Anonymous Referee #2

We would like to thank Anonymous Referee#2 for providing many detailed and substantive suggestions that greatly helped us improve the original manuscript.

"This paper builds upon current literature suggesting the necessity of a geographically varying dust lidar ratio for CALIPSO aerosol extinction retrievals. The authors focus on Saharan dust and dust mixtures over North Africa & Europe and provide substantial evidence for using a dust lidar ratio of 58 sr rather than 40 sr as implemented in the current CALIPSO algorithms based on comparisons with AERONET and co-located MODIS level 3 aerosol optical depth. They also propose two changes to the averaging strategy used in the production of the CALIPSO level 3 aerosol profile product with the goal to improve average dust extinction profiles by accounting for dust/polluted dust mixtures and accounting for other aerosol types differently. The metric for improved agreement between the current CALIPSO level 3 dust extinction climatology and the revised climatology algorithm proposed by the authors is the BSC-DREAM8b dust model.

A growing body of literature supports the notion that dust lidar ratio varies regionally and this should be accounted for in CALIPSO aerosol extinction algorithms and by that of future space-borne LIDAR systems. Since dust is one of the most influential natural aerosol sources with respect to global radiative forcing, this issue is important to address for improving global climate models. Therefore, this paper is well within the scope of Atmospheric Chemistry and Physics. Though the suggestion that the CALIPSO dust lidar ratio should be revised over northern Africa and Europe is not new with this manuscript, this is one of the first papers tackling the practical details of implementing such a revision in regional monthly averages. Further, the proposed methods of averaging non-dust aerosol into CALIPSO level 3 averages are new, making the content of the manuscript relevant, novel and noteworthy.

Altogether, this is a strong manuscript that is comprehensive and clearly written. Methodologies are thoroughly documented and sufficient background material is summarized to place this work in the context of current literature on the topic. Most of my specific comments request additional detail or provide suggestions to improve upon an already superb manuscript. I do not believe most comments will change the papers' conclusions, but I request the authors please pay special regard to the following specific comments as the first may influence quantitative conclusions and the second is a main point in the paper that needs elaboration.

• The CALIPSO AOD averaging methodology implemented in section 3.2 causes a low bias in the mean AOD when quality screening is used. Mean AOD should be computed by averaging the quality-screened aerosol extinction profiles and then vertically integrating the mean aerosol extinction profile to avoid this low-bias. (See specific comment 3).

• Please provide justification for assuming extinction = 0.0 /km for aerosol species that are not dust or polluted dust and show the impact of this assumption alone. (Specific comment 10).

Given the high quality of this novel manuscript, the timeliness of its content and its scientific value, I recommend this paper for publication in Atmospheric Chemistry and Physics after the authors please address my comments".

Specific comments:

"1. Page 12, line 19. Using only extinction QC = 0 may bias the averages low because extinction QC = 16 indicates cases where the layer is opaque which is a perfectly acceptable scenario that occurs for the densest layers. Additionally, it is important to quality screen by extinction uncertainty. This was implemented in Section 3.2, but not here. Why not? Does including/excluding these additional quality screening constraints change your conclusions or statistics upon which your conclusions are based?"

The filters applied in our analysis followed closely the methodology introduced in Schuster et al, (2012). Specifically, we used the following Schuster et al. (2012) CALIPSO and AERONET aerosol optical depth comparison criteria:

1. CALIPSO Level 2, Version 3 cloud-free column aerosol optical depths at the closest approach
2. AERONET Level 2, Version 2 (includes cloud screening and quality control)
3. CALIPSO laser footprint collocated within 80 km of AERONET site
4. CALIPSO overpass synchronized to within 30 min of AERONET AOD measurement
5. CALIPSO DEM_ surface elevation must be within 100m of AERONET site elevation
6. CALIPSO extinction QC flag=0 for all layers
7. CALIPSO CAD score less than −20 for all layers

Regarding the QC flag, Schuster et al. (2012) mentions that: "CALIPSO Extinction QC 532 = 0, indicates that a successful extinction solution was achieved with the default lidar ratio assigned to each layer. We do not screen for opaque aerosol layers in the CALIPSO product (no ground return), but there are no opaque layers in our analysis."

