Short Comments by Mark Vaughan and Stuart Young

Comments: This paper compares the aerosol extinction profiles and aerosol optical depths (AODs) retrieved by the CATS lidar with similar quantities retrieved by AERONET, MODIS, and CALIOP. To our knowledge, this is the first ever in-depth assessment of satellite-derived AODs measured/retrieved in the near-infrared, and thus should be of great interest to several different groups in the aerosol research community. Overall, the authors have done a good job in bringing multiple analyses together. However, we find several places where additional analyses are warranted and where more in-depth discussions will help strengthen the final manuscript.

Response: We thank Mark Vaughan and Stuart Young and we appreciate the suggestions and comments which we believe are shaping this paper into a better paper

General Remarks

Comments: When filtering the extinction coefficients retrieved from the CATS and CALIOP measurements, the authors say that candidate extinction coefficients were constrained to a "nominal range" of 0 to 1.25 km–1, and that "all near zero negative values" are set to zero (page 6, lines 114–119). Presumably these "near zero negative values" that were set to zero were not actually removed from consideration, but instead were included in subsequent data averaging operations (the writing in this section is not sufficiently clear on this point). Changing negative values to zeros prior to averaging is not statistically valid, as it guarantees high biases in the estimated mean values. Reporting these high biases erroneously improves the comparisons of lidar-derived optical depths with those obtained by other sensors. To avoid this, all CATS and CALIOP mean values should be correctly recalculated before the final version of this manuscript is published.

Response: This is an excellent point. First, when calculating AOD and AOD climatology, we used the CATS and CALIOP derived AOD values. Thus, this is no need for us to detailing with negative extinction coefficients and we have revised the paper to reflect the issue. We did, however, apply the constraint that AOD values which came from retrievals containing extinction coefficients greater than 1.25 km–1 be excluded to avoid using AOD values from what are likely cloud contaminated profiles. We have added the following discussions.

"In AOD related studies, CAT and CALIOP reported AOD values are used. However, although not derived in this study, only AOD values with corresponding aerosol vertical extinction that meet the QA criteria as mentioned in Sections 2.1 and 2.2 were used."

Still, the extinction coefficients are used in estimating aerosol vertical distributions. As suggested we have revised our calculations and included those negative values, instead of setting them to zero. Note that similar approaches have been adopted for passive-based AOD studies, where *negative AODs are used to reduce high bias in long term studies (Remer et al., 2005). We have made the changes in the text accordingly.*

"Extinction was also constrained using a threshold as provided in the CATS data catalog (Extincton_Coefficient_1064_Fore_FOV <= 1.25 km^{-1}), similar to several previous studies (Redemann et al., 2012; Toth et al., 2016). Only profiles with extinction coefficient values less than 1.25 km^{-1} are included in this study. Small negative extinction coefficient values, however, are included in aerosol profile related analysis, to reduce potential high biases in computed mean profiles. Note that a similar approach has also be conducted in deriving passive-based AOD climatology (e.g. Remer et al., 2005)."

Remer, L.A., Y.J. Kaufman, D. Tanré, S. Mattoo, D.A. Chu, J.V. Martins, R. Li, C. Ichoku, R.C. Levy, R.G. Kleidman, T.F. Eck, E. Vermote, and B.N. Holben, 2005: The MODIS Aerosol Algorithm, Products, and Validation. J. Atmos. Sci., ***62****, 947–973, https://doi.org/10.1175/JAS3385.1*

Comment: While the main body of the text emphasizes the correlations between the CATS retrievals and the other data sets (e.g., lines 186–208), the authors do not provide any quantitative statements about the magnitudes of the CATS AODs or the differences between the different AOD estimates. Given that this paper is (to our knowledge) the first ever in-depth look at 1064 nm AOD, tables showing global and regional mean values and quantifying the CATS AOD estimates relative to the other sensors would add significantly to the value delivered by this paper. Profiles of the relative CATS-CALIOP extinction coefficient differences (i.e., (CATS(z) - CALIOP(z)) / CALIOP(z)) would be especially interesting.

Response: We have added a table to include regional and global means. Still, we documented that the differences may also be introduced by sampling differences of the sensors.

"Table 2. CALIOP and CATS mean aerosol optical depth for regions as highlighted in Figure 6 and globally between +/- 52° latitude."

