Some comments on “Investigation of CATS aerosol products and application toward global diurnal variation of aerosols” by Logan Lee, Jianglong Zhang, Jeffrey S. Reid, and John E. Yorks

Mark Vaughan (mark.a.vaughan@nasa.gov) and Stuart Young (stuart.young01@gmail.com)

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.

General Remarks

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.

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.

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.)

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?

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.

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…)

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?

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?

Specific Comments

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

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?

page 5, line 113: "Extinction_Coefficient_Uncertainty_1064_Fore_FOV $\leq$ 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.

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

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

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

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.

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

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?

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.

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

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

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.

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.

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

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).

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”.

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.

References

Kim, M.-H. et al., 2017: Quantifying the low bias of CALIPSO’s column aerosol optical depth due to undetected aerosol layers, *J. Geophys. Res. Atmos.*, **122**, 1098–1113, doi:10.1002/2016JD025797.

Lopes, F. et al., 2013: Evaluating CALIPSO’s 532nm lidar ratio selection algorithm using AERONET sun photometers in Brazil, *Atmos. Meas. Tech.*, **6**, 3281–3299, doi:10.5194/amt-6-3281-2013.

Omar, A. et al., 2013: CALIOP and AERONET aerosol optical depth comparisons: One size fits none, *J. Geophys. Res. Atmos.*, **118**, 4748–4766, doi:10.1002/jgrd.50330.

Toth, T. D. et al., 2018: Minimum Aerosol Layer Detection Sensitivities and their Subsequent Impacts on Aerosol Optical Thickness Retrievals in CALIPSO Level 2 Data Products, *Atmos. Meas. Tech.*, **11**, 499–514, doi:10.5194/amt-11-499-2018.