In order to account for the reviewer's suggestion, we examined our dataset for opaque layers finding that no profile from the total of 1203 profiles used had an ext_QC equal to 16. By expanding our inquiry onto a larger domain, we found 38 opaque cases in a sample of 3305 dust profiles. As a consequence, the QC filter was found not to alter our analysis.

This is not true however when it comes to the extinction uncertainty criterion. We found 8 out of 1203 collocated cases that should be excluded due to the extinction uncertainty filter. These cases corresponded to low AODs (lower than 0.1).

Following the reviewers' comment, a new dataset has been created taking into account the filters suggested. The CALIPSO-AERONET comparison and related plots have been revised based on this comment and also taking into account the comments made by David Winker. A major re-evaluation has been made in our paper in order to acknowledge the points raised by both reviewers and we thank them for pointing out these important issues.

"2. Page 15, line 21. The CALIPSO extinction quality screening metrics different in Section 3.2 from those used in Section 3.1 (page 12, lines 16-21). It is recommended that the same quality screening metrics be used in both sections not only to avoid biasing statistics derived from the analyses (quality screening procedures are insufficient in Section 3.1), but it would also make the CALIPSO analyses consistent throughout the paper".

This comment has been taken into account in our revised analysis. The quality screening metrics are clearly presented and a common policy regarding the quality requirements has been kept in the manuscript.

"3. Page 16, lines 2-4. Mean AOD is computed here by integrating the quality screened dust extinction profiles and then averaging the AODs. Though it has not been documented in the literature, computing mean AOD this way will be biased low. The reason is because when the profiles are quality screened, the optical path length changes. For instance, say a profile has only one aerosol layer extending from 4 km to the surface which had a CAD score of -5 which would be removed by quality screening. Next to this profile is another single aerosol layer extending from 4 km to the surface with a CAD score of -100. The second profile passes quality screening. Mean AOD is computed in this paper by integrating both of these extinction profiles independently and then averaging the AODs. The result will be the average of zero AOD in the first profile and a non-zero AOD in the second profile which will lead to a mean AOD that is biased low because the optical path lengths are not equal solely due to quality screening.

The alternative way is to average the quality-screened dust extinction profiles into a mean dust extinction profile and then integrate to acquire a mean AOD. This method assumes that the extinction of the 'bad' aerosol layer that we removed in the example above is the same as the 'good' aerosol layer in the second profile. This is a reasonable assumption – that the layer is horizontally homogeneous and is used by models when assigning aerosol into grid cells.

The low bias in mean AOD computed by averaging the AODs can be quite large compared to the alternative method; sometimes by a factor of two or more. The attached Figure 1 shows CALIPSO level 3 mean AOD for every latitude/longitude grid cell in July 2007, night, all-sky sky condition, computed by averaging the AODs (method 1; vertical axis) and computed by integrating the mean aerosol extinction profile (method 2; horizontal axis). Colors represent the number of latitude/longitude grid cells on a logarithmic scale. A low bias is always present in the former method with respect to the latter method. It is recommended that AOD be computed by integrating the quality-screened dust extinction profiles rather than the method currently employed".

The reviewer is right here. We had experienced this underestimation when we compared the different product versions with the BSC-DREAM8b model (Section 3.3). However we neglected this difference in the AERONET comparison. Instead, we have filtered our dataset having a standard deviation criterion that excludes scene averages that are produced from high variability in the 5-km single AODs. We agree that the difference between the two averaging methods should maximize in the case of bad aerosol layers that interrupt a continuous scene and cause inhomogeneity, which however may come from atmospheric inhomogeneity as well.

