Response to Anonymous Referee #3

We thank the referee for the many useful comments, questions, suggestions, and typographical corrections. They have helped improve the clarity of the manuscript. Referee comments and original text are shown in black font below. Reponses to referee comments and revised text are shown in red font. Page numbers and line numbers refer to the original manuscript.

The authors describe the Level 3 CALIOP monthly mean AOD and extinction profile products, their quality screening and averaging methods. This paper is of excellent quality and worthy of publication. It is very well written and very well structured. Major comments: Any reference for these statements and assumptions would help the reader: . "The underlying assumption is that all aerosol layers below 250 m are in reality attached to the surface" . ". . . rarity of tropospheric aerosol detection above 12 km" . ". . . altitudes where shallow convective clouds are expected: above 8 km at the equator and lower towards the poles" . ". . . lidar ratio of a dust/marine mixture, which would fall in the range 20sr-40sr" Please consider adding a Table in section 5 that summarizes all the quality filters that were applied to the level 2 data (i.e., iso80km, CAD, cirrusFringe, extQC, extUnc, NSA filter).

The requested references have been added and are discussed in the detailed responses below. We agree that adding a table to summarize the quality filters is useful for readers, given that several other publications which apply quality filtering to CALIOP data do the same. The table has been added to the end of Sect. 3 (replicated below). We chose this location because it is necessary to list items discussed in Sect. 4 to fully describe methods used to generate level 3 output.

Table 1 is given here as a high-level summary of the averaging methods and quality filtering procedures detailed in the following two sections.

Table 1. Summary of averaging methods and quality filtering procedures used to generate the version 3 level 3 aerosol product. Details are discussed in the indicated sections. AGL and AMSL indicate "above ground level" and "above mean sea level", respectively.

Averaging method / quality filtering procedure	*Section*
Aerosol extinction for "clear-air" assigned ≡ 0 km^{-1}	*4.1*
Clear-air below aerosol layers with bases < 250 m AGL ignored	*4.2*
Isolated 80 km horizontal resolution aerosol layers rejected	*5.1*
CAD score outside [−100, −20] range rejected	*5.2.1*
Aerosol in contact with ice clouds (top temperature < 0° C) above 4 km AMSL rejected	*5.2.2*
Extinction QC flag ≠ 0, 1, 16, 18 rejected	*5.3.1*
Extinction uncertainty = 99.9 km^{-1} rejected, and all extinction below	*5.3.2*
All samples ≤ 60 m AGL excluded	*5.4*

It is not clear why the NSA filter is not part of Fig. 8 and 9

The NSA filter ignores all information within 60 m of the surface so any aerosol samples are not eligible for rejection. Hence, there is no frequency of rejection to report in Fig. 8 or 9.

Fig. 18 b is not explained in the text. What is its purpose? Similar comment for Fig. 20b.

Figure 18b is explained on P24 L13:

> *Figure 18(b) shows three aerosol extinction profiles retrieved from the attenuated backscatter in Fig 18(a). While the strongly negative values adjacent to the surface are readily apparent in this example, positive σ values that are biased low are not as easy to detect.*

Figure 20b shows the number of unfiltered aerosol samples. It is shown to provide context for the number of samples contributing to the other three panels in the figure. The following sentence in red was added to the manuscript on P26 L17.

> *Figure 20 summarizes the impact of quality filters on the global $\bar{\sigma}$ profile, the frequency of rejection, and filter aggressiveness. … For context, Fig. 20(b) shows the number of unfiltered aerosol samples, which decreases with increasing altitude.*

Consider giving two different names for variable "Agr" in Eq. A1 and Eq. A3.

Thank you for the suggestion. We find the two variations, $Agr(z)$ and $\boldsymbol{Agr}$ (now Eq. B1 and Eq. B3), work well in terms of distinguishability and consistency so we will retain the current names.

Detailed comments:

P4 L23: "The 12 km upper limit was selected due to the rarity of tropospheric aerosol detection above 12 km in the level 2 product". Would it be possible to quantify that statement? CALIOP Level 3 would then be missing any aerosol event above 12km?

The following statement in red was added to quantify the statement. Yes, CALIOP level 3 will miss these high-altitude events, but this is acceptable because the emphasis for this product is the lower troposphere.

> *The 12 km upper limit was selected due to the rarity of tropospheric aerosol detection above 12 km in the level 2 product (e.g., 0.04 % of tropospheric aerosol layers detected by CALIOP version 3 are above 12 km in 2010).*

P5 L2: consider adding "at night" to "Figure 2 depicts these sky conditions for an individual level 2 granule" if that is the case (and consider adding that info in the legend of Fig. 2)

The "nighttime granule" qualifier is now added to the Fig. 2 caption, but not to the main text because the day/night detail is not pertinent to the describing how the figure conveys the sky conditions.