Region	*Latitude*	*Longitude*	*Mean CATS AOD (DJFMAM/JJASON)*	*Mean CALIOP AOD (DJFMAM/JJASON)*
Global	*52°S-52°N*	*180°W-180°E*	*0.09/0.10*	*0.09/0.09*
India	*7.5°N - 32.5°N*	*65°E - 85°E*	*0.22/0.26*	*0.22 /0.28*
Africa - North	*2.5°N - 22.5°N*	*35°W - 20°E*	*0.26/0.23*	*0.30 /0.25*
Africa - South	*17.5°S - 2.5°N*	*0° - 30°E*	*0.14/0.22*	*0.15 /0.13*
Middle East	*12.5°N - 27.5°N*	*35°E - 50°E*	*0.22/0.33*	*0. 26/0.35*
China	*27.5°N - 37.5°N*	*110°E - 120°E*	*0.19/0.18*	*0.21 /0.16*

We have also added a plot of the difference (CATS(z)-CALIOP(z)) in Appendix A. As CALIOP extinction values become very small, the ratio of (CATS(z)-CALIOP(z))/CALIOP(z) has a tendency to grow very large from just a few data points and greatly impacts the standard deviation. Thus we plotted only the difference and did not include (CATS(z)-CALIOP(z))/CALIOP(z) and error bars with this particular plot.

Comment: In section 3.1.1., CATS observations are compared with other observations made within ±30 mins and ±0.4 degrees. For aerosols, this is probably not too much of a problem a lot of the time, but we have seen numerous cases where there can be large differences in the scenes being observed (e.g., see Omar et al., 2013: "In 45% of the coincident instances CALIOP and AERONET do not agree on the cloudiness of the scenes."). For AERONET, the comparisons may be improved by imposing another criterion, i.e., that the AERONET AODs made at the closest times preceding and following the CATS observations not vary by more than x%. A similar filter for potential spatial differences could include wind speed and direction (e.g., Lopes et al., 2013) and the spatial separations of the AERONET sites and the CATS observations. (This is likely to be quite a bit messier.)

Response: We have included the references as suggested and reminded readers that the collocation criteria may have impacts to the results due to the spatial and temporal sampling methods chosen.
"Note that as suggested by Omar et al., 2013, the choices of spatial and temporal collocation windows have an effect on collocation results. However, we consider this as a topic beyond the scope of this study"

Comment: While the authors point out a number of differences between the CATS retrievals and those derived from other sensors, they typically do not attempt to identify the causes of these differences. For example, based on the scaling factors in the linear regressions, the CATS AODs are lower than all of the AODs with which they are being compared (i.e., AERONET in Figure 1, MODIS in Figure 2, and CALIOP in Figure 3). This is perhaps not surprising for the AERONET and MODIS comparisons, but the cause for the CATS-CALIOP differences is not as obvious. Differences between CALIOP and MODIS at visible wavelengths are frequently explained by CALIOP's low daytime detection sensitivity and the missed detection of some of the vertical extent of the aerosol layer (e.g., Kim et al., 2017 and Toth et al., 2018). Can the authors enumerate the possible causes that would explain the disparities between CATS and CALIOP?

Response: *Slopes in linear regressions can often be biased by outliers. In Figure 6, which are spatial plots of AODs from CALIOP and CATS, differences are less noticeable for the DJFMAM season. For the JJASON season, CATS AODs are lower at certain regions (Middle East, India, and North Africa) and higher over other regions (South Africa). The cause of those discrepancies, however, is unclear to us. Also, Version 2 of the CATS data are used in this study, and we expect some difference with the version 3 of CATS data. To really explore the issue, it deserves a paper of its own. Thus, we leave this topic to a future paper.*

Comment: Furthermore, given the lower CATS AODs shown in Figure 2, it's surprising to see that the CATS extinctions coefficients shown in Figure 5 are typically larger than CALIOP at all altitudes, and that the closest agreement is over land (where CATS slightly underestimates CALIOP at lower altitudes). Again, some discussion of the possible causes of this paradox would be welcome.

Response: First, there is a call from the community to avoid using slopes from the regression analysis as they are prone to noisy data, and we are kind of agree with them. Statistically, we expect a high percentage of small AODs versus large AODs. Still, slopes are dominated by high AOD cases, while the averaged profiles may be more dominated by low AOD cases. This could explain the difference.