We acknowledge here that the standard deviation threshold may not filter the "bad" scenes in some cases and so we altered our analysis in order to apply more strict filters. To be more specific, in the revised version of our manuscript, we performed calculations using both methodologies (average of single AODs vs. integrated AODs from averaged profiles) in order to provide a plot similar with the one provided by the reviewer for our dataset (Figure 1 – lower, left plot). From this Figure, we do not observe large discrepancies between the two methods (integration of averaged extinction profile vs. averaging of single 5-km AODs). We believe that one reason for this is because we use cloud-free profiles only and homogeneous dust scenes. However, a small number of cases show inconsistency.

In the revised version, we use 2 criteria to screen bad retrievals and atmospheric inhomogeneity, i.e. (1) the two averaging methods should agree within 0.02 of the AOD and (2) the standard deviation for the 5-km AOD dataset should be less than 0.02. Moreover, we follow the reviewer's suggestion and use only integrated values in the revised version and moreover, we apply even more strict filters, adopting Level 3 constraints which are mentioned in detail in the revised manuscript. Furthermore, we have to mention here that a scene that passes all these criteria, should also have acceptable 5-km retrievals for the full length included in AERONET 80-km or MODIS 1x1 degree collocation requirements.

The conclusions of the comparisons are not significantly altered.

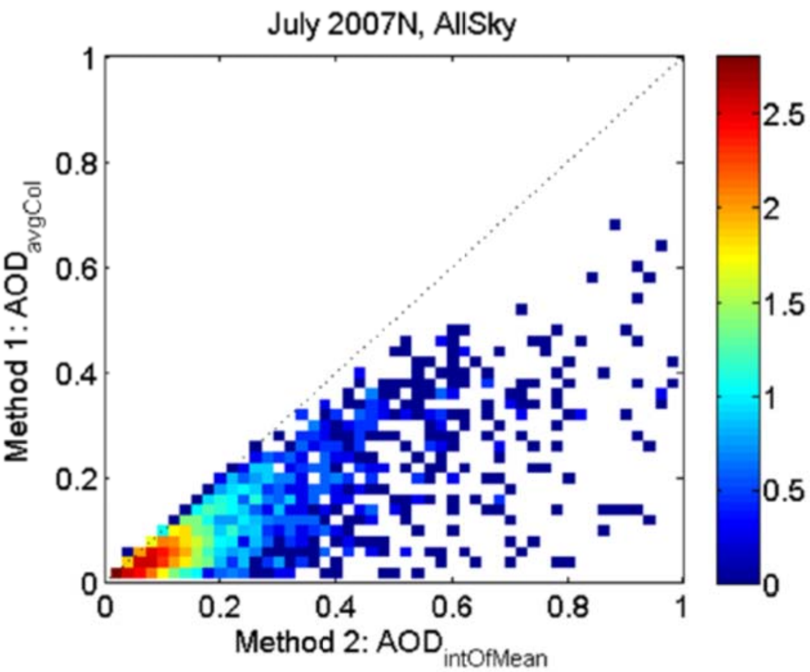

Fig. 1. *CALIOP level 3 mean aerosol optical depth computed by two methods (see specific comment 3)*

"4. Figure 3, top row. Suggestion. . .It would be better to visualize these as 2D histograms where the colors represent the number of cases within a certain CALIPSO AOD / MODIS AOD bin. That way the central tendency would become more evident. As it is displayed, the tendency of the outliers is highlighted. I want to know where samples are most frequent in the black blob near the origin where all of the points are bunched on top of each other".

We provide 2D histograms in the revised version of our manuscript and we include a comment on the central tendency which maximizes for AODs lower than 0.2. We don't comment further on that, since this finding needs more analysis and larger datasets to draw more general conclusions.

"5. Figure 3 caption. The caption says that the upper panel has no filters applied while the lower panel has "filters for dust presence,. . .". Does this imply that aerosol types other than dust are included in the top row? If so, why? The discussion is focused on dust and the impact of changing the dust lidar ratio is shown by comparing the left and right panels. With all of this, it seems like we should not be looking at all aerosol types in the top row, only dust."