P6 L3: MODIS needs to be defined

Done

P6 L13: consider adding “clouds” to “desert and snow”

Added the text in red:

> *In daytime, the signal-to-noise ratio (SNR) is lower relative to night, particularly over high albedo surfaces such as desert or snow or over clouds* (Hunt et al., 2009).

P7 L16: this is if you consider well-mixed aerosols in an atmospheric column of 1km. Is that, right? If so, you might want to add this information.

It assumes well-mixed aerosols in an atmospheric column of 10 km. The following statement in red was added:

> *Clarke and Kapustin (2002), for example, show background aerosol extinction levels of 10^{-4} km^{-1} to 10^{-3} km^{-1} in remote parts of the Pacific basin, implying a missing AOD ranging from 10^{-3} to 10^{-2} in the cleanest regions (assuming well-mixed aerosols in a 10 km deep column).*

P8 L10: “The underlying assumption is that all aerosol layers below 250 m are in reality attached to the surface”. Is this assumption based on any observation? Any reference here, even case studies, would be helpful.

The underlying assumption is that aerosol is well mixed within the planetary boundary layer at these altitudes. The sentence was rephrased to better express this assumption. Two references are given as examples which retrieved planetary boundary layer (PBL) heights from CALIOP data. Figure 3 of McGrath-Spangler and Denning (2013) shows maps of seasonal mean daytime PBL depths which are primarily above 500 m. Figures 3 and 7 of Luo et al., (2014) show PBL heights (labelled as BLH) for marine and continental sites also above 250 m.

> *The underlying assumption is that the atmosphere is well mixed below 250 m. Turbulent mixing within the daytime boundary layer tends to homogenize aerosol loading, and the planetary boundary layer is generally much deeper than 250 m for marine and continental conditions (e.g., McGrath-Spangler and Denning (2013); Luo et al. (2014)).*

P8 L14: “Consequently, global mean level 3 AOD is increased by a small amount, roughly 1 %.” This applies to this one case over the Arabian sea. Do you have more global statistics?

The AOD increase of 1 % is the global mean value, not just for the Arabian Sea case.

P11 L2: “sky” is repeated twice.

Corrected

P11 L28-29: “Corresponding seasonal totals and averages, defined as December–February (DJF), March–May (MAM), June–August (JJA) and September–November (SON), are also reported in supplementary material.” Which figure(s) does this refer to?

These figures are not included in the supplementary material. This line has been removed in the revised manuscript. Seasonal acronym definitions now appear as they are used.

Fig. 6 and 7: Are these gridded maps? 2◦ x 5◦? Consider adding this information in the legend.

These are maps of statistics reported by the level 3 product so they are 2 x 5 degree resolution. The legends for Figs 6-8 and Figs S1- S4 now include “as reported by the level 3 product” to make this clear.

Fig. 8: Do “All filters” include the “NSA” filter as well (not shown in Fig. 8 bur in Fig. 19 instead)?

Yes, the NSA filter is applied to all figures except in Fig. 19 which demonstrates the effect. The NSA filter is different than other filters because no samples are rejected, they are just ignored within 60 m of the surface at all times.

Fig. 8: consider adding the total percentage of aerosol samples rejected by each filter in the title of each of the 6 graphs

This is a very good idea. The percentages have been added to Figs. 8 and S3.

P15 L4: “altitudes where shallow convective clouds are expected: above 8 km at the equator and lower towards the poles”. Any reference here?

Added the following reference which shows typical altitudes of deep convection based on combined CloudSat/CALIPSO data. Also, replaced “shallow convective clouds” with “deep convective clouds” because that is the cloud type being discussed.

> *Mace, G.G. and F.J. Wrenn, 2013: Evaluation of the Hydrometeor Layers in the East and West Pacific within ISCCP Cloud-Top Pressure–Optical Depth Bins Using Merged CloudSat and CALIPSO Data. J. Climate, 26, 9429–9444, https://doi.org/10.1175/JCLI-D-12-00207.1*

P15 L5: delete “at”

Done

P15 L6: Fig. S4b instead of Fig. S4a

Corrected

P15 L30: "because there is no confidence in"

Corrected

P16 L1-3: "Low CAD scores also indicate a high probability of layer detection artifacts where noise spikes cause the feature finder to detect layers that do not actually exist." However, the authors select data with CAD scores between -100 and -20. These are "low" CAD scores. Consider rephrasing.

Indeed, the sentence is discussing no-confidence CAD scores rather than low-confidence CAD scores. The following changes were made to clarify.