Comment: The results shown in Figure 5 are a prime candidate for further investigation into the underlying causes of the differences. Except for the over land case, CATS extinction profiles consistently and significantly overestimate CALIOP extinction profiles. It seems that there are four likely suspects in causing this (always keeping in mind that that all four could be collaborating in various nefarious ways to bring this about): layer detection, cloud-aerosol discrimination (including inadequate boundary layer cloud clearing), lidar ratio selection, and calibration. Of these four, the easiest to investigate (at least at a superficial level) is lidar ratio. The table below shows the default lidar ratios assigned by each instrument.

Aerosol Type	CATS	CALIOP
Dust	40 sr	44 sr
Dust mixture (a)	40 sr	N/A
Polluted dust (a)	N/A	48 sr
Dusty marine (a)	N/A	37 sr
Marine	25 sr	23 sr
Clean/background	35 sr	30 sr
Polluted continental	35 sr	30 sr
Smoke	40 sr	30 sr
Volcanic (b)	35 sr	44 sr

a) CATS identified dust mixtures over land and water; CALIOP identifies 'polluted dust' over land only and 'dusty marine' over water only.

(b) For CATS, all aerosol above 10 km is classified as volcanic. For CALIOP, volcanic aerosol is identified in the stratosphere only.

Since the CATS marine lidar ratio is large than the CALIOP marine lidar ratio, and the CATS dust mixture lidar ratio is larger than the CALIOP dusty marine lidar ratio (and CATS smoke and polluted continental lidar ratios are greater than their CALIOP counterparts as well), then, *all other things being equal*, one should expect the CATS over-ocean extinction profiles to be uniformly larger than the CALIOP extinction profiles. (But are all other things actually equal?)

The case is less clear over land. But since the CATS dust lidar ratio is less than the CALIOP dust lidar ratio and the CATS dust mixture lidar ratio is less than the CALIOP polluted dust lidar ratio, if we assume that the over-land aerosols detected in this study are dominated by dust (which might not be a bad assumption?), then perhaps the over-land profile comparison makes sense too. (All other things being equal, that is…)

Response:
We have added a discussion of potential sources of CATS-CALIOP extinction and AOD differences in the text:

"In addition, due to the precessing orbit of the ISS, the CATS sampling is irregular and very different compared to the sun-synchronous orbits of the A-Train sensors. These orbital differences between CATS and CALIOP make comparing the data from these two sensors challenging since they are fundamentally observing different locations of the Earth at different times. Thus, we shouldn't expect the extinction profiles and AOD from these two sensors to completely agree. Additionally, there are other algorithm and instrument differences that can lead to differences in extinction coefficients and AOD. Over land where dust is the dominant aerosol type, differences in lidar ratios between the two retrieval algorithms (CATS uses 40 sr while CALIOP uses 44 sr), can cause CATS extinction coefficients that are up to 10% lower than CALIOP, potentially explaining the higher CALIOP extinction values in Figure 5e. Over ocean, especially during daytime, differences in CATS and CALIOP lidar ratios for marine and smoke aerosols, as well as issues with CATS cloud-aerosol discrimination in V2-01 for daytime observations, can cause CATS extinction coefficients that are as much as 25% higher than CALIOP (Figure 5b and 5d). Yorks et al. (2019) shows examples of these daytime cloud-aerosol discrimination issues in V2-01 data, which have been improved for CATS V3-00 data. A brief analysis using 3 months of CATS V3-00 data showed improvement in agreement for AOD, but some differences were still evident in the extinction vertical profiles. These remaining differences, as well as the differences observed in nighttime only profiles (Figure 5c) are likely attributed to differences in CATS and CALIOP 1064 nm backscatter calibration. Pauly et al. (2019) reports that CATS attenuated total backscatter is about 18% higher than CALIOP due to calibration uncertainties for both sensors."

Comment: The CATS extinction profiles shown in Figures 5 and 10 peak at altitudes some hundreds of meters higher than do CALIOP's, except over land. While CALIOP's profiles show almost no roll off until about the last range bin above the surface, the CATS profiles start dropping off below about 500 m, or at approximately 8 to 9 range bins above the surface. What is happening here? Is CATS altitude registration and/or surface detection the culprit? Or is the cloud filter too aggressive in the boundary layer (i.e., are strongly scattering aerosols being misclassified as clouds)? Irrespective of the underlying cause(s), is this behavior a major source of AOD differences between CATS and CALIOP?