Figure 3 and the related discussion have been revised to make our findings clearer. The reviewer is absolutely right in their comments and suggestions here, which helped us improving this section a lot. In any case, we show only dust cases in Figure 3. For the revised analysis we specifically show:

- in the upper panel of Figure 3 - only dust cases for which CALIPSO data are filtered based on Level 3 specifications
- in the lower panel of Figure 3 - the final 234 cells of our comparison, resulting from filters applied to account for the horizontal homogeneity and encounter for the spatial sampling differences of the two sensors. To be specific, we use only MODIS Level 3 AODs produced from at least 60 Level 2 records of 10 km spatial resolution out of a maximum of 121 pixel counts - to ensure the representativeness of the MODIS Level 3 product. Moreover, we use only the cases where the CALIPSO overpass has a length greater than 100 km within the MODIS cell. Since we use only dust CALIPSO retrievals, the latter prerequisite along with the criterion for the AOD average consistency between $AOD_{IntOfMean}$ vs $AOD_{AvgOfAods}$ ensures the dust presence domination in the 1x1 degree grid cell as well. Finally, the MODIS-retrieved cloudiness is set to be less than 20%.

The relevant paragraph in the manuscript has been re-written to make clearer the aforementioned data procedures and corresponding results.

"6. Figure 3 caption. The text says that the filters applied to the bottom row are for dust presence, cloudiness, and MODIS sampling. Are CALIPSO quality screening filters applied to all panels in this figure? They should be. Either way, it is recommended that the authors please clarify when these quality screening filters are or are not used".

The filters used for CALIPSO are now mentioned in Paragraph 3.2.1 (Comparison methodology). We have rephrased the text to be clearer in this part. Figure 3 caption now reads:

"Comparison of CALIPSO AODs (1x1 degree) versus collocated MODIS-Aqua Level 3 product using LR equal to 40 sr (left) and LR equal to 58 sr (right). Upper panels: 2D histograms representing the number of cases found for each of CALIPSO/MODIS AOD bin between 0 and 1.0 (bin step equal to 0.0125). MODIS data are not filtered, while CALIPSO data are filtered according to Level 3 specifications. Only CALIPSO overpasses that are cloud-free and for which the aerosol classification scheme reveals only dust presence are considered. Lower panels: data are screened to ensure horizontal homogeneity in the cell and CALIPSO data representativeness compared to MODIS spatial averages, as well as cloud-free conditions for the MODIS cell. "

"8. Page 17, line 8. The AOD bias of the original CALIPSO product is written as -0.07. It should be rounded to -0.08 based on the value in Table 2".

The reviewer is right. After the introduction of AOD calculation using integration of the averaged profiles, the new bias reported in Table 2 is -0.066 and we now quote in the text the -0.07 value.

"9. Page 17, line 9. The slope of the linear regression is written as 0.73, but it is listed in the top row of Table 2 as 0.704".

The slope is 0.73 after the new calculations (so we didn't revise the text). The revised statistics are written in detail in Table 2 and the text follows now the Table.

"10. Page 26, line 3. Provide an argument to justify assigning an extinction value of 0.0 /km to aerosol types other than polluted dust and dust. Lines 2-3 state that the extinction is lower below 0.5 km after making this assumption because the marine aerosol extinction was set to 0.0 /km instead of ignored in the average. However, no benefit or rationale is provided. One benefit that could be proposed is that it makes the vertical profile of extinction homogeneous all the way to the surface instead of becoming larger in the lowest 0.5 km, but how do we know the dust extinction does not increase in the lowest 500 meters? Many aerosol extinction profiles are largest near the surface. Please explain the rationale for making this assumption in the text. What does it mean physically or statistically? How do the AOD comparisons change with and without this assumption? Assigning extinction = 0.0 /km for all aerosol types other than polluted dust and dust will drive the Version III mean AOD down. On the other hand, ignoring all other aerosol types in the average will drive the mean up. Why is the former better than the latter?"

By examining the original CALIPSO L3 combined v1.0 product, there are some cases where the 2°x5° product of the dust monthly mean extinction profile can appear larger than the monthly mean extinction profile with reference to the total aerosol load. A case like this can be seen in the following figure which refers to the L3 Apro combined v1.0 product of 2008-February, for the cell with centroids -27.5 longitude & 10.0 latitude. The black line refersto the mean extinction at 532nm from all aerosols and the orange to the mean extinction at 532nm from only dust aerosols.