> *A CAD score of −100 indicates that the feature is very likely an aerosol layer, and a CAD score of +100 indicates that the feature is very likely a cloud. There is no confidence in cloud-aerosol discrimination for features with |CAD score| < 20. For the year 2010 at night in version 3, over 85 % of aerosol layers have CAD score < −90 and around 4 % have CAD score > −20. The remaining 11 % have intermediate levels of confidence.*
>
> *Aerosol layers having CAD scores outside the range of [−100, −20] are rejected because there is no confidence in discriminating aerosol from cloud…~~Low~~ No-confidence CAD scores also indicate a high probability of layer detection artifacts where noise spikes cause the feature finder to detect layers that do not actually exist.*

Fig. 10: Is this figure global? If not, you might want to provide the latitude/ longitude range

Amended the figure caption:

> *Figure 10. Median overlying integrated attenuated backscatter (IAB) for aerosol layers having the indicated CAD score for 2010 at night, global.*

P17 L22: "causing aerosol subtyping misclassifications". Is this the case for the aerosols near the surface in Fig.12? If so, are those corrected? It's not clear from the text.

Yes, this explains why the marine aerosol is misclassified as dust near the surface in Fig. 12. These misclassifications are not corrected because the cirrus fringe filter is only applied above 4 km. The text has been revised as shown below to clarify.

> *The aerosol classified as dust within the marine boundary layer (albeit infrequently, < 0.3 %) is likely associated with residual cloud layers detected at 1/3 km resolution affecting δ_v, causing aerosol subtyping misclassifications. However, the enhanced frequency of dust detection at higher altitudes is the main issue addressed by this filter: when the cirrus fringe filter is not applied (blue profile), the peak altitude of dust frequency appears at nearly 7 km.*

P19 L3: "and having a cloud top temperature less than 0◦C"

Corrected

P19 L11: consider adding "in this case" to "The reduction in full column dust AOD. . ."

Added

P19 L13: consider adding the dust AOD reduction in % in the text (i.e., values on Fig. 14)

The percent dust AOD reduction for Fig. 14 is given in the text on P19 L11:

> *"The reduction in full column dust AOD is small, about 7 %)."*

P21 L3: "These are the least frequent of all solutions for aerosol layers, however" Can this be quantified?

The following text in red was added with the percentage given in Table 1: "*These are the least frequent of all solutions for aerosol layers, however (~0.01 % of all retrievals).*"

P21 L28: ". . . lidar ratio of a dust/marine mixture, which would fall in the range . . ." Consider adding references for typical lidar ratios.

Added a citation to Omar et al. 2018 which discusses the CALIOP lidar ratio used for dust/marine mixtures and provides references to relevant CALIOP papers discussing typical lidar ratio values.

P23 L17: consider rephrasing this way: ". . . profile shape the least while still rejecting solutions that are untrustworthy.

Done. Good call.

P23 L24: "while having only a small impact on mean AOD" Consider quantifying this statement

The following in red was added: *"…while having only a small impact on global mean AOD (a reduction of ~10 %)."*

P24 L12: "or" is repeated twice

Corrected

P24 L14: What do the authors mean by "this" example? Fig 18a?

It refers to Fig. 18(b). The text was modified to clarify that the example is the extinction profiles (not the attenuated backscatter): *"While the strongly negative values adjacent to the surface are readily apparent for the extinction profiles in this example,…"*

Fig 18a is missing its color scale

The color scale is now added.

Fig. 18b is not explained in the text (and Profile #1-3 are not described). What is its purpose?

The description of this figure is revised as shown below. The extinction profiles show how the NSA in level 1B attenuated backscatter affects level 2 aerosol extinction retrievals by causing a strongly negative value adjacent to the surface. These are three extinction profiles

> *Figure 18(b) shows three aerosol extinction profiles retrieved from the attenuated backscatter in Fig 18(a) along separate 5 km segments containing the NSA. While the strongly negative values adjacent to the surface are readily apparent for the extinction profiles in this example, positive σ values that are biased low are not as easy to detect.*

Fig 19b: Consider adding AOD in the title.

The title has been re-labeled as "AOD ratio (filter/no filter)"

P25 L6: "AOD increases by 5–10 % in regions affected by the NSA". The red box in 19b seems to show values between ~0.8 and ~1.2? The authors might refer to the pixels that show only a certain percentage of NSA frequency on 19a. if so, this needs to be specified. "AOD decreases by 5 % in unaffected regions" but it looks like it varies on 19b. More explanation here would be appreciated.

The language below in red was added to clarify that the 5 – 10% values are rough estimates based only on the values > 1 within the red boxes. The 5 % value is based on the average.

> *AOD increases by roughly 5–10 % in level 3 profiles affected by the NSA (based on values > 1 within the red boxes) because strongly negative near-surface σ is rejected. Conversely, the NSA is also present in this example along the equator, and yet excluding these σ values does not increase AOD, illustrating the difficulty of predicting the influence of the NSA on retrieved extinction. AOD also decreases by roughly 5 % on average in unaffected regions, a consequence of this conservative strategy.*

P26 L16: instead of "more aggressive at changing", consider "more impact than others"

I prefer to use "aggressive" to maintain consistent language throughout this section when referring to this filter. The wording in red below was added to help clarify the statement, however.