Response: The 2 biggest issues in the CATS V2-01 data were the daytime calibration and the daytime cloud-aerosol discrimination. A CATS paper in preparation (Yorks et al., 2019) has included details about the cloud-aerosol discrimination issues, while Rebecca Pauly's 1064 nm calibration paper has a lot of details about the new daytime calibration. We have checked this *issue by reprocessing the analysis using 3 months of V3 data and we found an improvement in agreement for AOD, but with some differences still evident in the vertical profiles.*

Comment: The seasonal maps (Figure 6) show that the CALIOP AODs exceed those of CATS over the Arabian Peninsula, and to a smaller degree over the African region bordering the Gulf of Guinea. Can this also be explained by differences in lidar ratio selection, or are there other factors at work?

Response: We suspect the difference in retrieval method as mentioned above may contribute. Also, CALIOP provides early morning and afternoon overpasses while CATS can observe at near all solar hours, the differences may also be associated with these sampling differences.

Specific Comments

Comment: page 4, line 85: provide a reference for "Feature Type Score"

Response: We have added the reference to the text.

Comment: page 5, line 107: did the authors also consider potential sources of bias errors; e.g., unusually large or small calibration coefficients, or large values of overlying integrated attenuated backscatter?

Response: We have adopted the QA steps from a few previous papers such as Campbell et al., 2012; Toth et al., 2016; 2018. The thresholds for the above mentioned criteria are not mentioned and used in those previous papers, and thus we didn't include the check as suggested.

Comment: page 5, line 113: "Extinction_Coefficient_Uncertainty_1064_Fore_FOV ≤ 10 km-1"; despite the heritage from Campbell et al. (2012), using relative uncertainties still makes much, much more sense. Given the noise in the CATS daytime measurements, an uncertainty threshold of 10 km-1 might be reasonable for an estimated extinction coefficient of 1 km-1. However, for the substantially smaller extinction coefficients (e.g., 0.01 km-1 to 0.1 km-1) that make up a very large majority of the measurements, an uncertainty threshold of 10 km-1 seems prohibitively large.

Response: Agreed. Since we have to apply the thresholds to all observations, lowering the threshold may exclude heavy plumes that may indeed be valid. Also, other QA steps, along with this threshold are also used, as thus, we expect some of the issues as mentioned can be captured by other QA steps. Thus, the QA steps remain unchanged.

Comment: page 6, line 128: distinguish between laser spot size (~70 m) and receiver footprint diameter at the Earth's surface (~90 m).

Response: We have changed the sentence to "with a laser spot size of around 70 m"

Comment: page 6, line 129: say which version of the CALIPSO data products was used (version 4.1, right?)

Response: We have included "CALIOP Level 2.0 Version 4.1" in the sentence.

Comment: page 7, line 137: "signal-to-noise", not "single to noise"

Response: Done.

Comment: page 7, line 148: "Atmospheric_Volume_Description = 3 (aerosol only)"; note that in the CALIPSO version 4.1 data products, 3 indicates tropospheric aerosols and 4 indicates stratospheric aerosols. Were stratospheric aerosols excluded accidentally or deliberately? (Previous versions of the CALIPSO data products did not differentiate between tropospheric and stratospheric aerosols. In these earlier products, requiring the atmospheric volume description to equal 3 would correctly identify all aerosol data.) If accidentally, please correct the calculations. If deliberately, please explain why.

Response: We have updated this to include Atmospheric_Volume_Description = 4 as well, and updated the text accordingly.

"Atmospheric_Volume_Description = 3 or 4 (aerosol only)"

Comment: page 8, line 163: logarithmic interpolation, correct? Also, please state the actual value of the Ångström exponent given by Shi et al.

Response: Yes. The Angstrom exponent value is computed for each AOD retrieval. We have revised the discussion to avoid confusion. "Here we assume the angstrom exponent value, computed using instantaneous AOD retrievals at the 860 and 1240 nm, remains the same for the 860 to 1064 nm wavelength range, similar to what has been suggested by Shi et al., (2011; 2013)."

Comment: page 8, line 170: while "AERONET data are considered as the ground truth for evaluating CATS retrievals", it should be noted that there are very few AERONET sites in remote oceans. Do MODIS retrievals substitute as the gold standard in these places?

Response: Even though a better performance can be expected from MODIS aerosol retrievals over ocean versus over land, we still think that only AERONET data should be used for ground truth, as instantaneous retrievals from passive sensors suffer from various issues such as cloud contamination.

Comment: page 9, line 186–187: some discussion on the rationale for the choices of ±0.4° and ±30 minutes would be helpful in evaluating the strength of the comparisons.