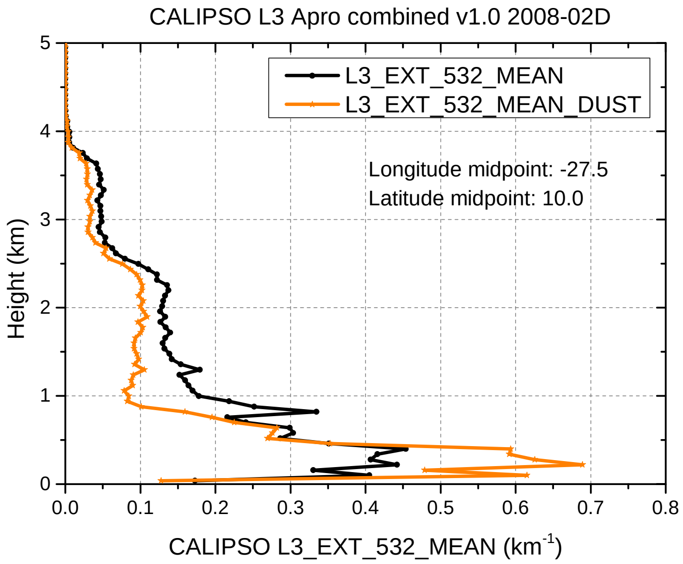

Figure RII-3: CALIPSO L3 Apro combined v1.0 2008-02D profiles of mean extinction at 532 nm (black) and mean dust extinction at 532 nm (orange) for the 2°x5° cell with longitude midpoint at -27.5° and latitude midpoint at 10.0 °.

It should be expected that the extinction due to all aerosols will be greater than the extinction due to dust aerosols only. While this is true above about 0.5 km, it reverses at lower heights. We believe that this problem is introduced because other aerosol types are ignored in dust averaging scheme. The inconsistency is larger for altitudes closer to the ground where the number of other aerosol observations is statistically significant in comparison with dust aerosols and clear air observations.

To demonstrate the differences between the averaging procedures, we present in the new version of our manuscript a synthetic scene like the one presented below. Assuming that half of the scene has a uniform dust curtain (with an extinction of 0.1 km^{-1} at all heights from 0-4 km) and the other half has a marine layer from 0-2 km and clear air from 2-4 km (Figure RII-4). The mean extinction dust profiles derived without the assumption of other aerosol types with 0.0 km^{-1} is shown in Figure RII-4(b), and the one created with the assumption of other aerosol types with 0.0 km^{-1} extinction is shown in Figure RII-4(c).