> *The aggressiveness metric $Agr(z)$ indicates the effectiveness of sample rejection on changing $\bar{\sigma}$, with larger values indicating the filter is more aggressive at changing $\bar{\sigma}$ than other filters.*

Fig 20b: The description of this graph is not clear. It does not seem to be used in the text either.

As commented previously, Figure 20b shows the number of unfiltered aerosol samples. It is shown to provide context for the number of samples contributing to the other three panels in the figure. The following sentence in red was added to the manuscript on P26 L17.

> *Figure 20 summarizes the impact of quality filters on the global $\bar{\sigma}$ profile, the frequency of rejection, and filter aggressiveness. ... For context, Fig. 20(b) shows the number of unfiltered aerosol samples, which decreases with increasing altitude.*

P27 L5: "dominate the aerosol sample rejection"

Done

P27 L11: four instead of three metrics

Corrected

P27 L12: delta_z_63 needs to be defined in the appendix, right after z_63 (eq. A2).

The following was added to Appendix A (now Appendix B):

> *The extinction scale height difference used in Sect. 6 is defined as*
>
> $$\Delta z_{63} = z_{63,\, all\ filters} - z_{63,\, no\ filters} \quad (B3)$$

P27 L13: consider two different names for "Agr" in Eq. A1 and Eq. A3.

Thank you for the suggestion. We find the two variations, $Agr(z)$ and ***Agr***, work well in terms of distinguishability and consistency so we will retain the current names.

P27 L22: "For the remaining filters, delta_z_63 is either zero or decreases by 60 m because these filters act upon layers at higher altitudes". This is over ocean according to Table 2. You might want to describe what happens over land as well.

This sentence applies to land and ocean because the "remaining filters" are the isolated 80 km, CAD, and cirrus fringe filters. The text in red was added to clarify.

> *For the remaining filters, Δz_{63} is zero or decreases by 60 m (over land and ocean) because these filters act upon layers at higher altitudes.*

Legend of Table 2: "and with no filters against all filters and with each filter applied independently". Consider rephrasing. Also, what is the reasoning behind the ocean and land separation in the Table?

Rephrased to:

> *Global metrics comparing changes in level 3 mean AOD when all filters are applied (top row) and when each filter is applied independently (remaining rows)*

Fig. 22: It is not clear why the authors show the surface elevation in this figure (not discussed in the text).

The following was added to P28 L16: *"Median surface elevations are shown to indicate altitudes where the number of samples averaged begins to decrease (often rapidly) relative to higher altitudes, thereby increasing the uncertainty in the mean values being compared."*

Fig. 23: Consider adding a mean AOD histogram as a fourth plot (instead of values on Fig 23a)

The mean values on Fig. 23a suite the analysis adequately by denoting which regions have higher AOD on average relative to others. Histograms would tell a similar story. However, adding 12 histograms onto the same plot (one for each region) would be difficult to interpret compared to the simple mean values. I prefer the keep the simple mean values reported in Fig 23a in lieu of adding the AOD histogram as a fourth plot.

P30 L6: "greater geometric depth". Am I not understanding this correctly? Doesn't this mean z_63_filtered > z_63_non_filtered? Which would mean a higher altitude under which 63 % of the aerosols reside (instead of a geometric depth)?

Yes, your interpretation is correct. This means the altitude under which 63 % of the aerosol reside is higher after filtering. The "greater geometric depth" phrase assumes that the base altitude remains the same. In that case, greater geometric depth means the same thing as higher altitude. However, I prefer the language you suggest, so the text is reworded as:

> *The change in extinction scale height Δz_{63} is positive for most regions (Fig. 23(b)), indicating that the altitude containing the bulk of the mean AOD is higher after quality filtering.*

P30 L16-18: "When aerosol loading is low, rejecting just a small number of aerosol samples may have a larger impact on sigma than in regions where aerosol loading is high since there are not many aerosol samples to begin with" This is not clear. Consider rephrasing.

Agreed. Rephrased to:

> *When aerosol loading is low, rejecting just a small number of aerosol samples may have a large impact on $\bar{\sigma}$ because there are not many aerosol samples to begin with.*

P31 L16: similar levels of uncertainty

Corrected

P31 L24: in order to reduce

Corrected

Appendix: Table B1 comes sooner in the text than appendix A: consider switching Appendix A and B.

Done. Appendices A and B are now switched. References to the tables/equations in the main text and supplementary material have been updated.

Fig. S1-S2 consider "quality screening" instead of "data filtering" to match the legend of Fig. 6-7
Done