Reponses: We picked this threshold following a few previous papers (e.g. Toth et al., 2018). We have added discussions in the text to further clear this issue:

"*Note that as suggested by Omar et al., 2013, the choices of spatial and temporal collocation windows have an effect on collocation results. However, we consider this as a topic beyond the scope of this study*"

Comment: page 9, line 193: how frequently do "profiles with all retrieval fill values" occur in the CATS data set?

Response: We have examined the dataset and found that profiles in which there were no cloud or aerosol made up about 5.4% (3583933/65792363) of all profiles. The text has been updated accordingly.

"Such profiles containing all retrieval fill values were found to make up approximately 5.4% of all CATS profiles in the dataset."

Comment: page 9, line 194: as a rule of thumb, how close to sunrise and sunset can reliable AERONET measurements be obtained?

Response: We are not aware if any study have been conducted on this issue. Because it is hard to "validate" AERONET observations. But it is an interesting topic for a future paper.

Comment: page 11, line 244: The authors say, "using over land (ocean) daytime data only, for a total of 171 (1207) collocated pairs." Here we echo the remarks of an anonymous reviewer commenting on a paper for which one of us (Mark Vaughan) is a coauthor (see https://doi.org/10.5194/acp-2018-1090-RC1).
Way back in 2010 Prof. Robock pleaded with us to end this misuse of parentheses [Robock, A. (2010), Parentheses are (are not) for references and clarification (saving space), Eos Trans. AGU, 91(45), 419–419, doi:10.1029/2010EO450004]. My understanding is that one of the publishers in our field has specifically written it out of their style guide. I read pretty widely and the only genre of writing where I have experienced this application of parentheses is in the atmospheric sciences journals. I hope the authors will consider rewriting this sentence.

Response: Done. We have rewritten the sentence.

Comment: page 11, line 245: The authors say, "daytime data from both CALIOP and CATS are expected to be nosier due to solar contamination". While this is true, the day-night differences at 1064 nm are very different for the two lidars. CATS daytime SNR is substantially worse than CATS nighttime SNR, whereas CALIOP daytime SNR is only marginally worse than CALIOP nighttime SNR. The primary reason for this is that CALIOP 1064 nm detector is an avalanche photodiode for which the dark counts contribute substantial amounts of noise irrespective of the external lighting conditions. Moreover, while CATS 1064 nm nighttime SNR is much higher than CALIOP 1064 nm nighttime SNR, for daytime measurements the CALIOP SNR is higher. This should be explained in greater detail in a forthcoming CATS calibration paper.

Response: Great comment. But we think those comments should be included in a future paper, hopefully written by one of the coauthors.

Comment: page 12, line 260: "it is speculated". Who's doing this speculating? If it's the authors, then come right out and say so!

Response: We have revised the sentence to "although we speculate"

Comment: page 14, line 311: The authors say, "the shapes of the CATS and the CALIOP nm extinction vertical profile are very similar for all three cases". This qualitative assessment would be much more meaningful if it was augmented by a set of quantitative metrics (e.g., profiles of (CATS(z) - CALIOP(z)) / CALIOP(z), with error bars to indicate the magnitude of the variability in the ratios).

Response: We have included a plot of CATS(z) – CALIOP(z) (Appendix A) for the mean CATS and CALIOP vertical profiles. As CALIOP extinction values become very small, the ratio of (CATS(z)-CALIOP(z))/CALIOP(z) has a tendency to grow very large from just a few data points and greatly impacts the standard deviation. Thus we plotted only the difference and did not include (CATS(z)-CALIOP(z))/CALIOP(z) with error bars with this particular plot.

Comment: page 18, line 405: The authors say, "nighttime retrievals from CATS *are considered to be less noisy* than daytime" (emphasis added). This sentence suggests that there might be some debate about day versus night noise magnitudes. There is no such debate. The fact is that "nighttime retrievals from CATS *are significantly and demonstrably less noisy* than daytime retrievals".

Response: We have used the wording as suggested. Thanks for the comment.

Comment: page 23, lines 514–517: The authors' conclusions reinforce the conventional wisdom. However, we think it's important to emphasize that at present these conclusions are highly tentative, and will remain so until a comprehensive analysis of the CATS calibration accuracy and stability is completed.

Response: We have added the comment as suggested:
"Still, at present these conclusions are tentative, and will remain so until a comprehensive analysis of the CATS calibration accuracy and stability is completed. "