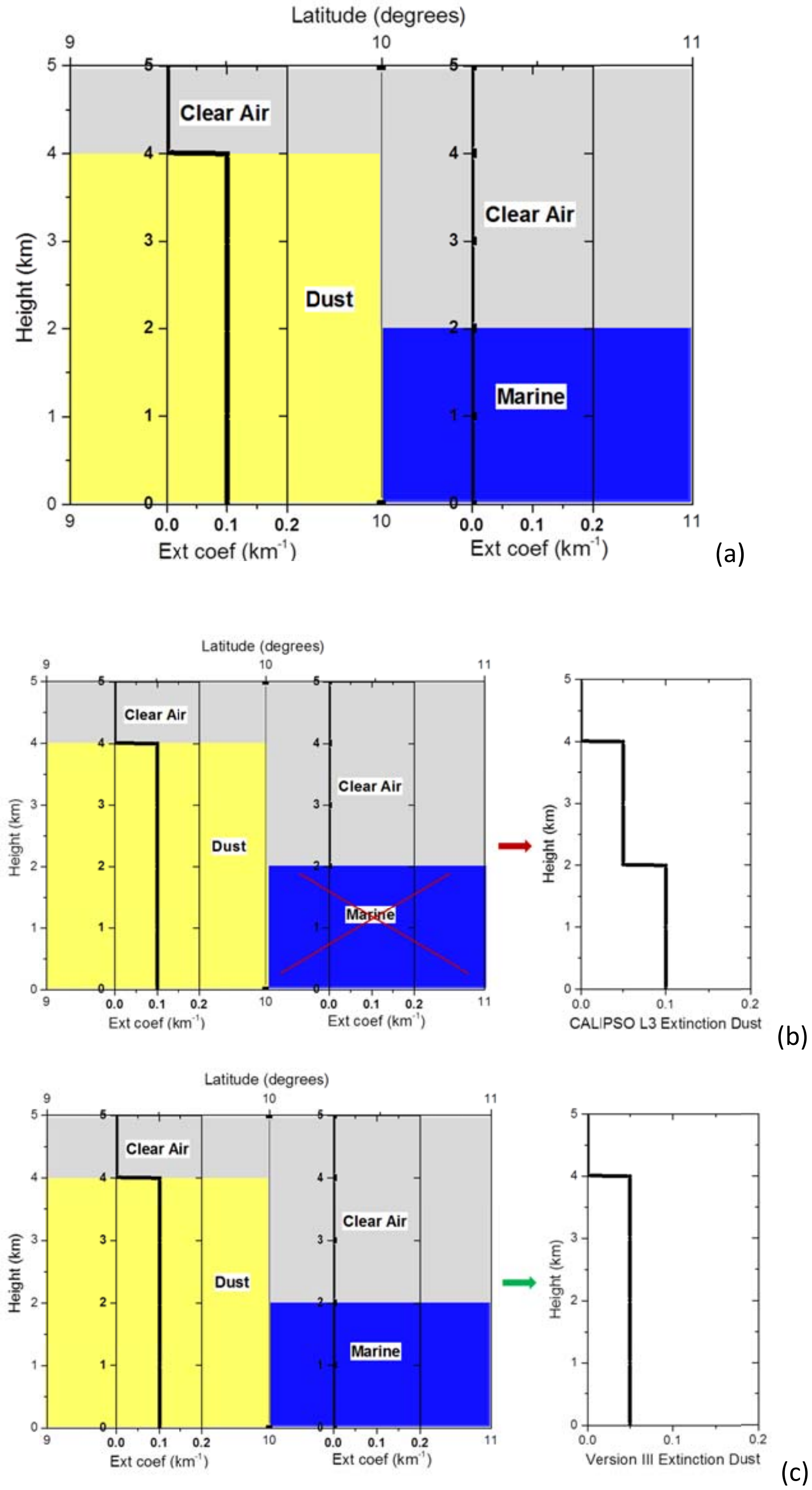

Figure RII-4: Case A: (a) a dust and marine scene in a CALIPSO overpass. (b) without the consideration of other aerosol types (CALIPSO L3, Version I), (c) with the consideration of other aerosol types equal to 0.0 km-1 dust extinction (Version III).

Another example of the same misrepresentation of the mean dust profile can be seen in Figure RII-5 where we see a scene with an elevated smoke plume between 2 and 4 km.

