclassifications and improve mean aerosol extinction quality. Based on our as-yet-unpublished analysis used to develop a CAD threshold for the level 3 aerosol profile product, we found that a more restrictive threshold (requiring CAD to be closer to -100) does little to reduce extinction uncertainty or improve mean aerosol extinction quality. We settled on -20 because it threw out most of the obvious misclassifications which occur mostly for $-5 < \mathrm{CAD} < 0$. A more restrictive threshold seemed to unnecessarily throw out “good” aerosol samples. Unfortunately this information is unpublished so we don’t have anything to cite in this paper.

Regarding the Yang et al. (2012) paper the reviewer mentions, they look at backscatter and color ratio near clouds as a function of aerosol CAD score. As you would expect, backscatter and color ratio are more cloud-like (larger) with CAD scores near 0 and more aerosol-like (smaller) with CAD score near -100. Their plots of backscatter and color ratio as a function of distance from cloud are very similar for the cases where all aerosol CAD scores are allowed and where CAD < -60, meaning that there isn’t much benefit from a less restrictive threshold than one around -60. More restrictive CAD scores reduces backscatter substantially which consequently reduces the aerosol extinction and at some point would bias the mean aerosol extinction low. This is why we don’t apply a very restrictive CAD score in the level 3 aerosol product. Still, the reviewer has a point that this could be related to the disparity between CALIPSO and AAO extinction if AAO extinction is always far from clouds and CALIPSO extinction is amongst the cloud field. However, setting a more restrictive CAD threshold does not always equate to CALIPSO extinction retrieved farther from clouds, especially when most of the layers we are dealing with are detected at 20 km and 80 km horizontal resolution.

We took a look at the impact of choosing a more restrictive CAD score on the mean extinction profile comparison. The two figures attached show the profiles where CAD < -20 (the way it is in the paper now) and with CAD < -80. When CAD is < -70, it is considered “high confidence aerosol”. There really is not a significant improvement in the agreement between the two instruments with a more restrictive CAD score and the altitude where we are concerned most with cloud contamination (~1.2 km) does not improve either. It looks like most of the aerosol layers we consider are high confidence aerosol and excluding those that have lower confidence does not change the mean extinction significantly.





18. (Figure 1 caption) Suggest changing to “Lettered areas are described in the text and in

Table 1”.

The authors concur with this recommendation and have made this change in the revised manuscript.

19. (Figure 5 and captions) It would be helpful if this caption indicated what the various solid lines represent.

The authors concur with this recommendation and have made this change in the revised manuscript.

Typographical comments:

1. (page 17197, line 14) “lose” should be “loose”

This has been corrected in the revised manuscript.