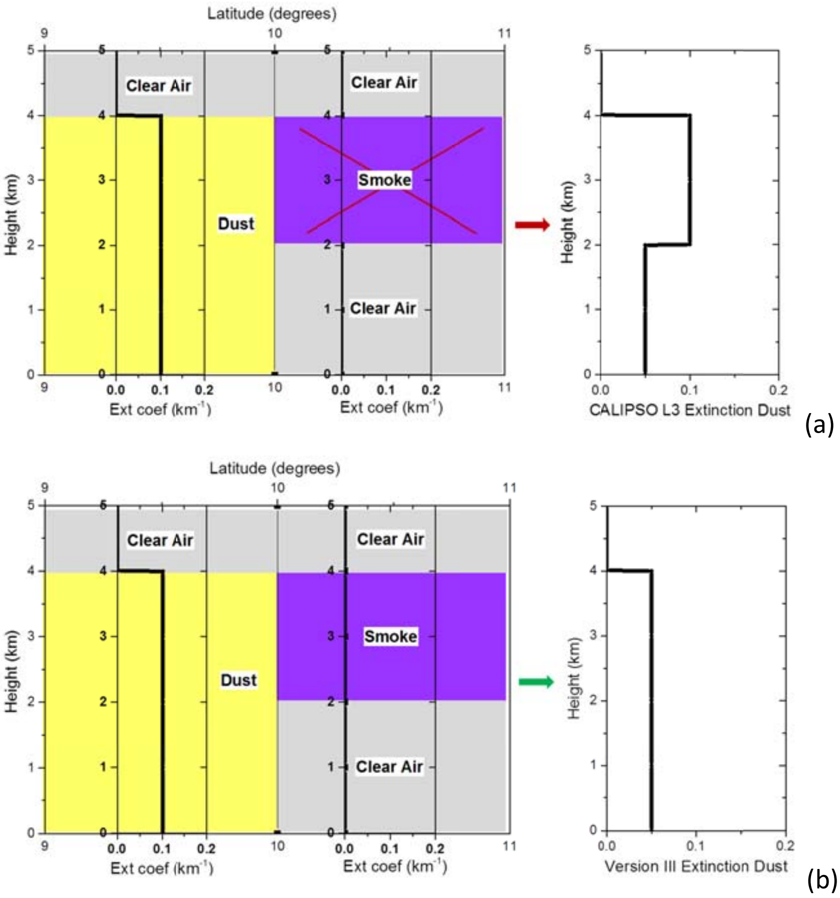

Figure RII-4: Case B: a dust and smoke scene in a CALIPSO overpass. (b) without the consideration of other aerosol types (CALIPSO L3, Version I), (c) with the consideration of other aerosol types equal to 0.0 km-1 dust extinction (Version III).

As we see in both cases (A and B), the mean dust load in the scene is the same, but due to the omission of non-dust aerosols in the scene a misrepresentation in the dust profiles is introduced in Version I (CALIPSO L3) profiles, - delivering different quantitative distributions of the same dust aerosol load.

The methodology proposed in our work is suitable for applications that acknowledge the contribution of dust in the total extinction profile. Such applications could include, for example, radiative-transfer simulations per aerosol type. Furthermore, the former knowledge of the aerosol vertical distribution of desert dust can help to improve the dust model over Europe through the estimation of the uncertainties with respect to production, particle properties, transport paths, and evolution during the particle's life time. In this sense, the new CALIPSO L3 Version III product presents larger differences at low levels in comparison with the standard CALIPSO L3 product would overestimate the dust extinction for these applications.

“11. Figure 5 caption. The text describing the green and red lines on the middle panel is inconsistent with the legend text on the figure. The caption text should label the mean layer depolarization reported by CALIPSO as the green line (not the red line) and the re-calculated particle depolarization (i.e., “corrected”) as the red line (not the green line)”.

This was indeed a typo. We correct it in the new version.

“12. Figure 9 color map. Consider using a color map that goes from blue to white to red. The current color map makes it difficult to separate positive and negative biases as well as changes in biases between the panels. This is only a suggestion”.

This is again a good suggestion; we provide now new maps introducing the colorbar suggested.

“13. Figure 8. The biases in Figure 9 change sign between North Africa (negative bias) and central Europe (positive). With this in mind, the profiles in Figure 8 should be evaluated for these two regions separately since we expect the profile shapes to differ since the vertical distribution and mixture of aerosols are different in these two regions. Pure dust would be more representative over Northern Africa perhaps making the Version II and Version III profiles similar while mixtures of dust and polluted dust over Europe could cause these profiles to be dissimilar”.

The suggestion has been followed and a new Figure (12) has been produced to demonstrate the profile shape differences regionally, over Europe and Northern Africa. The vertical distribution of the occurrence of different components averaged in each area is presented as well - to show the frequency of clear air, dust, polluted dust or other existing types. The respective description of Figure 12 has been added as the last paragraph in the Discussion section.